

---

Investigating the Regulation of *fos*  
During Embryonic Wound Healing  
and Morphogenesis in *Xenopus*

---

Helena West

MSc by Research

University of York

Biology

June 2022

# Abstract

Extracellular signal-related kinase (ERK) regulates gene expression through substrate phosphorylation. The conserved transcriptional repressor Capicua (CIC) is negatively regulated via ERK-mediated phosphorylation. Both fibroblast growth factor (FGF) signalling and embryonic wounding activate ERK, however the contribution of CIC to transcriptional regulation in these contexts is understudied. It was hypothesised that ERK relieves CIC-mediated repression downstream of FGF signalling and post-wounding during *Xenopus* development; this project aimed to identify targets of CIC/FGF and investigate their regulation in these contexts.

Here, 44 putative CIC/FGF target genes, including the AP-1 gene *fos*, were identified through gene-level differential expression analysis of RNA-seq data from FGF4 overexpressing and CIC knockdown *X. tropicalis* embryos. *Xenopus* embryos were treated with FGF (SU5402), MEK (PD0325901) or nitric oxide (NO) synthase (TRIM) chemical inhibitors and the expression of *fos* and activated ERK investigated via *in situ* hybridisation and immunostaining, respectively. Gastrula stage expression of *fos* was found to be FGF-dependent, and CIC binding motifs were identified within a conserved region of *fos* intron 1. Corroborating with CIC/ERK-mediated regulation, *fos* expression was reduced following ERK inhibition during embryonic wound healing. Inhibiting NO production elevated activated ERK expression post-wounding, thus NO may negatively regulate *fos* expression via ERK. Inhibiting ERK activation similarly reduced *fos* expression during neurulation, suggesting comparable mechanisms may govern wound healing and morphogenesis. Other CIC targets representing AP-1 genes, namely *jun* and *atf3*, were also found to be expressed post-wounding and during neurulation. Overall therefore, the data support the hypothesis, and propose a novel role for CIC in intron mediated enhancement of the CIC/FGF target gene *fos*, and a wider role for CIC in regulating AP-1 gene expression during wound healing and morphogenesis.

# Contents

<b>Abstract</b>	<b>2</b>
<b>Contents</b>	<b>3</b>
<b>List of Tables</b>	<b>7</b>
<b>List of Figures</b>	<b>8</b>
<b>List of Accompanying Material</b>	<b>10</b>
<b>Declaration</b>	<b>11</b>
<b>Acknowledgements</b>	<b>12</b>
<b>Chapter 1: Introduction</b>	<b>13</b>
1.1 ERK is the effector of the Ras-MAPK signalling pathway	13
1.2 Transcriptional regulation downstream of ERK	13
1.3 The transcriptional repressor CIC is negatively regulated by ERK	16
1.4 Phosphorylation may cause relocalisation or degradation of CIC	18
1.5 Co-repressors facilitate repression of CIC target genes	18
1.6 CIC functions as a tumour suppressor	19
1.7 Aims and hypotheses	19
<b>Chapter 2: Materials and Methods</b>	<b>21</b>
2.1 Embryological Methods	21
2.1.1 Ethics statement	21
2.1.2 <i>Xenopus</i> embryo culture	21
2.1.3 Wounding	21
2.1.4 Treatment with chemical inhibitors	21
2.1.5 Microinjection	23
2.1.6 Imaging	23
2.2 Molecular Biology Methods	23
2.2.1 Agarose gel electrophoresis	23
2.2.2 RNA extraction	23
2.2.3 First strand cDNA synthesis	24
2.2.4 PCR-based cloning of full-length <i>fos</i> , and <i>jun</i>	24

2.2.5 Plasmid transformation and purification	26
2.2.6 Messenger RNA synthesis	26
2.2.7 Cloning of <i>in situ</i> hybridisation probe template	26
2.2.8 Synthesis of <i>in situ</i> hybridisation antisense RNA probe	27
2.2.9 Whole-mount RNA <i>in situ</i> hybridisation	28
2.2.10 Modified <i>in situ</i> hybridisation	29
2.2.11 Whole-mount immunostaining	29
2.2.12 Western blotting	30
2.3 Bioinformatics	31
2.3.1 Gene-level differential expression analysis of RNA-Seq data	31
2.3.2 Intersection of gene lists	31
2.3.2 Gene ontology analysis	32
2.3.3 Conservation analysis of <i>fos</i> genomic sequence	32
2.3.4 Motif enrichment analysis	32
<b>Chapter 3: FGF-dependent Regulation of the Transcriptional Repressor CIC</b>	<b>33</b>
3.1 Introduction	33
3.1.1 FGF signalling has diverse roles during embryogenesis	33
3.1.2 FGF ligands and receptors	33
3.1.3 The FGF signalling cascade	34
3.1.4 Transcriptional regulators of FGF signalling downstream of the Ras-MAPK pathway	36
3.1.5 CIC may regulate of a subset of FGF target genes	36
3.1.6 Chapter hypothesis and aims	37
3.2 Results	38
3.2.1 Gene-level differential expression analysis of FGF4 overexpressing and CIC knockdown <i>X. tropicalis</i> embryos	38
3.2.2 A subset of FGF target genes are putatively regulated by CIC	40
3.2.3 Gene-ontology enrichment analysis for genes significantly upregulated following CIC knockdown	41
3.2.4 Spatial and temporal co-localisation of <i>fos</i> and dpERK in gastrula stage <i>X. tropicalis</i>	45
3.2.5 FGF and MEK inhibition prevent <i>fos</i> induction in gastrula stage <i>X. tropicalis</i>	46

3.2.6 CIC binding sites exist within a highly conserved region of <i>fos</i> intron 1	48
3.3 Discussion	51
3.3.1 Gene-level differential expression analysis output of CIC target genes	51
3.3.2 CIC putatively regulates a subset of FGF target genes	51
3.3.3 A role for CIC in feedback inhibition of FGF signalling	52
3.3.4 The AP-1 component gene <i>fos</i> is putatively a target of CIC/FGF	53
3.3.5 Early gastrula expression of <i>fos</i> is FGF and ERK-dependent in <i>Xenopus</i>	53
3.3.6 Candidate regulation of a <i>fos</i> intronic promoter by CIC	54
3.3.7 Co-operative transcriptional and post-translational regulation of Fos expression by dpERK	54
3.3.8 Summary	55
<b>Chapter 4: The Regulation of <i>fos</i> During Wound and Morphogenesis</b>	<b>56</b>
4.1 Introduction	56
4.1.1 The imperfect nature of adult wound healing	56
4.1.2 Embryonic wound healing – the gold standard of wound healing	56
4.1.3 The early phase of embryonic wound healing	56
4.1.4 The late phase of embryonic wound healing	57
4.1.5 Wound-induced changes in gene expression are relatively understudied	58
4.1.6 Transcription of the immediate early gene <i>fos</i> is induced rapidly in response to wounding	58
4.1.7 Candidate regulation of <i>fos</i> by CIC during embryonic wound healing and neurulation	58
4.1.8 Chapter hypothesis and aims	59
4.2 Results	60
4.2.1 Inhibition of MAPK signalling reduces <i>fos</i> expression during <i>X. tropicalis</i> wound healing	60
4.2.2 The nitric oxide synthesis inhibitor TRIM heightens dpERK expression during <i>X. laevis</i> wound healing	61
4.2.3 Detection of Fos by anti-Fos (2G9) antibody	63
4.2.4 Treatment with nitric oxide synthesis inhibitor TRIM and the expression of Fos protein during <i>X. laevis</i> wound healing	64
4.2.5 A wider role for CIC during embryonic wound healing	66

4.2.6 Candidate regulation of AP-1 component genes by CIC during wound healing	71
4.2.7 <i>rasl11b</i> is induced rapidly post-wounding	72
4.2.8 Spatial and temporal co-localisation of <i>fos</i> and dpERK in neurula stage <i>X. tropicalis</i>	73
4.2.9 MEK but not FGF inhibition reduces <i>fos</i> expression along the dorsal neural closure in <i>X. tropicalis</i>	74
4.2.10 AP-1 component gene expression during neurulation	77
4.3 Discussion	79
4.3.1 Reduction in <i>fos</i> expression during embryonic wound healing following ERK inhibition may reflect loss of intron-mediated enhancement	79
4.3.2 Nitric oxide may serve as a negative regulator of dpERK and resultantly <i>fos</i> during embryonic wound healing	80
4.3.3 CIC as a general regulator of wound-responsive genes	82
4.3.4 Similarities exist in the regulation of <i>fos</i> expression during wound healing and neurulation	83
4.3.5 AP-1 activity during neurulation	85
<b>Chapter 5: General Discussion</b>	<b>86</b>
5.1 Summary	86
5.2 Limitations and future work	87
5.3 Conclusions and implications	88
<b>Appendices</b>	<b>90</b>
Appendix A: Sequence of full-length <i>fos</i> and <i>jun</i> inserts in pCS2+ vector	90
Appendix B: Predicted protein sequence of Fos and Jun clones	92
Appendix C: R code for gene-level differential expression analysis	93
<b>Abbreviations</b>	<b>97</b>
<b>References</b>	<b>99</b>

# List of Tables

Table		Page
1	Details of primers used to amplify <i>X. tropicalis</i> full-length <i>fos</i> (accession: NM_001016200) and full-length <i>jun</i> (accession: XM_031900902.1).	25
2	Enzymes and buffers used for template linearisation, and polymerases used for <i>in situ</i> hybridisation probe synthesis.	27
3	Length of time (minutes) <i>Xenopus</i> embryos were subject to 10 µg/ml Proteinase K treatment during <i>in situ</i> hybridisation protocol, with respect to developmental stage.	29
4	GORilla gene-ontology analysis of biological processes for genes significantly upregulated following CIC knockdown.	43
5	GORilla gene-ontology analysis of biological components for genes significantly upregulated following CIC knockdown.	44
6	Motif scanning for CIC consensus motif in <i>X. tropicalis</i> , <i>H. sapien</i> , <i>M. musculus</i> and <i>G. gallus fos</i> intron 1.	50
7	Lists of genes significantly upregulated by both CIC knockdown in <i>X. tropicalis</i> embryos and wounding in <i>X. laevis</i> embryos at 30, 60 and 90 minutes post-wounding.	68
8	Phenotypic quantification of dpERK immunostaining along the dorsal neural closure and in anterior and posterior domains of late neurula stage <i>X. tropicalis</i> embryos.	76

# List of Figures

Figure		Page
1	X-ray crystallography structure of the human c-Fos/c-Jun heterodimer basic leucine zipper domain in complex with DNA.	15
2	Schematic of the domain structure of human Capicua (CIC) protein isoforms.	17
3	Schematic representation of overall project hypothesis.	20
4	Schematic representation of the Ras-MAPK pathway induced downstream of FGF signalling.	35
5	Volcano plots of RNA-seq gene-level differential expression analysis for FGF4 overexpressing and CIC knockdown stage 14 <i>X. tropicalis</i> embryos.	39
6	Venn diagram of overlap between significantly upregulated and downregulated genes following FGF4 overexpression and CIC knockdown in stage 14 <i>X. tropicalis</i> embryos.	40
7	GOrilla gene-ontology analysis of biological processes for genes significantly upregulated following CIC knockdown.	42
8	<i>X. tropicalis</i> gastrula stage series of dpERK and <i>fos</i> expression analysed via immunostaining and <i>in situ</i> hybridization, respectively.	46
9	SU5402 and PD0325901 inhibit early gastrula stage dpERK and <i>fos</i> expression in <i>X. tropicalis</i> embryos, as assayed via immunostaining and <i>in situ</i> hybridisation, respectively.	47
10	Conservation of <i>fos</i> genetic locus and schematic representation of candidate CIC binding sites for <i>X. tropicalis fos</i> .	49
11	PD0325901 inhibits dpERK expression and reduces <i>fos</i> expression 30 minutes post-wounding in <i>X. tropicalis</i> embryos, as assayed via immunostaining and <i>in situ</i> hybridisation, respectively.	61
12	Wounding time-course of dpERK expression in <i>X. laevis</i> embryos, analysed via immunostaining.	62

13	Expression of dpERK in <i>X. laevis</i> embryos at 50, 70 and 90 minutes post-wounding following culture in either NAM/10 or TRIM, as assayed via immunostaining.	63
14	Anti-Fos 2G9 can detect overexpressed exogenous <i>X. tropicalis</i> Fos via Western blotting and immunostaining in <i>X. laevis</i> embryos.	64
15	Expression of Fos protein in <i>X. laevis</i> embryos cultured in NAM/10 or TRIM as assayed by immunostaining with anti-Fos 2G9.	65
16	Venn diagram of overlap between genes significantly upregulated following CIC knockdown, and genes significantly upregulated at 30, 60 and 90 minutes post-wounding in <i>Xenopus</i> .	67
17	Heat-map depicting expression patterns of 48/51 genes during development which are significantly upregulated following CIC knockdown and post-wounding.	70
18	<i>In situ</i> hybridisation for <i>atf3</i> and <i>jun</i> post-wounding in <i>X. tropicalis</i> embryos.	72
19	<i>rasl11b</i> is induced rapidly post-wounding in early tailbud stage <i>X. tropicalis</i> embryos as analysed by <i>in situ</i> hybridisation.	73
20	<i>X. tropicalis</i> neurula stage series of dpERK and <i>fos</i> expression analysed via immunostaining and <i>in situ</i> hybridization, respectively.	74
21	dpERK immunostaining of late neurula stage <i>X. tropicalis</i> embryos treated with PD0325901 or SU5402.	75
22	<i>In situ</i> hybridisations for <i>fos</i> in late neurula stage <i>X. tropicalis</i> embryos treated with PD0325901 or SU5402.	77
23	<i>In situ</i> hybridisations for <i>atf3</i> and <i>jun</i> in neurula stage <i>X. tropicalis</i> embryos.	78
24	Hypothesised axis of regulation between nitric oxide, dpERK and <i>fos</i> during <i>Xenopus</i> embryonic wound healing.	81

# List of Accompanying Material

Additional File		Link
S1	Filtered (TPM > 1) and unfiltered gene-level differential expression analysis from FGF4 overexpressing and CIC knockdown <i>X. tropicalis</i> embryos.	<a href="#">Additional File S1</a>
S2	Genes significantly up- or downregulated following FGF4 overexpression ( $q$ -value < 0.1, effect size > 1.5, < 1/1.5).	<a href="#">Additional File S2</a>
S3	Genes significantly up- or downregulated following CIC knockdown ( $q$ -value < 0.1, effect size > 1.5, < 1/1.5)	<a href="#">Additional File S3</a>
S4	Lists of gene intersections for Venn diagram of genes significantly up- and downregulated by FGF4 overexpression and CIC knockdown.	<a href="#">Additional File S4</a>
S5	Gene ontology enrichment analysis for biological processes and components enriched following CIC knockdown.	<a href="#">Additional File S5</a>
S6	FASTA file of <i>fos</i> coding sequence alongside 5 kb upstream and 1 kb downstream from <i>H. sapien</i> , <i>M. musculus</i> , <i>G. gallus</i> , and <i>X. tropicalis</i> for mVISTA/LAGAN conservation analysis and FIMO motif scanning.	<a href="#">Additional File S6</a>
S7	FASTA file of <i>fos</i> intron 1 sequence from <i>H. sapien</i> , <i>M. musculus</i> , <i>G. gallus</i> and <i>X. tropicalis</i> for SEA motif enrichment analysis.	<a href="#">Additional File S7</a>
S8	Output of SEA motif enrichment analysis for CIC consensus (TSAATGRA) in <i>fos</i> intron 1 from <i>H. sapien</i> , <i>M. musculus</i> , <i>G. gallus</i> and <i>X. tropicalis</i> .	<a href="#">Additional File S8</a>
S9	Genes significantly upregulated at 30, 60 and 90 minutes post-wounding in <i>X. laevis</i> ( $q$ -value < 0.1, fold change > 1.5) and following CIC knock down in <i>X. tropicalis</i> ( $q$ -value < 0.1, effect size > 1.5). Original wounding data taken from Abaffey et al. (2019).	<a href="#">Additional File S9</a>

# Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References. Work was conducted under the supervision of Dr. Harry Isaacs and Professor Betsy Pownall. The *X. tropicalis* FGF4 overexpression and CIC knockdown RNA-seq data was collected by a former member of the Isaacs lab, Michael King, and processed by members of staff at the University of York Biology Technology Facility. John Davey of the University of York Biology Technology Facility performed initial RNA-quality control of the RNA-seq data and alignment of the reads to the reference genome. This project utilised the R Studio Salmon output from John Davey's analysis to perform a gene-level differential expression analysis. The *X. laevis* RNA-seq data and differential expression analysis concerning wound-induced changes in gene expression was taken from Abaffy et al. (2019). RNA-seq data from Owens et al. (2016) was used to obtain the normal expression pattern of 48 genes during *X. tropicalis* development.

# Acknowledgements

I would like to say a big thank you to my supervisors Dr. Harry Isaacs and Professor Betsy Pownall for all their support, expertise, and 'give it a go' attitude throughout this project. Alongside creating a friendly and welcoming lab environment, they provided me with invaluable guidance and also the freedom to explore my own ideas - I know I couldn't possibly have asked for better supervisors! I would also like to thank Dr. Andrew Holding for his advice, feedback and time as a member of my thesis advisory panel. Last but not least, I'd like to thank the rest of L2 Frog, Alex, Laura and Matt, who have helped me laugh through all the times things didn't quite go to plan.

# Chapter 1: Introduction

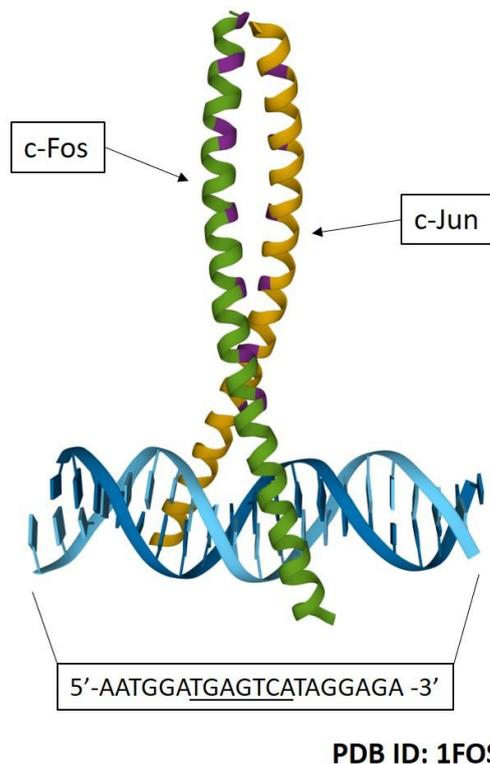
## 1.1 ERK is the effector of the Ras-MAPK signalling pathway

Whilst the advent of high-throughput transcriptomic analyses have illuminated many of the genes underlying distinct cell physiologies, many of the transcriptional regulators governing differential gene expression *in vivo* remain elusive. Extracellular signal-regulated kinase (ERK), a mitogen-activated protein kinase (MAPK), is able to directly regulate gene expression through phosphorylation of transcription factors (Yoon and Seger, 2006). In addition, ERK can also indirectly modulate gene expression through activation of other kinases, including members of the p90 ribosomal S6 kinase (RSK) family, which in turn phosphorylate and regulate the activity of transcription factors (Frödin and Gammeltoft, 1999). In mammals two isoforms of ERK exist, ERK1 and ERK2, however there is thought to be considerable redundancy and therefore ERK is used here to collectively refer to ERK1 and ERK2 (Frémin et al., 2015). ERK itself is commonly activated downstream of receptor tyrosine kinase (RTK) signalling in response to mitogenic stimuli (Lemmon and Schlessinger, 2010). RTK signalling activates the Ras-MAPK cascade which culminates in di-phosphorylation and activation of ERK (dpERK) by mitogen-activated protein kinase kinase (MEK) (Lemmon and Schlessinger, 2010). The Ras-MAPK pathway has gained particular attention due to its prominent role in oncogenesis, with constitutive activation of the pathway being a frequent feature of many cancers (Santarpia, Lippman and El-Naggar, 2012). As the effector of the pathway, uncovering targets of ERK is imperative to understanding aberrant gene expression downstream of pathological Ras-MAPK signalling. Whilst several phosphotargets of dpERK have already been identified, including activator protein 1 (AP-1) and E26 transformation-specific (Ets) transcription factors, due to the complexity and diversity of RTK signalling, knowledge is still putatively incomplete (Ünal, Uhlitz and Blüthgen, 2017).

## 1.2 Transcriptional regulation downstream of ERK

AP-1 is a dimeric transcription factor with multiple variants. Fos, Jun, Atf and Maf family members dimerise via basic leucine-zipper (bZip) domains to form the bipartite DNA-binding domain of AP-1 and confer sequence specificity (Chiu et al., 1988; Halazonetis et al., 1988; Glover and Harrison, 1995; Karin, Liu and Zandi, 1997). To exemplify, Fos family members, including Fos, FosB, and Fos-related antigen 1 and 2 (Fra-1/Fra-2), can heterodimerize with Jun family members, including Jun, JunB and JunD, to form AP-1 (figure 1) (Chiu et al., 1988; Halazonetis et al., 1988; Garces de Los Fayos Alonso et al., 2018). Unlike Fos family members, Jun family members can

also additionally homodimerise to form AP-1 (Karin, Liu and Zandi, 1997). Fos/Jun and Jun/Jun AP-1 transcription factors regulate gene expression through binding to the palindromic AP-1 sequence motif TGAGTCA, often referred to as a TPA-responsive element (TRE), in the promoter and/or enhancer regions of target genes (Glover and Harrison, 1995; Lee, Mitchell and Tjian, 1987). Alongside homodimerization and heterodimerization with Fos, Jun family members can also heterodimerize with Atf family members, including Atf2-7 and bAtf, with Atf family members themselves also capable of homodimerization to form AP-1 (Karin, Liu and Zandi, 1997; Garces de Los Fayos Alonso et al., 2018). Jun/Atf and Atf/Atf AP-1 transcription factors bind preferentially to the AP-1 sequence motif TGACGTCA, also known as the cAMP-responsive element (CRE) (Karin, Liu and Zandi, 1997). Evidently, transcriptional regulation by AP-1 is multifaceted and complex. AP-1 activity can be positively or negatively regulated through post-translational modification of constituent proteins by dpERK. For example, the stability of Fos, and hence AP-1 activity, is increased by phosphorylation of Fos at the C-terminus by dpERK and/or p90 RSK (Okazaki and Sagata, 1995; Chen et al., 1996; Murphy et al., 2002). In contrast, phosphorylation of Jun by dpERK adjacent to the bZip domain limits its DNA-binding potential (Gazon et al., 2017). Broadly, AP-1 is implicated in numerous cellular processes including proliferation, differentiation and apoptosis (Karin, Liu and Zandi, 1997)



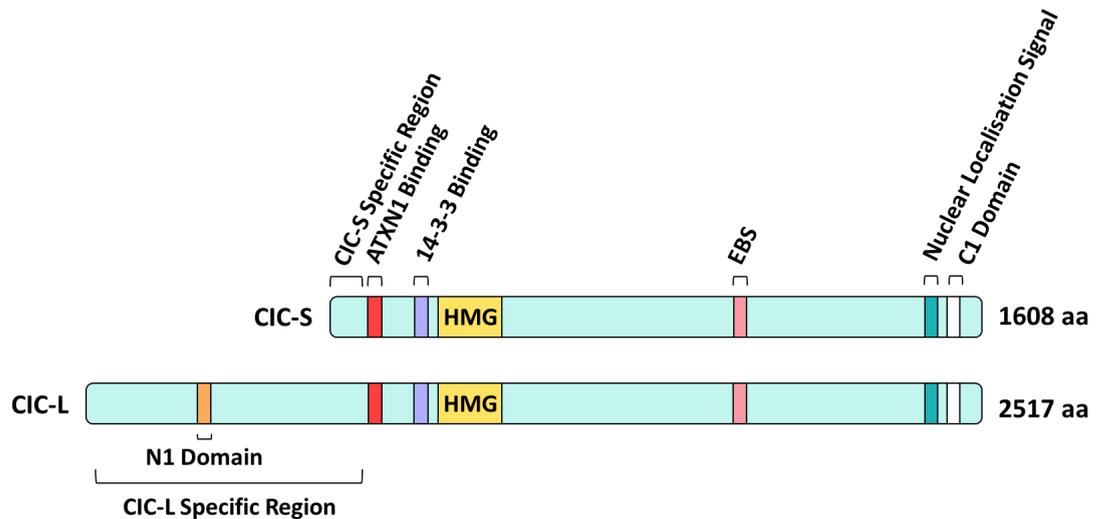
**Figure 1: X-ray crystallography structure of the human c-Fos/c-Jun heterodimer basic leucine zipper domain in complex with DNA.**  $\alpha$ -helical domains of c-Fos and c-Jun peptides heterodimerize through a leucine zipper to form the transcription factor AP-1. The AP-1 DNA-binding domain is bipartite and composed of basic regions of c-Fos and c-Jun peptides which probe into adjacent major grooves of the DNA. c-Fos peptide is highlighted in green, c-Jun peptide highlighted in yellow, and leucine residues, which represent every seventh amino acid within the zipper region, are highlighted in purple. The 5' to 3' sequence of darker blue DNA chain is indicated, the AP-1 binding sequence (TRE motif) is underlined. Note that human c-Fos and c-Jun are *Xenopus* Fos and Jun homologs. PDB ID: 1FOS (Sehna et al., 2021)

Other established phosphotargets of dpERK include ETS domain containing transcription factors. The ETS domain encodes a conserved winged helix-turn-helix tertiary structure which confers binding specificity for a core purine-rich GGA(A/T) DNA sequence motif (Karim et al., 1990; Nye et al., 1992; Wei et al., 2010). The transcription factor Elk-1 is a member of the ternary complex factor (TCF) subfamily of Ets transcription factors (Sharrocks, 2001). Phosphorylation of Elk-1 by dpERK enhances assembly of a ternary complex which encompasses Elk-1 and serum response factor (SRF), alongside other factors, and binds to promoter and/or enhancer regions containing a serum responsive element (SRE) (Gille et al., 1995; Cruzalegui, Cano and Treisman, 1999). In response to mitogenic stimuli, phosphorylation of Elk-1 by dpERK increases the induction potential of SRE regulated target genes, including notably the AP-1 component gene *fos* (Gille et al., 1995; Babu et al., 2000). Amongst others, the

polyomavirus enhancer activator 3 (PEA3) subfamily of Ets transcription factors, which includes ETV1, ETV4, ETV5, represent further Ets transcription factors known to be regulated downstream of Ras-MAPK signalling and dpERK (Qi et al., 2020). This subfamily has important roles in regulating cell proliferation and migration (Qi et al., 2020).

### 1.3 The transcriptional repressor CIC is negatively regulated by ERK

An additional relatively recently characterised phosphotarget of dpERK is the transcriptional repressor Capicua (CIC) (Jiménez et al., 2000). Originally identified as part of a P-element screen in *Drosophila*, CIC has since been shown to be highly conserved with homologs also present in vertebrates (Jiménez et al., 2000; Lee et al., 2002). Two main spliceoforms of CIC exist, a shorter isoform (CIC-S), and a longer isoform (CIC-L) (figure 2) (Lam et al., 2006; Forés et al., 2015; Simón-Carrasco et al., 2018; Lee, 2020). The protein isoforms have distinct N-termini, with the CIC-L isoform exclusively containing exon 1, and the CIC-S isoform exclusively containing exon 2 (Forés et al., 2015; Lee, 2020). Exon 1 contains an N1 domain, the function of which in CIC-L is currently not clear; all other exons are common to both isoforms (Simón-Carrasco et al., 2018). Through a high-mobility group box (HMG-box) DNA-binding domain, CIC is able to bind to conserved octameric motifs (T(G/C)AATG(A/G)A) in promoter and/or enhancer regions of its target genes (Jiménez et al., 2000; Kawamura-Saito et al., 2006; Ajuria et al., 2011; Weissmann et al., 2018). DNA-binding is thought to be additionally enhanced by the C1 domain (Forés et al., 2017).



**Figure 2: Schematic of the domain structure of human Capicua (CIC) protein isoforms.** Human CIC has two isoforms, CIC-short (CIC-S) and CIC-long (CIC-L) of lengths 1608 and 2517 amino acids, respectively. Each isoform has a differential first exon and therefore the extreme N-terminus (left on image) is unique to each isoform. HMG; high-mobility group box. aa; amino acid. EBS; ERK binding site.

Activated ERK interacts with the ERK binding site (EBS) of CIC, and may subsequently phosphorylate Ser<sup>1382</sup> and Ser<sup>1409</sup> of the human CIC homolog (Astigarraga et al., 2007; Futran et al., 2015; Dissanayake et al., 2011). As well as a direct interaction, dpERK may also indirectly mediate CIC phosphorylation through activation p90 RSK which can in turn phosphorylate human CIC at Ser<sup>173</sup> (Dissanayake et al., 2011). Phosphorylation of CIC dissociates CIC from its target sequence, and this relieves repression of CIC target genes (Astigarraga et al., 2007; Dissanayake et al., 2011; Futran et al., 2015). The capability of dpERK to dissociate CIC from its targets is elegantly demonstrated by ChIP-seq using an antibody against CIC, whereby a greater number of ChIP peaks were identifiable after treating human cell lines with a MEK inhibitor (Weissmann et al., 2018; Barrett et al., 2008). Notably however, dpERK may not be the sole regulator of CIC activity; amongst other ERK-independent mechanisms, tyrosine-phosphorylation of CIC by the non-receptor tyrosine kinase c-Src has also recently been demonstrated to inhibit the function of CIC in mammals (Papagianni et al., 2018; Yang et al., 2016; Bunda et al., 2020).

## 1.4 Phosphorylation may cause relocalisation or degradation of CIC

The fate of CIC following phosphorylation remains to be fully established. In *Drosophila*, CIC was initially shown to act as a transcriptional repressor downstream of Torso and Epidermal growth factor (EGF) RTK signalling and facilitate anteroposterior and dorsoventral embryonic axis patterning (Jiménez et al., 2000; Goff, Nilson and Morisato, 2001; Ajuria et al., 2011). Downstream of EGF RTK signalling, phosphorylation of CIC causes subcellular relocalisation of CIC from the nucleus to the cytoplasm, thus inhibiting its activity (Astigarraga et al., 2007). In contrast, phosphorylation of CIC downstream of Torso signalling in *Drosophila* results in CIC being targeted for degradation (Astigarraga et al., 2007). It is possible that both modes for CIC down-regulation also exist in mammals. Indeed, it has been proposed that phosphorylation of CIC in the C-terminus by dpERK impedes nuclear import of CIC by preventing an interaction with importin- $\alpha$ 4 (Dissanayake et al., 2011). Whilst p90 RSK phosphorylation of CIC adjacent to the HMG-box may create a docking site for 14-3-3 proteins and consequently limit the ability of CIC to bind DNA (Dissanayake et al., 2011). In *Xenopus*, it's been demonstrated that co-expression of fibroblast growth factor (FGF) ligands may target exogenous GFP-tagged CIC for degradation within certain regions of the embryo (King, 2018).

## 1.5 Co-repressors facilitate repression of CIC target genes

How CIC elicits gene repression is also not yet fully clear and appears to be context dependent (Forés et al., 2015). In the *Drosophila* early embryo, repression of genes including *tailless* and *huckebein* by CIC, necessary for successful axis patterning, relies on the presence of the co-repressor Groucho (Jiménez et al., 2000; Forés et al., 2015). Whereas, during *Drosophila* wing development, the repression of *argos* by CIC for example, is entirely independent of Groucho (Forés et al., 2015). In *Drosophila* and mammalian cells, an interaction between CIC and the co-repressor ATAXIN-1 has been found to modulate CIC repressive activity; this interaction is implicated in pathogenesis of the neurodegenerative disease spinocerebellar ataxia type I (SCA1) (Lam et al., 2006). Perhaps some of the most well established targets of CIC in mammals are members of the PEA3 subfamily of Ets transcription factors, *ETV1*, *ETV4* and *ETV5* (Dissanayake et al., 2011; Weissmann et al., 2018). Indeed, ERK putatively regulates the PEA3 subfamily of Ets transcription factors through negative regulation of CIC (Qi et al., 2020). An interaction between CIC and the SIN3 deacetylation complex, which encompasses the histone deacetylase enzymes HDAC1 and HDAC2, may partially account for repression of *ETV4* and *ETV5* by CIC (Weissmann et al., 2018). Loss of histone acetylation, an epigenetic mark often associated with active euchromatin, can

create a less permissive environment for transcription (Bannister and Kouzarides, 2011). Thus, CIC may elicit gene repression through recruitment of co-repressors and alteration of the epigenetic landscape.

