

Identification of IRF6 Downstream Target Genes in Zebrafish.

MA YANKUN

(Bachelor of Science, Zhejiang University)

**A THESIS SUBMITTED
FOR THE DEGREE OF MASTER OF SCIENCE
DEPARTMENT OF PAEDIATRICS
NATIONAL UNIVERSITY OF SINGAPORE**

2013

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety.

I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Ma Yankun

August 1, 2013

Acknowledgement

Foremost, I would like to express my sincere gratitude to my supervisor, Prof. Samuel Chong, for his continuous support of my MSc research and personal development, for his advice, patience, enthusiasm, and immense knowledge. His guidance helped me through this wonderful journey of research and learning. I am very grateful for the opportunity to work with him and it has been my privilege to learn from him. I could not have imagined having a better advisor and mentor for my MSc study.

Besides my supervisor, I would like to thank the rest of my thesis advisory committee: Prof. Heng Chew-Kiat and Prof. Lee Guat Lay Caroline for their encouragement, insightful comments, and precious advice.

Sincere gratitude also goes to Dr Felicia Cheah, for her effort in initiating the project, and the valuable suggestions regarding both my MSc study and living in Singapore. I would also like to express my utmost appreciation to all members working under Prof. Sam's group, for their friendship, support and effort in making the whole group feel like a big family: Mr. Arnold Tan, Ms. Chen Min, Ms. Indhu Shree, Mr. Eugene Saw, Ms. Zhao Mingjue, Ms. Phang Guiping, Ms Mulias and many others.

Financial support from the National University of Singapore is sincerely acknowledged.

I am deeply thankful to my family for their love, support, and sacrifice. Without their support, this thesis would never have been written.

Summary

Gastrulation is an important step in early embryogenesis. It involves a series of coordinated cell movements to organize the germ layers and establish the major body axes of the embryo (Lepage and Bruce, 2010; Wang and Steinbeisser, 2009). During the process of studying *interferon regulatory factor 6 (IRF6)* which is known to be involved in syndromic oral clefting, we found out a drastic and prominent knockdown phenotype leading by Morpholino targeting at the splice junction of exon 3 and intron 3 of *irf6* pre-mRNA (E3I3) in zebrafish that strongly suggests a critical role of Irf6 in proper gastrulation and early embryogenesis. In this study, we profiled the transcriptome of embryos lack of functional Irf6 leading by the injection of E3I3 using the Agilent zebrafish gene expression microarray. We identified and characterized *cyr61* and *mapkapk3* as target genes of Irf6 at gastrulation stage in zebrafish. The findings gathered from this study will provide novel insights into how IRF6 normally function in vertebrate embryogenesis and also contribute new knowledge into understanding gastrulation process.

Table of Contents

Contents

| | |
|--|----|
| Declaration..... | 1 |
| Acknowledgement | 2 |
| Summary | 4 |
| Table of Contents..... | 5 |
| List of Tables | 8 |
| List of Figures..... | 9 |
| Abbreviations..... | 10 |
| Chapter I: Introduction..... | 11 |
| 1.1 Early development of the zebrafish..... | 11 |
| 1.1.1 Zebrafish as a model organism for the study of vertebrate development..... | 11 |
| 1.1.2 Epiboly of zebrafish | 13 |
| 1.1.3 Gastrulation of zebrafish | 14 |
| 1.2 Role of IRF6 in development..... | 16 |
| 1.2.1 Interferon Regulatory Factor | 16 |
| 1.2.2 IRF6 is important in early development in zebrafish and Xenopus | 19 |
| 1.2.3 IRF6 and oral clefting..... | 23 |
| 1.2.4 IRF6 in mouse development..... | 25 |

| | |
|---|----|
| 1.3 Other functions and regulation of IRF6 | 26 |
| 1.3.1 IRF6 functions as a transcriptional factor | 26 |
| 1.3.2 IRF6 and cell proliferation and differentiation..... | 26 |
| 1.3.3 Regulation of IRF6..... | 28 |
| 1.4 Microarray..... | 29 |
| 1.5 Objectives of the project | 30 |
| Chapter II: Materials and Methods | 31 |
| 2.1: Ethics statement/ fish strain | 31 |
| 2.2: Morpholino injection | 31 |
| 2.3: Total RNA extraction from fish embryos | 32 |
| 2.4: Microarray sample preparation and hybridization..... | 33 |
| 2.5: Microarray analysis and statistics | 33 |
| 2.6: Semi-quantitative reverse-transcription PCR: | 34 |
| 2.7: pcDNA/His-Irf6-FL and pcDNA/His-Irf6-E3I3 plasmid construction | 35 |
| 2.7.1 Amplification of full length and truncated Irf6..... | 35 |
| 2.7.2 Plasmid digestion..... | 36 |
| 2.7.3 Creating blunt end | 36 |
| 2.7.4 Ligation..... | 36 |
| 2.7.5 Transformation | 37 |
| 2.8: <i>In vitro</i> protein expression | 37 |

| | |
|--|----|
| 2.8.1: Generation of DNA templates for full length and truncated Irf6 protein expression | 37 |
| 2.8.2: Protein expression using TNT wheat germ expression system | 38 |
| 2.9: Protein purification | 38 |
| 2.10: Western blotting..... | 39 |
| 2.11: Electrophoretic mobility shift assay (EMSA)..... | 41 |
| Chapter III: Results | 43 |
| 3.1: Genome-wide gene profiling microarray analysis of the E3I3 injected embryos..... | 43 |
| 3.2: Gene ontology study of differentiated expressed genes in E3I3 MO-injected embryos | 51 |
| 3.3: Microarray differential gene expression validation | 53 |
| 3.4: <i>cyr61</i> and <i>mapkapk3</i> are direct downstream targets of Irf6..... | 55 |
| 3.5: Preliminary morphology study of <i>cyr61</i> and <i>mapkapk3</i> MO blocked embryos | 57 |
| Chapter IV: Discussion | 62 |
| 4.1: Interpretation of expression profile of E3I3 MO-injected embryos: Irf6 functions as an essential transcriptional factor during early development..... | 62 |
| 4.2 The multi-function role of IRF6..... | 64 |
| 4.3 <i>cyr61</i> and <i>mapkapk3</i> are direct downstream targets of Irf6..... | 65 |
| 4.4 Conclusion and future work..... | 68 |
| Reference | 70 |

List of Tables

| Table No: | Page |
|--|-------|
| 1. A summary of IRF family member functions | 18 |
| 2. SDS-PAGE recipe | 40 |
| 3. Antibody used in western blotting | 41 |
| 4. Microarray gene expression analysis: E3I3 MO-injected embryos vs mock-MO injected embryos. | 45-49 |

List of Figures

| Figure No. | | Page |
|------------|--|-------|
| 1 | Structure of zebrafish embryo and progression of epiboly | 14 |
| 2 | The gastrulation period | 15 |
| 3 | Phylogenetic analysis of <i>irf</i> gene family and alignment of the predicted proteins from different species | 19-20 |
| 4 | Aberrant <i>irf6</i> transcript variants can cause early embryonic lethality | 23 |
| 5 | Genes differentially regulated by Irf6 during early embryogenesis | 50 |
| 6 | Gene ontology analysis of differentially expressed genes | 52-53 |
| 7 | Validation of differentially expressed <i>cyr61</i> and <i>mapkapk3</i> | 54 |
| 8 | <i>cyr61</i> and <i>mapkapk3</i> are directly bound by Irf6 and E3I3 truncated protein | 56 |
| 9 | <i>mapkapk3</i> MO-injected embryos show defects in the epithelial layer | 58 |
| 10 | <i>mapkapk3</i> MO does not cause a lethal phenotype for embryos | 59 |
| 11 | <i>cyr61</i> MO-injected embryos show gastrulation defects. | 60 |
| 12 | One-quarter of <i>cyr61</i> MO-injected embryos die after 24hours. | 61 |

Abbreviations

| | |
|------------|--|
| CL/P | Cleft lip with or without the palate |
| CPO | Cleft palate only |
| C-terminal | carboxyl-terminus |
| Cyr61 | Cysteine-rich 61 |
| DBD | DNA binding domain |
| E3I3 | Morpholino targeting at the splice junction of exon 3 and intron 3 of <i>irf6</i> pre-mRNA |
| EMSA | Electrophoretic mobility shift assay |
| EVL | Enveloping layer |
| GO | Gene ontology |
| IAD | IRF-associated domain |
| IFN | Interferon |
| IRF | Interferon regulatory factor |
| ISRE | Interferon-sensitive response element |
| Mapkapk3 | mitogen-activated protein kinase-activated protein kinase 3 |
| MH2 | Mad-homology 2 |
| MO | Morpholino |
| N-terminal | Amino-terminus |
| PID | Protein interaction domain |
| PPS | Popliteal pterygium syndrome |
| SCC | Squamous cell carcinoma |
| VWS | Van der Woude syndrome |
| YSL | Yolk syncytial layer |

Chapter I: Introduction

1.1 Early development of the zebrafish

1.1.1 Zebrafish as a model organism for the study of vertebrate development

With the gradual understanding of the mechanisms involved in development, developmental biology has become one of the most exciting and fast-growing fields of biology. As a complex branch of biology, understanding developmental processes requires combining information from molecular biology, physiology, anatomy, cancer research and even evolutionary studies (Gilbert, 1999). Hence, many discoveries that originated from investigating development defects, such as the Wnt (Klaus and Birchmeier, 2008), Hedgehog (Gupta et al., 2010), and Notch families (Bray, 2006), are now also known to play significant roles in cancer or are linked to other human diseases. Animal models are widely used in developmental studies. Among them, zebrafish is a well established animal model used especially to study early stage developmental processes.

The zebrafish (*Danio rerio*) belongs to the family *Cyprinidae* (Detrich et al., 1999), and serves a useful role in bridging the gap between *Drosophila/Caenorhabditis elegans* and mouse/human genetics. As early as the 1930s, this tropical fish was being used as a classical developmental and embryological model (Roosen-Runge, 1937). Beginning in the 1980s, the development of genetic techniques enabled the use of zebrafish for studies of developmental biology (Lieschke and Currie, 2007; Streisinger et al., 1981). The advent of

large-scale mutagenic screens (Amsterdam et al., 1999) cemented the zebrafish's role as an important vertebrate model in developmental biology.

Advantages of the zebrafish include its small size (up to 6 cm), short generation time (2~3 months), external fertilization, and large egg clutches (100-200 eggs per mating). Zebrafish embryos are transparent throughout early development, providing easy visual access to all developmental stages and facilitating embryological experiments and morphological screening (Detrich et al., 1999). Aside from these advantages, technically, the methodologies routinely applied to *Xenopus* embryos can also be successfully performed on zebrafish (Detrich et al., 1999; Eisen, 1996). Forward-genetic screening and reverse-genetic transient morpholino knockdowns allow for investigation of gene function. Nowadays precise genome editing becomes available by several methods, such as TALEN and CRISPR approaches (Auer et al., 2014; Bedell et al., 2012). With the availability of these techniques, we are able to use the zebrafish to model almost any genetic mutation that causes diseases in human.

The zebrafish genome has been sequenced and mapped. The genetic map has been continually improving, and currently more than 2000 microsatellite markers (Knapik et al., 1998; Shimoda et al., 1999) and more than 26,000 protein-coding genes have been defined (Collins et al., 2012) for the 1.412 gigabases (Gb) genome (Howe et al., 2013). The information is available on ZFIN, NCBI and ENSEMBL websites, further facilitating research using zebrafish.

1.1.2 Epiboly of zebrafish

Epiboly was first described in the teleost fish *Cyprinus* by von Baer in 1835 as the overgrowth of the yolk by the blastoderm (Betchaku and Trinkaus, 1978). The term epiboly has now been defined as the thinning and spreading of a sheet of cells to cover the embryo during gastrulation (Gilbert 2003).

Before the initiation of epiboly, the embryo is organized into three layers (Fig 1.): the enveloping layer (EVL), a single-layer epithelium; the deep cells layer, which eventually gives rise to embryonic tissues; and the yolk syncytial layer (YSL), an extra-embryonic syncytium populating the interface between the yolk and deep cells (Lepage and Bruce, 2010). When epiboly starts, the yolk cell domes and deep cells move radially outwards, forming a cap of cells over the yolk. With the progression of epiboly, the thinning blastoderm (EVL and deep cells) spreads vegetally, expanding its surface area to cover the yolk cell, past the equator of the embryo. When the embryo reaches 50% epiboly, the blastoderm begins to converge dorsally. In the end, the deep cells, EVL and YSL move towards the vegetal pole in a coordinated manner, eventually closing the blastopore (Lepage and Bruce, 2010).

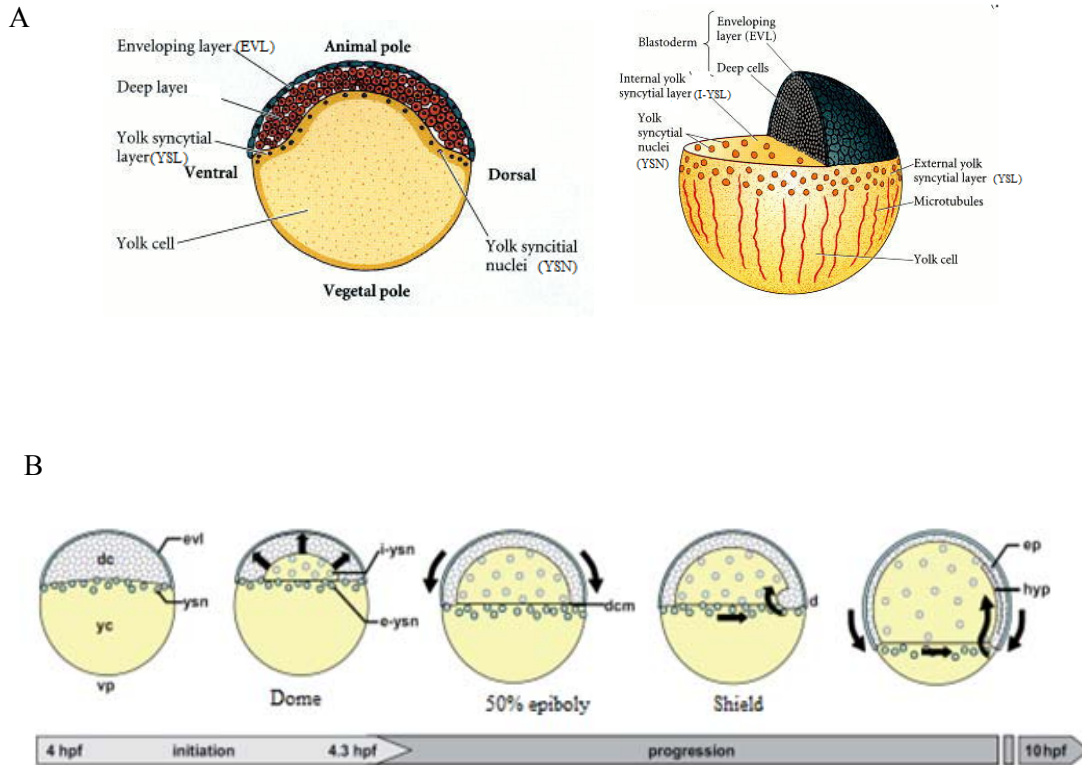


Figure1: Structure of zebrafish embryo and progression of epiboly

(A) Epiboly is organized into 3 layers: enveloping layer (EVL), yolk syncytial layer (YSL) and deep cells (Taken from Gilbert 2000).

(B) Schematic depiction of epiboly initiation and progression in the zebrafish embryo (Taken from Lepage and Bruce 2010).

1.1.3 Gastrulation of zebrafish

Gastrulation is a morphogenetic process that results in the formation and spatial separation of the embryonic germ layers: ectoderm, mesoderm, and endoderm and to sculpt the body plan (Rohde and Heisenberg, 2007). The gastrulation process includes three major features: epiboly, internalization and convergent extension (Warga and Kimmel, 1990), and these movements of the cells during gastrulation are conserved within vertebrates (Solnica-

Krezel, 2005). In zebrafish, the gastrula period extends from 5.5 hour to about 10 hour (Figure 2). At 50% epiboly (6 hour post-fertilization (hpf)), the rim of the blastoderm thickens to a bilayered germ-ring, which marks the beginning of gastrulation (H. William Dietrich, 1999). The inner layer or hypoblast forms the embryonic mesoderm and endoderm, whereas the outer layer or epiblast forms the embryonic ectoderm (Warga and Kimmel, 1990). Following gastrulation, cells in the organism are either organized into sheets of connected cells or as isolated cells, and the fate of these cells is determined (Brian K. Hall, 1998).

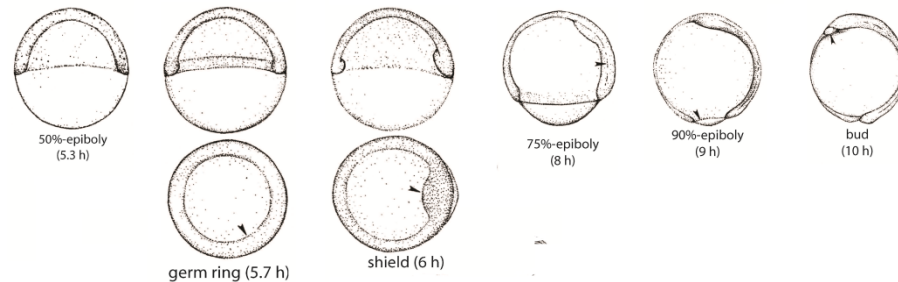


Figure 2: The gastrulation period.

Gastrulation starts at 50% epiboly stage, including three major features: epiboly, internalization of and convergent extension, results in the formation of ectoderm, mesoderm, and endoderm (Adapted and modified from Kimmel, Ballard et al. 1995).

To date, a number of genes have been shown to be involved in gastrulation in zebrafish, such as FoxH (Pei et al., 2007) and Mapkapk2 (Holloway et al., 2009). Among these genes, IRF6 is considered critical to early development since blocking IRF6 function causes a lethal phenotype during gastrulation (Sabel et al., 2009).

1.2 Role of IRF6 in development

1.2.1 Interferon Regulatory Factor

The interferon regulatory factor (IRF) family comprises nine transcription factors: IRF1, IRF2, IRF3, IRF4 (also known as LSIRF, PIP or ICSAT), IRF5, IRF6, IRF7, IRF8 (also known as ICSPB) and IRF9 (also known as ISGF3 γ) (Lohoff and Mak, 2005; Taniguchi et al., 2001).

All IRF proteins possess a highly conserved N-terminal DNA binding domain (DBD) of approximately 120 amino acids that forms a helix-turn-helix motif. This DBD recognizes a consensus DNA sequence - the interferon-stimulated response element (ISRE; ^A/_GNGAAANNGAAACT, also known as IRF-E) (Taniguchi et al., 2001). By contrast, the C-terminal regions of IRFs are less conserved protein interaction domains (PID) which mediate interactions with other protein factors thereby conferring specific activities of each IRF (Savitsky et al., 2010). All IRFs except IRF1 and IRF2 possess a PID showing homology to the Mad-homology 2 (MH2) domains of the Smad family (Mamane et al., 1999), whereas IRF1 and IRF2 share an IRF-associated domain 2 (IAD2) (Taniguchi et al., 2001). These C-terminal regions might function as regulatory regions, and specific protein-protein interaction mediated by these PIDs may determine whether the IRF protein functions as a transcriptional activator or repressor (Savitsky et al., 2010).

With the gene-disruption studies of most of the *IRF* genes being carried out, the functions of IRFs are becoming clearer. Through interaction with family members or other

transcription factors, IRFs have distinct roles in the regulation of host defense, such as innate and adaptive immune responses and the development of immune cells (Taniguchi et al., 2001). The functions of the IRFs have also expanded to distinct roles in biological processes such as pathogen response, cytokine signaling, cell growth regulation, oncogenesis and hematopoietic development (Table 1) (Tamura et al., 2008) .

