Investigation of the role of Capicua in the FGF signalling pathway and the wound response

Laura May-Seen Cowell

Master by Research

University of York

Biology

September 2019

Abstract

Fibroblast growth factor (FGF) signalling is critical for the initiation and regulation of multiple developmental processes including gastrulation, mesoderm induction and limb development. Despite extensive understanding of FGF signal transduction via tyrosine kinase receptors (RTKs), the specific mechanism responsible for regulation of target gene transcription is still not fully understood.

The protein Capicua (CIC) has been linked to transcriptional regulation in RTK signalling via the ERK pathway in multiple organisms. ERK signalling cascades also mediate wound signal transduction and transcription of associated genes. We hypothesise that transcription of a subset of FGF target genes, and genes involved in the wound response, rely on ERK mediated relief of CIC transcriptional repression.

The aims of this project were to establish if CIC operates downstream of FGF signalling through analysis and validation of RNA-seq data, and to investigate the relationship between CIC and ERK in FGF signalling and wound healing in *Xenopus* embryos.

The work in this thesis shows that CIC knockdown and FGF overexpressing embryos exhibit similar phenotypes and transcriptomes. Immunostaining for myc-tagged CIC indicates that CIC expression is reduced following activation of FGF signalling or the wound response. 75% of the putative CIC and FGF regulated genes analysed have enriched CIC binding sites and 75% were upregulated in RT-PCR analysis of CIC knockdown and FGF overexpressing embryos. Additionally, genes associated with wound healing (*fos and gadd45a*) are upregulated in CIC knockdown embryos. These data support the notion that CIC has a transcriptional regulatory role for a subset of FGF target genes and genes involved in the wound response.

Misregulation of the FGF signalling pathway and/or CIC repression is associated with a range of disorders and cancers. Consequently, understanding the molecular mechanisms involved in these pathways may allow development of more effective treatments for injury, neurodegenerative and developmental disorders, and cancer.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my supervisor Dr Harv Isaacs for his invaluable support, guidance and feedback on various reports and presentations throughout this project. I would also like to thank Dr Betsy Pownall, Lewis, Gideon, Ali, Alex and Monica for their help, knowledge and encouragement throughout the year. Thank you also to Dr Will Brackenbury for his time, advice, and constructive feedback and questions as a member of my thesis advisory panel.

Declaration

I declare that this is an original piece of work conducted under the supervision of H. V. Isaacs at the University of York. All data presented is my own work unless referenced otherwise. 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 Biology Technology Facility. Initial quality control of the RNA-seq data and mapping of sequences to the *Xenopus tropicalis* transcriptome was undertaken by John Davey at the University of York Technology Facility. With the exception of initial RNA-seq gene lists, work presented in this thesis has not previously been published or submitted for a qualification at this, or any other, University.

Chapter 1: Introduction

1.1 The fibroblast growth factor family

Fibroblast growth factor (FGF) signalling is critical for the initiation and regulation of multiple developmental processes including gastrulation, mesoderm induction and limb development (Rossant et al., 1997; Slack et al., 1996; ten Berge et al., 2008). FGFs also play essential roles in injury and tissue repair and regeneration in the adult (Maddaluno et al., 2017). In vertebrates, the FGF ligand family has 22 members which can be categorised into 7 phylogenetic subfamilies (Ornitz, 2000). These subfamilies can also be further classified into 3 groups: intracellular (FGF11-14), endocrine (FGF19, 21 and 23) and canonical (FGF1-10, 16-18, 20 and 22) (Itoh, 2010). Intracellular FGFs have high sequence homology with the FGF family but are not secreted and do not activate FGF receptors (FGFRs) (Beenken and Mohammadi, 2009). The canonical FGFs share a conserved core of 140 amino acids, have a strong affinity for heparin sulphate proteoglycans (HSPGs), and signal through the FGFR family of tyrosine kinase receptors (Dorey and Amaya, 2010). Endocrine FGF ligands target cells through the bloodstream and bind Klotho molecules as co-factors for receptor binding rather than HSPGs (Ornitz and Itoh, 2015).

1.2 The fibroblast growth factor receptor family

FGFR1, 2, 3 and 4 are composed of a single transmembrane domain, an intracellular domain with tyrosine kinase activity, and an extracellular ligand binding region (Böttcher and Niehrs, 2005). Immunoglobulin-like domains Ig1, Ig2 and Ig3 in the extracellular region are important for ligand and co-factor binding, receptor dimerisation, and ligand specificity as different isoforms of FGFR1-3 are generated through alternate splicing of Ig3 (Table 1) (Johnson et al., 1991). Binding specificity is primarily regulated by alternative splicing of exons IIIb and IIIc of the C-terminal half of Ig3. For example, FGFR2 and FGFR3 exclusively express the IIIb exon in epithelial cells and IIIc in mesenchymal lineages (Scotet and Houssaint, 1998; Wuechner et al., 1996). Ig2 also contributes to FGF ligand and receptor specificity as it rotates to form a greater number of contacts with ligands such as FGF10 in a ligand-induced conformation change (Yeh et al., 2003). In this manner, FGF signalling is tightly regulated in order to ensure correct embryonic development and adult tissue homeostasis. Misregulation of the FGF signalling pathway can lead to a wide range of disorders including skeletal abnormalities, hypophosphatemic rickets and deafness (Beenken and Mohammadi, 2009). FGFR mutations in particular are associated with a number of cancers as FGF signalling is important in regulating cell proliferation and angiogenesis (Turner and Grose, 2010).

In addition to the 4 classical tyrosine kinase receptors, FGFR-like (FGFRL) also has a strong affinity for FGF ligands but lacks a tyrosine kinase domain (Trueb et al., 2003). The receptor is otherwise structurally similar to FGFR1-4 and was initially thought to function only as a decoy receptor to further modulate FGF signalling (Steinberg et al., 2010). However, overexpression of FGFRL leads to increased ERK signalling, thus suggesting that the receptor may also act as a non-tyrosine kinase signalling molecule (Ornitz and Itoh, 2015).

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

Table 1: Ligand specificity of the fibroblast growth factor receptor family (Tiong et al., 2013; Steinberg et al., 2010)

1.3 Canonical fibroblast growth factor signal transduction

Binding of FGF and accessory molecule HSPG to FGFR1-4 leads to receptor dimerisation and autophosphorylation of intracellular tyrosine kinase domains (Schlessinger, 2000). This leads to further tyrosine phosphorylation within the receptor to create docking sites for secondary messengers. Phosphotyrosine residues bind Src homology 2 (SH2) domains found in various signalling proteins, allowing recruitment and activation of signalling complexes (Pawson et al., 1993). FGF signal transduction proceeds via 3 main pathways: the phospholipase Cγ (PLCγ), phosphoinositide-3 (PI3) kinase or extracellular signal-regulated kinase (ERK) pathway (Figure 1).

Binding of PLCγ to an FGFR phosphotyrosine residue at position 766, activates the enzyme leading to hydrolysis of phosphotidylinositol-4,5-diphosphate to inositol-1,4,5-trisphosphate (IP_3) and diacylglycerol (DAG) (Peters et al., 1992). IP_3 stimulates intracellular calcium release and DAG activates protein kinase C. This FGF signal transduction pathway is required for the modulation of planar cell polarity needed for cell movements during gastrulation (Sivak et al., 2005).

FGF signalling can also activate the PI3 kinase pathway through Frs2 (FGFR substrate 2) binding Gab1 (growth factor receptor bound protein 2-associated-binding protein 1) indirectly via Grb2 (growth factor receptor bound protein 2). Phosphorylation of Gab1 leads to the subsequent

recruitment and activation of PI3 kinase and downstream effector proteins (Ong et al., 2001). AKT/protein kinase B is an important downstream mediator of this pathway and is a proto-oncogene involved in regulating cell survival and growth (Nicholson and Anderson, 2002).

The most common FGF signalling pathway in development involves association of Frs2 with the activated receptor (Kouhara et al., 1997). This allows assembly of the protein complex required for initiation of the ERK signalling cascade (Hadari et al., 2001). Frs2 recruits the adaptor protein Grb2 which in turn binds the guanine nucleotide exchange factor son of sevenless (Sos) via its Src homology 3 (SH3) domain (Ong et al., 2000). Sos promotes dissociation of GDP from small GTP binding protein Ras to allow its conversion to the active GTP-bound form. Activation of Ras leads to a cascade of sequential phosphorylation of transducing proteins Raf, Mek and ERK (also known as MAPK) (Schlessinger, 2000). Activated serine/threonine kinase ERK induces expression of target genes by directly phosphorylating transcription factors or by phosphorylating and activating other kinases to do so.

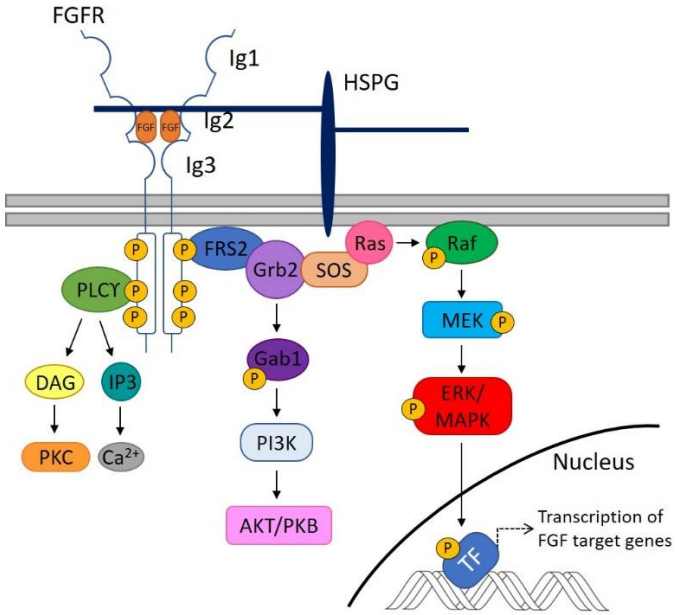

Figure 1: Schematic representation of FGF signal transduction via the PLCγ, PI3 kinase and ERK signalling pathways. Binding of extracellular fibroblast growth factors (FGFs) and accessory molecule heparin sulphate proteoglycan (HSPG) causes dimerisation of FGF receptors and autophosphorylation of intracellular tyrosine kinase domains. FGF signal transduction proceeds via 3 main pathways: the phospholipase Cγ (PLCγ), phosphoinositide-3 kinase (PI3K) or extracellular signal-regulated kinase (ERK) pathway.

1.4 Transcriptional regulation in the FGF signalling pathway

Many FGF target genes have been identified through transcriptomic screens involving the comparison of wild type *Xenopus* with those in which FGF signalling had been impaired. For example, a microarray screen of *Xenopus* explants treated with SU5402, a specific FGFR1 inhibitor,

highlighted 38 genes positively regulated and 5 genes negatively regulated by FGF signalling (Chung et al., 2004). Inhibition of FGF signalling by dominant negative mutants of FGFR1 and FGFR4a led to the identification of 67 genes significantly upregulated and 16 genes significantly downregulated by FGF signalling in normal *Xenopus laevis* development (Branney et al., 2009).

The caudal-related (Cdx) group of ParaHox genes are among the FGF targets downregulated in the presence of dominant negative FGFRs (Pownall et al., 1996; Isaacs et al., 1998; Branney et al., 2009). In vertebrates, the *Cdx* genes are required for normal antero-posterior patterning, and the specification and differentiation of gut endoderm (Beck and Stringer, 2010). *Cdx1*, *Cdx2* and *Cdx4* are initially expressed in the early mesoderm surrounding the blastopore region in *Xenopus* gastrulation (Pillemer et al., 1998). Normal mesodermal *Cdx* gene expression requires FGF signal transduction via activation of ERK, in addition to FGF dependent wnt8 expression (Keenan et al., 2006). Overexpressing FGF4 in late gastrula to late neurula *Xenopus* leads to an anterior extension of the *Cdx4* domain and a subset of homeobox (Hox) genes resulting in a posteriorised phenotype (Pownall et al., 1996).

Brachyury (*Xbra*) is another well-established target whose expression is upregulated when FGF4 is overexpressed (Isaacs et al., 1994). The T-box transcription factor is crucial for mesoderm and notochord development in vertebrates, as homozygous mice mutants fail to form these along with all posterior structures (Chesley, 1935). *Xbra* expression is also reduced in *Xenopus* embryos treated with SU5402 (Chung et al., 2004) or expressing dominant negative FGFRs (Isaacs et al., 1994). *Xbra* is part of a closed autocatalytic regulatory loop in the early mesoderm where FGF4 induced *Xbra* expression maintains FGF4 expression via positive feedback (Isaacs et al., 1994).

FGF signalling is also mediated by E26 transformation-specific (ETS) transcription factors. The polyoma enhancer activator 3 (PEA3) subfamily (also known as ETS variant transcription factor 4 (Etv4)) includes PEA3, Etv5 and Etv1. Members of this subfamily act as downstream effectors of FGF-Ras-ERK signalling (Garg et al., 2018). Transcription of PEA3 ETS factors is repressed by the protein Capicua (CIC) in the absence of modulation by activated ERK, $p90^{RSK}$ (p90 ribosomal S6 kinase) and 14-3-3 proteins (Dissanayake et al., 2011). PEA3 and Etv5 expression closely correlates with that of FGF8 and FGF3, is lost in zebrafish embryos treated with SU5402 and is ectopically expressed in the presence of FGF8/3 coated beads (Raible and Brand, 2001). Additionally, knockdown of the PEA3 subfamily in zebrafish produced phenotypes resembling FGF deficient embryos and significantly decreased FGF target gene expression (Znosko et al., 2010). PEA3 and Ets2 have also been shown to mediate FGF signalling in a negative feedback loop by directly binding

Dusp6 (dual-specificity phosphatase 6) promoters to inactivate ERK (Ekerot et al., 2008; Znosko et al., 2010).

MyoD (myoblast determination protein 1), another well-documented FGF target gene, encodes a basic-helix-loop-helix transcription factor able to induce differentiation of skeletal muscle cells through direct regulation of muscle-specific genes (Tapscott, 2005). Previous work has shown that treatment of *Xenopus laevis* animal caps with eFGF (also known as FGF4) directly activates *Xenopus myoD* (*XmyoD*) even in the presence of translational inhibitor cycloheximide (CHX) (Fisher, 2002). This indicates that activation of X*myoD* by FGF is an immediate early response not requiring translation or transcription. CHX treatment alone activates some transcription of XmyoD suggesting that FGF target gene expression involves inhibition of a labile transcriptional repressor, the levels of which rapidly decay in the absence of protein synthesis. Expression of Cdx genes is also increased following CHX treatment of animal caps indicating that *XmyoD* is not the only FGF target likely to be regulated in this manner (Keenan et al., 2006).

Other evidence for an unstable protein repressing FGF transcriptional target *XmyoD* includes the fact that following the midblastula transition, *XmyoD* is transiently expressed at a low level ubiquitously throughout the embryo (Rupp and Weintraub, 1991). This is quickly silenced, possibly due to zygotic gene activation and translation of a transcriptional repressor. The high levels of *XmyoD* necessary for myogenic development could then be achieved in specific cells through inactivation of the repressor by modification via FGF signalling. Furthermore, transfection of the myoD enhancer sequence fused to the chloramphenicol acetyltransferase (CAT) reporter gene into multiple cell lines significantly increases *myoD* expression. However, expression of the human *myoD* enhancer upstream of a *lacZ* reporter in transgenic mice results in intense staining of only skeletal muscle-forming regions. Transgene activation and endogenous *myoD* expression is restricted to limb buds and the myotomal compartments of somites in vivo. The difference between the myoD expression pattern in tissue culture and *in vivo* may be due to the presence of a transcriptional repressor necessary for silencing *myoD* in non-muscle cells within the embryo which is not present in derived cell lines (Goldhamer et al., 1992).

1.5 The transcriptional repressor Capicua and ERK signalling

High mobility group box protein Capicua (CIC) has been identified as a transcriptional repressor potentially involved in regulating expression of FGF gene targets.

CIC has been linked to transcriptional regulation in receptor tyrosine kinase (RTK) signalling via the ERK pathway in multiple organisms including mice and humans (Jimenez et al., 2012). In all cases

studied, binding of CIC to target protein promoters and enhancers leads to transcriptional repression in the absence of RTK signalling. For example, Torso is a tyrosine kinase receptor activated at the anterior and posterior poles of the *Drosophila melanogaster* syncytial embryo, where its ligand Trunk is produced. This creates a gradient of ERK activity at each embryonic pole. Phosphorylation of CIC by activated ERK (diphosphorylated ERK (dpERK)) leads to degradation of CIC, thus creating an opposing gradient of CIC protein levels decreasing towards the pole (Jimenez et al., 2000). This allows transcription of huckebein (*hkb*) and tailless (*tll*). CIC represses *hkb* more effectively than *tll*, leading to a narrower, nested *hkb* expression profile required for the differentiation of terminal region structures (Duffy and Perrimon, 1994). At the anterior terminus, CIC is also involved in establishing posterior boundaries of Bicoid target genes through binding of Bicoid responsive enhancers (Löhr et al., 2009)

CIC has also been shown to act as a vital regulatory switch in RTK signalling in the *Drosophila* ovarian follicle where EGFR signalling induces nuclear export and partial relocalisation of CIC to the cytoplasm following phosphorylation by ERK. CIC is important in establishing the dorsoventral axis as it represses homeodomain transcription factor Mirror in ventral follicle cells, thus allowing *pipe* transcription (Andreu, González-Pérez et al., 2012). EGFR-mediated relocalisation of nuclear CIC therefore contributes to defining the *pipe* expression border position by affecting the spatial distribution of Mirror in follicle cells. Loss of maternal CIC function therefore leads to dorsalisation of the embryo due to absence of pipe expression (Andreu, Ajuria et al., 2012).

Increasing evidence regarding the involvement of CIC in RTK signalling and the fact that CIC fits the MyoD model for a transcriptional repressor led to our hypothesis that FGF regulated gene transcription is mediated, at least in part, by CIC. FGF has been shown to be the sole activator of ERK signalling in early development in *Xenopus tropicalis* (Christen and Slack, 1999), so CIC was identified as a potential component of the FGF pathway downstream of ERK.

1.6 Relief of Capicua transcriptional repression via interactions with dpERK

Capicua exists in 2 main isoforms with different sizes and N-terminal regions. In *Xenopus tropicalis*, CIC-long (CIC-L) contains N-terminal exon 1 (2812bp), whereas CIC-short (CIC-S) contains exon 2 (49bp) instead (King, 2019). The remaining CIC exons (3-22) are common to both isoforms in *Xenopus tropicalis* and *Mus musculus* (Figure 2A). *Xenopus tropicalis* and *Mus musculus* CIC proteins both contain 7 functional domains: N1, ATXN-1, 14-3-3. HMG-box, C2, NLS and C1 (Astigarraga et al., 2007). The N1 domain is only present in the CIC-L isoform and not the shorter CIC-S protein (Figure 2B).

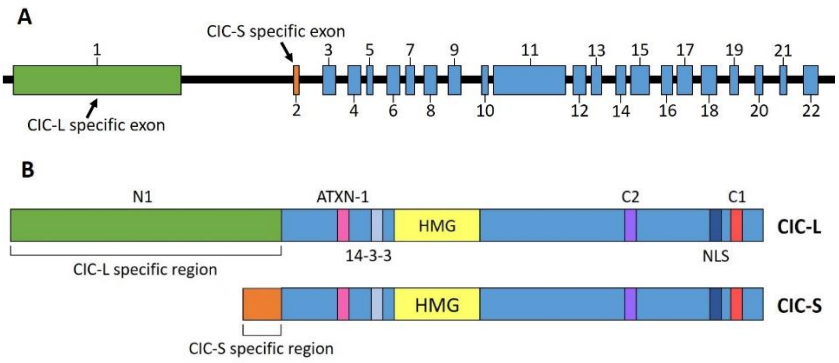

Figure 2: Schematic diagram of CIC exon structure and functional domains. A Capicua (CIC) consists of 22 exons. Exon 1 is specific to the CIC-long (CIC-L) isoform and exon 2 to the CIC-short (CIC-S) isoform. Both CIC-L and CIC-S share the common exons 3-22. **B** The CIC protein contains 7 functional domains, 6 of which are found in both the CIC-L and CIC-S isoforms. The N1 domain is only present in the CIC-L protein 2. The HMG domain is a high mobility group-box domain and the NLS encodes a nuclear localisation sequence.

Activated ERK acts via several mechanisms involving different functional domains to reduce nuclear levels or activity of CIC. For example, in *Drosophila*, dpERK binds to the C2 docking motif and directly phosphorylates CIC to induce degradation of the protein (Jimenez et al., 2012). dpERK binding of the C2 domain in fly and human model systems also mediates further phosphorylation of CIC in response to signalling. $CIC^{\Delta C2}$ mutant flies lacking the C2 domain exhibit a similar phenotype to Torso KO (Astigarraga et al., 2007). Without a functional C2 motif, CIC is insensitive to relief of repression by Torso signalling. This suggests that the C2 domain is essential for relief of CIC repression.

CIC binds DNA via its High Mobility Group-box (HMG-box) domain, which recognises octameric T(G/C)AATG(A/G)A sites in target gene enhancers and promoters (Jimenez et al., 2012). The 14-3-3 domain contains a serine or threonine residue which, once phosphorylated, recruits members of the 14-3-3 phosphoserine family of binding proteins. In human HEK-293 cell studies, ERK signal transduction activates p90 ribosomal S6 kinase ($p90^{RSK}$) expression. $p90^{RSK}$ phosphorylates the serine at the 14-3-3 motif to create a docking site for 14-3-3 regulatory proteins (Dissanayake et al. 2011). This blocks or reduces optimal binding efficiency of the neighbouring HMG-box domain to DNA leading to relief of repression by CIC.

Human CIC contains a common nuclear localisation sequence (NLS) at the C-terminal. This allows binding of Importin α4 (KPNA3) leading to import of CIC into the nucleus. Activated ERK phosphorylates the NLS site at Ser^{1382} and Ser^{1409} and/or KPNA3 to prevent binding of KPNA3 to CIC and subsequent nuclear import (Astigarraga et al., 2007).