## 1.6 CIC functions as a tumour suppressor

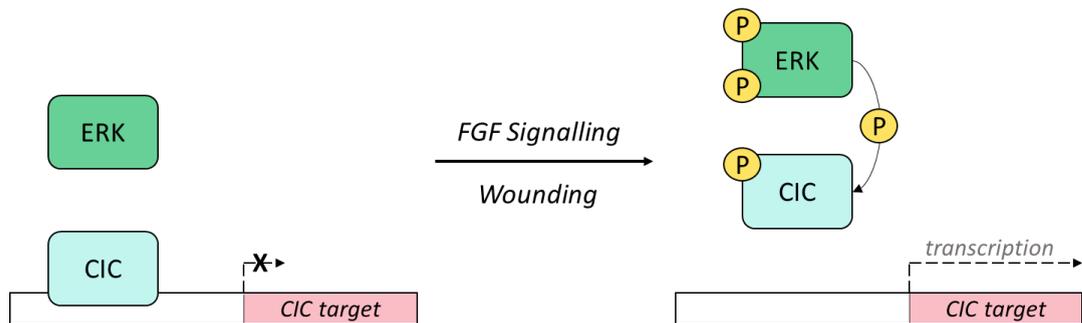
In *Drosophila*, alongside its role in anteroposterior and dorsoventral axis patterning during embryonic development, CIC also controls the growth of organs, such as the imaginal discs and midgut (Tseng et al., 2007; Jin et al., 2015). In mammals, one of the first implicated roles for CIC was in central nervous system (CNS) development (Lee et al., 2002). Since, CIC activity has been linked to lung development, T cell differentiation, and abdominal wall closure, amongst other diverse roles (Lee, 2020). Importantly, CIC functions as a tumour suppressor, with mutation and dysregulation of CIC documented in CNS tumours, including oligodendrogliomas and glioblastomas, alongside soft tissue sarcomas, such as E-wing like sarcomas (Bettegowda et al., 2011; Yip et al., 2012; Bunda et al., 2019; Kawamura-Saito et al., 2006). Due to the involvement of aberrant CIC function in disease progression, understanding both its regulation and identifying its target genes is imperative.

## 1.7 Aims and hypotheses

*Xenopus* species provide an attractive model organism for studying *in vivo* regulation of gene expression during development due to their macroscopic egg size, *ex utero* development, and genetic amenability (Blum and Ott, 2018). Two CIC isoforms also exist in *Xenopus* species, CIC-L (2510 amino acids) and CIC-S (1608 amino acids) (King, 2018). Whilst the CIC-L isoform represents the dominant maternally expressed transcript, both isoforms are zygotically expressed, including around the blastopore during early gastrulation, and within the neural plate during neurulation (King, 2018). Importantly, key protein domains (HMG-box, C1, EBS, nuclear localisation signal, ATXN-1 and 14-3-3 binding; see figure 2) are highly conserved in *Xenopus* CIC (King, 2018). Therefore, *Xenopus* provides a suitable model organism to investigate the regulation and targets of CIC. The majority of ERK activation during early *Xenopus* development has been demonstrated to be downstream of FGF signalling (Christen and Slack, 1999; Branney et al., 2009). Furthermore, in response to embryonic wounding in *Xenopus*, ERK is activated and recent work within the Isaacs lab has also demonstrated that Myc-tagged CIC protein is degraded (*manuscript in preparation*, King et al. (2022)). However, the contribution of FGF signalling and embryonic wounding to the regulation of CIC is relatively understudied. In this project, it was hypothesised that ERK relieves CIC-mediated repression downstream of FGF signalling and post-wounding during *Xenopus* development (figure 3).

The overall aims of this project are:

- 1) Perform gene-level differential expression analysis of RNA-seq data from FGF4 overexpressing and CIC knockdown *X. tropicalis* embryos (previously analysed at the transcript-level, (King, 2018)) to identify putative CIC/FGF target genes.
- 2) Investigate the regulation of putative CIC target genes in the context of FGF signalling and embryonic wound healing in *Xenopus*.



**Figure 3: Schematic representation of overall project hypothesis.** Downstream of FGF signalling and following embryonic wounding ERK becomes di-phosphorylated (dpERK) and activated. In turn, dpERK mediates phosphorylation of CIC, either directly or indirectly. This results in dissociation of CIC from promoter and/or enhancer regions of its target genes and relieves transcriptional repression.

# Chapter 2: Materials and Methods

## 2.1 Embryological Methods

### 2.1.1 Ethics statement

All animal work undertaken was governed by a UK Home Office licence issued to Professor Betsy Pownall and approved by the University of York Animal Welfare Ethical Review Body (AWERB).

### 2.1.2 *Xenopus* embryo culture

Dr. Harry Isaacs obtained *Xenopus tropicalis* and *Xenopus laevis* embryos by artificial *in vitro* fertilisation. *X. tropicalis* and *X. laevis* embryos were cultured in Modified Ringer's Solution (MRS; 0.1 M NaCl, 1.8 mM KCl, 2.0 mM CaCl<sub>2</sub>, 1.0 mM MgCl<sub>2</sub>, 5 mM HEPES, pH 7.6) and Normal Amphibian Medium (NAM) (NAM salts; 110 mM NaCl, 2 mM KCl, 1 mM Ca(NO<sub>3</sub>)<sub>2</sub>, 1 mM MgSO<sub>4</sub>, 0.1 mM EDTA dissolved in 1 mM NaHCO<sub>3</sub> and 5 mM HEPES pH 7.4 solution), respectively. At 10 minutes post-fertilisation, embryos were flooded with 1/9 strength MRS (MRS/9) or 1/3 strength NAM salts (NAM/3). Subsequently, *X. tropicalis* and *X. laevis* embryos were dejellied in 3% w/v L-cysteine/dH<sub>2</sub>O (Sigma), pH 7.8 and 2.5% w/v L-cysteine hydrochloride monohydrate/NAM (Sigma), pH 7.88, before culture was continued in 1/20 strength MRS (MRS/20) and 1/10 strength NAM salts (NAM/10), respectively. Embryonic stage was assigned according to Nieukoop and Faber (1967).

### 2.1.3 Wounding

*Xenopus* embryos were wounded centrally on the flank using a tungsten needle. Embryos were allowed to heal for a designated length of time, before being fixed in MEMFA (0.1 M MOPS, 2 mM EDTA, 1 mM MgSO<sub>4</sub>, 3.7% formaldehyde) for 1 hour at room temperature (RT) and then stored in 100% methanol at -20°C. Unless specified, the vitelline membrane was not removed from embryos to prevent accidental wounding, and subsequent *in situ* hybridisations were carried out as per the modified *in situ* hybridisation protocol (see materials and methods; 2.2.10).

### 2.1.4 Treatment with chemical inhibitors

*X. tropicalis* sibling embryos (stage 14-15) were transferred into an agarose-coated 12-well plate containing wells with 1 ml of either MRS/20, MRS/20 containing 0.1% v/v dimethyl sulfoxide (DMSO) or 25 µM PD0325901 (Cell Guidance Systems).

PD0325901 was diluted from 61.4 mM stock in MRS/20 containing 0.1% v/v DMSO. PD0325901 is an inhibitor of MAPK kinases MEK1 and MEK2 (Barrett et al., 2008). Embryos were incubated at 21°C in the 12-well plate for a minimum of two hours, in the dark, and until approximately stage 21. Following, embryos were transferred from the wells into pre-equilibrated 55 mm 1.5% agarose-coated plates containing 9 ml of the assigned culture media. Embryos were wounded before being transferred back into their original wells and left to heal for 30 minutes at RT in the dark. At 30 minutes post-wounding, embryos were fixed in MEMFA for 1 hour at RT and then stored in 100% methanol at -20°C.

At approximately stage 8 and stage 14, *X. tropicalis* sibling embryos were transferred into agarose-coated 12-well plates containing wells with 1 ml of either MRS/20, MRS/20 containing 0.2% v/v DMSO, 25 µM of PD0325901 (Cell Guidance Systems) or 200 µM SU5402 (Sigma). Stage 14 embryos to be used for *fos in situ* hybridisation had the vitelline membrane removed 1 hour prior to transfer. PD0325901 and SU5402 were prepared as 61.4 mM and 100 mM stocks in DMSO and diluted to their final concentrations in MRS/20 containing 0.2% v/v DMSO. SU5402 is an FGFR inhibitor, but has also been shown to inhibit other RTKs (Mohammadi et al., 1997). Embryos were incubated at 21°C in the 12-well plate for a minimum of two hours, in the dark, and until stage 10-10.5 and stage 18-19 for developmentally younger and older starting embryos, respectively. Embryos were then fixed in MEMFA for 1 hour at RT and stored in 100% methanol at -20°C.

Protocol adapted from Abaffy et al. (2019). The vitelline membrane was removed from *X. laevis* embryos using forceps at stage 22. At stage 24, embryos were transferred into an agarose-coated 12-well plate containing wells with 1.5 ml of either NAM/10 or 2 mM 1-(2-Trifluoromethylphenyl) imidazole (TRIM, Cambridge Bioscience Ltd) and incubated at RT for 1 hour in the dark. Throughout the course of the experiment, 1 ml of 2 mM TRIM was taken from the wells and replaced with 1 ml fresh every hour. TRIM was prepared as a 60 mM stock in DMSO and diluted to its final concentration in NAM/10. Following, embryos were transferred from wells into pre-equilibrated 55 mm 1.5% agarose-coated plates containing 9 ml of the assigned culture media. Embryos were wounded before being transferred back into their original wells and allowed to heal for 50, 70 or 90 minutes at RT in the dark. After the required healing time, embryos were fixed in MEMFA for 1 hour at RT and then stored in 100% methanol at -20°C.

### 2.1.5 Microinjection

Dr. Harry Isaacs performed *X. laevis* embryo injections in 5% w/v Ficoll / (NAM/3) using a Drummond microinjector with pulled needles (Drummond). Embryos were either injected at the one-cell stage with 4 ng of *fos* mRNA (see materials and methods; 2.2.4, 2.2.6), or unilaterally at the 2-cell stage with 2 ng of *fos* mRNA. After the embryos had healed, culture was continued in NAM/10. At the desired stage, embryos were flash frozen on dry ice and stored at -70°C for Western blotting, or fixed in MEMFA for 1 hour at RT and stored in 100% methanol at -20°C for immunostaining.

### 2.1.6 Imaging

SPOT Advanced software was used alongside a Leica MZ FLIII microscope with a SPOT 14.2 Colour Mosaic Camera (Diagnostic Instruments Inc.) to capture images of the embryos in a range of focal planes. In several instances, the stacking feature of Adobe photoshop CS3 was used to align and blend the different focal planes into a single image.

## 2.2 Molecular Biology Methods

### 2.2.1 Agarose gel electrophoresis

DNA and RNA samples were resolved by electrophoresis on 1% and 1.5 - 2% w/v Tris-Acetate-EDTA (TAE; 40 mM Tris pH 7.6, 20 mM acetic acid, 1 mM EDTA) agarose gels (containing 0.00001 % v/v 10 mg/ml ethidium bromide) respectively, and at 150 V and 180 V, respectively. Gels were buffered in TAE, and a 1 kb Plus DNA Ladder (New England Biolabs) used to evaluate product size.

### 2.2.2 RNA extraction

Flash frozen *X. tropicalis* and *X. laevis* embryos stored at -70°C were brought to RT and homogenised in 1 ml TRI-Reagent (Sigma-Aldrich). Samples were left on ice for 1 minute and then centrifuged at 13,000 rpm for 10 minutes at 4°C. The supernatant was collected and left at RT for 3 minutes, before the addition of 200 µl of chloroform (Sigma-Aldrich) followed by 15 seconds of manual shaking. Samples were left at RT again for 3 minutes before being centrifuged for 15 minutes at 13,000 rpm and 4°C. The upper aqueous phase containing RNA was isolated, a further 200 µl of chloroform added, and the samples centrifuged at 13,000 rpm and 4°C for 5 minutes. The upper aqueous phase was collected again, 500 µl of isopropanol added, and samples briefly vortexed and then stored at -20°C for 30 minutes. Following, RNA was pelleted by

centrifugation for 15 minutes at 13,000 rpm and 4°C and the supernatant discarded. The pellet was washed by addition of 200 µl ice cold 70% EtOH, vortexing and then re-centrifugation at 13,000 rpm and 4°C for 10 minutes. The supernatant was discarded and the RNA vacuum desiccated, and resuspended in 100 µl nuclease-free water and 100 µl lithium chloride (LiCl) precipitating solution (4M LiCl, 20 mM Tris pH 7.4 and 10 mM EDTA) and left at -70°C overnight. Samples were subsequently centrifuged at 13,000 rpm for 20 minutes at 4°C, and the supernatant discarded. The pellet was washed again in 200 µl ice cold 70% ethanol by vortexing and re-centrifugation at 13,000 rpm and 4°C for 5 minutes. The RNA was vacuum desiccated and resuspended in 15-20 µl nuclease-free water. RNA quality was assessed by resolving the sample on a 1.5 % w/v agarose gel, and RNA concentration and purity determined using Nanodrop™ 8000 Spectrophotometer (Thermo Scientific).

### 2.2.3 First strand cDNA synthesis

For PCR, cDNA was synthesised from 1 µg of RNA, 1 µl of 10 mM dNTP mix (Thermo Scientific), 1 µl of 0.2 µg/µl random hexamers (Thermo Scientific) and nuclease-free water upto a total volume of 13 µl. Samples were heated at 65°C for 5 minutes, and then 5X SuperScript™ IV Buffer (Invitrogen), 1 µl SuperScript™ IV reverse transcriptase (Invitrogen), and 1µl of 0.1M DTT added (total volume of 20 µl). Following, samples were heated for 10 minutes sequentially at 23°C, 55°C, and 80°C.

### 2.2.4 PCR-based cloning of full-length *fos*, and *jun*

Primers (table 1) were used to amplify full-length *fos* (accession: NM\_001016200) and full-length *jun* (accession: XM\_031900902.1) from total RNA extracted from *X. tropicalis* ovaries, and stage 14 *X. tropicalis* embryos, respectively. For *jun* amplification, the PCR reaction was set up with 2 µl cDNA, 38.5 µl nuclease-free water, 1 µl of 10 mM dNTP mix (Thermo Scientific), 1 µl of 10 µM forward primer, 1 µl of 10 µM reverse primer and 5 µl 10X reaction buffer with MgSO<sub>4</sub> (Promega). For *fos* amplification, 1 µl of nuclease-free water was substituted with 1 µl DMSO to reduce primer-dimer formation. A manual hot start was performed with 1.5 µl of *Pfu* DNA Polymerase (Promega) added at the end of the 2 minute initial denaturation (total volume 50 µl). Protocol for PCR was: 2 mins at 95°C; 30 cycles of 95°C for 30s, 60°C for 30s, 72°C for 2 mins, and a final 10 min at 72°C. Amplicons were cleaned using Monarch® PCR & DNA Cleanup Kit (5 µg) as per manufacturer's protocol (New England BioLabs). Sticky ends were created by combining 15 µl PCR amplicon with 10X buffer B (Roche), 1.5 µl XhoI (Promega), 1.5 µl BamHI (Promega) and nuclease-free water up to a total volume of 100 µl and incubating at 37°C for 2 hours. Alongside, 62.5 ng of the vector pCS2+ was digested under the same conditions. PCR

amplicon digests were gel purified on a 0.8% w/v agarose gel using Monarch® DNA Gel Extraction Kit as per manufacturer's protocol (New England BioLabs). Digested amplicon was unidirectionally cloned into pCS2+ by combining 5 µl digest amplicon, 1 µl digested pCS2+, 1 µl T4 DNA ligase (Promega), 10X T4 ligase buffer (Promega) and nuclease-free water to a total volume of 10 µl at 18°C for approximately 48 hours. Ligations were transformed into *dam*/*dcm* competent *E. coli* (New England BioLabs), and plasmid DNA subsequently purified (see materials and methods; 2.2.5). Colonies successfully containing the plasmid with insert were identified using a PCR colony screen. Individual colonies were lysed by pipetting in 2 µl nuclease-free water, and combined with 2X *Taq* PCR Master Mix (Promega), 1 µl of 10 µM SP6 primer, 1 µl of 10 µM *fos/jun* reverse primer (table 1) and nuclease-free water to a total volume of 20 µl. Protocol for PCR was; 1.5 min at 98°C; 30 cycles of 98°C for 30s, 50°C for 30s, 72°C for 1 min 30s. Successful amplification of insert was confirmed by running PCR products on a 1% w/v agarose gel. Plasmids were sent for sequencing (Eurofins Genomics) using SP6 and T7 primers to confirm absence of mutations in the full-length *fos/jun* sequence (Appendix A, B). Sequencing data was analysed using SeqMan software from the Lasergene Genomics Suite (DNA Star).

**Table 1: Details of primers used to amplify *X. tropicalis* full-length *fos* (accession: NM\_001016200) and full-length *jun* (accession: XM\_031900902.1).** In the forward primer, italicised bases represent spacer sequence to aid restriction enzyme binding, underlined sequence highlights a BamHI restriction site, and capitalised sequence denotes a partial kozak consensus sequence to enhance translation. In the reverse primer, italicised bases represent spacer sequence to aid restriction enzyme binding and underlined sequence highlights a XhoI restriction site. T<sub>m</sub>; melting temperature.

mRNA target	Primer	Sequence (5'-3')	T <sub>m</sub> (°C)
<i>fos</i>	Forward	<i>aga gag gga tcc</i> ACC ATG tat cac gcc ttc tcc agc	67.7
	Reverse	<i>aga gag ctc gag</i> cag agc caa aag ggt ggg aga	68.6
<i>jun</i>	Forward	<i>aga gag gga tcc</i> ACC ATG act gca agg atg gaa cct	67.6
	Reverse	<i>aga gag ctc gag</i> gaa tgt ttg cat ctg ctg tgt	65.1

### 2.2.5 Plasmid transformation and purification

pGEM®-T Easy plasmid with a 701 bp *fos* insert (pGEM®-T Easy *fos*) and 445 bp *rasl11b* insert (pGEM®-T Easy *rasl11b*) were kindly provided by L. Cowell (Cowell, 2019). Plasmid DNA was transformed into *dam*/*dcm*<sup>-</sup> competent *E. coli* (New England BioLabs) by combining 5 ng of plasmid DNA with 25 µl competent *E. coli* on ice. After 30 minutes on ice, the *E. coli* were heat shocked in a 40°C water bath for 90 seconds, and then placed on ice for a further 2 minutes. Following, 1 ml lysogeny broth (LB; 10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) was added, and the mixture shook vigorously (250 rpm) for 1 hour, before 100 µl and 300 µl of the bacterial suspension were plated on to separate LB-agar petri dishes containing 100 µg/ml ampicillin. Petri dishes were incubated overnight at 37°C, and unique colonies selected for liquid overnight culture in 3 ml LB containing 100 µg/ml ampicillin. During the logarithmic growth phase, plasmid DNA was isolated from the cultures using QIAprep® Spin Miniprep Kit as per manufacturer's protocol (QIAGEN). To confirm retention of pGEM-T Easy plasmid inserts, 5 µl of DNA was digested in a total volume of 20 µl containing 10X Buffer H (Roche), 1 µl EcoR1 (Promega) and dH<sub>2</sub>O for 90 minutes in a 37°C incubator. Digested samples were resolved on a 1% w/v agarose gel alongside undigested plasmid.

### 2.2.6 Messenger RNA synthesis

Plasmid DNA (pCS2+ *fos* and pCS2+ *jun*) was linearised to create a template for mRNA synthesis by digesting 2.5 µg of plasmid DNA in a total volume of 100 µl containing 10X buffer D (Roche), 2 µl NotI (Promega) and nuclease-free water for 2 hours at 37°C. Digested samples were resolved on a 1% w/v agarose gel alongside undigested plasmid. mRNA synthesis was carried out using the mMACHINE® Kit (Ambion) as per manufacturer's instructions for capped transcription reaction assembly. The optional TURBO DNase step was completed, and phenol:chloroform extraction and isopropanol precipitation used for recovery of the RNA. Synthesised mRNA was resolved on a 1.5% w/v agarose gel, and concentration determined using Nanodrop™ 8000 Spectrophotometer (Thermo Scientific).

### 2.2.7 Cloning of *in situ* hybridisation probe template

Plasmid DNA (pGEM-T Easy *fos*, pGEM-T Easy *rasl11b*, pCS2+ *jun*) was linearised to create a template for RNA probe synthesis by digesting 5 µl of plasmid DNA in a total volume of 100 µl containing 10X buffer H (Roche), 2 µl NcoI (Thermo Scientific), and dH<sub>2</sub>O for 2 hours at 37°C (table 2). Digested samples were resolved on a 1% w/v agarose gel. Linearised samples were brought up to 400 µl with nuclease-free water

and then 40  $\mu$ l 3M sodium acetate ( $C_2H_3NaO_2$ ) and 400  $\mu$ l water-saturated phenol-chloroform added, before centrifugation at 13,000 rpm for 5 minutes. The aqueous phase was isolated, 1 ml of 100% ethanol added, and the DNA left to precipitate for 1 hour (minimum) at  $-20^\circ C$ . DNA was pelleted by centrifugation at  $4^\circ C$  at 13,000 rpm for 15 minutes, and then washed through addition of 100  $\mu$ l ice cold 70% ethanol, vortexing, and re-centrifuging. The DNA was vacuum desiccated, resuspended in 20  $\mu$ l nuclease-free water and stored at  $-20^\circ C$ . Retention of linearised plasmid was confirmed by resolving samples on a 1% w/v agarose gel.

**Table 2: Enzymes and buffers used for template linearisation, and polymerases used for *in situ* hybridisation probe synthesis.** Templates for antisense transcription were created by linearising plasmid DNA using the stated enzymes and buffers. gBlock DNA is purchased linear. The polymerases used to synthesise the antisense probe following purification of the linearised template are indicated. N/A; not applicable.

DNA Template	Linearisation Enzyme	Linearisation Buffer	Polymerase
pGEM-T Easy <i>fos</i>	NcoI	H	SP6
pGEM-T Easy <i>rasl11b</i>	NcoI	H	SP6
pCS2+ <i>jun</i>	NcoI	H	T3
gBlock <i>atf3</i>	N/A	N/A	T7

### 2.2.8 Synthesis of *in situ* hybridisation antisense RNA probe

Linearised pGEM-T Easy *fos* and pGEM T-Easy *rasl11b* (1  $\mu$ g) was combined with 5X transcription buffer (Promega), 10X DIG rNTP mix (Roche), 1  $\mu$ l SP6 polymerase (Ambion), 2  $\mu$ l DTT (Promega) and  $dH_2O$  to a total volume of 20  $\mu$ l. Linearised pCS2+ *jun* (1  $\mu$ g) was combined with 5X transcription buffer (ThermoFisher Scientific), 10X DIG rNTP mix (Roche), 1  $\mu$ l T3 polymerase (ThermoFisher Scientific), and  $dH_2O$  to a total volume of 20  $\mu$ l. gBlock DNA (50 ng) as purchased from Integrated DNA Technology was used to synthesise the *atf3* antisense RNA probe. The reaction was set up as per synthesis of the *jun* antisense probe but with T7 polymerase (ThermoFisher Scientific) substituting for T3. Transcription reactions were allowed to proceed overnight at  $37^\circ C$ . The synthesis of *fos* (701 bp), *rasl11b* (445 bp), *atf3* (400 bp) and *jun* (predicted 671 bp total; 283 bp pCS2+ vector with 388 bp *jun*) RNA probes was confirmed by resolving the transcription reaction on a 1.5% w/v agarose gel. Antisense RNA probes were precipitated overnight at  $-20^\circ C$  through the addition of 50

$\mu\text{l}$  nuclease-free water, 50  $\mu\text{l}$  5M ammonium acetate ( $\text{NH}_4\text{CH}_3\text{CO}_2$ ) and 300  $\mu\text{l}$  100% ethanol. Subsequently, the RNA was pelleted by centrifugation at 4°C at 13,000 rpm for 15 minutes and then washed through addition of 100  $\mu\text{l}$  ice cold 70% ethanol, vortexing, and re-centrifuging. The RNA was vacuum desiccated, and resuspended in 50  $\mu\text{l}$  nuclease-free water through briefly vortexing and heating for 2 minutes at 80°C. Retention of the RNA probe following ethanol precipitation was confirmed by resolving the sample on a 1.5% w/v agarose gel. RNA probes were stored at -70°C ready for *in situ* hybridisation use. Antisense *xbra* RNA probe for *in situ* hybridisation was taken from a communal stock within the Isaacs lab (Smith et al., 1991).

### 2.2.9 Whole-mount RNA *in situ* hybridisation

Embryos (without vitelline) fixed in MEMFA and stored at -20°C, were rehydrated through sequential 10 minute washes in 75% MeOH/PBSAT, 50% MeOH/PBSAT, and phosphate buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM  $\text{Na}_2\text{HPO}_4$ , 1.8 mM  $\text{KH}_2\text{PO}_4$ ) containing 0.1% Tween (PBSAT). Embryos were treated with 10  $\mu\text{g}/\text{ml}$  Proteinase K (Roche, table 3), and washed twice for 5 mins in 5 ml of 0.1 M triethanolamine pH 7.8. At the end of the second wash, the addition of 12.5  $\mu\text{l}$  acetic anhydride ( $(\text{CH}_3\text{CO})_2\text{O}$ ) followed by 5 minutes of gentle agitation, was repeated twice. Embryos were washed twice for 5 minutes in PBSAT, refixed for 20 minutes in 3.65-3.8% v/v formaldehyde/PBS and washed five more times for 5 mins in PBSAT. Embryos were transferred into screw-top eppendorfs and 1 ml hybridisation buffer (50% formamide, 1 mg/ml total yeast RNA, 5x SSC pH7, 100  $\mu\text{g}/\text{ml}$  heparin, 1X Denhart's, 0.1% Tween, 0.1% CHAPS, 10 mM EDTA) heated to 60°C added. Embryos were left at 60°C for 10 minutes, before hybridisation buffer was replaced with fresh and embryos were horizontally rocked at 60°C for 2 hours (pre-hybridisation). Hybridisation buffer was replaced with fresh containing DIG-labelled RNA antisense probe (0.5 - 2  $\mu\text{l}$ ), and embryos rocked overnight at 60°C. Following, whilst maintaining temperature at 60°C, embryos were washed twice in hybridisation buffer for 10 mins, three times in 2X SSC containing 0.1% tween for 20 minutes, and 3 times in 0.2X SSC containing 0.1% tween for 30 minutes. At RT, embryos were washed twice in maleic acid buffer (MAB; 100 mM maleic acid, 150 mM NaCl, pH 7.8) for 15 minutes, and then pre-incubated in 1 ml MAB containing 2% Blocking Reagent (Roche) and 20% heat treated lamb serum (Thermo Fisher Scientific) for 2 hours whilst rolling. The solution was replaced with fresh containing 1/2000 dilution of anti-digoxigenin-AP fab fragments (Roche), and the embryos rolled overnight at 4°C. Subsequently, embryos were washed three times for 5 minutes, and then 3 times for 1 hour in MAB, and once in Alkaline Phosphatase (AP) Buffer (100 mM Tris pH 9.5, 50 mM  $\text{MgCl}_2$ , 100 mM NaCl, 0.1% Tween) for 10 mins. Embryos were left in 1 ml dilution of BM-Purple

(Roche) in AP Buffer to allow staining to develop, before being washed twice for 15 minutes in PBSAT, and then bleached in 5% v/v H<sub>2</sub>O<sub>2</sub>/PBSAT. Embryos were stored at RT in 3.65-3.8% v/v formaldehyde/PBS.

**Table 3: Length of time (minutes) *Xenopus* embryos were subject to 10 µg/ml Proteinase K treatment during *in situ* hybridisation protocol, with respect to developmental stage.** Developmental stage assigned according to (Nieuwkoop P D And Faber, 1967).

Developmental Stage	Proteinase K Treatment (min)
10-10.5	2.5
11	3.0
12-17	5.5
18	7.0
19	7.5

### 2.2.10 Modified *in situ* hybridisation

The modified *in situ* hybridisation protocol begins as per the steps of the immunostaining protocol upto and including the PBS washes after treatment with 5% H<sub>2</sub>O<sub>2</sub>/PBS (see materials and methods; 2.2.11). The protocol then proceeds as per the *in situ* hybridisation protocol starting from, and including the 0.1M triethanolamine washes (see materials and methods; 2.2.9). The 20 minute fix in 3.65-3.8% v/v formaldehyde/PBS and immediately following PBS washes were only included for the more delicate gastrula stage 10-10.5 embryos.

### 2.2.11 Whole-mount immunostaining

Embryos (with vitelline) fixed in MEMFA and stored at -20°C, were rehydrated through sequential 10 minute washes in 75% MeOH/PBS, 50% MeOH/PBS, and PBS. Embryos were rocked for 40 minutes at RT in 10 ml potassium dichromate (K<sub>2</sub>Cr<sub>2</sub>O<sub>7</sub>) solution containing 5 % acetic acid (CH<sub>3</sub>COOH), and washed three times for 5 minutes, and then 3 times for 20 minutes, in PBS. Following, embryos were rocked at RT in 5% H<sub>2</sub>O<sub>2</sub>/PBS for 45 minutes under a lamp, and then washed 3 times for 10 minutes in PBS. Embryos were washed twice for 15 minutes in BBT (PBS, 1% BSA, 0.1% Triton X-100), and blocked by a proceeding 1 hour wash in 1 ml BBT containing 5% horse serum (Vector Laboratories). The solution was then replaced with fresh containing a

1:10,000 dilution of monoclonal anti-MAP kinase activated (di-phosphorylated ERK-1&2) antibody produced in mouse (Sigma, M9692) or 1:10,000 dilution of mouse monoclonal anti c-Fos (Merck, 2G9). Embryos were left rolling in primary antibody overnight at 4°C. Afterwards, embryos were washed four times for 1 hour in BBT, once for 1 hour in BBT + 5% horse serum, and once in 1 ml fresh BBT + 5% horse serum containing a 1:1000 dilution of alkaline phosphatase anti-mouse IgG (H+L) made in horse (Vector, AP-2000). Embryos were rolled overnight at 4°C before being washed once for 1 hour in BBT, and then four times for 1 hour in PBSAT. Embryos were left in a 1 ml dilution of BM-Purple in AP buffer to allow staining to develop, and then washed twice for 15 minutes in PBS before being stored at RT in 3.65-3.8% v/v formaldehyde/PBS.

## 2.2.12 Western blotting

Embryos flash frozen on dry ice and stored at -70°C were homogenised by addition of 50 µl phosphosafe extraction reagent (Novagen) and gentle pipetting. Samples were centrifuged for 1 minute at 13,000 rpm, and 2X sample buffer (120mM Tris/Cl pH6.8, 20% glycerol, 4% SDS, 0.04% bromophenol blue, 10% β-mercaptoethanol) added to the supernatant. Following, samples were heated at 95°C for 5 minutes before being centrifuged again at 13,000 rpm for 1 minute. Samples representative of the protein content from 1 embryo (20 µl) were loaded onto a 10% SDS-Page gel and ran at 180V alongside PageRuler™ prestained protein ladder (Thermo Scientific) in running buffer (3 g/L Tris, 14.4 g/L glycine, 0.1% SDS). Once resolved, samples were transferred onto Immobilon-P PVDF membrane (Merck) by electroblotting in transfer buffer (3g/L Tris, 14.4 g/L glycine, 10% methanol) at 85V for 2 hours. The membrane was then washed four times for 20 minutes in PBSAT, and blocked in 5% w/v milk powder/PBS (blocking solution) for 1 hour at RT. Following, the protein ladder was used to direct cutting of the membrane at 55 kDa and the blocking solution was replaced with fresh containing primary antibody. Molecular weights above 55 kDa were probed with 1:500 dilution of mouse monoclonal anti c-Fos (Merck, 2G9), and molecular weights below with 1:40,000 dilution of anti-MAP Kinase (ERK-1, ERK-2) produced in rabbit (Sigma, M5670). Membranes were left in primary antibody overnight at 4°C, and subsequently washed four times for 20 minutes in PBSAT before being left in blocking solution for 30 minutes at RT. Blocking solution was then replaced with fresh containing the appropriate secondary antibody and membranes left at RT for 1 hour. A 1:20,000 dilution was used for goat anti-mouse IgG (H+L), HRP (Invitrogen, 31430) and a 1:2,000 dilution for anti-rabbit HRP-linked (Cell Signalling Technologies, 7074). Membranes were washed four times for 20 minutes in PBSAT and BM

chemiluminescence blotting substrate (Roche) and UltraCruz autoradiography film (Santa Cruz Biotechnology) used to detect protein.