Table 1: A summary of IRF family member functions

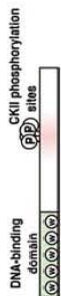
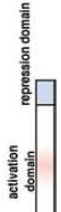
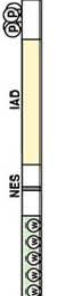
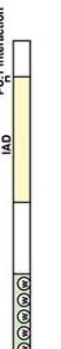





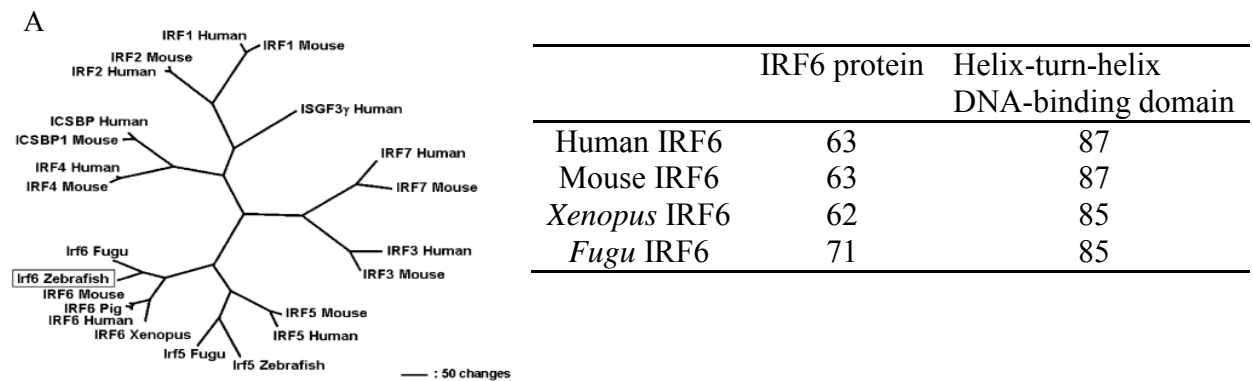
| Function domains of IRFs | | Expression | Roles in immune responses immune cells and other cells | Role in cell growth |
|--------------------------|---|--|--|---|
| IRF-1 |  | Constitutive and IFN-inducible in various cell types | <ul style="list-style-type: none"> Stimulates expression of IFN-inducible gene | <ul style="list-style-type: none"> Suppresses onco gene-induced transformation (Lysyl oxidase) Required for DNA damage-induced growth arrest Required for DNA damage-induced apoptosis |
| | | Inducible by DNA damage at transcriptional and posttranslational levels | <ul style="list-style-type: none"> Binds to MyD88 and enhances TLR-dependent gene induction in IFN-γ treated cells | |
| IRF-2 |  | Constitutive and IFN-inducible in various cell types | | <ul style="list-style-type: none"> Attenuates type I IFN responses by antagonizing IRF1 and IRF9 In some cases, cooperates with IRF1 to activate transcription |
| | | | | |
| IRF-3 |  | Constitutive in various cell types | <ul style="list-style-type: none"> Induces type I IFNs and chemokines upon virus infection, TLR stimulation and cytosolic DNA stimulation | <ul style="list-style-type: none"> Stimulates apoptosis in Mϕs upon bacterial infection May promote DNA damage-induced apoptosis |
| | | | | |
| IRF-4 |  | Constitutive in B cells, M ϕ s, CD11b+ DCs and pDCs | <ul style="list-style-type: none"> Binds to MyD88 and negatively regulates TLR-dependent induction of proinflammatory cytokine genes | <ul style="list-style-type: none"> May possess oncogenic potential |
| | | Inducible by antigen stimulation in T cells and by TLR signaling in M ϕ s | | |
| IRF-5 |  | Constitutive in B cells and DCs | <ul style="list-style-type: none"> Binds to MyD88 and positively regulates TLR-dependent induction of proinflammatory cytokine genes | <ul style="list-style-type: none"> Suppresses onco gene-induced transformation |
| | | Inducible by type I IFNs, TLR signaling and DNA damage in various cells | <ul style="list-style-type: none"> Induces type I IFNs and proinflammatory cytokines upon virus infection | <ul style="list-style-type: none"> Required for DNA damage-induced apoptosis |
| IRF-6 |  | Constitutive in skin | <ul style="list-style-type: none"> Unknown, but translocates from the cytoplasm to the nucleus upon poly(I:C) treatment | |
| | | | | |
| IRF-7 |  | Constitutive in B cells, pDCs and monocytes | <ul style="list-style-type: none"> Binds to MyD88 and induces type I IFNs upon TLR signaling | |
| | | Inducible by type I IFNs in various cell types | <ul style="list-style-type: none"> Induces type I IFNs upon virus infection | |
| IRF-8 |  | Constitutive in B cells, M ϕ s, CD8+ DCs and pDCs | <ul style="list-style-type: none"> Binds to TRAF6 and is required for TLR9 signaling in DCs | <ul style="list-style-type: none"> Inhibits myeloid cell growth Promotes apoptosis in myeloid cells Its absence leads to CML-like disease |
| | | Inducible by IFN- γ in M ϕ s and by antigen stimulation in T cells | <ul style="list-style-type: none"> Stimulates IFN-γ and PAMP-inducible genes | |
| IRF-9 |  | Constitutive and inducible by IFN- γ in various cell types | <ul style="list-style-type: none"> Binds to STAT1 and STAT2 to form ISGF3 and stimulates type I IFN-inducible genes | <ul style="list-style-type: none"> Mediates type I IFN induction of p53 (p53) |
| | | | | |

Table 1: A summary of IRF family members functions. Adapted and modified from Tamura et al., 2008; Savitsky, D et al., 2010; Honda K. et al., 2006 and Marmare Y. 1999

1.2.2 IRF6 is important in early development in zebrafish and *Xenopus*

Among these IRF proteins, IRF6 is a unique member as it is not involved in immune regulatory pathways. Instead, mutations in *IRF6* have been identified as causative of the allelic autosomal dominant clefting disorders Van der Woude syndrome (VWS; OMIM no. 119300) and popliteal pterygium syndrome (PPS; OMIM no. 119500) (Kondo et al., 2002). A more exciting finding was the observation that blocking IRF6 function in zebrafish and *Xenopus* causes a lethal phenotype during gastrulation, indicating a critical role in early vertebrate development (Sabel et al., 2009). Even though its function is not related to regulation of host defense, IRF6 still shares a highly-conserved N-terminal helix-turn-helix DNA-binding domain and a less conserved C-terminal protein-binding domain. A comparison of the protein sequences of IRF6 in human, mouse, *Xenopus*, zebrafish and Fugu reveals that their DNA-binding domains are highly conserved among all five species (Figure.3).



In zebrafish, *irf6* transcript is deposited as a maternal transcript (Ben et al., 2005). During the gastrulation period (~7-9 hpf), *irf6* expression is concentrated in the forerunner cells. From the bud stage to the 3-somite stage (~10–11 hpf), *irf6* is highly expressed in the Kupffer's vesicle and at the 14-somite (16 hpf), expression is observed in the otic placode. From 2-5 day post fertilization (dpf), *irf6* is expressed in the esophagus, pharynx, and mouth, as well as in the pharyngeal arches (Ben et al., 2005).

Gene function can be knocked down by using ATG-translation blocking morpholinos (MOs), which are antisense 25-base oligo nucleotides that target and bind sequences about 25 bases after the start codon, thus blocking translation initiation of transcripts (Summerton, 1999). *Irf6* knockdowns have produced grossly normal embryos without defects in skin, pectoral fins, or craniofacial cartilage after 4 days (Sabel et al., 2009). As *Irf6* is a maternal transcript and the abundant maternal *Irf6* protein may compensate for the reduction of zygotic *Irf6* expression, translation-blocking MOs may have limited effectiveness. Thus, a dominant negative *irf6* mRNA containing only the DNA binding domain of *irf6* (*irf6*DBD) was introduced into 1-2 cell stage zebrafish embryos to block translation of maternal *irf6* transcripts. With the existence of the *irf6*DBD, the embryonic development stalled and the embryo ruptured at 90% epiboly (~ 9hpf) (Sabel et al., 2009). Embryos injected with a lower dose of *irf6*DBD mRNA survived, and showed short pectoral fins, blistered skin and smaller, more disorganized cartilage elements of the craniofacial skeleton at 3 dpf (Sabel et al., 2009). The latter phenotypes are consistent with the *Irf6*-null mouse, which had shorter forelimbs, abnormal skin, and craniofacial defects (Ingraham et al., 2006; Richardson et al., 2006).

Independently, our group also generated an antisense MO (E3I3-MO) targeting the exon 3 - intron 3 splice junction of *irf6* pre-mRNA to investigate the role of zygotic *irf6* in early embryogenesis. Embryos injected with normal (1mM) or low (0.1mM) dose of E3I3-MO exhibited 100% lethality at the gastrula stage (Figure 4) (unpublished data). Time-lapse analysis of the injected embryos revealed developmental arrest at the epiboly stage (5 hpf), leading to embryonic rupture near the animal pole and spillage of the deep cells at around 9 hpf. The arrest of epiboly movement and subsequent rupturing of these embryos are reminiscent of the phenotypes described in Sabel et al. (2009). Both *irf6*DBD and E3I3-MO are thought to inhibit transcriptional activation of downstream target genes, some of which may play important roles in zebrafish early development.

In *Xenopus*, where two paralogues of *irf6* with identical expression patterns exist, *irf6* is maternally expressed, with later expression surrounding the blastopore and in the tailbud blastema (Hatada et al., 1997; Klein et al., 2002). *Irf6*-depleted embryos are delayed in gastrulation and exhibit a blastopore closure defect. Besides, the depleted embryos also fail to elongate fully, and exhibit epidermal and head defects (Sabel et al., 2009). Injection of zebrafish *irf6*DBD mRNA into *Xenopus* embryos also caused rupture of the embryo near the animal pole.

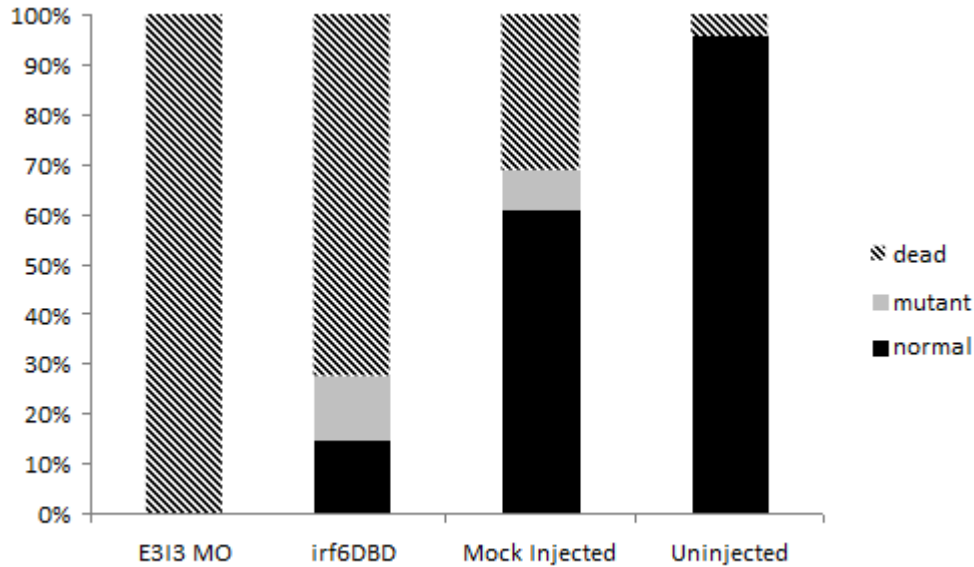


Figure 4: Aberrant *irf6* transcript variants can cause early embryonic lethality of zebrafish. Percentage of zebrafish embryos at 24 hpf of E3I3 MO, *irf6* DBD, Mock MO injected and uninjected sample class that were normal (black), or mutant (head and tail defects) (grey), or dead (striped) (unpublished data of our group).

1.2.3 IRF6 and oral clefting

Human *IRF6* mutations are responsible for Van der Woude syndrome (VWS) and popliteal pterygium syndrome (PPS), which show different degrees of cleft lip, cleft palate, lip pits, skin folds, syndactyly and oral adhesions (Kondo et al., 2002). Autosomal dominant Van der Woude syndrome (VWS) (OMIM no.119300) is the most common syndromic form of clefting, which is characterized by presence of bilateral lower lip pits and hypodontia (Rizos and Spyropoulos, 2004). Some patients have sensorineural hearing loss or otitis media (Kantaputra et al., 2002; Salamone and Myer, 2004). Popliteal pterygium syndrome (PPS) (OMIM no.119500) exhibits a similar phenotype to VWS, but may present with a mixture of oral adhesions, eyelid adhesions (ankyloblepharon), pterygia, webbing of the

lower limbs, bands of mucous membrane between the jaws, syndactyly, and genital anomalies as well (Froster-Iskenius, 1990; Stottmann et al., 2010). It was reported that a common haplotype associated with *IRF6* contains a mutation attributable to approximately 12% of common forms of cleft lip and palate (Zucchero et al., 2004).

Cleft lip and/or palate is one of the most common birth defects which is caused by multiple genetic and environmental factors (Murray, 2002). Patients with cleft lip and/or palate require surgical, nutritional, medical and dental treatment and impose a substantial economic and psychological burden (Strauss, 1999). The average worldwide incidence of cleft lip and/or palate is 1 in 700 births and this frequency varies among different racial populations and different economic status (Vanderas, 1987), 1 in 500 in Asians and Amerindians and 1 in 2500 in Caucasians and Africans. Clefts are most often divided into cleft lip with or without cleft palate (CL/P) and those that involve the palate only (CPO), as the mechanism of CL/P involves the primary (hard) palate but CPO affects only the secondary (soft) palate (Fraser, 1955). Studies of cleft cases suggest that about 70% of cases of CL/P and 50% of CPO are nonsyndromic as affected individuals have no other physical or developmental anomalies (Jones, 1988). The syndromic cases, who have significant physical or developmental defects, can be subdivided into chromosomal syndromes, Mendelian disorders (Online Mendelian Inheritance in Man, 2002), teratogen-induced and uncategorized syndromes (Murray, 2002). Non-syndromic oral clefting is a complex trait caused by multiple factors including environmental triggers like teratogens (e.g., smoking, pharmaceuticals and pesticides) (Little et al., 2004), infection, nutrients (e.g., vitamins or trace elements) and cholesterol metabolism. Besides, several genes have been found to be involved in the palate formation. Point mutations of *Msx1* and

Tgfb3 have been identified in cases of cleft lip and/or palate (Murray, 2002). Other genes (*P63*, *PVRL1*, *TGFA*, *TBX22* and *SATB2*) that play a role in human palate development were also reported (FitzPatrick et al., 2003) .

1.2.4 IRF6 in mouse development

In *Irf6*-null mice, embryos lack external ears and have snouts and jaws that are shorter and more rounded than their wild-type littermates (Ingraham et al., 2006; Richardson et al., 2006). This phenotype is consistent with the observation that *Irf6* is expressed at key stages of facial development, and especially high levels are present in the ectoderm covering the facial processes immediately prior to and during palatal fusion to form the lip and primary palate (Knight et al., 2006) .

Aside from the craniofacial defects, *Irf6*- null mice exhibit taut, shiny skin and an epidermis that is thicker than in wild-type mice. The skin also lacks the normal wrinkled appearance (Ingraham et al., 2006; Richardson et al., 2006). Cell proliferation and apoptosis experiments suggest that the suprabasal keratinocytes of *Irf6*- null mice fail to stop proliferating and fail to terminally differentiate (Ingraham et al., 2006). The severe defects in the *Irf6*-null mouse embryos emphasize the important role of IRF6 in mouse craniofacial development and keratinocyte differentiation.

1.3 Other functions and regulation of IRF6

1.3.1 IRF6 functions as a transcriptional factor

Even though all IRF proteins possess a highly conserved N-terminal DNA binding domain (DBD) and recognize the ISRE (Taniguchi et al., 2001), different members may act as transcriptional activator or repressor. IRF1, IRF3 and IRF9 usually act as transcriptional activators, whereas IRF8 acts as a repressor. IRF6 was reported to function as a transcriptional activator as it activated the expression of ISRE-containing promoter reporter constructs in transfected cells (Fleming et al., 2009; Savitsky et al., 2010). IRF6 itself has an identical binding site. Full length IRF6 failed to bind the known consensus sites in the electrophoretic mobility shift assays, but the IRF6-DBD showed specific, high affinity binding to the consensus sequence of AACCGAAAC^C/_T *in vitro* (Little et al., 2009). Furthermore, ChIP-seq of keratinocytes under differentiating conditions show the consensus binding site of full length IRF6 is more likely to be NAC^C/_TGAAACN (Botti et al., 2011). IRF6 knock-down in primary human keratinocytes cause down regulation of 269 genes. Gene ontology analysis shows that these down-regulated genes are significantly related to cell adhesion, cell motion, cell morphogenesis, regulation of cell death, and stem cell development (Botti et al., 2011) .

1.3.2 IRF6 and cell proliferation and differentiation

The cell cycle is an intricate, temporally organized system that allows for the tightly regulated process of cell division. This progress involves the precise control of many cell

cycle regulators, which express in different stages of the cell cycle and consists of checkpoints (Bailey et al., 2008). The entry or exit of the cell cycle plays an important role in regulation of cell proliferation and differentiation.

The re-induction of IRF6 in breast cancer cells induces cell cycle arrest, which suggests that IRF6 may act as a mediator of cellular proliferation and differentiation in mammary epithelial cells (Bailey et al., 2008). Recent findings also suggest IRF6 is involved in cell proliferation, as down-regulation of IRF6 can promote invasive behavior of squamous cell carcinoma (SCC) cells (Botti et al., 2011). Besides, several genes related to cell proliferation (NGF, VEGFC et al.) are directly regulated by IRF6 (Botti et al., 2011). These findings imply that IRF6 can play an important role in the regulation of cell proliferation.

Complete knockout of *Irf6* in the mouse results in severe skin abnormalities (Ingraham et al., 2006). Cell proliferation and cell death analysis of the skin showed over-proliferation in the spinous layer, and failure of termination of cell differentiation, contributing to the abnormal skin (Ingraham et al., 2006). This finding suggests that IRF6 is necessary for regulating proliferation and terminal differentiation of keratinocytes. An *in vitro* study of *Irf6*^{-/-} keratinocyte figures out that the absence of *Irf6* causes a defect of differentiation, whereas over expression of *Irf6* can't promote differentiation, indicating it is necessary but not sufficient to promote keratinocyte differentiation (Biggs et al., 2012). Recently, IRF6 is also reported to function as a primary downstream target of Notch in keratinocyte, and contribute to the regulation of differentiation and repression of tumor (Restivo et al., 2011).

1.3.3 Regulation of IRF6

Dysregulation of genes involved in cell proliferation are often related the carcinogenesis and IRF6 may show a similar link. The protein level of IRF6 is down-regulated in 71% of SCCs, and the amount of IRF6 is found to correlate with histological stage, the highest in well-differentiated tumors and the lowest in high-grade, poorly differentiated SCCs (Botti et al., 2011). The reduction of *IRF6* mRNA and protein is also observed in poorly aggressive human breast cancer cell lines (MCF-7, T47-D). In aggressive and metastatic breast cancer cell lines (MDA-MB-231 and HS578T), IRF6 is completely absent (Bailey et al., 2005). These findings suggest that IRF6 is strictly regulated in both RNA and protein level.

Methylation at CpG islands of tumor suppressor gene promoters is a common phenomenon in cancer cells. The presence of 5-methyl cytosine within the CpG island of SCCs has been confirmed, and inhibition of DNA methyl transferase activity can induce *IRF6* expression (Botti et al., 2011). These findings suggest that that repression of *IRF6* transcription in SCC may be caused by promoter methylation, and IRF6 may act as a tumor suppressor.

IRF6 protein level is regulated in a cell cycle-dependant pattern. Cell cycle arrest (stopping at G₀ phase) is associated with a significant increase in total amount of IRF6, and the non-phosphorylated IRF6 is the prominent isoform (Bailey et al., 2008). When cells enter the G1 phase, phosphorylated IRF6 begins to decrease, this decrease being mediated by ubiquitination and proteasome degradation (Bailey et al., 2008). These findings suggest that

IRF6 protein expression and phosphorylation are regulated by proteasome degradation in a cell cycle-dependant pattern.

IRF6 is also a direct target of p63. The p53-related transcriptional activator p63 plays a central role in maintaining cellular proliferation during development. As a result of the alternative usage of 2 promoters and of complex alternative splicing, the p63 gene encodes 6 isoforms (Moretti et al., 2010). Among these isoforms, Δ Np63 is the major isoform expressed in primary keratinocytes and the palatal epithelia (Thomason et al., 2010). During early differentiation, Δ Np63 promotes transcription of *IRF6*, and the IRF6 protein in turn promotes Δ Np63 degradation (Moretti et al., 2010). This feedback regulation may play an important role in controlling the proliferation and differentiation of keratinocytes.

1.4 Microarray

Microarray is a hybridization of a nucleic acid sample to large amount of oligonucleotide probes which are printed to a solid platform to determine gene sequence or to detect gene expression or for gene mapping (<http://www.ncbi.nlm.nih.gov/genome/probe/doc/TechMicroarray.shtml>). In a typical microarray to detect the expression level of different samples, the RNA samples of interested will be reverse-transcript into cDNA, followed by labeling with dyes (Cyanine3, Cyanine 5). After the hybridization to the chip printed with probes, those DNA with specific binding to the probes will be attached to the chip, whereas the others will be washed out. The signal of each probe will be scanned and further analysis. With the huge amount of information get from microarray, the process of understanding the functions of

genes or proteins is greatly accelerated. With the wide application of microarray, many useful tools and software, like Tools BioconductorGene Map Annotator and Pathway Profiler (GenMAPP), Spotfire DecisionSite for Functional Genomics, Genespring, are designed and facilitate the usage of microarray data (Hoheisel, 2006).

1.5 Objectives of the project

The objectives of this project were:

1. To identify differentially expressed genes in *irf6* knockdown morphants;
2. To validate putative downstream target genes of Irf6;
3. To perform preliminary knockdown analysis of differentially expressed genes.

The knowledge gathered from this study will provide novel insights into how Irf6 functions in vertebrate early embryogenesis.

Chapter II: Materials and Methods

2.1: Ethics statement/ fish strain

Singapore wild-type and AB strain (Eugene, Oregon) zebrafish were maintained in a life support system at 28 °C. Embryos were staged according to standard criteria as described (Kimmel et al., 1995). All animal work was performed and approved by the NUS Institutional Animal Care and Use Committee (IACUC).