1.7 ERK signalling and the wound response

ERK is also activated in response to injury. It has been demonstrated that 2 minutes after mechanical wounding of intestinal epithelial cell (IEC-6) monolayers, cells at the wound edge strongly express cytoplasmic dpERK. ERK signalling cascades then lead to further phosphorylation and transcription of genes needed for tissue repair such as early growth response-1 (Egr-1) and c-Fos, peaking at 20 minutes post-wounding (Dieckgraefe et al., 1997). Active ERK1/2 expression is also upregulated in tissue repair following wounding of lens epithelial cell (LEC) monolayers (Wang et al., 2003). Additionally, healing is inhibited in LEC monolayer wounds treated with U0126, a specific inhibitor of ERK1/2 activation, thus indicating that ERK activity is required for the wound response.

Pre-metamorphic Xenopus are able to heal epidermal wounds without scarring, as well as regenerate limbs, tails and lens' as required (Beck et al., 2009). This makes them an excellent model for the investigation of wound healing and regenerative mechanisms. ERK is activated rapidly and transiently in response to wounding or dissection of a *Xenopus* embryo (Christen and Slack, 1999; LaBonne and Whitman, 1997). *Xenopus* embryonic wound healing has 2 phases (Li et al., 2013). The first of which involves activated ERK suppressing PI3 kinase activity to allow the subsequent activation of Rho and myosin-2. This leads to assembly of a contractile actomyosin cable in epithelial cells surrounding the wound (Martin and Lewis, 1992). This cable, along with the contraction and ingression of exposed deep mesenchymal cells, draws the wound together like a 'purse string' (Bement et al., 1999). Actomyosin cables have also been observed in the healing of adult tissues such as the intestinal and cornea epithelia so this mechanism may not be restricted only to the embryonic wound response (Bement et al., 1993; Danjo and Gipson, 1998). ERK signalling decreases after phase 1 to allow PI3 kinase signalling to resume, thus leading to increased Rac and Cdc42 activity (Li et al., 2013). This allows filopodia formation at the wound leading edge to promote migration and wound closure. Unlike endogenous activation of ERK in *Xenopus* embryos, ERK activation in the wound response occurs independently of FGF and is unaffected by the presence of dominant negative FGF receptors (Christen and Slack, 1999).

Activation of ERK has been shown to relieve transcriptional repression by CIC in several signalling pathways (Jimenez et al., 2012). As ERK is also involved in the wound response, this may be another mechanism through which target gene expression is regulated by ERK and CIC. We hypothesise that CIC functions in transcriptional regulation of genes involved in the wound response, in addition to FGF pathway target genes (Figure 3).

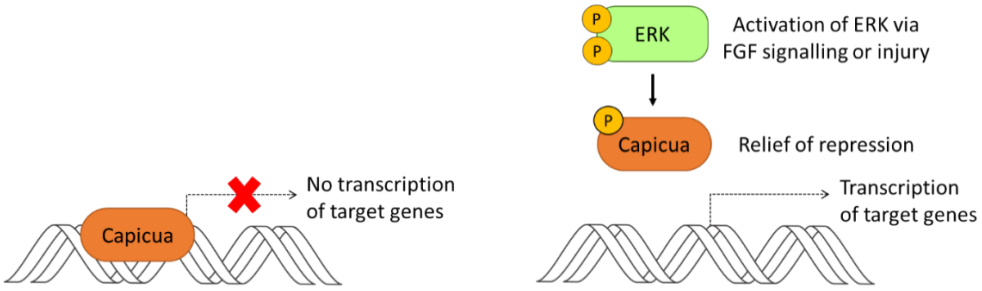

Figure 3: Hypothesised model for transcriptional regulation of FGF and wound response activated target genes. ERK (extracellular signal-regulated kinase) is phosphorylated when an injury is sustained or the FGF signalling pathway is activated. Phosphorylation of transcriptional repressor protein Capicua by diphosphorylated ERK leads to degradation or relocalisation of Capicua to the cytoplasm, thus allowing transcription of target genes.

1.8 Capicua expression in *Xenopus* development

CIC-L is maternally expressed at the 8-cell stage (stage 4) with enriched expression in the animal pole (Figure 4). CIC-S expression is initially lower but increases following the midblastula transition when zygotic gene transcription is activated. Both CIC isoforms are expressed in the marginal zone, at the dorsal blastopore lip, during early gastrulation (stage 10) (King, 2019). FGF3 and FGF8 are first expressed in early gastrula stage *Xenopus* around the blastopore in order to induce mesoderm in this CIC-expressing region (Lombardo et al., 1998; Christen and Slack, 1997). As development continues, expression of CIC-L and CIC-S becomes more widespread and is particularly enriched in the otic vesicles, branchial arches, neural tube, notochord, somites, forebrain, midbrain and hindbrain (King, 2019), all known regions of FGF signalling (Lea et al., 2009). Widespread CIC expression would be necessary at early tailbud stage (stage 25) to prevent non-specific transcription of FGF dependent genes if CIC is the downstream repressor in this signalling pathway.

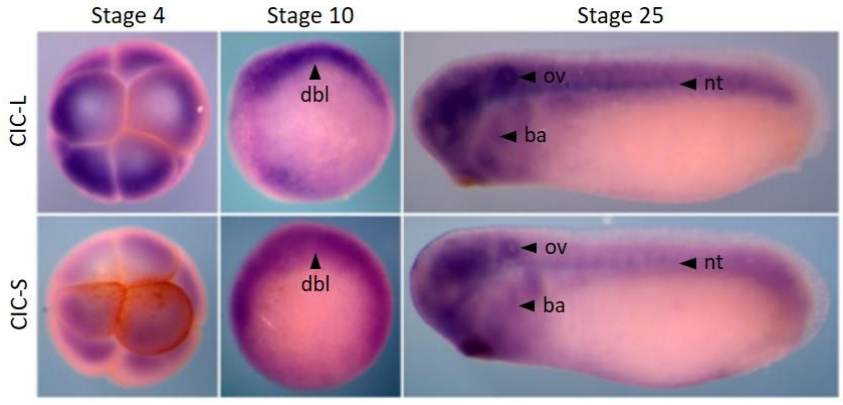

Figure 4: The spatial expression pattern of CIC-L and CIC-S in *Xenopus tropicalis* analysed using in situ hybridisation (King, 2019). 8-cell stage 4 (animal view), early gastrula stage 10 (vegetal view) and early tailbud stage 25 (lateral view). Both Capicua isoforms (CIC-long and CIC-short) are enriched at the dorsal blastopore lip (dbl) during gastrulation and later in the otic vesicles (ov), branchial arches (ba) and neural tube (nt). Embryos staged according to Nieuwkoop and Faber (1994) stages of *Xenopus* development.

1.9 Project aims

Although lots is known about the FGF transcriptome, the specific mechanism responsible for regulation of target gene transcription is still not fully understood. Evidence suggesting involvement of a labile transcriptional repressor (Fisher et al., 2002) led to identification of CIC as a strong candidate for this role, as the protein has been shown to function as a transcriptional repressor relieved by RTK signalling via ERK activation in multiple species (Jimenez et al., 2012). CIC is expressed in known regions of FGF activity, where activation of ERK in FGF signal transduction may lead to relief of target gene repression by CIC (King, 2019). ERK signalling cascades also mediate wound signal transduction and transcription of genes relevant to wound repair (Christen and Slack, 1999; Li et al., 2013). Consequently, CIC may also regulate transcription of genes involved in the wound response independently of FGF.

Hypothesis: Transcription of a subset of FGF target genes, and genes involved in the wound response, rely on ERK mediated relief of CIC transcriptional repression.

The overall aims of this project are:

- To establish if CIC operates downstream of FGF signalling through analysis and validation of RNA-seq data
- To investigate the relationship between CIC and ERK in FGF signalling and embryonic wound healing

Chapter 2: Materials and methods

2.1 Embryological methods

2.1.1 *Xenopus* embryo culture

Fertilised *Xenopus tropicalis* embryos were cultured in 1/9 strength Modified Ringers Solution (MRS)/9 (11.11mM NaCl. 0.2mM KCl, 0.22mM $CaCl_2$, 0.11mM $MgCl_2$, 5mM HEPES/NaOH (Tindall et al. 2007)) and staged according to Nieuwkoop and Faber (1994). Embryos were transferred to MRS/20 (5mM NaCl, 0.09mM KCl, 0.1mM $CaCl_2$, 0.05mM $MgCl_2$, 5mM HEPES) before gastrulation.

Fertilised *Xenopus laevis* embryos were cultured in 1/3 strength Normal Amphibian Medium (NAM)/3 (3.7mM NaCl, 0.067mM KCl, 0.033mM $Ca(NO_3)_2$, 0.33mM $MgSO_4$, 3.3µM EDTA, 5mM HEPES pH7.4, 1mM $NAHCO_3$ (Slack and Forman, 1980)) and staged according to Nieuwkoop and Faber (1994). Embryos were transferred to NAM/10 (1.1mM NaCl, 0.02mM KCl, 0.01mM $Ca(NO_3)_2$, 0.01mM $MgSO_4$, 1µM EDTA, 5mM HEPES pH 7.4 (Slack and Forman, 1980)) before gastrulation.

2.1.2 Microinjection

Xenopus tropicalis injections were carried out in MRS/9 + 3% ficoll using the gas PM 1000 Cell Microinjector with pulled needles (Narishige). *Xenopus laevis* injections were carried out in NAM/3 + 5% ficoll using the Drummond Microinjector with pulled needles (Drummond). Injected embryos were cultured in the ficoll solutions to allow healing before transfer to MRS/20 and NAM/10 respectively. Embryos were flash frozen on dry ice for RNA extraction and reverse transcription polymerase chain reaction (RT-PCR) or western blots, or fixed in MEMFA (0.1M MOPS pH 7.4, 2mM EGTA, 1mM $MgSO_4$, 3.7% formaldehyde (Guille, 1999)) for in situ hybridisation (after removal of the vitelline membrane) or immunostaining.

2.1.3 Wounding

The vitelline membranes of late neurula stage 20 *Xenopus laevis* embryos were removed before creating a wound on one side of each embryo using a tungsten needle. Embryos were cultured in NAM/10 at 21°C and fixed at a range of time points post-wounding.

The vitelline membranes of myc-CIC microinjected gastrula stage 10.5 *Xenopus laevis* embryos were removed from the vegetal pole to avoid unintentional damage to the animal pole. Embryos were then wounded at the microinjection site on the animal pole using a tungsten needle. The embryos were cultured in NAM/10 at 21°C for 30 minutes post-wounding before being fixed in MEMFA.

2.1.4 Photography

Images of embryos were taken using the SPOT 14.2 Colour Mosaic camera (Diagnostic Instruments Inc.) and SPOT Advanced software with a Leica MZ FLIII microscope. Adobe Photoshop CS3 was used to process images.

2.2 Molecular biology methods

2.2.1 Extraction of total RNA

In order to extract RNA, frozen *Xenopus tropicalis* embryos were homogenised in 1ml Tri-Reagent (Sigma-Aldrich) and left on ice for 1 minute. Samples were centrifuged at 13,000rpm for 10 minutes at 4°C and the supernatant placed at room temperature for 5 minutes. 200µl chloroform was added to the supernatant and left at room temperature for 5 minutes before being centrifuged at 13,000rpm for 15 minutes at 4°C. The aqueous phase was transferred to a new Eppendorf and 200µl chloroform added before being centrifuged at 13,000rpm for 5 minutes at 4°C. 500µl isopropanol was added to the aqueous phase, vortexed and placed at -20°C for 29 minutes. Samples were centrifuged at 13,000rpm for 15 minutes at 4°C and the supernatant discarded. 200µl ice cold 70% ethanol was added to the RNA pellet, vortexed and centrifuged at 13,000rpm for 10 minutes at 4°C. The supernatant was discarded and the pellet dried by desiccation. RNA was resuspended in 50µl 2.5M LiCl and placed at -20°C overnight to precipitate. Samples were then centrifuged at 13,000rpm for 20 minutes at 4°C. 200µl ice cold 70% ethanol was added to the RNA pellet, vortexed and centrifuged at 13,000rpm for 5 minutes at 4°C. The pellet was dried by desiccation and resuspended in 20µl dH_2O.

2.2.2 cDNA synthesis

cDNA was synthesised from 0.5µg RNA with 0.2µg random hexamers (Thermo Scientific) and 1µl 10mM dNTP (Roche) made up to 13µl with dH_2O. Reaction mixture was heated to 65°C for 5 minutes then placed on ice for 1 minute before addition of 4µl 5x SSIV buffer (Invitrogen), 1µl 100mM DTT (Invitrogen), 1µl 200U/µl SuperScript IV Reverse Transcriptase (Invitrogen) and 1µl dH_2O. The mixture was then incubated at 23°C for 10 minutes, 55°C for 10 minutes and then 80°C for 10 minutes.

IDT PrimerQuest (https://eu.idtdna.com/primerquest/home/index) was used to design gene specific primers with forward and reverse primers in different exons with amplicon sizes between 400bp and 800bp (Table 2).

Gene	Forward primer sequence 5'-3'	Reverse primer sequence 5'-3'
fos	CCAGATCTTCAGTGGCTTGT	CTATACAGTGGCTCCCATTCTG
fosl1	TCGCAAGGAGCTAACAGATTAC	CATGGACTTTGCTCTCCACTAC
frzb	GGCTGGTGCTCCTATCATTAC	CCTTCATGGGCTTGCATTTAC
rasl11b	ATCTTGCCATCCAGGTTCAG	CCTGCATGTTTGGTGACTTTG

Table 2: Forward and reverse gene specific primer sequences for amplification of *fos, fosl1, frzb and rasl11b.*

12.5µl PCR Master Mix 100rxn (Promega), 1.5µl 100µM forward and 1.5µl 100µM reverse gene specific primers and 7.5µl dH_2O were used to amplify 2µl cDNA via PCR. The reaction mixture was heated to 95°C for 2 minutes before 30 cycles of 95°C for 30 seconds, 58°C for 50 seconds and 72°C for 30 seconds before a final elongation at 72°C for 10 minutes.

PCR products were purified using the Quick-Start MinElute PCR Purification Kit (Qiagen) as per manufacturer's protocol.

2.2.3 Agarose gel electrophoresis

DNA and RNA samples were separated on 1.5% agarose gels in Tris-Acetate-EDTA buffer (40mM Tris pH7.6, 20mM acetic acid, 1mM EDTA) stained with ethidium bromide. The 1kb plus DNA ladder (New England BioLabs) was run alongside to predict product sizes.

2.2.4 Cloning of in situ hybridisation probe template

PCR products were ligated into the pGEM-T Easy vector and transformed into *E. coli* competent cells (DH5α, Invitrogen). Transformations were grown overnight at 37°C on LB agar plates containing 100µg/ml ampicillin. 5-10 colonies were selected from each bacterial plate and each subjected to PCR with gene specific primers (Table 2) to determine which colonies had taken up the insert successfully. For each gene, 3 colonies that had successfully taken up the insert were selected and cultured overnight in 3ml LB-broth containing 100µg/ml ampicillin, in a shaker at 37°C, 250rpm. Plasmids were isolated from the bacterial cultures using the QIAprep Spin Miniprep Kit as per manufacturer's protocol. DNA concentration was determined using the Nanodrop-8000 and an EcoR1 digest was carried out to confirm presence of the insert in the purified plasmids. Samples of the undigested plasmids were sequenced using the Eurofins Genomics postal sequencing service. Sequencing was analysed using SeqMan software from the Lasergene Genomics Suite (DNA Star) to confirm that the correct region had been amplified and identify in which orientation the insert had been incorporated into the plasmid. This determined which restriction enzyme and polymerase was used for each gene to produce antisense RNA probes.

2.2.5 In situ hybridisation probe synthesis

1µg of plasmid DNA was linearised by 2µl of the appropriate enzyme (Table 3) with 10µl buffer and 85µl dH_2O at 37°C for 1.5-2 hours.

Plasmid	Linearisation enzyme	Buffer	Polymerase
pGEM-T Easy fos	Nco1	H	SP6
pGEM-T Easy fosl1	Sal1	H	T7
pGEM-T Easy frzb	Apa1	A	SP6
pGEM-T Easy rasl11b	Nco1	H	SP6

Table 3: Enzymes, buffers and polymerases used for in situ hybridisation probe synthesis. Enzymes and buffers used to linearise *fos, fosl1, frzb and rasl11b* cDNA ligated into pGEM-T Easy plasmids. Polymerases used to generate RNA probes following linear plasmid purification.

Linear plasmids were made up to 400µl with dH_2O and 400µl phenol chloroform added. Samples were vortexed for 15 seconds then centrifuged at 13,000rpm for 5 minutes. The aqueous phase was transferred to a new Eppendorf and an equal volume of chloroform added. Samples were vortexed for 15 seconds, centrifuged at 13,000rpm for 5 minutes and the aqueous phase transferred to a new tube. DNA was precipitated with 30µl 3M sodium acetate, 600µl ethanol and 1µl GlycoBlue (Invitrogen) at -20°C overnight. Samples were centrifuged at 13,000rpm for 20 minutes at 4°C. 200µl ice cold 70% ethanol was added to the pellet before being vortexed and centrifuged at 13,000rpm for 5 minutes at 4°C. The pellet was dried by desiccation and resuspended in 23µl dH_2O.

Digoxigenin (DIG) labelled antisense RNA probes for in situ hybridisation (Table 4) were synthesised using 4µl 5x transcription buffer (NEB), 2µl 10x DIG NTP mix (Roche), 2µl 100mM DTT (Invitrogen), 1µl polymerase (Ambion), 3µl DNA, 7µl dH_2O at 37°C overnight. Plasmid DNA was degraded by incubating the reaction mixture with 1µl RNase-free DNase I (Promega) at 30°C for 20 minutes. Probes were precipitated overnight at -20°C with 50µl 5M ammonium acetate, 300µl ethanol, 50µl dH_2O and 1µl GlycoBlue (Invitrogen). Samples were centrifuged at 13,000rpm for 15 minutes at 4°C. 100µl ice cold 70% ethanol was added to the pellet before being vortexed and centrifuged at 13,000rpm for 5 minutes at 4°C. The pellet was dried by desiccation and resuspended in 50µl dH_2O.

Gene	Probe sequence
fos	CCAGATCTTCAGTGGCTTGTACAGCCAACCCTTATTTCTTCTGTAGCCCCATCACAGTCTCGGGCACACCCTTATGGGTCCACACC AGCTTACAGCCGATCTAGCGTTATGAAAGGATCTGCTGGAAGAGGTCAGAGCCTGGGAAGAAGAGGAAAAATGGAGCAGCTTT CTCCagaagaagaagaaaaaaggaaagtaagacgagaaaggaataagaTGGCAGCTGCCAAGTGTCGTAACCGCCGTCGGGAGTTAA CAGACACCCTTCAAGCGGAGACTGATGACCTGGAGGACCAGAAATCTGCCCTGCAGGCAGAGATTGCCGGCCTTCTAAAGGAG AAGGAAAAGCTGGAGTTTATACTTGCAGCTCACAAACCAGCTTGCAAAATTCCACATGATCTTGATGGAGCTTTTCAAGACTTGAC CTCATCTCTTGATCTGGGTCTGATTTCAGAGACCCCTTGTTCTTCCAGCTCTCAGGAGCCTGTAGCAGAGCCTCTGTTTCCCATTG GCCTTTCTCAGTCTTCCATGCCTGAAAAGGAGAACACCCACTGCAAGTCTCTATGGAACTCAAATCTGAACCACTGGATGATTTTC TGTTTAACTCTTCTCACACAGGTGTAACTGATGCAGCACGTTCTGTGCCAGATGTAGATCTTACTAGCTCTCTTTACACATCAGAAT GGGAGCCACTGTATAG
fosl1	TCGCAAGGAGCTAACAGATTACCTGCAGGCAGAGACAGACAAACTTGAAGAAGAGAAGTCATCCCTCCAGAAAGAAATTGCTGA GCTGCAGAAGCAGAAGGATAAGCTGGAACTCATCCTTGAGGCTCACCAGCCTATATGCAAGTTTCCTGACTCCCATCACAACATG CAGCAAGTGGACTCCTCCAGGCTGGTTAAGAAGGAACCACATGAAGAGTCACCCAGGGGACCTAAAGTCAACCTTCCCAGGATA GAGCTGAGCGACACAATCCTAGAGCCAGAGGCCCTTCACACCCCAACACTCATGAAGACACCATCCATTACTCCTTTTACGCCAA ATTTGATATTCACTTATCCTGGTCCACAAGAATCATGTTCTACAGCGCACCGAAGGCTGAGCAGAAGCAGCAGCAGTGGTAGTAG TGGAGAGCAAAGTCCATG
frzb	GGCTGGTGCTCCTATCATTACCCAACGCTCACTGTGCTTCATGTGAGCCTGTGCGGATTCCCATGTGCAAATCTATGCCATGGAAC ATGACTAAAATGCCCAACCATCTCCACCACAGCACTCAAGCCAATGCCATTTTGGCAATTGAACAGTTTGAAGGTTTGCTGACCAC TGAATGTAGCCAGGACCTTCTGTTCTTTCTGTGTGCTATGTATGCCCCCATTTGTACCATCGATTTCCAGCACGAGCCAATTAAGCC TTGCAAATCTGTATGTGAAAGGGCCAGGGCCGGCTGTGAGCCCATTCTCATCAAGTACCGGCACACTTGGCCAGAGAGCCTGGC ATGTGAGGAGCTCCCCGTATATGACAGAGGAGTCTGCATCTCCCCAGAGGCTATCATCACGGTGGAACAAGGAACAGATTCGAT GCCAGACTTCCCCATGGATTCAAACAACGGAAATTGTGGAAGCACGGCAGGGGAGCACTGTAAATGCAAGCCCATGAAGG
rasl11b	ATCTTGCCATCCAGGTTCAGGACACACCAGGAGTACAGATAAATGATCAGAATCTGGACTCTAATGAGCAGCTCAACAAATCCCT GAGATGGGCCGATGCTGTTGTGATCGTGTTCTCCATCACAGACTGTAAAAGCTTTGATCTTATCAGTCGCCTGCACCAGCACGCC CGGCAGCTTCACCCCGATAACAGAATCCCTATTGTCATTGTCGCCAATAAAGCAGATCTGCTGCACCTGAAACAGGTGGAACCAC AGCATGGACTTCAGCTGGCCAACATGTTGGGTTGCACTTTCTATGAAGTGAGTGTTAGCGAGAACTATATCGATGTGTACAATGC TTTCCAGGTACTGTGTAAAGAAATCAGCAAGCAGCAGAACACAGGAACCCCTGAAAGGCGGAAAAACTCGCTTATTCCACGTCCA AAGTCACCAAACATGCAGGA

Table 4: Digoxigenin labelled antisense RNA probe sequences for in situ hybridisation of *fos, fosl1, frzb and rasl11b.*

2.2.6 In situ hybridisation

For in situ hybridisation (Harland, 1991), MEMFA fixed de-membraned embryos stored in ethanol at -20°C were rehydrated through a graded series of ethanols and washed in PBS with 0.1% Tween (PBSAT). Embryos were then treated with 10µg/ml proteinase K at room temperature for 2-13 minutes depending on stage of development. They were then washed for 5 minutes in 0.1M triethanolamine pH 7.8 twice before two additions of 12.5µl acetic anhydride. Embryos were washed in PBSAT, re-fixed in 10% formaldehyde in PBSAT for 20 minutes, and washed in PBSAT again. Incubation in pre-hybridisation buffer (50% formamide, 5x SSC pH7, 100µg/ml heparin, 1x Denhart's, 0.1% Tween, 0.1% CHAPS, 10mM EDTA) at 60°C on a horizontal tube rocker for 2 hours was carried out before hybridisation with 3µl/ml DIG labelled antisense RNA probes at 60°C overnight. Embryos were kept above 60°C and washed in hybridisation buffer twice for 10 minutes, 2x SSC + 0.1% Tween three times for 20 minutes and 0.2x SSC + 0.1% Tween three times for 30 minutes. Two 15 minute MABT (100mM maleic acid, 150mM NaCl, pH 7.8, 0.1% Tween) washes were then carried out at room temperature. Embryos were pre-incubated in 1ml MAB + 2% BMB + 20% heat treated lamb serum at 60°C on a horizontal tube rocker for 2 hours. The solution was

replaced with fresh solution with the addition of a 1/2,000 dilution of anti-DIG antibody coupled to alkaline phosphatase (Roche) and rolled at 4°C overnight. Embryos were washed in MABT before a 3 minute and 10 minute wash in alkaline phosphatase buffer (100mM Tris pH 9.5, 50mM $MgCl_2$, 100mM NaCl, 0.1% Tween). Staining was visualised by addition of BM purple (Roche). Embryos were bleached in 5% H_2O_2 in PBSAT to aid gene expression pattern visualisation.