## 2.3 Bioinformatics

### 2.3.1 Gene-level differential expression analysis of RNA-Seq data

RNA samples were collected in triplicate by Michael King (King, 2018); *X. tropicalis* embryos were injected with either TALENs targeting the HMG-box DNA-binding domain of CIC, a plasmid CSKA-eFGF designed to overexpress FGF4 (Isaacs, Pownall and Slack, 1994), or water. Initial RNA quality control, library preparation and RNA-sequencing was performed by staff at the University of York Technology Facility. A 2100 bioanalyzer instrument (Agilent) was used for RNA quality control, and 9 Illumina next generation sequencing libraries created (water, FGF4, and CIC TALEN injected samples in triplicate). Sequencing was conducted using a single lane of the Illumina HiSeq 2000 platform with a coverage of  $\geq 40$  million reads per sample. Initial data analysis, including aligning raw reads to the *X. tropicalis* reference genome (version 9.1) was undertaken by John Davey using R package Salmon (Patro et al., 2017). This project utilised the Salmon output to perform gene-level differential expression analysis using the R package Sleuth (Pimentel and McGee, 2021) and the counts aggregation method. The reduced model was defined as '~batch', and the full model '~batch + treatment' to mitigate variation accounted for by embryo batch (Appendix C). Both the FGF overexpression and CIC knockdown gene-level RNA-seq data was filtered to remove genes without an average TPM value of  $> 1$  in either the control or treatment data (additional file S1). Genes which met the criteria of  $q$ -value  $< 0.1$  and effect size  $> 1.5$  were defined as significantly upregulated and genes which met the criteria  $q$ -value  $< 0.1$  and effect size  $< 1/1.5$  as significantly downregulated (additional file S2, S3).

### 2.3.2 Intersection of gene lists

Venn diagrams were produced using GeneVenn webserver (<http://genevenn.sourceforge.net/>). Statistical analysis of the overlap between two gene lists was determined using R package GeneOverlap which implements a Fisher's exact test (Shen and at Mount Sinai, 2020). Whole-genome size for comparison between FGF4 overexpression and CIC knockdown gene lists was set as 23,635, and genome size for comparison between CIC knockdown and wound-induced lists set as 8,713. Criteria used to determine gene lists, and complete lists of gene identities can be found in additional file S4, S9 and table 7.

### 2.3.2 Gene ontology analysis

Gene ontology enrichment analysis was performed using the GOrilla webserver (Eden et al., 2009, 2007); most recent update 06/03/21). Enrichment of biological processes was assessed using the human database as a reference. An unranked list of genes which were upregulated following CIC knockdown and met the criteria  $p$ -value  $< 0.01$ , effect size  $> 1.5$  and TMP  $> 1$  in either control or TALEN-treated embryos was used as the target list. The target gene list was isolated from the background list. Results were filtered to include  $p$ -values  $< 1e-4$  (additional file S5).

### 2.3.3 Conservation analysis of *fos* genomic sequence

UCSC Table Browser tool (Kent et al., 2002; Karolchik et al., 2004) was used to export the transcribed sequence of *fos* alongside 5 kb upstream and 1 kb downstream for *Homo sapiens* (accession: NM\_005252.4), *Mus musculus* (accession: NM\_010234.3), *Gallus gallus* (accession: NM\_205508.1) and *X. tropicalis* (accession: NM\_001016200.2). The table browser was also utilised to create sequence annotation files (UTR and intron location) for each of the orthologues (additional file S6). FASTA files were submitted to the Vista genome browser (Frazer et al., 2004; Mayor et al., 2000) and mVista and the LAGAN alignment programme (Brudno et al., 2003) used to align sequences with the *X. tropicalis* sequence set as the base. Percentage of conservation identity and minimum conservation window parameters were set as 70% and 100 bp, respectively, for all species.

### 2.3.4 Motif enrichment analysis

MEME Suite 5.4.1 (Bailey et al., 2015) was used to identify candidate CIC binding sites with the consensus TSAATGRA. The UCSC table browser tool (Kent et al., 2002; Karolchik et al., 2004) was used to generate FASTA files of sequence for input into the MEME suite (additional file S6, S7). Motif scanning using FIMO analysis (Grant, Bailey and Noble, 2011) was performed to identify locations of sequence matching the consensus, with matches filtered to have a  $p$ -value  $< 1e-4$ . Motif enrichment using SEA (Timothy L. Bailey and Charles E. Grant, 2021) was performed to assess for enrichment of the CIC consensus in *fos* intron 1 sequence from *H. sapien*, *M. musculus*, *G. gallus* and *X. tropicalis* (additional file S8). The E-value threshold for reporting enriched motifs was set as  $E \leq 10$ , and shuffled input sequences were used as a control.

# Chapter 3: FGF-dependent Regulation of the Transcriptional Repressor CIC

## 3.1 Introduction

### 3.1.1 FGF signalling has diverse roles during embryogenesis

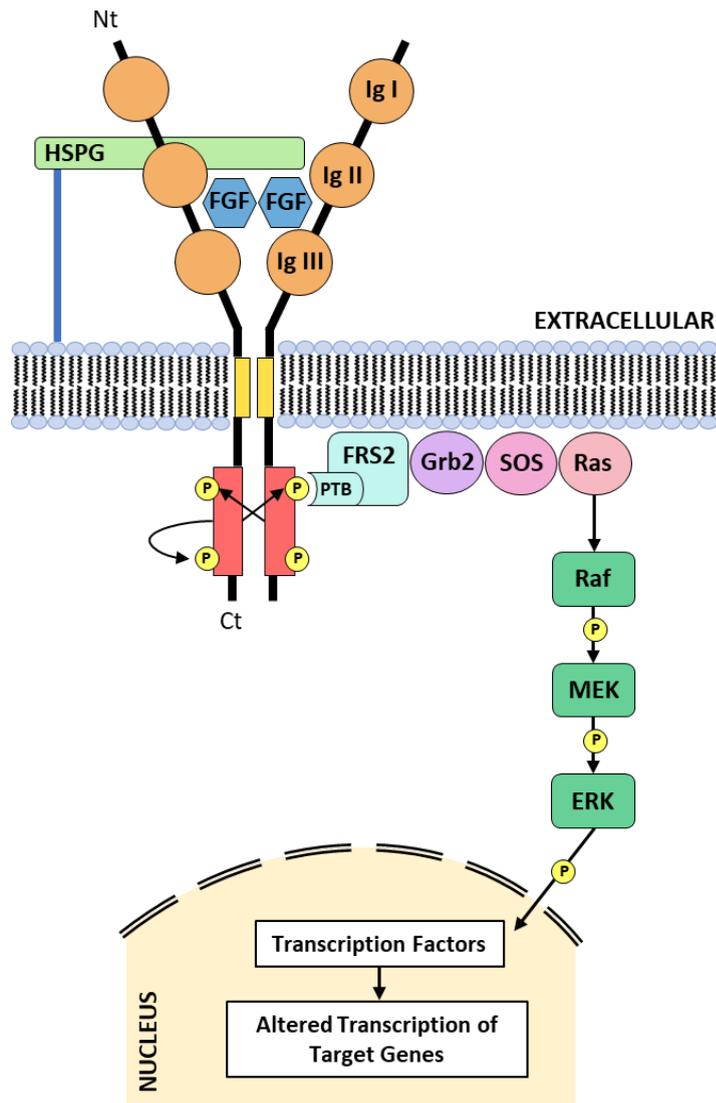
The changes in gene expression and cell behaviour which govern cell differentiation and tissue organisation during embryonic development are controlled by intricate signalling networks. Fibroblast growth factor (FGF) signalling, a member of the RTK signalling family, is a highly conserved signalling pathway with well-documented involvement in embryogenesis (Dorey and Amaya, 2010). Through regulation of *hox* genes, FGF signalling acts as a posteriorising agent and contributes to anteroposterior patterning of the *Xenopus* embryo (Pownall et al., 1996). A positive feedback loop exists between FGF signalling and the T-box transcription factor Xbra which is instrumental in the maintenance of mesoderm (Isaacs, Pownall and Slack, 1994; Fletcher and Harland, 2008). Further, amongst other contributions, FGF signalling also plays a key role in neural induction and limb development (Delaune, Lemaire and Kodjabachian, 2005; Min et al., 1998; Dorey and Amaya, 2010).

### 3.1.2 FGF ligands and receptors

As is prototypical of RTK signalling, FGF signalling can be stimulated through receptor dimerisation following ligand binding (Ullrich and Schlessinger, 1990). In total, 22 genes encoding FGF ligands and 4 genes encoding FGF receptors (*FGFR1-4*) have been characterised (Partanen, Vainikka and Alitalo, 1993). FGF ligands broadly fall into three categories: canonical, endocrine or intracellular (Ornitz and Itoh, 2015). Here, the focus is on canonical secreted FGF ligands (FGF1-10, 16-18, 20 and 22) which signal through FGFRs and represent the majority category (Ornitz and Itoh, 2015). The FGFR-ligand interactome is highly complex, with multiple ligands capable of binding to multiple different receptors. Ligand binding and specificity is largely attributed to three immunoglobulin-like domains (Ig1-3) within the extracellular portion of the FGFR (Partanen, Vainikka and Alitalo, 1993). Exons within the Ig3 domain are subject to alternative splicing, with the existence of multiple receptor isoforms adding further to the signalling complexity (Werner et al., 1992; Partanen, Vainikka and Alitalo, 1993). Heparan sulphate proteoglycans (HSPGs) in the extracellular matrix confer essential stability to the receptor-ligand and receptor-receptor interactions (Partanen, Vainikka and Alitalo, 1993; Lin et al., 1999).

### 3.1.3 The FGF signalling cascade

Following stimulatory ligand binding and receptor dimerisation, the intracellular tyrosine kinase domain mediates trans-autophosphorylation of the FGFR dimer (Lemmon and Schlessinger, 2010). The phospho-tyrosines subsequently act as docking sites for adaptors which relay the signal, through protein-protein interactions, to effector proteins which ultimately bring about the desired cellular response. To exemplify, the Ras-MAPK pathway activated downstream of FGFR signalling is initiated through the phosphotyrosine binding domain (PTB) containing adaptor protein Fibroblast growth factor receptor substrate 2 (FRS2; figure 4) (Kouhara et al., 1997). FRS2 interacts with Growth factor receptor-bound protein 2 (Grb2), which in turn interacts with Son of sevenless (SOS), and culminates in recruitment and activation of the small GTPase Ras at the cell surface membrane (Kouhara et al., 1997; Lowenstein et al., 1992; Ong et al., 2000). Following, Ras stimulates the MAPK cascade, through activation of serine/threonine kinase Raf, which phosphorylates and activates MEK, a serine/threonine/tyrosine kinase. Sequentially, MEK phosphorylates and activates the serine/threonine kinase ERK (Seger and Krebs, 1995). ERK is the effector of the Ras-MAPK pathway and through phosphorylation of transcriptional regulators and cytoskeletal components, is able to drive changes in gene expression and cell behaviour. Other eminent pathways activated downstream of FGFR signalling include the phospholipase C $\gamma$  (PLC $\gamma$ ) and phosphatidylinositol-3-kinase/Akt (PI3K/Akt) signalling pathways (Ornitz and Itoh, 2015).



**Figure 4: Schematic representation of the Ras-MAPK pathway induced downstream of FGF signalling.** Stimulatory ligand binding and receptor dimerisation stabilised by heparan sulphate proteoglycans (HSPGs) leads to trans-autophosphorylation of the intracellular tyrosine kinase domain of FGF receptors. Phosphotyrosine binding domain (PTB) containing proteins translate the signal through protein-protein interactions into activation of the MAPK signalling cascade, whereby the kinases Raf, MEK, and ERK sequentially phosphorylate and activate one another. Ultimately, ERK is the effector of the pathway and can drive changes in gene expression either directly through interaction with transcription factors, or indirectly through activation of alternative proteins which in turn regulate transcription factors. Other pathways activated downstream of FGF signalling include phospholipase C $\gamma$  (PLC $\gamma$ ) and phosphatidylinositol-3-kinase/Akt (PI3K/Akt) signalling pathways (not shown). FGF; fibroblast growth factor. PTB; phosphotyrosine binding domain. Ig; immuno-globulin-like domain.

### 3.1.4 Transcriptional regulators of FGF signalling downstream of the Ras-MAPK pathway

The Ras-MAPK pathway is perhaps the most well characterised signal transduction pathway downstream of FGF signalling. The targets of dpERK discussed in chapter 1 also represent targets of ERK as activated downstream of FGF signalling. For example, in *Xenopus*, FGF signalling through the Ras-MAPK pathway has been demonstrated to activate AP-1 activity which then synergises with the FGF/Xbra autocatalytic loop important for maintenance of the mesoderm (Kim et al., 1998). Whereas, phosphorylation of Elk-1, a TCF subfamily Ets transcription factor, by dpERK downstream of FGF signalling is important for differentiation of mesodermal derivatives, namely the notochord and muscle (Nentwich et al., 2009; Gille et al., 1995). Similarly, the regulation of Ets transcription factors downstream of FGF signalling is also implicated in neural induction and patterning (Hongo and Okamoto, 2022). The PEA3 subfamily of Ets transcription factors are significant targets of dpERK and FGF signalling (Garg et al., 2018; Znosko et al., 2010), as they mediate feedback inhibition of FGF signalling through induction of DUSP6 (Li et al., 2007; Ekerot et al., 2008). A common feature of signalling pathways is upregulation of their respective negative regulators to enable a transient response to stimuli; dual specificity phosphatase (DUSP) enzymes are able to downregulate the Ras-MAPK pathway through de-phosphorylation and inactivation of MAPKs (Ekerot et al., 2008).

### 3.1.5 CIC may regulate of a subset of FGF target genes

Alongside regulation downstream of FGF signalling, the PEA3 subfamily of Ets transcription factors are regulated by the transcriptional repressor CIC (Kawamura-Saito et al., 2006; Weissmann et al., 2018). Moreover, in mouse embryonic stem cells, CIC was found to associate with the genomic locus of FGF-target gene *DUSP6* (Weissmann et al., 2018). CIC itself is negatively regulated by dpERK, with phosphorylation of CIC by dpERK capable of relieving CIC-mediated gene repression (see chapter 1). Notably, as the majority of ERK activation during early *Xenopus* development is downstream of FGF signalling (Christen and Slack, 1999; Branney et al., 2009), FGF signalling may negatively regulate CIC activity through the Ras-MAPK pathway. Indeed, CIC expression has previously been demonstrated in domains of FGF activity (King, 2018). Furthermore, as the FGF-target gene *xmyod* is induced following cycloheximide treatment, with one explanation being that cycloheximide inhibits translation of a transcriptional repressor, this implies that FGF-target genes may be regulated by a labile transcriptional repressor (Fisher, Isaacs and Pownall, 2002). Hence, it is an attractive proposal that CIC negatively regulates a subset of

FGF-target genes and presents an additional phosphotarget of dpERK downstream of FGF signalling.

### 3.1.6 Chapter hypothesis and aims

Hypothesis:

During *Xenopus* development ERK relieves CIC-mediated repression downstream of FGF signalling, and by extension, CIC negatively regulates a subset of FGF target genes.

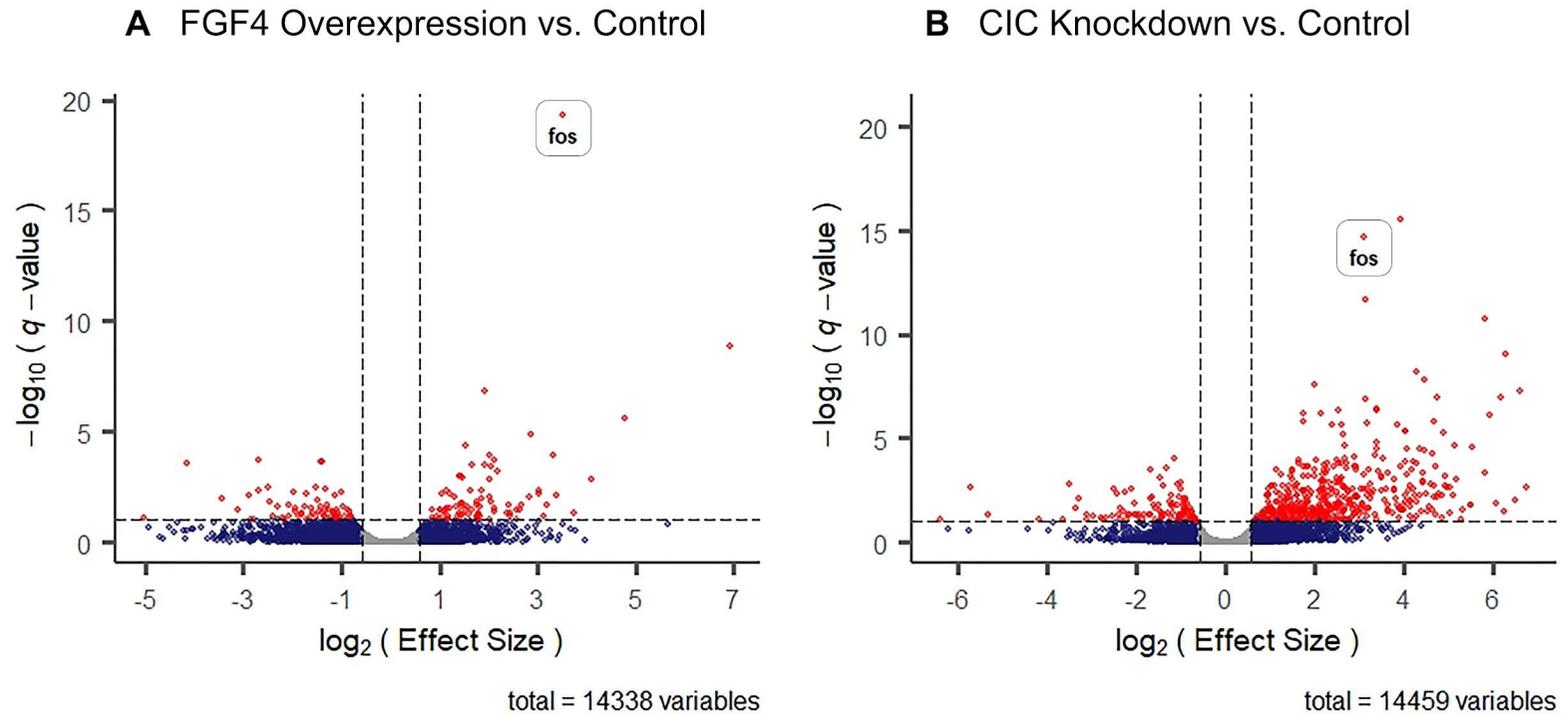
Aims of this chapter:

- Perform gene-level differential expression analysis on RNA-seq data from FGF overexpressing and CIC knockdown embryos to identify genes putatively regulated by both FGF signalling and CIC.
- Undertake gene ontology enrichment analysis to highlight potential biological processes and components associated with genes upregulated following CIC knockdown.
- Investigate putative FGF/CIC target genes by validating their regulation by FGF signalling and analysing their genomic loci for candidate CIC binding sites.

## 3.2 Results

### 3.2.1 Gene-level differential expression analysis of FGF4 overexpressing and CIC knockdown *X. tropicalis* embryos

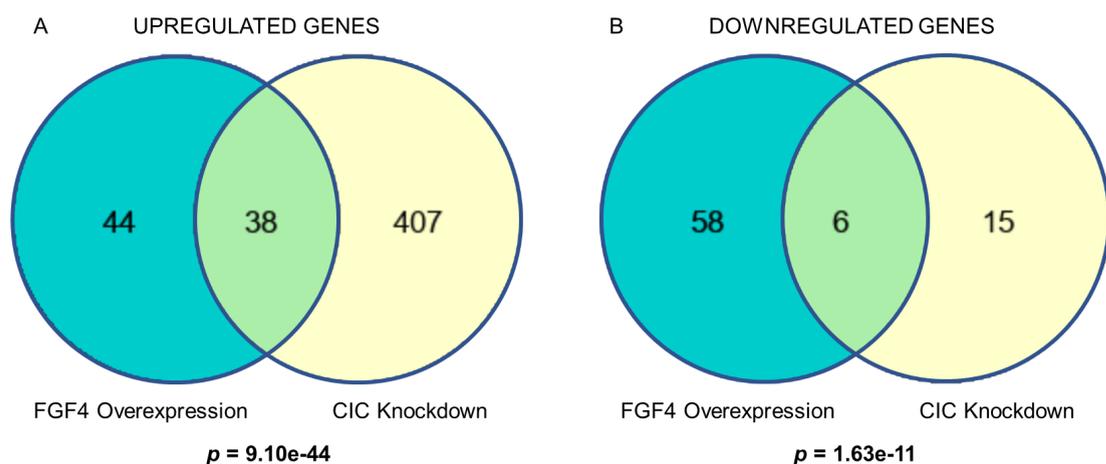
A RNA-seq study in *X. tropicalis* embryos, conducted as part of a PhD within the Isaacs lab, aimed to independently identify targets of FGF signalling and CIC through microinjection of a plasmid overexpressing FGF4 (CKSA-eFGF; (Isaacs, Pownall and Slack, 1994)), and transcription activator-like effector nucleases (TALENs) designed to knockdown CIC, respectively (King, 2018). An initial bioinformatics analysis was performed by the University of York Bioscience Technology Facility, which included aligning reads to the *X. tropicalis* reference genome using the R package Salmon (Patro et al., 2017), and producing a transcript-level statistical analysis. As an extension, this project utilised the Salmon output to perform a gene-level analysis through the R package Sleuth via the counts aggregation method (Pimentel and McGee, 2021). As principal component analysis indicated a batch effect (*data not shown*), both batch and treatment were modelled as variables. Both the FGF4 overexpression and CIC knockdown gene-level RNA-seq data was filtered to remove genes without an average transcripts per million (TPM) value of  $> 1$  in either the control or treatment data. This left a total of 14,338 and 14,459 genes for the FGF4 overexpression and CIC knockdown data, respectively (additional file S1). A gene significantly up or down-regulated following FGF4 overexpression was defined as having a  $q$ -value  $< 0.1$  and an effect size of  $> 1.5$ , or  $< 1/1.5$ , respectively. This isolated 146 genes in total, with the likelihood of  $< 15$  false positives, and outputted well characterised targets of FGF signalling including for example *pax6*, and *dusp6* (Bertrand, Médevielle and Pituello, 2000; Ekerot et al., 2008). Of the 14,338 genes in total, 82 genes met the criteria to be defined as significantly upregulated following FGF4 overexpression, and 64 genes were identified as significantly downregulated (figure 5A, additional file S2). For the CIC knockdown dataset, the same criteria were used to define significance; of the 14,459 genes in total, 445 met the criteria to be defined as significantly upregulated, and 21 genes as significantly downregulated (figure 5B, additional file S3).



**Figure 5: Volcano plots of RNA-seq gene-level differential expression analysis for FGF4 overexpressing and CIC knockdown stage 14 *X. tropicalis* embryos.** King (2019) performed triplicate RNA-seq experiments in stage 14 *X. tropicalis* embryos injected with either CSKA-eFGF (FGF4 overexpression), Capicua TALENs (CIC knockdown), or water (control). Initial quality control and read-mapping was undertaken by Dr. John Davey, University of York. Gene-level differential expression analysis was then performed from the Salmon output using Sleuth. **(A)** Volcano plot of statistical significance (FDR corrected  $p$ -value) against effect size for 14,338 annotated *X. tropicalis* genes in FGF4 overexpression versus control embryos. **(B)** Volcano plot of statistical significance (FDR corrected  $p$ -value) against effect size for 14,459 annotated *X. tropicalis* genes in CIC knockdown versus control embryos. Each point represents an individual gene. Genes highlighted in red represent those which meet criteria of both  $q$ -value  $< 0.1$  and an effect size of  $< 1/1.5$ , or  $> 1.5$ , whilst genes in blue meet the effect size criteria but not the  $q$ -value cut off. Genes in grey do not meet either criteria.

### 3.2.2 A subset of FGF target genes are putatively regulated by CIC

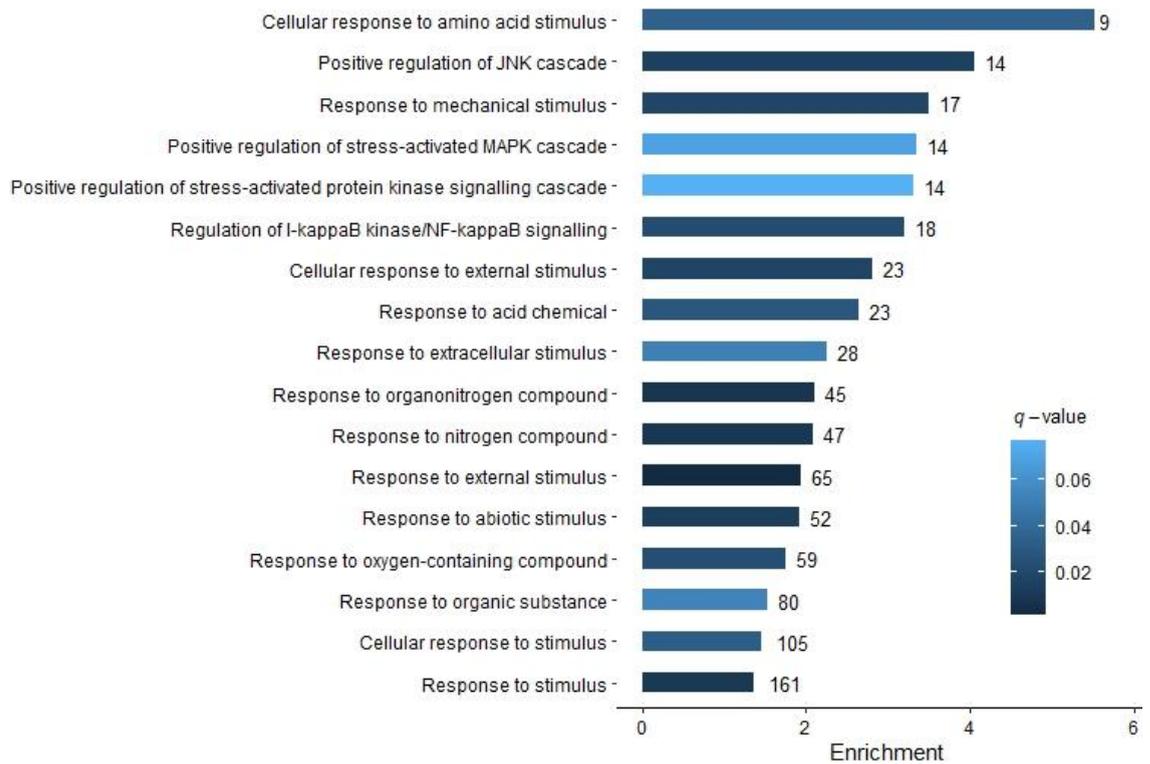
It is predicted that if CIC is acting as a negative regulator downstream of FGF signalling, then inhibiting CIC activity should upregulate FGF target genes. Indeed, 38 genes were identified as significantly upregulated upon both FGF4 overexpression and CIC knockdown, and a further 6 genes significantly downregulated by both CIC knockdown and FGF4 overexpression (figure 6, additional file S4). A Fisher's exact test determined the size of the intersection between genes upregulated by FGF4 overexpression and CIC knockdown, and genes downregulated by FGF4 overexpression and CIC knockdown to be highly significant in both cases ( $p=9.10e-44$ ,  $p=1.63e-11$ , respectively). Henceforth, these 44 genes are candidate target genes of both CIC and FGF signalling, and are supportive of the hypothesis that during early *Xenopus* development CIC acts as a transcriptional repressor downstream of FGF signalling.



**Figure 6: Venn diagram of overlap between significantly upregulated and downregulated genes following FGF4 overexpression and CIC knockdown in stage 14 *X. tropicalis* embryos. (A)** Criteria for defining significant upregulation following FGF4 overexpression and CIC knockdown: RNA-seq  $q$ -value  $< 0.1$ , effect size  $> 1.5$ . A significant number of genes (38) are upregulated by both FGF4 overexpression and CIC knockdown ( $p = 9.10e-44$ ) **(B)** Criteria for defining significant downregulation following FGF4 overexpression and CIC knockdown: RNA-seq  $q$ -value  $< 0.1$ , effect size  $< 1/1.5$ . A significant number of genes (6) are downregulated by both FGF4 overexpression and CIC knockdown ( $p = 1.63e-11$ )

### 3.2.3 Gene-ontology enrichment analysis for genes significantly upregulated following CIC knockdown

Gene-ontology analysis was used to investigate candidate biological processes and components regulated by CIC during embryonic development. Lower stringency criteria ( $p$ -value < 0.01 and effect size > 1.5) were used to define genes significantly upregulated following CIC knockdown. This increased the size of the input list for GOzilla ontology analysis and enabled identification of a greater number of ontology terms (additional file S5). Enrichment of biological processes was assessed using the human database as a reference. In total, 19 terms associated with biological processes achieved a  $p$ -value <  $1e-4$  (figure 7, table 4). Of particular interest are terms including 'positive regulation of JNK cascade' and 'positive regulation of stress activated MAPK cascade'. As both Jun N-terminal kinase (JNK) and MAPK signalling pathways are activated downstream of RTK signalling, including FGF signalling (Ornitz and Itoh, 2015), this is in keeping with a role for CIC downstream of FGF signalling. Only one biological component term, 'transcription factor AP-1 complex' achieved a  $p$ -value of <  $1e-4$ , with the four genes associated with this term being *fos*, *jun*, *junb* and *jund* (table 5). This implicates a role for CIC in regulating AP-1.



**Figure 7: GOrilla gene-ontology analysis of biological processes for genes significantly upregulated following CIC knockdown.** Genes which met the criteria of  $p$ -value  $< 0.01$  and effect size  $> 1.5$  following CIC knockdown were defined as significantly upregulated (additional file S5). Results were filtered to show ontology terms which achieved  $p$ -value  $< 1e-4$ , and these are displayed on the y-axis. The x-axis depicts fold enrichment relative to expected enrichment using the human database as a reference. A colour scale is used to indicate  $q$ -value.

**Table 4: GOrilla gene-ontology analysis of biological processes for genes significantly upregulated following CIC knockdown.** Genes which met the criteria of  $p$ -value  $< 0.01$  and effect size  $> 1.5$  following CIC knockdown were defined as significantly upregulated (additional file S5). Results were filtered to show ontology terms which achieved  $p$ -value  $< 1e-4$ . Enrichment of biological processes was assessed using the human database as a reference. FDR; false discovery rate. GO; gene ontology.

