MOLECULAR GENETIC STUDIES OF SEVERAL CONNEXIN GENES IN THE ZEBRAFISH (*DANIO RERIO*)

by

Liang Tao

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Biological Sciences

The University of Manitoba

Winnipeg, Manitoba, Canada

Molecular Genetic Studies of Several Connexin Genes in the Zebrafish (*Danio rerio*)

By

Liang Tao

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

Doctor of Philosophy

ABSTRACT

This study explores the molecular genetic mechanisms that regulate the expression and function of five connexin (Cx) genes (Cx43, Cx44.1, Cx45.6, Cx48.5, and Cx30.3) in zebrafish (*Danio rerio*) embryonic development and adult tissues.

The transcriptional start sites and 5'UTRs (untranslated regions) of these genes were identified using 5' RACE (rapid amplification of cDNA ends). Interestingly, most of these Cx genes have alternative 5' UTRs, differing only in their first exons, which are generated from different promoter usages and alternative pre-mRNA splicing. Comparative analyses of zebrafish and mouse orthologs suggest that these features of Cx gene expression are evolutionarily conserved. Preliminary *in silico* and *in vivo* comparative analyses of promoter activities of zebrafish Cx44.1 and mouse Cx50 suggest that transcription regulatory networks of homologous Cx genes might be conserved between species. *In silico* analyses confirmed very strong loop-stem secondary structures and frequent uORFs in their 5'UTRs. Sequence analyses also suggest that IRESs (internal ribosome enter sites) might be involved in their translational regulation. Interestingly, tandem alternative splicing was identified within 5' UTRs of Cx45.6, and this is the first report of this phenomenon in Cx genes. Overall the diversity of 5'UTRs reflects a complexity of post-transcriptional regulation for Cx genes, especially considering the fact that alternative transcripts actually encode the same protein product. Functional analyses using "transcript variant-specific morpholino knock-down" approaches in GFP (green fluorescent protein)-labelled zebrafish embryos are proposed

to elucidate this question.

A new zebrafish Cx gene, Cx30.3, was identified within the Cx48.5 promoter. Sequence similarity and biophysical analyses suggest that zebrafish Cx30.3 is the ortholog of mammalian Cx26. Its transcriptional start site and full-length gene organization were identified using RACE technique. RT-PCR (reverse transcriptase polymerase chain reaction), real time PCR, and *in situ* hybridization techniques revealed that Cx30.3 is expressed in early embryos and in the skin, inner ear, lateral line, fin, heart, and retina of adults. This expression pattern may suggest the role of Cx30.3 in the development and function of hearing and balance in zebrafish. For future studies, zebrafish Cx30.3 could be employed to model human diseases of hearing and disorders in the skin.

ACKNOWLEDGEMENTS

Firstly, and foremost, I would like to thank my supervisor, Dr. Gunnar Valdimarsson. I owe a huge debt of gratitude to him for his constant guidance, encouragement, and support throughout the course of my studies in his lab. Without his guidance and support, I could not have completed my Ph.D. research and this thesis. It is my honour to have had the opportunity to work with him, and I am proud to have met so nice a mentor. I cannot thank him enough. I also gratefully thank Lorna Jakobson for her help in many ways.

I would also like to thank other professors on my advisory committee, including Dr. Steve Whyard, Dr. R. Daniel Gietz, Dr. Jiuyong Xie, and Dr. Erwin Huebner. Special thank you should go to Dr. Steve Whyard for his efforts in reviewing and editing my thesis manuscripts, as the acting advisor, before my submission of this thesis to the examining committee. I sincerely thank him for his invaluable comments, discussions, and assistance to my thesis preparation. I am also very thankful to him for his active help in quantitative PCR techniques, as well as fruitful discussions for experimental design and result processing of some of my experiments when I first started. I would also like to thank him for allowing me extensive use of his real time PCR instruments and other laboratory facilities. Dr. Whyard has served on variable committees linked to my graduate studies. I would like to express appreciation to him for his constant support and encouragement through the years. Dr. Erwin Huebner was on my advisory committee when I started as a M.Sc. student, as well as on various committees. I would like to thank

him for his continuous encouragement and support throughout my graduate studies. Special thanks to him should also be added for allowing me extensive use of his micro-injection equipment, microtome, and other laboratory facilities. I would also like to express my deep gratitude to Dr. R. Daniel Gietz, and Dr. Jiuyong Xie, for their insightful discussions and advice on my research, as well as all their efforts in guiding and supporting me throughout my studies. Also special thanks to them should be added for their efforts in evaluating my thesis. In addition, I would like to express great appreciation to Dr. Jeffrey J. Essner (Iowa State University, USA) for his willingness to be my external thesis examiner, and for his efforts in evaluating my thesis and for his insightful comments and advice. I would also like to express grateful appreciation to Dr. William Diehl-Jones for being willing to examine my thesis and for his thoughtful comments and advice.

I would also like to thank other professors, including Dr. Gary Anderson, Dr. Jennifer McLeese, Dr. Gail Davoren, and Dr. Brenda Hann, for being on my Ph.D. entrance or candidacy examination committees. I would also like to thank all professors who taught my courses. Besides my supervisor Dr. Valdimarsson, these professors include Dr. Ted Wiens, Dr. Steve Whyard, Dr. Ivan Oresnik (Dept. of Microbiology), Dr. Deborah Court (Dept. of Microbiology), Dr. Brian Fristensky (Dept. of Plant Science), Dr. R. D. Gietz (Dept. of Biochemistry & Medical Genetics), and Dr. Robert Shiu (Dept. of Physiology), and other professors in the Faculty of Medicine. I would also like to thank Mrs. Tara Narayansingh for her constant encouragement and support when I

worked as her TA, as well as throughout my stay in the department. I would like to thank John Brubacher for his help in sectioning and microscopy techniques. I would like to thank Mr. Terry Smith and all other staff in the department's animal care facility, for their excellent everyday care of our zebrafish. I would also like to thank all my lab-mates for their friendship, support and assistance, particularly David Shearer, Tarek Bader, Ryan Szajkowski, Aditi Singh, Scott Read, Shaohong Cheng, Tara Christie, etc. I would also like to extend a general thank you to the Department of Biological Sciences (and the former Zoology Department), the department head, Dr. Judy Anderson, and former head Dr. Erwin Huebner, all professors and staff, research associates, post-docs, and fellow graduate students at the department, for all their friendship, support and assistance during my stay in this department.

The research in this thesis was supported by grants to Dr. Gunnar Valdimarsson from the National Science and Engineering Research Council (NSERC) and the MHRC, and in addition, a student grant to myself from Sigma Xi (Grants-in-Aid of Research, $1,500 USD). I was financially supported by the Manitoba Health Research Council (MHRC) studentship, and different fellowships from the University of Manitoba.

Finally, I would like to express a very special thank you to my family, my parents and in-laws, for all their understanding, support and encouragement. I also thank my daughter, Hannah Tao, for her love. Most of all I would thank my wife, Fang Nan, for her constant love and support. I owe a huge debt of love and support to my wife and my daughter through the years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Gap Junction and Connexins

1.1.1 Gap junctional intercellular communication (GJIC)

As individual cells are separated by plasma membranes in multicellular organisms, cell-to-cell communication across plasma membranes is critically important to coordinate the behaviour of cells and tissues during embryonic development, as well as maintain homeostasis and normal function of tissue/organs in adult life. A variety of mechanisms have evolved to support such intercellular communication. These include plasmadesmata in plants and gap junctions in animals.

Gap junctions are highly specialized membrane structures found in both vertebrates and invertebrates from mesozoa to mammals. They are thusly named because they provide physical contact and direct passages between the cytoplasms of adjacent cells having a 2 – 4 nm extracellular gap between them (Robertson 1963; Revel and Karnovsky 1967; Bennett and Goodenough 1978). Gap junctional intercellular communication (GJIC) is mediated via cell-to-cell diffusion of small metabolic molecules, ions, and second messengers, allowing both electrical and chemical coupling between neighbouring cells. The gating properties and functions of gap junctional channels are strictly regulated (Goodenough et al. 1996; Kumar and Gilula 1996; Wei et al. 2004; Herve et al. 2007a). While the cellular functions of gap junctions are dependent

on the cell type, several broadly defined functions can be identified. A primary function of gap junctions is to provide a mechanism for both electrical and chemical coupling between neighbouring cells by mediating direct exchange of ions and small molecules, including but not limited to $Ca^{2+,}$ K^{+}, cAMP, cGMP, ATP, inositol triphosphate (IP3), oligonucleotides, siRNAs, glucose, and glutamate (Goodenough et al. 1996; Kumar and Gilula 1996; Wei et al. 2004; Herve et al. 2007a). It has been well demonstrated that GJICs play essential roles in many biological processes, including impulse conduction, cell proliferation, growth and differentiation, metabolic coordination (particularly in avascular organs such as epidermis and lens), and embryonic development (Bruzzone et al. 1996b; Lo 1996; Saez et al. 2003; Mese et al. 2007). Loss or mis-regulation of gap junicational communication can consequently lead to a variety of human diseases or defects in the nervous, cardiac, aural and visual systems (White and Paul 1999; Kelsell et al. 2001; Vinken et al. 2003; Gerido and White 2004; Trosko 2007; Dobrowolski and Willecke 2009). The tight docking of gap junction proteins between adjacent cells can also provide an adhesive force between cells and interact with the internal cytoskeleton to provide structural integrity (Lin et al. 2002; Elias et al. 2007). While gap junctions typically function by allowing direct exchange of small molecules and signals between two neighbouring cells, unpaired hemichannels can also participate in mediating electrical and chemical signalling exchange between the cytoplasm and the extracellular environment (Goodenough and Paul 2003). Junction-independent cell signalling functions have also been suggested via the cytoplasmic C- terminals of the connexin

proteins (Giepmans 2004).

1.1.2 Gap junction proteins and basic structural organization of gap junctions

Gap junctional channels are formed by end-to-end apposition via non-covalent interactions of two hemi-channels, each of which is composed of six protein subunits embedded in the plasma membranes of two neighbouring cells. Apart from the gap junctional function, unpaired hemi-channels can function in single-cell membranes similar to regular ion channels (Goodenough and Paul 2003; Bennett et al. 2003; Shestopalov and Panchin 2008).

Two distinct protein families have evolved to fulfil this fundamental and highly conserved function of intercellular communication. In vertebrates, gap junction channels and hemi-channels are made of proteins encoded by a family of genes called "connexins" (Cxs) (Goodenough and Paul 2003; Goodenough 1974; Bruzzone et al. 1996a), whereas in invertebrates, gap junction proteins are encoded by another family of genes called "innexins" (Inxs) (Phelan et al. 1998; Phelan 2005). Historically, connexins were considered to be the only structural proteins of gap junction channels until the innexins were discovered in 1998 (Phelan et al. 1998). Connexins are not homologous to innexins according to their primary sequences, but interestingly, gap junctions formed from connexins or innexins share similarities in structure and function. More recently, another distinct group of proteins called "pannexins" (Panxs), which are homologous to innexins, has been identified in vertebrates, including mammals (Baranova et al. 2004; Herve et al.

2005; Panchin 2005; Barbe et al. 2006; Yen and Saier 2007). Based on analyses of their sequence and function similarities, innexins and pannexins belong to a single superfamily and therefore might have evolved from a common ancestor (Yen and Saier 2007). The lack of these proteins in some metazoans suggests that other still-undiscovered gap junction proteins could exist (Shestopalov and Panchin 2008).

All identified gap junction proteins share a common topology, with four α-helical transmembrane domains, two extracellular loops, an intracellular loop, and N- and C-termini located on the cytoplasmic side. Connexins have three highly conserved cysteines in each extracellular loop (thus six conserved cysteines in total) that form intramolecular disulfide bonds, and consequently play essential roles in protein folding and intercellular docking of opposing hemi-channels (Unger et al. 1999; Yeager et al. 1998; Rahman and Evans 1991; John and Revel 1991; Dahl et al. 1992). The only known exception is Cx23, which has only four cysteines (Iovine et al. 2008). Similarly, innexins and pannexins have two conserved cysteines in each extracelluar loop that serve similar functions to that of connexins (Baranova et al. 2004; Herve et al. 2005; Yen and Saier 2007; Phelan and Starich 2001). In addition, there is a highly conserved proline motif at the extracellular end of the second transmembrane domain in all gap junction proteins, which is also considered to be involved in the formation of gap junction channels. The capacity of two apposing hemi-channels (each from the neighbouring cells respectively) to couple together and form intercellular channels has been considered the fundamental characteristic of all gap junction proteins.

Because innexins and pannexins are relatively newly-discovered gap junction proteins, the structure and functions of gap junctions have primarily focussed on connexin-based channels. The present thesis research will focus on the regulation of expression and function of connexins using zebrafish as the model organism.

1.1.3 Regulation of connexin-based channels

The half channel "connexon" is assembled from six protein subunits, called "connexins", which form a cylinder structure with a pore in the centre. The four transmembrane domains in each connexin molecule contribute to the wall of the pore. The two extracellular loops play roles in inter-connexon recognition and the formation of gap junction channels. The transmembrane and the extracellular regions are highly conserved among different members of the connexin family, which to a certain degree allows different types of connexins to be able to assemble together. Connexons assembled by the same type of connexins are termed homomeric connexons, whereas those containing multiple types of connexins are termed heteromeric connexons. Channels formed by the docking of two identical connexons are termed homotypic channels, whereas channels formed by the docking of two different connexons are heterotypic channels. It seems that there are four possible assemblies of gap junction channels, among which only homomeric homotypic channels are made of identical connexins. However, not all types of connexins have the same capability to form heteromeric connexons or heterotypic gap junction channels – there is a highly regulated

code of compatibility among connexins. For example, Cx31 can only form homotypic channels, whereas Cx46 can interact with other connexins (Elfgang et al. 1995; White et al. 1995). Cx26 can form heteromeric channels with Cx30 and Cx32, whereas it cannot form functional channels with Cx40 (Bukauskas et al. 1995; Segretain and Falk 2004). Cx43 can form functional heteromeric channels with Cx46 but not with Cx50 (though all three Cxs are co-expressed in the lens). When the E2 domains of Cx46 and Cx50 were exchanged, the hybrid Cx50 was capable of forming heterotypic channels with Cx43, which suggested that the second extracellular loops (E2 domains) are important for the compatibility between connexins (White et al. 1995). Distinct electrophysiological and permeability-selective properties have been shown not only for different homotypic gap junction channels, but also between homotypic and heterotypic channels. The complex of interactions between different connexins expands the diversity of both channel types and channel properties, thus providing a mechanism for the regulation of channel activity and function (White and Bruzzone 1996).

The intracellular regions, including the N- and C-terminal segments and the intracellular loop, are important for the regulation of gating properties and channel activities. It seems that the N-terminal is often highly conserved among different connexins (Bruzzone et al. 1996b). For example, the zebrafish Cx43, Cx44.1 and Cx48.5, which are all expressed in the fish lens have only a few nucleotides of difference immediately downstream of the translational initiation codon (Cheng et al. 2003). In contrast, the cytoplasmic loop and the C-terminal of connexins vary significantly in

sequence and length. For example, the C-terminus of Cx26 has only a few residues, whereas Cx50 has a long C-terminal tail of about 200 residues (Zhang and Nicholson 1989; White et al. 1992).

It has been established that the gating mechanisms of gap junction channels include changes of cytoplasm membrane potential, changes of cytoplasmic pH and Ca^{2+} concentration, and phosphorylation of connexin peptides. The cytoplasmic regions of connexins are targets of a variety of intracellular signalling molecules, and potential sites for cytoplasmic protein-protein interactions, as well as post-translational modifications such as phosphorylation, thus offering complicated mechanisms for the regulation of channel property and function (Herve et al. 2007a; Giepmans 2004).

Initially, it was thought that the exchange of any ions or small metabolites (less than 1.2 kDa) through gap junction channels was completely free, passive, and non-specific. However, later evidence has demonstrated that the permeability of gap junction channels are in fact connexin-dependent and highly regulated. Different connexins possess unique properties with regard to unitary conductance, ionic and chemical selectivity, size cut-off, and gating activity, all of which can provide different signalling pathways in different cell types, dependent on the type of connexins present (Bruzzone et al. 1996a). Accordingly, studies on connexin-related diseases have also emphasized the non-redundance or differential functions of different connexins, where the loss of one connexin cannot be compensated by the presence of other connexins in the same tissue (or cell type) (White and Paul 1999; Kelsell et al. 2001; White 2003;

White 2002).

The biosynthesis, assembly, and trafficking of connexins is considered to follow the general secretory pathway for general membrane proteins, although most of the studies refer to Cx43 and there is no evidence that all connexins share the same pathways (Saez et al. 2003; Laird 2006; Salameh 2006). Briefly, connexin peptides are co-translationally integrated into the endoplamic reticulum (ER) membrane, followed by the oligomerization of six connexins into a connexon in the *trans*-Golgi network. Connexons are then delivered to the cell membrane within secretory vesicles, transported along microtubules. The connexons enter into the cell membrane without preselected localization, and they can move laterally in the plane of the plasma membrane. When reaching the outer rim of a gap junction cluster (or regions of cell-to-cell contact), connexons coming from two neighboring cells can interact with each other via their extracellular loops, thereby forming a new gap junction channel. Intercellular channels are then aggregated into gap junction plaques, each of which may contain from less than a dozen to many thousand individual channels.

In contrast to typical membrane proteins that form channels mediated by gating, the connexins have a surprisingly short half-life, ranging from 1 to 5 hours, which in fact makes gap junction plaques highly dynamic. With new channels being added to the periphery, old channels are removed from the center by a process termed "internalization", followed by degradation of connexins via either lysosomal or proteosomal pathways. It is thought that the rapid turnover of gap junction channels

provides a regulatory mechanism for the cells to rapidly adapt changes of physiological conditions or respond to environmental stimuli (Salameh 2006; Herve et al. 2007b).

1.1.4 Connexin gene family

Since the cloning of the first connexin gene from rats in 1986 (Paul 1986; Heynkes et al. 1986; Kumar and Gilula 1986), a large multi-gene family has been revealed to date, including 21 human connexin genes and 20 mouse genes (Sohl and Willecke 2003). One conventional nomenclature for connexin genes is based on the molecular mass of their corresponding proteins, as for example, Cx43 encodes a connexin (Cx) protein of about 43 kDa in size. This nomenclature frequently causes confusion, especially to those who are new or not in this research field, because connexin orthologues from different organisms might have different molecular weights, and connexins that have the same molecular weight are not necessarily orthologues in different species. Therefore, a suitable prefix is usually used to distinguish connexins from different organisms. For example, hCx43, mCx43, rCx43, and ZfCx43 represent human, mouse, rat, and zebrafish orthologues of connexin43 genes, respectively; ZfCx44.1 is the zebrafish ortholog of mCx50 (mouse connexin50); ZfCx48.5 is the zebrafish ortholog of mCx46; and ZfCx30.3 is the ortholog of mCx26. Connexins have been divided into three subgroups known as α, β, and γ according to the degree of peptide sequence homology and the length of their cytoplasmic domains (O'Brien et al. 1998; Sohl et al. 2001; Harris 2001; Willecke et al. 2002). Another conventional

nomenclature has been used, which takes such evolutionary relationships into consideration. For example, GJA8 represents Gap Junction Alpha 8, (otherwise known as Cx50), and belongs to the α subgroup; whereas GJB2 stands for Gap Junction Beta 2, (otherwise known as Cx26), and is a member in the β subfamily. Clearly, a suitable and consistent nomenclature is required, and this issue is still under in-depth discussion at international gap junction conferences.

Each connexin gene shows unique tissue- or cell type- specific distributions of expression, while most tissues and many cell types co-express multiple connexin species. Some connexins, such as Cx43 for example, are ubiquitously expressed in almost all tissues or cell types, whereas the expression of most connexins is restricted in specific tissues or cell types. For example, mouse Cx46 and Cx50 are lens-specific connexins. Even in the same tissue, different connexins may show distinct cell type-specific or development stage-specific expression patterns. The distinct but overlapping expression patterns indicated the presence of complicated and tight regulatory mechanisms for the expression and function of connexin genes (Saez et al. 2003; Oyamada et al. 2005). However, our knowledge of the gene structure and the regulation of connexin expression is still very limited. In fact, there have been relatively few studies in the literature, in sharp contrast with thousands of articles on expression patterns and biological function of connexin proteins; and studies on regulation of connexin gene expression have focused mainly on the transcriptional level (Table 1-1).