2.2.7 Immunostaining

For immunostaining (Christen and Slack, 1999), MEMFA fixed embryos stored in ethanol at -20°C were rehydrated through a graded series of ethanols and washed in PBS. Embryos were then treated with $K_2Cr_2O_7$ in 5% acetic acid at room temperature for 40 minutes. They were then washed in PBS before being bleached for 45 minutes in 5% H_2O_2 in PBS and washed in PBS again. Embryos were blocked in BBT (PBS, 1% BSA, 0.1% Triton X-100) + 5% horse serum for 1 hour before incubation in a 1/10,000 dilution of dpERK1+2 monoclonal mouse antibody (Sigma) or 1/5,000 dilution of αmyc 9E-10a monoclonal mouse antibody (Cell Signaling Technology) at 4°C overnight. Embryos were washed in BBT then BBT + 5% horse serum for 1 hour before incubation with BBT + 5% horse serum with a 1/1000 dilution of horse anti-mouse igG-AP conjugated secondary (VectorLab) overnight at 4°C. Immunostaining was visualised by addition of BM purple (Roche). If necessary, embryos were bleached in 5% H_2O_2 in PBS again to better visualise staining.

2.2.8 Western blots

Five *Xenopus laevis* embryos collected at gastrula stage 10.5, flash frozen and stored at -80°C were homogenised in 50µl Phosphosafe buffer (Novagen) and centrifuged at 13,000 rpm for 1 minute. The supernatant was added to 50µl of sample buffer (120mM Tris/Cl pH6.8, 20% glycerol, 4% SDS, 0.04% bromophenol blue, 10% β-mercaptoethanol) and heated at 90°C for 5 minutes. 20µl of sample was loaded per lane of 10% SDS-PAGE gel in addition to 10µl PageRuler prestained protein ladder (ThermoScientific) and run at 180V for 2 hours. Proteins were transferred to Immobilon-P Transfer Membrane (Millipore) by electroblotting wet transfer in 10% methanol transfer buffer at 85V for 2 hours. Membranes were washed in PBSAT and blocked in PBSAT + 5% milk powder (blocking solution) for 1 hour at room temperature. Membranes were then transferred to fresh blocking solution containing primary antibodies (dpERK at a 1:4,000 dilution) and left overnight at 4°C. Membranes were washed in PBSAT and blocked for 30 minutes at room temperature before transfer to blocking solution containing secondary antibody (anti-mouse diluted to 1:4,000) for 1 hour at room temperature. After further PBSAT washes, BM Chemiluminescence Blotting Substrate kit (Roche) and ECL Hyperfilm (Amersham) were used to detect proteins. Membranes were stripped

twice for 5 minutes in stripping buffer (137mM NaCl, 20mM glycine, pH2.5) and re-probed for total ERK (1:500,000) with secondary anti-rabbit antibody (1:2,000).

2.3 Data analysis

2.3.1 Differential expression analysis of RNA-seq data

Xenopus tropicalis mRNA samples of CIC knockdown by TALENs, water injected and FGF4 overexpressing embryos were collected in triplicate by Michael King, a PhD student in the Isaacs' lab (King, 2019). RNA was analysed using the bioanalyzer to confirm RNA quality. Library preparation and Illumina sequencing was carried out by Bioscience Technology Facility staff at the University of York. A single lane of the Illumina HiSeq 2000 platform was utilised for sequencing samples. First strand cDNA synthesis was carried out using 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 found 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 to take into account the varying number of reads sequenced for each sample, and the varying expression of transcripts of the whole transcriptome. For example, the TPM value for 'transcript A' should represent the number of transcripts of A that would be observed if one million transcripts were sequenced from the whole transcriptome.

2.3.2 Gene ontology analysis of RNA-seq data

Initial RNA-seq bioinformatics analysis was performed at the University of York Bioscience Technology Facility. Raw reads for each sample were aligned to the *Xenopus tropicalis* reference transcriptome (genome v9.1) (http://www.xenbase.org/common/displayJBrowse.do?data=data/xt9_1) using Salmon (http://salmon.readthedocs.io) to produce estimated read counts for each transcript for each sample. Sleuth (http://pachterlab.github.io/sleuth/) was then used to calculate differential expression (Q values and effect sizes) by fitting a statistical model to the estimated read counts.

43,558 transcripts were found for 23,635 genes. Transcripts with a q value < 0.05 and effect size > 1.5 in CIC knockdown and/or FGF4 overexpression were selected for gene ontology (GO) analysis using the Protein ANalysis THrough Evolutionary Relationships (PANTHER) classification tool (Mi et al., 2013) (http://www.pantherdb.org/). Genes were classified according to the PANTHER GO-Slim subset of GO terms. The PANTHER GO-Slims are GO terms identified by GO phylogenetic annotation and expert review as both informative of function and evolutionarily conserved (Mi et al., 2019).

2.3.3 Motif enrichment

MEME suite 5.0.5 was used to determine whether CIC binding sites were enriched at the genomic locus of genes potentially regulated by CIC and FGF (http://meme-suite.org/tools/ame). Analysis of motif enrichment was carried out using Fisher's exact test with the average odds score used for sequence scoring. 1004 control sequences were generated by randomly shuffling the input sequences conserving frequencies of word size 1. The E value threshold for reporting enriched motifs used was $E \leq 10$.

2.3.4 ImageJ analysis of myc immunostaining

Stained areas of equal size for 5 embryos per treatment were converted to 16-bit grayscale and each image duplicated to create an identical copy. The threshold in the copy was adjusted to create a binary image with stained areas to be measured highlighted. Measurements were set to redirect to the un-thresholded version and highlighted particles of the binary image analysed.

2.3.5 ImageJ analysis of western blots

Western blot lanes were selected and profile plots of each generated using ImageJ. A straight line was drawn along the base of each peak to enclose the peak and exclude background noise. The area of each peak was then measured and the relative density calculated by dividing each by the corresponding loading control peak area and then normalised un-injected control peak area.

Chapter 3: Transcriptomic analysis of FGF4 overexpression and Capicua knockdown in *Xenopus tropicalis*

3.1 Introduction

As discussed in the introduction (Chapter 1), we hypothesise that transcription of a subset of FGF target genes relies on ERK mediated relief of CIC transcriptional repression. If CIC is responsible for repressing gene transcription in the absence of FGF signalling, disruption of the CIC gene should have a similar effect on the resulting transcriptome as an overexpression of FGF4. Consequently, an RNA-seq experiment was carried out by a PhD student in the Isaacs' lab (King, 2019) in triplicate on control, FGF4 overexpressing, and CIC knockdown (via transcription activator-like effector nucleases (TALENs)) *Xenopus tropicalis* embryos.

TALENs are composed of a non-specific DNA-cleaving nuclease (*FokI*) fused to a TAL effector DNA-binding domain capable of being engineered to target a specific sequence (Lei et al., 2012). A pair of forward and reverse TALENs sequences are used to bind effector binding element upstream and downstream of the target region. This can be utilised for genome editing by inducing double-strand breaks in target DNA sequences, which cells respond to with error-prone repair mechanisms such as non-homologous end joining (NHEJ) (Moore and Haber, 1996) leading to production of non-functional gene products or nonsense mediated decay of the mRNA. In this case, TALENs were designed to target the HMG-box domain of CIC as this is the site responsible for binding octameric T(G/C)AATG(A/G)A sequences in target gene enhancers and promoters (Jimenez et al., 2012). Chromosomal rearrangement or introduction of nucleotide insertions or deletions following NHEJ in this region should result in any CIC protein produced being unable to function. In *Xenopus*, the HMG-box coding region spans exon 6 and 7. Consequently, the forward and reverse TALENs used were designed to flank and target exon 6 in order to induce double-strand breaks in this region. 91.8% (74/81) of *Xenopus tropicalis* embryos injected with a total of 1ng forward and reverse TALENs mRNA exhibited a mutant phenotype ranging in severity from reduced head/eye pigmentation or cyclopia (50.8%) to total head loss (16.1%) (King, 2019). Most embryos also displayed an enlarged proctodaeum. Sequencing of DNA extracted from 8 of the TALENs injected embryos at late tailbud stage (40-41) revealed a 100% targeting efficiency with indels successfully introduced adjacent to the target site (King, 2019). The range of severity of phenotype is due to the mosaicism associated with TALENs targeting. The TALENs system requires production of the TALEN proteins from injected mRNA to target and knockdown CIC. Although the TALEN mRNAs are translated as soon as they are injected, the protein levels take a while to accumulate to the level

necessary for targeting to occur. As the protein is accumulating, cell division is already occurring so multiple independent targeting events may occur in different cells leading to a mosaicism of daughter cells within the embryo. Additionally, TALENs can disrupt both alleles of a gene leading to production of 2 different mutations following NHEJ in somatic cells. Indel mutations are usually located in the spacer region between the TALENs effector binding elements so the TALENs pair may bind again, thus introducing additional mutations during development (Lei et al., 2012).

Until the midblastula transition (MBT), the zygotic genome is quiescent, and development is controlled by maternal factors within the cytoplasm (Lee et al., 2014). Consequently, there is no transcription of TALENs targeted loci until after the MBT, and any maternally deposited CIC mRNA is wild type. To ensure that very early development was not affected in either set of embryos, FGF4 was injected into embryos in a CSKA plasmid construct with a β-actin promoter which does not become active until after the MBT (Isaacs et al., 1994). Embryos were collected at neurula stage 14 for RNA-seq analysis using Illumina hi-seq. The abundance of 43,558 transcripts for 23,635 genes was analysed for 3 batches of FGF4 injected, CIC knockdown and water injected embryos.

The aims of this chapter are:

- Analyse RNA-seq data to identify genes upregulated and downregulated by FGF4 overexpression and CIC inhibition
- Undertake gene ontology enrichment analysis in order to identify biological processes associated with upregulated genes
- Establish the statistical significance of the overlap in genes upregulated and downregulated by both FGF4 overexpression and CIC inhibition
- Select putative FGF and CIC regulated target genes for validation of RNA-seq data
- Determine if target genes are expressed in regions of CIC and FGF expression
- Validate RNA-seq data using RT-PCR with primers specific to target genes
- Analyse level of enrichment of CIC binding sites around the genomic loci of target genes

3.2 Results

3.2.1 Changes in gene expression as a result of CIC inhibition and FGF4 overexpression

Initial quality control of the RNA-seq data and mapping of sequences to the *Xenopus tropicalis* transcriptome was undertaken by John Davey at the University of York Technology Facility. Data was then sorted to identify annotated transcripts with statistically significant changes in expression

by assessing calculated q value and effect size for each transcript. A q value is a measure of statistical significance of differential expression which allows for false positives. The smaller the value, the more significant the change in expression and the fewer genes expected to be false positives. Effect sizes are calculated by linear models representing the relative change in expression between treated and control samples. An effect size of 1 indicates no change in gene expression compared to the control group after any batch effects have been taken into account. An effect size greater than 1 indicates an upregulation of gene expression and an effect size less than 1 represents a downregulation of gene expression. For example, an effect size of 2 represents a 2 times change in gene expression and an effect size of 0.5 indicates a 0.5 times change in gene expression. 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. Out of 43,558 transcripts, 331 and 81 fit these criteria for upregulated genes in CIC knockdown and FGF4 overexpressing embryos respectively (Table 5). 100 transcripts had a q value > 0.05 and an effect size < 0.75 in CIC knockdown embryos and 92 in FGF4 overexpressing embryos (Table 6). The complete gene transcript list from this RNA-seq experiment is available on Google Drive (https://drive.google.com/open?id=13gbNFwAbD00BSEWvahELt1lYPjyTGHv7).

Transcript ID	Gene	CIC knockdown		FGF4 injection	
		q value	Effect size	q value	Effect size
NM_001078836.1	apold1	0.004	4.69	0.002	5.45
NM_001079231.1	arrdc2	$5.75x10^{-7}$	2.56	$3.10x10^{-6}$	2.47
XM_012965809.2	arrdc2	$2.17x10^{-4}$	2.05	$2.50x10^{-6}$	2.44
XM_002934698.4	atf3	$3.91x10^{-4}$	1.96	0.014	1.76
XM_018091031.1	b4galt1.1	0.030	2.21	0.018	2.44
XM_012970501.2	bmp7.2	$2.16x10^{-5}$	25.23	0.001	15.07
XM_012964062.2	c4bpa	0.041	2.17	0.008	2.63
NM_001102857.1	cbx4	0.028	2.22	0.001	2.87
XM_004916833.3	chic1	0.003	1.83	0.041	1.67
NM_203542.1	cldn6.1	0.003	1.85	0.015	1.78
XM_002931681.4	exoc3l1	0.005	5.81	$8.94x10^{-5}$	9.97
XM_012952528.1	fam83c	0.007	1.89	0.001	2.12
NM_001045662.1	fam83c	0.040	1.80	0.026	1.93
XM_004911187.3	fat1	$1.41x10^{-4}$	1.96	0.036	1.63
XM_004914180.3	fgd3	0.024	20.40	0.045	20.21
NM_001016200.2	fos	$1.42x10^{-12}$	4.43	$7.15x10^{-17}$	5.38
XM_002939331.4	fosl1	0.005	3.09	0.015	3.00
NM_001005438.1	frzb	0.001	4.49	0.045	3.21
XM_012961068.2	gadd45a	0.003	1.58	0.027	1.51
NM_001142145.1	gpcpd1	0.007	2.49	0.018	2.45
XM_002936205.4	htr1b	$5.14x10^{-6}$	24.11	0.027	7.75
XM_004919807.3	ier3	$4.83x10^{-5}$	2.63	0.001	2.33
XM_002932276.4	LOC100485153	0.041	11.79	0.004	24.59
XM_012962310.2	LOC100486038	0.019	2.24	0.003	2.66
XR_001924462.1	LOC101730746	0.041	1.80	0.017	1.99
XM_012953285.1	LOC101731310	$7.19x10^{-5}$	28.53	0.025	11.53
XM_004914317.3	LOC101732940	$5.75x10^{-7}$	4.61	0.001	3.24
XM_018090681.1	LOC101733948	0.008	2.22	0.001	2.68
XM_018090141.1	LOC105945708	0.048	1.96	0.001	2.75
XR_001170914.1	LOC105947461	0.001	2.11	$7.13x10^{-5}$	2.44
NM_001030330.1	mmp1	$2.65x10^{-4}$	8.19	0.043	4.65
XM_004915576.3	mst1	0.040	1.68	0.009	1.89
NM_001130266.1	nfkbiz	0.006	2.13	0.027	2.02
NM_001015693.1	pnpla3	0.002	1.64	0.008	1.60
NM_001015774.1	rasl11b	$3.45x10^{-7}$	2.60	0.001	1.99
XM_018091191.1	sat1	0.032	1.73	0.017	1.86
NM_001007996.1	sat1	$1.44x10^{-4}$	1.64	0.001	1.59
XM_012963073.2	sgk1	$2.26x10^{-4}$	2.45	$7.13x10^{-5}$	2.66
NM_001030422.1	sgk1	0.002	2.04	$1.86x10^{-5}$	2.51
NM_001097368.1	sox17b.2	$6.00x10^{-4}$	1.66	0.016	1.53
XM_012960662.2	tmcc1	0.005319	14.59	0.032	11.36
NM_001142914.1	tmcc1	0.014	1.59	0.018	1.62
NM_001142050.1	usp2	$1.74x10^{-5}$	3.48	0.045	2.22
NM_001017208.2	wnt8a	0.041	1.83	0.004	2.20

Table 5: Gene transcripts upregulated in both CIC knockdown and FGF4 injected *Xenopus tropicalis* embryos. Gene transcripts with an RNA-seq q value < 0.05 and effect size > 1.5 are classed as upregulated. Embryos collected at neurula stage 14 for RNA-seq analysis.

Transcript ID	Gene	CIC knockdown		FGF4 injection	
		q value	Effect size	q value	Effect size
XM_018095307.1	atp2a2	0.047	0.741	0.006	0.680
NM_001127035.1	axl	0.014	0.567	0.030	0.571
NM_001006762.1	cebpa	0.017	0.363	0.001	0.277
XM_012960859.1	celsr2	0.004	0.653	0.025	0.677
NM_001011044.1	cygb	0.024	0.545	0.005	0.481
NM_001126689.1	dpysl3	0.011	0.588	0.043	0.609
NM_001006869.1	efnb3	0.011	0.682	0.017	0.678
NM_001079128.1	irx3	$1.71x10^{-5}$	0.599	0.042	0.717
NM_001097188.1	LOC100485697	0.001	0.603	0.042	0.670
XM_002934275.4	msi1	$1.82x10^{-4}$	0.503	0.020	0.587
XM_012961432.2	nkain1	0.027	0.554	0.041	0.545
NM_001005637.1	notch3	$6.86x10^{-5}$	0.617	0.023	0.701
XM_004917080.2	pax6	0.018	0.316	0.009	0.273
XM_002939393.4	pax6	0.033	0.300	0.030	0.274
XM_002933915.4	ripply2.2	0.004	0.639	0.030	0.671
XM_002932136.4	s1pr5	0.016	0.702	0.020	0.691
NM_001113010.1	serpina1	0.016	0.670	0.030	0.668
XR_001923782.1	slc23a2	0.011	0.586	$2.41x10^{-5}$	0.459
NM_001127068.1	spib	0.005	0.183	0.001	0.137
XM_018092518.1	unc13d	$2.98x10^{-4}$	0.205	0.013	0.269
NM_001001216.1	znf219	0.001	0.493	0.045	0.583

Table 6: Gene transcripts downregulated in both CIC knockdown and FGF4 injected *Xenopus tropicalis* embryos. Gene transcripts with an RNA-seq q value < 0.05 and effect size < 0.75 are classed as downregulated. Embryos collected at neurula stage 14 for RNA-seq analysis.

3.2.2 Gene ontology enrichment analysis to aid identification of potential CIC and FGF regulated target genes

In order to investigate which biological processes genes upregulated in the RNA-seq data set are involved in, gene ontology (GO) enrichment analysis was carried out to allow classification of genes according to their functional characteristics. The PANTHER classification system (Mi et al., 2013) (www.pantherdb.org) was used to perform Fisher's exact statistical overrepresentation tests and calculation of false discovery rate (FDR). This analysis was carried out on all genes identified as significantly upregulated in the FGF4 overexpressing (Table 7) and CIC knockdown embryos (Table 8) respectively. Genes were classified according to the PANTHER GO-Slim biological process subset of GO terms which have been previously identified as both informative of function and evolutionarily conserved (Mi et al., 2019).

PANTHER GO-Slim Biological Process	Number of genes	Expected	Fold enrichment	+/-	Raw P value	FDR
Transmembrane receptor protein tyrosine kinase signalling pathway	5	0.31	16.35	+	$1.59x10^{-5}$	$1.27x10^{-3}$
ERK/MAPK cascade	8	0.81	9.9	+	$1.49x10^{-6}$	$1.79x10^{-4}$
Intracellular signal transduction	13	2.6	4.99	+	$1.28x10^{-6}$	$3.06x10^{-4}$
Developmental process	10	3.29	3.04	+	$1.39x10^{-3}$	$4.18x10^{-2}$
Signal transduction	16	5.63	2.84	+	$7.98x10^{-5}$	$4.79x10^{-3}$
Cell communication	16	6.42	2.49	+	$3.70x10^{-4}$	$1.48x10^{-2}$
Cellular process	32	18.38	1.74	+	$1.16x10^{-4}$	$5.58x10^{-3}$
Unclassified	13	25.69	0.51	-	$3.76x10^{-4}$	$1.29x10^{-2}$

Table 7: Biological processes associated with gene transcripts upregulated in FGF4 overexpressing *Xenopus tropicalis* embryos. Gene ontology processes identified using PANTHER Fisher's exact statistical overrepresentation test with false discovery rate (FDR). GO-Slim biological processes are a subset of gene ontology terms identified by the PANTHER classification system indicating the biological systems to which a protein contributes.