GO Term	Description	$p$ -value	FDR $q$ -value	Enrichment	Number of Genes
GO:0071230	Cellular response to amino acid stimulus	3.09E-05	0.0356	5.51	9
GO:0046330	Positive regulation of JNK cascade	8.59E-06	0.0161	4.06	14
GO:0009612	Response to mechanical stimulus	7.63E-06	0.019	3.49	17
GO:0032874	Positive regulation of stress-activated MAPK cascade	0.000075	0.0702	3.35	14
GO:0070304	Positive regulation of stress-activated protein kinase signalling cascade	8.77E-05	0.0773	3.31	14
GO:0043122	Regulation of I-kappaB kinase/NF-kappaB signalling	0.000014	0.0232	3.20	18
GO:0071496	Cellular response to external stimulus	8.42E-06	0.018	2.81	23
GO:0001101	Response to acid chemical	2.13E-05	0.029	2.65	23
GO:0009991	Response to extracellular stimulus	5.23E-05	0.0523	2.26	28
GO:0010243	Response to organonitrogen compound	1.69E-06	0.00844	2.11	45
GO:1901698	Response to nitrogen compound	1.35E-06	0.0101	2.09	47
GO:0009605	Response to external stimulus	1.36E-07	0.00204	1.94	65
GO:0009628	Response to abiotic stimulus	4.8E-06	0.0144	1.91	52

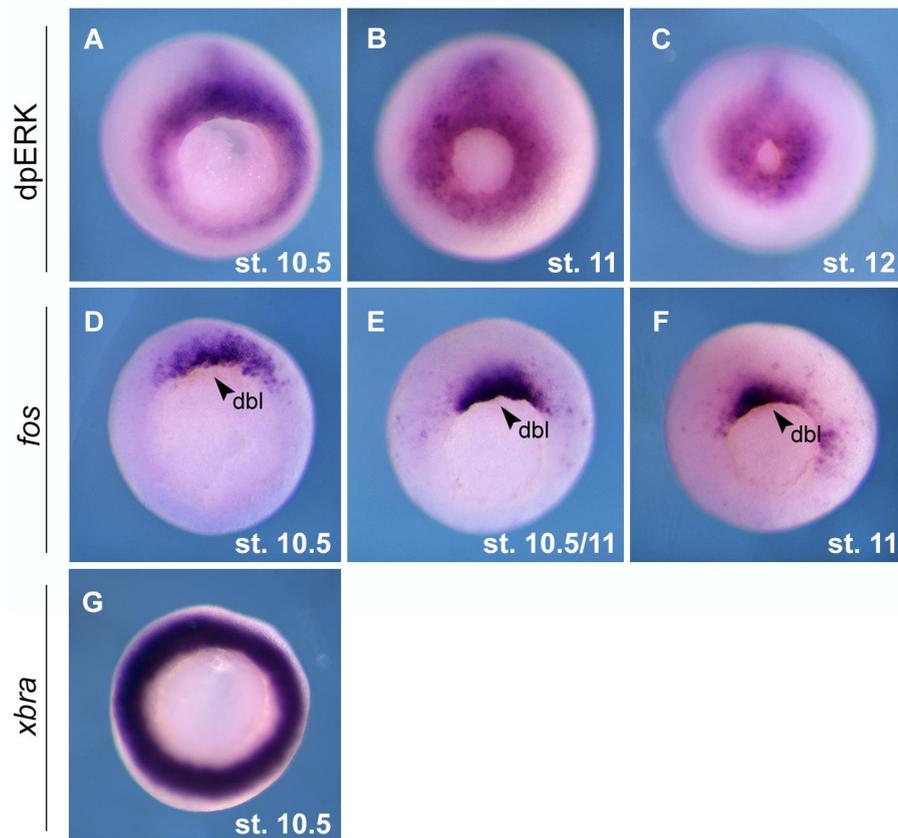
GO:1901700	Response to oxygen-containing compound	1.62E-05	0.0243	1.75	59
GO:0010033	Response to organic substance	5.05E-05	0.054	1.53	80
GO:0051716	Cellular response to stimulus	2.65E-05	0.033	1.45	105
GO:0050896	Response to stimulus	2.82E-06	0.0106	1.36	161

**Table 5: GOrilla gene-ontology analysis of biological components for genes significantly upregulated following CIC knockdown.** Genes which met the criteria of  $p$ -value  $< 0.01$  and effect size  $> 1.5$  following CIC knockdown were defined as significantly upregulated (additional file S5). Results were filtered to show ontology terms which achieved  $p$ -value  $< 1e-4$ . Enrichment of biological components was assessed using the human database as a reference. The four genes associated with the gene ontology term 'Transcription factor AP-1 complex' are *fos*, *jun*, *jund* and *junb*. FDR; false discovery rate. GO; gene ontology.

GO Term	Description	$p$ -value	FDR $q$ -value	Enrichment	Number of Genes
GO:0035976	Transcription factor AP-1 complex	8.26E-7	1.58E-3	33.06	4

### 3.2.4 Spatial and temporal co-localisation of *fos* and dpERK in gastrula stage *X. tropicalis*

The AP-1 component gene *fos* is highly significantly upregulated following both CIC knockdown ( $q = 2.20e-15$ , effect size = 8.56) and FGF4 overexpression ( $q = 5.21e-20$ , effect size = 11.33), and is therefore a strong candidate gene to be regulated by CIC downstream of FGF signalling. Indeed, in mouse and human cell lines, CHIP-seq with an antibody against CIC identified *fos* as a directly bound by CIC (Weissmann et al., 2018). Furthermore, Fos is a known phosphotarget of dpERK and, as part of AP-1, is implicated in FGF-mediated regulation of mesoderm induction (Okazaki and Sagata, 1995; Kim et al., 1998). Regulation of Fos at both the transcriptional and post-translation level by dpERK could prove an efficient way to fine-tune Fos activity to dynamic dpERK expression. To investigate whether FGF signalling might regulate *fos* through activation of ERK and subsequent negative regulation of CIC, the spatial and temporal expression of dpERK and *fos* were assessed via immunostaining and *in situ* hybridisation, respectively. *In situ* hybridisations were also undertaken for *xbra*, an exemplar FGF-responsive gene which shows expression around the blastopore (Schulte-Merker and Smith, 1995); figure 8G). For gastrula stage *X. tropicalis* embryos, *fos* expression localises at the dorsal blastopore lip (figure 8D-F). Likewise, dpERK expression encompasses the whole perimeter of the blastopore, and at dorsal regions extends further towards the animal hemisphere (figure 8A-C). As during early *Xenopus* development the majority of ERK activation is downstream of FGF signalling (Christen and Slack, 1999; Branney et al., 2009), the co-occurrence of *fos* and dpERK expression during gastrulation in known regions of FGF activity is supportive of *fos* being regulated downstream of FGF signalling and ERK.

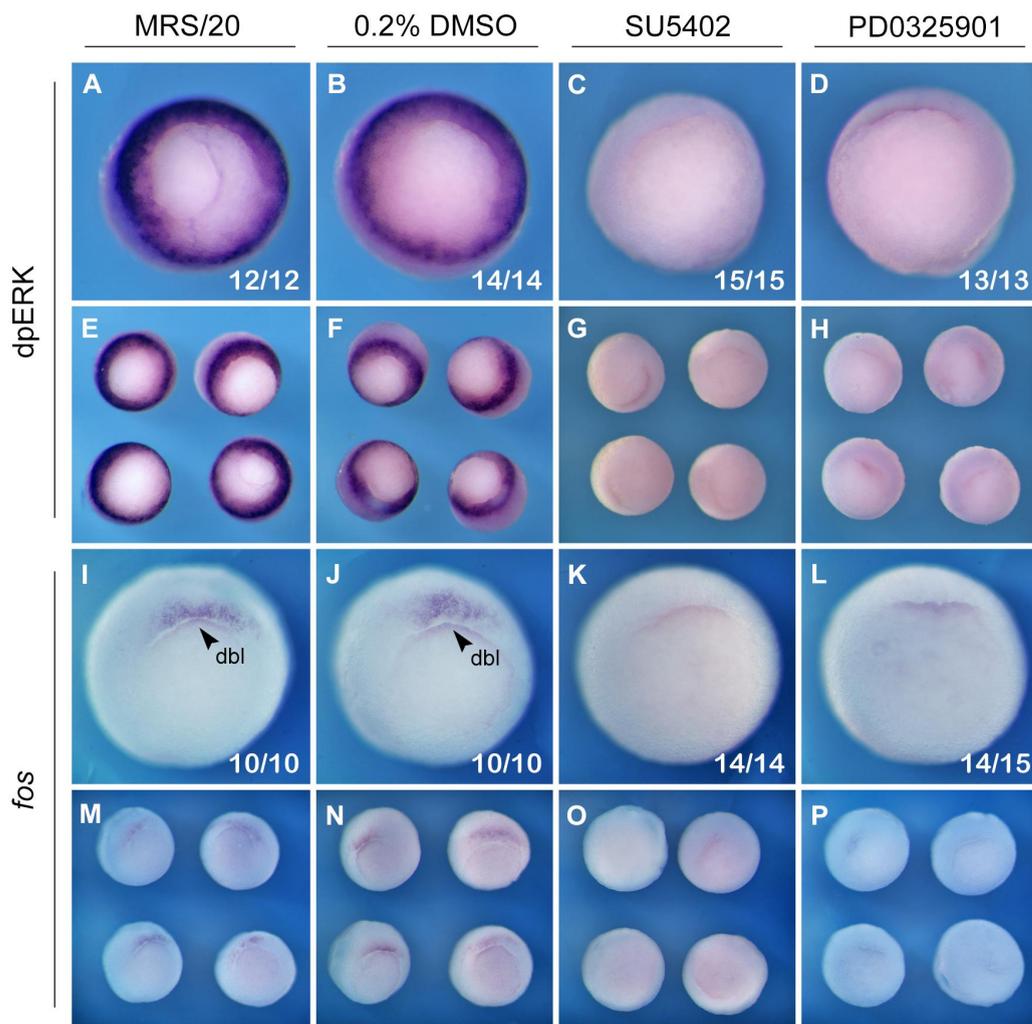


**Figure 8: *X. tropicalis* gastrula stage series of dpERK and *fos* expression analysed via immunostaining and *in situ* hybridisation, respectively. (A) stage 10.5, (B) stage 11, (C) stage 12, dpERK immunostaining, vegetal view, dorsal top. (D) stage 10.5, (E) stage 10.5 - 11, (F) stage 11, *fos in situ* hybridisation, vegetal view, dorsal top. (G) stage 10.5 *xbra in situ* hybridisation, vegetal view. Arrowhead indicates dorsal blastopore lip 'dbl'. St.; stage. Images are representative of n=10 embryos, and are not to scale.**

### 3.2.5 FGF and MEK inhibition prevent *fos* induction in gastrula stage *X. tropicalis*

To determine whether dpERK provokes dorsal *fos* expression during gastrulation, the chemical inhibitors SU5402 and PD0325901 were used to inhibit FGF signalling, and MAPK signalling, respectively (Mohammadi et al., 1997; Barrett et al., 2008). SU5402 inhibits FGF expression at the level of the receptor (Mohammadi et al., 1997), whilst PD0325901 prevents phosphorylation and activation of MEK, the kinase upstream of ERK, which is responsible for phosphorylation and activation of ERK (Barrett et al., 2008). Stage 8 *X. tropicalis* embryos were cultured for a minimum of 2 hours in either normal culture medium (MRS/20), MRS/20 containing 0.2% DMSO, 200  $\mu$ M SU5402 or 25  $\mu$ M PD0325901 diluted in MRS/20 containing 0.2% DMSO. At early gastrula stage 10, embryos were collected and fixed in MEMFA. Subsequently, dpERK and *fos* expression were assayed via immunostaining and *in situ* hybridisation, respectively. In agreement with the majority of ERK activation during early *Xenopus* development

being downstream of FGF signalling (Christen and Slack, 1999; Branney et al., 2009), treatment with SU5402 was sufficient to suppress ERK activation around the blastopore (figure 9C,G). PD0325901 treatment, which targets MAPK signalling specifically, was also able to prevent ERK activation around the blastopore (figure 9D,H). Both SU5402 and PD0325901 treatment reduced *fos* expression around the dorsal blastopore lip relative to MRS/20 and DMSO controls (figure 9I-P), thus implicating *fos* induction during early gastrulation as being both FGF-dependent and through the MAPK pathway.

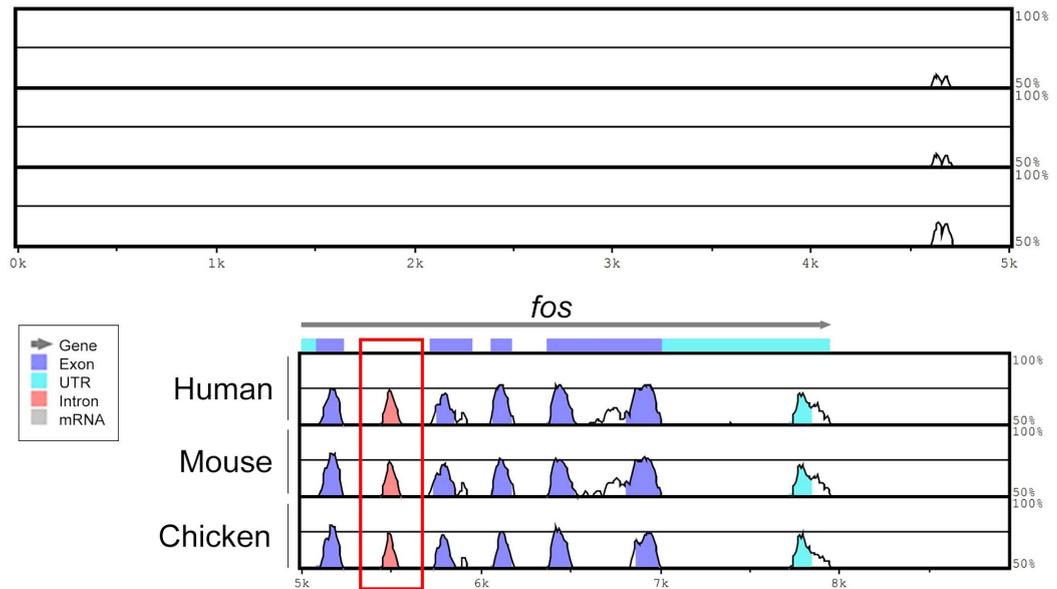


**Figure 9: SU5402 and PD0325901 inhibit early gastrula stage dpERK and *fos* expression in *X. tropicalis* embryos, as assayed via immunostaining and *in situ* hybridisation, respectively.** *X. tropicalis* sibling embryos were cultured from stage 8 for a minimum of 2 hours in either MRS/20 (A,E,I,M), MRS/20 containing 0.2% DMSO (B,F,J,N), 200  $\mu$ M SU5402 in 0.2% DMSO (C,G,K,O) or 25  $\mu$ M PD0325901 in 0.2% DMSO (D,H,L,P). At approximately early gastrula stage 10, embryos were fixed in MEMFA ready for subsequent immunostaining with anti-dpERK (M9692) (A-H) and *fos* *in situ* hybridisation (I-P). The number of embryos observed with the displayed phenotype is indicated in the bottom right hand corner of individual embryo images (A-D, I-L). Family pictures of four representative embryos are also shown (E-H, M-P). Black arrowhead indicates the dorsal blastopore lip 'dbl'. Images not to scale.

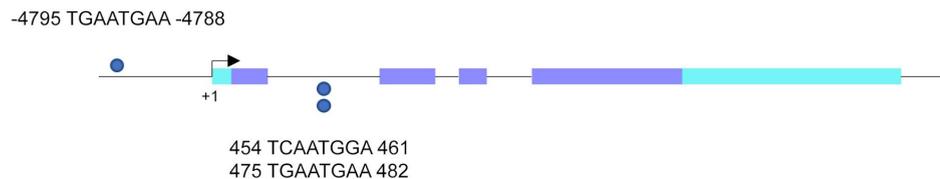
### 3.2.6 CIC binding sites exist within a highly conserved region of *fos* intron 1

Both the FGF4 overexpression RNA-seq data (figure 5A) and *fos in situ* hybridisations of gastrula stage *X. tropicalis* treated with FGF and MEK chemical inhibitors (figure 9) point towards a role for FGF signalling regulating *fos* expression through the Ras-MAPK pathway. To investigate the plausibility of this being via negative regulation of CIC by dpERK, the *fos* genetic locus was analysed for CIC binding motifs. As regulatory elements are often conserved, a LAGAN alignment was first performed to help identify any candidate promoter or enhancer elements (Brudno et al., 2003). The conservation of *H. sapien*, *M. musculus* and *G. gallus fos* coding sequence, alongside 5 kb upstream and 1 kb downstream, was investigated using the *X. tropicalis* sequence as a reference. Alongside exonic regions and a region within the 3' UTR, two other conserved regions were identified, one within 500 bp upstream of the transcriptional start site, and another centrally within intron 1 (figure 10A). FIMO motif scanning was performed for the CIC octameric consensus motif (TSAATGRA) within 5 kb upstream, 1 kb downstream, and the coding region of *X. tropicalis fos* (Grant, Bailey and Noble, 2011). Three matches to the CIC consensus were found with  $p$ -value  $< 1e-4$ , and interestingly two of these matches mapped roughly centrally to intron 1 coincident with the region of high conservation (figure 10A,B, table 6). When the *H. sapien*, *M. musculus* and *G. gallus* intron 1 regions of *fos* were also analysed, at least one CIC consensus was locatable in all 3 species (table 6). In corroboration with *X. tropicalis*, in all other species, the CIC motif matches mapped approximately centrally within intron 1 and onto the negative strand (table 6). A SEA motif enrichment analysis of *fos* intron 1 for all four species found a significant ( $q = 1.43e-2$ ) 5-fold enrichment of the CIC consensus with 100% true positives, and 0% false positives (additional file S8). The presence of CIC binding motifs is supportive of *fos* being a direct target of CIC, and it is intriguing to propose that a candidate site for CIC binding is within a conserved region of intron 1. Taken alongside the evidence that *fos* expression is also regulated by FGF signalling (figure 9), this is supportive of the hypothesis that CIC regulates the expression of a subset of FGF-target genes.

**A** Sequence conservation analysis of *fos* locus relative to *X. tropicalis*



**B** CIC motif scanning of *X. tropicalis fos* locus



**Figure 10: Conservation of *fos* genetic locus and schematic representation of candidate CIC binding sites for *X. tropicalis fos*.** (A) Image adapted from VISTA genome browser output (LAGAN alignment programme) (Brudno et al., 2003). Conservation of *H. sapien* (human), *M. musculus* (mouse) and *G. gallus* (chicken) *fos* coding region alongside 5 kb upstream and 1 kb downstream relative to *X. tropicalis* as a reference. Percentage of conservation identity and minimum conservation window parameters were set as 70% and 100 bp. Exonic regions with > 70% conservation are depicted in dark blue, UTRs in light blue, and introns in red. The direction of mRNA transcription is indicated by the grey arrow. The high conservation within intron 1 is particularly of interest and is highlighted with a red box. (B) Schematic representation of FIMO output for *X. tropicalis fos* coding region alongside 5 kb upstream and 1 kb downstream (Grant, Bailey and Noble, 2011). Sequence was scanned for the CIC octameric TSAATGRA motif. Exonic regions are depicted in dark blue, UTRs in light blue, and the transcriptional start site (+1) and direction of transcription is indicated. A match to the consensus with  $p < 1e-4$  is indicated by a dark blue circle, with the exact matched sequence and position relative to (+1) indicated. Matches on the positive strand are above the locus schematic, and matches on the negative strand below. Image is not to scale.

**Table 6: Motif scanning for CIC consensus motif in *X. tropicalis*, *H. sapien*, *M. musculus* and *G. gallus fos* intron 1.** FIMO was used to scan for the CIC consensus motif (TSAATGRA) in frog, human, mouse and chicken *fos* intron 1 (Grant, Bailey and Noble, 2011). Matches were filtered to have  $p$ -value  $< 1e-4$ . The sequence and position in intron 1 of the matches is indicated, whereby a position of 1 is representative of the start of intron 1. bp; base pairs.

<b>Species</b>	<b>Intron 1 length (bp)</b>	<b>Sequence and position relative to intron 1</b>	<b>Strand (+/-)</b>	<b><math>p</math>-value</b>	<b><math>q</math>-value</b>
<i>X. tropicalis</i> (Frog)	489	225 TCAATGGA 232	-	3.58e-05	0.0636
		246 TGAATGAA 253	-	7.93e-05	0.0636
<i>H. sapien</i> (Human)	753	420 TCAATGAA 427	-	7.93e-05	0.0636
<i>M. musculus</i> (Mouse)	754	415 TCAATGAA 422	-	7.93e-05	0.0636
<i>G. gallus</i> (Chicken)	438	236 TCAATGAA 243	-	7.93e-05	0.0636
		255 TCAATGAA 262	-	7.93e-05	0.0636

## 3.3 Discussion

### 3.3.1 Gene-level differential expression analysis output of CIC target genes

Gene-level differential expression analysis of RNA-seq data in which CIC had been knocked down in *X. tropicalis* embryos enabled identification of 466 putative CIC target genes. Genes which met the criteria of  $q$ -value  $< 0.1$  and effect size of  $> 1.5$  or  $< 1/1.5$  were defined as significantly up or downregulated following CIC knockdown, respectively (figure 5B, additional file S3). A number of the genes identified as CIC targets by RNA-seq in *X. tropicalis*, including *fos*, *fosb*, *fosl1* and *klf4*, were also outputted as CIC targets via ChIP-seq in mouse embryonic and/or human foetal neural and glioma stem cell lines (Weissmann et al., 2018). This is indicative of a conserved role for CIC across vertebrates. Interestingly however, the chosen criteria for significance for the RNA-seq data did not output known mammalian CIC targets from the PEA3 subfamily of Ets transcription factors, *ETV1*, *ETV4* and *ETV5* (Dissanayake et al., 2011; Weissmann et al., 2018). Perhaps this reflects the multifaceted nature of transcriptional regulation and that relief of repression alone may not be sufficient to actively induce expression of these genes at the developmental stage investigated. Alternatively, perhaps there is some disparity between CIC targets in different species, as Ets family genes are yet to be conclusively identified as CIC targets in *Xenopus*. Overall, the greater number of significantly upregulated genes (445) relative to downregulated genes (21) following CIC knockdown is consistent with CIC acting as a transcriptional repressor (Jiménez et al., 2000). Genes significantly downregulated following CIC knockdown likely represent indirect targets of CIC, although one study has suggested CIC may also function as a transcriptional activator (Yang et al., 2017).

### 3.3.2 CIC putatively regulates a subset of FGF target genes

In agreement with the hypothesis that during early *Xenopus* development CIC acts as a transcriptional repressor downstream of FGF signalling, a significant overlap was found to exist between genes upregulated by FGF4 expression and CIC knockdown, and between genes downregulated by FGF4 overexpression and CIC knockdown (figure 6). A total of 44 genes were found to be regulated by both FGF4 and CIC, however the intersection is not complete, with groups of genes appearing exclusively regulated by FGF4 or CIC. Overexpression of a single FGF ligand may not capture the full range of FGF-target genes regulated by CIC, as other targets may be induced through different canonical FGF ligands. Further, in later stages of *Xenopus* development, FGF signalling is not the sole mediator of ERK activation, with ERK

activation possible downstream of alternative RTK signalling pathways (Christen and Slack, 1999). In the context of embryonic wound healing, ERK activation is also FGF-independent (Christen and Slack, 1999) Therefore, negative regulation of CIC by dpERK may not be indefinitely reliant on FGF signalling, and so CIC may regulate genes outside of FGF-target genes. Likewise, alongside activating the Ras-MAPK pathway, FGF signalling also notably activates the PLC $\gamma$  and PI3K/Akt signalling pathways; these pathways may not impact upon CIC activity (Ornitz and Itoh, 2015).

Gene ontology analysis for enriched biological processes following CIC knockdown is also supportive of the hypothesis. Enriched biological process terms included 'positive regulation of JNK cascade' and 'positive regulation of stress activated MAPK cascade' (figure 7, table 4). As both JNK and MAPK signalling pathways are activated downstream of RTK signalling, including FGF signalling (Ornitz and Itoh, 2015), this is in keeping with a role for CIC downstream of FGF signalling. Broadly, many of the enriched biological process terms relate to stimulus response, with three terms specifically denoting response to external/extracellular stimuli. This is once again in agreement with CIC functioning downstream of RTK signalling, which is commonly activated in response to secreted extracellular mitogens. As ERK-independent regulation of CIC has been recently proposed (Papagianni et al., 2018; Yang et al., 2016; Bunda et al., 2020), it is possible some ontology terms may also reflect CIC functioning outside of Ras-MAPK signalling. Gene ontology analysis for biological components enriched following CIC knockdown outputted the transcription factor AP-1 complex (table 5). AP-1 has known involvement in FGF-mediated developmental processes, including for example mesoderm and otic induction (Tambalo et al., 2020; Kim et al., 1998). Therefore, the putative regulation of AP-1 component genes including *fos*, *jun*, *junb*, and *jund* by CIC is also supportive of a regulatory role for CIC downstream of FGF signalling.

### 3.3.3 A role for CIC in feedback inhibition of FGF signalling

It is interesting to note that several negative regulators of FGF signalling appear as targets of CIC. For example, following CIC knockdown *adamts1* is significantly upregulated (additional file S3); ADAMTS1 is able to inhibit the phosphorylation of ERK mediated by FGF4 in *Xenopus* animal caps and reduce the expression of the FGF-responsive and mesoderm-inducing gene *xbra* (Suga, Hikasa and Taira, 2006). Moreover, *adamts1* was also identified as a target of CIC via ChIP-seq in a human glioma stem cell line (Weissmann et al., 2018), hence implying it may be a conserved and direct target of CIC. Although not identified as targets of CIC in the CIC knockdown RNA-seq data, ChIP-seq identified *DUSP6*, *SPRED1*, *SPRED2* and *SPRED3* as

targets of CIC in mouse embryonic and human foetal neural stem cell lines (Weissmann et al., 2018). *Dusp6* and *Spred1-3* are all induced downstream of FGF signalling and act as negative regulators of ERK activity (Li et al., 2007; Wakioka et al., 2001; Hirata et al., 2016). A common feature of signalling pathways is often upregulation of their own negative feedback mechanisms. Thus, this provides a further interesting link between CIC and FGF signalling.

### 3.3.4 The AP-1 component gene *fos* is putatively a target of CIC/FGF

One gene identified as highly significantly upregulated following CIC knockdown was *fos* (figure 5B). Alongside the output of the gene-level differential expression analysis conducted here (additional file S3), *fos* has also been outputted as a target of CIC via ChIP-seq in human foetal neural stem cell lines and a human glioma stem cell line (Weissmann et al., 2018). Overall, this implies *fos* may be another direct and conserved target of CIC in vertebrates. Furthermore, *fos* is also an attractive target of FGF signalling. In addition to being highly significantly upregulated following FGF4 overexpression in *Xenopus* (figure 5A, additional file S2), in human embryonic stem cells, *fos* induction has been demonstrated downstream of FGF2-induced signalling (Kang et al., 2005). As a constituent of AP-1, Fos is also implicated in FGF-mediated regulation of mesoderm induction (Kim et al., 1998). Accordingly, *fos* is a strong candidate target of both FGF signalling and CIC, and provides a suitable gene to investigate the role of ERK in relieving CIC-mediated repression of FGF target genes.

### 3.3.5 Early gastrula expression of *fos* is FGF and ERK-dependent in *Xenopus*

In early gastrula stage *X. tropicalis* embryos there is co-localisation of *fos* expression and dpERK expression at the dorsal blastopore lip which is supportive of gastrula stage *fos* expression being regulated by dpERK (figure 8A-F). Further, this is a known region of FGF-activity, with the FGF-responsive gene *xbra* also showing strong expression around the blastopore (figure 8G) (Schulte-Merker and Smith, 1995). Importantly, both CIC isoforms are enriched at the dorsal blastopore lip at this stage (King, 2018). As both the FGFR inhibitor SU5402 and the MEK inhibitor PD0325901 inhibited *fos* expression at the dorsal blastopore lip, this contextually implies *fos* expression is both FGF-dependent and ERK-dependent (figure 9). Overall, this is supportive of FGF signalling regulating *fos* expression through ERK-mediated negative regulation of CIC. However, as SU5402 has known affinity for other RTKs, including vascular endothelial growth factor receptor 2 (VEGFR2) and platelet-derived growth factor receptor beta

(PDGFRB), a contribution from these RTKs to *fos* induction can't be completely dismissed. Nonetheless, any contribution is presumably minimal, as through use of a dominant negative FGFR designed to specifically inhibit FGF signalling, it has previously been demonstrated the majority of ERK activation during early *Xenopus* development is downstream of FGF signalling (Christen and Slack, 1999; Branney et al., 2009).

### 3.3.6 Candidate regulation of a *fos* intronic promoter by CIC

To better understand whether *fos* represents a direct target of CIC, the *fos* genomic locus was investigated for CIC binding sites. Interestingly, the CIC octameric consensus binding motif (TSAATGRA) was found to locate in a highly conserved region within *fos* intron 1 in human, mouse, chicken and frog species (figure 10, table 6). Previous studies have also denoted this region of high conservation and have implicated it in a phenomenon referred to as intron mediated enhancement (IME), whereby the presence of an intron is able to enhance gene expression through a variety of mechanisms (Coulon et al., 2010). In *M. musculus* (mouse), the identified CIC binding motif in intron 1 is +415 bp relative to the start of intron 1 (table 6) which corresponds to +706 bp relative to the transcriptional start site. This maps profoundly well onto a DNase resistant region previously identified in *fos* at +700 bp relative to the transcriptional start site in a mouse fibroblast cell line (Renz et al., 1985). Intriguingly, this intronic region in isolation has been demonstrated to have promoter activity, with other conserved regulatory elements, including a TATA box and AP-1 binding sites, having already been identified in *H. sapien*, *M. musculus*, *G. gallus*, *S. scrofa* and *X. laevis* (Coulon et al., 2010). Further, intron 1 of *fos* is also bound by p53 and necessary for transactivation of *fos* by p53 (Elkeles et al., 1999). This prompts the novel and exciting proposal that CIC functions as a transcriptional repressor for an intronic promoter within *fos* intron 1. Indeed, should CIC be some way involved in suppressing IME of *fos*, perhaps this could explain the massive effect size and upregulation of *fos* seen following CIC knockdown and FGF4 overexpression (figure 5).

### 3.3.7 Co-operative transcriptional and post-translational regulation of Fos expression by dpERK

Several transcriptional regulators of *fos* have already been characterised including the serum response factor (SRF) and Elk-1 (Misra et al., 1991; Cavigelli et al., 1995). Like CIC, the transcriptional activator Elk-1 is also regulated through phosphorylation by dpERK (Gille et al., 1995; Cruzalegui, Cano and Treisman, 1999). Therefore, activated ERK may simultaneously act to promote *fos* expression through relief of CIC-mediated

repression, alongside activating Elk-1 which promotes *fos* expression. It is interesting to speculate whether CIC and Elk-1 may act synergistically to drive the rapid immediate early gene expression of *fos*. Importantly, the regulation of both CIC and Elk-1 by dpERK makes it difficult to discern whether changes in *fos* expression following ERK inhibition (figure 9) reflect altered activity of Elk-1, CIC, or both transcriptional regulators.

In addition to the putative regulation of *fos* by dpERK at the transcriptional level through Elk-1 and CIC, dpERK also post-translationally regulates the stability of Fos through phosphorylation (Okazaki and Sagata, 1995; Chen et al., 1996; Murphy et al., 2002). Perhaps, the tight coupling of Fos expression to dpERK facilitates rapid on/off dynamics for Fos activity as the activity of dpERK itself can be rapidly controlled through post-translational phosphorylation. Indeed, as Fos is largely stress responsive, rapid on/off activity is crucial.

### 3.3.8 Summary

As 44 genes have been identified which are putatively FGF target genes regulated by CIC (figure 6, additional file S4), this supports the hypothesis that ERK relieves CIC-mediated repression downstream of FGF signalling. In further agreement with CIC acting downstream of FGF signalling, gene-ontology terms enriched following CIC knockdown have relevance to FGF signalling (figure 7, table 4), and CIC appears to regulate a number of genes involved in an FGF signalling-induced negative regulatory feedback loop. One candidate CIC/FGF target gene *fos*, has been validated to show FGF and ERK-dependent expression in gastrula stage *X. tropicalis* embryos (figure 9). However, further work is necessary to establish whether reduced *fos* expression following FGF and ERK signalling inhibition is directly due to CIC continuing to elicit repression, and not through other means, for example loss of Elk-1 induction. CIC binding motifs were found within a highly conserved region of *fos* intron 1 corresponding to an putative intronic promoter (figure 10, table 6), and in support of the hypothesis, this suggests alongside being a target of FGF signalling, *fos* may be a direct and conserved target of CIC.

# Chapter 4: The Regulation of *fos* During Wound and Morphogenesis

## 4.1 Introduction

### 4.1.1 The imperfect nature of adult wound healing

For most adults, minor wounds, such as a graze or superficial cut, pose little threat due to the body's wound healing response. Enacted following cellular injury, the response involves the directly damaged cell(s), and both proximal and systemically recruited populations of healthy cells (Eming, Martin and Tomic-Canic, 2014). The response is coordinated through a complex cell signalling network which ultimately elicits both structural changes, and changes in gene expression, at the wound site. Despite the obvious importance of wound healing, the response can be considered somewhat imperfect and is subject to dysfunction. The prevalence of chronic wounds in the elderly, and hypertrophic and keloid scarring, provide clear examples of where therapeutic intervention may be necessary to overcome inadequacies in a patient's wound healing response (Velnar, Bailey and Smrkolj, 2009).