Table 1-1. Studies on Gene Structure and Transcriptional Regulation of Cx Genes

Cxs	Gene Structure	Promoters & TTSs	TF sites or TFs, Other regulators	References
hCx26		P1	TA-box, Sp1/Sp3, DNA methylation	1, 2, 3, 4, 5
mCx26		P1	TA-box, Sp1/Sp3	6
hCx30	3 exons Cx	P1, P2	TA-box, Sp1, Egr	7
mCx31	Alternative exon-1	P1, P2	Egr2/Krox20	8, 9, 10
hCx32	Alternative exon-1	P1, P2	Sox10, Egr2/Krox20, DNA methylation	11, 12, 13, 14, 15
mCx32	Alternative exon-1	P1, P2, P3	Sp1, NF-1, HNF-1	6, 16, 17
rCx32	Alternative exon-1	P1, P2	Sp1/Sp3, YinYang-1, HNF-1, Mist1, DNA methylation	17, 18, 19, 20, 21, 22
mCx36	Spliced ORF			23
rCx36	Spliced ORF			23, 24
mCx39	Spliced ORF			25
mCx57	Spliced ORF			26
hCx40	Alternative exon-1	P1, P2	Sp1, Tbx3, Tbx5	27, 28, 29
mCx40	Alternative exon-1, 3 exons	P1, P2	Sp1,GATA4, Nkx2-5, Tbx5, Tbx2,	10, 30, 31, 32, 33, 34, 35
rCx40		P1	Sp1/Sp3, Nkx2-5, GATA4, Tbx5	32, 36
hCx43		P1	Sp1, AP-1, Tbx3, retinoids, estrogen, Ras-Raf-MAPK signalling, Wnt signalling, DNA methylation	29, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49
mCx43	Alternative exon-1	P1, P2, P3 (multi-TTSs)	Nkx2-5,Tbx2, Wnt signalling	30, 34, 35, 50, 51, 52, 53
rCx43	Alternative exon-1	P1, P2, P3 (multi-TTSs)	Sp1, AP-1, Nkx2-5, C-Fos, cAMP, estrogen, PGE2, TH/PTH, CNTF, Wnt signalling, DNA methylation	19, 40, 52, 54, 55, 56, 57, 58, 59, 60, 61
mCx45	3 exons	P1, P2, P3	Nkx2-5,	33, 62, 63
mCx46	Alternative exon-1, trans-splicing/mCx50	P1, P2		64
mCx47	Alternative exon-1	P1, 2, 3, 4		64

Table 1-1 References: [1] Tu and Kiang,1998. [2] Tu et al. 1998. [3] Tan et al. 2002. [4] Singal et al. 2000. [5] Loncarek et al. 2003. [6] Hennemann et al. 1992. [7] Essenfelder et al. 2005. [8] Hennemann et al. 1992. [9] Jungbluth et al. 2002. [10] Anderson et al. 2005. [11] Neuhaus et al. 1996. [12] Bondurand et al. 2001. [13] Hirai et al. 2003. [14] Yano et al. 2004. [15] Houlden et al. 2004. [16] Söhl et al. 1996. [17] Söhl et al. 2001. [18] Duga et al. 1999. [19] Piechocki et al. 1999. [20] Piechocki et al. 2000. [21] Field et al. 2003. [22] Koffler et al. 2002. [23] Söhl et al. 1998. [24] Condorelli et al. 1998. [25] von Maltzahn et al. 2004. [26] Hombach et al. 2004. [27] Dupays et al. 2003. [28] Fijnvandraat et al. 2003. [29] Hoogaars et al. 2004. [30] Kasahara et al. 2001. [31] Bruneau et al. 2001. [32] Linhares et al. 2004. [33] Jay et al. 2004. [34] Christoffels et al. 2004. [35] Dupays et al. 2005. [36] Bierhuizen et al. 2000. [37] Rogers et al. 1990. [38] Fishman et al. 1991. [39] Petrocelli et al. 1993. [40] Piersanti et al. 1995. [41] Geimonen et al. 1996. [42] Clairmont et al. 1996. [43] Clairmont et al. 1997. [44] Geimonen et al. 1998. [45] Carystinos et al. 2003. [46] Chen et al. 2003. [47] Oltra et al. 2003. [48] Vine et al. 2005. [49] Mitchell et al. 2005. [50] Chen et al. 1995. [51] Kasahara et al. 2003. [52] Pfeifer et al. 2004. [53] Chen et al. 2004. [54] Yu et al. 1994. [55] Civitelli et al. 1997. [56] Ai et al. 2000. [57] Stock et al. 2000. [58] Mitchell et al. 2001. [59] Fernandez-Cobo et al. 2001. [60] Teunissen et al. 2003. [61] Ozog et al. 2004. [62] Jacob and Beyer 2001. [63] Dupays et al. 2005. [64] Pfeifer et al. 2004.

1.2 Transcriptional Regulation of Connexin Genes

In general, gene expression is achieved by a multi-step program that is tightly controlled by multi-layered regulatory mechanisms. It starts in specific areas of the nucleus, where transcription factors (*trans*-acting proteins) bind to specific DNA sequences (*cis*-regulatory elements) in the proximity of the target genes activating or repressing the transcription via specific interactions with RNA polymerase. Newly formed pre-mRNAs are bound by a number of mRNA-binding proteins (mRBPs), and are modified by diverse RNA processing reactions including 5' capping, splicing, editing, 3' cleavage and polyadenylation. Mature, fully-processed mRNA transcripts are then exported out of the nucleus through nuclear pores; and in the cytoplasm, they are localized to specific subcellular regions by motor proteins or by the signal recognition particle, where they recruit translation factors and ribosomes for protein synthesis. The activity of proteins can be further regulated by folding, post-translational modification (such as phosphorylation), subcellular localization, and interactions with other proteins.

Similar to other genes, expression of connexin genes is highly regulated by multi-layered mechanisms including transcriptional control, pre-mRNA processing control, mRNA transportation and localization control, mRNA stability control, translational control, post-translational modification and localization control. Transcription, as the necessary first step of gene expression, is believed to be central to gene regulation, where promoters and transcription factors are key components regulating the initiation of transcription. Transcription factors work in groups, rather than

individually. There has been growing evidence that transcription factor binding sites in a promoter are organized into clusters called "*cis*-regulatory modules (CRMs)", which contain dozens of binding sites usually for ~3 – 10 transcription factors. CRMs can span a few hundred nucleotides in the sequence around the transcriptional start sites, which "read out" the concentrations of multiple transcription factors that are present in a particular cell context, thus determining the initial spatio-temporal expression patterns of the target gene (Davidson 2001; Davidson 2006; Istrail and Davidson 2005). Functional CRMs can be detected either experimentally or by *in silico* comparative promoter analysis (Klingenhoff et al. 1999; Fessele et al. 2002).

DNA sequence around the transcriptional start site is described as the basal (proximal) promoter of a gene, which is usually a few hundred base pairs (bp) in length and can often include the first non-coding exon (Werner 1999). For connexin genes, it was believed that the basal promoter P1 is located within the first 300 bp upstream of the transcriptional start site (Oyamada et al. 2005). Within this region, a number of transcription factor binding sites have been identified in some connexin genes, among which mammalian Cx43, Cx40, and Cx32 have been the most extensively studied connexin genes thus far (Table 1-1). Ubiquitous (or cell type-independent) transcription factors include the TATA-box-binding protein (TBP), a GC-box-binding protein such as Sp1/Sp3, and an activator protein 1 (AP-1). These general transcription factors are believed to facilitate the RNA polymerase II maintaining the basal level of transcription of connexin genes; however, it is unlikely that all these factors are necessary for

transcription of connexin genes. For example, some connexin genes might contain "TATA-less" promoters. Genes with this type of promoter are often transcribed from multiple transcription start sites spanning within decades of base pairs (Deng et al. 1986; Weis and Reinberg 1992). In our studies on zebrafish Cx43, multiple transcriptional start sites were shown to span approximately 50 bases, and no TATA-like elements were found upstream of this transcriptional initiation region (Chatterjee et al. 2005).

Special spatio-temporal expression of connexin genes is dependent on the presence of cell type-specific transcription factors. Studies on the molecular mechanisms for tissue-specific expression of connexin genes have focused mainly on mammalian Cx43, Cx40, Cx45, and Cx32 (also see Table 1-1). The cardiac conduction system specifically expresses Cx43, Cx40, and Cx45, with distinct cell type-specific and developmental stage-specific expression patterns (Gros et al. 2004). Cardiac-specific transcription factors, such as Nkx2-5, Tbx5, and GATA4, play crucial roles in the transcriptional regulation of these cardiac connexin genes, which was also confirmed by Nkx2-5 and Tbx5 knockout mice (Bruneau et al. 2001; Kasahara et al. 2001; Fijnvandraat et al. 2003; Kasahara et al. 2003; Jay et al. 2004; Linhares et al. 2004; Dupays et al. 2005). Another well studied connexin is mammalian Cx32, which is expressed mainly in hepatocytes, secretory acinar cells, and Schwann cells. Transcription factors such as HNF-1, Mist1, Sox10 and Egr2/Knox20 were demonstrated to control the transcription of Cx32 (Piechocki et al. 2000; Hennemann et al. 1992a; Koffler et al. 2002; Bondurand et al. 2001; Rukstalis et al. 2003).

In addition to the above transcription factors, cell signalling molecules known to affect transcription such as cAMP (Yu et al. 1994; Civitelli et al. 1997; Van der Heyden et al. 1998), retinoids (Rogers et al. 1990; Clairmont and Sies 1997; Clairmont et al. 1996; Vine et al. 2005), estrogen (Yu et al. 1994; Geimonen et al. 1996; Geimonen et al. 1998; Piersanti and Lye 1995; Petrocelli and Lye 1993; Oltra et al. 2003), and thyroid/parathyroid hormones (Stock and Sies 2000; Mitchell et al. 2001) play roles in the transcription of the Cx43 gene. Furthermore, several studies have suggested that the Wnt pathway (Van der Heyden et al. 1998; Geimonen et al. 1996; Chen et al. 1995; Ai et al. 2000) and Ras-Raf-MAPK pathway (Carystinos et al. 2003) are involved in the regulation of expression of Cx43. Studies on Cx26 (Tan et al. 2002; Singal et al. 2000; Loncarek et al. 2003), Cx32 (Piechocki et al. 1999; Hirai et al. 2003; Yano et al. 2004), and Cx43 (Piechocki et al. 1999; Chen et al. 2003) also suggested that DNA methylation of the CpG islands plays important roles for the transcription of these Cx genes.

While transcription is considered integral to the regulation of gene expression, it does not always directly correlate with the final concentration of the gene products that govern the phenotype. In recent years, gene regulation at the post-transcriptional level (for example, pre-mRNA splicing, microRNA, etc.) has been attracting more interest, as it can provide mechanisms that finely modulate the dynamic spatio-temporal expression of genes during embryonic development and a variety of cellular processes.

Recent studies on connexin gene structure have demonstrated that some connexins commonly contain variant 5' UTRs (alternative first exons) generated from

multiple promoter usage or pre-mRNA alternative splicing, suggesting a complex mechanism for regulation of connexin genes by selectively using different non-coding leading sequences within the transcripts. The next section will provide additional details about mRNA splicing and our current knowledge of structure and post-transcriptional regulation of connexin genes.

1.3 Pre-mRNA Splicing and Connexin Gene Structure

In general, pre-mRNA splicing is a RNA-catalyzed molecular process that occurs in the cell nucleus, and results in the formation of mature mRNA transcripts by removing the intron(s) from the primary transcripts and joining the exons together (Sharp 2005; Sharp 1994; Sharp 1985).

Splicing is guided by special sequence elements at the intron-exon junctions called splice sites. The 5' splice site, located at the 5' end of the intron, is a loosely-conserved consensus sequence, which includes a GU dinucleotide at the extreme 5' end of the intron; the 3' splice site has three conserved elements – a conserved adenosine nucleotide within the branch point consensus sequence, followed by a polypyrimidine tract, followed by a terminal AG dinucleotide at the extreme 3' end of the intron. Splicing is achieved through two trans-esterification reactions, which are catalyzed by a large complex called the spliceosome, resulting in the release of the intron as a "lariat" structure and the joining of two exons. Pre-mRNA splicing occurs in the cell nucleus and is closely coupled to transcription (Black 2003).

The spliceosome includes five highly conserved small nuclear ribonucleoprotein particles (snRNPs) and a large number of nuclear proteins bound to the pre-mRNA. It is the stepwise assembly of the spliceosome that brings the splice sites close, and it is the catalytic components (RNA molecules) in these snRNPs that catalyze the splicing reactions. The molecular events of pre-mRNA splicing have been well elucidated, where the stepwise assembling and rearrangement of the spliceosome are central. The trans-esterification reactions do not need an input of energy, but ATP hydrolysis is required for the activities of the spliceosome during splicing (Burge et al 1999).

For some multi-exon transcripts, exons might be selectively joined, thus resulting in different mRNA products. This process is termed "alternative splicing" of the pre-mRNA. Alternative splicing can be classified into five basic patterns: (i) a cassette exon can be either included in the mRNA or excluded; (ii) mutually exclusive exons; (iii) alternative 5' and 3' splice sites; (iv) alternative promoters and alternative poly-A sites switch the 5'- or 3'- exons; (v) a retained intron can be either removed or retained in the mature mRNA. In general, these are often used in a combinational manner to generate multiple patterns of inclusion of exons (Black 2003).

The determination of splice site selection depends on the assembly of the spliceosome (Matlin and Moore 2007). The splice site consensus sequence elements are necessary but not sufficient to determine a functional splice site (Black 2003). Like other regulatory mechanisms for gene expression (e.g., transcription), the regulation of alternative splicing involves variable *cis*-acting elements (regulatory sequences) and

trans-acting factors (RNA-binding proteins or RNAs). These auxiliary *cis*-acting elements are known as exonic and intronic "splicing enhancers" (ESEs, ISSs respectively) and "splicing silencers" (ESSs and ISSs respectively). Enhancers activate adjacent splice sites or counteract splicing silencers, whereas silencers repress splice sites or antagonize splicing enhancers (Blencowe 2000; Pozzoli and Sironi 2005). Some identified *trans*-acting regulatory factors include members of the SR protein family, and a large number of hnRNPs (heterogeneous nuclear ribonucleoproteins). SR proteins usually bind to ESEs and act as splicing activators, participating in the regulation and also the assembly of the spliceosome. Structurally, a SR protein molecule has one or two N-terminal RRM (RNA recognition motif) domains and C-terminal RS domains that are enriched in arginine and serine residues. Therefore, SR proteins can mediate both protein-protein and protein-RNA interactions, playing an essential role in spliceosome assembly (Valcarcel and Green 1996; Graveley 2000). Knockout mouse mutations demonstrated that different SR proteins show distinct alternative splicing activity and their roles are not overlapping (Xu et al. 2005). Other *trans*-acting factors, such as hnRNPs, usually bind to splicing silencers (ESSs and ISSs), and function as splicing repressors. Some well-studied repressors include hnRNPA1, and polypyrimidine tract binding proteins (PTB/hnRNPI), but there are numerous others (Krecic and Swanson 1999; Dreyfuss et al. 2002).

The regulation of alternative splicing is specific to cell types, developmental stages or signalling pathways (Black 2003; Xie 2008; Xie and Black 2001). The

regulation mechanism involves multi-dimensional interactions between variable positive and negative regulatory factors (*cis*-acting elements and *trans*-acting proteins or hnRNPs) in the assembly of the spliceosome. The splicing decisions are made depending on the precise combinations and relative ratios of these positive and negative factors. For example, the relative enrichment of ESEs or ESSs in the exon, and the relative concentrations of the RNA-binding activator or repressor proteins may determine the frequency with which the exon is included or skipped (Black 2003; Caceres and Kornblihtt 2002; Matlin et al. 2005; Soller 2006).

Other *cis*-acting influences include the secondary structure of the RNA, since secondary structure can either sequester *cis*-regulatory elements or bring them into closer juxtaposition (Buratti and Baralle 2004). In addition, as pre-mRNA splicing is highly coupled to transcription, it is believed that the transcriptional machinery also influences the decision of alternative splicing choice. During transcription, the emerging pre-mRNA is immediately coated with various RNA processing and general RNA-binding factors. It has been established that different portions of the C-terminal of RNA polymerase II large subunit recruit different RNA processing factors, including splicing factors, to the emerging (or nascent) mRNA transcript. Further, the elongation rate of transcription sometimes is also a significant factor affecting exon selection. In these cases, competing (strong or weak) splice sites are differentially selected by speeding up or slowing down the elongation of transcription (Caceres and Kornblihtt 2002; Kornblihtt 2006; Kornblihtt 2005; Maniatis and Reed 2002).

Alternative splicing provides the eukaryotic cells with an elaborate mechanism for the control of gene expression. Most alternative splicing events affect the coding region of a gene, thus different protein-coding elements are selectively joined to form different mRNA transcripts, leading to production of variable proteins with distinct structures, properties and functions. Therefore, alternative splicing is an important source of protein diversity in eukaryotic cells (Mendes Soares and Valcarcel 2006; Stetefeld and Ruegg 2005; Maniatis and Tasic 2002; Graveley 2001). Some alternative splicing (~25%) only occurs in the untranslated regions (UTRs) of a gene, such as has been observed in some connexin genes. Although such alternative splicing events do not result in different proteins from the same gene, UTRs frequently influence mRNA export, cytoplasmic localization, stability, and translational efficiency, and thus they might provide a fine modulation for regulation of gene expression in different cell types, at different developmental stages, or in response to different stimuli (Kornblihtt 2005; Hughes 2006).

Evolutionarily more complex organisms are thought to require both more diversity of proteomes and more complexity of regulatory networks, and alternative splicing could play an important role in evolution by expanding the coding capacity of a genome, as well as providing elegant mechanisms for the regulation of gene expression. In fact, alternative splicing is considered very prevalent in multicellular organisms – it is "the rule" rather than "the exception" (Zavolan and van Nimwegen 2006). It is estimated that more than 70% of human genes are alternatively spliced, and at least 15% (perhaps as many as 50%) of human genetic diseases are caused by defects in alternative splicing

(e.g., mutations either in splice site sequences or variable *cis*-regulatory elements (Modrek et al. 2001; Grasso et al. 2004). Adding to this complexity, it is now apparent that the introns are not all "junk DNA", as they were once considered, but rather they frequently contain variable *cis*-regulatory elements for the control of gene expression (Andolfatto 2005; Mattick 2004). Evolutionarily complex organisms may contain numerous introns in their genome that are amenable to the genetic recombination of exons from different genes, thus allowing new genes to be evolved more easily by using pre-existing DNA elements (Lareau et al. 2004).

Early studies suggested that connexin genes share a relatively simple structure consisting of two exons, with their 5' UTRs interrupted by an intron of variable length. That is, the first exon contains most of the 5' untranslated region (UTR), while the second exon contains the remaining portion of the 5' UTR (usually a few nucleotides), the uninterrupted coding region, and the 3' UTR. However, in recent years, various exceptions to this rule have been documented (summarized in Table 1-1) (Sohl and Willecke 2003; Willecke et al. 2002; Oyamada et al. 2005). Some connexin genes have been found with their coding regions interrupted by an intron. For example, coding regions of murine Cx36 and Cx39 are interrupted by an intron in their N-terminus (Condorelli et al. 1998; Söhl et al. 1998; von Maltzahn et al. 2004), while mouse Cx57 coding sequence is interrupted in its C-terminus (Hombach et al. 2004). A few exceptional connexin genes have been found to contain three exons, with their 5'UTRs interrupted by two introns, such as human Cx30 (Essenfelder et al. 2005), mouse Cx40

(Anderson et al. 2005) and Cx45 (Jacob and Beyer 2001), and zebrafish Cx48.5 (Tao and Valdimarsson, unpublished). Several connexin gene mRNAs have been observed to have alternative first exons in their 5' UTRs, arising from different promoter usage and pre-mRNA splicing in an alternative manner. These include human Cx30 gene (Essenfelder et al. 2005), mammalian Cx31 (Anderson et al. 2005; Gabriel et al. 2001; Hennemann et al. 1992b), Cx32 (Neuhaus et al. 1995; Sohl et al. 1996; Neuhaus et al. 1996; Duga et al. 1999; Söhl et al. 2001), Cx40 (Anderson et al. 2005; Dupays et al. 2003), mouse Cx43 (Pfeifer et al. 2004), and mouse Cx45 (Anderson et al. 2005). However, it should be noted that in all these cases, the transcript isoforms vary only in the first exons within their 5' UTRs, i.e., leaving their coding regions unaltered (Sohl and Willecke 2003). Therefore, even with alternative 5'UTRs, different transcript isoforms actually encode the same connexin proteins.

Several studies on connexin gene structure have demonstrated that alternative first exons are spliced in a tissue or cell-type specific manner, suggesting that alternative splicing of connexin genes may be finely regulated in a tissue or cell-type specific manner (Anderson et al. 2005; Sohl et al. 1996; Neuhaus et al. 1995; Neuhaus et al. 1996). In a transgenic mice study, it was found that Cx32 mRNA is alternatively spliced in different tissues; in nervous tissues, it is processed to remove a small 345-bp intron, while in the liver, it is processed to remove a larger intron of 6.1 kb. Sohl G. et al. (1996) reported another alternative isoform of Cx32 mRNA is expressed in mouse embryonic cells, oocytes, and liver. The human Cx32 gene consists of three exons that are

alternatively spliced to produce mRNAs with different 5' UTRs. Similar to rodent Cx32, transcription is initiated from two tissue-specific promoters. In the liver and pancreas, promoter P1 is used, which is located more than 8 kb upstream of the translation start site and the transcript is processed to remove a large intron. In contrast, in nerve cells, transcription is initiated from promoter P2, which is located 497 bp upstream from the translation start codon and thus, the transcript is processed to remove a small 355-bp intron (Neuhaus et al. 1996). Some bovine Cx32 transcripts have three exons, including a second exon (exon 1a) also in the 5' UTR, which is alternatively spliced. Thus bovine Cx32 encodes three different spliced transcripts: exon1- exon2; exon1- exon 1a –exon2; exon 1b- exon2 (Duga et al. 1999). Another example is the Cx45 gene, which contains three exons separated by two introns. Exon 1 and exon 2 contain only 5' UTR sequences and the third exon contains the remaining 5' UTR, the full coding region and 3' UTR. Cx45 transcripts also contain alternative first exons processed tissue-specifically, while both exons 2 and 3 are shared in all Cx45 transcripts (Jacob and Beyer 2001). More recently, Anderson and colleagues reported two novel first exons in mouse tissues, each of which is spliced to exon 2 and 3 in a tissue-specific manner (Anderson et al. 2005).