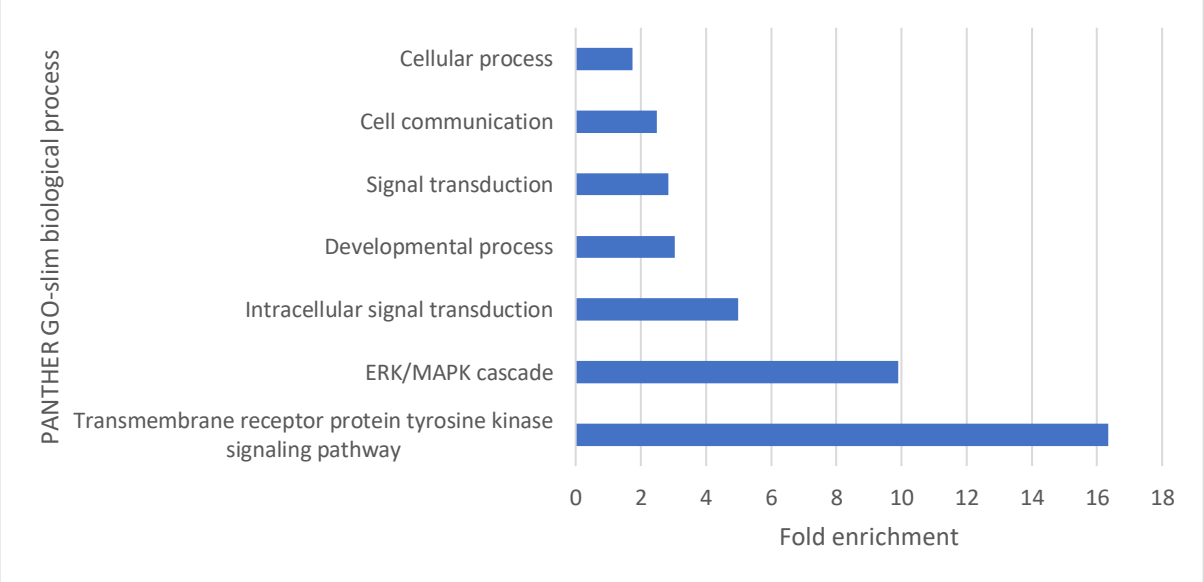

Figure 5: Bar chart showing fold enrichment for biological processes associated with gene transcripts upregulated in FGF4 overexpressing *Xenopus tropicalis* embryos. Gene ontology processes identified using PANTHER Fisher's exact statistical overrepresentation test with false discovery rate (FDR). GO-Slim biological processes are a subset of gene ontology terms identified by the PANTHER classification system indicating the biological systems to which a protein contributes.

PANTHER GO analysis identified 7 biological processes associated with genes significantly upregulated in FGF4 overexpressing embryos (Figure 5). The GO terms with the highest fold enrichment were the transmembrane receptor protein tyrosine kinase signalling pathway, ERK/MAPK cascade, intracellular signal transduction and developmental process. This further

supports the hypothesis that these genes are regulated by FGF as FGF4 typically acts via tyrosine kinase receptors to activate ERK cascades intracellularly to regulate multiple developmental processes.

PANTHER GO-Slim Biological Process	Number of genes	Expected	Fold enrichment	+/-	Raw P value	FDR
Cell proliferation	8	0.66	12.21	+	$6.39x10^{-7}$	$1.53x10^{-4}$
Regulation of cell cycle	11	2.22	4.95	+	$2.18x10^{-5}$	$1.31x10^{-3}$
Mitosis	8	2.16	3.71	+	$1.78x10^{-3}$	$3.89x10^{-2}$
Response to stress	17	5.00	3.40	+	$1.69x10^{-5}$	$1.35x10^{-3}$
Response to external stimulus	9	2.77	3.25	+	$2.24x10^{-3}$	$4.48x10^{-2}$
Cell death	10	3.39	2.95	+	$2.62x10^{-3}$	$4.50x10^{-2}$
Death	10	3.39	2.95	+	$2.62x10^{-3}$	$4.84x10^{-2}$
Cell adhesion	10	3.46	2.89	+	$2.98x10^{-3}$	$4.77x10^{-2}$
Biological adhesion	10	3.46	2.89	+	$2.98x10^{-3}$	$4.48x10^{-2}$
ERK/MAPK cascade	10	3.49	2.87	+	$3.15x10^{-3}$	$4.45x10^{-2}$
Regulation of phosphate metabolic process	14	5.37	2.61	+	$1.24x10^{-3}$	$3.30x10^{-2}$
Organelle organization	26	11.70	2.22	+	$2.10x10^{-4}$	$8.39x10^{-3}$
Intracellular signal transduction	24	11.23	2.14	+	$5.59x10^{-4}$	$1.92x10^{-2}$
Phosphate-containing compound metabolic process	31	16.32	1.90	+	$6.70x10^{-4}$	$2.01x10^{-2}$
Metabolic process	82	56.05	1.46	+	$1.30x10^{-4}$	$6.24x10^{-3}$
Cellular process	112	79.28	1.41	+	$7.87x10^{-6}$	$9.45x10^{-4}$
Unclassified	87	110.82	0.79	-	$1.43x10^{-3}$	$3.42x10^{-2}$

Table 8: Biological processes associated with gene transcripts upregulated in CIC knockdown *Xenopus tropicalis* embryos. Gene ontology processes identified using PANTHER Fisher's exact statistical overrepresentation test with false discovery rate (FDR). GO-Slim biological processes are a subset of gene ontology terms identified by the PANTHER classification system indicating the biological systems to which a protein contributes.

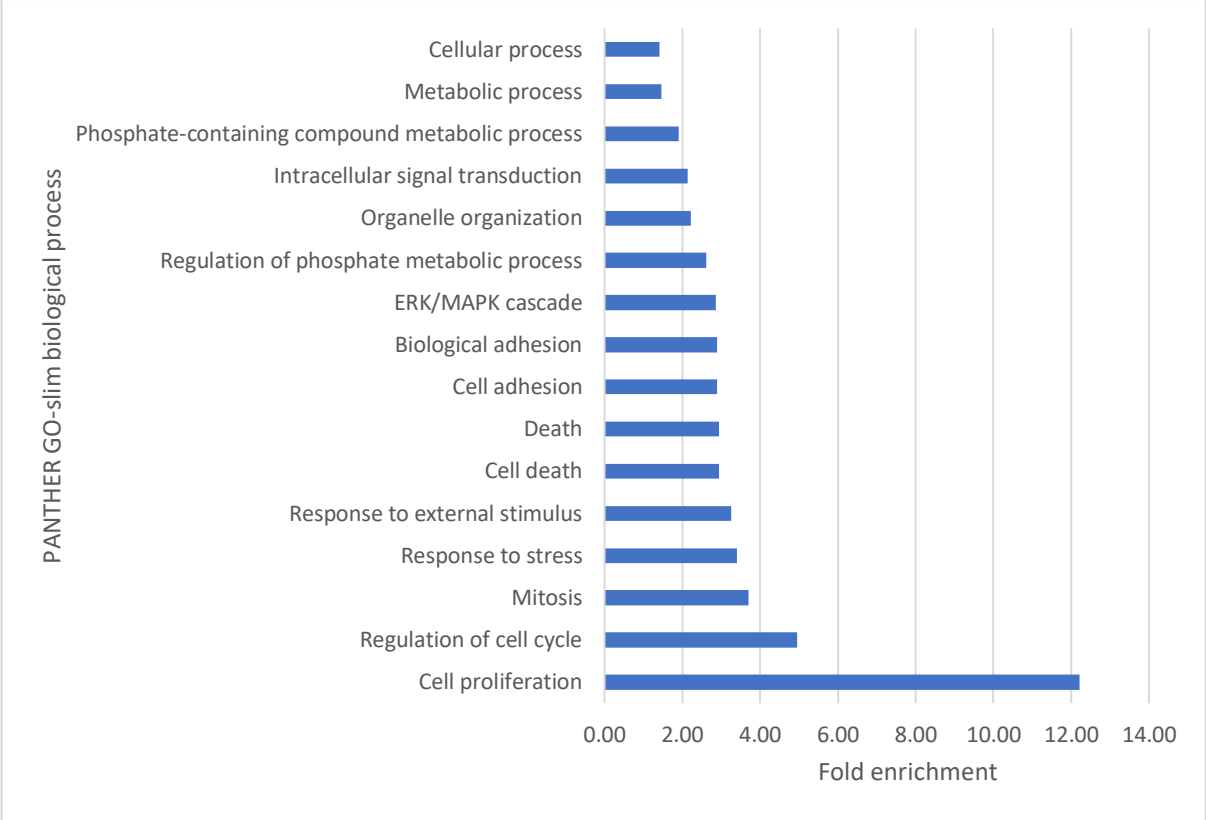

Figure 6: Bar chart showing fold enrichment for biological processes associated with gene transcripts upregulated in CIC knockdown *Xenopus tropicalis* embryos. Gene ontology processes identified using PANTHER Fisher's exact statistical overrepresentation test with false discovery rate (FDR). GO-Slim biological processes are a subset of gene ontology terms identified by the PANTHER classification system indicating the biological systems to which a protein contributes.

PANTHER GO analysis identified 16 biological processes associated with genes significantly upregulated in CIC knockdown embryos (Figure 6). The GO term with the highest fold enrichment was cell proliferation. This is encouraging as FGF is important in regulation of cell proliferation. Other genes were involved in the same processes as those upregulated in the FGF overexpression such as ERK/MAPK cascade, intracellular signal transduction and cellular process.

A strong overlap between significantly upregulated and downregulated gene transcripts was observed in CIC knockdown and FGF overexpressing embryos with 44 upregulated and 21 downregulated in both (Figure 7).

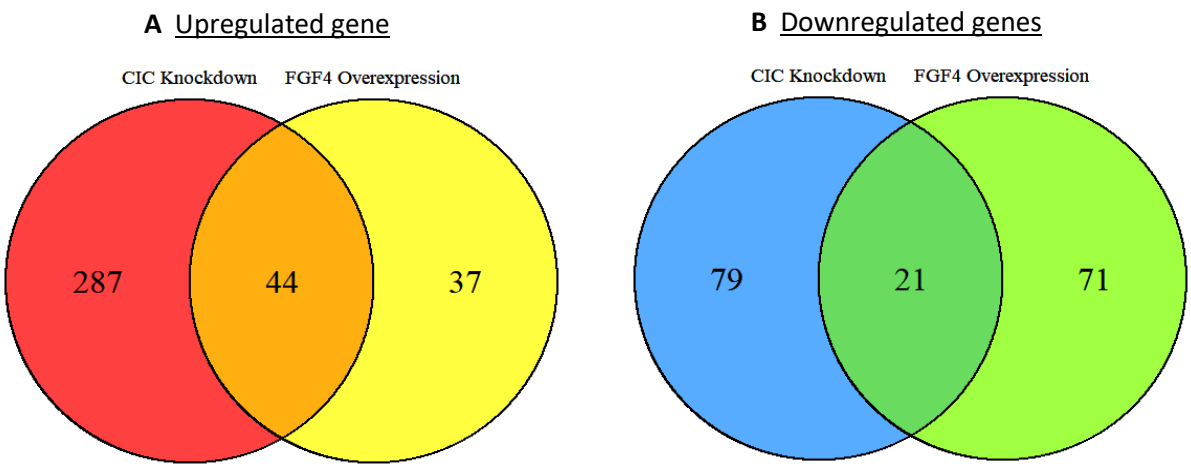

Figure 7: Venn diagrams showing the overlap between significantly upregulated and downregulated gene transcripts in CIC knockdown and FGF4 overexpressing *Xenopus tropicalis* embryos. (A) Upregulated genes: RNA-seq q value < 0.05 and effect size > 1.5 **(B)** Downregulated genes: RNA-seq q value < 0.05 and effect size < 0.75. Embryos collected at neurula stage 14.

3.2.3 Establishing the statistical significance of observed gene overlaps

In order to investigate the probability of the observed overlaps, Python was used to randomly sample sets of the number of upregulated genes (331 and 81) from numbers between 1 and 23,635 (number of genes in the *Xenopus tropicalis* genome) and the size of the overlap between the 2 sets recorded (Figure 8a). The highest number of overlaps in 100,000 runs was 8 so the probability of getting an overlap of size 9 or greater is $p < 0.00001$. Consequently, an overlap of 44 upregulated genes is statistically significant.

Sets of 100 and 92 numbers were randomly sampled for downregulated genes (Figure 8b). The highest number of overlaps in 100,000 runs was 5 so the probability of getting an overlap of size 6 or greater is $p < 0.00001$. Consequently, an overlap of 21 significantly downregulated genes is statistically significant.

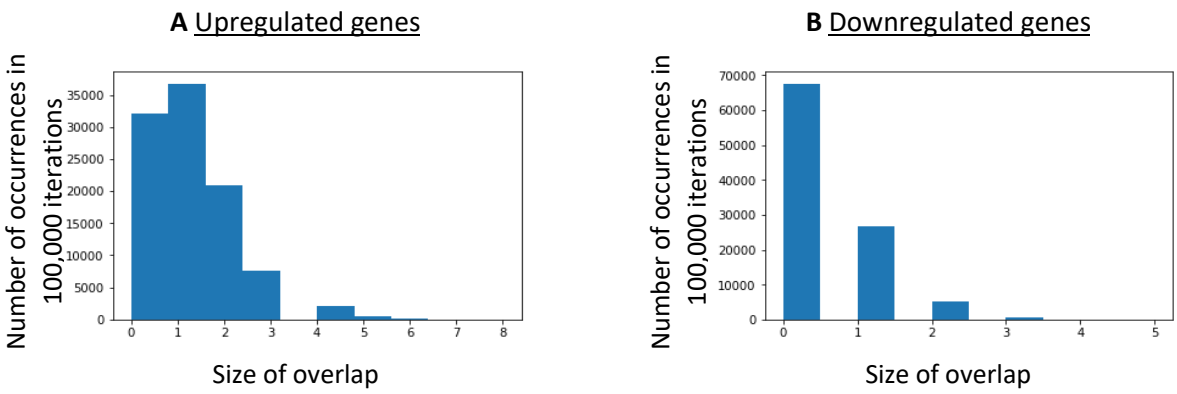

Figure 8: Histograms showing expected size of overlaps modelling RNA-seq data using Python. (A) Upregulated genes: 331 gene transcripts had an RNA-seq q value < 0.05 and an effect size > 1.5 in *Xenopus tropicalis* embryos injected with FGF4, and 81 gene transcripts in CIC knockdown embryos with an overlap of 44. Histogram shows size of overlaps generated by random sampling of sets of 331 and 81 numbers between 1 and 23,635 (number of genes in the *Xenopus tropicalis* genome). **(B)** Downregulated genes: 100 gene transcripts had an RNA-seq q value < 0.05 and an effect size < 0.75 in *Xenopus tropicalis* embryos injected with FGF4, and 92 gene transcripts in CIC knockdown embryos with an overlap of 21. Histogram shows size of overlaps generated by random sampling of sets of 100 and 92 numbers between 1 and 23,635 (number of genes in the *Xenopus tropicalis* genome).

3.2.4 Putative CIC and FGF regulated target genes selected for validation of RNA-seq data

Putative CIC and FGF regulated target genes were selected from those upregulated in both the CIC knockdown and FGF4 overexpressing embryos for use in validation of the RNA-seq data. To select candidate genes, the 44 genes were sorted by decreasing effect size and 4 genes involved in relevant processes selected working down from the top of this list (Table 9).

frzb was selected because of a high effect size, low q value and the fact that it is an antagonist for *wnt* (Kawano and Kypta, 2003), a gene crucial in many developmental processes. *frzb* is also known to be positively regulated by FGF signalling (Branney et al., 2009). *fos* and *fosl1* were identified as genes of interest due to high effect sizes and the fact that they encode transcription factors that bind Jun family members to form activator protein-1 (AP-1) heterodimers (Angel and Karin, 1991). AP-1 mediates FGF and BMP signalling during Xenopus development (Lee et al., 2011). *rasl11b* was chosen as a candidate gene due to a relatively high effect size and low q value out of those identified by gene ontology analysis to be involved in the ERK/MAPK cascade.

Gene	Description	CIC knockdown		FGF4 injection	
		q value	Effect size	q value	Effect size
frzb	wnt antagonist	$8.35x10^{-4}$	4.49	0.0447	3.21
fos	Transcription factor that binds Jun family members to form AP-1 heterodimers which mediate FGF and BMP signalling	$1.42x10^{-12}$	4.43	$7.15x10^{-17}$	5.38
fosl1	Transcription factor that binds Jun family members to form AP-1 heterodimers which mediate FGF and BMP signalling	0.00532	3.09	0.0154	3.00
rasl11b	GTPase involved in MAPK cascade	$3.45x10^{-7}$	2.60	0.00145	1.99

Table 9: Candidate genes selected from 44 gene transcripts upregulated in both FGF4 overexpressing and CIC knockdown *Xenopus tropicalis* embryos for use in validation of the RNA-seq data set. Genes selected due to relatively high effect sizes, low q values and involvement in relevant biological processes. Genes ordered by CIC knockdown effect size.

3.2.5 Candidate gene expression in wild type *Xenopus tropicalis* embryos

Having identified genes potentially regulated by both CIC and FGF, gene specific in situ hybridisation probes were designed and synthesised to determine if the genes are expressed in regions relevant to CIC expression and FGF signalling in normal *Xenopus tropicalis* development.

FGF is the sole activator of ERK in early *Xenopus* development (Branney et al., 2009). dpERK can therefore be used to indicate the presence of FGF signalling. For example, dpERK is present in the early mesoderm surrounding the blastopore at gastrula stage and is particularly enriched dorsally (Figure 9a). ERK is also later activated along the dorsal midline in neural tube formation (Figure 9b).

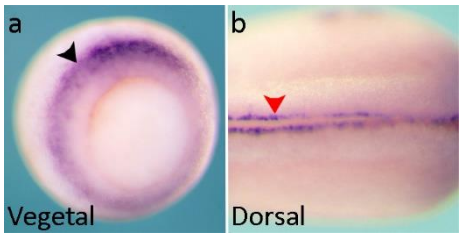

Figure 9: The spatial expression pattern of dpERK in *Xenopus* analysed using whole mount immunohistochemistry with an antibody against dpERK. (a) gastrula stage 10.5 *Xenopus tropicalis*, vegetal view, (b) neurula stage 18 *Xenopus laevis*, dorsal view, anterior to left. Black arrow indicates early mesoderm and red arrow indicates the neural folds fusing together during neural tube formation.

In situ hybridisation for *frzb, fos, rasl11b* and *fosl1* was carried out at a range of stages of *Xenopus tropicalis* development (Figures 10-13) to visualise their spatial expression patterns.

Distinct dots of *frzb* expression can be seen at one side of the vegetal pole at late blastula stage (Figure 10). *frzb* expression then expands to form a ring around the blastopore during early gastrulation with higher levels of expression dorsally. By gastrula stage 11, *frzb* is only expressed in the involuting prechordal mesoderm. At early neurula stage 14, *frzb* is expressed in the posterior presomitic mesoderm on both sides of the dorsal midline.

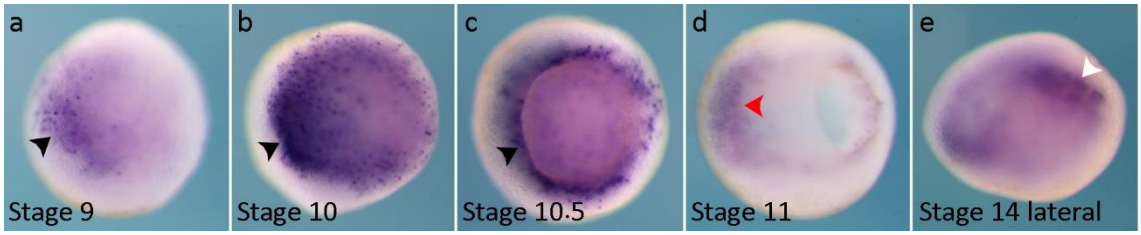

Figure 10: The spatial expression pattern of *frzb* during early *Xenopus tropicalis* development analysed using in situ hybridisation. (a) late blastula stage 9, vegetal view (b) early gastrula stage 10, vegetal view (c) gastrula stage 10.5, vegetal view (d) gastrula stage 11, vegetal to right, (e) early neurula stage 14, lateral view, anterior to left. Black arrows indicate dorsal early mesoderm, red arrow indicates prechordal mesoderm, and white arrow indicates presomitic mesoderm. Embryos staged according to Nieuwkoop and Faber (1994) stages of *Xenopus* development.

At late blastula stage, *fos* is expressed in a ring around the equator of the embryo when viewed from the animal hemisphere (Figure 11). During gastrulation, fos is only expressed in the vegetal pole in early mesoderm around the blastopore. Expression increases during early gastrulation in a ring around the blastopore with particularly enriched expression dorsally. By gastrula stage 11, *fos* expression has become restricted to the dorsal blastopore. At early neurula stage 14, *fos* is expressed along the dorsal midline where the neural tube is forming.

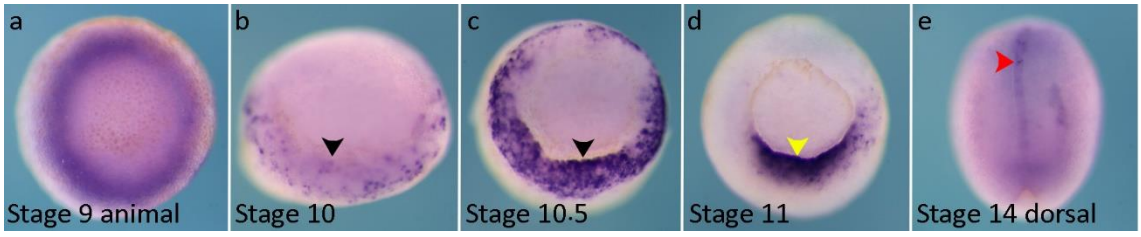

Figure 11: The spatial expression pattern of *fos* during early *Xenopus tropicalis* development analysed using in situ hybridisation. (a) late blastula stage 9, animal view, (b) early gastrula stage 10, vegetal view (c) gastrula stage 10.5, vegetal view, (d) gastrula stage 11, vegetal view, (e) early neurula stage 14, dorsal view, anterior to top. Black arrows indicate dorsal early mesoderm, yellow arrow indicates the dorsal blastopore and red arrow indicates forming neural tube. Embryos staged according to Nieuwkoop and Faber (1994) stages of *Xenopus* development.

At late blastula stage, *rasl11b* is weakly expressed in a ring around the equator of the embryo when viewed from the animal hemisphere (Figure 12). Rasl11b is then expressed dorsally in the vegetal pole. At gastrula stage 10.5, expression becomes more widespread but remains highest at the dorsal blastopore. By stage 11, *rasl11b* expression forms a neat ring around the blastopore in the

early mesoderm. From neurula stage, *rasl11b* is expressed in the otic vesicles, posterior presomitic mesoderm and proctodaeum.