### 4.1.2 Embryonic wound healing – the gold standard of wound healing

In the hope of highlighting potential therapeutic targets able to improve adult wound healing, much emphasis has been placed on understanding embryonic wound healing. This is largely driven by the knowledge that embryos, until a certain developmental stage, have the capacity to perform rapid scar-free healing (Lorenz et al., 1993; Colwell et al., 2006; Cass et al., 1997). Therefore, embryos provide a paradigm of an idealist wound healing response. Akin to human foetuses, which lose the ability to heal scarlessly after approximately 24 weeks of gestation (Larson, Longaker and Lorenz, 2010), as *Xenopus* froglets mature their ability to heal cutaneous wounds scarlessly also diminishes (Bertolotti, Malagoli and Franchini, 2013). Accompanied by their macroscopic egg size, *ex utero* development, and genetic amenability, this makes *Xenopus* an attractive model to study embryonic wound healing (Blum and Ott, 2018).

### 4.1.3 The early phase of embryonic wound healing

Embryonic wound healing is typically considered as two main phases: an initial fast contraction of the wound diameter followed by a slower, final wound closure. In the initial phase of wound healing, rapid assembly of an actin-myosin II ring in peripheral

cells around the wound drives closure of over two-thirds of the wound (Martin and Lewis, 1992; Davidson, Ezin and Keller, 2002; Yoshii et al., 2005; Wood et al., 2002). The rate of wound closure is correlated with wound size, with smaller wounds overall healing at a much faster rate retrospectively to larger wounds (Davidson, Ezin and Keller, 2002). In *Xenopus* embryos wounded by a transverse bisection, approximately 80% of the wound diameter can be closed within 30 minutes (Yoshii et al., 2005). As parallels can be drawn between the movements of cells during this initial wound closure and the visual appearance of pulling on a string to close a purse, the movement has been widely termed a 'purse-string' contraction (Martin and Lewis, 1992). Assembly of the actin-myosin II ring, and subsequent contraction, is mediated by dpERK through activation of the small GTPase Rho which is necessary for myosin II activation (Li et al., 2013; Abreu-Blanco, Verboon and Parkhurst, 2014). The mechanism of ERK activation itself during wound healing isn't completely clear. Recent work within the Isaacs lab has demonstrated that wounding embryos in presence of calcium chelating agents and calcium-free media attenuates ERK activity, thus suggesting calcium signalling may be in part responsible for ERK activation (personal communication with Dr. Harry Isaacs).

#### 4.1.4 The late phase of embryonic wound healing

The second phase of embryonic wound healing restores the integrity of the epithelium and includes *de novo* induction of healing-specific gene expression. Complete closure of the wound is achieved via a 'zippering' process in which filopodia and lamellipodia protrusions from ectodermal cells at the wound perimeter reseal opposing epithelial sheets (Davidson, Ezin and Keller, 2002; Wood et al., 2002). For *Xenopus* embryos wounded by a transverse bisection, complete wound closure is achieved within 3 to 6 hours post-wounding (Yoshii et al., 2005). Whilst ERK signalling drives assembly of the actin-myosin II ring during the early phase of wound closure, PI3K signalling is responsible for assembly of the filopodia and lamellipodia actin-based protrusions through activation of Rac and Cdc42 in the later phase of wound closure (Li et al., 2013; Wood et al., 2002). During the early phase of wound closure ERK is able to suppress PI3K signalling, with attenuation of ERK signalling then enabling PI3K signalling to commence (Li et al., 2013). Notably, a further even later phase in embryonic wound healing has also been recently described, beyond the visible closure of the wound, and is thought to involve tissue remodelling - likely through secretion of a specific subset of matrix metalloproteinases (MMPs) (Abaffy et al., 2019).

#### 4.1.5 Wound-induced changes in gene expression are relatively understudied

Due to the rapidity of the structural changes which occur during embryonic wound healing, the relevance of *de novo* transcription has largely been overlooked. Instead, much research has focused on post-translational mechanisms, relatively quicker than *de novo* protein synthesis, that may be driving wound healing. Indeed, subjecting embryos to  $\alpha$ -amanitin treatment to inhibit transcription has been shown to have little visible effects on wound closure (Li et al., 2013). Nonetheless, it is evident that changes in gene expression do occur following embryonic wounding (Abaffy et al., 2019; Ding et al., 2017). Indeed, nitric oxide putatively represents one key regulator of differential gene expression associated with embryonic wound healing, as inhibiting nitric oxide production alters wound-induced patterns of gene expression (Abaffy et al., 2019).

#### 4.1.6 Transcription of the immediate early gene *fos* is induced rapidly in response to wounding

Immediate early genes (IEGs), often stress responsive genes, are genes which can be rapidly transcribed and translated post-stimulation (Bahrami and Drabløs, 2016). Rapid transcription has been attributed to a bivalent epigenetic landscape, and strong RNA Pol II association at promoter regions (Bahrami and Drabløs, 2016). Transcription of IEG mRNA is undisturbed in the presence of protein synthesis inhibitors as the protein components necessary for IEG expression already exist within the cell (Bahrami and Drabløs, 2016). Due to the independence of IEG mRNA expression on *de novo* protein synthesis, it's speculated that  $\alpha$ -amanitin treatment may affect IEG induction to a lesser extent than other genes. Transcripts of the IEG *fos* can be detected rapidly after wounding (Greenberg and Ziff, 1984; Abaffy et al., 2019; Ding et al., 2017); recent work within the Isaacs lab found *fos* mRNA to be detected as early as 10 minutes post-wounding in *X. laevis* embryos (Cowell, 2019).

#### 4.1.7 Candidate regulation of *fos* by CIC during embryonic wound healing and neurulation

As the transcriptional repressor CIC is regulated by dpERK, and *fos* is a putative target gene of CIC (see chapter 3), this by extension postulates *fos* as a downstream target of dpERK. In the context of wound healing, this axis of regulation is attractive as recent work within the Isaacs lab has demonstrated activated ERK and *fos* mRNA spatially

coincide at the wound site (Cowell, 2019), and Myc-tagged CIC is degraded upon wounding (*manuscript in preparation*, King et al. (2022)).

On numerous occasions parallels have been drawn between wound healing and embryonic morphogenesis (Martin et al., 1994; Wood et al., 2002; Nodder and Martin, 1997). Neurulation is the morphogenetic process by which the neural plate folds into the neural tube (Lowery and Sive, 2004). Of particular interest is that whilst *fos* mRNA and dpERK coincide in the context of wound healing, recent work from the Isaacs lab suggests that this is also the case during neurulation (Cowell, 2019). Therefore, it is also enticing to propose that dpERK might similarly regulate *fos* expression through CIC in the context of both wound healing and neurulation.

#### 4.1.8 Chapter hypothesis and aims

Hypothesis:

ERK relieves CIC-mediated repression of *fos* during both embryonic wound healing and neurulation in *Xenopus*.

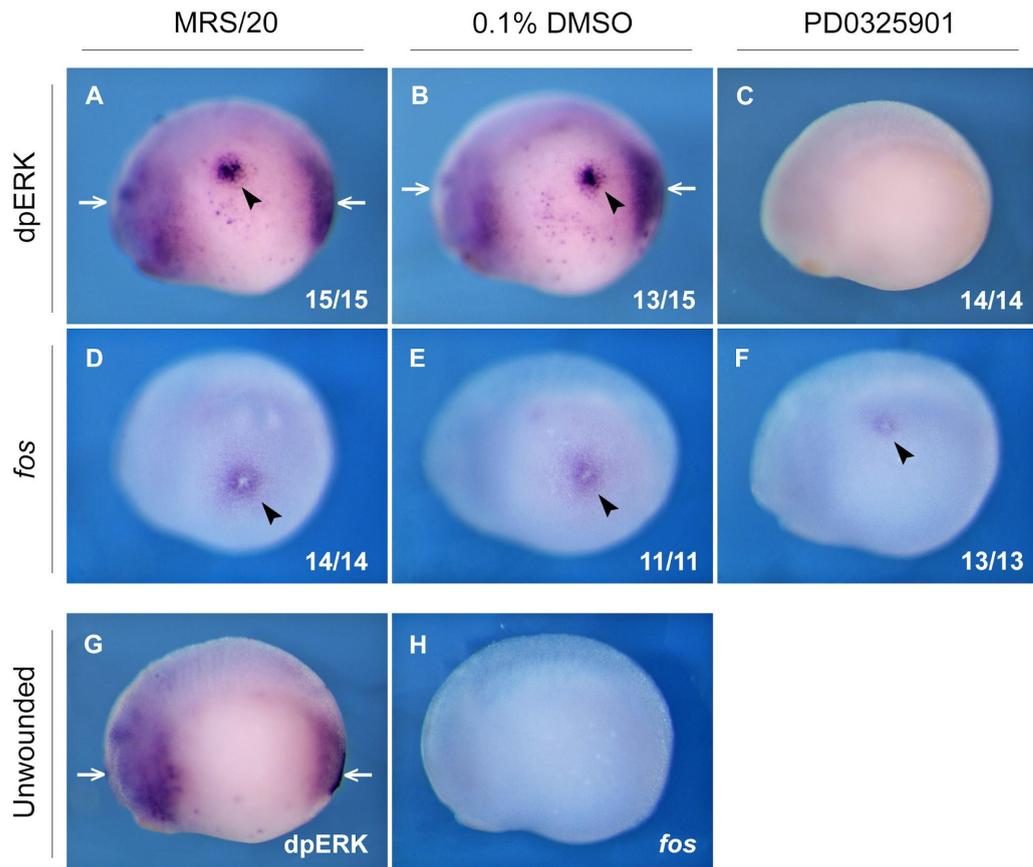
Aims of this chapter:

- Investigate the effects of ERK inhibition on *fos* expression during *X. tropicalis* embryonic wound healing and neurulation.
- Assess the ability of anti-Fos 2G9 antibody to detect *Xenopus* Fos protein.
- Investigate the effects of inhibiting nitric oxide production on dpERK and Fos protein expression during *X. laevis* embryonic wound healing.
- Identify other candidate targets of CIC during embryonic wound healing, validate these as wound responsive genes, and investigate their expression during neurulation.

## 4.2 Results

### 4.2.1 Inhibition of MAPK signalling reduces *fos* expression during *X. tropicalis* wound healing

To determine whether expression of *fos* during embryonic wound healing is downstream of ERK activation as predicted by the hypothesis, *X. tropicalis* embryos were wounded in the presence of the MEK inhibitor PD0325901 (Barrett et al., 2008). Sibling embryos were cultured for a minimum of 2 hours in either MRS/20, MRS/20 containing 0.1% DMSO or 25  $\mu$ M PD0325901 in 0.1% DMSO. At early tailbud stage 20, embryos were wounded centrally on the flank with a tungsten needle and subsequently allowed to heal for 30 minutes before being fixed in MEMFA. Previous studies within the Isaacs lab in *X. laevis* indicated both dpERK and *fos* expression are readily detectable at 30 minutes post-wounding (Cowell, 2019). Immunostaining for dpERK confirmed PD0325901 successfully inhibited dpERK expression 30 minutes post-wounding (figure 11A-C,G). Alongside loss of expression at the wound site, treatment with the MEK inhibitor also prevented endogenous anterior and posterior domains of ERK activation. On the contrary, *fos* expression at the wound site was still detectable via *in situ* hybridisation 30-minutes post wounding in the presence of PD0325901 (figure 11D-F,H). However, relative to culture in MRS/20 and 0.1% DMSO, treatment with PD0325901 does appear to somewhat reduce *fos* expression at the wound site. This implies dpERK is at least partly involved in *fos* induction during embryonic wound healing.

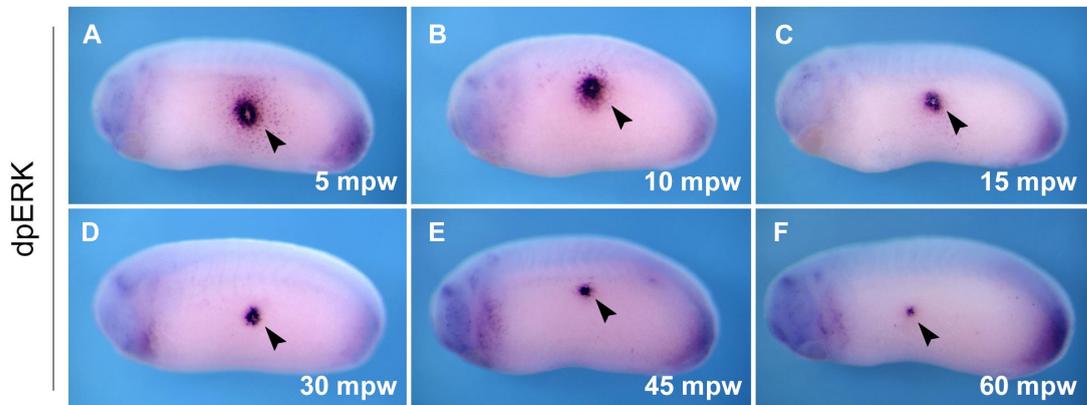


**Figure 11: PD0325901 inhibits dpERK expression and reduces *fos* expression 30 minutes post-wounding in *X. tropicalis* embryos, as assayed via immunostaining and *in situ* hybridisation, respectively.** At stage 15, *X. tropicalis* sibling embryos were cultured for a minimum of 2 hours in either MRS/20 (A,D,G,H), MRS/20 containing 0.1% DMSO (B,E) or 25  $\mu$ M PD0352901 in 0.1% DMSO (C,F). At approximately stage 20, embryos were wounded centrally on the flank, and allowed to heal for 30 minutes, before being fixed in MEMFA for subsequent immunostaining with anti-dpERK (M9692) (A-C, G) and *fos in situ* hybridisation (D-F,H). Unwounded embryos are displayed for comparison (G,H). The number of embryos observed with the displayed phenotype is indicated in the bottom right hand corner (A-F). Black arrowheads indicate observable expression at the wound site, white arrows point to endogenous dpERK expression. Images not to scale.

#### 4.2.2 The nitric oxide synthesis inhibitor TRIM heightens dpERK expression during *X. laevis* wound healing

Abaffy et al. (2019) put forward a role for nitric oxide in regulating *fos* expression during *X. laevis* embryonic wound healing. To investigate whether nitric oxide may also regulate ERK, the normal expression pattern of dpERK over a wounding time-course was first explored. At approximately stage 24, *X. laevis* embryos cultured in NAM/10 were wounded and allowed to heal for a designated length of time before being fixed in MEMFA for subsequent immunostaining. The induction of dpERK at the wound site was rapid (detectable at 5 minutes post-wounding), with maximal expression occurring

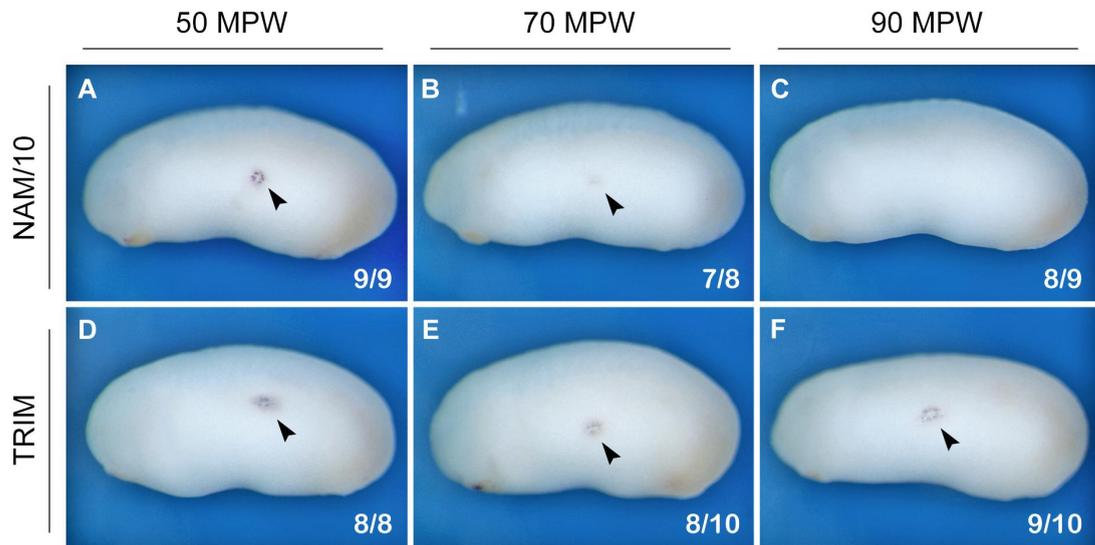
somewhere between 5-15 minutes post-wounding (figure 12). Within 10-15 minutes post-wounding, dpERK expression appears to begin to decrease, and continues to decrease up until the final investigated time point (60 minutes post-wounding) (figure 12).



**Figure 12: Wounding time-course of dpERK expression in *X. laevis* embryos, analysed via immunostaining.** At approximately stage 24 *X. laevis* embryos cultured in NAM/10 were wounded centrally on the flank and allowed to heal for either 5 minutes (**A**), 10 minutes (**B**), 15 minutes (**C**), 30 minutes (**D**), 45 minutes (**E**) or 60 minutes (**F**) before being fixed in MEMFA for immunostaining with anti-dpERK (M9692). Black arrowhead indicates expression at the wound site. Images are representative of n=10 embryos. mpw; minutes post-wounding.

As inhibiting nitric oxide production has been demonstrated to increase *fos* expression post-wounding (Abaffy et al., 2019), and inhibiting ERK activation appears to decrease *fos* expression (figure 11), it was hypothesised nitric oxide may negatively regulate dpERK expression during embryonic wound healing. To investigate this, *X. laevis* embryos were cultured in the presence of the nitric oxide synthase inhibitor 1-(2-Trifluoromethylphenyl) imidazole (TRIM). It has previously been demonstrated that 2 mM TRIM is sufficient to inhibit nitric oxide production in *X. laevis* (Abaffy et al., 2019). *X. laevis* embryos, with the vitelline membrane removed prior, were cultured for 1 hour in NAM/10 alone or 2 mM TRIM in NAM/10. At late tailbud stage 24, embryos were wounded centrally on the flank with a tungsten needle and subsequently allowed to heal for 50, 70 or 90 minutes before being fixed in MEMFA for subsequent immunostaining. The time points assayed were chosen primarily for two reasons: firstly, between 60 and 90 minutes post-wounding reflects the period during which Abaffy et al. (2019) found TRIM treatment to increase *fos* expression; secondly, as dpERK expression is relatively low at 60 minutes post-wounding (figure 12F) any increase in dpERK expression as a consequence of TRIM treatment may be more readily detectable. Indeed, *X. laevis* embryos treated with TRIM show increased dpERK expression at the wound site most notably at 90 minutes post-wounding (figure 13C,F).

On the contrary, at 50 minutes post-wounding an obvious difference between embryos cultured in NAM or TRIM is not apparent (figure 13,A,D). The increased dpERK expression seen at 90 minutes post-wounding following TRIM treatment is concurrent with the increased *fos* expression found in Abaffey et al. (2019), and is hence supportive of dpERK regulating *fos* expression during embryonic wound healing.

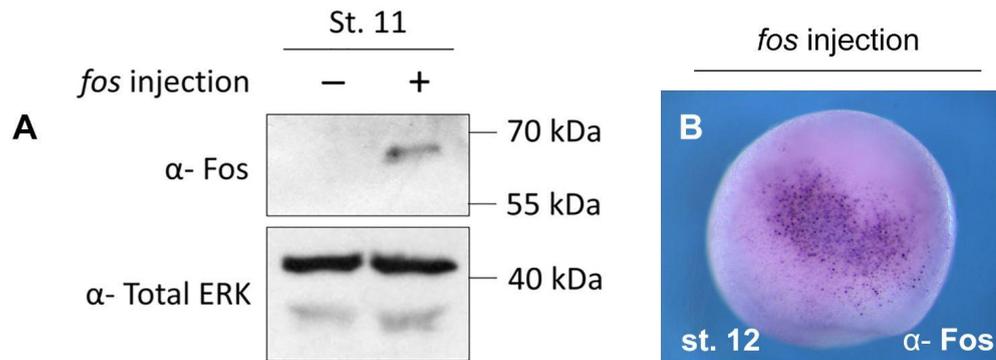


**Figure 13: Expression of dpERK in *X. laevis* embryos at 50, 70 and 90 minutes post-wounding following culture in either NAM/10 or TRIM, as assayed via immunostaining.** At approximately stage 24, *X. laevis* embryos with the vitelline membrane removed were cultured for 1 hour in either NAM/10 (A,B,C) or 2 mM TRIM in NAM/10 (D,E,F). Embryos were then wounded centrally on the flank and allowed to heal for either 50 minutes (A,D), 70 minutes (B,E), or 90 minutes (C,F), before being fixed in MEMFA for subsequent immunostaining with anti-dpERK (M9692). The number of embryos observed with the displayed phenotype is indicated in the bottom right hand corner. Black arrowhead indicates observable expression at the wound site. MPW; minutes post-wounding. Images not to scale.

#### 4.2.3 Detection of Fos by anti-Fos (2G9) antibody

Analyses so far have concerned expression of *fos* mRNA, due to the fast turn-over and short half-life of *fos* mRNA it is important to assess whether *fos* mRNA expression directly translates into Fos protein expression (Shyu, Greenberg and Belasco, 1989). To confirm the specificity of 2G9 antibody (Merck) against *Xenopus* Fos, *X. laevis* embryos were injected with 4 ng of *X. tropicalis fos* mRNA at the one-cell stage. At stage 11, five injected and five sibling control uninjected embryos were flash frozen and collected for Western blotting. Western blot analysis with the anti-Fos 2G9 antibody detected a protein between 55 and 70 kDa in injected but not uninjected samples (figure 14A). This is similar in size to Fos detected in other *Xenopus* samples using the same antibody (Leclerc, Duprat and Moreau, 1999).

At the 2-cell stage, *X. laevis* embryos were also unilaterally injected with 2 ng *X. tropicalis fos* mRNA, and subsequently fixed at stage 12 in MEMFA. Immunostaining with anti-Fos 2G9 on unilaterally injected embryos was able to detect ectopic patches of Fos expression (figure 14B). Henceforth, both the Western blot analysis and immunostaining testify the ability of anti-Fos 2G9 to detect overexpressed Fos as translated from *X. tropicalis fos* mRNA.

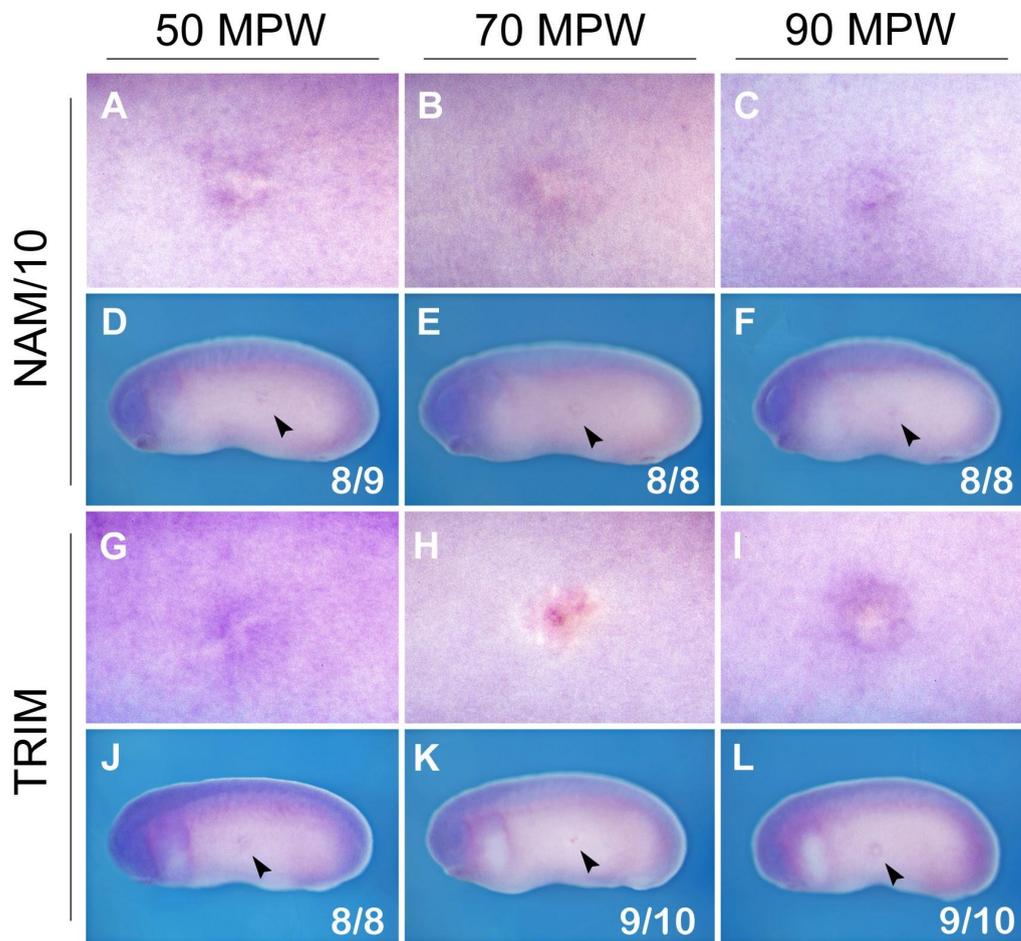


**Figure 14: Anti-Fos 2G9 can detect overexpressed exogenous *X. tropicalis* Fos via Western blotting and immunostaining in *X. laevis* embryos. (A)** Western blot analysis of *X. laevis* embryos injected at the one-cell stage with 4 ng mRNA encoding *X. tropicalis fos* alongside control uninjected sibling embryos. Embryos were flash frozen on dry ice at stage 11. Membranes were probed with anti-Fos (2G9), and anti-MAP Kinase (ERK1 and ERK2, M5670) as a loading control. Each band is representative of the protein content from approximately one embryo. **(B)** Immunostaining with anti-Fos 29G for *X. laevis* embryos unilaterally injected at the 2-cell stage with 2 ng *X. tropicalis fos* mRNA. At stage 12 embryos were fixed in MEMFA. St.; stage.

#### 4.2.4 Treatment with nitric oxide synthesis inhibitor TRIM and the expression of Fos protein during *X. laevis* wound healing

Abaffy et al. (2019) demonstrated that inhibiting nitric oxide production is able to reverse the decrease in *fos* expression seen between 60 and 90 minutes post-wounding in *X. laevis* embryos, with instead *fos* expression being higher at 90 minutes post-wounding than at 60 minutes. It was predicted that this increase in *fos* expression following TRIM treatment would translate into increased Fos protein expression. To extend the analysis undertaken by Abaffy et al. (2019), *X. laevis* embryos, with the vitelline membrane removed prior, were cultured for 1 hour in NAM/10 or NAM/10 containing 2 mM TRIM. At late tailbud stage 24, embryos were wounded centrally on the flank with a tungsten needle and subsequently allowed to heal for 50, 70 or 90 minutes before being fixed in MEMFA for subsequent immunostaining with anti-Fos 2G9. For both treatments, and at all time-points, endogenous Fos protein appeared detectable at the wound site (figure 15).

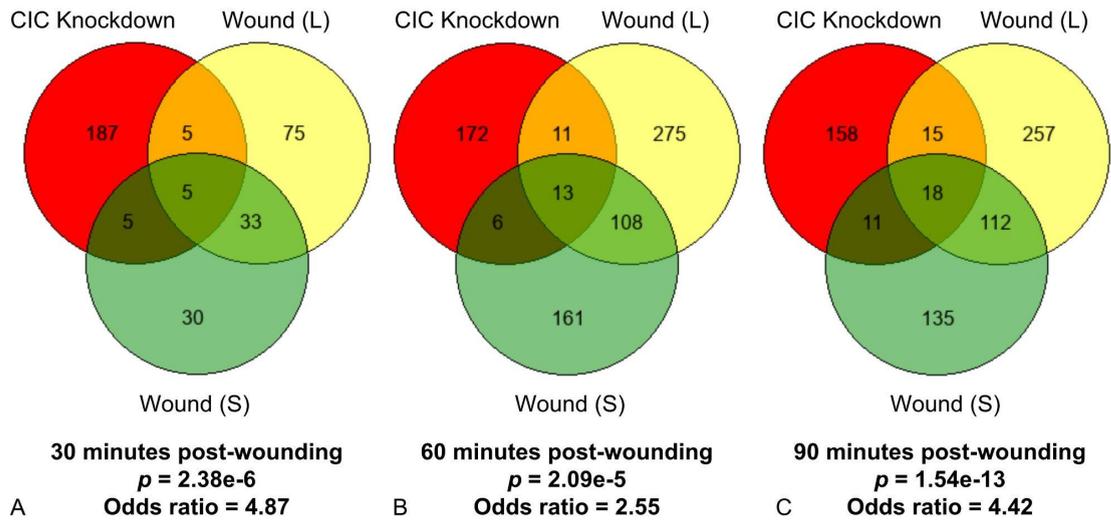
Interestingly, Fos protein expression does appear slightly heightened at 90 minutes post-wounding for TRIM treated embryos relative to NAM/10 control embryos (figure 15C,F,I,L). However, quantification of expression at the wound site across multiple technical and biological replicates is required to eliminate any subjectivity and bias.



**Figure 15: Expression of Fos protein in *X. laevis* embryos cultured in NAM/10 or TRIM as assayed by immunostaining with anti-Fos 2G9.** At approximately stage 24, *X. laevis* embryos with the vitelline membrane removed were cultured for 1 hour in either NAM/10 (**A-F**) or 2 mM TRIM in NAM/10 (**G-L**). Embryos were then wounded centrally on the flank and allowed to heal for either 50 minutes (**A,D,G,J**), 70 minutes (**B,E,H,K**) or 90 minutes (**C,F,I,L**) before being fixed in MEMFA for subsequent immunostaining with anti-Fos 2G9. The number of embryos observed with the displayed phenotype is indicated in the bottom right hand corner of whole embryo images (**D-F, J-L**). Higher magnification images of the wound site for the displayed whole embryo images are shown (**A-C,J-K**). Black arrowheads indicate observable expression at the wound site. MPW; minutes post-wounding.

#### 4.2.5 A wider role for CIC during embryonic wound healing

To uncover whether CIC may play a wider role in regulating gene expression during wound healing, genes which were significantly upregulated following CIC knockdown in *X. tropicalis* embryos were compared to those significantly upregulated at 30, 60 and 90 minutes post-wounding in *X. laevis* embryos (figure 16) (Abaffy et al., 2019). To overcome differences in annotation of the *X. tropicalis* and *X. laevis* genomes, the RNA-seq data was first reduced to lists of 8,713 genes with common annotation in both organisms and across both *X. laevis* subgenomes. Furthermore, the *X. laevis* S and L subgenomes were evaluated independently, as this enabled identification of genes upregulated from both S and L subgenomes and may postulate a gene dosage mechanism. Significantly upregulated genes following CIC knockdown were defined as those with a  $q$ -value  $< 0.1$  and effect size  $> 1.5$  (additional file S9). As dictated in Abaffy et al. (2019), significantly upregulated genes following wounding were defined as those with a  $q$ -value  $< 0.1$  and a fold-change  $> 1.5$  (additional file S9). Including *fos*, 51 genes were found to be significantly upregulated following both CIC knockdown and wounding (table 7). The number of genes upregulated by both CIC knockdown and wounding becomes larger as time post-wounding increases, and it is interesting to speculate that as *fos*, *fosl1*, *jun*, *jund* and *atf3* appear upregulated early (30 mins post-wounding) the increasing number of genes at later time points may reflect expression of AP-1 target genes.