2.2: Morpholino injection

Gene knock down analysis was carried out by Morpholino injection (explained in the introduction part) to study the functions of target genes. Morpholinos were purchased from Gene Tools LLC (Philomath, OR). They were injected into the embryos at the one- to four-cell stage at a concentrations of 1.0 mM in 1X Danieau's buffer (58 mM NaCl; 0.7 mM KCl; 0.4 mM MgSO₄; 0.6 mM Ca(NO₃)₂ and 5.0 mM Hepes, pH 7.6). Approximately 2 nl of morpholino was injected into each embryo by using a FemtoJet[®] Microinjector (Eppendorf) under a dissection microscope (MZ FL III, Leica). The morpholino was designed to block the *irf6* pre-mRNA splicing: E3I3, 5'-ctg tgt gtg tgt tac CAG GGT TGC T-3' (exon sequence capitalized). A generic morpholino oligo was used as the morpholino toxicity control: STD, 5'-CCT CTT ACC TCA GTT ACA ATT TAT A-3'. Two morpholinos to knockdown the *cyr61* and *mapkapk3* genes were: *cyr61* MO, 5'- GCC TGG ACA GCC ACG AGA CAT CTC T-3' and *mapkapk3* MO, 5'-TCT GAG ACT TTC CAT TCT GGA GCA T-3'.

2.3: Total RNA extraction from fish embryos

To search for genes regulated by Irf6, we conducted a whole transcriptome microarray analysis of zebrafish embryos subjected to dominant-negative Irf6 perturbation. The total RNA was used as the biological sample to perform the microarray analysis. Total RNA of fifteen to twenty zebrafish embryos were collected and homogenized in 0.5ml TRIZOL[®] RNA isolation reagent (Invitrogen, catalog no.15596-026) using a plastic pestle. The samples were then incubated for five minutes at room temperature for complete dissociation of the nucleoprotein complex. 0.1 ml of chloroform (EMD Chemicals Inc, CX1055) was added and shaken vigorously for 15 seconds and then incubated at room temperature for two to three minutes. The sample was then centrifuged at 16,000 g for 15 minutes at 4°C. The aqueous phase was transferred to a new 2 ml microfuge tube and 0.25 ml of isopropyl alcohol was added to precipitate the RNA at room temperature for 10 minutes. After that, the sample was centrifuged at 16,000 g for 10 minutes at 4°C and the supernatant was discarded. The RNA pellet was washed in 0.5 ml of 70% ethanol and centrifuged at 8000 g for five minutes at 4°C. The supernatant was discarded and the air-dried RNA pellet was dissolved in 0.1% DEPC (Sigma, D5758) water. The RNA concentration was determined by using the Nanodrop Spectrophotometer (Thermo Scientific).

2.4: Microarray sample preparation and hybridization

Four biological replicates of the E3I3-MO injected and mock injected embryos were harvested at the 1k cell stage and 40% epiboly stage. The total RNA was extracted by using TRIZOL[®] RNA isolation reagent (Invitrogen, catalog no.15596-026) and quantified. The total RNA was subsequently sent for zebrafish gene expression microarray analysis (Agilent). Briefly, cDNA reversely-transcribed from the total RNA was used for the synthesis of Cyanine-3 labeled cRNA by using Agilent Low Input Quick Amp Labeling kit (Agilent). After purification, the labeled cRNA was used for the hybridization with the slides (Agilent SurePrint G3 (Zebrafish), one color, 8x60K format). The slides were then scanned and the raw data was extracted using Agilent Feature Extraction Software for further analysis.

2.5: Microarray analysis and statistics

The raw data extracted by the Agilent Feature Extraction Software was included in the final analysis to detect differentially expressed genes by using GeneSpring software (Agilent, USA). Briefly, the raw data were subjected to summarization, normalization and filtering. After that, the one-way ANOVA was subsequently used to detect the p-value for the respective gene expression fold changes. The criteria for a gene to be considered differentially expressed were set at $p \leq 0.05$ and a minimal fold change of two. Gene Ontology analysis was performed using the GO analysis function within GeneSpring (Agilent).

2.6: Semi-quantitative reverse-transcription PCR:

To validate the result of the microarray data, semi-quantitative reverse-transcription PCR was carried out. The first strand cDNA was generated using SuperScript™ II Reverse Transcriptase (Invitrogen, 18064-014). 0.5 µg of total RNA (100 ng/ µl), 1 µl Oligo-(dT) primer (500 µg/ml), 1 µl dNTP (10 mM each) and 13 µl nuclease free water were mixed together in a 200 µl PCR tube and incubated for five minutes at 65 °C. After the incubation, the mixture was quickly chilled on ice. 4 µl of 5 X First-Strand Buffer and 2 µl of 0.1 M DTT were added into the PCR tube and incubated at 42 °C for two minutes. Subsequently, 1 µl (200 units) of SuperScript™ II RT was added into the reaction followed by incubation at 42°C for 50 minutes. After the incubation, the whole reaction was stopped by heating at 70°C for 15 minutes. RNA was removed from the cDNA by adding 1 µl (2 units) of RNase H (Invitrogen, 18021-071) and incubated at 37°C for 20 minutes.

100 ng of the cDNA template was used for PCR amplification using Hotstart Taq Polymerase (Qiagen, 203203). The primers pairs: *cyr61* F/R 5'-AGT GAC CAA CAG TAA CGC TCA GTG C -3' / 5'-CCG GCT TAC GAG GTC TTG TTG TAC G -3' and *mapkapk3* F/R 5'-GAG GAG CCG TCG CAC CTG -3' / 5'-GCC ACT CGG ATC TTA TTC AC-3' were used for the amplification of *cyr61* and *mapkapk3* respectively. Another primer pair: *β-actin* F/R 5'- TGA CCC TGA AGT ACC CAA TTG AG -3' / 5'- GGC AAC ACG CAG CTC ATT G-3' was used to amplify the internal control *β-actin*.

The PCR cycling conditions were set as follows:

| | | | |
|----------------------|------|---------|-------------|
| Initial denaturation | 95°C | 15 mins | |
| Denaturation | 95°C | 30 sec | ← 35 cycles |
| Annealing | 60°C | 30 sec | |
| Extension | 72°C | 30 sec | |
| Final extension | 72°C | 10 mins | |

Amplified products were then analyzed by agarose gel electrophoresis.

2.7: pcDNA/His-Irf6-FL and pcDNA/His-Irf6-E3I3 plasmid construction

2.7.1 Amplification of full length and truncated Irf6

The pcDNA/His-Irf6-FL and pcDNA/His-Irf6-E3I3 plasmid were constructed to express His-tagged IRF6 full length and truncated proteins. The full length and truncated *irf6* were amplified from the first strand cDNA that was reversely-transcribed from the total RNA extracted from wild-type and E3I3 MO-injected embryos respectively. The primers used for full length *irf6* (around 1.5 kb) amplification were *irf6* F 5'-ATG TCG TCT CAT CCA CGG CG -3' and *irf6* FL R 5'-TTA CTG CGT GTG TGC AGG GCG G -3', whereas the primers for truncated *irf6* (426 bp) amplification were *irf6* F (mentioned above) and *irf6* E3I3 R 5'- TCA TGC CAT GTG ATG CAT AT-3'. For PCR reaction, 40.6 µl of nuclease-free water, 5 µl of 10 X reaction buffer, 0.4 µl of dNTPs (25 mM each), 1.25 µl of each primers (10 µM), 1 µl of *Pfu* DNA polymerase (2.5 U/ µl) (Stratagene, 600135) and 0.5 µl of DNA template (100 ng/ µl) were mixed together in a 200 µl PCR tube. The PCR condition was: 95 °C for 15 minutes, 35 cycles of 95 °C for 30 seconds, 60 °C for 30 seconds and 72 °C for two minutes, followed by a final extension of 72°C for 10 minutes.

Amplified products were analyzed by agarose gel electrophoresis and the target bands were purified using illustra GFX PCR DNA and Gel Band Purification kit (GE Healthcare, 28-9034-70).

2.7.2 Plasmid digestion

The vector pcDNATM 3.1/His A (Invitrogen, 350512) was digested with *Kpn* I (Fermentas, ER0521) at 37 °C for four hours. The digestion reaction mixture consisted of 2 µl of *Kpn* I enzyme (10 U/µl), 2 µl of 10X Buffer *Kpn* I, 1µl of plasmid (1µg/µl) and nuclease - free water.

2.7.3 Creating blunt end

The blunt ended pcDNATM 3.1/His A plasmid was created by treating the linearized plasmid with T4 polymerase (Fermentas, EP0061). The 20 µl reaction mixture consisted of 4 µl of 5X reaction buffer, 1 µg linearized plasmid, 0.2 µl of T4 DNA Polymerase (5U/µl), 2 µl dNTP (25 mM each) and nuclease - free water. The reaction was carried out at 11 °C for 20 minutes and was stopped by heating at 75 °C for 10 minutes.

2.7.4 Ligation

The blunt ended vector pcDNATM 3.1/His A was ligated with full length and truncated *irf6* fragment using T4 ligase (Fermentas, EL0014) respectively. The insert fragment was 5:1 molar ratio over vector in a 20 µl of reaction mixture. Ligation was performed at 4 °C overnight.

2.7.5 Transformation

Fifty microliters of Subcloning Efficiency™ DH5α™ Competent Cells (Invitrogen, 18265-017) were removed from –85°C freezer, and thawed on ice. 5 µl of the DNA ligation reaction was added directly to tube containing 50 µl competent cells. The mixture was incubated on ice for 30 minutes and then heat-shocked for 20 seconds at 42°C without shaking. After incubation on ice for two minutes, 0.95 ml of room temperature S.O.C. medium (Invitrogen, 15544-034) was added, and the tube was incubated one hour at 37°C in a shaker at 225 rpm. Thereafter, 100 µl of the reaction was spread on LB agar plates containing 100 µg/ml ampicillin. The plate was incubated overnight at 37°C (16 hours) and the colonies were picked randomly. Colony PCR was carried out to check the insert. The constructed plasmids containing full length *irf6* and truncated *irf6* sequence were recorded as pcDNA/His-Irf6-FL and pcDNA/His-Irf6-E3I3 respectively.

2.8: *In vitro* protein expression

2.8.1: Generation of DNA templates for full length and truncated Irf6 protein expression

As TNT® SP6 High-Yield Wheat Germ Protein Expression System (Promega, L3261) was used as the *in vitro* protein expression system, the SP6 promoter is necessary for the protein expression. Thus, a SP6 promoter was added to the full length and truncated His-tag Irf6 DNA sequence. The DNA template for the expression of His-tagged full length Irf6 protein was amplified from pcDNA/His-Irf6-FL plasmid using SP6 plus primer: 5'- GCG

AAA TTA TAT TTA GGT GAC ACT ATA GAA CAG ACC ACC ATG GGG GGT TCT CAT CAT-3' and *irf6* FL R primer: 5'- TTA CTG CGT GTG CAG GGC GG-3'. The DNA template for His-tagged truncated protein expression was amplified from pcDNA/His-Irf6-E3I3 plasmid using SP6 plus primer and *irf6* E3I3 R primer 5'-TCA TGC CA CAT GTG ATG CAT AT-3'. The PCR condition was: 95 °C for 15 minutes, 35 cycles of 95 °C for 30 seconds, 60 °C for 30 seconds and 72 °C for two minutes, followed by a final extension of 72°C for 10 minutes. Amplified products were analyzed by agarose gel electrophoresis and the target bands were gel purified (GE Healthcare).

2.8.2: Protein expression using TNT wheat germ expression system

TNT[®] SP6 High-Yield Wheat Germ Protein Expression System (Promega, L3261) was used to express the recombinant His-tagged Irf6 full length protein and His-tagged E3I3 truncated protein. Thirty microliters of wheat germ mixture was removed from -80°C and thawed on ice, and 1mg purified DNA was added into the mixture and incubated at 25°C for two hours to express the target protein. A reaction without any DNA template was carried out in parallel as a negative control. The results of translation were checked by SDS-PAGE.

2.9: Protein purification

His-tagged protein purification was carried out by using Dynabeads[®] His-Tag Isolation & Pulldown system (Invitrogen, 10103D). 50 µl (2 mg) well-mixed Dynabeads were transferred to a microcentrifuge tube and place on a magnet for two minutes, then the

supernatant was discarded. The protein lysate generated from TNT[®] SP6 High-Yield Wheat Germ Protein Expression System was prepared with 700 µl of 1X Binding buffer / Wash Buffer (50 mM Sodium- Phosphate, 300 mM NaCl, 0.01% Tween-20 pH 8.0) and incubated with Dynabeads for 10 minutes at room temperature with rotation. After the incubation, the Dynabeads were washed 4 times with 300 µl 1X Binding/Wash Buffer by placing the tube on a magnet for two minutes, and the supernatant was discarded. 50 µl of His-Elution Buffer (300 mM Imidazole, 50 mM Sodium-phosphate, 300 mM NaCl and 0.01% Tween-20; pH 8.0) was added to the Dynabeads and incubated on a roller for 5 minutes at room temperature to elude his-tagged protein.

2.10: Western blotting

15% SDS-PAGE gels were used in this study. The gels were prepared as follows:

Table 2: SDS-PAGE gel recipe

| | Component | Volume |
|--------------------------|--------------------------|---------------|
| 15% Resolving gel | ddH ₂ O | 1.8 ml |
| | 30% Acrylamide | 4 ml |
| | 1.5M Tris pH8.8 | 2 ml |
| | 10% SDS | 80 ul |
| | 10% Ammonium persulphate | 80 ul |
| | TEMED | 8 ul |
| Total Volume | | 8 ml |
| 6% Stacking gel | ddH ₂ O | 2.6 ml |
| | 30% Acrylamide | 1 ml |
| | 0.5M Tris pH6.8 | 1.25 ml |
| | 10% SDS | 50 ul |
| | 10% Ammonium persulphate | 50 ul |
| | TEMED | 5 ul |
| Total Volume | | 5 ml |

Protein samples were loaded and run on a MiniProtean II system (Biorad) at 80V until the sample passed the stacking gel, followed by 120V for two hours. After an electrotransfer for one hour at 100V, the PVDF membranes were blocked overnight in blocking buffer at 4 °C (5% skim milk, 10 mM Phosphate buffer, 137 mM NaCl, 2.7 mM KCl, 0.1% Tween-20, pH7.4). After blocking, the membranes were placed in primary antibody diluted in 10 ml blocking buffer for two hours at room temperature. The membranes were then washed for 3 times with PBST buffer (10mM Phosphate buffer, 137mM NaCl, 2.7mM KCl, 0.1% Tween-20, pH7.4), 10 minutes each. After washing, the membranes were incubated with diluted secondary antibody for one hour at room temperature. The blots were developed with substrate for one minute and exposed with CL-XPosure (TM) Film (Pierce).

Table 3: Antibodies used in western blotting

| | Dilution | Host | Company |
|---------------------------|-----------------|-------------|-----------------------------------|
| Primary Antibody | | | |
| Anti-His tag antibody | 1:2000 | Mouse | Invitrogen, 372900 |
| Anti-Irf6 antibody | 1:2000 | Rabbit | Abcam, ab58915 |
| Secondary antibody | | | |
| Goat anti-mouse antibody | 1:40,000 | Goat | Santa Cruz Biotechnology,sc-2031 |
| Goat anti-rabbit antibody | 1:40,000 | Goat | Santa Cruz Biotechnology, sc-2030 |

2.11: Electrophoretic mobility shift assay (EMSA)

The electrophoretic mobility shift assay (EMSA) has been used extensively for studying DNA-protein interactions (Hellman and Fried, 2007). The DNA-protein complexes migrate slower than non-bound DNA in a native polyacrylamide or agarose gel, resulting in a “shift” in migration of the labeled DNA band. Double stranded oligonucleotide probes containing the IRF6 binding site: 5'-TTC CAA ATG GAC CGA AAC ATA TAA ATT TTG-3' for *mapkapk3* and 5'-GCG ATG ACG CTA ACC GAA ACT TGC TAG ATG-3' for *cyr61* were labeled with biotin using Biotin 3'-DNA Labeling Kit (Pierce, 89818). The labeling reaction was carried out by mixing 10 µl of 5X terminal deoxynucleotidyl transferase (TdT) reaction buffer, 5 µl unlabeled oligo (1 µM), 5 µl biotin-11-UTP (5 µM), 5 µl TdT (2U/µl) and 25 µl nuclease free water and incubated at 37°C for one hour. Thereafter, 2.5 µl of 0.2

M EDTA was added to stop the reaction. 50 μ l chloroform:isoamyl alcohol (24:1) was added to each reaction to extract the TdT. The mixture was vortex briefly and centrifuge for two minutes at 16,000 g to separate the phases. The top (aqueous) phase was saved. The probes were annealed by mixing equal amounts of labeled complementary oligos, denatured at 95°C for one minute, and then slowly cooled (1°C/6 minutes), and incubated at the melting temperature 50°C for two hours.

The binding reaction was tested using LightShift Chemiluminescent EMSA Kit (Pierce, 20148). The biotin labeled probe was incubated at room temperature for 30 min with Irf6 full length or truncated protein in the presence of binding reaction mixture (1X binding buffer; 2.5% glycerol; 5 mM MgCl₂; 50 ng/ul bovine serum albumin and 50 ng/ul poly(dI:dI)-poly(dI:dC)]. The DNA-protein complexes were resolved on a 6% non-denaturing polyacrylamide gel in 0.5X TBE (45 mM Tris-HCl, 45 mM Boric Acid, 1 mM EDTA, pH 8.3) for 1.5 h at 120 V. Proteins and bound probes were transferred to a positively charged nylon membrane (Pierce, 0077016) in 0.5X TBE at 380 mA for 30 minutes. The transferred membrane was UV cross-linked at 120mJ/cm² for one minute (Stratagene). The biotin-labeled DNA was detected with a Chemiluminescence Nucleic Acid Detection Module (Pierce, 0089880), the film was developed and exposed to with CL-XPosure (TM) Film (Pierce, 0034090) after which.

Chapter III: Results

3.1: Genome-wide gene profiling microarray analysis of the E3I3 injected embryos

Perturbation of *Irf6* either by injection of *Irf6* mRNA encoding only its DNA-binding domain (Sabel et al., 2009) or a splice-modifying Morpholino-E3I3 (our unpublished data) leads to distinct gastrulation defects and subsequent rupture of the injected zebrafish embryos at the animal pole, strongly suggesting that *Irf6* is critical for early development. Since *Irf6* is a transcriptional factor, identifying genes regulated by *Irf6* during early development can aid in understanding the role of *Irf6* in early development and gastrulation in particular. Thus, transcriptome profiling was performed to identify genes differentially expressed after injection of E3I3 morpholino into zebrafish embryos using the Agilent zebrafish gene expression microarray system.

At 40% epiboly (~5 hpf), four biological replicates (40% epiboly set one) of both the mock MO-injected and the E3I3 MO-injected embryos were harvested and a genome-wide microarray analysis was performed to identify differentially expressed genes. Another three biological replicates (40% epiboly set two) were analyzed separately to further confirm the expression array result. To identify genes that may be expressed at earlier stage, embryos were also harvested at the 1k cell stage. Gene expression profiles were analyzed using GeneSpring software (Agilent). After normalization and appropriate filtering, only those genes with significant changes of more than two fold ($p \leq 0.05$) in all replicates were classified as differentially expressed.

Datasets of differentially regulated genes were generated for the 1k cell stage (222 genes, 62 up-regulated and 160 down-regulated), 40% epiboly set one (577 genes, 251 up-regulated and 326 down-regulated) and 40% epiboly set two (552 genes, 238 up-regulated and 314 down-regulated). Among the hundreds of differentially regulated genes identified at 40% epiboly, 172 genes (125 down-regulated and 47 up-regulated) (Table 4) were consistently detected in both 40% epiboly datasets. Of note, 49 of the differentially expressed genes at the 40% epiboly stage (two up-regulated and 47 down-regulated) were also differentially expressed at 1k cell stage, as opposed to genes that were differentially expressed only at the 1k cell stage or the 40% epiboly stage.