In most cases, connexin genes contain a single intron of variable length. As mentioned above, introns are not simply "junk" DNA sequences, but rather, they often contain *cis*-regulatory elements for pre-mRNA splicing (Black 2003). Introns also contain regulatory elements for transcription and several studies have shown the importance of intron sequences for efficient transcription rates (Brinster et al. 1988;

Whitelaw et al. 1991; Warnecke et al. 1999). In the rat Cx31 gene, the intron sequence was found indispensable for basic promoter function (Gabriel et al. 2001). The rat Cx31 contains two exons separated by an intron of 2.6 kb. There is a "TATA-less" promoter located upstream of the first exon. Transcription is initiated from two alternate sites, resulting in the first exon consisting of 105 and 139 nucleotides respectively. The upstream 935-bp promoter region was predicted to contain five putative binding sites for the GATA transcription factors as well as a NF-kappa B binding site, a CAAT-box and E-box/E-box-related sequences. In transient transfection experiments, only constructs including exon-1 and the entire intron showed high promoter activity, whereas deletion of 1.1 kb of the intron from its 5' end sequence caused the loss of promoter function. Experiments also demonstrated that the intron sequence can increase the stability of the intron sequence. Studies on rat Cx31 intron suggest that the presence of the intron sequence provides a mechanism expanding the complexity of the regulation of connexin gene expression.

1.4 Zebrafish as a Model Organism

The zebrafish (*Danio rerio*) is a small tropical freshwater fish that originated from rivers throughout south-east Asia in India, Northern Pakistan, Nepal and Bhutan. It was established as a model organism for embryonic developmental studies by George Streisinger and co-workers in the late 1970s and early 1980s (Streisinger et al. 1981). As a vertebrate organism, the zebrafish has many organs and cell types similar to that of

mammals, although the fine details of anatomy and physiology are different in many aspects. Embryogenesis in zebrafish is also similar to that of higher ordered vertebrates, making zebrafish a useful model for vertebrate development. It is widely accepted that the molecular mechanisms underlying embryonic developmental and adult tissue/organ function are highly conserved in vertebrates, including human and fish. The zebrafish has become an increasingly favored model organism to investigate mechanisms underlying embryonic development, adult tissue function, and disease (Nüsslein-Volhard and Gilmour 2002).

The zebrafish genome has 25 chromosome pairs, and it has a relatively small haploid size, with 1.7×10^9 base pairs, which is about half the size of the human genome (Hinegardner and Rosen 1972). It was believed that almost every mammalian gene has an ortholog in the zebrafish genome, and most of the genes in zebrafish have the same function as in mammals (Postlethwait et al. 1998). In fact, the zebrafish has been described as "the canonical vertebrate" due to such genetic similarities (Fishman 2001). The genetic similarities between zebrafish and mammals allow mechanisms for embryonic development and the functions of genes in adult tissues/organs to be studied in a fish model, which as will be outlined below, can be more tractable than mammalian models in certain applications. In addition, zebrafish, like many fish species, have portions of their genome duplicated (Postlethwait et al. 1999; Postlethwait et al. 2000), and it is estimated that about 20% of mammalian genes have two zebrafish orthologues, many with distinct functions and expression domains (Postlethwait et al. 2000; Carpio

and Estrada 2006). This allows different aspects of the functions of mammalian orthologues to be investigated in isolation using the zebrafish model system, by examining for example, the orthologues' functions at different developmental stages and/or in different cell types (Dahm and Geisler 2006).

In recent years, the zebrafish model system have been exploited to model human diseases and health conditions, including cancer (Amatruda et al. 2002; Berghmans et al. 2005; Lam and Gong 2006; Goessling et al. 2007), infection and immunological disorders (Traver et al. 2003; van der Sar et al. 2004; Yoder et al. 2002), lung inflammation (Renshaw et al. 2007), aging (Keller and Murtha 2004; Gerhard 2007), cardiac development and regeneration disorders (Glickman and Yelon 2002; Poss 2007), kidney development and disease (Drummond 2005), tissue regeneration (Keating 2004), and muscle diseases (Bassett and Currie 2003; Guyon et al. 2007). Since modern drug discovery needs both *in vitro* and *in vivo* assays for the identification and validation of new therapeutics, the zebrafish model system has also been of interest to the pharmaceutical industry and used for drug discovery and evaluation (Rubinstein 2003; den Hertog 2005; Parng 2005; Zon and Peterson 2005; Kari et al. 2007; Martins et al. 2007; Silva et al. 2008).

The zebrafish has many advantages that make it an ideal experimental animal. The zebrafish is very small in size – adult fish measure about 3 to 5 cm in length. Therefore, zebrafish are very easy to maintain, minimizing the cost and effort to keep large populations of animals in the laboratory. The embryonic development of zebrafish

is extremely rapid. It takes only a few days to develop larvae from fertilized eggs. When incubated at 28.5°C, the gastrulation occurs around 6 hpf (hours post fertilization), and at 24 hpf, most vertebrate-specific body features and organs are recognizable. Over the next 24 hours, the organs gradually begin to function, and at 2 dpf (days post fertilization), embryos can start swimming. At 5 dpf, zebrafish larvae can swim well and search for food independently. This rapid embryonic development allows experiments to be completed within only a few hours or a few days (Haffter et al. 1996). The zebrafish larvae generally take about three months to reach sexual maturity. This relatively short generation time facilitates genetic analyses and has enabled the development of genetic transformation techniques.

In addition to rapid development, the zebrafish embryos have many properties that make the zebrafish model outstanding from other vertebrate model organisms. The zebrafish embryos develop outside the female (*ex utero*), they are relatively large (about 1 mm in diameter), and they are physically robust. These features allow easy observation and manipulation of the embryos during experiments. More importantly, the zebrafish embryos are completely transparent during early development. The external development and transparency of embryos allow real-time observation of the developing embryos (even deep inside) without the need for invasive intervention. In combination with fluorescent indicators, therefore, it is possible to visualize various cellular and/or embryonic developmental processes in a live and developing vertebrate embryo. In recent years, green fluorescent protein (GFP) labelled transgenic zebrafish has been

increasingly used to study the regulation of spatio-temporal gene expression, as well as cell migration during embryonic development (Udvadia and Linney 2003). The capability of real-time imaging in a natural developing environment makes the zebrafish an increasingly favourite organism to study embryonic development of vertebrates.

Zebrafish can generate large numbers of offspring. A single female adult fish can lay up to 200 eggs per mating per week, and embryos from the same batch develop synchronously. This feature, combined with rapid embryonic development, greatly facilitates statistical analyses as well as high-throughput approaches, including for example, mutagenesis screening, which is a forward genetics technique that is usually practical only in invertebrate model organisms. The capability of large-scale mutagenesis screening in the zebrafish is a powerful feature of this model organism.

In addition to its availability for forward genetics studies through which novel genes with essential functions could be screened and identified, analyses of the function of predicted or known genes can be carried out in the zebrafish model system using reverse genetics strategies, such as "morpholino knock-down" (Nasevicius and Ekker 2000), and TILLING (targeting induced local lesions in genomes) (Wienholds et al. 2002; Wienholds et al. 2003). In fact, with the unique advantages of zebrafish embryonic development, the zebrafish has indeed become a perfect organism to carry out morpholino analyses. The next section will discuss the methodologies for genetics studies in zebrafish, and will focus mainly on the morpholino knockdown technique and transgenic technologies in zebrafish.

1.5 Methodologies for Genetics Studies Using Zebrafish Model

Mutagenesis screening (forward genetics) is, to date, the most successful strategy to identify novel genes with an essential function in the zebrafish. The first large-scale genetic screens in zebrafish were employed in the early 1990s, and several hundred mutations were identified (Dahm and Geisler 2006; Haffter et al. 1996; Driever et al. 1996; Frohnhoefer 2002). In mutagenesis screening, mutations are randomly introduced by specific chemicals (usually ENU, *N*-ethyl-*N*-nitrosourea) or radiation (mainly γ-rays), followed by phenotypic analyses and searches for individuals with a mutant phenotype, thus leading to the identification of novel genes with unique or at least partially non-redundant functions. A more recent strategy is "insertional mutagenesis" (Raz et al. 1998; Amsterdam et al. 1999; Amsterdam and Hopkins 1999; Kawakami et al. 2000; Chen et al. 2002; Golling et al. 2002; Kawakami et al. 2004), where a known sequence (e.g., a transposon or a pseudotyped retroviral vector) is inserted into the zebrafish genome to obtain mutations. In this method, the inserted DNA sequence can serve as a starting point to identify the mutated gene (Dahm and Geisler 2006).

Genome sequencing projects have resulted in the identification of an increasing number of novel genes. A major challenge for biological research in the post-genome era will be to determine the functions of the newly identified genes, and in fact, the biological functions of many of the previously studied genes, which are not fully characterized. For example, *in situ* hybridization screens of zebrafish ESTs (expressed sequence tags) have identified the embryonic gene expression patterns of numerous

genes, but the functions of these genes have remained unknown (Crosier et al. 2001; Thisse et al. 2004). Reverse genetics approaches, such as knock-out or knock-down of specific genes, have offered a rapid characterization of the functions of these predicted or known genes. In the recent years, the gene specific knock-out (or knock-in) mouse has been the most successful protocol where loss-of-function (or gain-of-function) mutation genotypes can be attainable with homologous recombination in embryonic stem (ES) cells *in vitro* (Austin et al. 2004). Although many attempts have been made in recent years, the ES cell-based knock-out technique has not yet been established for zebrafish (Fan et al. 2004a; Fan et al. 2004b; Fan et al. 2006). In the past few years, however, an alternate method termed TILLING (targeting induced local lesions in genomes) has been successfully applied to establish stable knock-out zebrafish lines with functional mutations in specific genes (Wienholds et al. 2002; Wienholds et al. 2003; Wienholds and Plasterk 2004). This strategy was initially developed for the model plant *Arabidopsis thaliana* (Till et al. 2003; Till et al. 2006). In contrast to the homologous recombination-based knock-out technique in mice, TILLING uses ENU-induced random mutagenesis to generate point mutations in the genome of the F1 generation, and is followed by a subsequent screening for mutations in target genes. Once a mutation of interest is identified, this individual F1 fish is out-crossed to a wild-type fish in order to get heterozygous F2 fish. The heterozygous individuals of F2 generation are then in-crossed to generate the homozygous F3 fish, the stable knock-out zebrafish lines, for further phenotypic analyses. TILLING has become a powerful reverse genetics method

in zebrafish, however, since it is based on random mutagenesis, knock-in genotypes cannot be created using this method. In addition, the large-scale genetic screenings for F1 mutations make this technique less practical for smaller laboratories (Wienholds and Plasterk 2004; Draper et al. 2004).

In the past decade, several studies have attempted to establish RNAi-mediated knock-down techniques for the zebrafish model (Wargelius et al. 1999; Li et al. 2000). RNA interference (RNAi) has become an increasingly favourite gene silencing technique in the past ten years – the 2006 Nobel Prize in Physiology or Medicine was awarded to Andrew Z. Fire and Craig C. Mello for their pioneering work of RNAi in *C. elegans* (Fire et al. 1998). In RNAi experiments, small double-stranded RNA (dsRNA) is introduced into the cell cytoplasm. DsRNA acts through a conserved cellular machinery, which is in fact naturally employed by the micro RNA (miRNA) processing, leading to a sequence-specific translational inhibition and degradation of the base-paired endogenous mRNA and thus resulting in knock-down of the expression of the target gene (Hannon 2002; Murchison and Hannon 2004). The RNAi technique has been successfully applied to a wide range of species including fungi, plants, and animals (Hannon 2002). But in the zebrafish, the efficiency and specificity of RNAi remains unconvincing according to early studies (Oates et al. 2000; Zhao et al. 2001). Fortunately, another anti-sense gene silencing technique, morpholino knock-down, has been proven to work ideally in the zebrafish model (Nasevicius and Ekker 2000; Ekker 2000).

1.5.1 Morpholino knock-down

Morpholinos are chemically modified anti-sense oligonucleotides, which were initially developed for therapeutic applications in 1997 and then introduced for genetics studies in 2000 (Nasevicius and Ekker 2000; Ekker 2000; Summerton and Weller 1997; Arora et al. 2000; Qin et al. 2000; Heasman et al. 2000). Chemically, two modifications were introduced to make the morpholino oligos (MOs) more efficient and more stable for anti-sense applications. First, the ribose sugar in the backbone is replaced with a morpholine moiety. Second, the phosphodiester linkage is replaced by a phosphorodiamidate non-ionic linkage. With these modifications, the MOs still retain the same base stacking ability as natural nucleotides (DNA and RNA), but have a modified, neutral charged backbone, which enhances RNA-binding affinity and conveys excellent resistance to a wide range of nucleases and proteases. Therefore, MOs are extremely stable within biological systems, and ensure excellent anti-sense efficacy. In addition, MOs have a very broad range of working doses, and at low doses they show little or no cellular toxicity. These characteristics make MOs excellent anti-sense reagents (Summerton and Weller 1997).

Morpholinos are designed as short oligonucleotides (generally 18 to 25 nts long) to base-pair the mRNA of a target gene of interest (Ekker 2000; Corey and Abrams 2001). In contrast to traditional anti-sense approaches that utilize RNase-H-dependant degradation of mRNA as the mechanism of action, MOs binding to the target mRNA cannot trigger mRNA degradation since the RNA-morpholino hybrids are not substrates

for the RNase-H enzyme (Corey and Abrams 2001; Summerton 1999). It is believed that MOs act by blocking the 40s ribosome subunit scanning process, thus preventing the initiation of translation. For this reason, effective MOs must be designed to target the 5' UTR or immediately downstream (<25 nt) of the translational initiation codon in the coding region (Summerton 1999; Heasman 2002). Unlike traditional anti-sense oligos that are frequently designed to target the coding regions, MOs that target the coding regions do not work, because a fully assembled ribosome seems to be able to replace the bound MO when passing along the mRNA (Summerton et al. 1997; Sumanas and Larson 2002).

Morpholinos can also be designed to target against the exon/intron junctions (i.e., 5' or 3' splice sites) to block normal pre-mRNA splicing (Draper et al. 2001). This is an alternate gene knock-down mechanism of morpholinos, in which morpholino oligos interfere with pre-mRNA splicing events, resulting in either non-functional mRNA products or alterations in alternative splicing patterns. Since alternative splicing seems to be the rule rather than the exception in gene expression and regulation, morpholinos have provided a powerful tool for analyses of the regulation and function of alternative splicing as well as a variety of potential gene therapeutic applications (Lacerra et al. 2000; Mercatante et al. 2001; Woodley and Valcarcel 2002). More recently, morpholinos have also been demonstrated to block microRNA (miRNA) maturation, thereby preventing the degradation of target mRNA by RNA-induced silencing complexes (RISCs) (Kloosterman et al. 2007; Woltering and Durston 2008).

Zebrafish embryos are ideally suited for morpholino analysis due to the ease of delivery by microinjection and the ease of handling and subsequent observations (Nasevicius and Ekker 2000; Sumanas and Larson 2002; Ekker and Larson 2001). MOs are microinjected into the yolk of zebrafish embryos at early developmental stages (usually 1 to 8 cell stages). In zebrafish embryonic development, the yolk does not participate in the cleavage divisions but remains distinct from the embryo. This feature facilitates morpholino oligos to be distributed evenly over the whole embryo throughout early development. The anti-sense silencing effects of morpholino oligos are most efficient during the first few days during embryonic development. Their efficiencies progressively decrease later, possibly due to the dilution effects as cells divide and the MOs are possibly excreted (Sumanas and Larson 2002). But because most of the patterning, morphogenesis, and organogenesis in zebrafish occur during the first two or three days after fertilization, morpholino knock-down can reliably cover the full duration of embryonic development. Therefore, the zebrafish is indeed a perfect model organism for morpholino analyses (Sumanas and Larson 2002; Ekker and Larson 2001).

While MOs have been successfully applied in some model organisms, including zebrafish, frogs (*Xenopus*), sea urchin, chick, and mouse, etc., there are also many limitations worth considering. First, difficulties in the delivery of MOs into different experimental organisms restrict its applications in a wider range of organisms. In fact, to date, the MO technique seems to work best in systems where MOs can be delivered into the early embryos by microinjection, for example, frogs and zebrafish. MOs were first

developed as human therapeutic reagents (Summerton and Weller 1997), but the delivery of MOs into human cells is challenging, and despite considerable effort, delivery is still a big issue to be considered when applying MOs to various systems (Sumanas and Larson 2002).

Secondly, like other anti-sense techniques, the possibility of non-specific targeting and non-specific phenotypes is also a big concern, although MOs are more efficient than other anti-sense reagents (Sumanas and Larson 2002). Non-specific phenotypes could be the result of toxicity of the foreign morpholino oligos (especially at high injection doses), however, a more likely source is mis-targeting. When unanticipated phenotypes are observed following MO delivery, it is important to consider whether genes other than the one of interest have been affected. For these reasons, proper controls are a must for morpholino experiments (Sumanas and Larson 2002). Many researchers use random or unrelated sequence MOs as their controls. While these serve as controls for general toxicity, they may not address issues involving mis-targeting. Some researchers use the four-base mismatched MOs as controls, but these do not directly address mis-targeting issues. Western blots, immuno-histochemistry or reporter constructs are often used to demonstrate the inhibition of gene expression. Although it is very helpful to know that the gene of interest was targeted by the MOs, these types of analyses also do not address concerns about mis-targeting. For these reasons, RNA rescue experiments have been considered to be a must to confirm the specificity in anti-sense gene silencing techniques. Although RNA rescue is convincing, it is often

complicated due to RNA over-expression artefacts as well as many difficulties in mimicking the endogenous spatio-temporal gene expression patterns. Luckily for morpholino knockdown, a mediatory strategy is available where multiple non-overlapping MOs are used simultaneously. If phenotypes separately obtained from these two single MOs are identical, this confirms the specificity of MOs because it is much less likely that two different MOs produce the same phenotype by mis-targeting (Sumanas and Larson 2002). Furthermore, it was reported that non-overlapping MOs sequences targeting the same mRNA can produce a synergistic effect when co-injected. Therefore this synergistic effect can be used to reduce the dosage of morpholinos injected, thus reducing the possibility of mis-targeting as well as cell toxicity (Ekker and Larson 2001; Sumanas et al. 2001).

Thirdly, MOs only knock down gene expression, and there has been no evidence that gene expression of target genes can be completely turned off. In fact, to start morpholinos experiments, a series of dosage trials are absolutely needed in order to determine a suitable concentration which should be as low as possible to avoid toxicities but sufficient enough to obtain expected phenotypes (Sumanas and Larson 2002). Also, MOs only produce loss-of-function phenotypes rather than genotypes. For these reasons, a morpholino knock-down phenotype is less reliable and less reproducible than that of a genetic mutant. In contrast with the knock-out mice technique, therefore, morpholino knock-down is indeed transient in nature, and loss-of-function phenotypes can not be transmitted into the following generations.

Finally, the progressive dilution and excretion of the MOs in the developing embryo restrict its applications only within the early embryonic stages. But many genes may have different roles (including different protein products due to alternative splicing) in the adults rather than in the embryos. To fully elucidate the functions of such genes requires more complicated techniques.

Many of these drawbacks of morpholino knock-down have been overcome by the TILLING technique (mutagenesis-based knock-out zebrafish) in recent years. But without a doubt, morpholino knock-down is a powerful tool to address the functions of specific genes during zebrafish embryonic development. Compared with TILLING, morpholino knock-down is technically simple, cost-effective, and time-efficient – it only takes a few hours or a few days for the determination of the function of a single gene in the developing zebrafish embryos.

1.5.2 Zebrafish transgenesis

Transgenesis, the introduction of exogenous genes into the host genome, has been widely used to study the expression and function of specific genes *in vivo* as well as endow the host organism with properties of interest. The potential for transgenic zebrafish was first demonstrated about 20 years ago by microinjecting plasmid DNA into the cytoplasm of fertilized eggs and subsequently generating transgenic zebrafish lines in which the injected gene was able to be stably transmitted to next generations (Stuart et al. 1988; Stuart et al. 1990). Taking advantage of the accessibility and optical clarity of

zebrafish embryos, fluorescent reporter genes can be assayed in live and developing embryos, making it possible to visualize and track the details of dynamic gene expression as well as cellular processes during embryonic development. In the past two decades, zebrafish transgenesis has become a powerful tool for a variety of genetics and developmental studies including analysis of regulatory elements, tissue-restricted gene expression, targeted mis-expression, tracking of cellular linage, and insertional mutagenesis (Udvadia and Linney 2003; Muller et al. 2002; Amsterdam and Becker 2005; Amsterdam 2006; Sivasubbu et al. 2007).