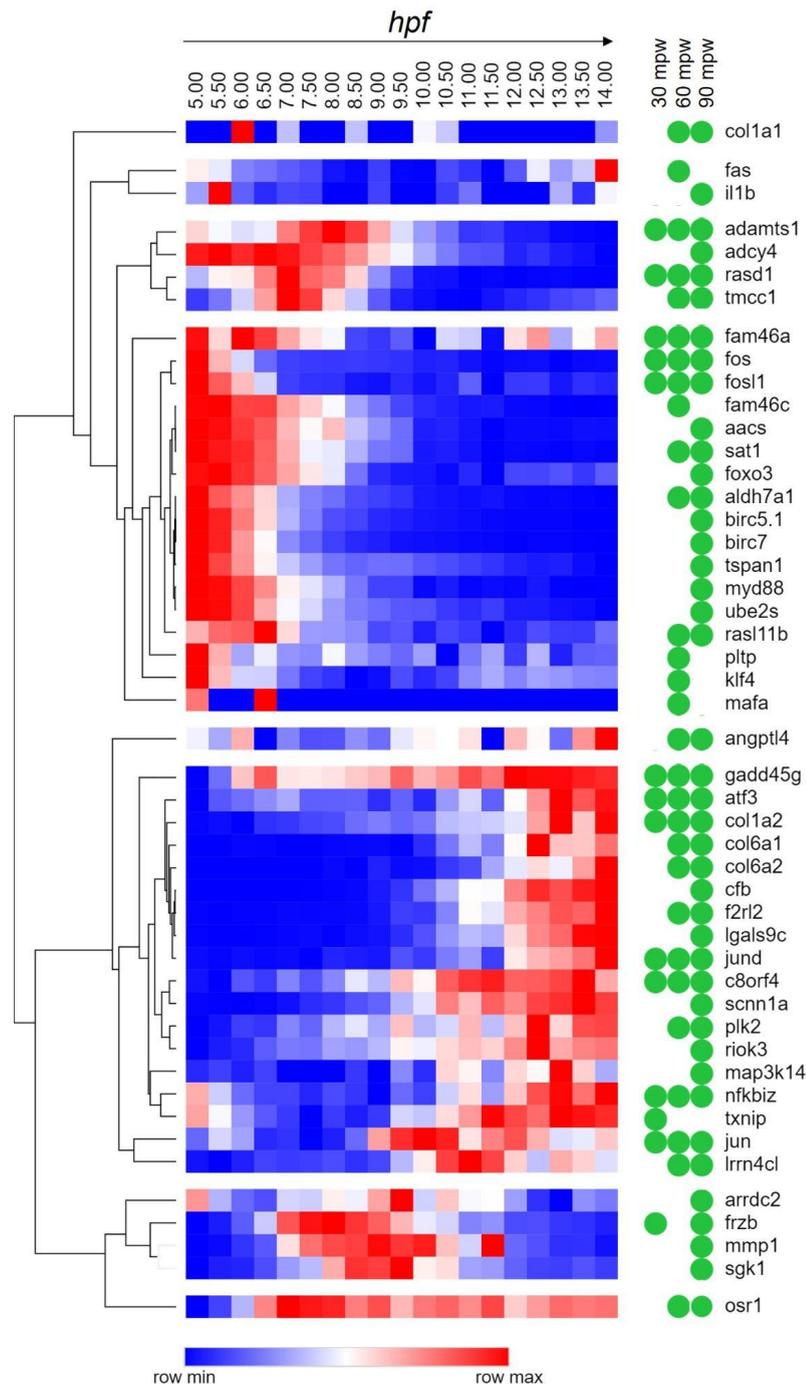


**Figure 16: Venn diagram of overlap between genes significantly upregulated following CIC knockdown, and genes significantly upregulated at 30, 60 and 90 minutes post-wounding in *Xenopus*.** CIC knockdown significantly upregulated: RNA-seq  $q$ -value < 0.1, effect size > 1.5. Wound (S) and Wound (L) significantly upregulated: RNA-seq  $q$ -value < 0.1, fold change > 1.5. Overlap between genes significantly upregulated at **(A)** 30 minutes post-wounding **(B)** 60 minutes post-wounding **(C)** 90 minutes post-wounding. There is a significant intersection between genes upregulated by wounding (S + L combined) and CIC knockdown at 30 minutes post-wounding ( $p = 2.38e-6$ ), 60 minutes post-wounding ( $p = 2.09e-5$ ) and 90 minutes post-wounding ( $p = 1.54e-13$ ). Odds ratios for 30 minutes (4.87), 60 minutes (2.55) and 90 minutes (4.42) post-wounding are all > 1 and so suggest an association between genes significantly upregulated by CIC knockdown and genes significantly upregulated by wounding. Note that the CIC knockdown RNA-seq was conducted in *X. tropicalis* embryos, whilst the wound RNA-seq was conducted in *X. laevis* embryos by Abaffy et al. (2019) and hence the wounding data is split into S and L subgenomes.

**Table 7: Lists of genes significantly upregulated by both CIC knockdown in *X. tropicalis* embryos and wounding in *X. laevis* embryos at 30, 60 and 90 minutes post-wounding.** Data relates to the Venn diagrams in figure 16. Capicua knockdown significantly upregulated: RNA-seq  $q$ -value < 0.1, effect size > 1.5. Post-wound significantly upregulated: RNA-seq  $q$ -value < 0.1, fold change > 1.5 (additional file S9). ‘S’ represents genes only upregulated on the *X. laevis* S subgenome following wounding, ‘L’ represents genes only upregulated on the *X. laevis* L subgenome following wounding, ‘S + L’ represents genes upregulated from both S and L subgeomes following wounding. MPW; minutes post wounding. CIC KD; Capicua knockdown.

30 MPW and CIC KD			60 MPW and CIC KD			90 MPW and CIC KD		
S	L	S + L	S	L	S + L	S	L	S + L
fam46a	c8orf4	adamts1	f2r12	aldh7a1	adamts1	adcy4	aacs	adamts1
gadd45g	col1a2	atf3	fam46a	c8orf4	angptl4	birc5.1	aldh7a1	arrdc2
jund	frzb	fos	fam46c	col1a2	atf3	col6a1	angptl4	atf3
nfkbiz	rasd1	fosl1	jund	col6a2	col1a1	ecm1	birc7	c8orf4
txnip	usp2	jun	osr1	fas	col6a1	fam46a	cfb	col1a1
			sat1	lrrn4cl	fos	il1b	col6a2	col1a2
				plk2	fosl1	jund	foxo3	f2r12
				rasd1	gadd45g	osr1	frzb	fos
				rasl11b	jun	riok3	lgals9c	fosl1
				tmcc1	klf4	sgk1	myd88	gadd45g
				usp2	mafa	wee2	plk2	jun
					nfkbiz		rasl11b	lrrn4cl
					pltp		scnn1a	map3k14
							tspan1	mmp1
							ube2s	nfkbiz
								rasd1
								sat1
								tmcc1

To look for groups of genes which may be co-regulated, the expression patterns during *X. tropicalis* normal development of 48 of the 51 genes identified as upregulated by both CIC knockdown and following wounding were subject to hierarchical clustering (figure 17). Normal patterns of expression were analysed from 5 hours post fertilisation (hpf), with the aim of mitigating maternal transcripts and capturing the onset zygotic transcription at the mid-blastula transition, up until 14 hpf. This captures the expression patterns of the genes during the major early morphogenetic movements of gastrulation and neurulation. It was speculated that genes which cluster together via their normal developmental expression might similarly cluster via their expression following wounding. Data indicating at which time points post-wounding the genes showed significantly increased expression is displayed adjacent to their developmental expression, however it is not obvious whether a relationship exists between expression post-wounding and developmentally regulated expression.

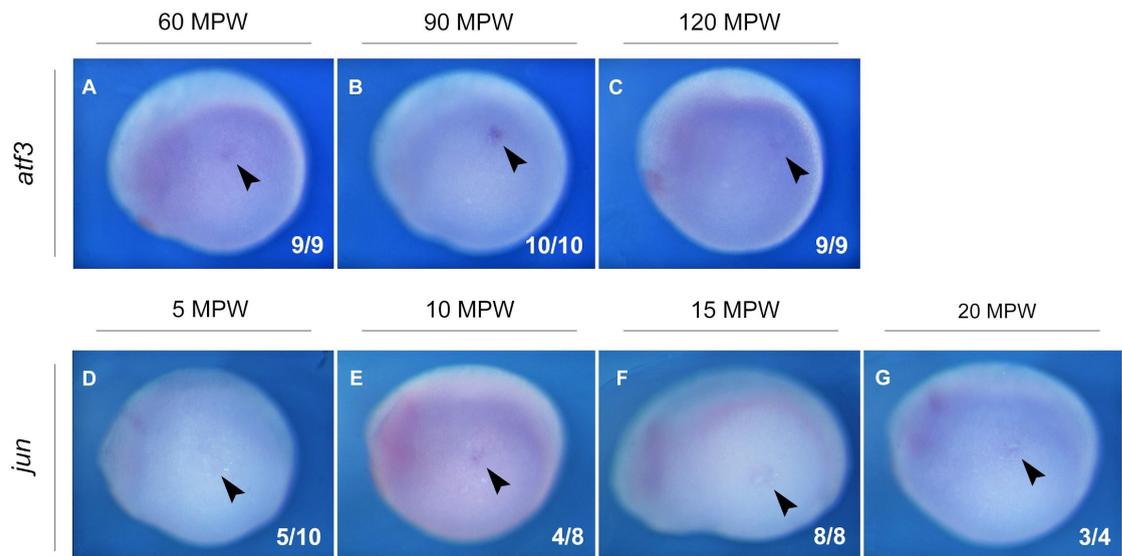


**Figure 17: Heat-map depicting expression patterns of 48/51 genes during development which are significantly upregulated following CIC knockdown and post-wounding.** The heat map was created using RNA-seq data from Owens et al. (2016) and Morpheus software (<https://software.broadinstitute.org/morpheus>). Red depicts higher levels of expression and blue lower levels of expression for each individual gene relative to itself (see colour bar), therefore expression levels are not directly comparable between genes. To the right of the heat map, expression of the genes at 30, 60 and 90 minutes post-wounding is indicated, with green corresponding to significant upregulation at that time point. Mpw; minutes post wounding, hpf; hours post fertilisation.

#### 4.2.6 Candidate regulation of AP-1 component genes by CIC during wound healing

Both *atf3* and *jun* are significantly upregulated following CIC knockdown (additional file S3), FGF4 overexpression (additional file S2), and from both *X. laevis* S and L subgenomes at 30, 60 and 90 minutes post-wounding (table 7, additional file S9) - this is identical to *fos*. Indeed, as *fos*, *jun* and *atf3* are all basic leucine zipper transcription factors, components of AP-1, and immediate early genes (Webster, Discher and Bishopric, 1994; Xie et al., 2005), it was postulated that *jun* and *atf3* may be expressed in a similar spatial and temporal manner to *fos* during wound healing.

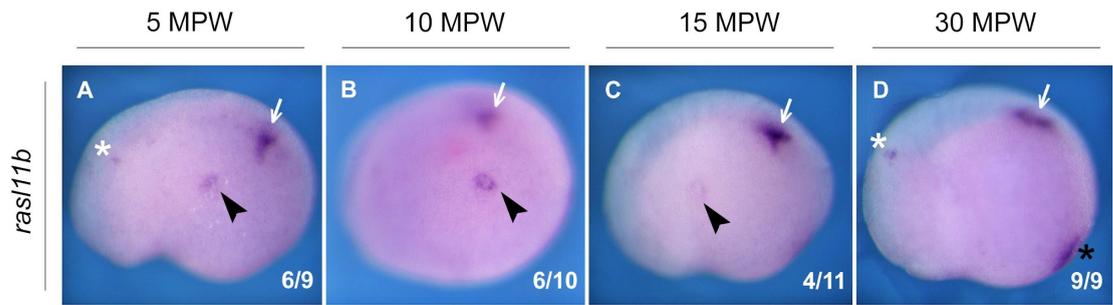
*In situ* hybridisations were undertaken on wounded embryos at time points ranging from 5-120 minutes post-wounding for *atf3* (figure 18A-C), *not all data shown*, and from 5-20 minutes post-wounding for *jun* (figure 18G-J). Both *atf3* and *jun* were detectable at the wound site via *in situ* hybridisation; *atf3* expression was maximally detected at 90 minutes post-wounding, and *jun* expression was weakly detectable at the wound site within the first 20 minutes post-wounding. For tested time points between 5-60 minutes post-wounding (*data not shown*), *atf3* expression was not detected. However, as the *atf3 in situ* hybridizations were left in the colour reaction for a suboptimal length of time, this may reflect poor detection opposed to absence of expression at these earlier time points.



**Figure 18: *In situ* hybridisation for *att3* and *jun* post-wounding in *X. tropicalis* embryos. (A-C) *In situ* hybridisations for *att3*. At approximately stage 20, *X. tropicalis* embryos cultured in MRS/20 were wounded centrally on the flank and allowed to heal for (A) 60 minutes, (B) 90 minutes or (C) 120 minutes, before being fixed in MEMFA. (D-G) *In situ* hybridisations for *jun*. At approximately stage 20, *X. tropicalis* embryos cultured in MRS/20 were wounded centrally on the flank and allowed to heal for (D) 5 minutes, (E) 10 minutes, (F) 15 minutes or (G) 20 minutes, before being fixed in MEMFA. Black arrowheads indicate expression at the wound site. The number of embryos observed with the displayed phenotype is indicated in the bottom right of the image.**

#### 4.2.7 *rasl11b* is induced rapidly post-wounding

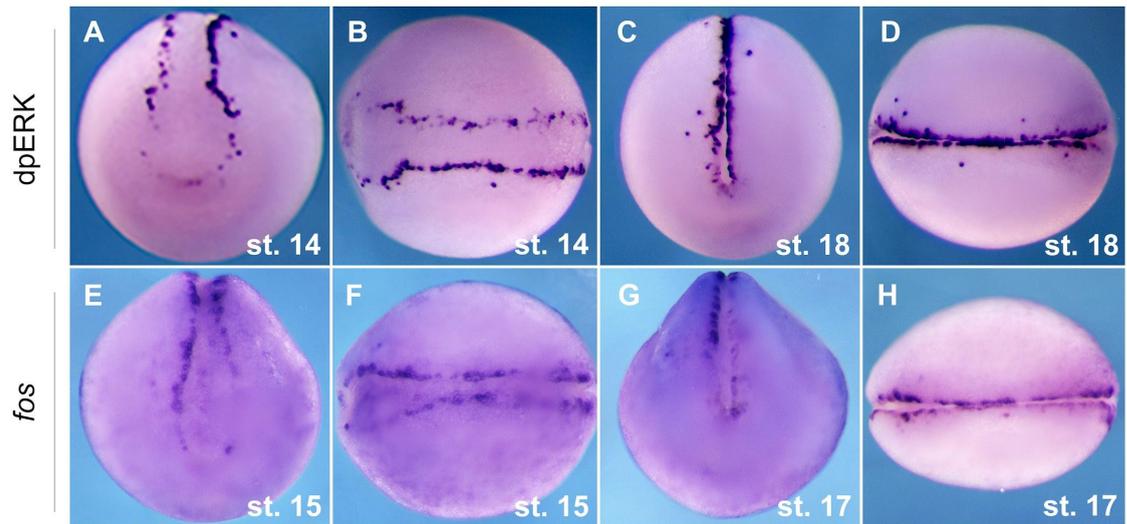
One gene clustering with *fos* via its developmental expression pattern and chosen for further analysis was *rasl11b* (figure 17); a gene significantly upregulated following FGF4 overexpression (additional file S2), CIC knockdown (additional file S3), and at 60 and 90 minutes post wounding (additional file S9). Interestingly, *in situ* hybridisation for *rasl11b* in *X. tropicalis* embryos detected *rasl11b* rapidly at the wound site, appearing at 5 minutes, and disappearing within 30 minutes (figure 19). Time points from 30-120 minutes post wounding were also analysed, however for these time points *rasl11b* was not readily detectable at the wound site (*data not shown*). Endogenous *rasl11b* expression is seen in the posterior presomitic mesoderm, the otic vesicle, and around the closed blastopore (figure 19). As a putative CIC target gene, the rapid induction of *rasl11b* post-wounding suggests CIC may play a wider role in regulating immediate early gene induction post-wounding beyond regulation of solely AP-1 component genes.



**Figure 19: *ras/11b* is induced rapidly post-wounding in early tailbud stage *X. tropicalis* embryos as analysed by *in situ* hybridisation.** *X. tropicalis* embryos, at approximately stage 20, were wounded in MRS/20, and allowed to heal for (A) 5 minutes, (B) 10 minutes, (C) 15 minutes or (D) 30 minutes, before being fixed in MEMFA. Following, *ras/11b* expression was assayed by *in situ* hybridisation. Black arrowheads indicate *ras/11b* expression at the wound site. White arrows and asterisk highlight regions of endogenous *ras/11b* expression in the posterior presomitic mesoderm and otic vesicles, respectively. Black asterisk indicates *ras/11b* expression around the closed blastopore. The number of embryos observed with the displayed phenotype is indicated in the bottom right of each panel. Images not to scale.

#### 4.2.8 Spatial and temporal co-localisation of *fos* and dpERK in neurula stage *X. tropicalis*

Due to observable similarities between morphogenesis and wound healing, it was hypothesised that *fos* expression during both wound healing and neurulation may be dependent on negative regulation of CIC by activated ERK. To investigate this, the spatial and temporal expression of dpERK and *fos* during neurulation were first assessed via immunostaining and *in situ* hybridisation, respectively (figure 20). Supportive of a role for dpERK regulating *fos* expression during neurulation, *fos* and dpERK co-localise along the edge of the dorsal neural closure.

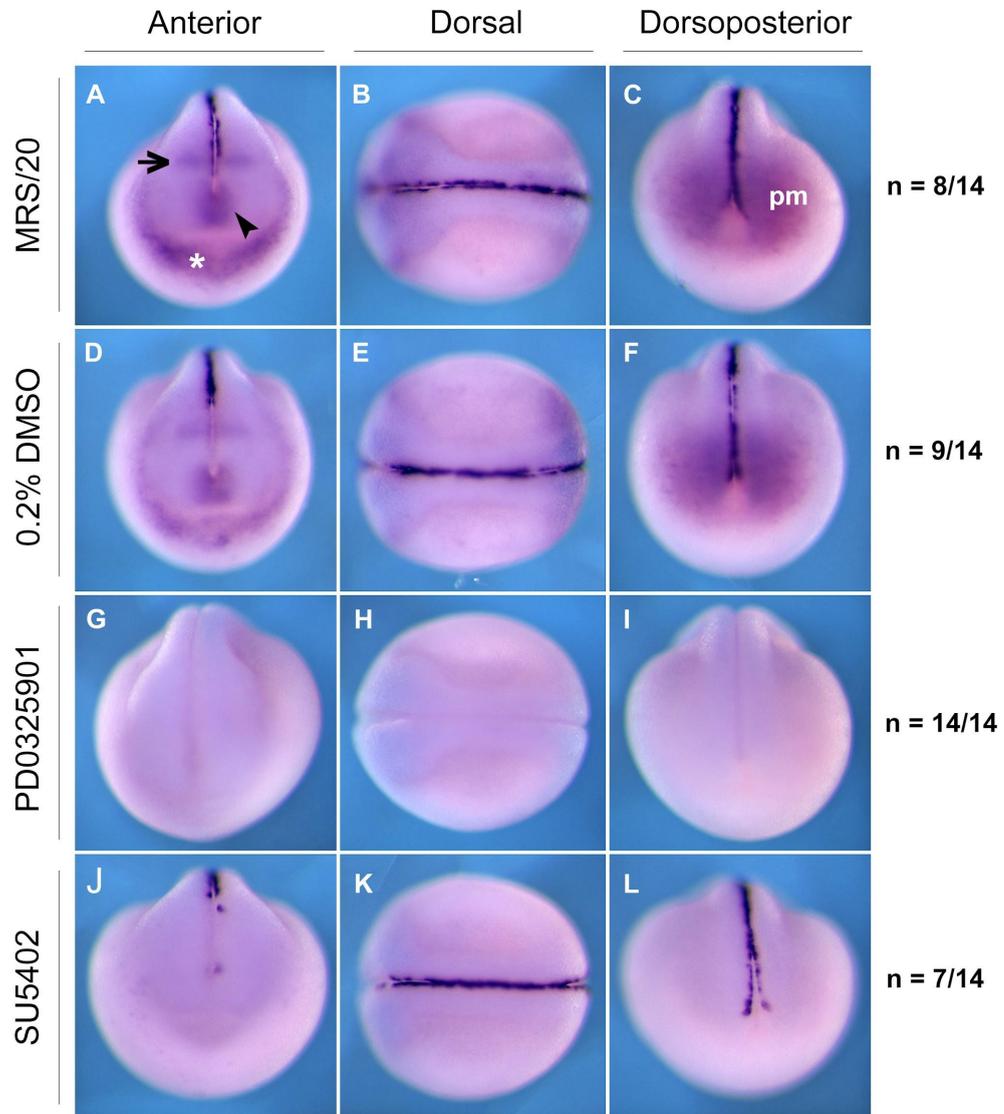


**Figure 20: *X. tropicalis* neurula stage series of dpERK and *fos* expression analysed via immunostaining and *in situ* hybridization, respectively. (A-D) dpERK immunostaining, (A) early neurula stage 14, anterior view, dorsal top (B) early neurula stage 14 dorsal view, anterior left (C) late neurula stage 18, anterior view, dorsal top (D) late neurula stage 18, dorsal view, anterior left. (E-H) *fos in situ* hybridisation (E) early neurula stage 15 anterior view, dorsal top (F) early neurula stage 15 dorsal view, anterior left (G) late neural stage 17 anterior view, dorsal top (H) late neurula stage 17 dorsal view, anterior left. Images are representative of n=10 embryos.**

#### 4.2.9 MEK but not FGF inhibition reduces *fos* expression along the dorsal neural closure in *X. tropicalis*

To determine whether dpERK induces *fos* expression during neurulation, the chemical inhibitors SU5402 and PD0325901 were again used to inhibit FGF signalling, and MAPK signalling, respectively (Mohammadi et al., 1997; Barrett et al., 2008). Early neurula stage 14-15 *X. tropicalis* embryos were cultured for a minimum of 2 hours in either MRS/20, MRS/20 containing 0.2% DMSO, 200  $\mu$ M SU5402 or 25  $\mu$ M PD0325901 diluted in MRS/20 containing 0.2% DMSO. At late neurula stage 18-19, embryos were collected and fixed in MEMFA. Subsequently, dpERK and *fos* expression were assayed via immunostaining and *in situ* hybridisation, respectively. For neurula stage embryos cultured in MRS/20, dpERK expression is seen in an anterior region of the neural plate corresponding to the midbrain-hindbrain boundary, a region corresponding to the forebrain, and also in the prospective gill region (figure 21A) (Christen and Slack, 1999). Dorsally, dpERK is present either side of the closing neural tube, marking the dorsal midline at late neural stages, and posteriorly dpERK expression can be seen in the posterior mesoderm (figure 21B,C) (Christen and Slack, 1999). Treatment with 25  $\mu$ M PD0325901 successfully inhibited all domains of dpERK activation, whilst treatment with 200  $\mu$ M SU5402 successfully inhibited anterior and

posterior domains of dpERK activity but dpERK expression along the dorsal neural closure appeared unaffected (figure 21G-L). This suggests that unlike other regions of dpERK expression in late neurula stage *X. tropicalis* embryos, dpERK expression along the dorsal neural closure is FGF-independent.



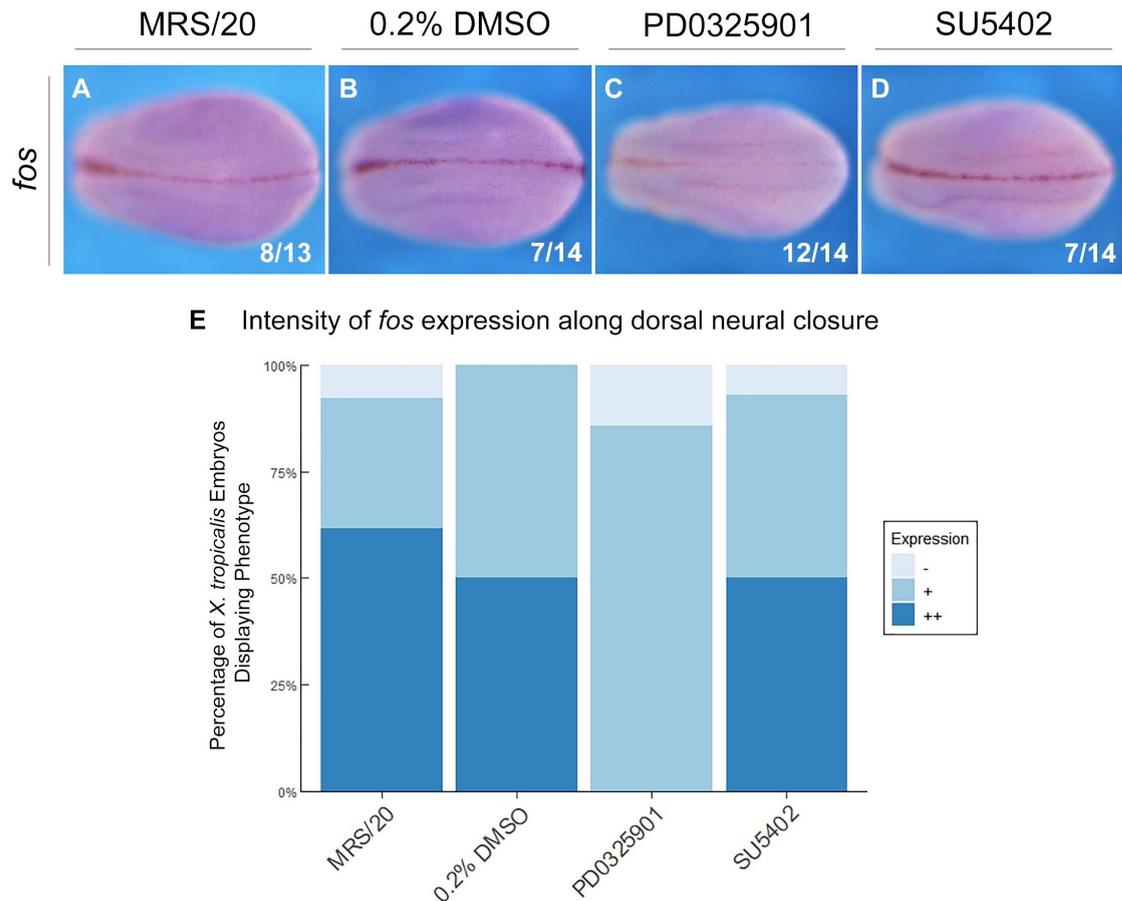
**Figure 21: dpERK immunostaining of late neurula stage *X. tropicalis* embryos treated with PD0325901 or SU5402.** *X. tropicalis* embryos were cultured from stage 15 in either (A-C) MRS/20 (D-F) MRS/20 containing 0.2% DMSO (G-I) 25  $\mu$ M PD0325901 in 0.2% DMSO (J-L) 200  $\mu$ M SU5402 in 0.2 % DMSO. At approximately stage 18 embryos were collected and fixed in MEMFA for immunostaining with anti-dpERK (M9692). (A,D,G,J) Anterior view, dorsal top (B,E,H,K) Dorsal view, anterior left (C,F I, L) Dorsoposterior view, dorsal top. Images in each row are taken from the same individual embryo and are not to scale. Black arrow indicates midbrain-hindbrain boundary in the neural plate, black arrowhead forebrain region, and white asterisk prospective gill region; pm, posterior mesoderm.

**Table 8: Phenotypic quantification of dpERK immunostaining along the dorsal neural closure and in anterior and posterior domains of late neurula stage *X. tropicalis* embryos.** For the dorsal neurula closure, ‘++’, ‘+’ and ‘-’ represent strong, weak and absence of detectable dpERK staining respectively. For anterior and posterior domains of dpERK expression ‘+’ and ‘-’ represent presence and absence of dpERK staining in both domains, respectively. Anterior and posterior domains of dpERK expression were not mutually exclusive and consistently concurrent. Dorsal neural closure staining, and anterior and posterior straining, were evaluated independently.

	Dorsal Neural Closure			Anterior and Posterior		<i>n</i> total
	++	+	-	+	-	
<b>MRS/20</b>	8	2	4	14	-	14
<b>0.2% DMSO</b>	9	2	2	13	-	13
<b>25 <math>\mu</math>M PD0325901</b>	-	-	14	-	14	14
<b>200 <math>\mu</math>M SU5402</b>	7	6	1	-	14	14

The expression of *fos* along the dorsal midline is comparable for MRS/20, 0.2% DMSO, and SU5402 treatments (figure 22A,B,D,E). On the contrary, treatment with PD0325901 appears to reduce the intensity of *fos* expression along the dorsal neural closure, with some residual expression still detectable (figure 22C,E). As the reduction but not complete loss of *fos* expression following PD0325901 treatment during neurulation mirrors the effects of PD0325901 treatment on *fos* expression post-wounding, this is supportive of similar underlying mechanisms governing embryonic wound healing and morphogenesis.

For both dpERK immunostaining and *fos in situ* hybridisation, there was considerable variation in the intensity and presence of staining along the dorsal neural closure. However, it was observed that despite embryos being siblings and fertilised concurrently, there was discrepancy in the stages of embryos subject to identical treatments. This natural variation in developmental progression is a caveat of working with living organisms, and could account for the variation in dpERK and *fos* expression along the dorsal neural closure for embryos subject to the same treatment. Indeed, embryos which did not match the dominant phenotype did frequently appear to be visibly at a different developmental stage (*observation, data not shown*).

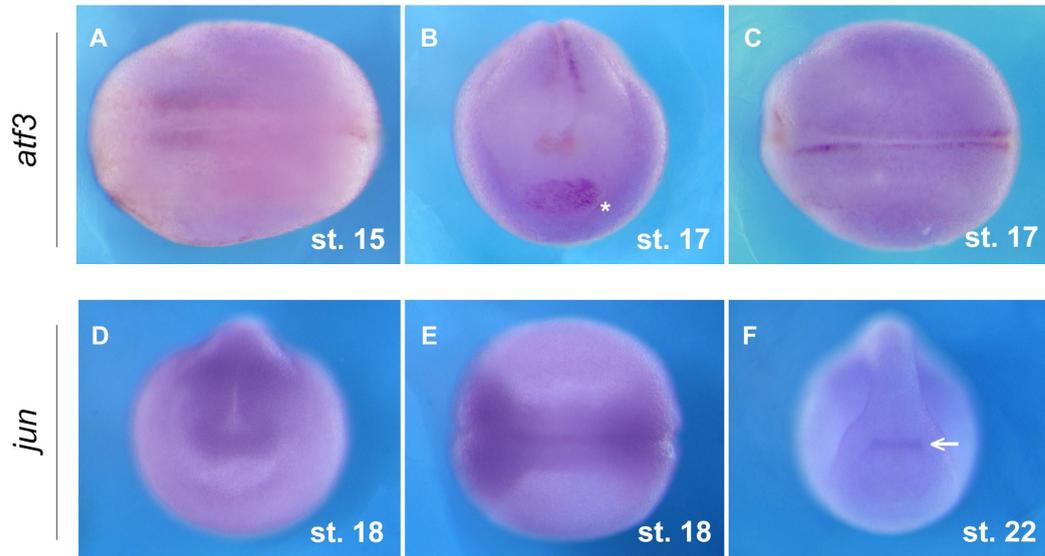


**Figure 22: *In situ* hybridisations for *fos* in late neurula stage *X. tropicalis* embryos treated with PD0325901 or SU5402.** *X. tropicalis* embryos with the vitelline membrane removed were cultured from stage 15 in either **(A)** MRS/20 **(B)** MRS/20 containing 0.2% DMSO **(C)** 25  $\mu$ M PD0325901 in 0.2% DMSO or **(D)** 200  $\mu$ M SU5402 in 0.2 % DMSO. At approximately stage 18/19, embryos were collected and fixed in MEMFA for *fos in situ* hybridisation. **(A,B,C,D)** Dorsal view, anterior left, images shown represent the dominant phenotype, or strong expression phenotype for 0.2% DMSO and SU5402, with the number of embryos observed indicated in the bottom right hand corner. **(E)** Percentage of embryos assigned to have strong ‘++’ /darkest blue, weak ‘+’ /mid-blue, or absence of ‘-’ /lightest blue, detectable *fos* expression along the dorsal neural closure. Note that embryos in figure (22) are not sibling embryos with those in figure (21).

#### 4.2.10 AP-1 component gene expression during neurulation

To investigate whether *atf3* and *jun* expression also mirror *fos* expression during neurulation, *in situ* hybridisations were performed on neural stage *X. tropicalis* embryos. In early neurula stage 15 embryos, regions of *atf3* expression were detectable on either side of the dorsal midline within the anterior half of the embryo (figure 23A). Reflective of *fos* expression, in neurula stage 17 embryos, *atf3* expression was detectable along the edge of the dorsal neural closure (figure 23C). An additional region of *atf3* expression was also detected in an anterior ectodermal domain (figure

23B). Contrastingly, in late neural stage embryos, unlike for *atf3* and *fos*, *jun* was not finely localised along the edge of the neural fold, but was more broadly expressed in the neuroectoderm (figure 23E). Interestingly, in early tailbud stage embryos, *jun* expression in the anterodorsal region of the neuroectoderm looks comparable to the expression of dpERK at the midbrain-hindbrain boundary (figure 23F, figure 21A).