Table 4: Microarray Gene Expression Analysis: E.313 MO-injected Embryos vs mock-MO injected Embryos

| Probe ID | NCBI Accession | Gene symbol | Gene name | 40%epiboly set 1 | | 40%epiboly set 2 | | 1k cell stage | |
|---|----------------|------------------|--|------------------|----------|------------------|----------|---------------|----------|
| | | | | fold change | P-value | fold change | P-value | fold change | P-value |
| Down-regulation in E.313 MO-injected embryos | | | | | | | | | |
| A_15_P163446 | NM_001080987 | cyf61 | cysteine-rich, angiogenic inducer, 61 | 123.1004 | 0.014054 | 15.0038 | 0.027466 | | |
| A_15_P161666 | NM_001080079 | mupkpk3 | mitogen-activated protein kinase-activated protein kinase 3 | 109.3555 | 0.014807 | 23.0102 | 0.017192 | | |
| A_15_P772116 | 793282//793282 | sidkeyp-178e17.1 | XM_001332970 | 52.1676 | 0.013111 | 12.9851 | 0.008160 | | |
| A_15_P103414 | BC078312 | | | 43.9603 | 0.039423 | 7.1892 | 0.027136 | | |
| A_15_P470700 | NM_131320 | lcp1 | lymphocyte cytosolic plastin 1 | 38.8929 | 0.046696 | 12.9826 | 0.003741 | | |
| A_15_P629031 | NM_001006094 | rab34 | RAB34, member RAS oncogene family | 28.6312 | 0.019530 | 7.5190 | 0.001060 | | |
| A_15_P108869 | NM_001006094 | rab34 | RAB34, member RAS oncogene family | 14.3480 | 0.031133 | 3.9206 | 0.019043 | | |
| A_15_P244106 | NM_001115059 | tgfb1b | transforming growth factor, beta receptor 1 b | 25.6549 | 0.013950 | 3.6605 | 0.002006 | | |
| A_15_P748421 | NM_001045401 | zgc:153411 | zgc:153411 | 23.0679 | 0.042884 | 8.5262 | 0.010043 | | |
| A_15_P119965 | NM_200591 | ghitm | growth hormone inducible transmembrane protein | 15.6594 | 0.019192 | 9.7635 | 0.001184 | 7.9928 | 0.006403 |
| A_15_P668011 | BC066450 | ghitm | growth hormone inducible transmembrane protein | 14.1390 | 0.009756 | 8.1682 | 0.000875 | 6.6663 | 0.002098 |
| A_15_P407150 | GT597797 | wurfb98b04 | wurfb98b04 | 14.6766 | 0.049105 | 2.5845 | 0.045889 | | |
| A_15_P177851 | BC125906 | hlcs | holocarboxylase synthetase (biotin-(propionyl)-Coenzyme A-carboxylase (ATP-hydrolysing)) | 14.3214 | 0.001311 | 7.1820 | 0.000016 | 8.7875 | 0.000003 |
| A_15_P716066 | BC154426 | hlcs | holocarboxylase synthetase (biotin-(propionyl)-Coenzyme A-carboxylase (ATP-hydrolysing)) | 3.0002 | 0.017839 | 3.2669 | 0.003188 | 3.5579 | 0.000006 |
| A_15_P194511 | NM_201289 | pkfr | pyruvate kinase, liver and RBC | 14.2975 | 0.017699 | 2.9484 | 0.028749 | | |
| A_15_P444325 | XM_002662593 | LOC100333117 | ubiquitin carboxyl-terminal hydrolase CYLD-like | 12.5674 | 0.001311 | 7.7852 | 0.000361 | 10.6936 | 0.000160 |
| A_15_P151201 | NM_205664 | zgc:77398 | zgc:77398 | 12.4060 | 0.002801 | 7.1271 | 0.000126 | 4.5389 | 0.001679 |
| A_15_P720926 | NM_205664 | zgc:77398 | zgc:77398 | 4.9967 | 0.033095 | 4.7654 | 0.000001 | 2.9498 | 0.002999 |
| A_15_P208476 | NM_001076772 | kfl13 | Kruppel-like factor 13 | 12.3711 | 0.009667 | 9.6416 | 0.003994 | | |
| A_15_P177836 | BC125850 | | | 12.3563 | 0.027901 | 3.8764 | 0.045075 | | |
| A_15_P398240 | CF594940 | sidkeyp-73d8.9 | sidkeyp-73d8.9 | 12.1880 | 0.017839 | 7.0973 | 0.002103 | | |
| A_15_P625701 | NM_001130572 | sidkeyp-73d8.9 | sidkeyp-73d8.9 | 10.5408 | 0.013111 | 7.1155 | 0.006261 | | |
| A_15_P763196 | EV756349 | | | 11.8379 | 0.004559 | 8.4780 | 0.003838 | 6.7452 | 0.000038 |
| A_15_P473790 | EV756349 | | | 10.5011 | 0.003278 | 8.8804 | 0.001027 | 10.1219 | 0.000265 |
| A_15_P115932 | NM_207085 | zgc:77287 | zgc:77287 | 11.7713 | 0.013910 | 21.9261 | 0.000080 | 10.4987 | 0.000002 |
| A_15_P631921 | NM_001002213 | tadn1 | TatD DNase domain containing 1 | 10.6814 | 0.012940 | 9.3501 | 0.002626 | 7.6132 | 0.000029 |
| A_15_P116245 | NM_001165918 | tbeta-a | beta thymosin-like protein 2 | 10.4980 | 0.032246 | 2.4784 | 0.039308 | | |
| A_15_P646396 | NM_001136333 | znf1 | zinc finger-like gene 1 | 10.3810 | 0.025711 | 2.7802 | 0.037537 | 2.8303 | 0.004678 |
| A_15_P705671 | NM_001136333 | znf1 | zinc finger-like gene 1 | 8.8966 | 0.014111 | 2.7670 | 0.004206 | 3.1170 | 0.000330 |
| A_15_P698071 | NM_001136333 | znf1 | zinc finger-like gene 1 | 6.6253 | 0.028752 | 2.6659 | 0.003530 | 2.6181 | 0.000696 |
| A_15_P718893 | NM_001136333 | znf1 | zinc finger-like gene 1 | 5.6366 | 0.025711 | 2.1153 | 0.017462 | 2.6863 | 0.000218 |
| A_15_P653021 | NM_001171032 | znf1 | zinc finger-like gene 1 | 5.4619 | 0.019192 | 2.1236 | 0.016445 | 2.3303 | 0.000828 |
| A_15_P625051 | NM_194389 | znf1 | zinc finger-like gene 1 | 4.9467 | 0.031647 | 2.1460 | 0.009204 | 2.0764 | 0.000371 |
| A_15_P671356 | NM_001076764 | zgc:153384 | zgc:153384 | 9.8298 | 0.011802 | 7.4895 | 0.001588 | 10.7554 | 0.000142 |
| A_15_P620581 | NM_001076764 | zgc:153384 | zgc:153384 | 5.0375 | 0.042884 | 3.0437 | 0.000508 | 6.7267 | 0.000055 |
| A_15_P334799 | NM_214693 | lgals3l | lectin, galactoside-binding, soluble, 3 (galectin 3)-like | 9.1440 | 0.031268 | 2.6220 | 0.018158 | | |
| A_15_P141796 | NM_001080042 | zgc:158346 | zgc:158346 | 9.0587 | 0.045299 | 6.8333 | 0.015635 | | |
| A_15_P175236 | NM_212593 | kiaa0907 | KIAA0907 protein | 8.9072 | 0.012940 | 6.8988 | 0.000302 | 7.9817 | 0.000019 |
| A_15_P141876 | NM_001017583 | zgc:110726 | zgc:110726 | 8.6952 | 0.016673 | 3.9397 | 0.000105 | 5.0382 | 0.000122 |
| A_15_P763091 | NM_001017583 | zgc:110726 | zgc:110726 | 6.1215 | 0.031832 | 3.2785 | 0.000254 | 4.4582 | 0.000068 |
| A_15_P613857 | NM_200367 | zgc:64031 | zgc:64031 | 8.5983 | 0.040615 | 2.9110 | 0.009103 | 3.7065 | 0.000170 |
| A_15_P113362 | NM_200367 | zgc:64031 | zgc:64031 | 6.4956 | 0.042884 | 4.6528 | 0.000770 | | |

| Probe ID | NCBI Accession | Gene symbol | Gene name | 40% epiboly set 1 | | 40% epiboly set 2 | | 1k cell stage | |
|---------------|----------------|-------------------|---|-------------------|----------|-------------------|----------|---------------|----------|
| | | | | fold change | P-value | fold change | P-value | fold change | P-value |
| A_15_P1757180 | NM_194390 | zmf2a | zinc finger-like gene 2a | 8.5963 | 0.018770 | 2.9864 | 0.022246 | 2.7034 | 0.015033 |
| A_15_P109056 | NM_194390 | zmf2a | zinc finger-like gene 2a | 6.6999 | 0.023198 | 2.7353 | 0.021054 | 2.0275 | 0.025479 |
| A_15_P764461 | | | | 8.5189 | 0.022972 | 2.9963 | 0.023516 | 2.1107 | 0.006195 |
| A_15_P632516 | NM_001130697 | zgc:195154 | zgc:195154 | 8.4440 | 0.043268 | 4.0610 | 0.008965 | 2.2938 | 0.045329 |
| A_15_P177486 | BC122238 | zgc:195154 | zgc:195154 | 8.0716 | 0.042891 | 4.2646 | 0.017100 | | |
| A_15_P696106 | NM_001130697 | zgc:195154 | zgc:195154 | 7.8883 | 0.045997 | 4.3800 | 0.014341 | | |
| A_15_P212551 | NM_001130697 | zgc:195154 | zgc:195154 | 7.7275 | 0.042690 | 4.3475 | 0.016449 | | |
| A_15_P174386 | NM_001098767 | pts | 6-pyruvoyltetrahydropterin synthase | 8.4268 | 0.017839 | 3.1981 | 0.001950 | 5.9774 | 0.000060 |
| A_15_P638926 | NM_001098767 | pts | 6-pyruvoyltetrahydropterin synthase | 2.1029 | 0.046787 | 3.1919 | 0.008144 | | |
| A_15_P509842 | NM_001126437 | syn1 | synapsin I | 8.1798 | 0.017883 | 3.0871 | 0.011344 | | |
| A_15_P169626 | NM_200156 | nup155 | nucleoporin 155 | 8.1690 | 0.015524 | 4.4981 | 0.000558 | 6.4075 | 0.000001 |
| A_15_P184531 | XM_686777 | si:ch211-245h14.1 | si:ch211-245h14.1 | 7.9474 | 0.017839 | 7.7343 | 0.000253 | | |
| A_15_P758961 | XM_001333472 | | | 7.5689 | 0.027729 | 2.3694 | 0.041671 | | |
| A_15_P116411 | NM_197942 | rass7b | Ras association (RalGDS/AF-6) domain family (N-terminal) member 7b | 7.5619 | 0.012940 | 2.6676 | 0.030830 | 4.9414 | 0.017132 |
| A_15_P112048 | NM_197942 | rass7b | Ras association (RalGDS/AF-6) domain family (N-terminal) member 7b | 7.2729 | 0.012886 | 2.7119 | 0.013887 | 5.6594 | 0.022454 |
| A_15_P321531 | XM_001336538 | | | 7.3033 | 0.044900 | 2.4866 | 0.002406 | | |
| A_15_P744431 | XM_001336538 | | | 7.0714 | 0.033095 | 2.6450 | 0.024942 | | |
| A_15_P164101 | NM_001089472 | ulk3 | unc-51-like kinase 3 (C. elegans) | 7.2625 | 0.002801 | 5.2318 | 0.000305 | 5.7540 | 0.000002 |
| A_15_P150336 | NM_001003867 | mybl2 | myeloblastosis oncogene-like 2 | 7.0810 | 0.031647 | 13.3836 | 0.002330 | | |
| A_15_P560967 | EH445542 | LOC793702 | hCG2038584-like | 6.9720 | 0.036766 | 2.8235 | 0.024464 | | |
| A_15_P246116 | XM_002666572 | LOC100333131 | hypothetical protein LOC100333131 | 6.5270 | 0.028717 | 3.9660 | 0.017355 | 4.7007 | 0.000335 |
| A_15_P112839 | NM_205709 | bub3 | BUB3 budding uninhibited by benzimidazoles 3 homolog (yeast) | 6.4015 | 0.016673 | 3.5734 | 0.028511 | | |
| A_15_P674571 | NM_205709 | bub3 | BUB3 budding uninhibited by benzimidazoles 3 homolog (yeast) | 4.6743 | 0.012940 | 3.4277 | 0.001800 | | |
| A_15_P113516 | NM_001161488 | pcp41l | Purkinje cell protein 4 like 1 | 6.3353 | 0.032228 | 2.7434 | 0.048213 | | |
| A_15_P628473 | NM_001145701 | LOC792757 | similar to zinc finger-like gene 1 | 6.1376 | 0.010082 | 2.4297 | 0.012731 | 2.3790 | 0.000254 |
| A_15_P666461 | NM_213522 | zgc:55347 | zgc:55347 | 5.9804 | 0.009667 | 5.1860 | 0.000189 | 3.3453 | 0.001594 |
| A_15_P663496 | NM_213522 | zgc:55347 | zgc:55347 | 5.6901 | 0.009667 | 5.0803 | 0.000240 | 3.0659 | 0.000816 |
| A_15_P624926 | NM_001020651 | zgc:112102 | zgc:112102 | 5.9285 | 0.030955 | 5.5531 | 0.001348 | 2.5046 | 0.017785 |
| A_15_P711561 | NM_001020651 | zgc:112102 | zgc:112102 | 5.1395 | 0.028973 | 4.7065 | 0.000278 | 2.1836 | 0.005588 |
| A_15_P118147 | NM_205659 | crkr1 | connector enhancer of kinase suppressor of Ras 1 | 5.9123 | 0.019485 | 2.6998 | 0.009063 | | |
| A_15_P621551 | NM_001002381 | zgc:92489 | zgc:92489 | 5.5272 | 0.027895 | 2.4150 | 0.001102 | | |
| A_15_P143981 | NM_200472 | pygmb | phosphorylase, glycogen (muscle) b | 5.5268 | 0.024369 | 2.3253 | 0.028842 | 2.8243 | 0.004768 |
| A_15_P263721 | EV759803 | | | 5.5261 | 0.046679 | 2.3395 | 0.024765 | | |
| A_15_P113225 | NM_213206 | rad51 | RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae) | 5.5231 | 0.034087 | 2.6241 | 0.026109 | | |
| A_15_P105981 | NM_200548 | pro1 | partner of NOB1 homolog (S. cerevisiae) | 5.4109 | 0.009667 | 2.5286 | 0.000755 | 3.2099 | 0.000344 |
| A_15_P721596 | NM_200548 | pro1 | partner of NOB1 homolog (S. cerevisiae) | 3.6226 | 0.009667 | 2.1807 | 0.000300 | 2.4289 | 0.000550 |
| A_15_P212006 | NM_200538 | slc2a12 | solute carrier family 2 (facilitated glucose transporter), member 12 | 5.4058 | 0.040678 | 4.1676 | 0.023598 | | |
| A_15_P106200 | NM_001001838 | top1mt | mitochondrial topoisomerase I | 5.3979 | 0.035627 | 4.1606 | 0.020788 | 5.3036 | 0.000001 |
| A_15_P108018 | AW941357 | | | 5.1797 | 0.046801 | 2.3120 | 0.026623 | | |
| A_15_P143446 | NM_001082926 | si:ch211-107n13.1 | si:ch211-107n13.1 | 5.1728 | 0.012886 | 5.9081 | 0.006269 | | |
| A_15_P705601 | | | | 5.1713 | 0.046418 | 2.5782 | 0.010313 | 2.3406 | 0.000712 |
| A_15_P118538 | NM_001025516 | zgc:114097 | zgc:114097 | 5.0765 | 0.002801 | 3.9090 | 0.000022 | 3.2179 | 0.000029 |
| A_15_P667166 | BC027700 | LOC407704 | hypothetical protein LOC407704 | 4.9070 | 0.045492 | 3.1285 | 0.000203 | 3.3402 | 0.000756 |
| A_15_P609427 | BC068425 | LOC407704 | hypothetical protein LOC407704 | 3.9943 | 0.044013 | 2.3347 | 0.000086 | | |
| A_15_P106825 | BC093424 | si:dkfy-273g18.1 | si:dkfy-273g18.1 | 4.8389 | 0.039423 | 2.3719 | 0.044745 | | |
| A_15_P130391 | NM_131714 | ctbp1 | C-terminal binding protein 1 | 4.5658 | 0.015165 | 3.0543 | 0.024246 | | |
| A_15_P738871 | NM_213184 | nfkbia | nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha a | 4.4084 | 0.027387 | 2.3192 | 0.004300 | 2.5766 | 0.000001 |
| A_15_P163066 | NM_213184 | nfkbia | nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha a | 4.3159 | 0.043704 | 2.8164 | 0.018079 | | |
| A_15_P663191 | NM_200820 | uqcr1 | ubiquinol-cytochrome c reductase core protein I | 4.1259 | 0.048021 | 2.7593 | 0.016176 | | |
| A_15_P150166 | NM_200820 | uqcr1 | ubiquinol-cytochrome c reductase core protein I | 4.2987 | 0.013557 | 3.5086 | 0.000015 | 2.5709 | 0.000150 |
| | | | | 3.4917 | 0.022489 | 3.2650 | 0.000644 | 2.2627 | 0.000209 |

| Probe ID | NCBI Accession | Gene symbol | Gene name | 40% epiboly ser. 1 | | 40% epiboly ser. 2 | | 1k cell stage | |
|--------------|----------------|----------------|---|--------------------|----------|--------------------|----------|---------------|----------|
| | | | | fold change | P-value | fold change | P-value | fold change | P-value |
| A_15_P483645 | NV_200820 | uqcrc1 | ubiquinol-cytochrome c reductase core protein 1 | 2.2800 | 0.028752 | 2.1452 | 0.000517 | 2.0564 | 0.003212 |
| A_15_P172296 | NV_001080683 | zgc158647 | zgc158647 | 4.2956 | 0.034096 | 4.3250 | 0.000008 | | |
| A_15_P116590 | NV_200669 | ptpn12 | protein tyrosine phosphatase, non-receptor type 12 | 4.2120 | 0.027245 | 3.3165 | 0.000100 | 2.5639 | 0.000063 |
| A_15_P688951 | NV_212890 | ift122 | intraflagellar transport 122 homolog (Chlamydomonas) | 4.2072 | 0.022764 | 3.0486 | 0.006474 | 2.5090 | 0.000057 |
| A_15_P117799 | NV_212890 | ift122 | intraflagellar transport 122 homolog (Chlamydomonas) | 4.1006 | 0.025711 | 3.0371 | 0.005423 | 2.4954 | 0.000026 |
| A_15_P763341 | NV_212890 | ift122 | intraflagellar transport 122 homolog (Chlamydomonas) | 2.6954 | 0.019962 | 2.6625 | 0.000044 | 2.0257 | 0.001654 |
| A_15_P651386 | NV_001114315 | zgc:174178 | zgc:174178 | 4.1307 | 0.023452 | 2.8680 | 0.004252 | 3.5424 | 0.037606 |
| A_15_P579202 | CK869927 | dus2l | dihydropyridine synthase 2-like, SVM1 homolog (S. cerevisiae) | 4.0580 | 0.032910 | 4.8394 | 0.000010 | | |
| A_15_P631946 | NV_130924 | ptpn1 | protein tyrosine phosphatase, non-receptor type 1 | 3.9014 | 0.020080 | 2.4376 | 0.009445 | 4.7908 | 0.000002 |
| A_15_P758926 | NV_001089466 | sreb12 | sterol regulatory element binding transcription factor 2 | 3.8852 | 0.031647 | 4.9350 | 0.005328 | 2.0985 | 0.001940 |
| A_15_P575067 | FM986818 | upf2 | UPF2 regulator of nonsense transcripts homolog (yeast) | 3.8817 | 0.027245 | 2.6247 | 0.007462 | 2.0844 | 0.024912 |
| A_15_P265551 | BC091967 | im:7154473 | im:7154473 | 3.8499 | 0.022764 | 3.6419 | 0.000822 | | |
| A_15_P373490 | A_15_P373490 | | | 3.6531 | 0.022972 | 2.8200 | 0.000055 | | |
| A_15_P106358 | NV_001923521 | | | 3.6278 | 0.033843 | 3.1762 | 0.003973 | | |
| A_15_P207056 | NV_001045076 | hspa14 | heat shock protein 14 | 3.6152 | 0.027068 | 2.1286 | 0.002370 | 2.0786 | 0.000406 |
| A_15_P557942 | NV_001337197 | LOC796836 | hypothetical LOC796836 | 3.5900 | 0.038506 | 3.1305 | 0.008005 | | |
| A_15_P363385 | NV_001100057 | LOC796505 | putative ATP-dependent RNA helicase DHX33-like | 3.5657 | 0.016014 | 2.1253 | 0.029137 | 2.3448 | 0.000038 |
| A_15_P174341 | NV_001080204 | cyp11b2 | cytochrome P450, family 11, subfamily B, polypeptide 2 | 3.4817 | 0.034764 | 3.6625 | 0.003748 | | |
| A_15_P599097 | NV_679659 | LOC560758 | hypothetical LOC560758 | 3.4652 | 0.029819 | 2.2233 | 0.001884 | | |
| A_15_P109888 | BC045923 | | | 3.4628 | 0.035063 | 2.4164 | 0.013203 | | |
| A_15_P117951 | NV_201016 | brpf1 | bromodomain and PHD finger containing, 1 | 3.3656 | 0.029286 | 3.8239 | 0.000738 | 2.2399 | 0.013519 |
| A_15_P688971 | EU486162 | brpf1 | bromodomain and PHD finger containing, 1 | 3.1365 | 0.042884 | 4.3448 | 0.000646 | 2.2817 | 0.000112 |
| A_15_P625141 | EU486162 | brpf1 | bromodomain and PHD finger containing, 1 | 3.0850 | 0.041105 | 3.7735 | 0.002462 | 2.1669 | 0.000608 |
| A_15_P225751 | NV_001128240 | srch211-51n9.7 | srch211-51n9.7 | 3.3227 | 0.036724 | 3.3515 | 0.002352 | | |
| A_15_P173058 | NV_207058 | setd2 | SET domain, bifurcated 2 | 3.2375 | 0.026434 | 4.4855 | 0.002258 | | |
| A_15_P156941 | NV_680769 | LOC561717 | RAB27B, member RAS oncogene family-like | 3.1694 | 0.034901 | 3.6337 | 0.024574 | | |
| A_15_P48060 | NV_201287 | alas1 | aminolevulinic acid, delta-, synthetase 1 | 3.1398 | 0.018770 | 3.2390 | 0.001263 | 2.2629 | 0.000099 |
| A_15_P110209 | NV_199811 | rap1 | replication protein A1 | 3.0261 | 0.022489 | 2.2664 | 0.009545 | 3.2345 | 0.000021 |
| A_15_P735146 | NV_001098740 | sdhb | succinate dehydrogenase complex, subunit B, iron sulfur (lp) | 2.8262 | 0.014150 | 2.0932 | 0.000065 | | |
| A_15_P140766 | NV_001098740 | sdhb | succinate dehydrogenase complex, subunit B, iron sulfur (lp) | 2.6956 | 0.009667 | 2.1842 | 0.000371 | | |
| A_15_P175171 | NV_001080627 | zgc:158458 | zgc:158458 | 2.8173 | 0.037247 | 2.3710 | 0.046211 | | |
| A_15_P177871 | BC127558 | | | 2.7555 | 0.043781 | 2.6902 | 0.000893 | 2.1147 | 0.000691 |
| A_15_P206581 | NV_001020640 | zgc:111991 | zgc:111991 | 2.7439 | 0.009667 | 2.8453 | 0.000416 | 2.4998 | 0.000131 |
| A_15_P154566 | BC076187 | zgc:92707 | zgc:92707 | 2.6912 | 0.035063 | 2.3230 | 0.001109 | 2.2831 | 0.000350 |
| A_15_P118450 | NV_205658 | tmem85 | transmembrane protein 85 | 2.6894 | 0.030955 | 2.6079 | 0.000371 | | |
| A_15_P335214 | NV_001045108 | fam65a | family with sequence similarity 65, member A | 2.6817 | 0.031476 | 4.5847 | 0.000782 | | |
| A_15_P119087 | NV_200756 | rwdd3 | RWD domain containing 3 | 2.6383 | 0.024849 | 2.3633 | 0.015122 | | |
| A_15_P113919 | NV_001098481 | atp5b | ATP synthase, H+ transporting, mitochondrial FO complex, subunit e, duplicate b | 2.6199 | 0.012940 | 2.5968 | 0.016895 | | |
| A_15_P676161 | NV_001098481 | atp5b | ATP synthase, H+ transporting, mitochondrial FO complex, subunit e, duplicate b | 2.5918 | 0.039698 | 2.2782 | 0.017036 | | |
| A_15_P105730 | NV_200037 | vps29 | vacuolar protein sorting 29 (yeast) | 2.5144 | 0.026434 | 2.4834 | 0.000482 | | |
| A_15_P774386 | NV_213722 | slc25a3b | solute carrier family 25 (mitochondrial carrier, phosphate carrier), member 3b | 2.5127 | 0.013604 | 2.0360 | 0.003157 | | |
| A_15_P735516 | NV_213722 | slc25a3b | solute carrier family 25 (mitochondrial carrier, phosphate carrier), member 3b | 2.4751 | 0.016673 | 2.1085 | 0.000131 | | |
| A_15_P347760 | NV_213722 | slc25a3b | solute carrier family 25 (mitochondrial carrier, phosphate carrier), member 3b | 2.4717 | 0.010082 | 2.0843 | 0.000260 | | |
| A_15_P671851 | NV_213722 | slc25a3b | solute carrier family 25 (mitochondrial carrier, phosphate carrier), member 3b | 2.4379 | 0.013111 | 2.0883 | 0.000061 | | |
| A_15_P383430 | | | | 2.4741 | 0.031539 | 2.2135 | 0.003156 | | |
| A_15_P264151 | EE321734 | tsc2 | tuberous sclerosis 2 | 2.4493 | 0.030782 | 2.3275 | 0.002185 | | |
| A_15_P149911 | NV_001100147 | zgc:158376 | zgc:158376 | 2.4108 | 0.013604 | 2.3436 | 0.000523 | | |
| A_15_P728111 | NV_131031 | bactin1 | bactin1 | 2.4081 | 0.034901 | 2.1849 | 0.000262 | | |
| A_15_P364815 | NV_001007307 | pcf11 | cleavage and polyadenylation factor subunit, homolog (S. cerevisiae) | 2.3799 | 0.026223 | 2.0522 | 0.020272 | 3.2450 | 0.002701 |
| | | | | 2.3559 | 0.027245 | 2.0154 | 0.001883 | | |