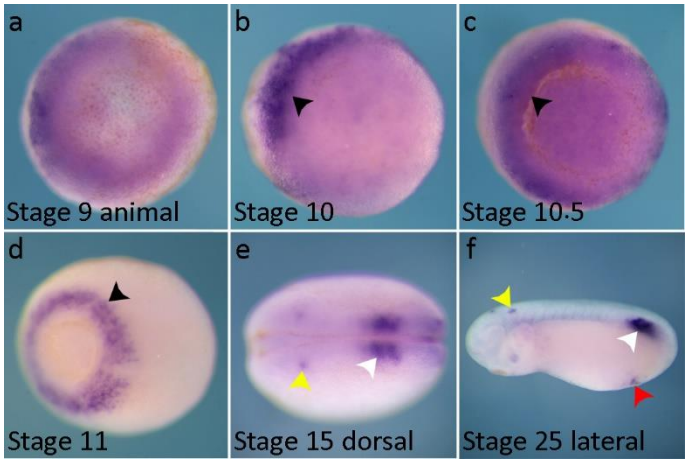

Figure 12: The spatial expression pattern of *rasl11b* during *Xenopus tropicalis* development analysed using in situ hybridisation. (a) late blastula stage 9, animal view, (b) early gastrula stage 10, vegetal view, (c) gastrula stage 10.5, vegetal view, (d) gastrula stage 11, vegetal view, (e) early neurula stage 15, dorsal view, anterior to left, (f) early tailbud stage 25, lateral view, anterior to left. Black arrows indicate early mesoderm, yellow arrows indicate otic vesicles, white arrows indicate presomitic mesoderm and red arrow indicates proctodaeum. Embryos staged according to Nieuwkoop and Faber (1994) stages of *Xenopus* development.

fosl1 is weakly expressed in the animal pole at stage 9 and 10 (Figure 13). At gastrula stage 10.5, *fosl1* expression can be seen in the early mesoderm at the dorsal blastopore before becoming more widespread around the blastopore region. At early neurula stage 14, fosl1 is expressed in the neural plate.

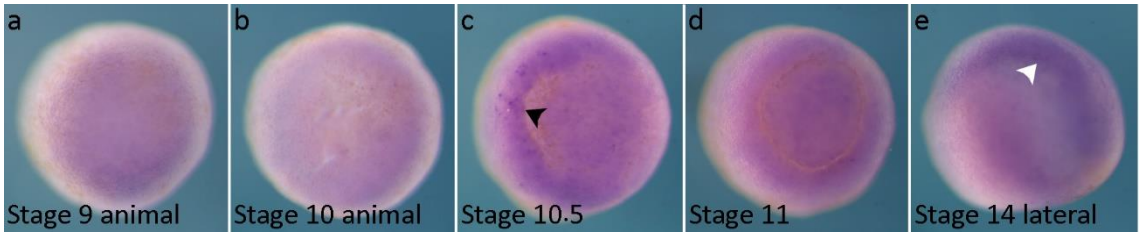

Figure 13: The spatial expression pattern of *fosl1* during early *Xenopus tropicalis* development analysed using in situ hybridisation. (a) late blastula stage 9, animal view, (b) early gastrula stage 10, animal view, (c) gastrula stage 10.5, vegetal view, (d) gastrula stage 11, vegetal view, (e) early neurula stage 14, lateral view, anterior to left. Black arrow indicates dorsal early mesoderm and white arrow indicates neural plate. Embryos staged according to Nieuwkoop and Faber (1994) stages of *Xenopus* development.

As predicted, in situ hybridisation for *frzb, fos, rasl11b* and *fosl1* indicates that the four candidate genes are expressed in known regions of FGF activity and CIC expression.

3.2.6 CIC knockdown and FGF overexpressing embryos exhibit a similar phenotype

Having shown that candidate genes are expressed in known regions of FGF signalling in keeping with our hypothesis, *Xenopus tropicalis* embryos were injected with 1ng CIC targeted TALENs mRNA, 5pg CSKA-FGF4 or 1nl H_2O, for use in validation of the RNA-seq data set.

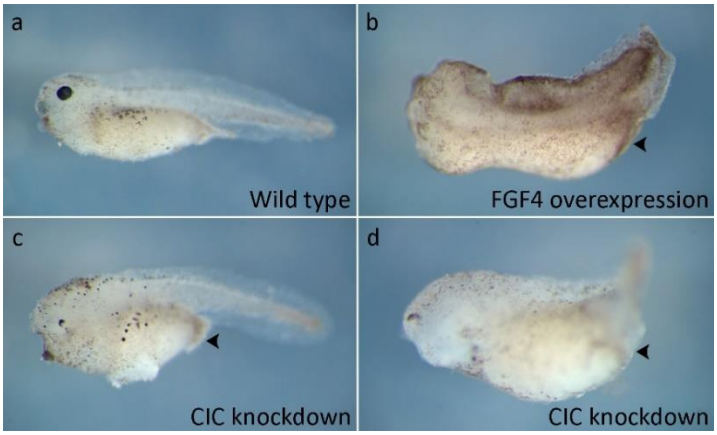

Figure 14: Wild type, FGF4 overexpresssing and CIC knockdown *Xenopus tropicalis* embryo phenotypes. Late tailbud stage embryos, lateral view, anterior to left. (a) wild type, un-injected embryo, (b) embryo injected with 5pg CSKA-FGF4 at the one-cell stage showing FGF4 overexpression phenotype of loss of head structures and an enlarged proctodaeum, (c) embryo injected with 1ng CIC targeted TALENs mRNA at the one-cell stage exhibiting CIC knockdown phenotype of reduced eye pigmentation and an enlarged proctodaeum, (d) embryo injected with 1ng CIC targeted TALENs mRNA at the one-cell stage displaying a more severe CIC knockdown phenotype with loss of head structures and an enlarged proctodaeum. Arrows indicate enlarged proctodaeum.

CIC knockdown via TALENs produced phenotypes with varying degrees of severity. CIC knockdown embryos exhibited phenotypes ranging from reduced eye pigmentation, a reduction in eye size, eye loss, and a reduction or total loss of head structures, in addition to an enlarged proctodaeum. FGF4 overexpressing embryos appeared similar with an enlarged proctodaeum and a reduction or total loss of head structures leading to a posteriorised phenotype (Figure 14). Not all CIC knockdown or FGF4 overexpressing embryos successfully completed gastrulation. CIC knockdown embryos had a death rate of 32.3% (21/65) which was higher than FGF4 overexpressing embryos at 20.7% (24/116) and un-injected embryos at 7.1% (5/70) (Table 10).

***Xenopus tropicalis* embryos**	**Alive**	**Dead**
CIC knockdown	44	21
FGF4 overexpression	92	24
Un-injected	65	5

Table 10: Death rates of injected Xenopus tropicalis embryos. Death rates of un-injected embryos and embryos injected with 1ng CIC targeted TALENs mRNA or 5pg CSKA-FGF4 at the one-cell stage.

3.2.7 Validation of RNA-seq data

In order to validate the RNA-seq data set, RT-PCR was carried out using primers for the target genes on RNA extracted from CIC knockdown, FGF4 overexpressing, H_2O injected and un-injected control embryos flash frozen at neurula stage 14.

Fos, fosl1 and rasl11b bands appeared more intense for CIC knockdown and FGF4 injected embryo cDNA compared to H_2O injected and un-injected controls (Figure 15). However RT-PCR bands for frzb were not more intense in CIC knockdown and FGF4 overexpressing embryos compared to controls.

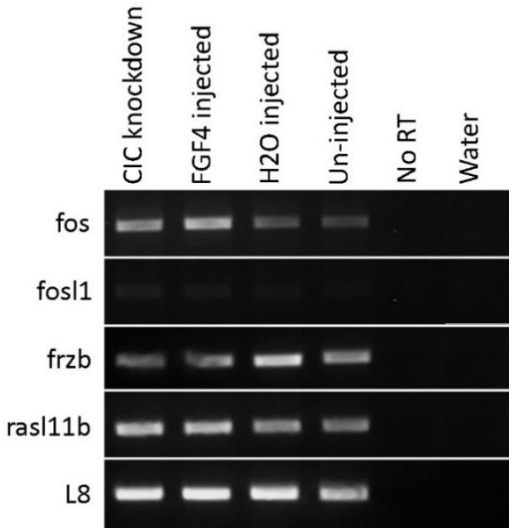

Figure 15: RT-PCR for target gene expression in CIC knockdown, FGF4 injected, H_2O injected and un-injected *Xenopus tropicalis* embryos. RT-PCR carried out using gene specific primers *for fos, fosl1, frzb, rasl11b* and *L8* (as a loading control) on RNA extracted from neurula stage 14 *Xenopus tropicalis* embryos injected with 1ng CIC targeted TALENs RNA, 5pg CSKA-FGF4 or 1nl H_2O, and un-injected embryos.

3.2.8 Enrichment of CIC binding sites

CIC binds octameric T(G/C)AATG(A/G)A sites in promoters and enhancers via its HMG-box domain to repress target gene transcription. In order to determine whether CIC binding sites are enriched around the genomic loci of target genes, MEME suite 5.0.5 (http://meme-suite.org/tools/ame) was used to analyse motif enrichment for the target genes and 1004 shuffled control sequences using Fisher's exact test (Table 10). Enrichment of CIC binding sites was analysed within 100 and 200 base pairs upstream and downstream of each target gene's transcribed sequence. The input sequence was then expanded to include DNA sequences upstream and downstream up to adjacent genes, as regulatory units can be distant from target transcriptional units.

CIC binding sites were enriched within 200 base pairs of the transcribed region for fos, fosl1 and frzb. When a greater region of DNA sequence up to adjacent genes was included, only fos and frzb were identified as genes with enriched CIC binding sites and the number of randomly shuffled control sequences with enriched CIC binding sites (false positives) was also lower. Putative CIC binding sites are located upstream of the transcribed region for fos and within the first intron (Figure 16). Binding motifs were more spread out throughout the gene sequence for fosl1 and frzb.

Input sequences	True Positives	False Positives
Transcribed sequence + 100bp upstream and downstream	3/4 – fos, fosl1 and frzb (75%)	87/1004 (8.7%)
Transcribed sequence + 200bp upstream and downstream	3/4 – fos, fosl1 and frzb (75%)	87/1004 (8.7%)
Sequence up to adjacent gene sequences	2/4 – fos and frzb (50%)	31/1004 (3.1%)

Table 11: Motif enrichment analysis of CIC binding sites. Enrichment of the T(G/C)AATG(A/G)A motif for *fos, fosl1, frzb* and *rasl11b* and 1004 control sequences using Fisher's exact test.

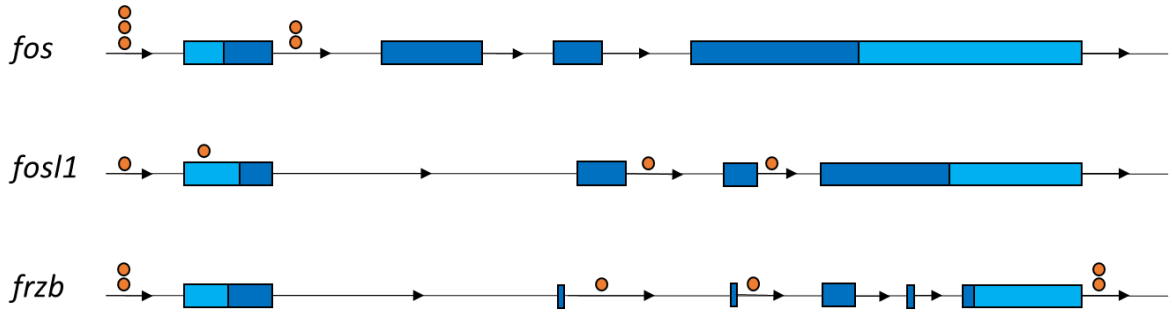

Figure 16: Schematic representation of putative CIC binding site locations within the *fos, fosl1* and *frzb* genes. Orange circles indicate locations of CIC binding motif (T(G/C)AATG(A/G)A). Light blue rectangles indicate 5' and 3' UTRs and dark blue rectangles represent coding sequences.

3.3 Discussion

3.3.1 CIC knockdown and FGF overexpressing embryos possess similar transcriptomes

RNA-seq data showed a statistically significant overlap between changes in gene expression in *Xenopus tropicalis* embryos injected with FGF4 and CIC targeted TALENs with the probability of observed overlaps occurring by chance $p < 0.00001$. Not all genes were significantly upregulated or downregulated in both sets of embryos as CIC may also function in other pathways independently of FGF. For example, ERK is activated in the wound response (Dieckgraefe et al., 1997) and may relieve CIC transcriptional repression of genes required for this FGF independent process following injury. Additionally, out of 88 genes significantly upregulated in FGF4 overexpressing embryos, 37 were not upregulated in CIC knockdown embryos. This may be due to the fact that FGF signal

transduction occurs via 3 main pathways, only one of which involves the ERK/MAPK cascade (Schlessinger, 2000). Consequently, we would expect the transcriptome for the two sets of embryos to be similar but not identical. The statistically significant overlap observed supports the notion that CIC operates in the same pathway as FGF.

3.3.2 Gene ontology enrichment analysis

PANTHER gene ontology enrichment analysis revealed that genes significantly upregulated in FGF4 injected embryos are involved in very relevant biological processes such as the transmembrane receptor protein tyrosine kinase signalling pathway, cellular process, ERK/MAPK cascade, intracellular signal transduction and developmental process. Genes significantly upregulated in CIC knockdown embryos were associated with a greater number of biological processes though many were also involved in cellular process, ERK/MAPK cascade and intracellular signal transduction. The most enriched gene ontology process for CIC knockdown transcripts was cell proliferation. This further supports our hypothesis as FGFs are growth factors which control cell proliferation and migration via different signal transduction pathways. Cell migration usually involves pathways such as Src and p38 MAPK whereas cell proliferation tends to upregulate ERK (Boilly et al., 2000). We would expect FGF target genes associated with cell proliferation to be expressed following relief of CIC repression via ERK (or in this case inhibited by TALENs knockdown).

3.3.3 Target genes are expressed in known regions of CIC expression and FGF signalling

In situ hybridisation for these genes at various stages of *Xenopus tropicalis* development showed that they are expressed in regions of known CIC and FGF activity. All 4 genes selected as candidate genes potentially regulated by both CIC and FGF signalling were expressed in the early mesoderm surrounding the blastopore during gastrulation. FGF2, FGF3, FGF4, FGF8 and FGF20 are expressed in this ring of early mesoderm around the blastopore during gastrulation (Christen and Slack, 1997; Lombardo et al., 1998; Lea et al., 2009). The target genes, both isoforms of CIC, and FGF4 are all particularly enriched at the dorsal blastopore lip during gastrulation (King, 2019; Isaacs et al., 1995). As the sole activator of ERK in early *Xenopus* development (Branney et al., 2009), FGF signalling in the early mesoderm may lead to relief of CIC transcriptional repression of FGF target genes in this region via dpERK-CIC interactions. Other genes previously established to be upregulated by FGF signalling such as *Cdx4* (Keenan et al., 2006)*, MyoD* (Fisher et. al, 2002) and *brachyury* (Smith et al., 1991) are also expressed in this region in gastrula stage 10.5 *Xenopus* embryos.

Candidate genes *frzb* and *rasl11b* are expressed in the presomitic mesoderm on both sides of the dorsal midline from early neurula stage. Presomitic mesoderm gives rise to somites and expresses

high levels of FGF8 to maintain the cells in an immature state posteriorly (Delfini et al., 2005). FGF signalling in this region may prevent CIC inhibiting *rasl11b* and *frzb* expression at this point in development. Both CIC-L and CIC-S show striated expression at tailbud stage indicating expression within somites (King, 2019). In situ hybridisation of *rasl11b* also showed staining on the ventral side in the proctodaeum, which gives rise to the anus. Overexpression of FGF4 in *Xenopus laevis* embryos leads to loss of anterior structures and an enlarged proctodaeum (Isaacs et al., 1994) so FGF signalling may regulate transcription of genes such as *rasl11b* in this region. *rasl11b* is also expressed in the developing otic vesicles where CIC-L, CIC-S, FGF2 and FGF10 and FGF2 are expressed (King, 2019; Lea et al., 2009; Branney et al., 2009). *fosl1*, CIC-L and CIC-S are all expressed in the neural plate where FGFR activity is required for neural induction (Delaune et al., 2005).

If FGF is responsible for regulation of CIC expression, FGF and CIC must be expressed at the same time in FGF responsive cells. FGF1, FGF2, FGF13, FGF22 and CIC-L are maternally expressed in early development (Lea et al., 2009). Both CIC isoforms are expressed at the marginal zone in the early mesoderm during early gastrulation (King, 2019), as are FGF4 and FGF8 (Lea et al., 2009). CIC expression becomes more widespread during development to include regions of known FGF signalling such as the dorsal somites, neural tissue, otic vesicles and branchial arches (King, 2019). Candidate genes upregulated in CIC knockdown and FGF4 overexpressing embryos are expressed in regions of both CIC and FGF expression. Consequently, signalling pathways involving both FGF and CIC may well regulate their expression.

3.3.4 CIC knockdown and FGF overexpressing embryos exhibit similar phenotypes

Targeting of the CIC gene by TALENs results in loss of head structures ranging from complete head loss to reduced eye pigmentation in *Xenopus tropicalis* embryos. These phenotypes are in keeping with previous observations by Michael King (2019) due to the mosaic nature of TALENs (Lei et al., 2012). CIC knockdown embryos resembled the posteriorised phenotype with an enlarged proctadaeum previously observed in FGF4 and FGF8 overexpressing embryos (Isaacs et al., 1994). The fact that CIC knockdown and FGF overexpressing embryos exhibit similar phenotypes supports our hypothesis that CIC operates in the same pathway as FGF.

3.3.4 Validation of RNA-seq data

RT-PCR on RNA extracted from CIC knockdown, FGF4 overexpressing and water injected *Xenopus tropicalis* embryos produced more intense bands for *fos*, *fosl1* and *rasl11b* in both CIC knockdown and FGF4 injected embryos. However the intensity of bands may not directly relate to input quantity due to the fact that a non-quantitative endpoint PCR was carried out. This experiment was

only carried out once and more quantitative results could be obtained in the future through repeats using q-PCR. Whole-mount in situ hybridisation could also be utilised to visualise changes in gene expression following FGF4 overexpression or CIC knockdown.

Many of the genes identified as being significantly upregulated in both CIC knockdown and FGF overexpressing embryos have previously been linked to positive regulation by FGF signalling. Consequently, it was expected that the candidate genes would show differential expression in FGF overexpressing embryos compared to water-injected controls in RT-PCR. For example, the spatial expression of *frzb* is dramatically reduced by the presence of dominant negative FGFR4 (Branney et al., 2009), and expression of both *fos* and *fosl1* were shown to be increased in microarray analysis of bovine ovarian granulosa cells treated with FGF8 (Jiang et al., 2013). During mesoderm induction, FGF2 treatment of *Xenopus* animal cap explants lead to increased AP-1 (c-Fos and c-Jun heterodimer) activity (Kim et al., 1998). Although AP-1 activity has been associated with FGF signalling, this is the first time linking FGF to the transcriptional regulation of *fos*. This is also the first study to show localised expression of *fos* in the developing amphibian embryo.

Other genes upregulated in the RNA-seq for both conditions but not selected for RT-PCR analysis have also been previously associated with FGF signalling. For example, FGF23 induces phosphaturia in proximal tubular epithelial cells by activating SGK1 via ERK (Andrukhova et al., 2012); FGF2 and Wnt3a have been shown to modulate expression of Rho activator *fgd3* in rat chondrosarcoma chondrocytes (Buchtova et al., 2015); and treatment with FGF2 increases *mmp1* expression in primary human osteoarthritis chondrocytes (Nummenmaa et al., 2015). This suggests that the genes highlighted in the RNA-seq were upregulated as described.

3.3.5 CIC binding site enrichment

CIC binding sites are enriched within 200bp of the transcribed sequence for *fos, fosl1* and *frzb*. This indicates that there are a greater number of sites available for CIC binding around the genomic locus for these genes compared to control sequences. CIC may bind enhancers or promoters in these regions in order to repress transcription of these FGF target genes. Inclusion of DNA sequences up to adjacent genes led to identification of only *fos* and *frzb* as genes with enriched the CIC binding motif and fewer false positives suggesting that binding sites are definitely enriched in these regions.

CIC binding sites were not found to be enriched around the genomic locus for *rasl11b* using MEME suite 5.0.5. *Rasl11b* had the lowest effect size out of selected candidate genes from the CIC knockdown RNA-seq data at 2.60, though this was still relatively high. It was predicted that *rasl11b*

would show enriched binding sites due to the gene being identified from the RNA-seq, expressed in regions of known CIC activity and upregulated in CIC knockdown embryos via RT-PCR. CIC may bind enhancers further away from the transcribed *rasl11b* sequence to regulate its expression.

The fact that the CIC knockdown phenotype resembles that of FGF overexpression, along with the statistically significant overlap between genes upregulated or downregulated in the 2 sets of embryos, and that genes are expressed in known regions of CIC and FGF activity supports the hypothesis that CIC and FGF do function in the same pathway, though further experiments are required to solidify this link.

Chapter 4: Manipulation of FGF and ERK signalling pathways in *Xenopus laevis*

4.1 Introduction

As discussed in the introduction (Chapter 1), we hypothesise that transcription of a subset of FGF target genes, and genes involved in the wound response, rely on ERK mediated relief of CIC transcriptional repression. In *Drosophila*, relief of CIC repression is carried out via different mechanisms depending on the RTK signalling pathway activated. Activation of ERK by Torso signalling in the embryonic poles leads to degradation of CIC whereas EGFR signal transduction via ERK in the ovarian follicle induces nuclear export and partial relocalisation of CIC to the cytoplasm (Jimenez et al., 2000; Andreu, Ajuria et al., 2012). Consequently, one of these mechanisms may also be utilised in the ERK mediated wound response and FGF signal transduction.