Targeted over-expression has been used to analyze the function of the gene of interest. For this purpose, it is particularly important to obtain tight spatio-temporal expression of the transgene in order to study its roles in the development of a specific cell type or at a particular developmental stage. This is achieved by using tissue- or stage-specific promoters, as well as inducible promoters (Udvadia and Linney 2003). One example is the use of a heat-shock promoter, which allows temporal control of transgene expression at any given time simply by increasing the ambient temperature (Adam et al. 2000; Halloran et al. 2000; Yeo et al. 2001; Bajoghli et al. 2004; Grabher and Wittbrodt 2004; Wu et al. 2008). Studies have shown that heat-shock induced expression of transgenes can also be spatially controlled (Halloran et al. 2000; Ramos et al. 2006; Hardy et al. 2007). Other genetic tools to regulate gene expression have also been exploited to achieve tight control. These include the Gal4-UAS system (Scheer and Campos-Ortega 1999; Scheer et al. 2002; Koster and Fraser 2001; Davison et al. 2007;

Scott et al. 2007), the tetracycline (Tet-On) system (Huang et al. 2005), and the Cre/loxP system (Langenau et al. 2005; Feng et al. 2007).

Another major task of zebrafish transgenesis is to analyze the function of putative transcriptional regulatory sequences. To this end, the transient expression assay is very useful for rapidly testing the capacity of a specific regulatory DNA sequence to drive reporter gene expression. A number of studies have shown that the transient expression assay can be used *in vivo* to analyze the cis-acting promoter/enhancer elements of specific genes that regulate spatio-temporal expression (Westerfield et al. 1992; Moav et al. 1993; Reinhard et al. 1994; Marsh-Armstrong et al. 1995; Kim et al. 1996; Moss et al. 1996; Long et al. 1997; Meng et al. 1997; Chen 1998; Hieber et al. 1998; Ju et al. 1999; Meng et al. 1999; Muller et al. 1999; Higashijima et al. 2000; Motoike et al. 2000; Park et al. 2000; Chen et al. 2001; Du and Dienhart 2001; Huang et al. 2001; Udvadia et al. 2001; Picker et al. 2002). For example, the analyses of deletion mutations and point mutations within the GATA-1 and GATA-2 gene promoters using GFP transient expression assays resulted in the isolation of regulatory elements responsible for tissue-specific expression (Meng et al. 1997; Meng et al. 1999). Another study has identified different enhancers that can control tissue-specific expression of zebrafish *sonic hedgehog (shh)* in either notochord or floorplate tissues (Muller et al. 1999). The transient assay has also been employed in over-expression studies to analyze the function of the transgene *in vivo*. One early example was the transient expression of chicken *gicerin* in zebrafish embryos under control of the neurofilament gene promoter in which

neuron-specific expression patterns were preferentially achieved (Kim et al. 1996).

However, one notable drawback of the transient transgenic assay is that the expression of the transgene in the injected embryos is typically ectopic and highly mosaic. The transient assay cannot reliably reproduce the expression patterns of endogenous genes, thus making it very challenging to pool expression information from such a mosaic and ectopic expression background (Udvadia and Linney 2003). In order to circumvent these problems, researchers have proposed to use artificial chromosomal vectors (BAC or PAC) that can accommodate much larger DNA insertions than plasmids. Reporter constructs that carry larger fragments (e.g., 20 – 55 kb) are less likely to be missing crucial regulatory elements and thus less likely to display ectopic expression (Jessen et al. 1998). This method is helpful to reproduce more reliable expression patterns of transgenes in injected embryos, but it is not suitable for mapping promoters because crucial *cis*-acting elements are also frequently located in the proximal region of the promoter. Thermes et al (2002) domenstrated one method in medaka fish (*Oryzias latipes*), by injecting a transgene construct flanked with the *S. cerevisiae* homing endonuclease *I-SceI* sites, which can significantly enhance promoter-dependent expression. Another study showed that flanking the transgene with adeno-associated virus inverted terminal repeats (AAV-ITRs) could significantly reduce mosaic expression of transgenes driven by ubiquitous promoters in injected embryos. But it remains to be tested whether it is effective for analyzing weakly expressed and developmentally regulated gene promoters (Hsiao et al. 2001). More recently, vectors based on the

transposon *Tol2* (derived from medaka fish) have also been used to reduce F0 mosaic expression as well as increase germline transmission of transgenes (Fisher et al. 2006; Kawakami et al. 2004). In addition, a dual fluorescence strategy has been developed, which allows direct comparison of the expression of a regulated transgene with that of a ubiquitously expressed transgene, thus facilitating the evaluation of transient expression of the transgene (Szeto and Kimelman 2004).

In contrast, stable transgenic lines can readily provide spatial and/or temporal gene expression control, while removing the inter-embryo variability that results from the microinjection and mosaic expression in injected embryos. Generating stable transgenic lines is generally much more time-consuming, but provides more reliable enhancer analysis as well as valuable genetic screening tools. For example, the generation of several *islet1-GFP* transgenic zebrafish lines has been successfully employed to map the regulatory elements controlling *islet1* branchimotor neuron expression (Higashijima et al. 2000; Uemura et al. 2005). These transgenic lines have also facilitated the genetic screening for identification of mutants affecting neuron migration and axon guidance (Wada et al. 2006; Tanaka et al. 2007). In the past two decades, numerous transgenic zebrafish lines have been established for studies on spatio-temporal gene control, as well as studies of cell migration and targeted over-expression (Udvadia and Linney 2003; Shafizadeh et al. 2002).

Microinjection of either plasmid DNA or bacterial artificial chromosomes (BACs) into one-cell stage zebrafish embryos has been the most commonly used method for

making transgenic zebrafish (Stuart et al. 1988; Culp et al. 1991; Amsterdam et al. 1995; Jessen et al. 1999). The foreign DNA is usually injected into the egg cytoplasm (or in some cases, the yolk), rather than one of the pronuclei, which are not actually visible in fertilized zebrafish eggs. High copy numbers (10^5 - 10^6) of DNA constructs are generally necessary in order to achieve reliable nuclear uptake and chromosome integration. The integration of foreign DNAs into the zebrafish chromosomes occurs randomly and the integrated DNA usually exists as tandem concatemers of many transgene copies, and may cause chromosomal rearrangements (Stuart et al. 1988; Stuart et al. 1990; Culp et al. 1991; Cretekos and Grunwald 1999). It is believed that enzymes that normally function in DNA repair and recombination are involved in the integration process. However, the detailed molecular mechanism for the integration remains unclear.

Transgenesis by direct injection of foreign DNAs is the most convenient and cost-effective technique, however, the expression of transgenes in F0 founders is highly mosaic and the efficiency of transgene transmission to the F1 generation is generally low (approximately 5 – 20%). The mosaic expression pattern of the transgene and overall low efficiency of germ line transmission might be due to extra-nuclear injection as well as rapid embryonic cell divisions (Stuart et al. 1988; Culp et al. 1991). In the past decades, new methods have been developed to improve the efficiency of integration of the transgene into the zebrafish genome. One of the methods relies on flanking the transgene construct with the *S. cerevisiae* homing endonuclease *I-SceI* sites, which was demonstrated in medaka fish (*Oryzias latipes*) (Thermes et al. 2002). The recognition site

of *I-SceI* meganuclease is 18 bp in length, and thus, only one cut might be found in about 7×10^{10} bp of random sequence. The *I-SceI* does not cut the medaka genome (~7×10^8 bp), but it can release the transgene fragment from the circular plasmid as they are co-injected into the one-cell stage fish embryos. It is thought that the co-injected *I-SceI* meganuclease likely counteracts the endogenous ligase activity, preventing the formation of long concatemers and thus provides more recombinogenic ends that facilitate efficient integration. The early integration facilitates a very high germline transmission rate of nearly 50% (Thermes et al. 2002). Although to date, there has been no report on its application in zebrafish transgenesis, this method should be very promising because the zebrafish genome is also relatively small (~3.5×10^9 bp) and it's more likely that the *I-SceI* meganuclease does not cut the zebrafish genome.

Another method uses mobile DNA elements known as transposons to enhance the efficiency of integration. This idea was originated from the applications of transposable P elements in *Drosophila* (fruit flies) germline transgenesis in the early 1980s (Rubin and Spradling 1983; Clough et al. 1985). Transposon vectors contain the *cis*-acting elements (e.g., short inverted DNA repeats) from natural transposons, but the transposase gene is replaced by the transgene of interest that can be up to 10 kb in length. The transposase necessary for the integration reaction is supplied by co-injecting it with the modified transposon either on a separate DNA vector or as synthetic mRNA (Sivasubbu et al. 2007; Balciunas et al. 2004; Kawakami 2005; Korzh 2007). Although a number of transposon vectors have been tested in zebrafish, only the *Sleeping Beauty* (an ancient fish

transposon) and *Tol2* (an *Ac*-like transposable element) systems are currently in use (Raz et al. 1998; Kawakami et al. 2000; Fadool et al. 1998; Kawakami et al. 1998; Kawakami and Shima 1999; Davidson et al. 2003). The *Sleeping Beauty* transposon system improved the rate of germ line transmission to about 35% (Davidson 2003), while the *Tol2*-mediated transgenesis can allow nearly 50% of the F0 zebrafish to pass the transgene to their progeny (Kawakami et al. 2004). In addition, DNA integration using the transposon mechanism provides a chance to excise the integrated transgene and re-integrate it at new genomic sites by injection of corresponding transposase. This characteristic is also very useful in insertional mutagenesis and gene and enhancer trap screens (Sivasubbu et al. 2007; Kawakami 2005; Korzh 2007). However, as a DNA-injection-based method, the major concern of this method is that many integration events are not transpositions but rather random integrations of concatemerized plasmid DNA (Raz et al. 1998; Kawakami et al. 2000).

Pseudotyped retroviral vectors have also been applied to zebrafish transgenesis in recent years (Chen et al. 2002; Amsterdam and Becker 2005; Lin et al. 1994; Gaiano et al. 1996a; Linney et al. 1999; Kikuta et al. 2007). In order to integrate foreign DNA into the target cells, retrovirus vectors are engineered to include the *cis*-acting sequences necessary for reverse transcription and integration, but do not encode viral proteins. High-titer engineered viruses are prepared in cultured cells with the viral proteins being supplied in *trans*. This method can ensure that the transgene is efficiently integrated into the host genome, but also prevents assembling additional virus in the host cells after

infection. The retroviral vector system for zebrafish transgenesis was derived from the Moloney murine leukemia virus (MoMuLV). The genome of MoMuLV is well suited for DNA insertions, but its envelope protein restricts infection only to mammalian cells. This problem was solved by pseudotyping a MoMuLV vector with the envelope glycoprotein (G) originated from the vesicular stomatitis virus (VSV). The MoMuLV(VSV) pseudotyped virus thus offers an extensive range of host species, including zebrafish. The engineered virus is generally injected into blastula-stage zebrafish embryo (500 – 2,000 cells) (Chen et al. 2002; Lin et al. 1994; Gaiano et al. 1996a). Different cells are independently infected, and a single copy of the transgene can be integrated at different chromosomal sites in different cells, thus resulting in a highly mosaic organism in the F0 generation. In blastula-stage zebrafish embryo, a few cells are differentiated into primordial germ cells (PGCs), and the infected PGCs can reliably transmit the transgene to the F1 generation at a relatively high rate (1–20%) (Amsterdam and Hopkins 1999; Chen et al. 2002; Gaiano et al. 1996a). In addition to efficient germ line transmission of the transgene, retroviral-mediated DNA integration typically introduces only a single copy of the transgene without causing chromosomal rearrangement. These features are valuable for insertional mutagenesis and gene- (and enhancer-) trap screens in zebrafish (Amsterdam et al. 1999; Golling et al. 2002; Gaiano et al. 1996b; Amsterdam and Hopkins 2004; Ellingsen et al. 2005). However, one disadvantage of this method is the frequent difficulty in obtaining virus preparations with sufficiently high titer, which are essential for zebrafish transgenesis (Kawakami 2005). Because different insertion

sequences in the virus vector can significantly affect the titer by unknown mechanisms, the design of specific DNA constructs usually requires extensive testing to find one that can work (Chen et al. 2002). The construction, packing, titering, and infection required by retroviral-mediated transgenesis are relatively time-consuming activities and therefore make this method less practical for the routine generation of transgenic zebrafish (Udvadia and Linney 2003).

Some other methods have been used to facilitate the nuclear uptake of plasmid DNA by binding nuclear proteins or nuclear localization signal (NLS) peptides (Collas and Alestrom 1997; Collas and Aleström 1998; Liang et al. 2000). This improvement is significant, especially considering the fact that zebrafish embryos divide very rapidly. The rapid and efficient nuclear localization of the injected DNA can allow integration events to occur at an earlier cell cleavage stage, thus reducing the frequency of mosaic expression in the F0 founders as well as increasing the rate of germ-line transmission. Another study has demonstrated gene transfer by sperm nuclear transplantation, in which sperm nuclei were pre-incubated with linearized plasmid DNA and then injected into the animal pole region of the unfertilized zebrafish eggs (Jesuthasan and Subburaju 2002). This technique generates non-mosaic expression of the transgene in the F0 founder zebrafish. This approach allows experiments to be completed in a few days, without the need to establish stable lines. However, this technique demands much more technical expertise, compared with other methods, and is therefore less practical for most laboratories (Jesuthasan and Subburaju 2002).

1.6 Thesis Objectives

Dr. Valdimarsson is one of the early pioneers to use zebrafish as a model to study connexins and gap junctional communication. This lab has made substantial contributions to the gap junction community by cloning and characterizing novel zebrafish connexin genes and by analyzing their expression and function during embryonic development (Cheng et al. 2003; Chatterjee et al. 2005; Cason et al. 2001; McLachlan et al. 2003; Valiunas et al. 2004; Christie et al. 2004; Cheng et al. 2004; Shearer et al. 2008). When I started my thesis research in this lab in 2003, a large group of zebrafish connexin genes had been previously isolated and characterized, including Cx43, Cx44.1, Cx48.5, Cx45.6, Cx35, Cx32.3, and Cx35.5. The expression patterns of these connexin genes had been intensively investigated and the functional analyses of both Cx44.1 and Cx48.5 during zebrafish embryonic development had been examined using morpholino oligo knockdown techniques (Cheng et al. 2004). The functions of these connexin genes during zebrafish embryonic development and the regulatory mechanisms for connexin gene expression have been two important aspects of the research in this lab. My thesis projects have been developed in this background.

As previously mentioned, the distinct but overlapping expression patterns of connexin genes suggest complicated regulatory mechanisms for their expression and function, but nevertheless, our knowledge of many connexin gene structures and the mechanisms for the regulation of gene expression is still limited. Therefore, one of the main objectives for this thesis was to dissect the gene structure and regulatory

mechanisms for the expression of several zebrafish connexin genes isolated in this lab. Another proposed objective for my thesis research was to analyze the function of connexin genes during zebrafish embryonic development using "morpholino knockdown". As a reverse genetic strategy, "morpholino" analysis relies on the information of mRNA sequences as well as their spatio-temporal expression. As discussed previously, in order to block the initiation of translation, effective morpholino oligos must be designed to target the 5'UTR of the gene of interest. For these objectives, my thesis research has focused on the following aspects:

1) mapping the transcriptional start sites and obtaining the promoter sequences of these connexin genes;
2) identifying the 5'UTR sequences and dissecting the gene organizational structure of these connexin genes;
3) analysing the *cis*-regulatory sequences and activation for transcriptional regulation of some of these connexins;
4) analysing mRNA expression patterns of these connexin genes;
5) designing Cx-promoter driven GFP expression constructs and transgenic zebrafish lines for both transcriptional studies and functional analyses of some of these connexin genes;
6) designing morpholino oligos and functional studies for some of these Cx genes.

CHAPTER TWO

ZEBRAFISH CONNEXIN 44.1

2.1 Introduction

The vertebrate lens is a highly specialized organ with relatively simple tissue organization that comprises only two cell types – cuboidal epithelial cells on the anterior surface, and elongated fibre cells forming the bulk of the lens tissue. The epithelial cells are metabolically active cells in the lens, whereas the lens fibres, which are devoid of nuclei and organelles to reduce light scattering, are metabolically inactive. Because the lens is an avascular organ, maintenance of lens cell homeostasis relies upon an elaborate system of gap junctions that allow ions, second messengers, and small metabolites to be exchanged between cells in the epithelial layer through to the metabolically inert fibre cells deep within the lens tissue (Gerido and White 2004; White et al. 1992; Cason et al. 2001; Paul et al. 1991; Rup et al. 1993; Jiang et al. 1994; Jiang and Goodenough 1996; Rae et al. 1996; Gao and Spray 1998; Berthoud et al. 1999; Ebihara et al. 1999; Francis et al. 1999; Kistler et al. 1999; Dahm et al. 1999; Martinez-Wittinghan et al. 2004; Gong et al. 2007; Mathias et al. 2007; Berthoud and Beyer 2008). For these reasons, the lens is an excellent model for studies on the expression and function of gap junctional proteins.

In the vertebrate lens, three different connexins have been identified. Mammalian Cx43 is expressed in the surface epithelial cells (Beyer et al. 1989; Musil et al. 1990; Yancey et al. 1992), whereas Cx50 and Cx46 are expressed in the fibre cells (White et al.

1992; Paul et al. 1991). Mouse knockout studies have demonstrated that Cx50 plays essential roles in lens development and adult lens tissue homeostasis, and the ablation of Cx50 in knockout mice resulted in the formation of cataracts (White et al. 1998; White et al. 2001; Sellitto et al. 2004). The Valdimarsson lab has previously isolated and characterized the zebrafish Cx44.1 gene, the ortholog of the mammalian Cx50 (Cheng et al. 2003; Cason et al. 2001). In order to elucidate the functional roles of zebrafish Cx44.1 during lens development, morpholino oligo (MO) knock-down analyses were previously performed in this lab. The MO was designed to target the sequence of the Cx44.1 mRNA immediately downstream the translational start codon AUG, according to general guidelines from Gene Tools, LLC (Philomath, OR). Morpholino-injected embryos showed reduced lens growth and the formation of cataracts (Cheng et al., unpublished data). However, this morpholino was later found to also target another lens connexin gene, ZfCx48.5, which shared a high degree of identity with Cx44.1 in the 5' end of the gene (there was only one nucleotide difference between the two genes in their first 24 nucleotides). Although this was a perfect morpholino oligo for the potential "double knockdown" studies of Cx44.1 and Cx48.5 in zebrafish lens (Cheng et al., unpublished data), morpholino analyses for zebrafish Cx44.1 would require appropriate gene-specific morpholino oligos. As effective MOs should bind to the 5'UTR or close to the start codon, it was essential to determine the 5'UTR sequence in order to conduct functional analyses of Cx44.1 by MO knock-down.

In this study, the 5'UTRs of Cx44.1 mRNA were successfully identified using 5'

RACE techniques. With the 5'UTR sequences of zebrafish Cx44.1, we can now design effective morpholino oligos to study the functional roles of Cx4.1 during zebrafish lens development, as well as molecular and genetic mechanisms for certain lenticular diseases, such as cataracts. More interestingly, our 5'RACE results showed that transcripts of Cx44.1 contain alternative 5'UTRs that are derived from different promoter usages and alternative pre-mRNA splicing. To look into the possibility of whether these features of Cx44.1 mRNA expression are evolutionarily conserved, I also performed 5'RACE experiments to identify the transcriptional start sites and the 5'UTR sequence of mouse Cx50. Alternative 5'UTRs of mCx50 were subsequently identified that are also generated by different promoters and alternative splicing. These findings suggest that the expression of Cx44.1 and Cx50 could be finely modulated by complicated mechanisms at both the transcriptional and the post-transcriptional levels, and to some degree, the regulatory mechanisms might be conserved. Comparisons between the proximal promoters of both connexin genes were performed by both *in silico* prediction of potential transcription factor binding sites and *in vivo* GFP transient expression assays. These analyses have provided preliminary evidence that transcriptional regulatory mechanisms for both Cx genes could be evolutionarily conserved.

The discovery of multiple transcript variants of zebrafish Cx44.1 has intrigued us to further look into the functional roles of each individual Cx44.1 transcript variant during lens development. Morpholino oligos that specifically target different splicing variants can be used for this purpose. However, because all transcript variants share the

same coding sequence (therefore produce the same connexin protein), more detailed expression patterns of these transcript variants are necessary before we could successfully determine the functional roles of each types of transcript of Cx44.1 using morpholino techniques. In this study, I performed real-time RT-PCR to characterize the expression patterns of these transcript variants during zebrafish embryonic development as well as in adult lens tissues. The comprehensive analyses of the expression and regulation of both ZfCx44.1 and mCx50 presented in this study have provided necessary information for us to conduct functional studies in the future, as well as studies on the regulatory mechanisms for the expression of both connexin genes.

2.2 Materials and Methods

2.2.1 Fish care.

Wild type zebrafish (*Danio rerio*) were purchased from a local pet store and maintained in an aquarium facility on a daily light/dark cycle of 14 h light and 10 h dark at 28.5 °C. The embryos were incubated in Egg water (60 µg/ml Instant Ocean sea salts) in a 28.5 °C water bath, and staged using time elapsed since fertilization. Embryos used for EGFP expression studies were treated with 0.003% 1-phenyl-2-thiourea (Sigma, Mississauga, ON, Canada) in Egg water beginning at ~24 hpf (hours post fertilization) to prevent pigment formation. Zebrafish maintenance, breeding and embryo collection followed standard protocols (Westerfield, 1995). The use of experimental animals was approved by the University of Manitoba Institutional Animal Care Committee.