| Probe ID | NCBI Accession | Gene symbol | Gene name | 40%epiboly set 1 | | 40%epiboly set 2 | | 1k cell stage | |
|--|----------------|-------------------|--|------------------|----------|------------------|----------|---------------|----------|
| | | | | fold change | P-value | fold change | P-value | fold change | P-value |
| A_15_P320741 | XM_701870 | im:6893790 | im:6893790 | 2.3048 | 0.043547 | 2.1848 | 0.006681 | 2.5482 | 0.003193 |
| A_15_P349800 | EG575269 | im:6893790 | im:6893790 | 2.1338 | 0.034618 | 2.5505 | 0.020406 | 2.4071 | 0.024028 |
| A_15_P261032 | NM_001114729 | limch1 | lim and calponin homology domains 1 | 2.2360 | 0.034087 | 2.6212 | 0.005500 | | |
| A_15_P176151 | NM_001037409 | itgb1bp1 | integrin beta 1 binding protein 1 | 2.2312 | 0.025846 | 2.9947 | 0.000008 | | |
| A_15_P193961 | NM_001077327 | zgc:154086 | zgc:154086 | 2.1914 | 0.016605 | 2.1771 | 0.000931 | | |
| A_15_P485928 | EG578017 | LOC799656 | hypothetical LOC799656 | 2.1790 | 0.043784 | 2.1693 | 0.003845 | | |
| A_15_P194971 | NM_001044871 | selo | selenoprotein O | 2.1718 | 0.031539 | 2.6206 | 0.000854 | | |
| A_15_P391660 | XM_002664763 | | | 2.1610 | 0.009667 | 2.3006 | 0.003649 | | |
| A_15_P507537 | BC131855 | ctnna2 | catenin (cadherin-associated protein), alpha 2 | 2.1284 | 0.019962 | 2.0640 | 0.003465 | | |
| A_15_P169496 | NM_001098772 | zgc:165502 | zgc:165502 | 2.0340 | 0.037197 | 2.0576 | 0.035735 | | |
| Up-regulation in E313 MO-injected embryos | | | | | | | | | |
| A_15_P183191 | XM_681611 | LOC58408 | similar to Sid182N13.3 (novel tropomyosin) | 27.9846 | 0.018560 | 11.2010 | 0.000346 | | |
| A_15_P105149 | NM_201008 | pfkfb2 | 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2 | 26.2851 | 0.022510 | 30.7103 | 0.000061 | | |
| A_15_P195421 | NM_001044330 | sid:keyp-59c12.1 | sid:keyp-59c12.1 | 16.1991 | 0.032376 | 5.0195 | 0.018835 | | |
| A_15_P679396 | CN508271 | bcl3 | B-cell CLL/lymphoma 3 | 15.2980 | 0.031288 | 6.5004 | 0.045988 | | |
| A_15_P112117 | NM_198067 | mmp2 | matrix metalloproteinase 2 | 9.4234 | 0.046418 | 7.8065 | 0.049198 | | |
| A_15_P412605 | XM_001332529 | LOC792922 | hypothetical LOC792922 | 8.4515 | 0.006969 | 3.5632 | 0.014372 | | |
| A_15_P104381 | NM_205569 | fos | v-fos FBJ murine osteosarcoma viral oncogene homolog | 8.2649 | 0.016605 | 3.5203 | 0.028547 | | |
| A_15_P118741 | NM_001002352 | rspo1 | R-spondin homolog (Xenopus laevis) | 7.3995 | 0.023838 | 4.7144 | 0.016413 | | |
| A_15_P176731 | NM_001002660 | sesn1 | sestrin 1 | 7.3171 | 0.014111 | 2.9703 | 0.034726 | | |
| A_15_P597407 | XM_685829 | | | 7.1357 | 0.013111 | 4.8619 | 0.013586 | | |
| A_15_P661791 | NM_001040250 | zgc:136874 | zgc:136874 | 7.1293 | 0.030778 | 2.4948 | 0.020158 | | |
| A_15_P148566 | NM_001040250 | zgc:136874 | zgc:136874 | 6.7082 | 0.028752 | 2.3390 | 0.016868 | | |
| A_15_P741971 | | | | 6.6983 | 0.044180 | 3.5917 | 0.060023 | | |
| A_15_P700161 | XM_001202025 | LOC100149324 | hypothetical LOC100149324 | 6.4997 | 0.042884 | 22.4438 | 0.001564 | | |
| A_15_P171926 | NM_001030197 | aplnrb | apelin receptor b | 6.4174 | 0.030782 | 2.6800 | 0.027899 | | |
| A_15_P115138 | NM_001003551 | zgc:101136 | zgc:101136 | 6.4000 | 0.033283 | 5.3993 | 0.001237 | | |
| A_15_P674801 | | | | 6.1171 | 0.023838 | 2.9235 | 0.002429 | | |
| A_15_P592557 | CK866640 | wu:ff5h10 | wu:ff5h10 | 6.0873 | 0.042884 | 3.0039 | 0.024404 | | |
| A_15_P106342 | BC127394 | | | 5.8454 | 0.014111 | 2.8627 | 0.035466 | | |
| A_15_P224171 | NM_001005938 | park7 | parkinson disease (autosomal recessive, early onset) 7 | 5.3456 | 0.018392 | 3.2869 | 0.002546 | | |
| A_15_P153716 | NM_001080586 | zgc:158319 | zgc:158319 | 5.2729 | 0.010082 | 4.1895 | 0.020383 | | |
| A_15_P741576 | NM_001080586 | zgc:158319 | zgc:158319 | 3.6165 | 0.037820 | 2.6731 | 0.032191 | | |
| A_15_P121276 | NM_131340 | thrb | thyroid hormone receptor beta | 4.8449 | 0.037247 | 4.1662 | 0.037558 | | |
| A_15_P109196 | BC134934 | LOC559239 | hypothetical LOC559239 | 4.5953 | 0.033095 | 2.3114 | 0.035011 | | |
| A_15_P736731 | NM_200150 | ankrd12 | ankyrin repeat domain 12 | 4.5779 | 0.031647 | 2.3993 | 0.007278 | | |
| A_15_P105747 | NM_200150 | ankrd12 | ankyrin repeat domain 12 | 4.3051 | 0.033095 | 2.3107 | 0.009262 | | |
| A_15_P110074 | NM_200150 | ankrd12 | ankyrin repeat domain 12 | 4.1144 | 0.028752 | 2.3358 | 0.005853 | | |
| A_15_P101258 | NM_205630 | zgc:77228 | zgc:77228 | 4.5760 | 0.029330 | 2.7734 | 0.029168 | | |
| A_15_P438810 | NM_213038 | zgc:73230 | zgc:73230 | 4.4661 | 0.017883 | 2.3311 | 0.004534 | | |
| A_15_P207341 | NM_213038 | zgc:73230 | zgc:73230 | 4.0325 | 0.030955 | 2.1963 | 0.009293 | | |
| A_15_P713198 | | | | 4.3395 | 0.034751 | 2.7055 | 0.045604 | | |
| A_15_P310476 | NM_00110837 | il13ra1 | interleukin 13 receptor, alpha 1 | 4.2096 | 0.034764 | 2.8258 | 0.030394 | | |
| A_15_P100834 | NM_001002431 | rgs20 | regulator of G-protein signaling 20 | 4.1142 | 0.013736 | 2.5872 | 0.027562 | | |
| A_15_P685295 | CT585688 | | | 4.0816 | 0.043894 | 4.7882 | 0.004868 | | |
| A_15_P465735 | NM_001045008 | si:ch211-129c21.1 | si:ch211-129c21.1 | 3.9680 | 0.030156 | 3.7943 | 0.001734 | | |
| A_15_P266736 | BC100129 | | | 3.9393 | 0.032524 | 2.3384 | 0.016849 | | |
| A_15_P100697 | BC047842 | pcm1 | pericentriolar material 1 | 3.8351 | 0.019192 | 2.0144 | 0.000448 | | |

| Probe ID | NCBI Accession | Gene symbol | Gene name | 40%epiboly set 1 | | 40%epiboly set 2 | | 1k cell stage | |
|--------------|----------------|---------------|--|------------------|----------|------------------|----------|---------------|-----------------|
| | | | | fold change | P-value | fold change | P-value | fold change | P-value |
| A_15_P750511 | BC122445 | sidakey-7n6.2 | sidakey-7n6.2 | 3.8160 | 0.016673 | 2.0164 | 0.046145 | | |
| A_15_P673671 | | | | 3.7958 | 0.028698 | 4.2517 | 0.001333 | | |
| A_15_P164136 | NM_001044766 | klf71 | Kruppel-like factor 7 (ubiquitous), like | 3.7208 | 0.035171 | 2.5570 | 0.041920 | | |
| A_15_P268891 | NM_001163832 | LOC100002616 | BTB (POZ) domain containing 9-like | 3.5647 | 0.028752 | 2.7015 | 0.002783 | | |
| A_15_P170906 | NM_001043321 | mif | macrophage migration inhibitory factor | 3.1589 | 0.039423 | 2.7685 | 0.001705 | | |
| A_15_P109354 | | | | 3.1225 | 0.043547 | 2.5151 | 0.007911 | | |
| A_15_P532687 | BC090702 | imi:7140827 | imi:7140827 | 3.0924 | 0.033095 | 2.1890 | 0.013032 | | |
| A_15_P145721 | NM_001126395 | ankle2 | ankyrin repeat and LEM domain containing 2 | 2.9609 | 0.046679 | 2.1792 | 0.045645 | | |
| A_15_P319766 | NM_001077296 | zgc:153351 | zgc:153351 | 2.7644 | 0.031276 | 2.9114 | 0.018084 | | |
| A_15_P114423 | NM_213316 | adhgap11a | Rho GTPase activating protein 11A | 2.6403 | 0.041915 | 2.2719 | 0.013258 | | |
| A_15_P104538 | NM_001003481 | zgc:91811 | zgc:91811 | 2.5907 | 0.033095 | 2.0118 | 0.049386 | | |
| A_15_P686816 | XM_002667627 | | | 2.3823 | 0.042690 | 2.0937 | 0.004989 | | |
| A_15_P710631 | XM_002662404 | prkca | protein kinase C, alpha | 2.2206 | 0.031543 | 3.1916 | 0.044802 | | |
| A_15_P117599 | NM_001002470 | surf4l | surfeit gene 4, like | 2.0362 | 0.044649 | 2.1640 | 0.013924 | | |
| | | | | | | | | | 2.3306 0.002198 |

Table 4: Microarray Gene Expression Analysis: E313 MO-injected Embryos vs mock-MO injected Embryos. Differentially expressed genes with a fold change ≥ 2 ($p \leq 0.05$) from E313 MO-injected embryos compared to the mock MO-injected embryos. List is shown in the order of the down-regulated, to the up-regulated genes. The high-light part are the genes detected by different probes.

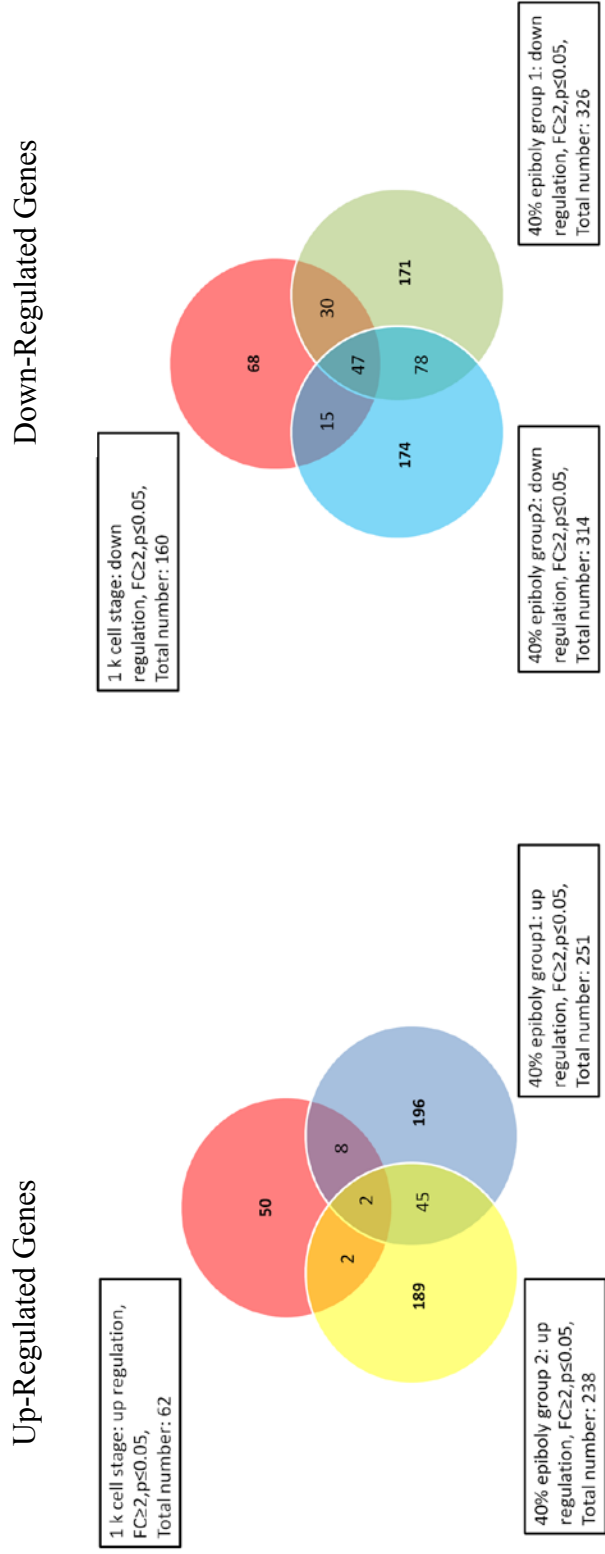


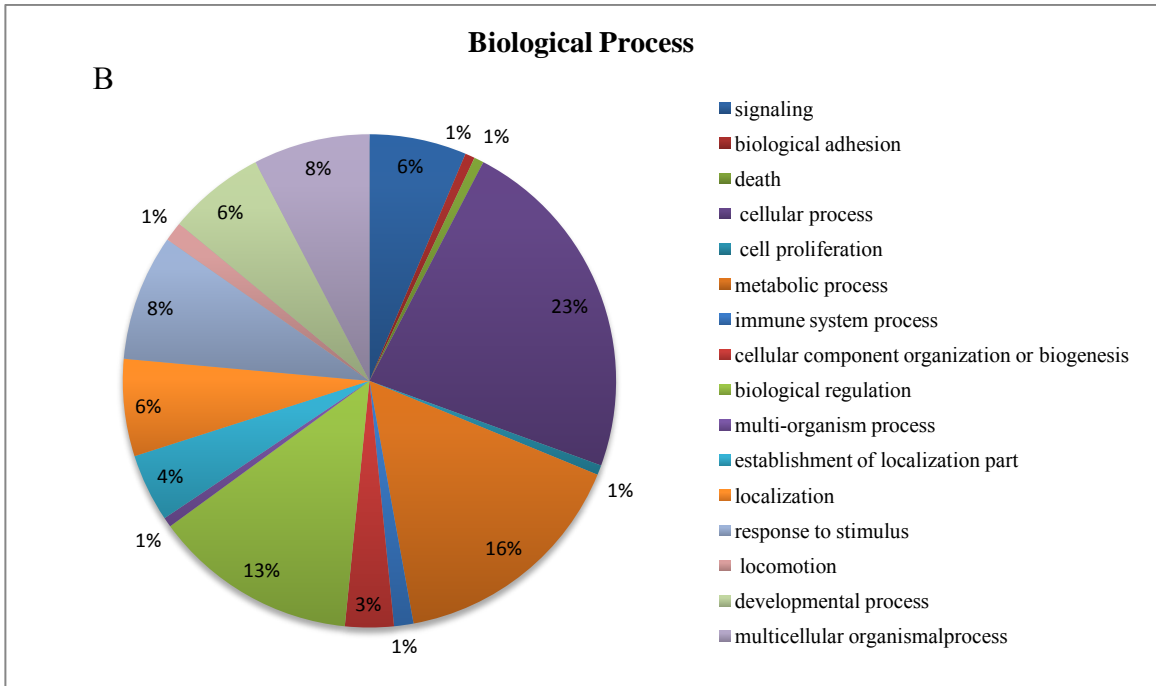
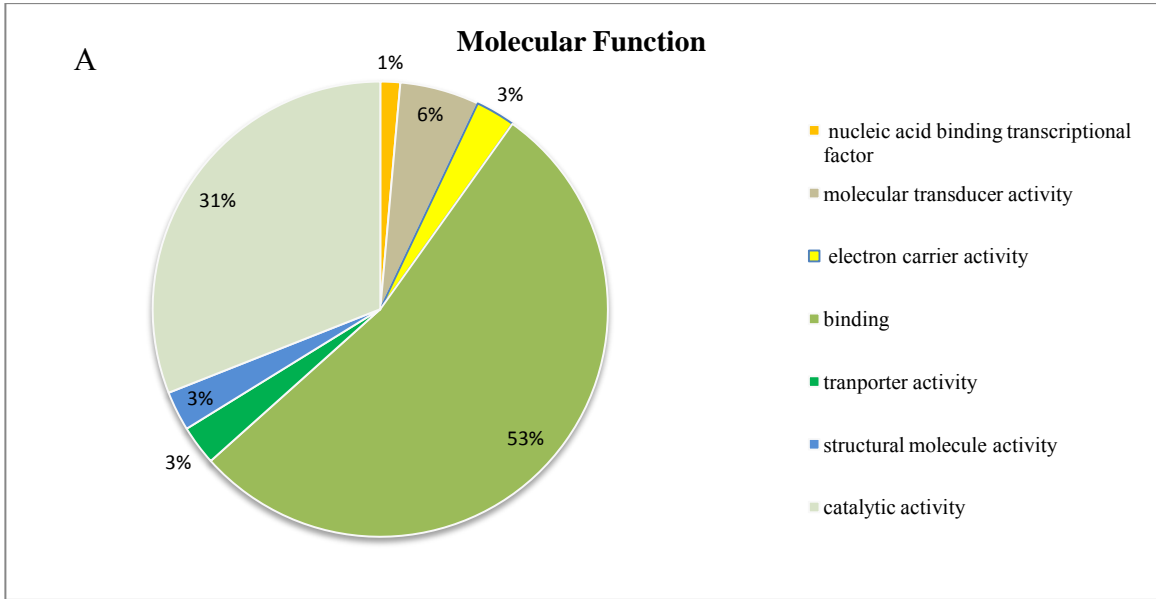
Figure 5: Genes differentially regulated by IRF6 during early embryogenesis. Venn diagram shows the number of differentially regulated genes detected in *irf6* knockdown morphants at 1K cell and 2 groups of 40% epiboly stage. 125 down-regulated and 47 up-regulated genes were consistently detected in 2 groups of 40% epiboly stage. Among them, 2 up-regulated and 47 down-regulated genes were also detected differentially expressed at 1k stage

3.2: Gene ontology study of differentiated expressed genes in E3I3 MO-injected embryos

The 172 genes found differentially expressed in both 40% epiboly data sets were subjected to Gene Ontology (GO) analysis to allow an interpretation on their putative functions, as gene ontology provides the consistent description of attributes of genes and gene products across species (Consortium, 2000). The GO analysis includes molecular function, biological process and cellular component catalogs, each of which includes several sub categories.