**Figure 23: *In situ* hybridisations for *atf3* and *jun* in neurula stage *X. tropicalis* embryos. (A-C) *In situ* hybridisations for *atf3*. (A) Stage 15, dorsal view, anterior left, (B) stage 17, anterior view, dorsal top, (C) stage 17, dorsal view, anterior left (D-F) *In situ* hybridisations for *jun*. (D) Stage 18, anterior view, dorsal top, (E) stage 18, dorsal view, anterior left, (F) stage 22, dorsoanterior view, dorsal top. White asterisk highlights an ectodermal domain of expression. White arrowhead points to a region within anterodorsal neuroectoderm. Images are representative of n=10 embryos.**

## 4.3 Discussion

### 4.3.1 Reduction in *fos* expression during embryonic wound healing following ERK inhibition may reflect loss of intron-mediated enhancement

Chapter 3 identified *fos* as a putative target of CIC, and CIC is negatively regulated by dpERK (see chapter 1). Hence, it was hypothesised that during embryonic wound healing dpERK induces *fos* expression through negative regulation of CIC. In support of this, recent work within the Isaacs lab found dpERK to be maximally expressed at the wound site within 15 minutes post-wounding, whilst maximal *fos* expression was detected sequentially, peaking around 30-45 minutes post-wounding (Cowell, 2019), thus satisfying the temporal requirements of *fos* being regulated downstream of dpERK. Importantly, it has also been demonstrated within the Isaacs lab that Myc-tagged CIC is degraded upon embryonic wounding in *Xenopus* (*manuscript in preparation*, (King et al. (2022))). Through use of the MEK inhibitor PD0325901, in keeping with the hypothesis, here it was shown that ERK inhibition partially reduces *fos* expression 30 minutes post-wounding in *X. tropicalis* embryos (figure 11). This partial reduction in *fos* expression post-wounding in response to ERK inhibition has also been reported for rat epithelial monolayers treated with the MEK inhibitor PD98059 (Dieckgraefe and Weems, 1999). Notably however, as *fos* expression is positively regulated by Elk-1, which is also activated by dpERK (see chapter 3 discussion; 3.3.7), it cannot be ruled out that the reduction in *fos* expression following ERK inhibition fully or partially reflects loss of Elk-1 mediated induction.

As ERK inhibition did not completely eliminate *fos* expression post-wounding, this implies that ERK does not exclusively regulate *fos* expression during embryonic wound healing. Phosphorylation of CIC by Src is able to relieve CIC-mediated repression, and Src family kinases are activated in response to wounding (Tsarouhas, Yao and Samakovlis, 2014; Yoo et al., 2012). Perhaps Src is able to somewhat compensate for the absence of dpERK and relieve CIC-mediated repression of *fos* post-wounding. Alternatively, based on the proposal that CIC may act as a repressor of a *fos* intronic promoter (see chapter 3 discussion; 3.3.6), it is exciting to speculate that the reduction but not complete loss of *fos* expression following ERK inhibition could reflect continued transcription from the *fos* canonical upstream promoter, but loss of transcription from the intronic promoter due to CIC continuing to elicit repression. To reiterate, perhaps ERK inhibition leads to a partial but not complete reduction in *fos* expression due to loss of the IME effect only.

### 4.3.2 Nitric oxide may serve as a negative regulator of dpERK and resultantly *fos* during embryonic wound healing

Abaffy et al. (2019) demonstrated that inhibiting nitric oxide production was able to reverse a decrease in *fos* expression normally seen between 60 and 90 minutes post-wounding in *X. laevis* embryos. Based on the hypothesis that dpERK induces *fos* expression during embryonic wound healing, it was questioned whether the effects of nitric oxide inhibition on *fos* expression could be explained by altered dpERK expression. Indeed, adding further strength to the hypothesis, dpERK expression was higher in *X. laevis* embryos treated with the nitric oxide synthase inhibitor TRIM than in control embryos at 90 minutes post-wounding (figure 13C,F). Any difference in dpERK expression between TRIM treated and control embryos is less obvious at 50 minutes post-wounding (figure 13A,D), suggesting nitric oxide does not impact the initial induction of dpERK post-wounding. In corroboration, Abaffy et al. (2019) found *fos* expression at 30 minutes post-wounding to be comparable between control embryos and those in which nitric oxide production had been inhibited. Overall, this postulates nitric oxide as a 'delayed' negative regulator of dpERK during embryonic wound healing and as such, nitric oxide may mediate the switch between ERK and PI3K signalling in the early and late phases of wound closure, respectively (figure 24). It is interesting to speculate whether negative regulation of ERK by nitric oxide might cooperate with upregulation of MAPK negative regulators downstream of CIC/ERK (see chapter 3 discussion; 3.3.3) to ensure a unidirectional transition from ERK to PI3K signalling and progression through the stages of embryonic wound healing. However, any conclusions drawn from this experiment are inherently limited due to the absence of a control to confirm inhibition of nitric oxide production; instead the knowledge that a similar assay was successful in Abaffy et al. (2019) was relied upon.

Several other lines of evidence add weight to the proposal that nitric oxide may function as a negative regulator of dpERK activity during embryonic wound healing. Firstly, the peak in nitric oxide production seen at 15 minutes post-wounding seemingly coincides with the onset in reduction of dpERK expression during *X. laevis* embryonic wound healing (Abaffy et al., 2019; Cowell, 2019). Secondly, inhibition of nitric oxide production delays initial wound closure, whilst inhibiting ERK activation has a complementary phenotype and increases the initial speed of wound closure (Abaffy et al., 2019; Li et al., 2013). Finally, S-nitrosylation of ERK by nitric oxide is a phenomenon known to inhibit ERK phosphorylation and hence provides a possible mechanism as to how nitric oxide may inhibit ERK activity (Feng et al., 2013).



**Figure 24: Hypothesised axis of regulation between nitric oxide, dpERK and fos during *Xenopus* embryonic wound healing.** During the early phase of wound healing ERK becomes activated and may induce fos expression through negative regulation of the transcriptional repressor CIC. Activated ERK in the early phase of wound healing also suppresses PI3K signalling. A burst of nitric oxide (NO) may subsequently attenuate ERK activity, enabling PI3K signalling and a transition to the later phase of wound healing, alongside reducing fos expression to basal levels.

To extend the analysis undertaken by Abaffy et al. (2019), the anti-Fos 2G9 antibody was used to probe whether the change in fos mRNA expression following nitric oxide synthesis inhibition also translated to altered Fos protein expression during embryonic wound healing. As inhibiting nitric oxide production prolonged heightened dpERK expression post-wounding (figure 13), and Fos acts as a read-out for ERK signal duration (Murphy et al., 2002), it was predicted that Fos protein would be present at higher levels in TRIM treated embryos than control embryos. Furthermore, the regulation of Fos by nitric oxide isn't completely novel; the DNA-binding ability of Fos/Jun AP-1 dimers can be inhibited by nitric oxide through a reaction with conserved cysteine residues in Fos and Jun (Nikitovic, Holmgren and Spyrou, 1998). Should nitric oxide also inhibit Fos at the level of transcription through negative regulation of dpERK, this would nicely complement the negative post-translational regulation of Fos by nitric oxide. Whilst anti-Fos 2G9 is capable of detecting endogenous *Xenopus* Fos protein (figure 14), and Fos protein levels do appear higher in TRIM treated embryos at 90 minutes post-wounding than in control embryos (figure 15C,F,I,J), the assay requires further optimisation to obtain a more conspicuous result. Perhaps reducing the dilution of anti-Fos 2G9 may allow for better detection of endogenous Fos protein. In addition, examining time points beyond 90 minutes post-wounding may allow more time for sustained ERK activation in response to inhibition of nitric oxide production to be translated into elevated Fos protein expression.

### 4.3.3 CIC as a general regulator of wound-responsive genes

To investigate whether CIC may regulate other wound-responsive genes alongside *fos*, genes significantly upregulated following CIC knockdown were compared to those significantly upregulated at 30, 60 and 90 minutes post-wounding in Abaffy et al. (2019) (figure 16, additional file S9). As the CIC knockdown RNA-seq and wounding RNA-seq data were conducted in *X. tropicalis* and *X. laevis* embryos, respectively, the list of genes investigated had to be reduced to 8,713 genes with common annotation in both species. Therefore, from the start the analysis was somewhat limited due to the small number of investigable genes - ideally, RNA-seq data sets from the same species and at the same developmental stage should have been compared. Nonetheless, at all time points there was a significant intersection between genes significantly upregulated by CIC knockdown and significantly upregulated following wounding (figure 16). In total, 51 genes were identified as regulated by both CIC and during embryonic wounding and implies CIC may be a more general regulator of wound-responsive genes. A number of genes are significantly upregulated by CIC knockdown but not following wounding, and this probably denotes CIC's involvement in regulating other developmental processes, including for example lung development (Lee et al., 2002). Likewise, a number of genes are significantly upregulated following wounding but not CIC knockdown and this likely reflects wound-induced signalling through other CIC-independent pathways. Particularly interesting is that multiple AP-1 component genes including *jun*, *jund*, *fosl1*, *atf3* and *fos* are all upregulated at 30 minutes post-wounding and by CIC knockdown. Therefore, genes expressed at later time points post-wounding may represent AP-1 target genes. Indeed, *mmp-1* upregulated at 90 minutes post-wounding is a well-documented AP-1 target gene (table 7) (Lin, Georgescu and Evans, 1993; Kimura et al., 2011). The upregulation of MMPs post-wounding is supportive of a role for *de novo* gene induction in the final proposed stages of embryonic wound healing involving tissue remodelling (Abaffy et al., 2019).

Congruent with the RNA-seq data from Abaffy et al. (2019) both *jun* and *atf3* were detectable at the wound site in *X. tropicalis* embryos. The induction of *jun* was detectable within 20 minutes post-wounding, and this might reflect that like *fos*, *jun* is an immediate early gene (Greenberg and Ziff, 1984; Bahrami and Drabløs, 2016). As *fos* is inactive as a monomer (Halazonetis et al., 1988), expression of its dimeric partner *jun* conceptually makes sense should Fos play a functional role in embryonic wound healing. The detection of *jun* at the wound site was rather inconsistent, and it is not obvious as to why. One possible explanation could be a technical fault during the *in situ* hybridisation process as the analysis was only completed once. Alternatively, the data could represent a true result and perhaps *jun* expression is highly sensitive to the

size and/or depth of wound which, due to manual wounding, would be slightly varied from embryo to embryo.

Although *atf3* was not readily detectable at the wound site until 60 minutes post-wounding, due to a suboptimal colour reaction during the *in situ* hybridisation process, it is not deemed sufficient to negate earlier *atf3* expression post-wounding. Indeed, Abaffy et al. (2019) detected an increase in *atf3* expression at 30 minutes post-wounding. Further, as *atf3* represents yet another example of an immediate early gene (Xie et al., 2005), it is hypothesised that its expression will be detected in a similar manner to *fos* and *jun* post-wounding. As multiple AP-1 genes are upregulated post-wounding and following CIC knockdown, it is interesting to question whether the different possible AP-1 variants may show redundancy to produce a robust wound-healing response or have different functional roles.

One gene clustering with *fos* based upon its developmental expression pattern, and significantly upregulated following CIC knockdown and post-wounding was *rasl11b*, a gene encoding a small Ras-like GTPase (figure 17) (Pézeron et al., 2008). Like *fos*, *rasl11b* expression at the wound site was detected rapidly at 5 minutes post-wounding (figure 19). As such, *rasl11b* might represent an additional novel IEG in the wound response regulated by CIC, and suggests a role for CIC beyond regulation of solely AP-1 component genes. Despite *rasl11b* being identified as significantly upregulated at 60 and 90 minutes post-wounding in Abaffy et al. (2019), at these time points *rasl11b* was not readily detectable at the wound site via *in situ* hybridisation (*data not shown*). This may reflect the sensitivity of RNA-seq in detecting differential gene expression. The detection of *rasl11b* expression between 5-20 minutes post-wounding via *in situ* hybridisation also highlights that the analysis in Abaffy et al. (2019) may fail to produce a comprehensive list of wound-responsive genes by assaying changes in gene expression from 30 minutes post-wounding onwards. Genes which show rapid and/or transient induction (< 30 minutes) in response to wounding may be missed.

#### 4.3.4 Similarities exist in the regulation of *fos* expression during wound healing and neurulation

The observation has been made several times that natural morphogenetic movements during embryogenesis, for example closure of the blastopore in *Xenopus*, superficially resemble embryonic wound healing (Martin et al., 1994; Wood et al., 2002; Nodder and Martin, 1997). In *Drosophila*, assembly of the actin contractile machinery during both wound healing and dorsal closure has been shown to be similar (Wood et al., 2002). Akin to wound healing, neurulation also appears to rely in part on actin contractile

machinery (Morriss-Kay and Tuckett, 1985). Henceforth, the hypothesis was extended to suggest that dpERK may induce *fos* expression through negative regulation of CIC during both wound healing and morphogenesis. In support, mirroring the colocalization of *fos* and dpERK during wound healing, *fos* and dpERK also co-localise along the dorsal neural closure during neurulation (figure 20). Similarly, activation of ERK in the context of wound healing is FGF-independent (Christen and Slack, 1999), and due to the absence of detectable FGF ligands along the dorsal neural closure, it has been proposed that this also represents a region of FGF-independent ERK activation (personal communication with Dr. Harry Isaacs). In agreement, unlike treatment with the MEK inhibitor PD0325901, treatment with the FGFR inhibitor SU5402 was unable to inhibit dpERK expression along the dorsal neural closure in *X. tropicalis* embryos (figure 21). Importantly, this does not reflect a lack of efficacy of SU5402 as known FGF-dependent domains of ERK activation, including the posterior mesoderm, midbrain-hindbrain boundary, prospective gill region and forebrain region were all inhibited by SU5402 (figure 21).

Once again reflective of embryonic wound healing, treatment with PD0325901 was found to largely reduce *fos* expression along the dorsal neural closure, but not completely eliminate all expression (figure 22). This is in an FGF-independent manner, as expression of *fos* along the dorsal neural closure is comparable to wildtype in embryos treated with SU5402 (figure 22). Overall this supports the hypothesis, but evokes that dpERK is not the sole regulator of *fos* expression during both neurulation and embryonic wound healing. The same explanations delineated as to why ERK inhibition may not completely inhibit *fos* expression in the context of wound healing are also applicable to the context of neurulation. For example, Src family kinases are also expressed during neurulation and so may phosphorylate CIC and compensate for loss of ERK-mediated phosphorylation (Bradley, 2016; Bunda et al., 2020). Alternatively, following ERK inhibition, *fos* expression may be reduced along the dorsal neural closure due to continued repression of a *fos* intronic promoter by CIC, or loss of Elk-1 mediated induction (see chapter 4 discussion; 4.3.1). Although the involvement of CIC is not directly investigated, as CIC-L long is broadly expressed along the neural plate, excluding the neural groove (King, 2018), this is in keeping with a role for CIC regulating *fos* expression during neurulation. Interestingly, as nitric oxide also contributes to neurulation (Nachmany et al., 2006), it is questioned whether in this context nitric oxide might also serve as a negative regulator of dpERK and *fos* expression.

### 4.3.5 AP-1 activity during neurulation

In further support of conservation between wound healing and morphogenesis, as with *fos*, alongside being expressed during embryonic wound healing, *jun* and *atf3* also show expression within the neural plate during neurulation (figure 23). Whilst *atf3* expression closely mirrors that of *fos*, and occurs along the edge of the dorsal neural closure (figure 23C), the domain of *jun* expression is much broader within the neural plate (figure 23E). This may reflect that as Fos and Jun, and Jun and Atf3 heterodimerize to form AP-1, the restricted expression of *fos* and *atf3* alone would be sufficient to localise the activity of Fos/Jun and Jun/Atf3 AP-1 heterodimers.

The role of AP-1 during neurulation is unclear. However, on the basis that conservation may exist between wound healing and morphogenesis it is tempting to propose AP-1 may serve a similar function in both contexts. Here, it was proposed that AP-1 may be involved in the later stages of embryonic wound healing which concern tissue remodelling through regulating the expression of MMPs, such as *mmp-1* (see chapter 4 discussion; 4.3.3). MMPs have also been extensively linked to neural development (Small and Crawford, 2016). MMP-9, for example, has been associated with neural crest cell delamination and migration in avian embryos and also represents a known AP-1 target gene (Monsonogo-Ornan et al., 2012; Mittelstadt and Patel, 2012). As the AP-1 component genes investigated here, *fos*, *jun* and *atf3*, are expressed closely to where neural crest cells begin their delamination and migration following neural tube closure, perhaps AP-1 is involved in regulating this process. Indeed, it would be interesting to see if *mmp* genes show similar patterns of expression to AP-1 component genes along the dorsal neural closure. As *fos*, *jun* and *atf3* were also all identified as CIC targets, and the CIC-L isoform is broadly expressed within the neural plate in *Xenopus* (King, 2018), by extension CIC may also be involved in regulating neural crest cell delamination and migration. CIC has already been implicated in several other aspects of neurodevelopment including neuronal stem cell proliferation and differentiation (Ahmad et al., 2019; Hwang et al., 2020). Interestingly, in keeping with the proposal similar mechanisms govern wound healing and morphogenesis, MMP-9 shows transient expression post-wounding in murine models, and it has been proposed elevated MMP-9 expression is involved in chronic wound pathogenesis through its deleterious effects on keratinocyte migration (Reiss et al., 2010).

# Chapter 5: General Discussion

## 5.1 Summary

The activity of the transcriptional repressor CIC is regulated by dpERK, with phosphorylation of CIC by dpERK able to dissociate CIC from its target genes and relieve repression (Astigarraga et al., 2007; Jiménez et al., 2000). ERK becomes activated downstream of FGF signalling and during embryonic wound healing, however the regulation of CIC by dpERK in these contexts is relatively understudied. In this project it was hypothesised that ERK relieves CIC-mediated repression downstream of FGF signalling and post-wounding during *Xenopus* development.

Here, 44 putative CIC/FGF target genes have been identified in *Xenopus*. It was subsequently validated that one such CIC/FGF target gene, *fos*, shows FGF and ERK-dependent expression during early gastrulation. Candidate CIC binding sites were found within a highly conserved region of *fos* intron 1 corresponding to a putative intronic promoter. Novelty, this postulates a role for CIC in IME of *fos*. Overall, the findings from this project are supportive of the hypothesis ERK relieves CIC-mediated repression downstream of FGF signalling, as alongside being a target of FGF signalling, *fos* may represent a conserved and direct target of CIC.

In support of the hypothesis that ERK also relieves CIC-mediated repression in the context of embryonic wound healing, ERK inhibition partially reduced *fos* expression post-wounding. The regulation of *fos* by dpERK during wound healing is further supported by the finding that inhibiting nitric oxide production elevated dpERK expression post-wounding; Abaffey et al. (2019) has previously demonstrated that this same treatment was also able to increase *fos* expression. As such, nitric oxide may be a negative regulator of dpERK and may facilitate the switch between ERK signalling and PI3K signalling during the early and late phases of embryonic wound healing, respectively.

Multiple AP-1 component genes, including *fos*, *jun*, and *atf3* are significantly upregulated following both embryonic wounding and CIC knockdown. Overall, this suggests CIC may be a general regulator of wound-responsive genes through regulation of AP-1. Previously the involvement of *de novo* gene expression to embryonic wound healing has been overlooked due to the rapidity of the process. However, based on the number of genes showing differential expression post-wounding, it seems unlikely *de novo* gene expression has an irrelevant role. Immediate early genes, for example the AP-1 component genes *fos*, *jun* and *atf3*, may

be particularly influential to embryonic wound healing as they can be rapidly induced. In turn, downstream AP-1 target genes, such as matrix metalloproteinases, may play a role in the later stages of embryonic wound healing involving tissue remodelling.

The observation that wound healing superficially resembles morphogenesis has been drawn numerous times. Consequently, the hypothesis was extended to suggest that ERK relieves CIC-mediated repression in the context of both embryonic wound healing and neurulation. Here it was found that *fos* and dpERK co-localise during both wound healing and neurulation, and in both contexts *fos* expression is partially reduced following ERK inhibition. Other AP-1 component genes, *atf3* and *jun*, which are induced post-wounding, are also expressed during neurulation. Overall therefore, the findings of this project support the notion that similar underlying mechanisms may govern wound healing and morphogenesis, and suggest it may be possible to transfer understanding between contexts.

## 5.2 Limitations and future work

Perhaps the main caveat to the evidence presented in this project is that *fos* is putatively regulated by both Elk-1 and CIC. The reduction in *fos* expression following ERK inhibition during morphogenesis and post-wounding is supportive of the hypothesis as this could reflect CIC continuing to elicit repression. However, as both Elk-1 and CIC are regulated by ERK, it is impossible to discern whether the reduction in *fos* expression following ERK inhibition reflects altered Elk-1 or CIC activity, or a combination of both. As such, it is imperative to directly link *fos* expression to CIC/ERK. Phosphomimetic forms of CIC, specifically concerning ERK phosphorylation sites, could be generated, and it could be investigated whether these mitigate the reduction in *fos* expression seen following ERK inhibition. Similarly, a phospho-resistant form of CIC could be used to see whether this ablates the increase in *fos* expression seen following FGF4 overexpression. These experiments would enable a direct link between FGF/ERK/CIC/*fos* during gastrulation, and ERK/CIC/*fos* during embryonic wound healing and neurulation, to be established.

In this project CIC binding motifs were identified within a conserved region of *fos* intron 1. Future work could determine whether these motifs correspond to *in vivo* CIC binding sites in *Xenopus* by performing ChIP-PCR with an antibody against CIC. Should this be the case, it would be interesting to investigate the contribution of CIC to *fos* IME. *In vitro* reporter assays could investigate whether the addition of purified CIC eliminates the intronic promoter activity of *fos* intron 1. Further, sequencing technologies could be utilised to probe whether *in vivo* *fos* transcription is driven from two different promoter

regions, a canonical upstream promoter, and an intronic promoter, and whether the presence of CIC alters the proportions of these two transcripts.

This project provides evidence that nitric oxide negatively regulates dpERK expression during embryonic wound healing. As such, the burst of nitric oxide released during wound healing *in vivo* may act as the switch between ERK and PI3K signalling. To explore this further, it could be investigated whether TRIM treatment also delays PI3K signalling activity during embryonic wound healing through immunostaining with an antibody against phospho-Akt. To examine a potential mechanism of nitric oxide inhibiting ERK activity, the levels of S-nitrosylated ERK could be compared between control and TRIM treated embryos post-wounding via a biotin switch assay and Western blotting. Broadly, due to the observed similarities between embryonic wound healing and morphogenesis, the whole analysis could be extended to investigate the regulation of *fos* and/or dpERK by nitric oxide during neurulation.

As CIC may be a more general regulator of AP-1 expression, the regulation of *atf3* and *jun* by CIC/ERK during embryonic wound healing and neurulation could be investigated. Furthermore, as both full-length *fos* and *jun* mRNA have been synthesised, microinjection experiments could be utilised to examine the effects of AP-1 overexpression on embryonic wound healing and neurulation. Inhibiting ERK activity has been shown to increase the initial speed of wound closure (Li et al., 2013), it would be interesting to investigate whether overexpression of AP-1 as a downstream target of ERK has a reciprocal effect and delays wound closure.

### 5.3 Conclusions and implications

The findings of this project support the hypothesis that ERK relieves CIC-mediated repression downstream of FGF signalling and during embryonic wound healing. As such, CIC putatively regulates FGF target genes and wound-responsive genes including, notably, *fos*. Further investigation into the proposed ERK/CIC/*fos* regulatory axis is crucial as misregulation of each component is linked to cancer. Indeed, CIC functions as a tumour suppressor and abnormal CIC activity has been associated with oligodendrogliomas, glioblastomas, and E-wing like sarcomas, amongst other cancers (Bettegowda et al., 2011; Yip et al., 2012; Bunda et al., 2019; Kawamura-Saito et al., 2006) The oncogenic activity of *fos* is also well-documented; overexpression of *fos* is able to transform cells in *in vitro* transformation assays and induce tumours *in vivo* (Miller, Curran and Verma, 1984; Grigoriadis et al., 1993). Further, constitutive activation of the Ras-MAPK pathway, and hence ERK, is a rife feature of many cancers, with gain-of-function Raf and Ras mutations being particularly common

(Santarpia, Lippman and El-Naggar, 2012; Dhillon et al., 2007). Therefore, detailed characterisation of an ERK/CIC/*fos* regulatory axis might accentuate potential therapeutics to counteract aberrant Ras-MAPK signalling during cancer. Similarly, establishing the existence and function of this pathway downstream of FGF signalling and during wound healing might highlight therapeutics to benefit wound healing disorders, such as chronic wounds, and disorders associated with mis-regulation of FGF signalling, for example growth conditions including congenital craniosynostosis and dwarfism syndromes, alongside various cancers and chronic kidney disease (Xie et al., 2020). Furthermore, as nitric oxide therapy is a current treatment for certain wound healing pathologies (Malone-Povolny, Maloney and Schoenfisch, 2019), identifying dpERK as negatively regulated by nitric oxide during embryonic wound healing could even postulate MEK inhibitors as an alternative treatment. Overall therefore, the work presented in this project has the potential to be extended into medically relevant knowledge, which in turn may assist development of therapeutics for numerous and diverse pathophysiologicals.

# Appendices

## Appendix A: Sequence of full-length *fos* and *jun* inserts in pCS2+ vector

Full-length *fos* and *jun* were cloned into the vector pCS2+ (see materials and methods; 2.2.4, 2.2.5). Plasmid samples were sent for sequencing with SP6 and T7 primers (Eurofins Genomics). Analysis of sequencing data in SeqMan software (DNA Star) confirmed both full-length *fos* and *jun* had been cloned without any mutations. The 5' - 3' sequence of the *fos* and *jun* inserts is below. The 5' end of the insert contains a BamHI site, and the 3' end of the insert contains an XhoI site; the cut versions of these restriction sites are indicated by lower case letters. The ATG start codon is highlighted in bold.

Insert	Sequence (5' - 3')
<i>fos</i>	gatccACC <b>ATG</b> TATCACGCCTTCTCCAGCAGCACCGAATATGATGCAG CTTCTTCCCGTTGCAGTAGTGCCTCTCCAGCCGGGGACAGCCTGAC CTACTACCCGTCCCCTGCAGCTTCTTCTCTAGCATGGGGTCACCT GTTTCTCCACAGGACTTCGGTGGTGATTCAAGCAGCAGTTTTGTAC CCACAGTCACTGCCATTTCCACCTCACCAGATCTTCAGTGGCTTGTA CAGCCAACCCTTATTTCTTCTGTAGCCCATCACAGTCTCGGGCACA CCCTTATGGGTCCACACCAGCTTACAGCCGATCTAGCGTTATGAAAG GATCTGCTGGAAGAGGTCAGAGCCTGGGAAGAAGAGGAAAAATGG AGCAGCTTTCTCCAGAAGAAGAAGAAAAAAGGAAAGTAAGACGAGA AAGGAATAAGATGGCAGCTGCCAAGTGTCGTAACCGCCGTCGGGA GTAAACAGACACCCTTCAAGCGGAGACTGATGACCTGGAGGACCAG AAATCTGCCCTGCAGGCAGAGATTGCCGGCCTTCTAAAGGAGAAGG AAAAGCTGGAGTTTATACTTGCAGCTCACAAACCAGCTTGCAAATT CCACATGATCTTGATGGAGCTTTTCAAGACTTGACCTCATCTCTTGAT CTGGGTCTGATTTTCAGAGACCCCTTGTCTTCCAGCTCTCAGGAGC CTGTAGCAGAGCCTCTGTTTCCCATTGGCCTTTCTCAGTCTTCCATG CCTGAAAAGGAGAACACCCAGCTGCAAGTCTCTATGGAECTCAAAT CTGAACCACTGGATGATTTTCTGTTTAACTCTTCTCACACAGGTGTAA CTGATGCAGCACGTTCTGTGCCAGATGTAGATCTTACTAGCTCTCTT TACACATCAGAATGGGAGCCACTGTATAGCACTTTATCTGCAGACAT GGAGCCTCTGTGCACACCAGTTGTTACCTGCACTCCAACATGCACT ACCTATACAACATCATTGTCTTACATACCCAGAATCTGACCACTTC CCCAACTGTGGAGCCGCACATCGGAGGGGAAGCAGCAGCAATGAG CAGTCATCAGACTCTCTAAACTCTCCACCCTTTTGGCTCTGgagct

*jun*

gatccACCATGACTGCAAGGATGGAACCTACTTTCTACGAAGATGCCCTGAGTGCTGCTTTTCGCCCAGCACGATGCTACTCCCTATGGTTACAACAAGGTGCTGAAGCAAAGCATGACCCTGAACCTGTCGGACCCCAGCAGCGCCATCAAACCTCACCTGAGAAACAAGGCAGCTGAACTGCTCACCTCCCCTGATGTTGGGCTTCTCAAGCTCGCCTCCCCGGAGTTGGAGAGGCTTATCATCCAGTCCAGCAATGGCATGATTACCACCACCCCGACCCCAACCAGTTCCTGTGCCCAAGAATGTCACAGACGAGCAAGAAGGCTTTGCAGAGGGGTTTGTAGGGCACTGGCAGAACTCCACCATCAAACAACCTTGCCAAGTGTAACCACTGCCACCCAACCTGCCAGCACGGGACTGACACCTGTATCTACTATTGCTGGAAACACTGGCTTCAACAATAGTCTTCACAACGAGCCCCCTGTGTATGCCAATCTAAGTAACTTTAACCCAAGTACCATCACCCACATCATCGCCTTTTAACAGCAACACCATGGGGTATACCGGCCAACATCAGAGTAACCCACCAATACCCGTGCAGCACCCAGGTTACAGGCTCTGAAAGAGGAACCACAGACTGTACCTGAAATGCCTGGGGAGACTCCTCCACTGTCCCCTATTGACATGGAGTCCAGGAAAGGATAAAGGCTGAAAGGAAGCGTATGAGGAATAGAATCGCAGCATCTAAATGCAGAAAAAGGAAACTGGAGAGGATTTCCAGGTGGAAGACAAAGTTAAAACTTAAATCCCAGAACTCTGAACTGGCATCCACTGCCAACATGCTTAGAGAGCAAGTAGCCCAGCTCAAACAAAAGTCATGAATCATGTCAACAGTGGGTGTCAGCTAATGTTAACACAGCAGATGCAAACATTCgagct

## Appendix B: Predicted protein sequence of Fos and Jun clones

Predicted protein sequence of Fos and Jun as translated from mRNA synthesised from pCS2+ *fos* and pCS2+ *jun* (see materials and methods; 2.2.6). Note that the endogenous stop codon for Fos and Jun was not included in the reverse primers designed to amplify the full-length clones (table 1). Hence, 6 additional amino acids (Leu, Glu, Pro, Leu, Glu, Leu) are predicted to be incorporated onto the C-terminus of Fos and Jun before two adjacent stop codons are reached within the pCS2+ vector sequence. Additional pCS2+ encoded amino acids are underlined. Nt; N-terminus, Ct; C-terminus.