The first intracellular event associated with wounding is a rapid influx of calcium ions (Ca^{2+}) which radiates out from the site of injury through several rows of cells depending on the severity of the wound (Drumheller and Hubbell, 1991). Calcium activates Rho as well as MAPKs such as ERKs, c-jun-N-terminal protein kinases (JNKs) and the p38 kinases via Ras (Whitmarsh and Davis, 1996; Benink and Bement, 2005; Sosnowski et al., 1993). Activation of ERK leads to repression of phosphoinositide 3-kinase (PI3K) and subsequently further activation of Rho, which phosphorylates myosin-2 (Li et al., 2013; Kimura et al., 1996). The influx of extracellular calcium and Rho GTPase activity are required for assembly of a contractile actomyosin cable around the wound site, which aids closure by drawing the edges together like a 'purse string' (Figure 17) (Bement et al., 1999; Martin and Lewis, 1992). Healing of lens epithelial monolayers is inhibited by the presence of U0126 (MAPKK/MEK inhibitor that prevents activation of ERK1/2) indicating that ERK activation is also crucial for initiation of the wound response (Wang et al., 2003).

In addition to Rho activation in the wound response, dpERK has also been linked to transcription of immediate early response gene *fos* (Dieckgraefe and Weems, 1999). Immediate early response genes (IEGs) such as *fos* and *early growth response (Egr)-1* can be transcribed within minutes of stimulation (Bahrami and Drablos, 2016). This rapid response is possible due to the fact that IEG transcription does not require protein synthesis as the proteins and transcription factors needed are already available in the cell (Herschman, 1991). Expression of *fos* has been reported in epidermal wound edge cells and superficial exposed wound mesenchymal cells in chick and rodent model organisms within 15 minutes of wounding (Martin et al., 1994), and *fos* expression in IEC-6 monolayers peaks 20 minutes after mechanical wounding (Dieckgraefe et al., 1997). Fos binds Jun

transcription factors to form AP-1 heterodimers which have been associated with cell motility, wound re-epithelialisation and regulation of wound healing genes when induced in the wound response (Figure 17) (Tran et al., 1999; Yates and Rayner, 2002; Martin et al., 1994). ERK activation, *fos* and *Egr-1* expression, and wound healing is reduced or completely inhibited by MAPKK inhibitor PD-98059 (Dieckgraefe and Weems, 1999). CIC may therefore play a role in the transcriptional aspect of the wound response downstream of dpERK.

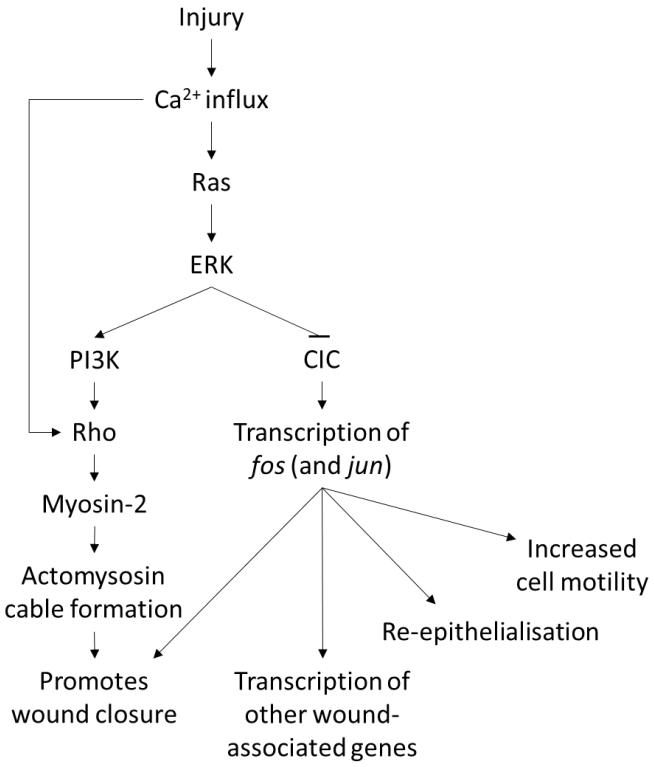

Figure 17: Schematic diagram of the hypothesised role of CIC within phase 1 of the wound response.

The aims of this chapter are:

- Determine if CIC is degraded or relocated to the cytoplasm following activation of FGF signalling
- Investigate ERK activation and *fos* transcription in the wound response over time in *Xenopus*
- Determine if CIC is degraded or relocated to the cytoplasm in the wound response

4.2 Results

4.2.1 Capicua expression and ERK activation in the presence and absence of FGF signalling

In order to determine whether CIC is degraded or relocated to the cytoplasm in the presence of FGF, *Xenopus laevis* embryos were unilaterally injected at the 2-cell stage with 50pg or 100pg amino terminus myc-tagged *Mus musculus* CIC-S mRNA (Kim et al., 2013; King, 2019), with or without 10pg FGF4 mRNA. Immunostaining for dpERK was carried out to indicate ERK activation, and immunostaining for myc undertaken to show presence of myc-CIC.

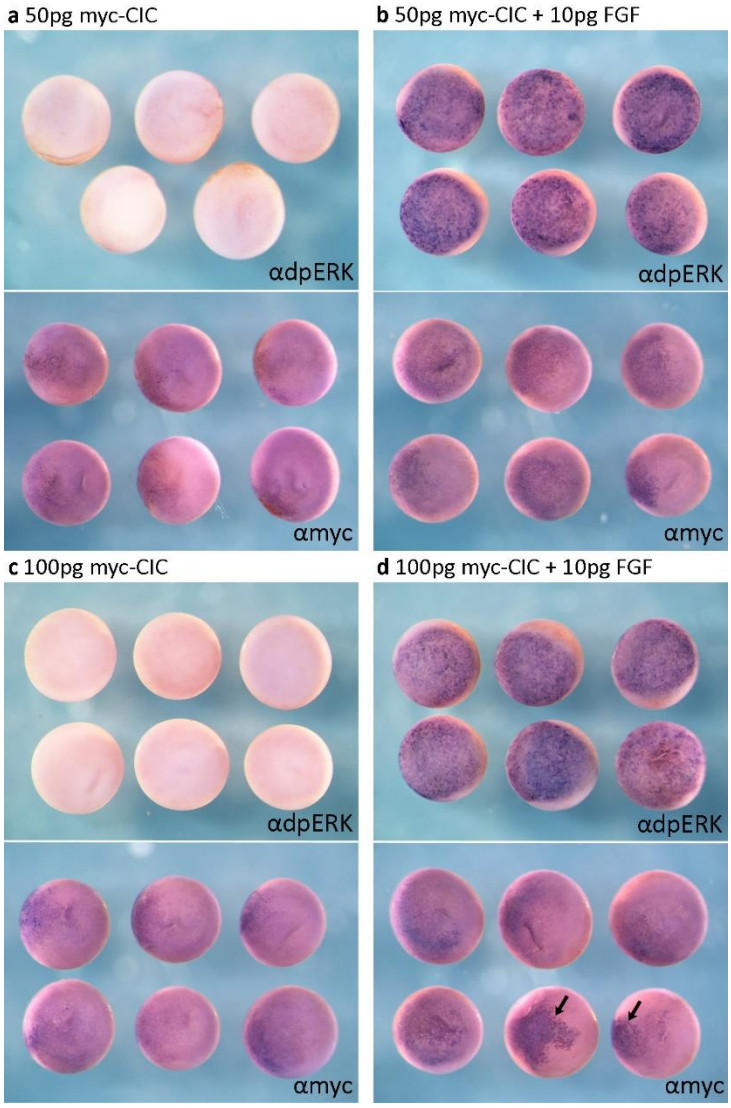

Figure 18: Immunostaining for dpERK and myc on *Xenopus laevis* embryos injected with myc-CIC mRNA or myc-CIC + FGF4 mRNA. Early gastrula stage 10.5 *Xenopus laevis* embryos unilaterally microinjected at the 2-cell stage with (a) 50pg myc-CIC mRNA (b) 50pg myc-CIC + 10pg FGF4 mRNA (c) 100pg myc-CIC mRNA (d) 100pg myc-CIC + 10pg FGF4 mRNA, arrows indicate 2 significantly more stained embryos. 10 embryos per treatment. Animal view.

ERK was activated in the animal pole of embryos overexpressing FGF due to activation of the FGF signalling pathway (Figure 18b and 18d).

Staining for myc appears more intense in 10 embryos injected with 50pg myc-CIC + 10pg FGF4 (Figure 18b) than 50pg myc-CIC alone (Figure 18a). 2 out of 10 embryos injected with 100pg myc-CIC + 10pg FGF4 and immunostained for myc (Figure 18d) were clearly more stained than 100pg myc-CIC alone (Figure 18c). There was no notable difference in staining between the other 8 embryos (Table 12).

	Number of embryos	
	More intense staining when co-injected with 10pg FGF	**No difference in staining with or without co-injection of 10pg FGF**
50pg myc-CIC	10	0
100pg myc-CIC	2	8

Table 12: Difference in myc immunostaining between embryos injected with myc-CIC or myc-CIC and FGF. Number of embryos showing more intense or no difference in staining when co-injected with 10pg FGF alongside 50pg or 100pg myc-CIC. 10 embryos per treatment.

In order to see if higher quantities of CIC and FGF4 had the same effect, a similar experiment was carried out with 10 *Xenopus laevis* embryos injected with 500pg myc-CIC mRNA and 10 co-injected with 500pg myc-CIC mRNA + 20pg CKSA-FGF4. The CSKA plasmid construct was used as it was hypothesised that not enough FGF was being translated for a level of ERK activation sufficient to degrade or relocate the high level of myc-CIC being expressed. Use of the plasmid meant that new mRNA continued to be made while injected FGF mRNA would have decayed away.

ERK was activated in a small region where FGF was overexpressed in myc-CIC and CSKA-FGF4 injected embryos (Figure 19c).

Myc staining looks nuclear in embryos injected with and without CSKA-FGF4 (Figure 19).

Myc staining was less intense in embryos co-injected with CSKA-FGF4 compared to myc-CIC alone. ImageJ was utilised to determine percentage area stained with a mean of 13.71% for FGF expressing embryos compared to 35.39% for those injected with only myc-CIC (Table 13, n=5). An un-paired T test showed that the reduction in percentage staining was statistically significant with a p value < 0.000001 (Figure 20)

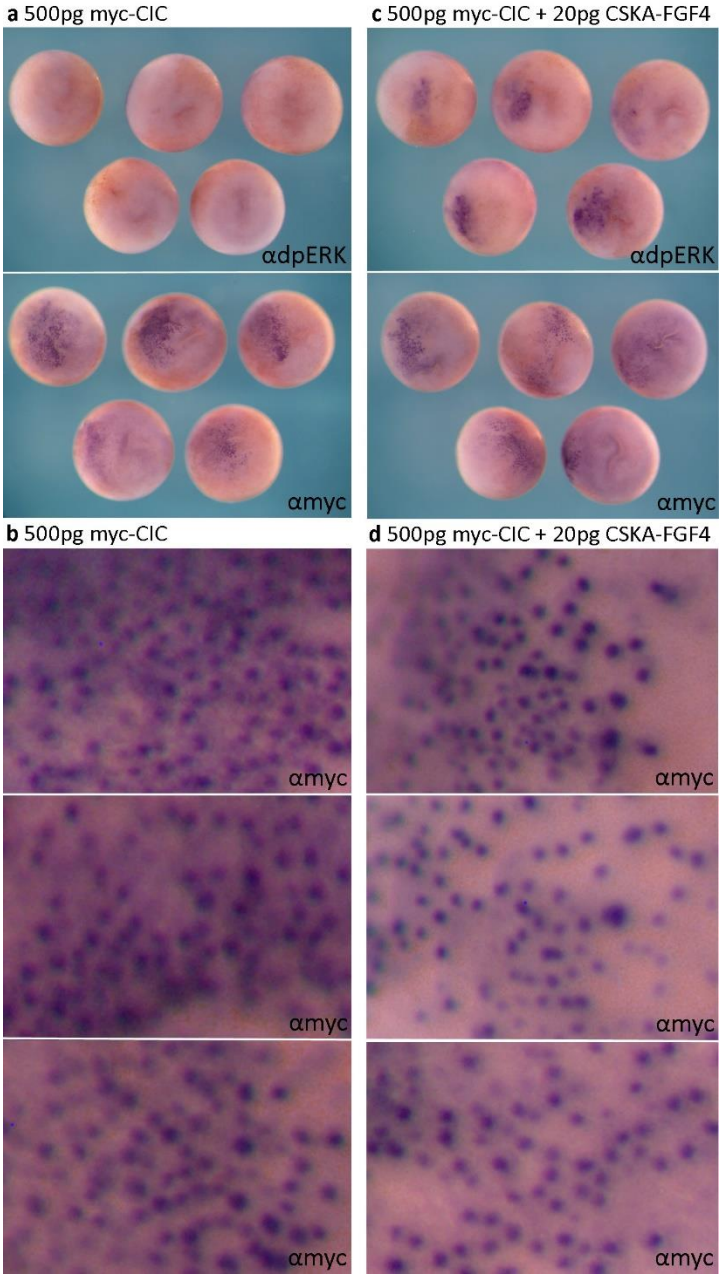

Figure 19: Immunostaining for dpERK and myc on *Xenopus laevis* embryos injected with myc-CIC mRNA or myc-CIC mRNA + CSKA-FGF4. Early gastrula stage 10.5 *Xenopus laevis* embryos unilaterally microinjected at the 2-cell stage with (a, b) 500pg myc-CIC mRNA (c, d) 500pg myc-CIC mRNA + 20pg CSKA-FGF4. 10 embryos per treatment. Animal view.

	Percentage area stained	
	Myc-CIC	**Myc-CIC + FGF**
1	38.14	11.19
2	33.69	12.01
3	36.70	14.54
4	36.93	14.87
5	31.47	15.96
Mean	35.39 (SEM= 1.22)	13.71 (SEM= 0.90)

Table 13: Percentage area stained in myc-CIC and myc-CIC + FGF injected embryos. 500pg myc-CIC mRNA and 500pg myc-CIC mRNA + 20pg CSKA-FGF4 injected embryos immunostained for myc at gastrula stage 10.5 and analysed using ImageJ. n=5 per treatment. SEM= standard error of the mean.

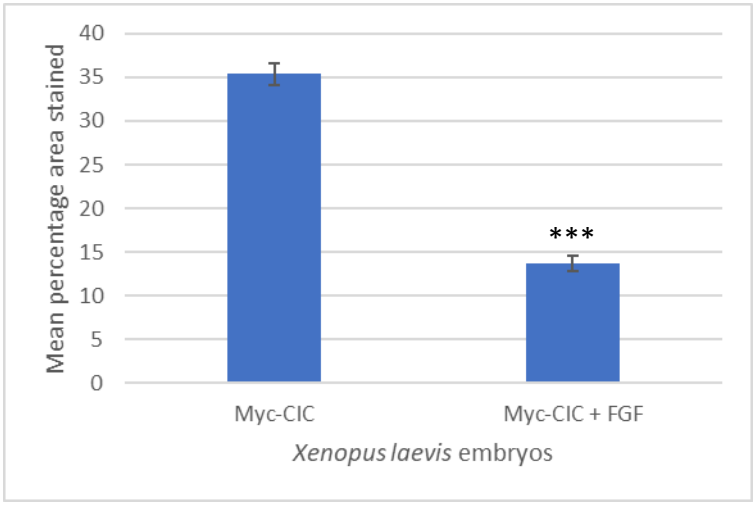

Figure 20: Percentage area stained in myc-CIC and myc-CIC + FGF injected embryos. 500pg myc-CIC mRNA and 500pg myc-CIC mRNA + 20pg CSKA-FGF4 injected embryos immunostained for myc at gastrula stage 10.5 and analysed using ImageJ. n=5 per treatment. p= 0.0000006. Error bars indicate standard error of the mean.

4.2.2 Western blot analysis of myc-CIC and FGF expressing embryos

A western blot for dpERK and total ERK was carried out on 500pg myc-CIC, 500pg myc-CIC + 20pg CSKA-FGF4 and un-injected *Xenopus laevis* embryos in order to show activation of ERK by FGF (Figure 21).

dpERK is present in injected and un-injected embryos as ERK is activated by endogenous FGF signalling in the early mesoderm surrounding the blastopore at the vegetal pole at this stage of development. Higher levels of dpERK are seen in FGF overexpressing embryos compared to un-injected and myc-CIC injected embryos. ImageJ analysis showed that the relative density of dpERK bands for myc-CIC and myc-CIC + CSKA-FGF4 injected embryos are 2.08 and 3.81 respectively when normalised to loading control and un-injected embryos.

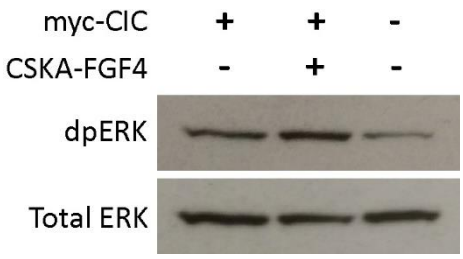

Figure 21: Western blot detecting dpERK and total ERK in *Xenopus laevis*. Embryos unilaterally injected with 500pg myc-CIC mRNA or 500pg myc-CIC mRNA + 20pg CSKA-FGF4 at the 2-cell stage, or un-injected. 5 embryos were collected at gastrula stage 10.5 for each condition and protein for approximately 1 embryo loaded for each.

In order to obtain a more quantitative result than immunostaining, western blots were carried out with the aim of determining if there was a reduction of overall levels or size shift in myc-CIC protein. A change in protein size would indicate partial degradation or proteolytic cleavage of CIC following FGF signalling. Due to the fact that CIC is a large protein with a molecular weight of 250kDa, various conditions were tested to optimise transfer of the protein. SDS-PAGE gels were made with 7.5% and 7% acrylamide, various concentrations of sample protein loaded, and gels run for increased lengths of time. Transfer buffer was altered to include 0.05% SDS instead of 10% methanol and proteins transferred overnight at 30V at 4°C. Anti-myc antibodies were used at a 1:4000 and 1:5000 dilution and 2 different antibodies tested. Unfortunately, no clear bands were observed using any of these conditions. Ponceau S staining of the membrane and Coommassie blue staining of the gel also revealed no bands.

4.2.3 ERK activation and *fos* transcription in the wound response

In order to investigate ERK activation and gene transcription in the wound response, wild type *Xenopus laevis* embryos were wounded at late neurula stage 20 with a tungsten needle. Immunostaining for dpERK was carried out to indicate ERK activation and in situ hybridisation undertaken to show *fos* expression.

ERK is activated extremely rapidly post-wounding with dpERK expression radiating across the embryo from the wound site after 5 minutes (Figure 22A). The presence of dpERK is transient with no staining at the wound site after 60 minutes.

fos is rapidly and transiently expressed following activation of ERK at the wound site (Figure 22). Transcription of *fos* takes longer than ERK activation post-wounding with expression beginning to show at 10 minutes post-wounding and strong expression at 30 minutes. *fos* is no longer expressed at the wound site 120 minutes after wounding and the embryo appears completely healed and scar-free.

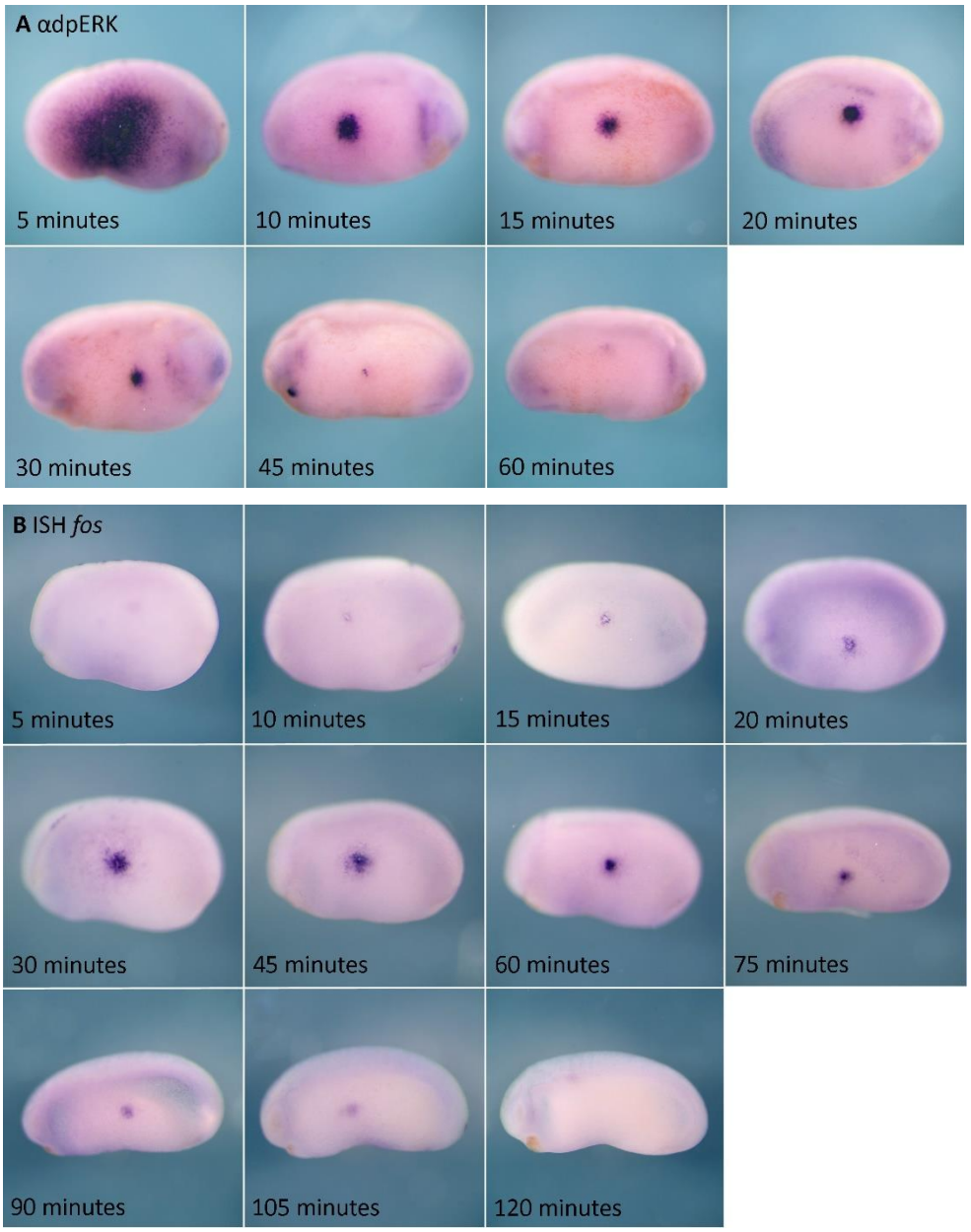

Figure 22: *Xenopus laevis* embryos wounded at late neurula stage 20 and fixed at a range of time points post-wounding for immunostaining for dpERK and in situ hybridisation of *fos*. (A) Immunostaining with an antibody against dpERK, 6 embryos per time point (B) In situ hybridisation of *fos*, 10 embryos per time point. Lateral view, anterior to left.