For molecular function, binding was the most outstanding function, with 53% of genes related to binding activities (Figure 6). *Irf6* is a transcriptional factor and likely regulates the expression of many down-stream targets. As approximately half of the putative candidate targets identified in the microarray are related to binding activities, *Irf6* could significantly affect interactions between molecules (e.g. DNA-protein binding). Catalytic activity was also a noticeable function as 31% of the genes have functions related to catalytic activity.

Approximately 20% of genes were tagged as processing cellular and metabolic functions. Cellular processes include cell communication, cellular senescence, and programmed cell death. Metabolic processes include transformation of small molecules and macromolecular processes, such as DNA repair and replication, protein synthesis and degradation.



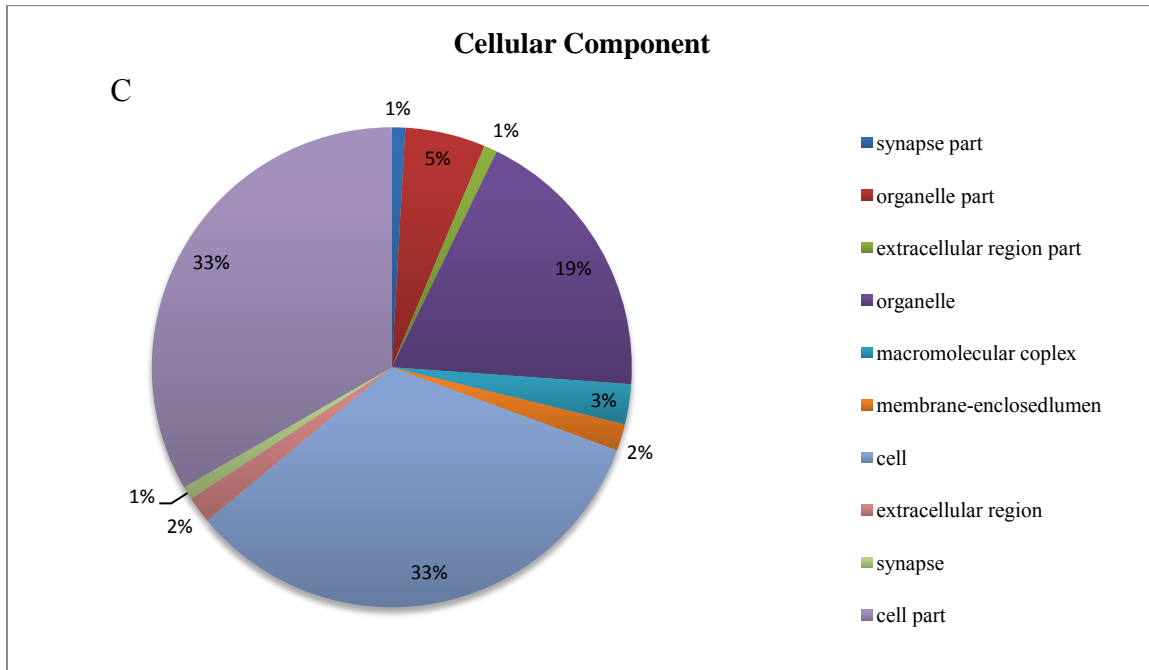


Figure 6: Gene ontology analysis of differentially expressed genes. The GO analysis includes (A) molecular function, (B) biological process and (C) cellular component catalogs, each of which includes several sub categories.

3.3: Microarray differential gene expression validation

Among the differentially expressed genes, *cyr61* and *mapkapk3* were the most highly down-regulated, with 123 and 109 fold reduction respectively. CYR61 is an extracellular matrix-associated protein involved in cell adhesion, cell migration and cell proliferation (Tatiana, 2001). MAPKAPK3 is a member of the Ser/Thr protein kinase family, known to interact with E47, which is involved in the regulation of tissue-specific gene expression and cell differentiation (Neufeld et al., 2000).

Reverse transcription PCR (RT-PCR) of biological quadruplicates was performed to validate the differential expression of *cyr61* and *mapkapk3*. The RT-PCR results confirm the down-regulation of these two genes, consistent with the microarray expression data from both the 40% epiboly data sets (1 and 2). Both *cyr61* and *mapkapk3* were not found differentially expressed at 1k cell stage. This is consistent with an mRNA deep sequencing analysis of the transcriptome dynamics during zebrafish embryonic stages, where *cyr61* and *mapkapk3* transcripts were detected at 3.5h or later (Aanes et al., 2011), which is after the 1k cell stage.

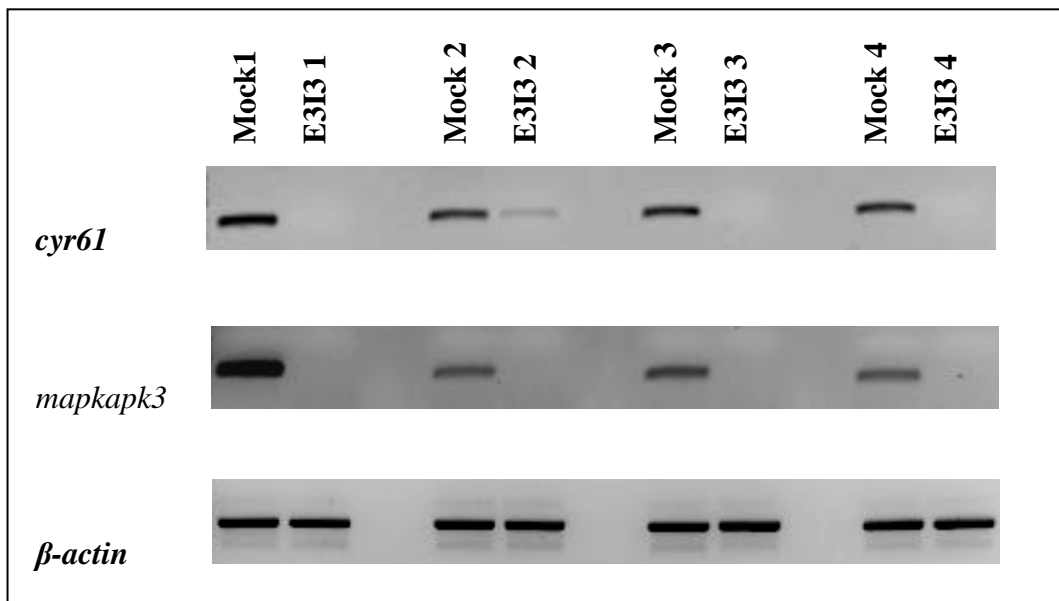


Figure 7: Validation of differentially expressed *cyr61* and *mapkapk3*

The down-regulation of *cyr61* and *mapkapk3* are confirmed with 4 replicates of E3I3 MO-injected and mock MO-injected samples by using reverse-transcription PCR.

3.4: *cyr61* and *mapkapk3* are direct downstream targets of Irf6

In addition to the remarkable down-regulation of *cyr61* and *mapkapk3* expression in E3I3-MO injected embryos, I identified sequence matching the canonical IRF6 binding site NAC^C/T_TGAAACN (Botti et al., 2011) in both genes, AACCGAAACT and GACCGAAACA respectively. To determine whether *cyr61* and *mapkapk3* are direct downstream targets of Irf6 and directly bound by Irf6 protein, electrophoretic mobility shift assays (EMSA) were performed by using both Irf6 full length protein and E3I3 truncated protein. His-tagged recombinant Irf6 proteins were generated using the TNT wheat germ expression system, followed by His-tag protein purification. Successful protein expression was detected using anti-his tag antibody (Figure 8A). For the EMSA reaction, a purified TNT wheat germ lysate without any template was used as the negative control. Purified Irf6 full length protein and E3I3 protein truncated protein were mixed with a double-stranded DNA containing the putative IRF6-binding sequence of *cyr61* and *mapkapk3*. Both Irf6 full length protein and E3I3 truncated protein showed binding and gel-shifted migration of the *cyr61* and *mapkapk3* DNA fragments (Figure 8B). It is noteworthy that the E3I3 truncated protein – DNA complex run at the same height as the full-length Irf6 protein. Considering the protein and DNA complex were run in a native gel, this shift may be caused by the oligomerization of the E3I3 truncated protein. The competition with a non-labeled oligo was done to demonstrate that the binding of full-length and truncated Irf6 protein to both *cyr61* and *mapkapk3* were specific (Figure 8 C and D).

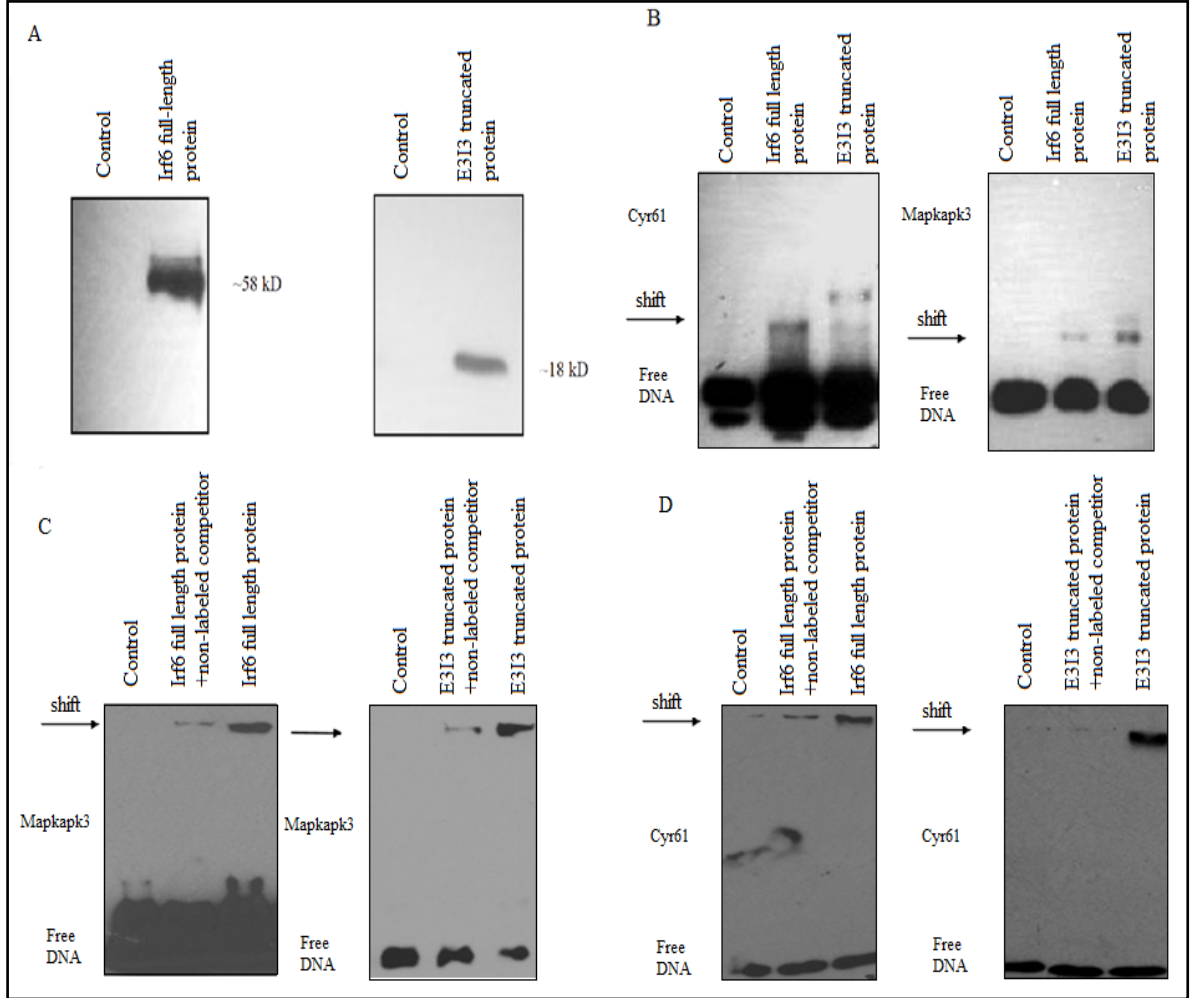


Figure 8: *cyr61* and *mapkapk3* are directly bound by Irf6 and E3I3 truncated protein. (A). Western blotting of His-tagged Irf6 full length protein and E3I3 are expressed by TNT wheat germ expression system and purification. The control is the TNT wheat germ lysate without any DNA template and purified under the same condition. Protein samples are blotted with anti-his tag antibody. (B). EMSA showing binding of Irf6 full length and E3I3 truncated protein to putative IRF6 binding site of *cyr61* and *mapkapk3*. (C) and (D) EMSA with non-labeled competitors. Arrows indicate gel-shifted DNA band.

3.5: Preliminary morphology study of *cyr61* and *mapkapk3* MO blocked embryos

A preliminary assessment of the role of *Cyr61* and *Mapkapk3* in early embryonic development was undertaken. Transcriptional blocking MO targeting *cyr61* and *mapkapk3* were injected into the embryos at 1 mM concentration, with uninjected and 1mM standard MO injected embryos used as controls.

mapkapk3 MO-injected embryos were grossly normal but displayed a kinked notochord and an aberrant epithelial layer of the skin (Figure 10). Survival ratios showed no significant difference compared to the STD MO-injected embryos ($p=0.92$).

For *cyr61*, injection of the translational MO resulted in 25% embryonic lethality 24hpf, which was significantly different from the STD MO-injected group ($p \leq 0.05$). Surviving embryos showed gross development defects and tissue disorganization of the cephalic region (Figure 11). All the embryos died by 3dpf.

The preliminary morphology study of *cyr61* and *mapkapk3* MO perturbed embryos did not reproduce the early rupture phenotype of *irf6* E3I3-MO injected embryos. However, the aberrant skin epithelial layer in *mapkapk3* MO-injected embryos is suggestive of a defect of skin epithelial development, reminiscent of *Irf6* knock-out mouse skin defects. This observation suggests that the epithelial defect caused by loss of function of *Irf6* may be mediated through *mapkapk3*. The gross developmental defects observed in *cyr61* knockdown embryos are also generally consistent with the severe developmental defects

observed in *Irf6* knockdown mice, suggesting that they may be mediated through *cyr61*.

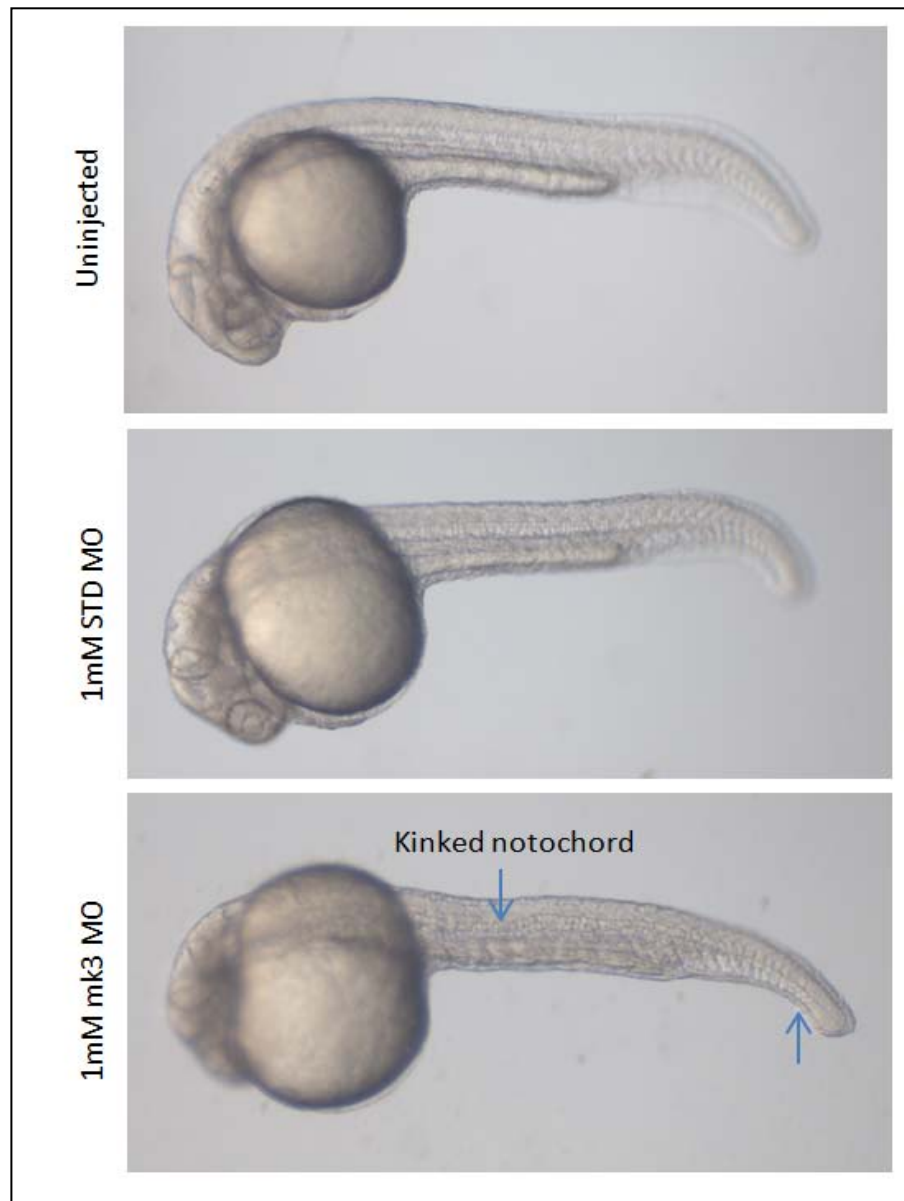
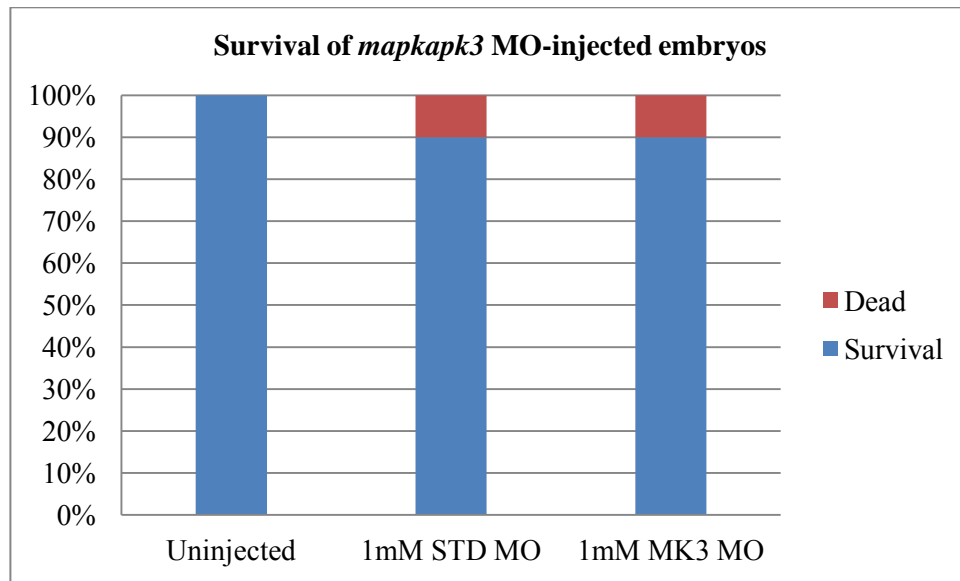


Figure 9: *mapkapk3* MO-injected embryos show a defect of epithelial layer. Embryos injected with 1mM *mapkapk3* MO developed grossly normally except for the kinked notochord and an aberrant epithelial layer at 24 hpf (arrow).



| | Number of injected /uninjected at 1-cell stage | Number of survival embryos at 24hpf | Survival Rate |
|------------------------|---|--|---------------|
| Uninjected | 43 | 43 | 100% |
| 1mM STD MO | 52 | 47 | 90% |
| 1mM <i>mapkapk3</i> MO | 49 | 44 | 90% |

Figure 10: *mapkapk3* MO does not cause lethal phenotype for embryos.