Plasmid	Sequence (Nt - Ct)
pCS2+ <i>fos</i>	<p>           MYHAFSSSTEYDAASSRCSSASPAGDSLTYYPSPAASFSSMG            SPVSPQDFGGDSSSSSFVPTVTAISTSPDLQWLWQPTLISSVAPS            QSRAHPYGSTPAYSRSSVMKGSAGRGQSLGRRGKMEQLSPE            EEEKRKVRERERNKMAAAKCRNRRRELDTLQAETDDLEDQKS            ALQAEIAGLLKEKEKLEFILAAHKPACKIPHDLDGAFQDLTSSLD            LGLISETPCSSSSQEPVAEPLFPIGLSQSSMPEKENTQLQVSM            ELKSEPLDDFLFNSSHTGVTDAAARSVPDVLDTSSLYTSEWEPL            YSTLSADMEPLCTPVVTCTPTCTTYTTSFVFYTPESDHFPNCG            AAHRRGSSSNEQSSDSLNSPTLLALL<u>LGPLGL</u> </p>
pCS2+ <i>jun</i>	<p>           MTARMEPTFYEDALSAAFAQHDATPYGYNKVLKQSMTLNLSD            PSSAIKPHLRNKAELLTSPDVGLLKLASPELERLIQSSNGMIT            TTPTPTQFLCPKNVTDEQEGFAEGFVRALAEHHQNNLPSVTT            ATQPASTGLTPVSTIAGNTGFNNSLHNEPPVYANLSNFPSTIT            TSSPFNSNTMGYTGQHQSNNPIPQHPRLQALKEEPQTVPEM            PGETPPLSPIDMESQERIKAEKRMNRNRIAASKCRKRKLERISR            LEDKVKNLKSQNSELASTANMLREQVAQLKQKVMNHVNSGCQ            LMLTQQMQTFL<u>LGPLGL</u> </p>

## Appendix C: R code for gene-level differential expression analysis

Link to ["samples\\_subset3.txt"](#)

Link to ["transcripts.csv"](#)

```
# Clear workspace
rm(list= ls())

# Load Sleuth
library(sleuth)
library(dplyr)

# Set the working directory
setwd("~/RNA Seq Analysis")

# Load table containing data on samples.
samples<-read.table("samples_subset3.txt", sep="\t", header=TRUE,
stringsAsFactors=FALSE)
summary(samples)

# Load table containing data on transcripts.
transcripts <- read.csv("transcripts.csv", header = TRUE, sep=",")
summary(transcripts)
head(transcripts)

# Create a Sleuth object
so <- sleuth_prep(samples, target_mapping=transcripts, aggregation_column =
"gene_id",
  read_bootstrap_tpm = TRUE, extra_bootstrap_summary = TRUE,
  gene_mode = TRUE)

# Fit to distribution
so <- sleuth_fit(so, ~genotype + batch, "full")

so <- sleuth_fit(so, ~batch, "reduced")

so <- sleuth_lrt(so, 'reduced', 'full')

# Perform Wald test
so <- sleuth_wt(so, "genotypeFGF4")
so <- sleuth_wt(so, "genotypeCICTALEN")

models(so)

# Visualise analysis live
sleuth_live(so)
# Create results table
results_table_genotypeFGF4 <- sleuth_results(so, 'genotypeFGF4', test_type = 'wt')
```

```

results_table_genotypeCICTALEN <- sleuth_results(so, 'genotypeCICTALEN',
test_type = 'wt')
head(results_table_genotypeFGF4)
head(results_table_genotypeCICTALEN)

# Save table
write.table(results_table_genotypeFGF4, file="sleuth_wald_test_genotypeFGF4.csv",
sep=",", row.names = FALSE)
write.table(results_table_genotypeCICTALEN,
file="sleuth_wald_test_genotypeCICTALEN.csv", sep=",", row.names=FALSE)

table_genotypeFGF4 <- read.csv(file = 'sleuth_wald_test_genotypeFGF4.csv', sep =
',', header = TRUE)
table_genotypeCICTALEN <- read.csv(file =
'sleuth_wald_test_genotypeCICTALEN.csv', sep = ',', header = TRUE)
head(table_genotypeFGF4)
head(table_genotypeCICTALEN)

# Remove duplicated rows from table
table_genotypeCICTALEN1 <-
table_genotypeCICTALEN[!duplicated(table_genotypeCICTALEN$target_id),]
table_genotypeFGF41 <-
table_genotypeFGF4[!duplicated(table_genotypeFGF4$target_id),]
head(table_genotypeFGF41)
head(table_genotypeCICTALEN1)

# Merge genotypeFGF41 and genotypeCICTALEN1 tables
table <- merge(table_genotypeFGF41, table_genotypeCICTALEN1, by = c('target_id',
'transcript_type', 'transcript_name'))

head(table)

# Extract TPM data from kallisto
TPM <- kallisto_table(so, use_filtered = FALSE, normalized = TRUE,
include_covariates = TRUE)
head(TPM)

# Subset based on sample and then remove unwanted columns

MK1TPM <- subset(TPM, TPM$sample == 'MK1')
MK1TPM <- MK1TPM[,-c(1,3,5,6)]
colnames(MK1TPM)[2] <- "MK1_TPM"
head(MK1TPM)

MK2TPM <- subset(TPM, TPM$sample == 'MK2')
MK2TPM <- MK2TPM[,-c(1,3,5,6)]
colnames(MK2TPM)[2] <- "MK2_TPM"
head(MK2TPM)

MK3TPM <- subset(TPM, TPM$sample == 'MK3')

```

```

MK3TPM <- MK3TPM[,-c(1,3,5,6)]
colnames(MK3TPM)[2] <- "MK3_TPM"
head(MK3TPM)

MK4TPM <- subset(TPM, TPM$sample == 'MK4')
MK4TPM <- MK4TPM[,-c(1,3,5,6)]
colnames(MK4TPM)[2] <- "MK4_TPM"
head(MK4TPM)

MK5TPM <- subset(TPM, TPM$sample == 'MK5')
MK5TPM <- MK5TPM[,-c(1,3,5,6)]
colnames(MK5TPM)[2] <- "MK5_TPM"
head(MK5TPM)

MK6TPM <- subset(TPM, TPM$sample == 'MK6')
MK6TPM <- MK6TPM[,-c(1,3,5,6)]
colnames(MK6TPM)[2] <- "MK6_TPM"
head(MK6TPM)

MK10TPM <- subset(TPM, TPM$sample == 'MK10')
MK10TPM <- MK10TPM[,-c(1,3,5,6)]
colnames(MK10TPM)[2] <- "MK10_TPM"
head(MK10TPM)

MK11TPM <- subset(TPM, TPM$sample == 'MK11')
MK11TPM <- MK11TPM[,-c(1,3,5,6)]
colnames(MK11TPM)[2] <- "MK11_TPM"
head(MK11TPM)

MK12TPM <- subset(TPM, TPM$sample == 'MK12')
MK12TPM <- MK12TPM[,-c(1,3,5,6)]
colnames(MK12TPM)[2] <- "MK12_TPM"
head(MK12TPM)

# Merge kallisto TPM data onto end of table containing FGF4 and CICTALEN Wald
data

table_1 <- merge(table, MK1TPM, by = 'target_id')
table_1 <- merge(table_1, MK2TPM, by = 'target_id')
table_1 <- merge(table_1, MK3TPM, by = 'target_id')
table_1 <- merge(table_1, MK4TPM, by = 'target_id')
table_1 <- merge(table_1, MK5TPM, by = 'target_id')
table_1 <- merge(table_1, MK6TPM, by = 'target_id')
table_1 <- merge(table_1, MK10TPM, by = 'target_id')
table_1 <- merge(table_1, MK11TPM, by = 'target_id')
table_1 <- merge(table_1, MK12TPM, by = 'target_id')
head(table_1)

# Save final table

```

```
write.table(table_1, file="sleuth_wald_test_full_table.csv", sep=",", row.names =  
FALSE)
```

# Abbreviations

AP-1	Activator protein 1
bZip	Basic leucine zipper
ChIP-seq	Chromatin immunoprecipitation sequencing
CIC / CIC-L / CIC-S	Capicua / Capicua-long isoform / Capicua-short isoform
CNS	Central nervous system
CRE	cAMP-responsive element
DUSP	Dual-specificity phosphatase
EBS	ERK binding site
EGF	Epidermal growth factor
ERK / dpERK	Extracellular signal-regulated kinase / di-phosphorylated extracellular signal-regulated kinase
Ets	E26 transformation-specific
FGF / FGFR	Fibroblast growth factor / fibroblast growth factor receptor
FRS2	Fibroblast growth factor receptor substrate 2
Grb2	Growth factor receptor-bound protein 2
HMG-box	High-mobility group box
HPF	Hours post-fertilisation
HSPG	Heparan sulphate proteoglycan
Ig	Immunoglobulin-like
IME	Intron-mediated enhancement
JNK	Jun N-terminal kinase
MAPK	Mitogen-activated protein kinase

MEK	Mitogen-activated protein kinase kinase
MMP	Matrix metalloproteinase
MRS	Modified Ringer's solution
NAM	Normal amphibian medium
NLS	Nuclear localisation signal
NO	Nitric oxide
PDGFRB	Platelet-derived growth factor receptor beta
PEA3	Polyomavirus enhancer activator 3
PI3K	Phosphatidylinositol 3-kinase
PLC $\gamma$	Phospholipase C $\gamma$
PTB	Phosphotyrosine-binding domain
RSK	Ribosomal S6 kinase
RT	Room temperature
RTK	Receptor tyrosine kinase
SOS	Son of sevenless
SRE	Serum response element
SRF	Serum response factor
TALEN	Transcription activator-like effector nucleases
TCF	Ternary complex factor
TPM	Transcripts per million
TRE	TPA-responsive element
VEGF/ VEGFR	Vascular endothelial growth factor / vascular endothelial growth factor receptor

# References

- Abaffy, P. et al. (2019). The role of nitric oxide during embryonic wound healing. *BMC genomics*, 20 (1), p.815.
- Abreu-Blanco, M. T., Verboon, J. M. and Parkhurst, S. M. (2014). Coordination of Rho family GTPase activities to orchestrate cytoskeleton responses during cell wound repair. *Current biology: CB*, 24 (2), pp.144–155.
- Ahmad, S. T. et al. (2019). Capicua regulates neural stem cell proliferation and lineage specification through control of Ets factors. *Nature communications*, 10 (1), p.2000.
- Ajuria, L. et al. (2011). Capicua DNA-binding sites are general response elements for RTK signaling in Drosophila. *Development*, 138 (5), pp.915–924.
- 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.
- Babu, G. J. et al. (2000). Phosphorylation of elk-1 by MEK/ERK pathway is necessary for c-fos gene activation during cardiac myocyte hypertrophy. *Journal of molecular and cellular cardiology*, 32 (8), pp.1447–1457.
- Bahrani, S. and Drabløs, F. (2016). Gene regulation in the immediate-early response process. *Advances in biological regulation*, 62, pp.37–49.
- Bailey, T. L. et al. (2015). The MEME Suite. *Nucleic acids research*, 43 (W1), pp.W39–W49.
- Bannister, A. J. and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell research*, 21 (3), pp.381–395.
- Barrett, S. D. et al. (2008). The discovery of the benzhydroxamate MEK inhibitors CI-1040 and PD 0325901. *Bioorganic & medicinal chemistry letters*, 18 (24), pp.6501–6504.
- Bertolotti, E., Malagoli, D. and Franchini, A. (2013). Skin wound healing in different aged *Xenopus laevis*. *Journal of morphology*, 274 (8), pp.956–964.
- Bertrand, N., Médevielle, F. and Pituello, F. (2000). FGF signalling controls the timing of Pax6 activation in the neural tube. *Development*, 127 (22), pp.4837–4843.
- Bettegowda, C. et al. (2011). Mutations in CIC and FUBP1 contribute to human oligodendroglioma. *Science*, 333 (6048), pp.1453–1455.
- Blum, M. and Ott, T. (2018). *Xenopus*: An Undervalued Model Organism to Study and Model Human Genetic Disease. *Cells, tissues, organs*, 205 (5-6), pp.303–313.
- Bradley, I. (2016). *An analysis of N-*Src* splice variants in *Xenopus* development*. MSc by Research, University of York.
- Branney, P. A. et al. (2009). Characterisation of the fibroblast growth factor dependent transcriptome in early development. *PloS one*, 4 (3), p.e4951.
- Brudno, M. et al. (2003). LAGAN and Multi-LAGAN: Efficient Tools for Large-Scale Multiple Alignment of Genomic DNA. *Genome research*, 13 (4), pp.721–731.
- Bunda, S. et al. (2019). CIC protein instability contributes to tumorigenesis in glioblastoma. *Nature communications*, 10 (1), p.661.

- Bunda, S. et al. (2020). c-Src Phosphorylates and Inhibits the Function of the CIC Tumor Suppressor Protein. *Molecular cancer research: MCR*, 18 (5), pp.774–786.
- Cass, D. L. et al. (1997). Wound size and gestational age modulate scar formation in fetal wound repair. *Journal of pediatric surgery*, 32 (3), pp.411–415.
- Cavigelli, M. et al. (1995). Induction of c-fos expression through JNK-mediated TCF/Elk-1 phosphorylation. *The EMBO journal*, 14 (23), pp.5957–5964.
- Chen, R. H. et al. (1996). Phosphorylation of c-Fos at the C-terminus enhances its transforming activity. *Oncogene*, 12 (7), pp.1493–1502.
- Chiu, R. et al. (1988). The c-Fos protein interacts with c-Jun/AP-1 to stimulate transcription of AP-1 responsive genes. *Cell*, 54 (4), pp.541–552.
- Christen, B. and Slack, J. M. (1999). Spatial response to fibroblast growth factor signalling in *Xenopus* embryos. *Development*, 126 (1), pp.119–125.
- Colwell, A. S. et al. (2006). An in vivo mouse excisional wound model of scarless healing. *Plastic and reconstructive surgery*, 117 (7), pp.2292–2296.
- Coulon, V. et al. (2010). A novel mouse c-fos intronic promoter that responds to CREB and AP-1 is developmentally regulated in vivo. *PloS one*, 5 (6), p.e11235.
- Cowell, L. (2019). *Investigation of the role of Capicua in the FGF signalling pathway and the wound response*. MSc by Research, University of York.
- Cruzalegui, F. H., Cano, E. and Treisman, R. (1999). ERK activation induces phosphorylation of Elk-1 at multiple S/T-P motifs to high stoichiometry. *Oncogene*, 18 (56), pp.7948–7957.
- Davidson, L. A., Ezin, A. M. and Keller, R. (2002). Embryonic wound healing by apical contraction and ingression in *Xenopus laevis*. *Cell motility and the cytoskeleton*, 53 (3), pp.163–176.
- 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.
- Dhillon, A. S. et al. (2007). MAP kinase signalling pathways in cancer. *Oncogene*, 26 (22), pp.3279–3290.
- 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.
- Ding, Y. et al. (2017). Genome-wide analysis of dorsal and ventral transcriptomes of the *Xenopus laevis* gastrula. *Developmental biology*, 426 (2), pp.176–187.
- 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.
- Eden, E. et al. (2007). Discovering motifs in ranked lists of DNA sequences. *PLoS computational biology*, 3 (3), p.e39.

- Eden, E. et al. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC bioinformatics*, 10, p.48.
- 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.
- Elkeles, A. et al. (1999). The c-fos proto-oncogene is a target for transactivation by the p53 tumor suppressor. *Molecular and cellular biology*, 19 (4), pp.2594–2600.
- Eming, S. A., Martin, P. and Tomic-Canic, M. (2014). Wound repair and regeneration: mechanisms, signaling, and translation. *Science translational medicine*, 6 (265), p.265sr6.
- Feng, X. et al. (2013). S-nitrosylation of ERK inhibits ERK phosphorylation and induces apoptosis. *Scientific reports*, 3, p.1814.
- Fisher, M. E., Isaacs, H. V. and Pownall, M. E. (2002). eFGF is required for activation of XmyoD expression in the myogenic cell lineage of *Xenopus laevis*. *Development*, 129 (6), pp.1307–1315.
- Fletcher, R. B. and Harland, R. M. (2008). The role of FGF signaling in the establishment and maintenance of mesodermal gene expression in *Xenopus*. *Developmental dynamics: an official publication of the American Association of Anatomists*, 237 (5), pp.1243–1254.
- Forés, M. et al. (2015). Origins of context-dependent gene repression by capicua. *PLoS genetics*, 11 (1), p.e1004902.
- 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.
- Frazer, K. A. et al. (2004). VISTA: computational tools for comparative genomics. *Nucleic acids research*, 32 (Web Server issue), pp.W273–W279.
- Frémin, C. et al. (2015). Functional Redundancy of ERK1 and ERK2 MAP Kinases during Development. *Cell reports*, 12 (6), pp.913–921.
- Frödin, M. and Gammeltoft, S. (1999). Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction. *Molecular and cellular endocrinology*, 151 (1-2), pp.65–77.
- Futran, A. S. et al. (2015). Mapping the binding interface of ERK and transcriptional repressor Capicua using photocrosslinking. *Proceedings of the National Academy of Sciences of the United States of America*, 112 (28), pp.8590–8595.
- Garces de Los Fayos Alonso, I. et al. (2018). The Role of Activator Protein-1 (AP-1) Family Members in CD30-Positive Lymphomas. *Cancers*, 10 (4). [Online]. Available at: doi:10.3390/cancers10040093.
- Garg, A. et al. (2018). FGF-induced Pea3 transcription factors program the genetic landscape for cell fate determination. *PLoS genetics*, 14 (9), p.e1007660.
- Gazon, H. et al. (2017). Hijacking of the AP-1 Signaling Pathway during Development of ATL. *Frontiers in microbiology*, 8, p.2686.
- Gille, H. et al. (1995). ERK phosphorylation potentiates Elk-1-mediated ternary

- complex formation and transactivation. *The EMBO journal*, 14 (5), pp.951–962.
- Glover, J. N. and Harrison, S. C. (1995). Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature*, 373 (6511), pp.257–261.
- Goff, D. J., Nilson, L. A. and Morisato, D. (2001). Establishment of dorsal-ventral polarity of the *Drosophila* egg requires capicua action in ovarian follicle cells. *Development*, 128 (22), pp.4553–4562.
- Grant, C. E., Bailey, T. L. and Noble, W. S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics*, 27 (7), pp.1017–1018.
- Greenberg, M. E. and Ziff, E. B. (1984). Stimulation of 3T3 cells induces transcription of the c-fos proto-oncogene. *Nature*, 311 (5985), pp.433–438.
- Grigoriadis, A. E. et al. (1993). Osteoblasts are target cells for transformation in c-fos transgenic mice. *The Journal of cell biology*, 122 (3), pp.685–701.
- Halazonetis, T. D. et al. (1988). c-Jun dimerizes with itself and with c-Fos, forming complexes of different DNA binding affinities. *Cell*, 55 (5), pp.917–924.
- Hirata, Y. et al. (2016). Interaction between a Domain of the Negative Regulator of the Ras-ERK Pathway, SPRED1 Protein, and the GTPase-activating Protein-related Domain of Neurofibromin Is Implicated in Legius Syndrome and Neurofibromatosis Type 1\*♦. *The Journal of biological chemistry*, 291 (7), pp.3124–3134.
- Hongo, I. and Okamoto, H. (2022). FGF/MAPK/Ets signaling in *Xenopus* ectoderm contributes to neural induction and patterning in an autonomous and paracrine manner, respectively. *Cells & development*, 170, p.203769.
- Hwang, I. et al. (2020). CIC is a critical regulator of neuronal differentiation. *JCI insight*, 5 (9). [Online]. Available at: doi:10.1172/jci.insight.135826.
- 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.
- 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.
- Jin, Y. et al. (2015). EGFR/Ras Signaling Controls *Drosophila* Intestinal Stem Cell Proliferation via Capicua-Regulated Genes. *PLoS genetics*, 11 (12), p.e1005634.
- Kang, H. B. et al. (2005). Basic fibroblast growth factor activates ERK and induces c-fos in human embryonic stem cell line MizhES1. *Stem cells and development*, 14 (4), pp.395–401.
- Karim, F. D. et al. (1990). The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes & development*, 4 (9), pp.1451–1453.
- Karin, M., Liu, Z. g. and Zandi, E. (1997). AP-1 function and regulation. *Current opinion in cell biology*, 9 (2), pp.240–246.
- Karolchik, D. et al. (2004). The UCSC Table Browser data retrieval tool. *Nucleic acids research*, 32 (Database issue), pp.D493–D496.
- 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.

- Kent, W. J. et al. (2002). The human genome browser at UCSC. *Genome research*, 12 (6), pp.996–1006.
- 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, R. et al. (2011). Phosphorylated c-Jun and Fra-1 induce matrix metalloproteinase-1 and thereby regulate invasion activity of 143B osteosarcoma cells. *Biochimica et biophysica acta*, 1813 (8), pp.1543–1553.
- King, M. et al. (2022) *The identification of novel mediators of Fgf/Erk signalling in early Xenopus development*.  
Manuscript in preparation.
- King, M. G. (2018). *Investigating the role of Capicua in mediating FGF transcriptional regulation in X. tropicalis*. phd, University of York. [Online]. Available at: <https://etheses.whiterose.ac.uk/24117/> [Accessed 6 June 2021].
- 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.
- 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.
- Larson, B. J., Longaker, M. T. and Lorenz, H. P. (2010). Scarless fetal wound healing: a basic science review. *Plastic and reconstructive surgery*, 126 (4), pp.1172–1180.
- Leclerc, C., Duprat, A. M. and Moreau, M. (1999). Noggin upregulates Fos expression by a calcium-mediated pathway in amphibian embryos. *Development, growth & differentiation*, 41 (2), pp.227–238.
- Lee, C.-J. et al. (2002). CIC, a member of a novel subfamily of the HMG-box superfamily, is transiently expressed in developing granule neurons. *Molecular Brain Research*, 106 (1), pp.151–156.
- Lee, W., Mitchell, P. and Tjian, R. (1987). Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. *Cell*, 49 (6), pp.741–752.
- Lee, Y. (2020). Regulation and function of capicua in mammals. *Experimental & molecular medicine*, 52 (4), pp.531–537.
- Lemmon, M. A. and Schlessinger, J. (2010). Cell signaling by receptor-tyrosine kinases. *Cell*, 141 (7), pp.1117–1134.
- Li, C. et al. (2007). Dusp6 (Mkp3) is a negative feedback regulator of FGF-stimulated ERK signaling during mouse development. *Development*, 134 (1), pp.167–176.
- 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.
- Lin, C. W., Georgescu, H. I. and Evans, C. H. (1993). The role of AP-1 in matrix metalloproteinase gene expression. *Agents and actions*, 39 Spec No, pp.C215–C218.
- Lin, X. et al. (1999). Heparan sulfate proteoglycans are essential for FGF receptor signaling during *Drosophila* embryonic development. *Development*, 126 (17), pp.3715–3723.

- Lorenz, H. P. et al. (1993). Fetal wound healing. The ontogeny of scar formation in the non-human primate. *Annals of surgery*, 217 (4), pp.391–396.
- Lowenstein, E. J. et al. (1992). The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell*, 70 (3), pp.431–442.
- Lowery, L. A. and Sive, H. (2004). Strategies of vertebrate neurulation and a re-evaluation of teleost neural tube formation. *Mechanisms of development*, 121 (10), pp.1189–1197.
- Malone-Povolny, M. J., Maloney, S. E. and Schoenfisch, M. H. (2019). Nitric Oxide Therapy for Diabetic Wound Healing. *Advanced healthcare materials*, 8 (12), p.e1801210.
- 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.
- Mayor, C. et al. (2000). VISTA : visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics*, 16 (11), pp.1046–1047.
- Miller, A. D., Curran, T. and Verma, I. M. (1984). c-fos protein can induce cellular transformation: a novel mechanism of activation of a cellular oncogene. *Cell*, 36 (1), pp.51–60.
- Min, H. et al. (1998). Fgf-10 is required for both limb and lung development and exhibits striking functional similarity to Drosophila branchless. *Genes & development*, 12 (20), pp.3156–3161.
- Misra, R. P. et al. (1991). The serum response factor is extensively modified by phosphorylation following its synthesis in serum-stimulated fibroblasts. *Molecular and cellular biology*, 11 (9), pp.4545–4554.
- Mittelstadt, M. L. and Patel, R. C. (2012). AP-1 mediated transcriptional repression of matrix metalloproteinase-9 by recruitment of histone deacetylase 1 in response to interferon  $\beta$ . *PloS one*, 7 (8), p.e42152.
- Mohammadi, M. et al. (1997). Structures of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors. *Science*, 276 (5314), pp.955–960.
- Monsonogo-Ornan, E. et al. (2012). Matrix metalloproteinase 9/gelatinase B is required for neural crest cell migration. *Developmental biology*, 364 (2), pp.162–177.
- Morriss-Kay, G. and Tuckett, F. (1985). The role of microfilaments in cranial neurulation in rat embryos: effects of short-term exposure to cytochalasin D. *Journal of embryology and experimental morphology*, 88, pp.333–348.
- Murphy, L. O. et al. (2002). Molecular interpretation of ERK signal duration by immediate early gene products. *Nature cell biology*, 4 (8), pp.556–564.
- Nachmany, A. et al. (2006). Neural tube closure depends on nitric oxide synthase activity. *Journal of neurochemistry*, 96 (1), pp.247–253.
- Nentwich, O. et al. (2009). Downstream of FGF during mesoderm formation in Xenopus: the roles of Elk-1 and Egr-1. *Developmental biology*, 336 (2), pp.313–326.
- Nieuwkoop P D And Faber. (1967). *Normal Table of Xenopus laevis*. Amsterdam, The

Netherlands: North Holland Publishing Company.

Nikitovic, D., Holmgren, A. and Spyrou, G. (1998). Inhibition of AP-1 DNA binding by nitric oxide involving conserved cysteine residues in Jun and Fos. *Biochemical and biophysical research communications*, 242 (1), pp.109–112.

Nodder, S. and Martin, P. (1997). Wound healing in embryos: a review. *Anatomy and embryology*, 195 (3), pp.215–228.

Nye, J. A. et al. (1992). Interaction of murine ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes & development*, 6 (6), pp.975–990.

Okazaki, K. and Sagata, N. (1995). The Mos/MAP kinase pathway stabilizes c-Fos by phosphorylation and augments its transforming activity in NIH 3T3 cells. *The EMBO journal*, 14 (20), pp.5048–5059.

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.

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

Papagianni, A. et al. (2018). Capicua controls Toll/IL-1 signaling targets independently of RTK regulation. *Proceedings of the National Academy of Sciences of the United States of America*, 115 (8), pp.1807–1812.

Partanen, J., Vainikka, S. and Alitalo, K. (1993). Structural and functional specificity of FGF receptors. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 340 (1293), pp.297–303.

Patro, R. et al. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature methods*, 14 (4), pp.417–419. [Accessed 29 April 2022].

Pézeron, G. et al. (2008). Rasl11b knock down in zebrafish suppresses one-eyed-pinhead mutant phenotype. *PloS one*, 3 (1), p.e1434.

Pimentel, H. and McGee, W. (2021). *sleuth: Tools for investigating RNA-Seq*. [Online]. Available at: <https://github.com/pachterlab/sleuth>.

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.

Qi, T. et al. (2020). Function and regulation of the PEA3 subfamily of ETS transcription factors in cancer. *American journal of cancer research*, 10 (10), pp.3083–3105.

Reiss, M. J. et al. (2010). Matrix metalloproteinase-9 delays wound healing in a murine wound model. *Surgery*, 147 (2), pp.295–302.

Renz, M. et al. (1985). Regulation of c-fos transcription in mouse fibroblasts: identification of DNase I-hypersensitive sites and regulatory upstream sequences. *The EMBO journal*, 4 (13B), pp.3711–3716.

Santarpia, L., Lippman, S. M. and El-Naggar, A. K. (2012). Targeting the MAPK-RAS-RAF signaling pathway in cancer therapy. *Expert opinion on therapeutic targets*, 16 (1), pp.103–119.

- Schulte-Merker, S. and Smith, J. C. (1995). Mesoderm formation in response to Brachyury requires FGF signalling. *Current biology: CB*, 5 (1), pp.62–67.
- Seger, R. and Krebs, E. G. (1995). The MAPK signaling cascade. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 9 (9), pp.726–735.
- Sehnal, D. et al. (2021). Mol\* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic acids research*, 49 (W1), pp.W431–W437. [Accessed 3 May 2022].
- Sharrocks, A. D. (2001). The ETS-domain transcription factor family. *Nature reviews. Molecular cell biology*, 2 (11), pp.827–837.
- Shen, L. and at Mount Sinai, I. S. of M. (2020). *GeneOverlap: Test and visualize gene overlaps*. [Online]. Available at: <http://shenlab-sinai.github.io/shenlab-sinai/>.
- Shyu, A. B., Greenberg, M. E. and Belasco, J. G. (1989). The c-fos transcript is targeted for rapid decay by two distinct mRNA degradation pathways. *Genes & development*, 3 (1), pp.60–72.
- Simón-Carrasco, L. et al. (2018). The Capicua tumor suppressor: a gatekeeper of Ras signaling in development and cancer. *Cell cycle*, 17 (6), pp.702–711.
- Small, C. D. and Crawford, B. D. (2016). Matrix metalloproteinases in neural development: a phylogenetically diverse perspective. *Neural Regeneration Research*, 11 (3), pp.357–362.
- 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.
- Suga, A., Hikasa, H. and Taira, M. (2006). Xenopus ADAMTS1 negatively modulates FGF signaling independent of its metalloprotease activity. *Developmental biology*, 295 (1), pp.26–39.
- Tambalo, M. et al. (2020). Enhancer activation by FGF signalling during otic induction. *Developmental biology*, 457 (1), pp.69–82.
- Timothy L. Bailey and Charles E. Grant. (2021). SEA: Simple Enrichment Analysis of motifs. *BioRxiv*. [Online]. Available at: doi:10.1101/2021.08.23.457422.
- Tsarouhas, V., Yao, L. and Samakovlis, C. (2014). Src kinases and ERK activate distinct responses to Stitcher receptor tyrosine kinase signaling during wound healing in Drosophila. *Journal of cell science*, 127 (Pt 8), pp.1829–1839.
- 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.
- Ullrich, A. and Schlessinger, J. (1990). Signal transduction by receptors with tyrosine kinase activity. *Cell*, 61 (2), pp.203–212.
- Ünal, E. B., Uhlitz, F. and Blüthgen, N. (2017). A compendium of ERK targets. *FEBS letters*, 591 (17), pp.2607–2615.
- Velnar, T., Bailey, T. and Smrkolj, V. (2009). The Wound Healing Process: An Overview of the Cellular and Molecular Mechanisms. *The Journal of international medical research*, 37 (5), pp.1528–1542.

- Wakioka, T. et al. (2001). Spred is a Sprouty-related suppressor of Ras signalling. *Nature*, 412 (6847), pp.647–651.
- Webster, K. A., Discher, D. J. and Bishopric, N. H. (1994). Regulation of fos and jun immediate-early genes by redox or metabolic stress in cardiac myocytes. *Circulation research*, 74 (4), pp.679–686.
- Wei, G.-H. et al. (2010). Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *The EMBO journal*, 29 (13), pp.2147–2160.
- Weissmann, S. et al. (2018). The Tumor Suppressor CIC Directly Regulates MAPK Pathway Genes via Histone Deacetylation. *Cancer research*, 78 (15), pp.4114–4125.
- Werner, S. et al. (1992). Differential splicing in the extracellular region of fibroblast growth factor receptor 1 generates receptor variants with different ligand-binding specificities. *Molecular and cellular biology*, 12 (1), pp.82–88.
- Wood, W. et al. (2002). Wound healing recapitulates morphogenesis in Drosophila embryos. *Nature cell biology*, 4 (11), pp.907–912.
- Xie, J. et al. (2005). Transcript profiling of immediate early genes reveals a unique role for activating transcription factor 3 in mediating activation of the glycoprotein hormone alpha-subunit promoter by gonadotropin-releasing hormone. *Molecular endocrinology*, 19 (10), pp.2624–2638.
- Xie, Y. et al. (2020). FGF/FGFR signaling in health and disease. *Signal transduction and targeted therapy*, 5 (1), p.181.
- Yang, L. et al. (2016). Minibrain and Wings apart control organ growth and tissue patterning through down-regulation of Capicua. *Proceedings of the National Academy of Sciences of the United States of America*, 113 (38), pp.10583–10588.
- Yang, R. et al. (2017). Cic Loss Promotes Gliomagenesis via Aberrant Neural Stem Cell Proliferation and Differentiation. *Cancer research*, 77 (22), pp.6097–6108.
- Yip, S. et al. (2012). Concurrent CIC mutations, IDH mutations, and 1p/19q loss distinguish oligodendrogliomas from other cancers. *The Journal of pathology*, 226 (1), pp.7–16.
- Yoon, S. and Seger, R. (2006). The extracellular signal-regulated kinase: multiple substrates regulate diverse cellular functions. *Growth factors*, 24 (1), pp.21–44.
- Yoo, S. K. et al. (2012). Early redox, Src family kinase, and calcium signaling integrate wound responses and tissue regeneration in zebrafish. *The Journal of cell biology*, 199 (2), pp.225–234.
- Yoshii, Y. et al. (2005). Wound healing ability of *Xenopus laevis* embryos. I. Rapid wound closure achieved by bisecting half embryos. *Development, growth & differentiation*, 47 (8), pp.553–561.
- 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.