2.2.2 RNA extraction:

RNA samples in the present study were isolated using Trizol reagent (Invitrogen Corp., Carlsbad, CA) according to the manufacturer's general instructions. RNA concentration was determined by UV spectrometry. The A260/A280 ratios were between 1.8 and 2.0 for all samples. The quality of RNA samples was also tested by RNA gel electrophoresis. For RACE-PCR and real time PCR experiments, no DNase treatments were performed so as to avoid as much as possible any loss and degradation of RNA.

2.2.3 5'RACE (Rapid Amplification of cDNA Ends)

5' RACE experiments were performed using a GeneRacer^{TM} Kit (Invitrogen) according to the manufacturer's recommended protocols. In brief, RNA samples for 5'RACE were isolated from adult zebrafish lens tissues and adult mouse lens tissues respectively. RNA samples were treated with calf intestinal phosphatase (CIP) in order to de-phosphorylate the 5'end of truncated mRNA and non-mRNA molecules. The de-phosphorylated RNA samples were then treated with tobacco acid pyrophosphatase (TIP) to remove the 5' cap structure from mature full-length mRNA, leaving a 5' phosphate group available. Using T4 RNA ligase, a universal RNA oligonucleotide (GeneRacer^{TM} RNA oligo, 5'-CGA CUG GAG CAC GAG GAC ACU GAC AUG GAC UGA AGG AGU AGA AA-3') was then ligated to the de-capped mRNA, providing a priming site for the following PCR reactions. First strand cDNAs were then reverse

transcribed using SuperScript III reverse transcriptase with a mix of random hexamers and the GeneRacer™ oligo dT primer (5'-GCT GTC AAC GAT ACG CTA CGT AAC GGC ATG ACA GTG $(T)_{24}$-3'). Primers used for the first round and nested PCR are listed in Table 2-1.

The nested PCR products were cloned into the pCR4-TOPO vector (Invitrogen). The ligation mixture was then used to transform One Shot® TOP10 chemically competent *E. Coli* cells (Invitrogen). Positive colonies were randomly isolated and grown in LB medium at 37°C with shaking (250 rpm) overnight. Independent plasmids were extracted and then sent out for sequencing (DNA Centre, University of Calgary). The 5'UTR sequences were then aligned against genomic DNA sequences to determine their transcriptional start sites and exon-intron structures. Using RNA structure prediction software mFOLD (Zuker 2003), we also analyzed the secondary structure of 5' UTR of zebrafish Cx44.1 and mouse Cx50.

2.2.4 Sequence analyses of ZfCx44.1 and mCx50 promoters

Analyses of zebrafish Cx44.1 (ZfCx44.1) and mouse Cx50 (mCx50) promoters were performed using the MatInspector program (Cartharius et al. 2005; Quandt et al. 1995), a tool that utilizes a library of matrix descriptions for transcription factor binding sites to locate matches in sequences of unlimited length. Potential transcription factor binding sites were predicted using the MatInspector program from Genomatix web service (http://www.genomatix.de).

2.2.5 Real Time RT-PCR

RNA samples for real-time RT-PCR of ZfCx44.1 were isolated from zebrafish embryos (0, 1 – 5 dpf stages) and adult lens tissues, while for mouse Cx50, RNA samples were from adult lens tissues. For each sample, DNase pre-treatment was not performed in order to maintain as much as possible the RNA integrity, as some DNases are suspected of having contaminating RNase activity. Total RNA (3 μg) was reverse transcribed by Superscript III (Invitrogen), with oligo-$dT_{(18)}$ (Invitrogen), for 1 hour at 52°C, in a reaction volume of 20μl. The resulting cDNA was then diluted to 100μl with ddH_2O for further use. A reverse transcriptase negative control was also used for monitoring genomic DNA contamination.

Primer pairs for ZfCx44.1 and mCx50 were designed to specifically target splice variants (Figure 2-4, and 2-6). In order to eliminate the amplification of potential contaminating genomic DNA, all primer pairs have at least one primer spanning an exon-exon boundary. All primers were manually designed, and their properties were checked by free software OligoAnalyzer 3.0 (http://www.idtdna.com/ANALYZER/Applications/OligoAnalyzer/default.aspx?err=true, Integrated DNA Technologies, Coralville, IA, USA). Zebrafish elongation factor 1α (zEF1α) and mouse elongation factor 1α-1 (mEF1α1) were used as reference genes for analyses of ZfCx44.1 and mCx50 respectively. Similarly, both primer pairs of the reference genes were designed to span their first introns respectively, so as to avoid

amplifying genomic sequences in real-time PCR conditions. Standard-curve tests were performed according to general quantitative PCR guidelines in order to validate the specificity and efficiency of all primer pairs. All primer pairs gave acceptable high specificities and efficiencies, thus allowing the comparative cycle threshold (ΔCt) method of analysis to be used (Pfaffl 2001). All primer pairs and their anticipated amplicon size are summarized in Table 2-2.

PCR reactions were performed on an iQ5 Real-Time PCR detection system (BIO-RAD, Hercules, CA, USA). cDNA samples (2μl per reaction) were mixed with 250 nM of each primer and 10μl of SYBR Green Supermix (BIO-RAD) reagent in a final volume of 20μl. In addition, a reverse transcriptase negative control and a template-negative control were also set up in parallel. PCR conditions were as follows: 95°C for 3 min followed by 40 cycles at 95°C for 10 sec, 62°C for 10 sec, and 72°C for 10 sec. After amplification, melting curve analyses were used to check PCR specificity, and/or monitor genomic DNA contamination.

Each sample was tested in at least duplicate reactions and all PCR runs were performed twice with cDNAs prepared from independent batches of RNA samples. Mean Ct values were calculated, and then used for the calculation of ΔCt. Results were expressed as a ratio of mRNA expression level to that of the EF1α references. The ratios were calculated by calculating ΔCt between mean Ct values of test and reference genes, and were presented as column graphs using MS Excel program.

2.2.6 EGFP transient expression driven by connexin promoters in fish embryos

EGFP expression constructs were prepared to assess the function of the zebrafish and mouse connexin gene promoters. The recombinant DNA constructions were prepared by PCR method, and Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen) was used to ensure the sequence accuracy of PCR products. The detailed method is described in Appendix I. In brief, by two rounds of PCR reactions, the zebrafish Cx44.1 promoter (a ~1.1 kb genomic DNA fragment immediately upstream of the translational start codon) was linked to the EGFP coding sequence followed by a SV40 poly-A tail (Figure 2-1, C). Similar strategies were applied to making the mouse Cx50-EGFP construction, and two different Cx50-EGFP constructs were prepared (Figure 2-2, E and F). For one promoter (P2), the entire 5'UTR sequence was retrieved from mouse cDNA by a first round PCR, and then linked to a ~300 bp promoter fragment immediately upstream of the 5' UTR, which was derived from genomic DNA also by PCR (Figure 2-2, A & B). By a third round of PCR, the resulted recombinant DNA fragment was then linked to the above EGFP expression component resulting in the final construct (Figure 2-2, E). For the other promoter (P3), which is located within the intron, a ~1 kb promoter fragment upstream the translational start codon was PCR amplified from genomic DNA (Figure 2-2, C), and then linked to the EGFP coding sequence and polyadenylation sequence by a second round of PCR, resulting the final EGFP construct (Figure 2-2, F). All PCR primers used for making these recombinant constructs are listed in Table 2-3.

The final PCR products (i.e., the recombinant DNA constructs) were purified

using the GENECLEAN® Spin Kit (BIO-101 Inc., Vista, CA,USA), and then the concentration was adjusted to about 100 ng/μl in 100 mM KCl solution with 0.05% of phenol red (as the injection indicator). DNA constructs were injected into 1-cell stage zebrafish embryos. Injected and non-injected control embryos were separately incubated in egg water in a water-bath incubator with the temperature set at 28.5°C. Egg water was regularly changed every 4 - 8 hours. At about 1 dpf stage, 0.003% of PTU (1-phenyl-2-thiourea) was added into the egg water in order to repress the pigment formation. Embryos were dechorionated by forceps and relocated to 24-well plates, with 1.5 - 2 ml of fresh egg water, and one embryo in each well. Each injected embryo was then assigned an identification number. Fluorescence expression was observed and recorded from this stage until day 5. During observations, live embryos were temporarily anaesthetized in 4% Tricaine (Sigma). Images were captured and recorded using a Zeiss Axioskop FS microscope equipped with a Sony DXC-950 CCD camera and Northern Eclipse software (Empix, Mississauga, ON).

2.3 Results

2.3.1 Promoter usage and alternative splicing of ZfCx44.1 and mCx50

Using the 5'RACE method, I identified the transcription start site and 5' UTR sequence of zebrafish Cx44.1 and the corresponding mouse ortholog Cx50. This is a RLM- (RNA ligase-mediated) 5'RACE strategy, which ensures that only cDNAs produced from mature capped mRNA transcripts are selected for the following PCR

amplification, and thus only the full-length of the 5' ends can be amplified and sequenced. This feature makes it possible to identify the transcriptional start sites when the full-length 5' UTR sequence is aligned against its genomic sequence. In fact, this method has been demonstrated to be comparable and, in many cases, more accurate than other approaches, for example, S1 nuclease protection, RNase protection, and primer extension (Liu and Gorovsky 1993; Schaefer 1995; Gum et al. 2003). The 5' RACE Nested PCR products were cloned into pCR4-TOPO vector (Invitrogen, CA), and positive clones were randomly isolated for sequencing. For zebrafish Cx44.1, 21 independent positive clones in total were sequenced, while for mouse Cx50, 16 independent positive clones were sequenced. Sequencing results were then aligned against the genomic DNA sequence, thus resulting in the identification of the transcriptional start site and the intron-exon structure of zebrafish Cx44.1 and mouse Cx50.

Using RNA samples from zebrafish adult lens tissues, 5' RACE experiments identified multiple transcriptional start sites (TSSs) of zebrafish Cx44.1. The sequence results are presented in the Figure 2-3. For convenience of description, the adenine site in ATG codon was assigned as position +1, with position -1 representing its 5' nucleotide and position +2 its 3' nucleotide. Since the core promoter is typically situated within about 100 bp surrounding the transcriptional start site (TSS), TSSs separated by more than 50 base pairs are considered to be controlled by distinct core promoters. Two distinct promoters (P1 and P2) were identified in the present study according to this criterion. As shown in Figure 2-3, each promoter can generate multiple TSSs. For P1,

transcription of Cx44.1 can be initiated from position -454 and -444. No TATA-like element was found in the vicinity of these TSSs. For P2, transcription can be initiated from nucleotide -560, -549, -536, -533, and -518, and a search for a TATA-box revealed a TATA-like element at nucleotide -590. As also shown in Figure 2-4, transcript I is transcribed from P1, whereas IIa and IIb are transcribed from P2. Figure 2-4 shows three different transcripts that arise from both the different promoter usage and pre-mRNA alternative splicing. Transcript I is transcribed from P1 while both IIa and IIb are from P2. Interestingly, both transcript I and IIa share the same splicing event which is distinct from that of transcript II. Pre-mRNA splicing follows the conserved "GU-AG" rule, and alternative splicing only occurs at the 5' splicing sites (or donor sites), as both are spliced to the same 3' site (or acceptor site) hundreds of base pairs downstream (nucleotide -16). Transcripts I and IIa use the 5' splicing site at nucleotide -391, generating an intron of 376 bp, whereas transcript IIb uses the 5' splicing site at position -479, generating a longer intron that is 464 bp in size (Figure 2-3 and connotation 2-4). As pre-mRNA splicing does not affect the coding region of Cx44.1, variable transcripts of Cx44.1 only differ in the 5'UTR. Therefore, alternative 5' UTRs resulting from different promoter usage and alternative pre-mRNA splicing is a significant feature of mRNA expression of zebrafish Cx44.1 in adult lens tissues.

In order to investigate whether the mouse ortholog Cx50 has similar features of mRNA expression, I also performed 5' RACE analyses using RNA samples from adult mouse lenses. As shown in Figure 2-5 (and connotation 2-6), multiple transcriptional

start sites were also identified, which were classified into three groups (I, II, and III), in terms of different promoter usage (P1, P2, and P3) and different splicing events. Among 16 transcripts identified in the 5' RACE experiment, twelve were produced from P1; three were transcribed from P2; and only one was detected that was generated from P3. This, to some degree, suggests that P1 is the dominant promoter for mouse Cx50 in lens (also see below, real-time RT-PCR analyses). As shown in the Figure 2-5, transcription driven by P1 may be initiated from alternate TSSs at positions -5586 and -5581, however no alternative TSSs were detected for transcripts driven by P2 and P3. Transcripts I and II share both 5' and 3' splicing sites which splice out an intron of 5,515 bp in size. Interestingly, P3 is situated within the intron, and its transcripts do not undergo pre-mRNA splicing. Similar to zebrafish Cx44.1, the generation of alternative 5' UTRs due to multiple promoter usage and different pre-mRNA splicing is also a significant feature of mRNA expression of mouse Cx50 in adult lens tissues.

2.3.2 5'UTRs of ZfCx44.1 and mCx50 are highly structured

Using mFOLD software (Zuker 2003), the secondary structure of 5' UTRs of zebrafish Cx44.1 and mouse Cx50 were analyzed (Figure 2-7, to 2-12). The most thermodynamically stable structures predicted by mFOLD are shown with calculated free energy indicated in the figures (e.g., ΔG°_{37} = -16.60 kcal/mol for Cx44.1-I). The alternative 5' UTRs of zebrafish Cx44.1 have significant variable length, and all transcripts have remarkably strong secondary structures, with stable stem-loops

throughout the UTRs.

A search for upstream AUG codons (uAUGs) revealed several uAUGs within the 5'UTR-III transcript of mCx50, but no uAUGs were found for all other transcripts of both Cx44.1 and Cx50. All of these uAUGs within mCx50 5'UTR-III are followed by in-frame stop codons prior to the normal AUG codon, resulting in a couple of uORFs (upstream open reading frame) of variable length (Figure 2-13, red arrows). It has been suggested that upstream AUGs can inhibit cap-dependent translation (Morris and Geballe 2000; Iacono et al. 2005). The existence of uAUGs might provide a mechanism that allows fine control of translational initiation from the correct AUG codon.

2.3.3 Expression Patterns of ZfCx44.1 and mCx50 Transcript Variants

The expression pattern of zebrafish Cx44.1 mRNA was previously characterized in this lab, using Northern blot analyses and *in situ* hybridization techniques, on embryos at various developmental stages. The zebrafish Cx44.1 gene is expressed in the lens of the developing embryo, and its expression begins during the early stages of lens morphogenesis (Cheng et al. 2003; Cason et al. 2001). In the present study, using real-time RT-PCR, we further studied tissue distribution of Cx44.1 mRNA expression in zebrafish adult tissues (Figure 2-14). RNA samples were isolated from zebrafish adult tissues including lens, retina, heart, liver, ovary, testis, spleen, gill, gut, brain, and swim bladder. No DNase pre-treatments were performed in order to achieve as much as possible of mRNA integrity. The zebrafish EF1α gene was selected as a reference gene in

these real-time PCR experiments, since EF1α has been demonstrated as the most suitable reference gene for zebrafish real-time PCR analyses, both in time-course and tissue based experiments (Tang et al. 2007). One of the EF1α-specific primers, zEF1α/q/F1, was designed to span the junction of exon 1 and exon 2, while zEF1α/q/R1 is situated in exon 2. This primer pair produces an amplicon of 94 bp in size (Table 2-2). To measure the combined level of expression of the three alternative Cx44.1 transcripts (two derived from the P1 promoter and two from the P2 promoter), a set of three primers were designed and used together in the real-time PCR reactions. The reverse PCR primer, Cx441/q/R1, is situated in the common coding region of all transcripts. Two forward primers, Cx441/q/1/F1 (which anneals to transcript I and IIa) and Cx441/q/2b/F1 (which anneals to transcript IIb), were each designed to span one of the two alternate 5' splicing sites. The design of these two primer pairs not only ensured that all types of transcripts of Cx44.1 were being properly targeted, but it also made sure that only cDNA templates were amplified during PCR. The elongation time for real-time PCR reaction was very short (10 sec.), which also significantly suppressed the amplification of genomic templates that contains relatively long intron fragments. After PCR amplification, melting curve analysis was employed to monitor for genomic DNA contamination. In addition, a reverse transcriptase negative control and a template-negative control were also set up in parallel for each batch of PCR experiment. It has been demonstrated that this idea of primer design plus proper experimental controls can efficiently eliminate genomic DNA contamination when doing RT-PCR without DNase pre-treatments.

As shown in Figure 2-14 (Tissue distribution of zebrafish Cx44.1 total mRNAs), zebrafish Cx44.1 is primarily expressed in the adult lens tissues. In the testis, we also detected a relatively low level of Cx44.1 expression. The present study has provided further supporting evidence that zebrafish Cx44.1 is a lens-specific connexin gene that is expressed (and functioning) during the lens development as well as in the functional maintenance of adult lens tissues. Its expression in the testis has not been reported previously in the literature. At this stage, it is not possible to claim conclusively that Cx44.1 has an essential role in the testis, but detection of its transcripts in this tissue should warrant further analyses.

We also performed real time RT-PCR to characterize the expression pattern of the alternative transcripts of Cx44.1 in the adult lens. As mentioned, the primers Cx441/q/1/F1 and Cx441/q/2b/F1 were designed to specifically target the two alternative 5' splicing sites. It is worth noting that the Cx441/q/1/F1 primer actually targets both transcript I and IIa (Figure 2-4). To address this problem, we designed an additional primer pair Cx441/q/2a/F1 and Cx441/q/2a/R1, which specifically targets transcript IIa. The results regarding transcript-I can therefore be corrected by calculation. Since this additional primer pair produces an amplicon of 128 bp in size, we further designed zEF1α/q/R2 for the reference, with zEF1α/q/F1 together, which can produce a 133 bp PCR fragment. The real-time RT-PCR analyses show that the three types of transcripts of Cx44.1 are expressed at a very similar abundance, about 4% of EF1α transcripts, in the lens (Figure 2-15).

Since all three types of transcript variants were detected in adult zebrafish lens tissues, it was of interest to determine whether they are also expressed and functioning during embryonic development. To test this, I also performed real-time RT-PCR analyses for different transcript variants at early developmental stages. Results show that all three transcript variants are indeed expressed during zebrafish early embryonic development, although the relative abundance of Cx44.1 transcripts in the embryo is very low relative to EF1α expression levels (Figure 2-16). In contrast, expression of these Cx44.1 transcripts in adult lenses is expressed with a much higher ratio to that of EF1α transcripts. It also appears that all three types of transcripts are generally expressed at a similar abundance at specific stages during the development. The minor differences in expression among transcripts I, IIa, and IIb could simply be the result of experimental variation. Further analyses would need to be performed to assess whether these changes in expression are biologically relevant. Expression of Cx44.1 was not detected at 0.5 dpf, but relatively strong expression was detected at 1 dpf. This observation is in agreement with previous studies, where expression of Cx44.1 was detected from 1 dpf and later on (Cheng et al. 2003; Cason et al. 2001). Therefore, it is speculated that expression of Cx44.1 could be initiated somewhere between 0.5 and 1 dpf, upon the time of zebrafish eye development.

We also performed expression analyses for alternative transcripts of mCx50 in the mouse adult lens. As shown in Figure 2-6, all three primer pairs were designed to specifically target three different transcript variants respectively (also see Table 2-2). The

relative amounts of the transcript variants are presented in Figure 2-17. It appears that transcript II and III, with a similar expression level, are both the minority among all transcripts, while transcript I is much more strongly expressed in the adult lens tissues.

2.3.4 Promoter Analyses of Zebrafish Cx44.1 and Mouse Cx50

In order to investigate whether the zebrafish Cx44.1 and mouse Cx50 DNA share similar molecular mechanisms for their transcriptional regulation, we conducted a sequence similarity analysis between the Cx44.1 and Cx50 promoters, but no significant sequence conservation was found. Using MatInspector web-based software (Cartharius et al. 2005; Quandt et al. 1995), a search for potential transcription factor binding sites within the proximal promoters of Cx44.1 and Cx50 was then performed. This analysis revealed a couple of transcriptional regulatory motifs when using the eye as the tissue filter for transcription factors. As shown in Figure 2-18, we observed multiple sites for several conserved transcription factors, for example, AP1R (Maf and AP1 related factors), BRNF (Brn POU domain factors), FKHD (Fork head domain factors), HOXF (Homeobox factors), NR2F (Nuclear receptor factors), PAX6 (PAX-4/PAX-6 paired domain binding factors), SORY (SOX/SRY-sex/testis determining and related HMG box factors). These transcription factors are all involved in embryonic development, and essential for the development of the eye. For example, PAX6 has been established to be the "master control" of eye development (Gehring 1996; Cvekl and Piatigorsky 1996; Gehring and Ikeo 1999; Wawersik et al. 2000; Chow and Lang 2001; Tsonis and Fuentes

2006; Kozmik 2008). SOX proteins are regulators for crystalline expression in the lens (Chow and Lang 2001), although interestingly, SOX proteins are members of the SORY family of factors, which also regulate testis formation, and this motif may help explain why some Cx44.1 expression was observed in the zebrafish testis (Figure 2-14). Transcription factors in BRNF family are highly restricted in central nervous system neurons, and within the retina, these factors are present only within ganglion cells. It was demonstrated that BRNF transcription factors are required for development of retina ganglion cells (Gan et al. 1996).