Survival rate of *mapkapk3* MO-injected embryos does not show a significant difference comparing to the mock MO-injected embryos at 24 hours post fertilization.

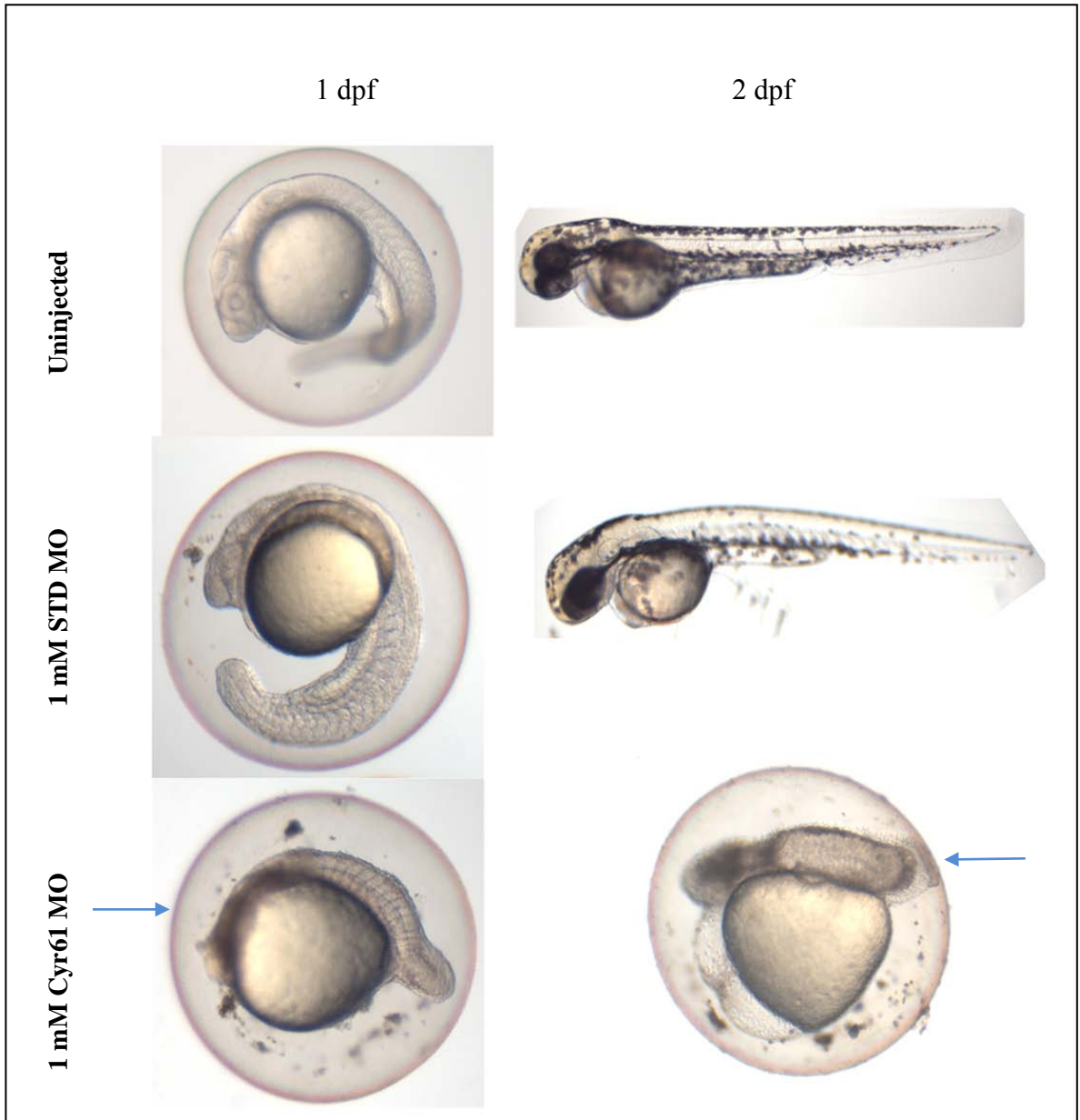
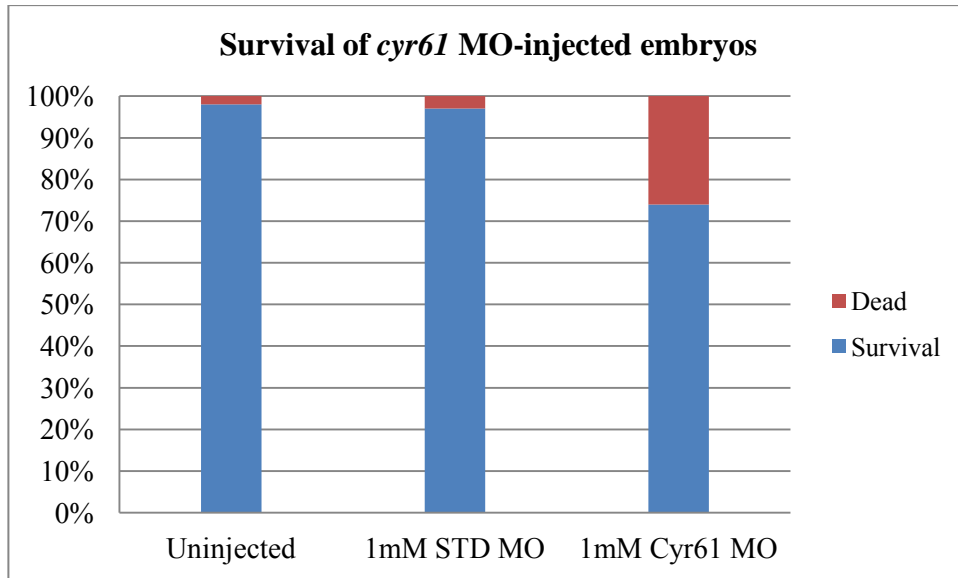


Figure 11: *cyr61* MO-injected embryos show gastrulation defects.

At 24 hpf, *cyr61* MO-injected embryos show severe development defects and an obvious cell death around head region is observed (arrow).



| | Number of injected /uninjected at 1-cell stage | Number of survival embryos at 24hpf | Survival Rate |
|----------------------------|--|-------------------------------------|---------------|
| Uninjected | 142 | 139 | 98% |
| 1mM STD MO | 60 | 58 | 97% |
| 1mM <i>cyr61</i> MO | 65 | 48 | 74% |

Figure 12: One-quarter of *cyr61* MO-injected embryos die after 24 hours. Survival rate of *cyr61* MO-injected embryos (74%) is significant different from the mock MO-injected embryos (97%) (Chi-square test, $p \leq 0.05$) at 24 hour post fertilization.

Chapter IV: Discussion

4.1: Interpretation of expression profile of E3I3 MO-injected embryos: Irf6 functions as an essential transcriptional factor during early development

IRF6 is a unique member of the IRF family of transcription factor genes. Although it shares a highly conserved helix-turn-helix DNA binding domain (DBD) and a less conserved protein interaction domain (PID) of the IRF family, it is neither involved in any regulatory pathways nor known functions of other *IRF* family members. Mutations in the *IRF6* gene have been identified as causative of the allelic autosomal dominant clefting disorders Van der Woude syndrome (VWS; OMIM no. 119300) and popliteal pterygium syndrome (PPS; OMIM no. 119500) (Kondo et al., 2002). *Irf6* is critical in zebrafish development as the introduction of a putative dominant negative *Irf6* containing only the *Irf6* DNA binding domain produces early embryonic lethality (Sabel et al., 2009). An antisense MO (E3I3-MO) targeting at the splice junction of exon 3 and intron 3 of *irf6* pre-mRNA leads to a rupture phenotype of the embryos during gastrulation (unpublished data), reminiscent of the phenotype described in Sabel et al. (2009) and confirms an extremely important role of *Irf6* in gastrulation. Thus, identification of genes that are affected by absence of functional *Irf6* will aid in our understanding of the role of *Irf6* in the regulation of gastrulation.

A genome-wide transcriptome microarray analysis was performed to detect the differential gene expression between mock MO-injected and E3I3 MO-injected embryos at 40% epiboly. 172 genes (125 down-regulated and 47 up-regulated) were identified as differentially regulated at 40% epiboly, just before E3I3 MO-injected embryos start

exhibiting gastrulation stalling. Since the E3I3 MO-injected embryos begin to stall at the 40% epiboly stage, we hypothesized that the downstream molecules disruptions causing this effect must have occurred earlier. Hence, a microarray analysis was also performed at the 1k cell stage and 222 genes (62 up-regulated and 160 down-regulated) were identified.

Among the differentially regulated genes identified at 40% epiboly, 49 (2 up-regulated and 47 down-regulated) were also found to be already differentially expressed at the 1k cell stage. Thus, the differentially regulated genes identified at 40% epiboly can be divided into two groups: the genes dys-regulated as early as the 1k cell stage and genes only dys-regulated at 40% epiboly. The first group of early-expressed genes indicated that *Irf6* has started activating expression of other genes by the 1k cell stage. Given this early time-point, it is also likely that many of these genes are downstream targets directly transactivated by *Irf6*. For genes differentially regulated only at 40% epiboly but not at the 1k cell stage, it is likely that only some of them are direct *Irf6* targets, while others represent genes regulated by the *Irf6* targets.

An important observation from the *Irf6* disruption expression array results was the fact that there were many more down-regulated genes than up-regulated ones, especially for genes at both the 1k cell stage and 40% epiboly stage, 47 down-regulated vs 2 up-regulated, strongly suggesting that *Irf6* acts predominantly as a transcriptional activator rather than a repressor, supporting the findings of an earlier study (Fleming et al., 2009).

4.2 The multi-function role of IRF6

Since the discovery of IRF6 as the causative gene for VWS and PPS (Kondo et al., 2002), numerous functional studies have been carried out to elucidate its role in development, and regulation of cellular processes, such as differentiation, proliferation, apoptosis, autophagy and oncogenesis.

The role of IRF6 in development and disease has been studied in different species. In human, a common IRF6-linked haplotype contains an unobserved mutation attributable to approximately 12% to all common forms of cleft lip and palate (Zuccherro et al., 2004). In mice, loss of IRF6 causes a craniofacial defect with absent external ears, shorter and more rounded snouts and shorter jaws (Ingraham et al., 2006; Richardson et al., 2006). *Irf6*-null mice also exhibit severe skin defects caused by the over proliferation and failure of differentiation of the epithelial layer (Ingraham, Kinoshita et al. 2006, Richardson, Dixon et al. 2006). In zebrafish and *Xenopus*, the introduction of a dominant negative *Irf6* causes gastrulation defects and embryonic rupture near the animal pole (Sabel et al., 2009). The rupture of the embryos during epiboly was postulated to be caused by a failure of EVL integrity. Our GO analysis of differentially regulated genes supports this hypothesis as many cell and cell part component genes were observed.

The observation that *Irf6*-null mice exhibit severe skin defects led to further investigation on the role of IRF6 in differentiation and proliferation of keratinocytes. An *in vitro* study of *Irf6*^{-/-} keratinocytes observed that absence of IRF6 caused a defect of differentiation, while over expression did not promote differentiation, indicating it is necessary but not sufficient

to promote keratinocyte differentiation (Biggs et al., 2012). IRF6 has been reported to function as a primary downstream target of Notch in keratinocytes, and contribute to the regulation of differentiation and repression of tumors (Restivo et al., 2011). Besides, several genes related to cell proliferation and cell differentiation in keratinocyte are directly regulated by IRF6 (Botti et al., 2011). These findings imply that IRF6 likely functions to regulate cell proliferation (Botti et al., 2011). Even though the genes detected differentially expressed in IRF6 knock-down keratinocyte did not show up in our differentially expressed gene list, considering that keratinocytes are highly differentiated, the genes differentially regulated in the keratinocytes may not be exactly the same as those in a pluripotent embryo cell. Our expression array data show that 20% of the differentially expressed genes are involved in cellular processes (cell communication, cellular senescence, and programmed cell death), which is in line with this hypothesis.

4.3 *cyr61* and *mapkapk3* are direct downstream targets of Irf6

Among the differentially regulated genes caused by the induction of E3I3 MO, *cyr61* and *mapkapk3* were outstanding as they were highly down-regulated in E3I3 MO-injected embryos, and further confirmed by RT-PCR analysis. As the expression of these two genes were activated at 3.5 or later (Aanes et al., 2011), they were not detected as differentially expressed genes at 1k cell stage. An electrophoretic mobility shift assays (EMSA) with both Irf6 full length protein and Irf6 E3I3 truncated protein confirmed the direct binding of Irf6 to upstream elements of these two genes containing the canonical DNA-binding sequence. Together, these results provide compelling evidence that *cyr61* and *mapkapk3* are the direct down-stream targets of Irf6.

CYR61 is a multifunctional matricellular protein belonging to the CCN protein family, whose members also include CTGF, Nov, WISP-1, WISP-2, and WISP-3 (Lau, 2011). As a matricellular protein, CYR61 is involved in the regulation of inflammation and wound repair (Chiodoni et al., 2010). The main functions of CYR61 differ depending on the cell type having distinct interaction with integrins and heparan sulfate proteoglycans (HSPGs), (Lau, 2011). With the characters of extracellular matrix, CYR61 is tightly but non-covalently associated with the cell surface, as a result, CYR61 can support cell adhesion and induce adhesive signaling in many types of adherent cell (Chen and Lau, 2009). In human skin fibroblasts, CYR61 supports cell adhesion and leads to the formation of structures critical for cell motility (Chen et al., 2001). With the formation of these critical structures for motility, CYR61 stimulates cell migration in fibroblasts, smooth muscle cells (Grzeszkiewicz et al., 2002) and endothelial cells (Leu et al., 2002). Given the importance of CYR61 in cell adhesion and cell mobility, the down-regulation of *cyr61* in E3I3 MO-injected embryos may be a contributing factor leading to the final rupture in these embryos.

Aside from functions in cell adhesion and cell mobility, CYR61 is also reported to induce cell apoptosis in fibroblasts (Todorovic et al., 2005) and prostate carcinoma cells (Franzen et al., 2009), whereas it is involved in cell survival in endothelial cells (Leu et al., 2002) and breast cancer cells (Lin et al., 2004). Expression of *Cyr61* during mouse embryogenesis is accompanied by development of the skeletal, cardiovascular, and neuronal systems (O'Brien and Lau, 1992), and CYR61 has also been reported to regulate osteoblastic differentiation (Su et al., 2010) and affect cell adhesion (Lau, 2011). CYR61 is also important in embryonic development (Mo and Lau, 2006; Mo et al., 2002). In *Xenopus*,

Cyr61 knockdown causes defects in gastrulation resulting in delay of blastopore closure (Latinkic et al., 2003). In our preliminary morphology study of *cyr61* MO-injected embryos, injection of the translation blocking *cyr61* MO caused around 25% of the embryos to die by 24hpf, which was significantly higher than the STD MO-injected group ($p \leq 0.05$). Surviving embryos showed severe gross developmental defects, and all embryos were dead by 3dpf. It should be noted that unlike *Xenopus*, there are two other paralogs of *cyr61* existing in zebrafish (Fernando et al., 2010), which are not significantly affected by the loss of functional *Irf6* (tested by reverse-transcription PCR).

The other putative direct target of *Irf6* identified in this study, *Mapkapk3*, is a member of the mitogen-activated protein kinase (MAPK) family. *Mapkapk3* is targeted by all 3 cascades of MAPK, ERK, p38, and JNK, and mainly activated by the first two (Luig et al., 2010). *Mapkapk3* and its family member, *Mapkapk2* are bifunctional switches with multiple functions (Gaestel, 2006).

Mapkapk3 and *Mapkapk2* are reported to interact with E47, a helix-loop-helix transcription factor, to repress its transcriptional activity (Neufeld et al., 2000). E47 contains more than 100 potential phosphorylation sites and is known to be phosphorylated in many cell types (Neufeld et al., 2000). It is involved in regulation cell cycle progression, cytokine-mediated signaling, T lineage development and other functions (Schwartz et al., 2006). Since *Mapkapk3* and *Mapkapk2* can phosphorylate E47 and repress its function, they can also conceivably regulate the above processes.

Mapkapk3 and Mapkapk2 are also involved in the phosphorylation of the epithelial keratins, Keratin 18-Ser⁵² and Keratin20-Ser¹³ (Menon et al., 2010). As the Keratin 18-Ser⁵² is a hotspot of phosphorylation modification during the S and G2/M phases of the cell cycle (Liao et al., 1995), Mapkapk3 and Mapkapk2 could affect cell cycle function via their phosphorylation function.

Mapkapk3 MO-injected zebrafish embryos developed grossly normally except for a kinked notochord and an aberrant skin epithelial layer, and embryo survival was unaffected. Given that *Irf6* knock-out mouse show significant skin epithelial defects (Ingraham et al., 2006), the aberrant skin epithelial layer of *mapkapk3*-MO injected embryos suggests that the epithelial defect after *Irf6* perturbation may be mediated via down-regulation of *mapkapk3*.

4.4 Conclusion and future work

In this study, we identified and characterized *cyr61* and *mapkapk3* as target genes of *Irf6* at gastrulation stage in zebrafish by profiling the transcriptome of embryos lack of functional *Irf6* leading by the injection of E3I3 morpholino. The findings gathered from this study will provide novel insights into how IRF6 normally function in vertebrate embryogenesis and also contribute new knowledge into understanding gastrulation process. Moreover, as IRF6 is the causative factor of VWS and PPS, the identification of IRF6 downstream targets which may affect the differentiation of epithelium (*mapkapk3*) will contribute new knowledge into understanding the pathogenesis of human oral clefting.

For the future work, characterization other strongly regulated putative target identified from the Irf6 perturbation screen will enable precise dissection of the contribution of each of these Irf6-regulated genes to both early and late embryonic development, and construct a net work of how Irf6 functions in development. Except for the methodology used here, the chromatin immunoprecipitation (ChIP) will be useful to validate the actual binding of Irf6 to its downstream targets. Besides, Irf6 functions as a transcriptional activator, a luciferase assay needs to be performed to demonstrate the activate effect of the downstream targets.

As the fundamental mechanism for the analysis of the function of a protein translated from a specific gene in vivo, gene modification allows testing the specific functions of the particular protein and to observe the processes that the particular protein could regulate. Constructing of an Irf6 knockin model to recapitulate the mutations identified in human VWS and PPS (eg., R84C) by using clustered, regularly interspaced, short palindromic repeats (CRISPR)–CRISPR-associated (Cas) systems By constructing this mode, we will have a clinically relevant zebrafish orofacial cleft model and have the chance to dissect the mechanism of the pathogenesis of oral clefting.

Reference

Aanes, H., Winata, C.L., Lin, C.H., Chen, J.P., Srinivasan, K.G., Lee, S.G., Lim, A.Y., Hajan, H.S., Collas, P., Bourque, G., Gong, Z., Korzh, V., Alestrom, P., Mathavan, S., 2011. Zebrafish mRNA sequencing deciphers novelties in transcriptome dynamics during maternal to zygotic transition. *Genome research* 21, 1328-1338.

Amsterdam, A., Burgess, S., Golling, G., Chen, W., Sun, Z., Townsend, K., Farrington, S., Haldi, M., Hopkins, N., 1999. A large-scale insertional mutagenesis screen in zebrafish. *Genes & development* 13, 2713-2724.

Auer, T.O., Duroure, K., De Cian, A., Concordet, J.P., Del Bene, F., 2014. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome research* 24, 142-153.

Bailey, C.M., Abbott, D.E., Margaryan, N.V., Khalkhali-Ellis, Z., Hendrix, M.J., 2008. Interferon regulatory factor 6 promotes cell cycle arrest and is regulated by the proteasome in a cell cycle-dependent manner. *Molecular and cellular biology* 28, 2235-2243.

Bailey, C.M., Khalkhali-Ellis, Z., Kondo, S., Margaryan, N.V., Seftor, R.E., Wheaton, W.W., Amir, S., Pins, M.R., Schutte, B.C., Hendrix, M.J., 2005. Mammary serine protease inhibitor (Maspin) binds directly to interferon regulatory factor 6: identification of a novel serpin partnership. *The Journal of biological chemistry* 280, 34210-34217.

Bedell, V.M., Wang, Y., Campbell, J.M., Poshusta, T.L., Starker, C.G., Krug, R.G., 2nd, Tan, W., Penheiter, S.G., Ma, A.C., Leung, A.Y., Fahrenkrug, S.C., Carlson, D.F., Voytas, D.F., Clark, K.J., Essner, J.J., Ekker, S.C., 2012. In vivo genome editing using a high-efficiency TALEN system. *Nature* 491, 114-118.

Ben, J., Jabs, E.W., Chong, S.S., 2005. Genomic, cDNA and embryonic expression analysis of zebrafish IRF6, the gene mutated in the human oral clefting disorders Van der Woude and popliteal pterygium syndromes. *Gene expression patterns : GEP* 5, 629-638.

Betchaku, T., Trinkaus, J.P., 1978. Contact relations, surface activity, and cortical microfilaments of marginal cells of the enveloping layer and of the yolk syncytial and yolk cytoplasmic layers of fundulus before and during epiboly. *The Journal of experimental zoology* 206, 381-426.

Biggs, L.C., Rhea, L., Schutte, B.C., Dunnwald, M., 2012. Interferon regulatory factor 6 is necessary, but not sufficient, for keratinocyte differentiation. *The Journal of investigative dermatology* 132, 50-58.