4.2.4 Capicua expression and *fos* transcription in the wound response

In order to determine if CIC is degraded or relocated to the cytoplasm following ERK activation in the wound response, *Xenopus laevis* embryos were unilaterally injected at the 2-cell stage with 500pg myc-CIC mRNA. 14 embryos were wounded at the site of injection at early gastrula stage 10.5 and fixed 30 minutes post-wounding. Early gastrula stage was chosen so that results would be comparable to those obtained in Figure 19. Immunostaining for myc was carried out to visualise myc-CIC expression in 7 wounded and 10 control embryos. In situ hybridisation for *fos* was also carried out to see if there was a difference in *fos* expression between 7 myc-CIC expressing and 7 un-injected embryos in the wound response.

Myc staining was less intense in wounded myc-CIC injected embryos (Figure 23b) compared to controls (Figure 23a). ImageJ was utilised to determine percentage area stained with a mean of 9.19% for wounded myc-CIC injected embryos compared to 35.39% for control myc-CIC embryos (Table 14, n=5). An un-paired T test showed that the reduction in percentage staining was statistically significant with a p value < 0.00001 (Figure 24)

Two out of six un-injected embryos showed stronger *fos* expression at the wound site (Figure 23c) compared to six myc-CIC injected embryos (Figure 23d). However, there was not a detectable difference in the other embryos, and they did not all remain intact during the in situ hybridisation process.

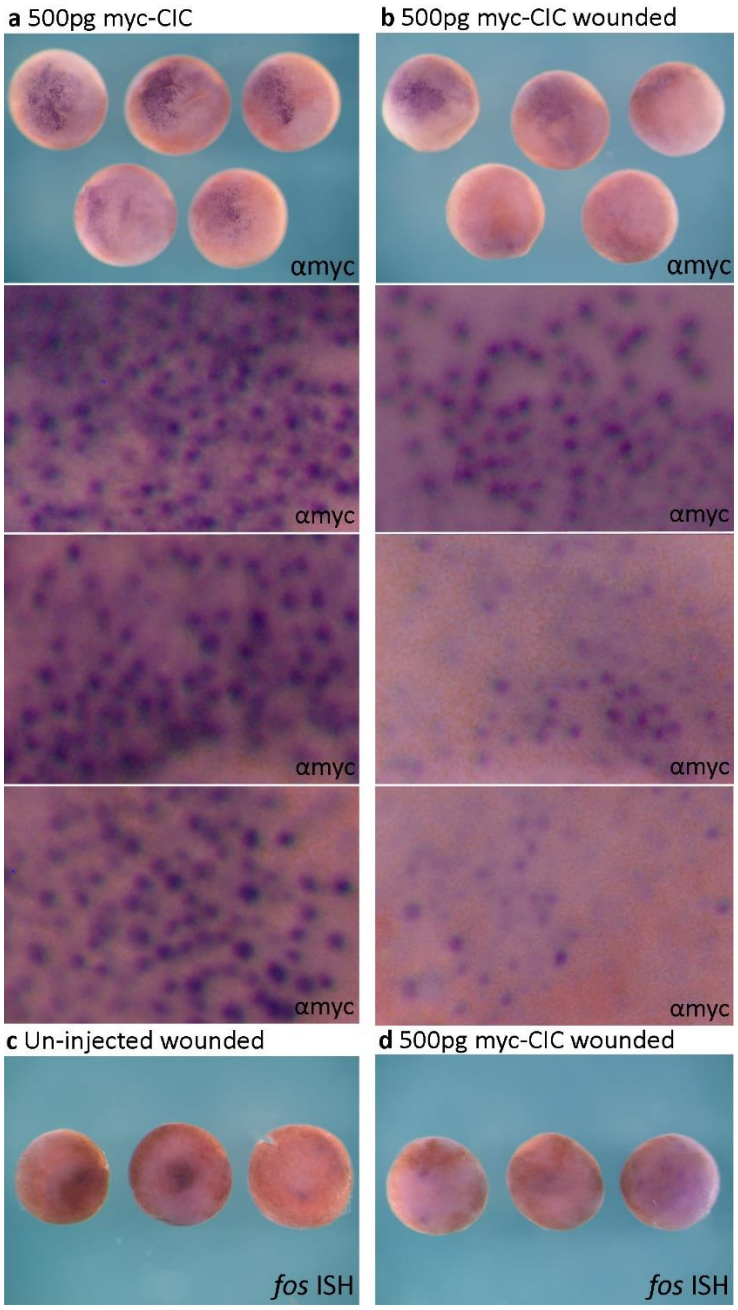

Figure 23: Immunostaining for myc and in situ hybridisation for *fos* in wounded and control myc-CIC expressing or un-injected *Xenopus laevis* embryos. (a) Immunostaining for myc in early gastrula stage *Xenopus laevis* embryos unilaterally injected at the 2-cell stage with 500pg myc-CIC mRNA (b) Immunostaining for myc in *Xenopus laevis* embryos unilaterally injected at the 2-cell stage with 500pg myc-CIC mRNA then wounded at the site of injection at early gastrula stage 10.5 and fixed 30 minutes post-wounding (c) in situ hybridisation for *fos* in un-injected *Xenopus laevis* embryos wounded at early gastrula stage 10.5 and fixed 30 minutes post-wounding (d) in situ hybridisation for *fos* in *Xenopus laevis* embryos unilaterally injected at the 2-cell stage with 500pg myc-CIC then wounded at the site of injection at early gastrula stage 10.5 and fixed 30 minutes post-wounding. Animal view.

	Percentage area stained	
	Myc-CIC	Wounded myc-CIC
1	38.14	12.40
2	33.69	8.61
3	36.70	6.78
4	36.93	13.16
5	31.47	5.00
Mean	35.39 (SEM= 1.22)	9.19 (SEM= 1.58)

Table 14: Percentage area stained in wounded and control myc-CIC injected embryos. 500pg myc-CIC mRNA injected embryos immunostained for myc at gastrula stage 10.5 and analysed using ImageJ. n=5 per treatment. SEM= standard error of the mean.

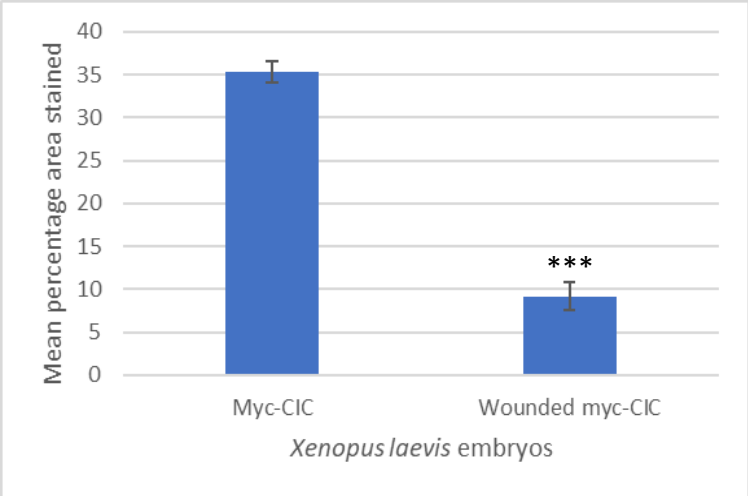

Figure 24: Percentage area stained in wounded and control myc-CIC injected embryos. 500pg myc-CIC mRNA injected embryos immunostained for myc at gastrula stage 10.5 and analysed using ImageJ. n=5 per treatment. p= 0.000001. Error bars indicate standard error of the mean.

4.3 Discussion

4.3.1 CIC expression is reduced following FGF signalling

It was predicted that myc-CIC would be degraded in embryos co-injected with FGF4 mRNA however *Xenopus laevis* embryos injected with 50pg and 100pg myc-CIC, with and without 10pg FGF4 mRNA did not conclusively show this result. This may be due to the fact that insufficient levels of FGF were present to activate enough ERK to completely degrade or relocate non-endogenous myc-CIC. Embryos injected with 500pg myc-CIC and 20pg CSKA-FGF4 did however show a reduction in staining compared to 500pg myc-CIC injected embryos and staining in both sets of embryos appeared nuclear. This suggests that myc-CIC was degraded by activation of ERK in FGF signal transduction when a greater amount of FGF4 was expressed for a longer period of time, in a similar manner to ERK mediated CIC degradation in the *Drosophila* Torso pathway (Jimenez et al., 2000). However, not all embryos showed the same level of staining indicating that unequal amounts of

FGF4 may have been produced. Mosaic expression is common in amphibians with injection of DNA constructs giving rise to some variation in the proportion of expressing cells between individual embryos. It is unlikely that the biological effects of FGF4 would be restricted due to the fact that FGF4 is a secreted molecule (Isaacs et al., 1994) but it may be worth repeating this experiment in the future with a larger number of embryos. In spite of the slight variation between levels of staining between embryos, ImageJ analysis and T test results indicate that the FGF4 induced reduction in staining was statistically significant.

Western blots carried out to quantitatively determine whether myc-CIC was partially degraded or proteolytically cleaved following activation of FGF signalling did not produce any bands. However, the 250kDa CIC protein has been previously observed in western blots for GFP tagged *Homo sapiens* CIC-S mRNA injected *Xenopus laevis* embryos and is not present in embryos co-injected with FGF8 mRNA (King, 2019). Embryos injected with GFP-CIC-S mRNA and FGF4 mRNA produced no 250kDa protein but instead revealed an 85kDa protein which could be a product of CIC protein degradation (King, 2019). This suggests that FGF signalling may be linked to the initiation of CIC protein degradation with FGF8 possessing a greater potency than FGF4.

RTK signalling in *Drosophila* has also been shown to lead to downregulation of CIC expression so it is logical that a similar mechanism may be employed in the FGF RTK pathway. For example, Torso signalling in the embryonic poles and EGFR signal transduction in the imaginal wing and eye discs both lead to reduced levels of CIC (Jimenez et al., 2000; Tseng et al., 2007). The fact that embryos overexpressing FGF show decreased immunostaining for CIC-S and CIC-S protein degradation in western blots is in keeping with mechanisms observed in other RTK pathways. This suggests a role for FGF mediated regulation of CIC expression levels.

4.3.2 Fos is expressed rapidly and transiently following ERK activation in the wound response

Activation of ERK at the wound site was rapid and transient with dpERK expressed within 5 minutes but gone within the hour. This is in line with previous findings that ERK activation is detectable in both superficial and deep wounds by western blot analysis from 2 minutes post-wounding with phosphorylation decreased by 1 hour (Li et al., 2013).

Ca^{2+} influx and ERK activation have been established as critical components of the first phase of wound closure for Rho activation and assembly of the actomyosin cable (Stanistreet, 1982; Wang et al., 2003). *fos* has also been strongly associated with wound healing in multiple species and cell cultures. In this study, immediate early response gene *fos* was expressed at the wound site

following activation of ERK with the strongest expression at 30 minutes post-wounding. Previous experiments have reported a greater than 30-fold increase in c-Fos mRNA by northern blot analysis of wounded IEC-6 cell monolayers peaking at 20 minutes (Dieckgraefe et al., 1997). *fos* expressed at the wound site may play a role in regulating transcription of genes involved in the healing process. For example, cells that express c-*fos* during the wound response subsequently upregulate transforming growth factor beta-1 (TGFβ1) mRNA 1 hour post-wounding and secrete TGFβ1 protein into the wound mesenchyme (Martin et al., 1994). c-*fos* may upregulate TGFβ1 at the wound site in order to initiate connective tissue contraction of the wound mesenchyme.

In contrast to this, it has been shown that *Xenopus* embryos injected with α-amanitin (Xenopus polymerase II transcription inhibitor) do not show a significant healing defect indicating that transcription is not important for the initial response to wounding (Li et al., 2013). However, bioinformatic analysis of four epidermal wound enhancers rapidly activated at *Drosophila* wound sites revealed evolutionarily conserved sequences matching binding sites for Jun/Fos and Grainy head (Grh) transcription factors in each (Pearson et al., 2009). Fos, Jun and Grh have been linked to epidermal wound repair, barrier development and differentiation in *Xenopus* and mammals (Ting et al., 2005; Yates and Rayner, 2002; Tao et al., 2005) The combination of Fos and Grh may be part of an ancient wound-response pathway conserved in both vertebrates and invertebrates, however other additional mechanisms have evolved alongside this to result in a similar output.

Additionally, interference with proteins upstream of wound-induced transcription lead to an inhibition of *fos* transcription. For example, calcium is required for wound-induced transcriptional activation of AP-1 via Ras and ERK activation (Tran et al., 1999). *c-fos* expression, cell motility and DNA synthesis are all inhibited by microinjection of a dominant-negative Ras mutant protein (Sosnowski et al., 1993). Treatment of wounded keratinocytes with ERK and p38 MAPK inhibitors reduces the activity of a response element controlled by AP-1 (Jaakkola et al., 1998).

4.3.3 CIC expression is reduced in wounded embryos

As ERK is activated rapidly in the wound response, it was predicted that wounding would lead to a decrease in CIC expression as activated ERK has been shown to relieve CIC transcriptional repression in multiple species (Jimenez et al., 2012). Myc-CIC-S expression was reduced in wounded embryos compared to control embryos injected with the same amount of myc-CIC-S. ImageJ analysis and T test results indicate that the wound induced reduction in staining was statistically significant. This is in keeping with our hypothesis that CIC is sent for degradation in the ERK mediated wound response.

As *fos* is expressed following ERK activation in the wound response (Dieckgraefe et al., 1997) and may be transcriptionally regulated by CIC in the FGF signalling pathway (Chapter 3), it was predicted that *fos* expression may also be regulated by CIC in the wound response. It was thought that overexpression of CIC may reduce *fos* transcription at the wound site, however only 2 un-injected embryos showed higher *fos* expression than myc-CIC-S injected embryos. No detectable difference in *fos* expression was observed between the other 5 myc-CIC-S injected and un-injected wounded embryos at gastrula stage 10.5. This could be due to the fact that dpERK at the wound site was sufficient to relieve CIC repression in the majority of the embryos in spite of the higher level of CIC protein expression.

Another gene strongly activated at wound sites in *Drosophila* and murine models is growth arrest and DNA damage inducible *gadd45* (Stramer et al., 2008). Gadd45 mediates activation of the p38/JNK pathway via MTK1/MEKK4 kinase following exposure to physiological and environmental stressors (Salerno et al., 2012). A cluster of Fos and Grh consensus sites are located upstream of the *gadd45* transcription start site indicating that *fos* expression may be important for regulation of other wound healing genes (Pearson et al., 2009). Both *fos* and *gadd45a* were upregulated in CIC knockdown and FGF overexpressing embryos (Chapter 3). As CIC expression is reduced in wounded embryos following rapid activation of ERK, and wound-associated genes have been identified as potential targets of CIC regulation, ERK mediated relief of CIC repression via wounding may allow transcription of *fos* and subsequently other wound healing enhancers such as *gadd45a*.

Chapter 5: General discussion

5.1 Summary

Despite extensive understanding of FGF signal transduction, the specific mechanism responsible for regulation of target gene transcription is still not fully understood. We hypothesise that transcription of a subset of FGF target genes, and genes involved in the wound response, rely on ERK mediated relief of CIC transcriptional repression.

Findings in this study support the notion that CIC operates in the same pathway as FGF, as CIC knockdown and FGF overexpressing embryos exhibit similar phenotypes and possess similar transcriptomes. CIC expression is reduced following activation of FGF signalling, and previous data indicates that CIC protein is partially degraded by injection of FGF4 mRNA and completely degraded by FGF8 (King, 2019). Additionally, 75% of the putative CIC and FGF regulated genes further analysed have enriched CIC binding sites around their genomic locus.

In keeping with our hypothesis, CIC expression is also reduced in wounded embryos. *fos* and *gadd45a* are both expressed in the wound response (Martin et al., 1994; Dieckgraefe et al., 1997; Stramer et al., 2008) and were also upregulated in CIC knockdown embryos. Consequently, wound-induced ERK mediated relief of CIC repression may allow transcription of *fos* and subsequently other wound healing enhancers such as *gadd45a*.

5.2 The role of Capicua in development and disease

5.2.1 Fibroblast growth factor signalling

FGF signalling is crucial for the initiation and regulation of multiple developmental processes including gastrulation, neural and mesoderm induction, limb development and anteroposterior patterning (Rossant et al., 1997; Slack et al., 1996; ten Berge et al., 2008). Correct regulation of this signalling pathway is extremely important for normal development as several FGFR mutations lead to cancer or skeletal abnormalities such as craniosynostosis, achondroplasia and Crouzon syndrome (Teven et al., 2014). Misregulation of FGF ligands can also lead to a range of disorders including hypophosphatemic rickets, lacrimo-auriculo-dento-digital syndrome and Kallmann syndrome (White et al., 2000; Milunsky et al., 2006; Falardeau et al., 2008). Furthering our understanding of the molecular mechanisms involved in FGF dependent transcription will be a significant benefit in fields such as stem cell research, regenerative medicine and treatment of cancers associated with misregulated FGF signalling. If CIC functions as a transcriptional repressor downstream of FGF, Torso and EGFR, it may also be an important component of other RTK signalling pathways.

5.2.2 Capicua and cancer

FGFR mutations are associated with a number of cancers due to the fact that correct FGF signalling is important for appropriate cell proliferation and angiogenesis (Turner and Grose, 2010). CIC is a known tumour suppressor gene as it is a critical regulator of cell proliferation.

Most cancer-causing CIC mutations identified so far are missense mutations in the HMG-box domain which is responsible for DNA binding (Tanaka et al., 2017). Missense and insertion/deletion mutations in the C1 motif are the next most common as the C1 domain normally interacts with the HMG-box to stabilise DNA binding (Fores et al., 2017). Loss of function mutations such as these are frequently observed in human neoplasms such as oligodendroglioma as loss of repression of target genes such as the PEA3 family promotes cell proliferation and migration (Tanaka et al., 2017). Alternatively, if CIC fuses with DUX4, the chimera is a potent transcriptional activator which significantly upregulates PEA3 family transcription factors (such as Etv1, Etv4 and Etv5) and Ccnd1/d2 resulting in Ewing-like sarcoma development (Kawamura-Saito et al., 2006; Yoshimoto et al., 2017). Chromosomal translocation of PEA3 transcription factors is also associated with development of Ewing and prostate tumours through the upregulation of matrix metalloproteases and other targets involved in extracellular matrix remodelling and metastatic behaviour (de Launoit et al., 2006; Oh et al., 2012). As PEA3 genes are known mediators of FGF signalling and their transcription is repressed by CIC in the absence of ERK signalling (Garg et al., 2018; Dissanayake et al., 2011), furthering our understanding of this pathway may allow development of improved cancer treatments in the future.

A common target of molecular targeted therapy is the RTK-Ras-ERK/MAPK pathway. Unfortunately, acquired resistance for these therapies occurs frequently so downstream modifiers such as CIC have the potential to be utilised as effective alternative targets (Schmitt et al., 2016).

5.2.3 Capicua and neurodegenerative disease

In vertebrates, the CIC ATXN1 binding domain enables binding of ATAXIN-1 (ATXN1) or ATAXIN-1-LIKE (ATXN1L) proteins to form a protein repressor complex (Lam et al., 2006). In spinocerebellar ataxia type 1 (SCA1) patients, a polyglutamine expansion-induced gain of function in ATXN1 results in loss of motor co-ordination, slurred speech and cognitive impairments (Zoghbi and Orr, 2009). Neurotoxic glutamine expanded ATXN1 binds CIC less efficiently leading to a significantly increased level of expression of CIC targets such as Etv1, Etv5 and Ccnd1 due to a reduction in CIC-ATXN1 co-repressive activity (Crespo-Barreto et al., 2010; Lim et al., 2008). In addition to this, glutamine expanded ATXN1 also causes stronger binding of CIC to promoters of certain genes leading to hyper-

repression of those targets (Fryer et al., 2011). A genetic reduction of CIC levels substantially improves the neuropathy, learning and memory of mice expressing glutamine expanded ATXN1. Physical exercise also improves the SCA1 phenotype as enhanced EGFR signalling in brainstem downregulates CIC in the tissue (Fryer et al., 2011). It has been hypothesised that SCA1, along with other neurodegenerative diseases including Huntington, prion and Alzheimer disease, involve both increased function of a specific endogenous protein complex, in addition to loss of function of another (Lim et al., 2008; Crespo-Barreto et al., 2010). Therefore, understanding the role of ATXN1-CIC protein complexes and mutations leading to simultaneous loss and gain of function mutations may allow discovery of critical pathways involved in polyglutamine expanded neurodegenerative pathologies.

Additionally, in mouse models, the progressive pathogenesis of SCA1 has been shown to require the continuous expression of glutamine expanded ATXN1 (Zu et al., 2004). This suggests that identification of drugs able to target the mutant ATXN1-CIC complexes in the correct manner may allow some degree of functional recovery for SCA1 patients (Zoghbi and Orr, 2009). Consequently, it is important that we understand how CIC functions and interacts with other proteins and promoters both in normal development and disease conditions to be able to develop effective therapeutics.

5.3 Future work

This study presents evidence towards a transcriptional regulatory role for CIC for a subset of FGF target genes and genes involved in the wound response. Further validation of the RNA-seq data using qPCR would provide more quantitative results but could not be carried out within the scope of this project. Alternatively, in situ hybridisation for transcripts identified in the RNA-seq could be carried out at a range of developmental stages in wild type, CIC knockdown and FGF overexpressing embryos, and gene expression compared.

Repeating the experiment in which immunostaining against myc was carried out on myc-CIC, myc-CIC + CSKA-FGF4 and wounded myc-CIC expressing embryos to obtain higher numbers of replicates would be advantageous in strengthening the reliability of statistics. Additionally, dry transfer of the western blot using the iBlot 2 dry blotting system (Thermo Fisher Scientific) with a higher concentration of antibody may allow quantitative determination of whether myc-CIC is partially degraded or proteolytically cleaved following activation of ERK by FGF signalling or wounding.

Following on from the encouraging MEME analysis carried out on the 4 genes investigated in this study, chromatin immunoprecipitation sequencing (ChIP-seq) with tagged constructs could be

utilised to identify global binding sites for CIC to aid identification of other CIC target genes. Comparison to the RNA-seq data set could be carried out to determine whether the other upregulated transcripts also possess enriched binding sites around their genomic locus. An additional approach would be identification of direct binding targets of CIC through electrophoretic mobility shift assays (EMSA).