In addition, we also found one INRE site (core promoter initiator element) situated between Cx44.1_P1 and P2, as well as Cx50_P1 and P2, respectively. It is noted that P1 and P2 are alternative core promoters for Cx44.1 and Cx50 respectively. A further search for a TATA box revealed one TATA-like element in Cx44.1_P2 and Cx50_P2 respectively, but no TATA-like element was found for Cx44.1_P1 and Cx50_P1. This structural arrangement of the INRE and TATA box, probably together with other *cis*-elements or *trans*-factors, might provide the switch for the selection of core promoters of both Cx44.1 and Cx50.

While software predictions can identify potential transcription factor binding sites *in vitro*, these are not enough to discriminate between functional and non-functional sites in the complex cellular environment (Cartharius et al. 2005; Wasserman and Sandelin 2004). Nevertheless, these common binding sites predicted by MatInspector in promoters of Cx44.1 and Cx50 suggest that the regulatory mechanism for transcription may be

conserved between fish and mouse. In order to test this, promoter activities of zebrafish Cx44.1 and mouse Cx50 were tested by an EGFP transient expression construct injected into zebrafish embryos. Expression constructs contained the proximal promoters of either (Cx44.1 or Cx50) connexin gene. For the analyses of Cx44.1, we tested ~1 kb genomic DNA immediately upstream of the translational start codon ATG. As this Cx44.1-EGFP recombinant construct carries the intact 5'UTR and intron sequences, *in vivo* transcription and pre-mRNA splicing of this expression construct in injected embryos will produce the same leading 5'UTR sequence as endogenous Cx44.1 transcripts. Figure 2-19 (A) shows that these Cx44.1 promoter fragments could drive EGFP transient expression in the retina of fish embryos. For mouse Cx50, the P3 is situated within the intron region, and two promoter constructs were therefore designed. Promoter constructs for both P1 and P2 contain approximately 300 bp of the promoter sequence upstream of the TSS of P2, along with the entire 5' UTR (i.e., including the first exon but excluding the intron), fused in frame to EGFP coding sequence (Figure 2-2, E). The promoter construct for P3 carries ~ 1 kb fragment that is located within the intron, immediately upstream of the translational initiation codon ATG (Figure 2-2, F). This design ensures that effects of P3 can be studied separately from that of P1 and P2. Figure 2-19 (C and D) shows that both mouse Cx50 promoters could also drive EGFP expression in the zebrafish embryos in patterns similar to that driven by the Cx44.1 promoter.

2.4 Discussion

In this study, multiple transcriptional start sites and alternative splicing in the 5'UTRs of both zebrafish Cx44.1 and mouse Cx50 have been identified using 5'RACE techniques. It is notable that different promoter usage and alternative splicing result in transcript variants that differ only in their 5'UTR. In general, 5'UTRs are important in the post-transcriptional regulation of gene expression, including mRNA transportation out of the nucleus, sub-cellular localization, mRNA stability, and efficiency of translation (Vilela et al. 1999; Wang et al. 1999; Kindler et al. 2005; Gebauer and Hentze 2004; Pickering and Willis 2005; Mignone et al. 2002). It is probable that alternative 5'UTRs of these connexin genes allow their expression and function to be finely modulated by post-transcriptional regulatory mechanisms.

Most cellular mRNAs use this cap-dependent mechanism for protein synthesis, and the 5'UTRs are essential for the initiation of mRNA translation. Most eukaryotic mRNAs contain short 5'UTRs (usually fewer than 50 nts), and thus, cap-mediated translation of such mRNAs is efficient enough. A longer 5' UTR usually increases efficiency of translation by facilitating to recruit more pre-initiation complexes, but further increases in 5' UTR length usually down-regulate the translational efficiency due to secondary structure or alternative initiation of translation at upstream AUG codons (Kozak 1991). About 10% of all mRNAs contain atypically long 5' UTRs, and their protein products typically regulate vital cellular processes, including cell growth, cell proliferation, apoptosis, etc. A long 5' UTR gives the propensity to form stable secondary structure that may inhibit translation initiation by impeding the ribosomal scanning

process (Pickering and Willis 2005a). In situations where cap-dependent translation is impaired or suppressed, an alternative mechanism can be used by such long 5'UTRs, as translational initiation can be mediated by a secondary structure called the internal ribosome entry site (IRES) (Kaminski et al. 1994; Jackson 2005). Recent estimates suggest that up to 10% of all cellular mRNAs have the potential to initiate translation via this mechanism (Pickering and Willis 2005).

The prediction of secondary structure of the 5'UTR showed that the 5'UTR variants of ZfCx44.1 and mCx50 are highly structured. It's also notable that the secondary structures show high diversity among different transcript variants. These features of the 5'UTR secondary structure make it possible for the expression of different transcripts to be regulated with distinctive pathways. We further performed real-time RT-PCR analyses for these transcript variants of ZfCx44.1 and mCx50 respectively. Although there is no linear relationship between mRNA abundance and yield of protein products, the relative amount of transcripts can be an important factor for the regulation of translation. This study's analyses suggest that expression of ZfCx44.1 and mCx50 might be tightly regulated at the post-transcriptional level. Complicated diversities of 5'UTRs could allow their expression to be finely modulated during the embryonic development or responding to physiological stimulus in adult tissues.

Expression pattern analyses have demonstrated that zebrafish Cx44.1 is predominantly expressed in the ocular lens, which is consistent with previous studies (Cheng et al. 2003; Cason et al. 2001). It is known that mouse Cx50 (ortholog of

zebrafish Cx44.1) is also predominantly expressed in the lens. The tissue expression patterns of both connexin genes appear to be highly conserved. Mice knockout studies have demonstrated that Cx50 plays essential roles in lens development and adult lens tissue homeostasis, as the ablation of Cx50 in knockout mouse resulted in the formation of cataracts (White et al. 1998). Knockout techniques have not yet been available for the zebrafish model, however, the use of morpholino knock-down techniques is ideally suited for genetic studies in injected zebrafish embryos. Previous morpholino studies of zebrafish Cx44.1 in this lab have observed reduced lens growth and the formation of cataracts during early embryonic development, even though the morpholino oligos were not perfect (Cheng and Valdimarsson., unpublished data). Nevertheless, these functional analyses certainly suggested that lens connexins play essential roles during lens development and the knock-down of connexin gene expression can lead to lenticular diseases such as cataracts. With the identification of the 5'UTR sequences of zebrafish Cx44.1, we can now design effective morpholino oligos, thus allowing genetic analyses of the roles that Cx44.1 plays during the lens development, as well as molecular and genetic mechanisms for lenticular diseases.

The expression of alternative transcripts of Cx44.1 raised an interesting question for our future studies – what are specific functional roles of each individual transcript variant in zebrafish lens development and the homeostasis of adult lens tissues, or why different transcript variants are needed. For this purpose, morpholino oligos have been designed to target each specific individual transcript variant of ZfCx44.1 that has been

derived from differential pre-mRNA alternative splicing. In our future studies, the use of morpholino knock-down techniques could provide a powerful strategy for elucidating the specific functional roles of alternative splicing of ZfCx44.1 during lens development in zebrafish embryos. However, all three transcript variants of ZfCx44.1 encode the same connexin protein, which brings a big challenge for us to determine their functional roles individually. To this end, more detailed expression patterns of these transcript variants are necessary in order to link splicing variants to potential phenotypes. In this study, I performed real-time RT-PCR to analyze the relative abundance of these transcript variants during zebrafish embryonic development as well as in adult lens tissues. These results can be very useful when functional studies for transcript variants of ZfCx44.1 are performed in our future studies.

In this study, preliminary promoter analyses were also performed with Cx44.1 and Cx50. The prediction of potential transcription factor binding sites in the proximal promoters of both genes suggested that common transcription factor networks might be conserved. We have used EGFP transient assays to test this hypothesis. Recombinant constructs were made by PCR method to fuse ~ 1 kb promoter fragments in frame to the EGFP coding sequence. This method ensures that the transcripts carry the same 5' UTR sequences as the endogenous genes. Despite examining about 50 injected embryos, it was surprising that GFP expression was not observed in the lens using either connexin promoter. Although more analyses could be needed, we suggest that some key regulatory elements might have been missed in these promoter constructs, such as, for example,

potential distal enhancers that are essential for lens-specific expression. In addition, some silencing elements might have also been missed in these promoter constructs, given the fact that EGFP was strongly expressed in the retina in injected embryos. To test these speculations, another Cx44.1 promoter construct was built using the PCR strategy, and it carries the entire 3'UTR fragment of Cx44.1 replacing the SV40 poly-A sequence in the previously used construct (Figure 2-1, D). Upon injecting this construct, about half of the injected embryos could express EGFP in the lens, though EGFP expression in the retina was still observed (Figure 2-19, B). Although the mechanisms seem to be unknown at this stage, this observation suggested that the 3'UTR of Cx44.1 probably has important regulatory roles for its expression, such as a lens-specific enhancer sequence.

The promoter constructs used in the present study can not perfectly mimic the endogenous expression patterns of ZfCx44.1 and mCx50, but nevertheless, the analyses in this study suggested that Cx44.1 and Cx50 share some common regulatory elements (e.g., transcription factor binding sites) in their proximal promoters. In general, the mechanisms of eye morphogenesis are evolutionally conserved throughout the animal kingdom, and there exist similar genetic networks of eye transcription factors. Also, similar inductive signal systems may be used to initiate the self-regulatory network of eye specification (Chow and Lang 2001; Ogino and Yasuda 2000; Zuber et al. 2003). The expression of connexin proteins is established early in the process of eye development, and connexins play essential roles in the coordination of eye development through signal conduction and exchange of substrates. Although the detailed mechanisms are still

unknown, analyses in the present study suggest that the regulatory mechanism for the initiation of connexin gene transcription can be conserved between zebrafish and mouse.

It is possible that endogenous expression patterns could be more perfectly mimicked if using much larger promoter fragments. Researchers have proposed to use artificial chromosomal vectors (BAC or PAC) that can accommodate much larger DNA insertions than regular plasmids. This method is helpful to reproduce reliable expression patterns in injected embryos since larger fragments are less likely to be missing crucial regulatory elements (Jessen et al. 1998). In our future studies, further analyses are needed in order to identify distal enhancers and silencers, thus allowing further investigation into the regulatory conservations between zebrafish and mouse.

2.5 Summary

In this chapter, 5' RACE experiments have demonstrated that alternative pre-mRNA splicing and different promoter usage are common features of zebrafish Cx44.1 and mouse Cx50 gene expression, and transcript variants of Cx44.1 and Cx50 only differ in their 5'UTRs. These results suggest that expression of both genes is tightly regulated at both the transcriptional and the post-transcriptional level. Promoter analyses for Cx44.1 and Cx50 were performed through both *in silico* prediction of transcription factor binding sites and *in vivo* GFP transient expression assays in zebrafish embryos. Preliminary comparisons between proximal promoters of Cx44.1 and Cx50 suggested that both connexin genes could share conserved transcriptional regulatory networks. The

expression patterns of transcript variants of Cx44.1 were examined using real-time PCR and in most cases, each transcript variant was expressed equally. With the identification of different transcript isoforms of zebrafish Cx44.1, the use of morpholinos that specifically target splicing variants Cx44.1 could provide a powerful strategy to investigate the functional roles of these alternative variants during embryonic development.

Figure 2-1. Zebrafish Cx44.1-eGFP expression constructs. Gray shadow boxes represent the 5'UTR sequences of Cx44.1 cDNA. A ~1.1 kb genomic DNA fragment immediately upstream of the translation start codon was amplified by PCR (**A**), The EGFP coding sequence plus SV40 poly-A tail were retrieved by PCR from an EGFP expression vector (**B**), which was then linked to the EGFP coding sequence by PCR, resulting in Cx44.1-eGFP recombinant DNA (**C**). An additional construct was made by replacing the SV40 poly-A tail with Cx44.1 endogenous 3'UTR tail (**D**).

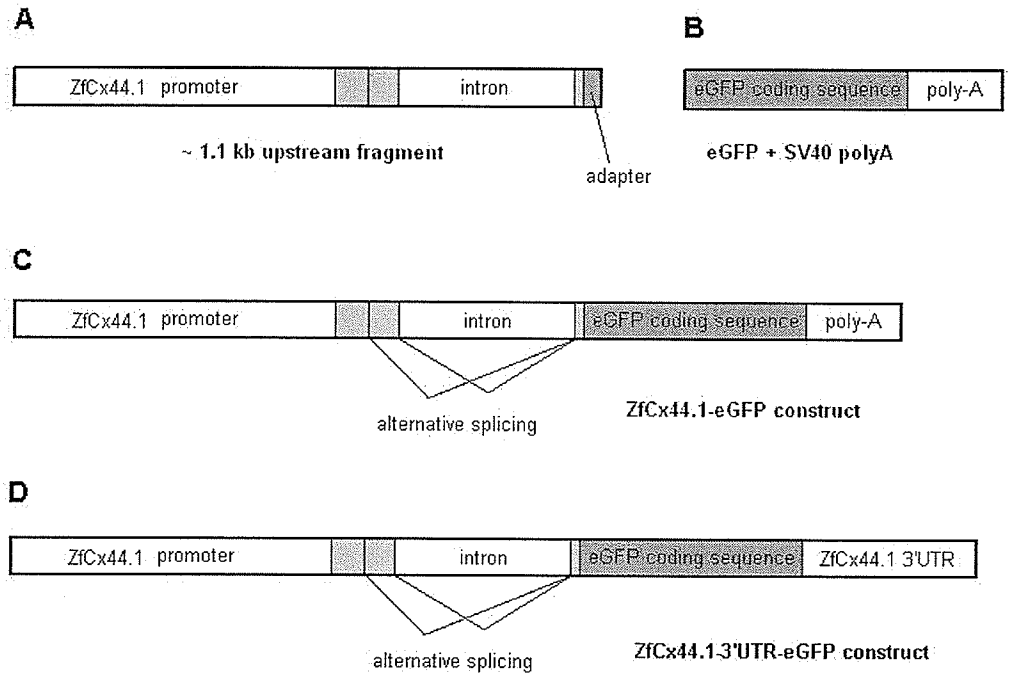


Figure 2-2. Mouse Cx50-eGFP expression constructs. Gray shadow boxes represent the 5'UTR sequences of Cx50 cDNA. Fragment (**A**) is the genomic region immediately upstream the translational start site. Fragment (**B**) contains ~ 300 bp promoter fragment (P2), the 5'UTR sequence of Cx50, and an adapter for eGFP coding sequence. Fragment (**C**) was amplified by PCR from genomic sequence which carries the promoter (P3) and an eGFP adapter. Both promoters (P2 and P3) were fused with EGFP and SV40 poly-A tail (**D**) by PCR strategy. The final PCR products (**E**, and **F**) are Cx50-eGFP recombinant construct used for transient expression assays after DNA purification.

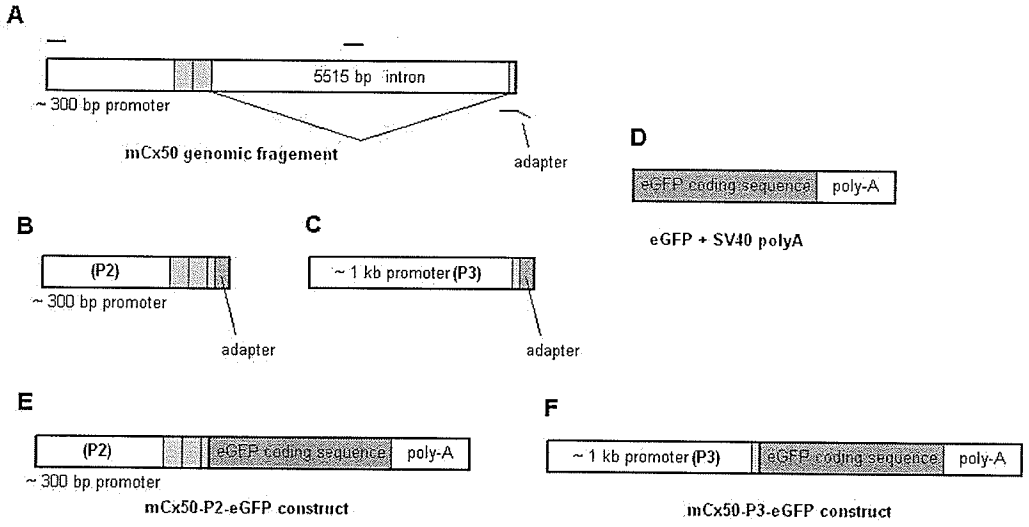


Figure 2-3. Alignment of zebrafish Cx44.1 transcript variants. Three transcript variants are derived from different promoter usage and alternative pre-mRNA splicing. Transcript I is transcribed from P1, whereas both IIa and IIb are from P2. Transcription can start from multiple transcriptional start sites (TSSs, indicated by * above the specific nucleotides). Transcript I and IIa share the same splicing event, but differ from that of transcript IIb. Splicing events follow the conserved "GU-AG" rule. The adenine site in ATG codon was assigned as position +1.

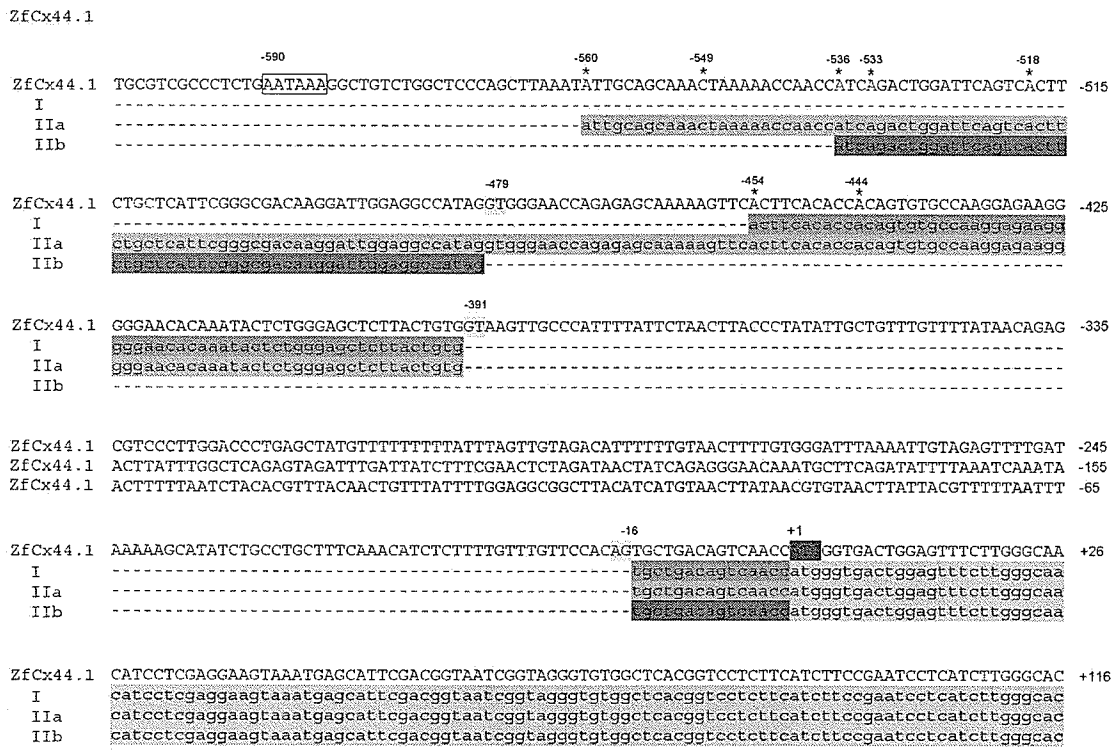


Figure 2-4. Sketch of ZfCx44.1 transcript variants. Boxes represent the exons of Cx44.1 transcripts, while bold lines (or blank boxes) are genomic sequences or introns. All three transcripts share the same coding sequence and 3'UTR, and alternative 5'UTRs are derived from different promoter usage (P1 or P2) and pre-mRNA alternative splicing (alternative 5' sites). Real-time PCR primer pairs were designed to target specific transcript variants of Cx44.1. Forward primers are indicated by red arrows, while reverse primers by blue arrows. All primer pairs have at least one primer spanning an exon-exon boundary.