Botti, E., Spallone, G., Moretti, F., Marinari, B., Pinetti, V., Galanti, S., De Meo, P.D., De Nicola, F., Ganci, F., Castrignano, T., Pesole, G., Chimenti, S., Guerrini, L., Fanciulli, M., Blandino, G., Karin, M., Costanzo, A., 2011. Developmental factor IRF6 exhibits tumor suppressor activity in squamous cell carcinomas. *Proceedings of the National Academy of Sciences of the United States of America* 108, 13710-13715.

Bray, S.J., 2006. Notch signalling: a simple pathway becomes complex. *Nature reviews. Molecular cell biology* 7, 678-689.

Chen, C.C., Chen, N., Lau, L.F., 2001. The angiogenic factors Cyr61 and connective tissue growth factor induce adhesive signaling in primary human skin fibroblasts. *The Journal of biological chemistry* 276, 10443-10452.

Chen, C.C., Lau, L.F., 2009. Functions and mechanisms of action of CCN matricellular proteins. *The international journal of biochemistry & cell biology* 41, 771-783.

Chiodoni, C., Colombo, M.P., Sangaletti, S., 2010. Matricellular proteins: from homeostasis to inflammation, cancer, and metastasis. *Cancer metastasis reviews* 29, 295-307.

Collins, J.E., White, S., Searle, S.M., Stemple, D.L., 2012. Incorporating RNA-seq data into the zebrafish Ensembl genebuild. *Genome research* 22, 2067-2078.

Consortium, T.G.O., 2000. Gene Ontology: tool for the unification of biology. *Nature genetics volume* 25, 25-29.

Detrich, H.W., 3rd, Westerfield, M., Zon, L.I., 1999. Overview of the Zebrafish system. *Methods in cell biology* 59, 3-10.

Eisen, J.S., 1996. Zebrafish make a big splash. *Cell* 87, 969-977.

Fernando, C.A., Conrad, P.A., Bartels, C.F., Marques, T., To, M., Balow, S.A., Nakamura, Y., Warman, M.L., 2010. Temporal and spatial expression of CCN genes in zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists* 239, 1755-1767.

FitzPatrick, D.R., Carr, I.M., McLaren, L., Leek, J.P., Wightman, P., Williamson, K., Gautier, P., McGill, N., Hayward, C., Firth, H., Markham, A.F., Fantes, J.A., Bonthron, D.T., 2003. Identification of SATB2 as the cleft palate gene on 2q32-q33. *Human molecular genetics* 12, 2491-2501.

Fleming, J.A., Song, G., Choi, Y., Spencer, T.E., Bazer, F.W., 2009. Interferon regulatory factor 6 (IRF6) is expressed in the ovine uterus and functions as a transcriptional activator. *Molecular and cellular endocrinology* 299, 252-260.

Franzen, C.A., Chen, C.C., Todorovic, V., Juric, V., Monzon, R.I., Lau, L.F., 2009. Matrix protein CCN1 is critical for prostate carcinoma cell proliferation and TRAIL-induced apoptosis. *Molecular cancer research : MCR* 7, 1045-1055.

Fraser, F.C., 1955. Thoughts on the etiology of clefts of the palate and lip. *Acta genetica et statistica medica* 5, 358-369.

Froster-Iskenius, U.G., 1990. Popliteal pterygium syndrome. *Journal of medical genetics* 27, 320-326.

Gaestel, M., 2006. MAPKAP kinases - MKs - two's company, three's a crowd. *Nature reviews. Molecular cell biology* 7, 120-130.

Grzeszkiewicz, T.M., Lindner, V., Chen, N., Lam, S.C., Lau, L.F., 2002. The angiogenic factor cysteine-rich 61 (CYR61, CCN1) supports vascular smooth muscle cell adhesion and stimulates chemotaxis through integrin $\alpha(6)\beta(1)$ and cell surface heparan sulfate proteoglycans. *Endocrinology* 143, 1441-1450.

Gupta, S., Takebe, N., Lorusso, P., 2010. Targeting the Hedgehog pathway in cancer. *Therapeutic advances in medical oncology* 2, 237-250.

H. William Dietrich, I., Monte Westerfield, and Leonard I. Zon, eds 1999. *Methods in Cell Biology, Volume 59, The Zebrafish: Biology* (Boston: Academic Press, 1999)

Hatada, S., Kinoshita, M., Takahashi, S., Nishihara, R., Sakumoto, H., Fukui, A., Noda, M., Asashima, M., 1997. An interferon regulatory factor-related gene (xIRF-6) is expressed in the posterior mesoderm during the early development of *Xenopus laevis*. *Gene* 203, 183-188.

Hellman, L.M., Fried, M.G., 2007. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature protocols* 2, 1849-1861.

Hoheisel, J.D., 2006. Microarray technology: beyond transcript profiling and genotype analysis. *Nature reviews. Genetics* 7, 200-210.

Holloway, B.A., Gomez de la Torre Canny, S., Ye, Y., Slusarski, D.C., Freisinger, C.M., Dosch, R., Chou, M.M., Wagner, D.S., Mullins, M.C., 2009. A novel role for MAPKAPK2 in morphogenesis during zebrafish development. *PLoS genetics* 5, e1000413.

Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M., Churcher, C., Scott, C., Barrett, J.C., Koch, R., Rauch, G.J., White, S., Chow, W., Kilian, B., Quintais, L.T., Guerra-Assuncao, J.A., Zhou, Y., Gu, Y., Yen, J., Vogel, J.H., Eyre, T., Redmond, S., Banerjee, R., Chi, J., Fu, B., Langley, E., Maguire, S.F., Laird, G.K., Lloyd, D., Kenyon, E., Donaldson, S., Sehra, H., Almeida-King, J., Loveland, J., Trevanion, S., Jones, M., Quail, M., Willey, D., Hunt, A., Burton, J., Sims, S., McLay, K., Plumb, B., Davis, J., Clee, C., Oliver, K., Clark, R., Riddle, C., Elliot, D., Threadgold, G., Harden, G., Ware, D., Mortimore, B., Kerry, G., Heath, P., Phillimore, B., Tracey, A., Corby, N., Dunn, M., Johnson, C., Wood, J., Clark, S., Pelan, S., Griffiths, G., Smith, M., Glithero, R., Howden, P., Barker, N., Stevens, C., Harley, J., Holt, K., Panagiotidis, G., Lovell, J., Beasley, H., Henderson, C., Gordon, D., Auger, K., Wright, D., Collins, J., Raisen, C., Dyer, L., Leung, K., Robertson, L., Ambridge, K., Leongamornlert, D., McGuire, S., Gilderthorp, R., Griffiths, C., Manthravadi, D., Nichol, S., Barker, G., Whitehead, S., Kay, M., Brown, J., Murnane, C., Gray, E., Humphries, M., Sycamore, N., Barker, D., Saunders, D., Wallis, J., Babbage, A., Hammond, S., Mashreghi-Mohammadi, M., Barr, L., Martin, S., Wray, P., Ellington, A., Matthews, N., Ellwood, M., Woodmansey, R., Clark, G., Cooper, J., Tromans, A., Grafham, D., Skuce, C., Pandian, R., Andrews, R., Harrison, E., Kimberley, A., Garnett, J., Fosker, N., Hall, R., Garner, P., Kelly, D., Bird, C., Palmer, S., Gehring, I., Berger, A., Dooley, C.M., Ersan-Urun, Z., Eser, C., Geiger, H., Geisler, M., Karotki, L.,

Kirn, A., Konantz, J., Konantz, M., Oberlander, M., Rudolph-Geiger, S., Teucke, M., Osoegawa, K., Zhu, B., Rapp, A., Widaa, S., Langford, C., Yang, F., Carter, N.P., Harrow, J., Ning, Z., Herrero, J., Searle, S.M., Enright, A., Geisler, R., Plasterk, R.H., Lee, C., Westerfield, M., de Jong, P.J., Zon, L.I., Postlethwait, J.H., Nusslein-Volhard, C., Hubbard, T.J., Roest Crolius, H., Rogers, J., Stemple, D.L., Begum, S., Lloyd, C., Lanz, C., Raddatz, G., Schuster, S.C., 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496, 498-503.

Ingraham, C.R., Kinoshita, A., Kondo, S., Yang, B., Sajan, S., Trout, K.J., Malik, M.I., Dunnwald, M., Goudy, S.L., Lovett, M., Murray, J.C., Schutte, B.C., 2006. Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (*Irf6*). *Nature genetics* 38, 1335-1340.

Jones, M.C., 1988. Etiology of facial clefts: prospective evaluation of 428 patients. *The Cleft palate journal* 25, 16-20.

Kantaputra, P.N., Yamasaki, K., Ishida, T., Kishino, T., Niikawa, N., 2002. A dominantly inherited malformation syndrome with short stature, upper limb anomaly, minor craniofacial anomalies, and absence of *TBX5* mutations: report of a Thai family. *American journal of medical genetics* 111, 301-306.

Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., Schilling, T.F., 1995. Stages of embryonic development of the zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists* 203, 253-310.

Klaus, A., Birchmeier, W., 2008. Wnt signalling and its impact on development and cancer. *Nature reviews. Cancer* 8, 387-398.

Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W., Richardson, P., 2002. Genetic and genomic tools for *Xenopus* research: The NIH *Xenopus* initiative. *Developmental dynamics : an official publication of the American Association of Anatomists* 225, 384-391.

Knapik, E.W., Goodman, A., Ekker, M., Chevrette, M., Delgado, J., Neuhauss, S., Shimoda, N., Driever, W., Fishman, M.C., Jacob, H.J., 1998. A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nature genetics* 18, 338-343.

Knight, A.S., Schutte, B.C., Jiang, R., Dixon, M.J., 2006. Developmental expression analysis of the mouse and chick orthologues of IRF6: the gene mutated in Van der Woude syndrome. *Developmental dynamics : an official publication of the American Association of Anatomists* 235, 1441-1447.

Kondo, S., Schutte, B.C., Richardson, R.J., Bjork, B.C., Knight, A.S., Watanabe, Y., Howard, E., de Lima, R.L., Daack-Hirsch, S., Sander, A., McDonald-McGinn, D.M., Zackai, E.H., Lammer, E.J., Aylsworth, A.S., Ardinger, H.H., Lidral, A.C., Pober, B.R., Moreno, L., Arcos-Burgos, M., Valencia, C., Houdayer, C., Bahuau, M., Moretti-Ferreira, D., Richieri-Costa, A., Dixon, M.J., Murray, J.C., 2002. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. *Nature genetics* 32, 285-289.

Latinkic, B.V., Mercurio, S., Bennett, B., Hirst, E.M., Xu, Q., Lau, L.F., Mohun, T.J., Smith, J.C., 2003. Xenopus Cyr61 regulates gastrulation movements and modulates Wnt signalling. *Development* 130, 2429-2441.

Lau, L.F., 2011. CCN1/CYR61: the very model of a modern matricellular protein. *Cellular and molecular life sciences : CMLS* 68, 3149-3163.

Lepage, S.E., Bruce, A.E., 2010. Zebrafish epiboly: mechanics and mechanisms. *The International journal of developmental biology* 54, 1213-1228.

Leu, S.J., Lam, S.C., Lau, L.F., 2002. Pro-angiogenic activities of CYR61 (CCN1) mediated through integrins α v β 3 and α 6 β 1 in human umbilical vein endothelial cells. *The Journal of biological chemistry* 277, 46248-46255.

Liao, J., Lowthert, L.A., Ku, N.O., Fernandez, R., Omary, M.B., 1995. Dynamics of human keratin 18 phosphorylation: polarized distribution of phosphorylated keratins in simple epithelial tissues. *The Journal of cell biology* 131, 1291-1301.

Lieschke, G.J., Currie, P.D., 2007. Animal models of human disease: zebrafish swim into view. *Nature reviews. Genetics* 8, 353-367.

Lin, M.T., Chang, C.C., Chen, S.T., Chang, H.L., Su, J.L., Chau, Y.P., Kuo, M.L., 2004. Cyr61 expression confers resistance to apoptosis in breast cancer MCF-7 cells by a

mechanism of NF-kappaB-dependent XIAP up-regulation. *The Journal of biological chemistry* 279, 24015-24023.

Little, H.J., Rorick, N.K., Su, L.I., Baldock, C., Malhotra, S., Jowitt, T., Gakhar, L., Subramanian, R., Schutte, B.C., Dixon, M.J., Shore, P., 2009. Missense mutations that cause Van der Woude syndrome and popliteal pterygium syndrome affect the DNA-binding and transcriptional activation functions of IRF6. *Human molecular genetics* 18, 535-545.

Little, J., Cardy, A., Munger, R.G., 2004. Tobacco smoking and oral clefts: a meta-analysis. *Bulletin of the World Health Organization* 82, 213-218.

Lohoff, M., Mak, T.W., 2005. Roles of interferon-regulatory factors in T-helper-cell differentiation. *Nature reviews. Immunology* 5, 125-135.

Luig, C., Kother, K., Dudek, S.E., Gaestel, M., Hiscott, J., Wixler, V., Ludwig, S., 2010. MAP kinase-activated protein kinases 2 and 3 are required for influenza A virus propagation and act via inhibition of PKR. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 24, 4068-4077.

Mamane, Y., Heylbroeck, C., Genin, P., Algarte, M., Servant, M.J., LePage, C., DeLuca, C., Kwon, H., Lin, R., Hiscott, J., 1999. Interferon regulatory factors: the next generation. *Gene* 237, 1-14.

Menon, M.B., Schwermann, J., Singh, A.K., Franz-Wachtel, M., Pabst, O., Seidler, U., Omary, M.B., Kotlyarov, A., Gaestel, M., 2010. p38 MAP kinase and MAPKAP kinases

MK2/3 cooperatively phosphorylate epithelial keratins. *The Journal of biological chemistry* 285, 33242-33251.

Mo, F.E., Lau, L.F., 2006. The matricellular protein CCN1 is essential for cardiac development. *Circulation research* 99, 961-969.

Mo, F.E., Muntean, A.G., Chen, C.C., Stolz, D.B., Watkins, S.C., Lau, L.F., 2002. CYR61 (CCN1) is essential for placental development and vascular integrity. *Molecular and cellular biology* 22, 8709-8720.

Moretti, F., Marinari, B., Lo Iacono, N., Botti, E., Giunta, A., Spallone, G., Garaffo, G., Vernersson-Lindahl, E., Merlo, G., Mills, A.A., Ballaro, C., Alema, S., Chimenti, S., Guerrini, L., Costanzo, A., 2010. A regulatory feedback loop involving p63 and IRF6 links the pathogenesis of 2 genetically different human ectodermal dysplasias. *The Journal of clinical investigation* 120, 1570-1577.

Murray, J.C., 2002. Gene/environment causes of cleft lip and/or palate. *Clinical genetics* 61, 248-256.

Neufeld, B., Grosse-Wilde, A., Hoffmeyer, A., Jordan, B.W., Chen, P., Dinev, D., Ludwig, S., Rapp, U.R., 2000. Serine/Threonine kinases 3pK and MAPK-activated protein kinase 2 interact with the basic helix-loop-helix transcription factor E47 and repress its transcriptional activity. *The Journal of biological chemistry* 275, 20239-20242.

O'Brien, T.P., Lau, L.F., 1992. Expression of the growth factor-inducible immediate early gene *cyr61* correlates with chondrogenesis during mouse embryonic development. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* 3, 645-654.

Pei, W., Noushmehr, H., Costa, J., Ouspenskaia, M.V., Elkhouloun, A.G., Feldman, B., 2007. An early requirement for maternal FoxH1 during zebrafish gastrulation. *Dev Biol* 310, 10-22.

Restivo, G., Nguyen, B.C., Dziunycz, P., Ristorcelli, E., Ryan, R.J., Ozuysal, O.Y., Di Piazza, M., Radtke, F., Dixon, M.J., Hofbauer, G.F., Lefort, K., Dotto, G.P., 2011. IRF6 is a mediator of Notch pro-differentiation and tumour suppressive function in keratinocytes. *The EMBO journal* 30, 4571-4585.

Richardson, R.J., Dixon, J., Malhotra, S., Hardman, M.J., Knowles, L., Boot-Handford, R.P., Shore, P., Whitmarsh, A., Dixon, M.J., 2006. *Irf6* is a key determinant of the keratinocyte proliferation-differentiation switch. *Nature genetics* 38, 1329-1334.

Rizos, M., Spyropoulos, M.N., 2004. Van der Woude syndrome: a review. Cardinal signs, epidemiology, associated features, differential diagnosis, expressivity, genetic counselling and treatment. *European journal of orthodontics* 26, 17-24.

Rohde, L.A., Heisenberg, C.P., 2007. Zebrafish gastrulation: cell movements, signals, and mechanisms. *International review of cytology* 261, 159-192.

Roosen-Runge, E., 1937. Observations of the early development of the zebrafish. *Brachydanio rerio*. . *Anat. Rec.* 70, s103.

Sabel, J.L., d'Alencon, C., O'Brien, E.K., Van Otterloo, E., Lutz, K., Cuykendall, T.N., Schutte, B.C., Houston, D.W., Cornell, R.A., 2009. Maternal Interferon Regulatory Factor 6 is required for the differentiation of primary superficial epithelia in *Danio* and *Xenopus* embryos. *Dev Biol* 325, 249-262.

Salamone, F.N., Myer, C.M., 3rd, 2004. Van der Woude syndrome: the most common cleft syndrome. *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery* 131, 141.

Savitsky, D., Tamura, T., Yanai, H., Taniguchi, T., 2010. Regulation of immunity and oncogenesis by the IRF transcription factor family. *Cancer immunology, immunotherapy : CII* 59, 489-510.

Schwartz, R., Engel, I., Fallahi-Sichani, M., Petrie, H.T., Murre, C., 2006. Gene expression patterns define novel roles for E47 in cell cycle progression, cytokine-mediated signaling, and T lineage development. *Proceedings of the National Academy of Sciences of the United States of America* 103, 9976-9981.

Shimoda, N., Knapik, E.W., Ziniti, J., Sim, C., Yamada, E., Kaplan, S., Jackson, D., de Sauvage, F., Jacob, H., Fishman, M.C., 1999. Zebrafish genetic map with 2000 microsatellite markers. *Genomics* 58, 219-232.

Solnica-Krezel, L., 2005. Conserved patterns of cell movements during vertebrate gastrulation. *Current biology* : CB 15, R213-228.

Stottmann, R.W., Bjork, B.C., Doyle, J.B., Beier, D.R., 2010. Identification of a Van der Woude syndrome mutation in the cleft palate 1 mutant mouse. *Genesis* 48, 303-308.

Strauss, R.P., 1999. The organization and delivery of craniofacial health services: the state of the art. *The Cleft palate-craniofacial journal : official publication of the American Cleft Palate-Craniofacial Association* 36, 189-195.

Streisinger, G., Walker, C., Dower, N., Knauber, D., Singer, F., 1981. Production of clones of homozygous diploid zebra fish (*Brachydanio rerio*). *Nature* 291, 293-296.

Su, J.L., Chiou, J., Tang, C.H., Zhao, M., Tsai, C.H., Chen, P.S., Chang, Y.W., Chien, M.H., Peng, C.Y., Hsiao, M., Kuo, M.L., Yen, M.L., 2010. CYR61 regulates BMP-2-dependent osteoblast differentiation through the $\alpha\beta3$ integrin/integrin-linked kinase/ERK pathway. *The Journal of biological chemistry* 285, 31325-31336.

Summerton, J., 1999. Morpholino antisense oligomers: the case for an RNase H-independent structural type. *Biochimica et biophysica acta* 1489, 141-158.

Tamura, T., Yanai, H., Savitsky, D., Taniguchi, T., 2008. The IRF family transcription factors in immunity and oncogenesis. *Annual review of immunology* 26, 535-584.

Taniguchi, T., Ogasawara, K., Takaoka, A., Tanaka, N., 2001. IRF family of transcription factors as regulators of host defense. *Annual review of immunology* 19, 623-655.

Thomason, H.A., Zhou, H., Kouwenhoven, E.N., Dotto, G.P., Restivo, G., Nguyen, B.C., Little, H., Dixon, M.J., van Bokhoven, H., Dixon, J., 2010. Cooperation between the transcription factors p63 and IRF6 is essential to prevent cleft palate in mice. *The Journal of clinical investigation* 120, 1561-1569.

Todorovic, V., Chen, C.C., Hay, N., Lau, L.F., 2005. The matrix protein CCN1 (CYR61) induces apoptosis in fibroblasts. *The Journal of cell biology* 171, 559-568.

Vanderas, A.P., 1987. Incidence of cleft lip, cleft palate, and cleft lip and palate among races: a review. *The Cleft palate journal* 24, 216-225.

Warga, R.M., Kimmel, C.B., 1990. Cell movements during epiboly and gastrulation in zebrafish. *Development* 108, 569-580.

Zucchero, T.M., Cooper, M.E., Maher, B.S., Daack-Hirsch, S., Nepomuceno, B., Ribeiro, L., Caprau, D., Christensen, K., Suzuki, Y., Machida, J., Natsume, N., Yoshiura, K., Vieira, A.R., Orioli, I.M., Castilla, E.E., Moreno, L., Arcos-Burgos, M., Lidral, A.C., Field, L.L., Liu, Y.E., Ray, A., Goldstein, T.H., Schultz, R.E., Shi, M., Johnson, M.K., Kondo, S., Schutte, B.C., Marazita, M.L., Murray, J.C., 2004. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. *The New England journal of medicine* 351, 769-780.