In this study, wound-induced dpERK and *fos* expression was investigated following wounding at late neurula stage 20 with clear results. Unfortunately, *fos* expression at the wound site at gastrula stage 10.5 was not as easily detectable. Consequently, expression of wound-associated genes such as *fos* could be analysed by in situ hybridisation with a greater number of embryos at a later stage of development in myc-CIC injected and wild type embryos to contribute to determining whether CIC overexpression affects transcription. Use of an additional model such as Zebrafish may further confirm these wound healing results.

5.4 Conclusions and implications

Evidence presented in this thesis supports the hypothesis that transcription of a subset of FGF target genes, and genes involved in the wound response, rely on ERK mediated relief of CIC transcriptional repression. Tight regulation of these target genes is essential for normal development as misregulation of the FGF signalling pathway and/or CIC repression is associated with a range of disorders and cancers. Understanding the molecular mechanisms involved in CIC regulated pathways may allow more effective approaches to treatment of developmental disorders, neurodegenerative diseases and cancer. Expanding our knowledge of the wound healing process may also ultimately lead to the development of new, more effective therapeutics and procedures to promote accelerated wound healing.

Abbreviations

AP-1	Activator protein-1
ATXN1	Ataxin-1
Xbra	Brachyury
CIC	Capicua
CIC-L	Capicua-long isoform
CIC-S	Capicua-short isoform
Cdx	Caudal type homeobox
Cdc42	Cell division cycle 42
ChIP-seq	Chromatin immunoprecipitation sequencing
Ccnd1/d2	Cyclin d1/d2
DAG	Diacylglycerol
DIG	Digoxigenin
dpERK	Diphosphorylated ERK
DUX4	Double homeobox 4
Dusp6	Dual-specificity phosphatase 6
ETS	E26 transformation-specific transcription factors
Etv	ETS variant transcription factor
Egr-1	Early growth response-1
EMSA	Electrophoretic mobility shift assay
EGFR	Epidermal growth factor receptor
ERK	extracellular signal-regulated kinase
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
Frs2	Fibroblast growth factor receptor substrate 2
FGFRL	Fibroblast growth factor receptor-like
GO	Gene ontology
Grb2	Growth factor receptor bound protein 2
Gab1	Growth factor receptor bound protein 2-associated-binding protein 1
GDP	Guanosine diphosphate
GTP	Guanosine triphosphate
HSPG	Heparin sulphate proteoglycan
HMG-box	High Mobility Group-box

Hox	Homeobox
Hkb	Huckebein
JNK	c-Jun N-terminal kinases
KPNA3	Importin α4
IP_3	Inositol-1,4,5-trisphosphate
IEC	Intestinal epithelial cells
LEC	Lens epithelial cells
MMP	Matrix metalloprotease
MBT	Midblastula transition
MAPK	Mitogen-activated protein kinase
Mek/MAPKK	Mitogen-activated protein kinase kinase
MRS	Modified Ringers Solution
MyoD	Myoblast determination protein 1
NHEJ	Non-homologous end joining
NAM	Normal Amphibian Medium
NLS	Nuclear localisation sequence
$p90^{RSK}$	p90 ribosomal S6 kinase
PEA3	Polyoma enhancer activator 3
PI3	Phosphoinositide-3
PLCγ	Phospholipase Cγ
PCR	Polymerase chain reaction
PANTHER	Protein analysis through evolutionary relationships
AKT/PKB	Protein kinase B
PKC	Protein kinase C
RTK	Receptor tyrosine kinase
SGK1	Serum/Glucocorticoid Regulated Kinase 1
Sos	Son of sevenless
SCA1	Spinocerebellar ataxia type 1
SH2	Src homology 2
SH3	Src homology 3
Tll	Tailless
TPM	Transcripts per million

References

ADHR Consortium. (2000). Autosomal dominant hypophosphataemic rickets is associated with mutations in FGF23. Nature genetics, 26 (3), pp.345–348.

Andreu, M. J. et al. (2012a). EGFR-dependent downregulation of Capicua and the establishment of Drosophila dorsoventral polarity. Fly, 6 (4), pp.234–239.

Andreu, M. J. et al. (2012b). Mirror represses pipe expression in follicle cells to initiate dorsoventral axis formation in Drosophila. Development , 139 (6), pp.1110–1114.

Andrukhova, O. et al. (2012). FGF23 acts directly on renal proximal tubules to induce phosphaturia through activation of the ERK1/2-SGK1 signaling pathway. Bone, 51 (3), pp.621–628.

Angel, P. and Karin, M. (1991). The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. Biochimica et biophysica acta, 1072 (2-3), pp.129–157.

Astigarraga, S. et al. (2007). A MAPK docking site is critical for downregulation of Capicua by Torso and EGFR RTK signaling. The EMBO journal, 26 (3), pp.668–677.

Bahrami, S. and Drabløs, F. (2016). Gene regulation in the immediate-early response process. Advances in biological regulation, 62, pp.37–49.

Beck, C. W., Izpisúa Belmonte, J. C. and Christen, B. (2009). Beyond early development: Xenopus as an emerging model for the study of regenerative mechanisms. Developmental dynamics: an official publication of the American Association of Anatomists, 238 (6), pp.1226–1248.

Beck, F. and Stringer, E. J. (2010). The role of Cdx genes in the gut and in axial development. Biochemical Society transactions, 38 (2), pp.353–357.

Beenken, A. and Mohammadi, M. (2009). The FGF family: biology, pathophysiology and therapy. Nature reviews. Drug discovery, 8 (3), pp.235–253.

Bement, W. M., Forscher, P. and Mooseker, M. S. (1993). A novel cytoskeletal structure involved in purse string wound closure and cell polarity maintenance. The Journal of cell biology, 121 (3), pp.565–578.

Bement, W. M., Mandato, C. A. and Kirsch, M. N. (1999). Wound-induced assembly and closure of an actomyosin purse string in Xenopus oocytes. Current biology: CB, 9 (11), pp.579–587.

Benink, H. A. and Bement, W. M. (2005). Concentric zones of active RhoA and Cdc42 around single cell wounds. The Journal of cell biology, 168 (3), pp.429–439.

ten Berge, D. et al. (2008). Wnt and FGF signals interact to coordinate growth with cell fate specification during limb development. Development , 135 (19), pp.3247–3257.

Boilly, B. et al. (2000). FGF signals for cell proliferation and migration through different pathways. Cytokine & growth factor reviews, 11 (4), pp.295–302.

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

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

Buchtova, M. et al. (2015). Fibroblast growth factor and canonical WNT/β-catenin signaling cooperate in suppression of chondrocyte differentiation in experimental models of FGFR signaling in cartilage. Biochimica et biophysica acta, 1852 (5), pp.839–850.

Chesley, P. (1935). Development of the short-tailed mutant in the house mouse. The Journal of experimental zoology, 70 (3), pp.429–459.

Christen, B. and Slack, J. M. (1997). FGF-8 is associated with anteroposterior patterning and limb regeneration in Xenopus. Developmental biology, 192 (2), pp.455–466.

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

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

Crespo-Barreto, J. et al. (2010). Partial loss of ataxin-1 function contributes to transcriptional dysregulation in spinocerebellar ataxia type 1 pathogenesis. PLoS genetics, 6 (7), p.e1001021.

Danjo, Y. and Gipson, I. K. (1998). Actin 'purse string' filaments are anchored by E-cadherin-mediated adherens junctions at the leading edge of the epithelial wound, providing coordinated cell movement. Journal of cell science, 111 (Pt 22), pp.3323–3332.

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

Delfini, M.-C. et al. (2005). Control of the segmentation process by graded MAPK/ERK activation in the chick embryo. Proceedings of the National Academy of Sciences of the United States of America, 102 (32), pp.11343–11348.

Dieckgraefe, B. K. et al. (1997). ERK and p38 MAP kinase pathways are mediators of intestinal epithelial wound-induced signal transduction. Biochemical and biophysical research communications, 233 (2), pp.389–394.

Dieckgraefe, B. K. and Weems, D. M. (1999). Epithelial injury induces egr-1 and fos expression by a pathway involving protein kinase C and ERK. The American journal of physiology, 276 (2), pp.G322–G330.

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

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

Drumheller PD, H. J. A. (1991). Local modulation of intracellular calcium levels near a single-cell wound in human endothelial monolayers. Arterioscler Thromb, 11, pp.1258–1265.

Duffy, J. B. and Perrimon, N. (1994). The torso pathway in Drosophila: lessons on receptor tyrosine kinase signaling and pattern formation. Developmental biology, 166 (2), pp.380–395.

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

Falardeau, J. et al. (2008). Decreased FGF8 signaling causes deficiency of gonadotropin-releasing hormone in humans and mice. The Journal of clinical investigation, 118 (8), pp.2822–2831.

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

Forés, M. et al. (2017). A new mode of DNA binding distinguishes Capicua from other HMG-box factors and explains its mutation patterns in cancer. PLoS genetics, 13 (3), p.e1006622.

Fryer, J. D. et al. (2011). Exercise and genetic rescue of SCA1 via the transcriptional repressor Capicua. Science, 334 (6056), pp.690–693.

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

Goldhamer, D. J. et al. (1992). Regulatory elements that control the lineage-specific expression of myoD. Science, 256 (5056), pp.538–542.

Guille, M. (Ed). (1999). Molecular Methods in Developmental Biology: Xenopus and Zebrafish. Humana Press.

Hadari, Y. R. et al. (2001). Critical role for the docking-protein FRS2 alpha in FGF receptor-mediated signal transduction pathways. Proceedings of the National Academy of Sciences of the United States of America, 98 (15), pp.8578–8583.

Herschman, H. R. (1991). Primary response genes induced by growth factors and tumor promoters. Annual review of biochemistry, 60, pp.281–319.

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

Isaacs, H. V., Pownall, M. E. and Slack, J. M. (1995). eFGF is expressed in the dorsal midline of Xenopus laevis. The International journal of developmental biology, 39 (4), pp.575–579.

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

Itoh, N. (2010). Hormone-like (endocrine) Fgfs: their evolutionary history and roles in development, metabolism, and disease. Cell and tissue research, 342 (1), pp.1–11.

Jaakkola, P. et al. (1998). Wound reepithelialization activates a growth factor-responsive enhancer in migrating keratinocytes. FASEB journal: official publication of the Federation of American Societies for Experimental Biology, 12 (11), pp.959–969.

Jiménez, G. et al. (2000). Relief of gene repression by torso RTK signaling: role of capicua in Drosophila terminal and dorsoventral patterning. Genes & development, 14 (2), pp.224–231.

Jiménez, G., Shvartsman, S. Y. and Paroush, Z. 'ev. (2012). The Capicua repressor--a general sensor of RTK signaling in development and disease. Journal of cell science, 125 (Pt 6), pp.1383–1391.

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

Kawamura-Saito, M. et al. (2006). Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation. Human molecular genetics, 15 (13), pp.2125–2137.

Kawano, Y. and Kypta, R. (2003). Secreted antagonists of the Wnt signalling pathway. Journal of cell science, 116 (Pt 13), pp.2627–2634.

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

Kim, E. et al. (2013). Structural basis of protein complex formation and reconfiguration by polyglutamine disease protein Ataxin-1 and Capicua. Genes & development, 27 (6), pp.590–595.

Kim, J. et al. (1998). Mesoderm induction by heterodimeric AP-1 (c-Jun and c-Fos) and its involvement in mesoderm formation through the embryonic fibroblast growth factor/Xbra autocatalytic loop during the early development of Xenopus embryos. The Journal of biological chemistry, 273 (3), pp.1542–1550.

Kimura, K. et al. (1996). Regulation of myosin phosphatase by Rho and Rho-associated kinase (Rho-kinase). Science, 273 (5272), pp.245–248.

King, M. G. (2019) Investigating the role of Capicua in mediating FGF transcriptional regulation in X. tropicalis. PhD thesis, University of York.

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

LaBonne, C. and Whitman, M. (1997). Localization of MAP kinase activity in early Xenopus embryos: implications for endogenous FGF signaling. Developmental biology, 183 (1), pp.9–20.

Lam, Y. C. et al. (2006). ATAXIN-1 interacts with the repressor Capicua in its native complex to cause SCA1 neuropathology. Cell, 127 (7), pp.1335–1347.

de Launoit, Y. et al. (2006). The Ets transcription factors of the PEA3 group: transcriptional regulators in metastasis. Biochimica et biophysica acta, 1766 (1), pp.79–87.

Lea, R. et al. (2009). Temporal and spatial expression of FGF ligands and receptors during Xenopus development. Developmental dynamics: an official publication of the American Association of Anatomists, 238 (6), pp.1467–1479.

Lee, M. T., Bonneau, A. R. and Giraldez, A. J. (2014). Zygotic genome activation during the maternal-to-zygotic transition. Annual review of cell and developmental biology, 30, pp.581–613.

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

Lei, Y. et al. (2012). Efficient targeted gene disruption in Xenopus embryos using engineered transcription activator-like effector nucleases (TALENs). Proceedings of the National Academy of Sciences of the United States of America, 109 (43), pp.17484–17489.

Li, J. et al. (2013). ERK and phosphoinositide 3-kinase temporally coordinate different modes of actin-based motility during embryonic wound healing. Journal of cell science, 126 (Pt 21), pp.5005–5017.

Lim, J. et al. (2008). Opposing effects of polyglutamine expansion on native protein complexes contribute to SCA1. Nature, 452 (7188), pp.713–718.

Löhr, U. et al. (2009). Antagonistic action of Bicoid and the repressor Capicua determines the spatial limits of Drosophila head gene expression domains. Proceedings of the National Academy of Sciences of the United States of America, 106 (51), pp.21695–21700.

Lombardo, A., Isaacs, H. V. and Slack, J. M. (1998). Expression and functions of FGF-3 in Xenopus development. The International journal of developmental biology, 42 (8), pp.1101–1107.

Maddaluno, L., Urwyler, C. and Werner, S. (2017). Fibroblast growth factors: key players in regeneration and tissue repair. Development , 144 (22), pp.4047–4060.

Martin, P. et al. (1994). Repair of excisional wounds in the embryo. Eye , 8 (Pt 2), pp.155–160.

Martin, P. and Lewis, J. (1992). Actin cables and epidermal movement in embryonic wound healing. Nature, 360 (6400), pp.179–183.

Mi, H. et al. (2013). Large-scale gene function analysis with the PANTHER classification system. Nature protocols, 8 (8), pp.1551–1566.

Mi, H. et al. (2019). PANTHER Version 14: More Genomes, a New PANTHER GO-Slim and Improvements in Enrichment Analysis Tools. Nucleic Acids Research, 47 (D1): pp.D419–26.

Milunsky, J. M. et al. (2006). LADD syndrome is caused by FGF10 mutations. Clinical genetics, 69 (4), pp.349–354.

Moore, J. K. and Haber, J. E. (1996). Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in Saccharomyces cerevisiae. Molecular and cellular biology, 16 (5), pp.2164–2173.

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

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

Nummenmaa, E. et al. (2015). Effects of FGF-2 and FGF receptor antagonists on MMP enzymes, aggrecan, and type II collagen in primary human OA chondrocytes. Scandinavian journal of rheumatology, 44 (4), pp.321–330.

Oh, S., Shin, S. and Janknecht, R. (2012). ETV1, 4 and 5: an oncogenic subfamily of ETS transcription factors. Biochimica et biophysica acta, 1826 (1), pp.1–12.

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

Ong, S. H. et al. (2001). Stimulation of phosphatidylinositol 3-kinase by fibroblast growth factor receptors is mediated by coordinated recruitment of multiple docking proteins. Proceedings of the National Academy of Sciences of the United States of America, 98 (11), pp.6074–6079.

Ornitz, D. M. (2000). FGFs, heparan sulfate and FGFRs: complex interactions essential for development. BioEssays: news and reviews in molecular, cellular and developmental biology, 22 (2), pp.108–112.

Ornitz, D. M. and Itoh, N. (2015). The Fibroblast Growth Factor signaling pathway. Wiley interdisciplinary reviews. Developmental biology, 4 (3), pp.215–266.

Pawson, T. et al. (1993). Proteins with SH2 and SH3 domains couple receptor tyrosine kinases to intracellular signalling pathways. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 340 (1293), pp.279–285.

Pearson, J. C. et al. (2009). Multiple transcription factor codes activate epidermal wound-response genes in Drosophila. Proceedings of the National Academy of Sciences of the United States of America, 106 (7), pp.2224–2229.

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

Pillemer, G. et al. (1998). Nested expression and sequential downregulation of the Xenopus caudal genes along the anterior-posterior axis. Mechanisms of development, 71 (1-2), pp.193–196.

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

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

Rossant, J., Ciruna, B. and Partanen, J. (1997). FGF signaling in mouse gastrulation and anteroposterior patterning. Cold Spring Harbor symposia on quantitative biology, 62, pp.127–133.

Rupp, R. A. and Weintraub, H. (1991). Ubiquitous MyoD transcription at the midblastula transition precedes induction-dependent MyoD expression in presumptive mesoderm of X. laevis. Cell, 65 (6), pp.927–937.

Salerno, D. M. et al. (2012). Gadd45a and Gadd45b modulate innate immune functions of granulocytes and macrophages by differential regulation of p38 and JNK signaling. Journal of cellular physiology, 227 (11), pp.3613–3620.

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

Schmitt, M. W., Loeb, L. A. and Salk, J. J. (2016). The influence of subclonal resistance mutations on targeted cancer therapy. Nature reviews. Clinical oncology, 13 (6), pp.335–347.

Scotet, E. and Houssaint, E. (1998). Exon III splicing switch of fibroblast growth factor (FGF) receptor-2 and -3 can be induced by FGF-1 or FGF-2. Oncogene, 17 (1), pp.67–76.

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

Slack, J. M. et al. (1996). The role of fibroblast growth factors in early Xenopus development. Biochemical Society symposium, 62, pp.1–12.

Slack, J. M. and Forman, D. (1980). An interaction between dorsal and ventral regions of the marginal zone in early amphibian embryos. Journal of embryology and experimental morphology, 56, pp.283–299.

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

Sosnowski, R. G., Feldman, S. and Feramisco, J. R. (1993). Interference with endogenous ras function inhibits cellular responses to wounding. The Journal of cell biology, 121 (1), pp.113–119.

Stanisstreet, M. (1982). Calcium and wound healing in Xenopus early embryos. Journal of embryology and experimental morphology, 67, pp.195–205.

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

Stramer, B. et al. (2008). Gene induction following wounding of wild-type versus macrophage-deficient Drosophila embryos. EMBO reports, 9 (5), pp.465–471.

Tanaka, M., Yoshimoto, T. and Nakamura, T. (2017). A double-edged sword: The world according to Capicua in cancer. Cancer science, 108 (12), pp.2319–2325.

Tao, J. et al. (2005). BMP4-dependent expression of Xenopus Grainyhead-like 1 is essential for epidermal differentiation. Development , 132 (5), pp.1021–1034.

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

Teven, C. M. et al. (2014). Fibroblast growth factor (FGF) signaling in development and skeletal diseases. Genes & diseases, 1 (2), pp.199–213.

Tindall, A. J. et al. (2007). Expression of enzymes involved in thyroid hormone metabolism during the early development of Xenopus tropicalis. Biology of the cell / under the auspices of the European Cell Biology Organization, 99 (3), pp.151–163.

Ting, S. B. et al. (2005). A homolog of Drosophila grainy head is essential for epidermal integrity in mice. Science, 308 (5720), pp.411–413.

Tiong, K. H., Mah, L. Y. and Leong, C.-O. (2013). Functional roles of fibroblast growth factor receptors (FGFRs) signaling in human cancers. Apoptosis: an international journal on programmed cell death, 18 (12), pp.1447–1468.

Tran, P. O. et al. (1999). A wound-induced [Ca2+]i increase and its transcriptional activation of immediate early genes is important in the regulation of motility. Experimental cell research, 246 (2), pp.319–326.

Trueb, B. et al. (2003). Characterization of FGFRL1, a novel fibroblast growth factor (FGF) receptor preferentially expressed in skeletal tissues. The Journal of biological chemistry, 278 (36), pp.33857–33865.

Tseng, A.-S. K. et al. (2007). Capicua regulates cell proliferation downstream of the receptor tyrosine kinase/ras signaling pathway. Current biology: CB, 17 (8), pp.728–733.

Turner, N. and Grose, R. (2010). Fibroblast growth factor signalling: from development to cancer. Nature reviews. Cancer, 10 (2), pp.116–129.

Wang, E. et al. (2003). Electric fields and MAP kinase signaling can regulate early wound healing in lens epithelium. Investigative ophthalmology & visual science, 44 (1), pp.244–249.

White, K. E. et al. (2000). Autosomal dominant hypophosphataemic rickets is associated with mutations in FGF23. Nature Genetics, 26, pp.345–348

Whitmarsh, A. J. and Davis, R. J. (1996). Transcription factor AP-1 regulation by mitogen-activated protein kinase signal transduction pathways. Journal of molecular medicine , 74 (10), pp.589–607.

Wuechner, C. et al. (1996). Developmental expression of splicing variants of fibroblast growth factor receptor 3 (FGFR3) in mouse. The International journal of developmental biology, 40 (6), pp.1185–1188.

Yates, S. and Rayner, T. E. (2002). Transcription factor activation in response to cutaneous injury: role of AP-1 in reepithelialization. Wound repair and regeneration: official publication of the Wound Healing Society [and] the European Tissue Repair Society, 10 (1), pp.5–15.

Yeh, B. K. et al. (2003). Structural basis by which alternative splicing confers specificity in fibroblast growth factor receptors. Proceedings of the National Academy of Sciences of the United States of America, 100 (5), pp.2266–2271.

Yoshimoto, T. et al. (2017). CIC-DUX4 Induces Small Round Cell Sarcomas Distinct from Ewing Sarcoma. Cancer research, 77 (11), pp.2927–2937.

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

Zoghbi, H. Y. and Orr, H. T. (2009). Pathogenic mechanisms of a polyglutamine-mediated neurodegenerative disease, spinocerebellar ataxia type 1. The Journal of biological chemistry, 284 (12), pp.7425–7429.

Zu, T. et al. (2004). Recovery from polyglutamine-induced neurodegeneration in conditional SCA1 transgenic mice. The Journal of neuroscience: the official journal of the Society for Neuroscience, 24 (40), pp.8853–8861.