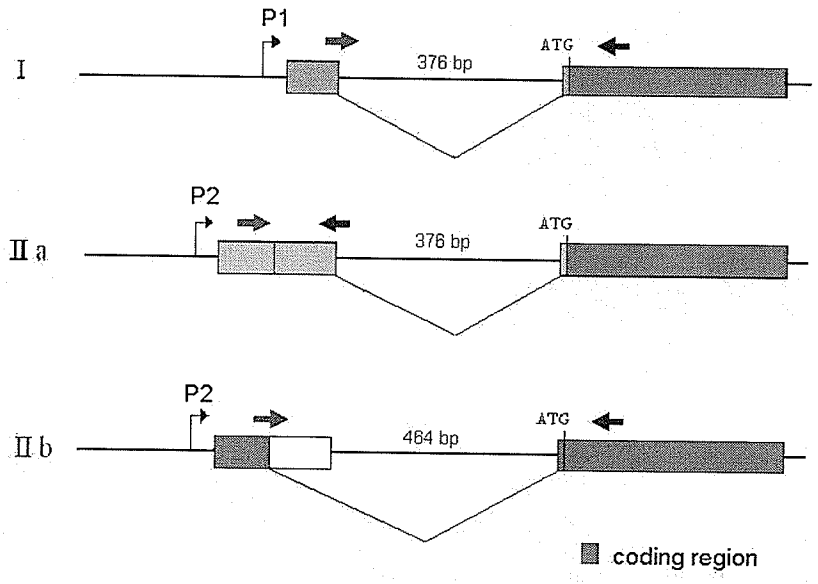


Figure 2-5. Alignment of mCx50 transcript variants. Three transcript variants are derived from different promoter usage (P1, P2, or P3) and pre-mRNA splicing. Transcription from P2 can start from multiple transcriptional start sites (TSSs), which are indicated by * above the alignment. Splicing events follow the conserved "GU-AG" rule (highlighted with yellow color). Transcript I and II share the same splicing event, whereas transcript III does not undergo pre-mRNA splicing and its promoter and 5'UTR are situated in the intron region.

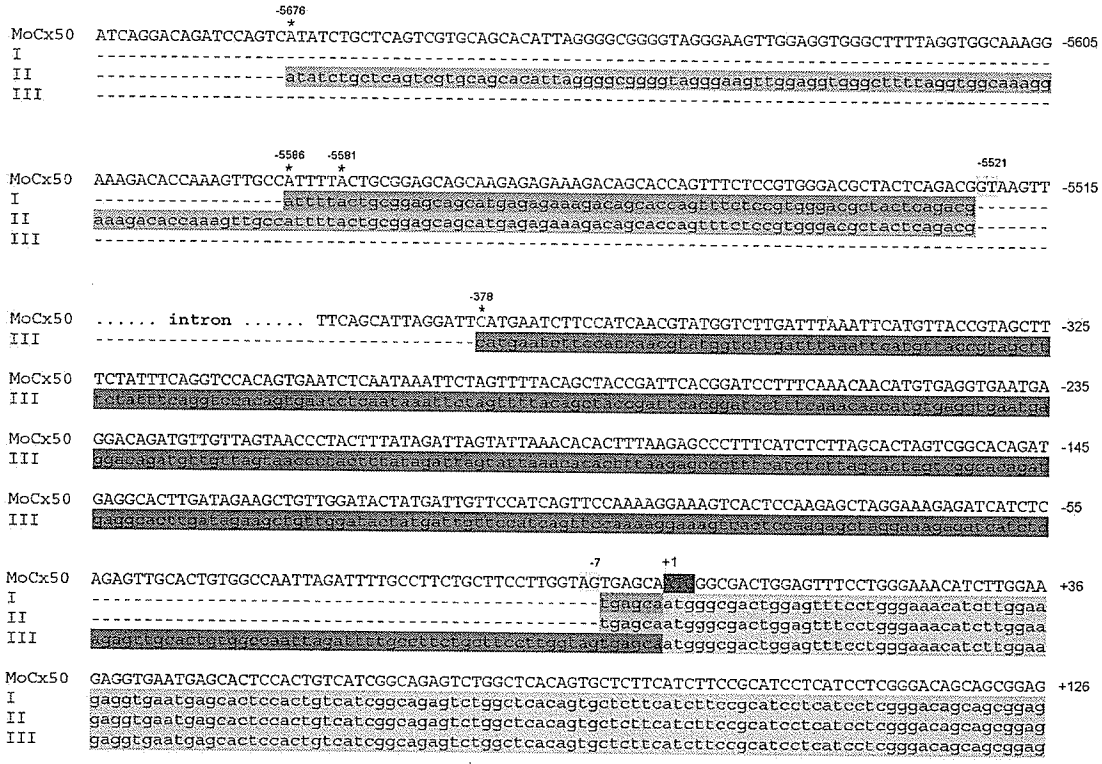


Figure 2-6. Sketch of mCx50 transcript variants. Boxes represent the exons of Cx50 transcripts, while bold lines (or blank boxes) are genomic sequences or introns. All three transcripts (I, II, and III) share the same coding sequence and 3'UTR, and alternative 5'UTRs are derived from different promoter usage (P1, P2, or P3) and pre-mRNA splicing (either undergoing splicing or not). Real-time PCR primer pairs were designed for targeting specific transcript variants of Cx50. Forward primers are indicated by red arrows, while reverse primers by blue arrows. All primer pairs have at least one primer spanning an exon-exon boundary.

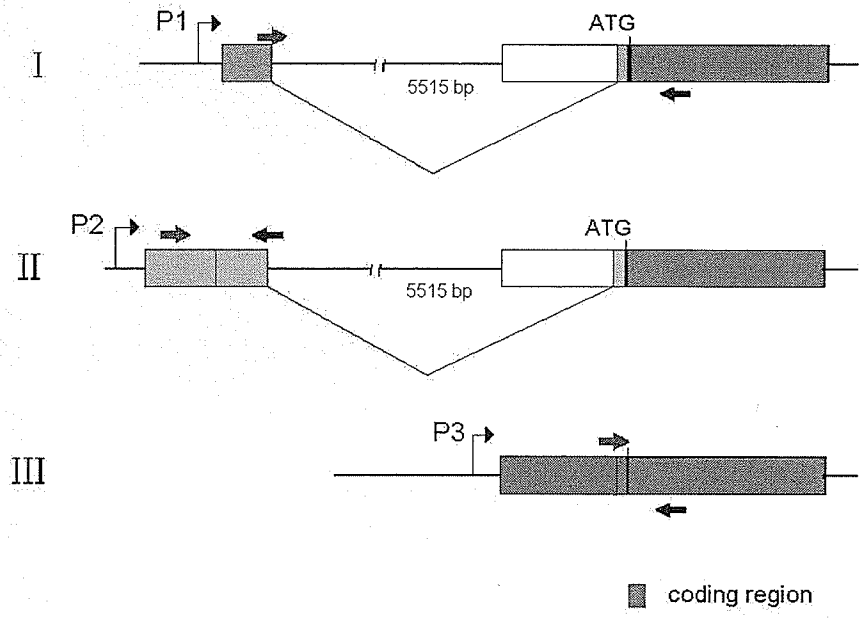


Figure 2-7. Secondary structure of 5' UTR of ZfCx44.1 transcript I. The most thermodynamically stable structures predicted by mFOLD program are shown with calculated free energy ($\Delta G°_{37}$ = -16.60 kcal/mol).

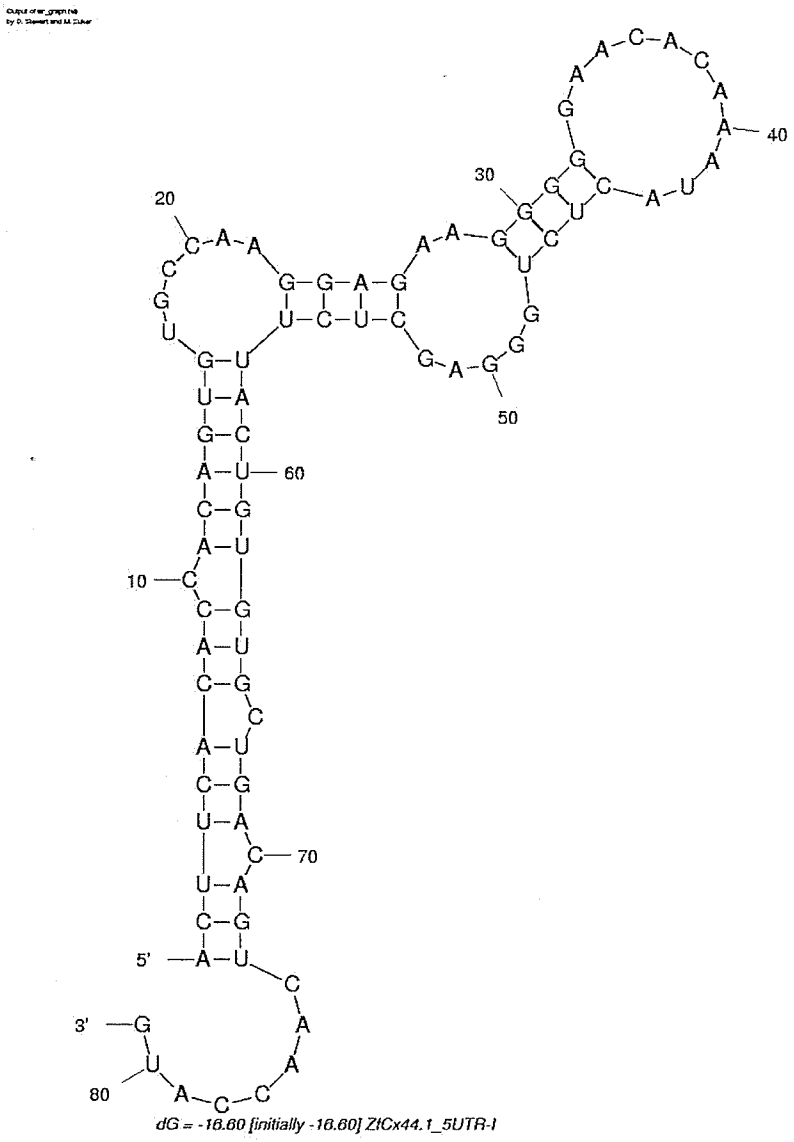


Figure 2-8. Secondary structure of 5' UTR of ZfCx44.1 transcript IIa. The most thermodynamically stable structures predicted by mFOLD program are shown with calculated free energy ($\Delta G°_{37}$ = -44.90 kcal/mol).

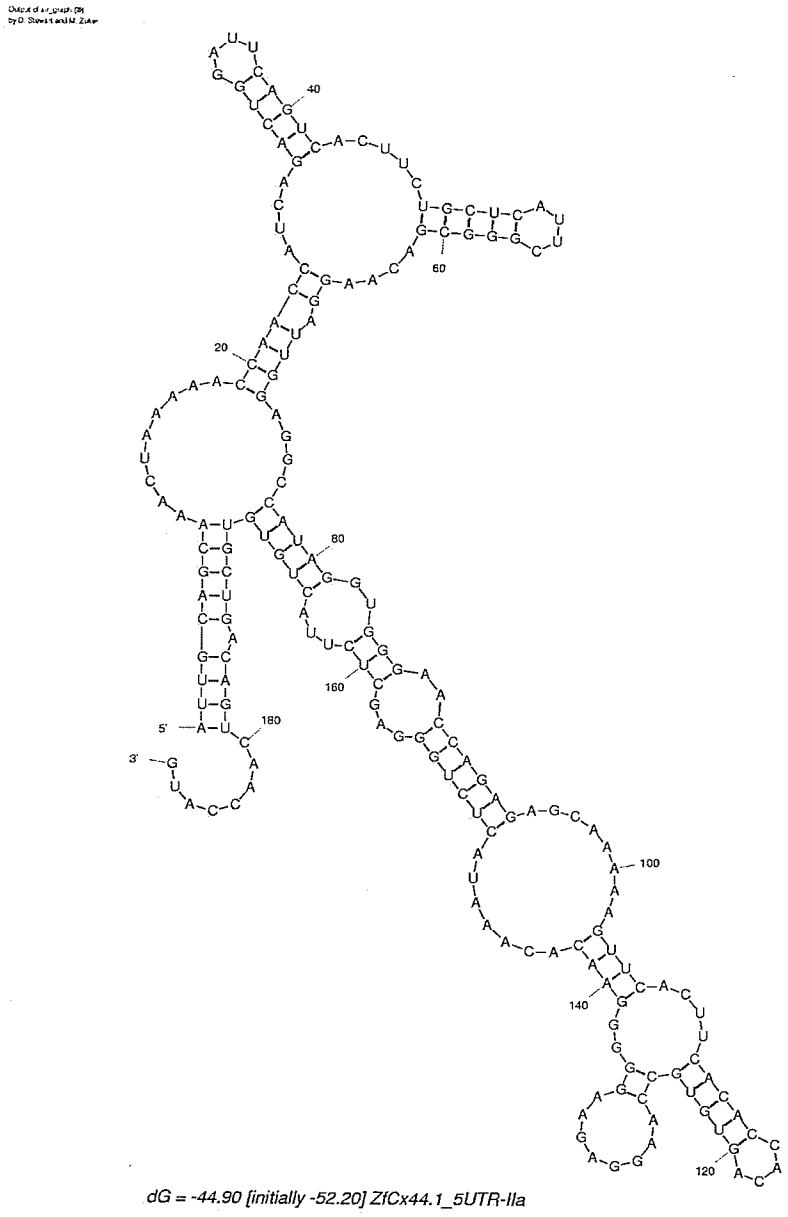


Figure 2-9. Secondary structure of 5' UTR of ZfCx44.1 transcript IIb. The most thermodynamically stable structures predicted by mFOLD program are shown with calculated free energy ($\Delta G°_{37}$ = -17.60 kcal/mol).

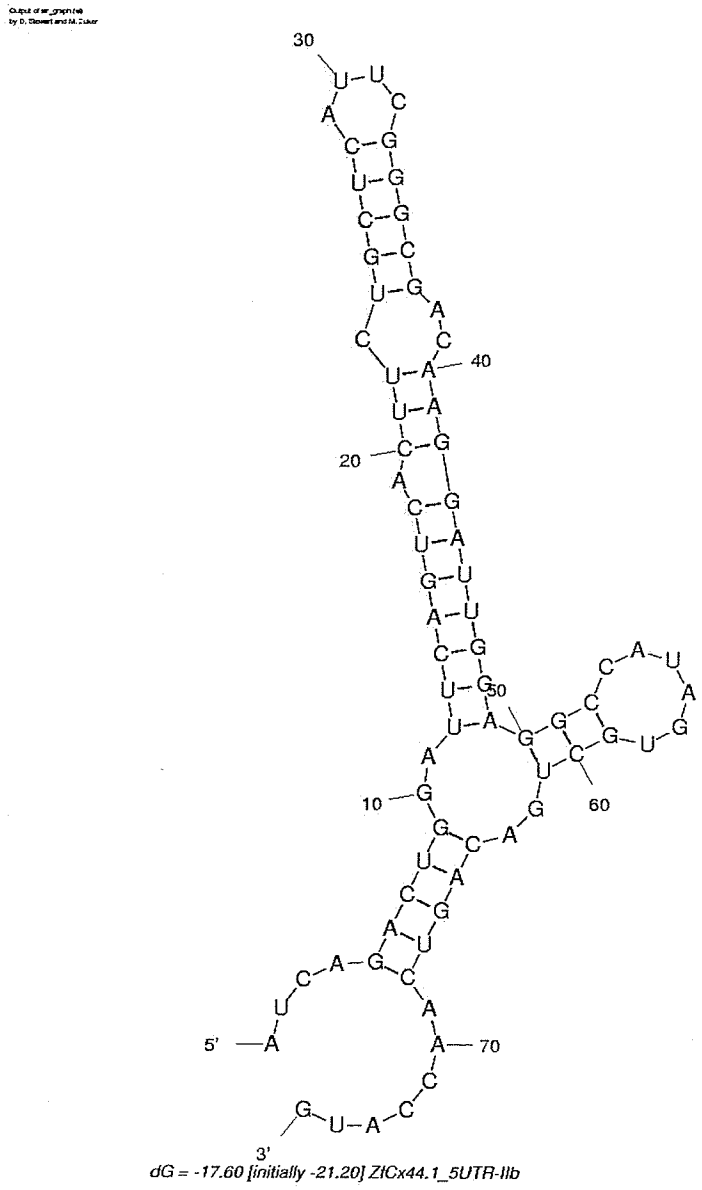


Figure 2-10. Secondary structure of 5' UTR of mCx50 transcript I. The most thermodynamically stable structures predicted by mFOLD program are shown with calculated free energy ($\Delta G°_{37}$ = -15.90 kcal/mol).

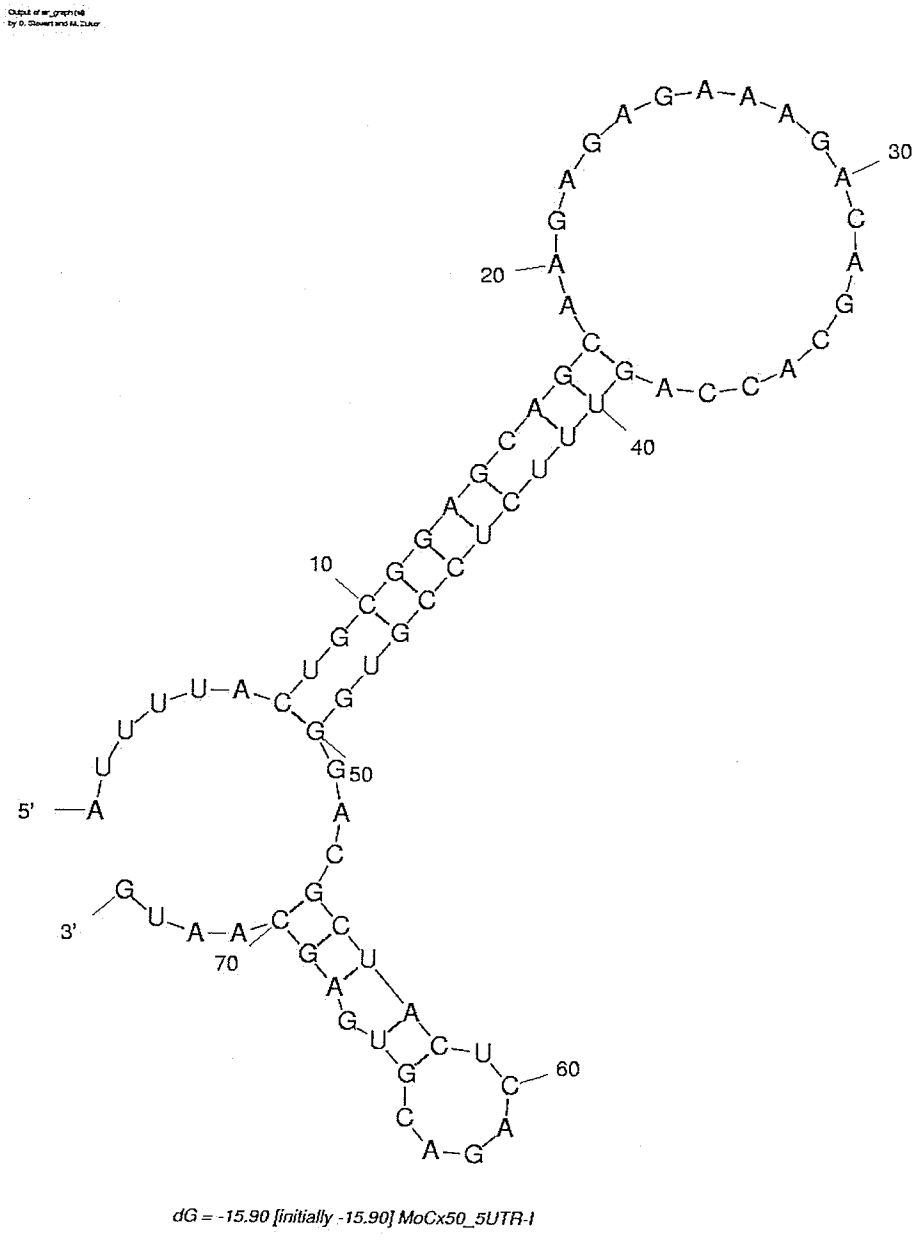


Figure 2-11. Secondary structure of 5' UTR of mCx50 transcript II. The most thermodynamically stable structures predicted by mFOLD program are shown with calculated free energy ($\Delta G°_{37}$ = -38.90 kcal/mol).

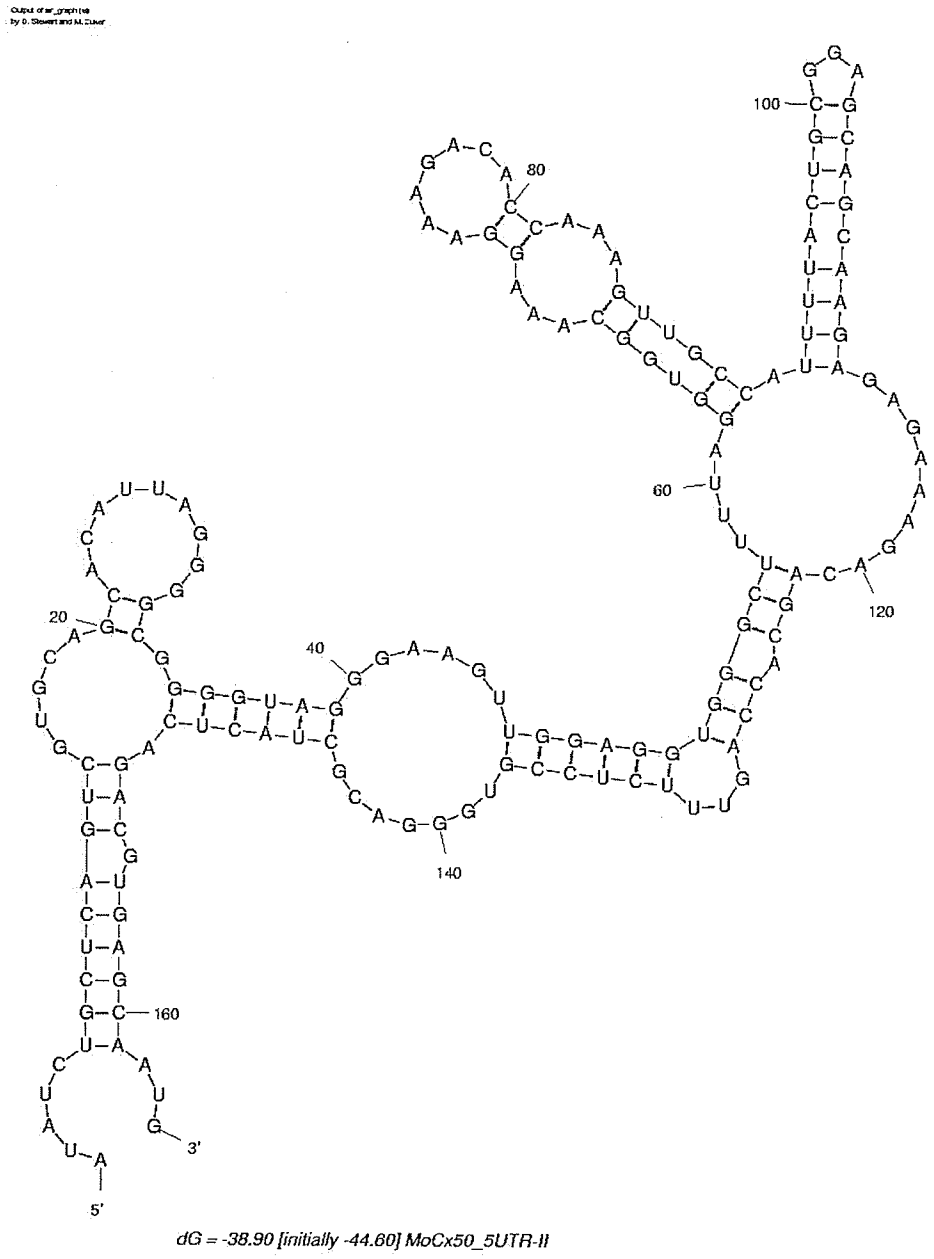


Figure 2-12. Secondary structure of 5' UTR of mCx50 transcript III. The most thermodynamically stable structures predicted by mFOLD program are shown with calculated free energy ($\Delta G°_{37}$ = -80.01 kcal/mol).

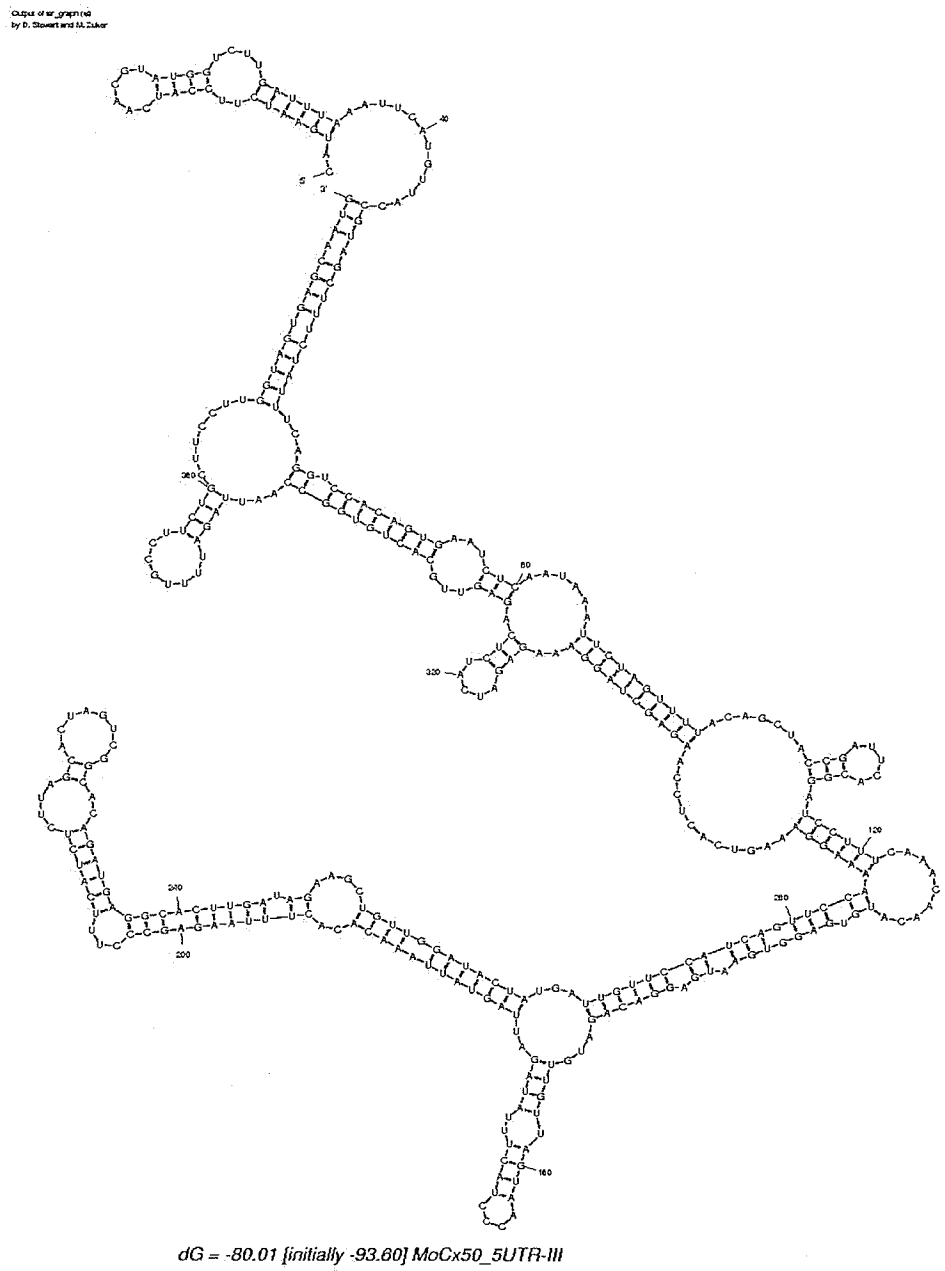


Figure 2-13. uORFs of mCx50 transcript III. The normal translational start codon is indicated by the green triangle (Frame +1). All of the uAUGs of Cx50 5'UTR-III are followed by in-frame stop codons prior to the normal AUG codon, resulting in a couple of uORFs of variable length (indicated by red arrows).

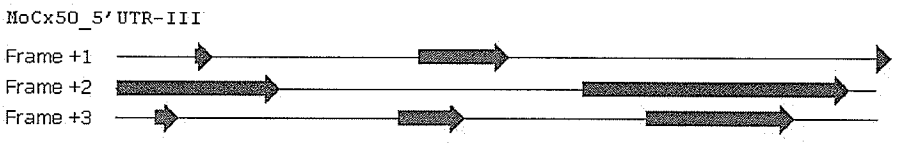


Figure 2-14. Tissue distribution of ZfCx44.1 total mRNAs. Cx44.1 mRNA is primarily expressed in the adult lens tissues, while a relatively low level of Cx44.1 expression was also detected in the testis. The ratios were calculated by calculating ΔCt between mean Ct values of test and EF1α genes in two replicate experiments, and were presented as a column graph.

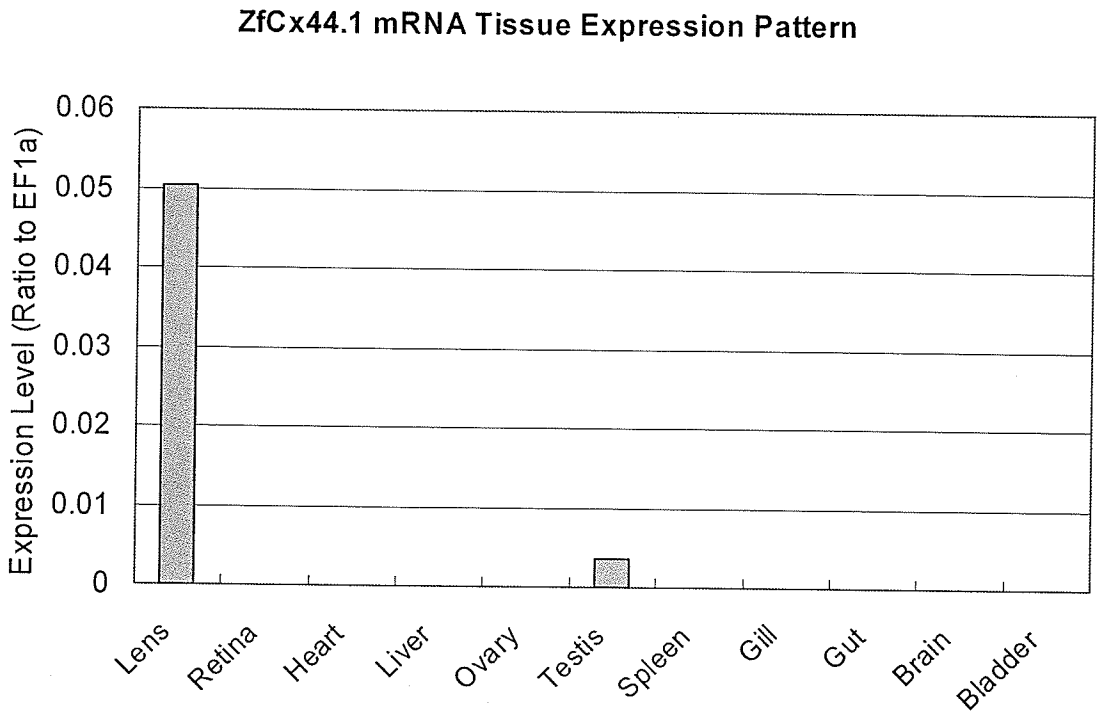


Figure 2-15. Expression of ZfCx44.1 transcripts in the adult lenses. Three transcript variants have similar expression level (about 4% of EF1α transcripts). The ratios were calculated by calculating ΔCt between mean Ct values of test and EF1α genes in two replicate experiments, and the results were presented as a column graph.

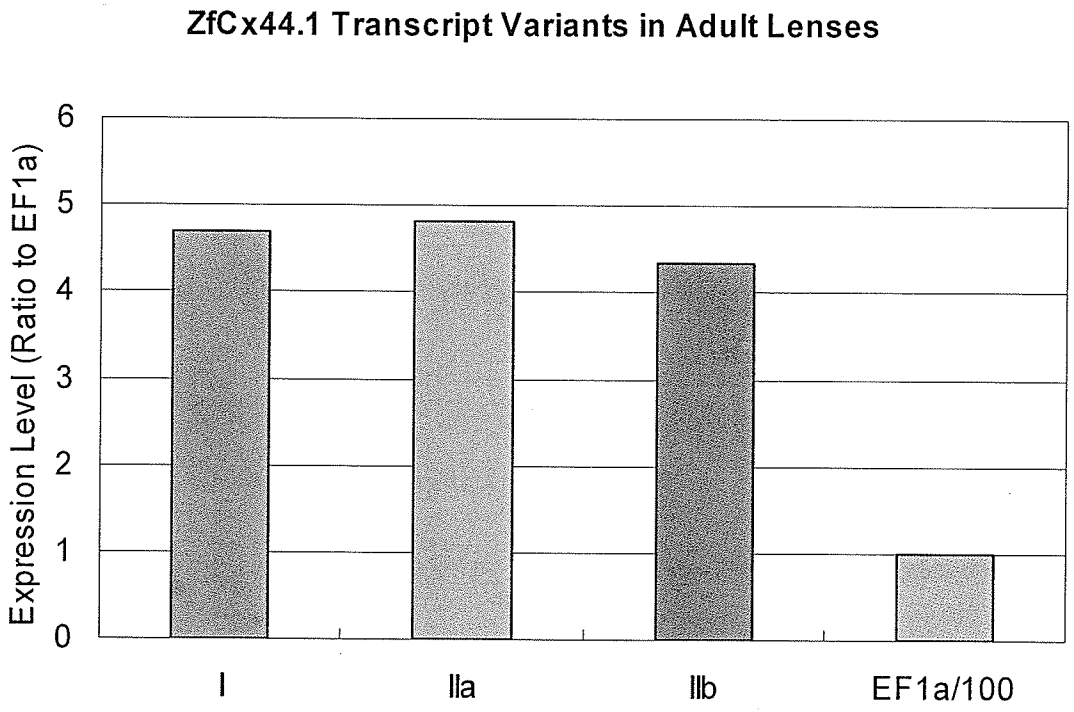


Figure 2-16. Expression of ZfCx44.1 transcripts during development. Analyses were performed using about 50 embryos at different developmental stages indicated by dpf (days post fertilization). The ratios were calculated by calculating ΔCt between mean Ct values of test and EF1α genes in two replicate experiments, and the results were presented as a column graph. Expression of Cx44.1 was not detected at 0.5 dpf, but relatively strong expression was detected at 1 dpf. All three types of transcripts are expressed at a similar abundance at specific stages during embryonic development.

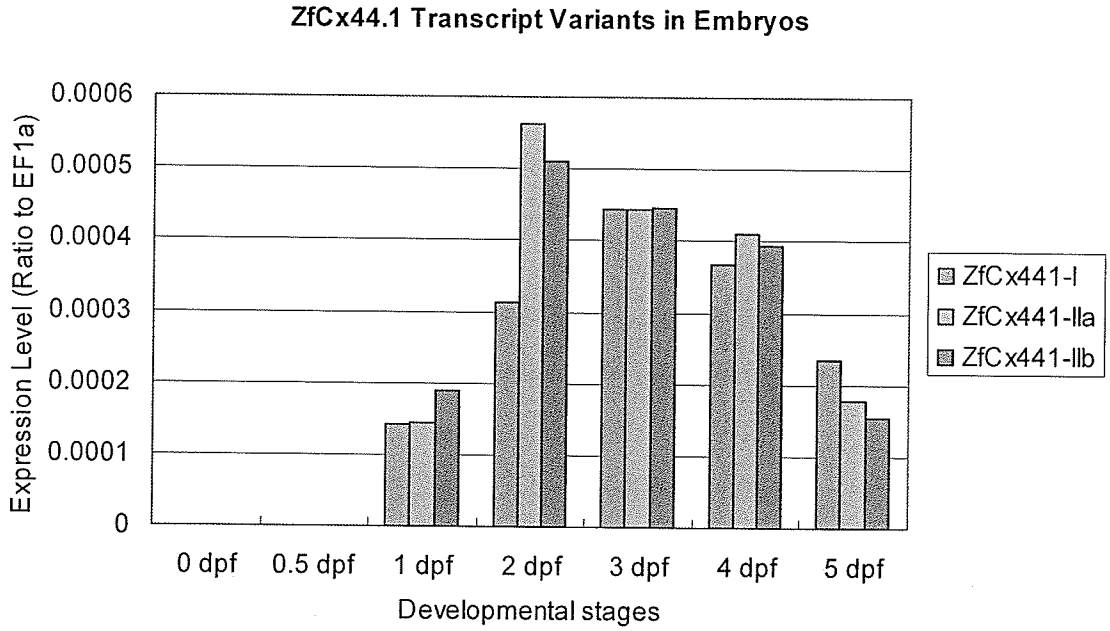


Figure 2-17. Expression of mCx50 transcripts in the adult lenses. Transcript I is much more strongly expressed than transcript II and III which are both the minority among all transcripts in the adult lens tissues. The expression ratios were calculated by calculating ΔCt between mean Ct values of test and EF1α genes in two replicate experiments, and the results were presented as a column graph.

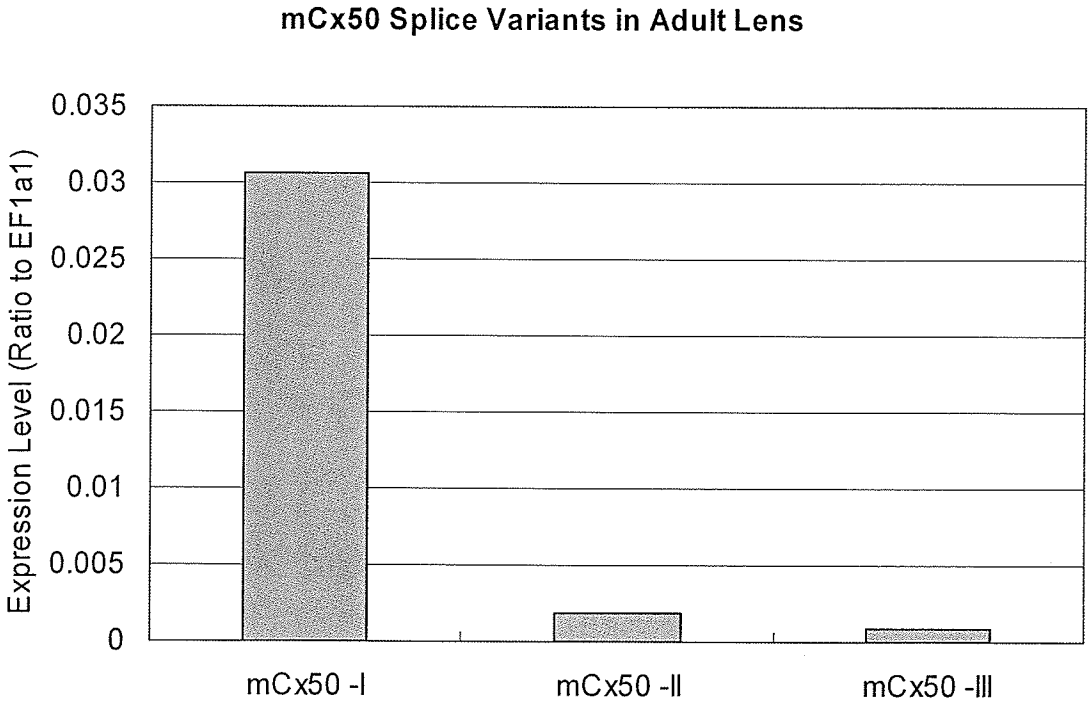


Figure 2-18. Prediction of potential transcription factor binding sites within the proximal promoters of both zebrafish Cx44.1 and mouse Cx50. Analyses were performed by MatInspector web-based software using the eye as the tissue filter with default parameters. Resultant figures were downloaded from the MatInspector web server, and then annotated manually. Different core promoters are indicated by arrows at their corresponding transcriptional start sites (TSSs), whereas the translational start codons are indicated by small vertical lines. Potential transcription factor binding sites are marked by solid dots with different colors surrounding the promoter sequences.

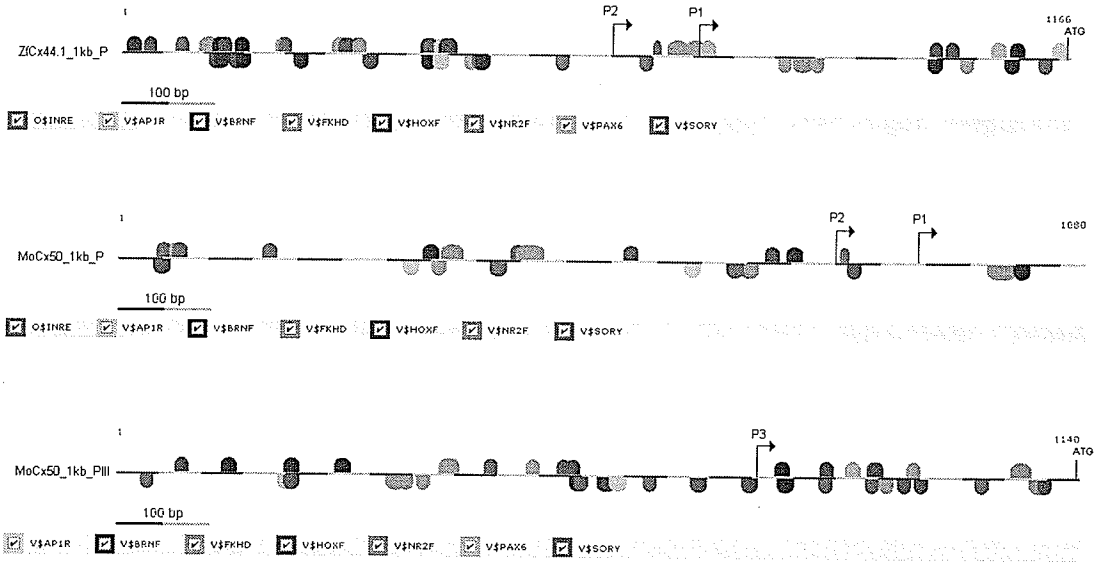


Figure 2-19. EGFP expression driven by ZfCx44.1 and mCx50 promoters. A ~ 1kb promoter fragment immediately upstream of the AUG codon of zebrafish Cx44.1 could drive EGFP transient expression in the retina of fish embryos (**A**). When the SV40 poly-A tail was replaced by the endogenous 3'UTR in the above Cx44.1-EGFP construct, EGFP expression was observed in the lens of fish embryos (**B**). Both mouse Cx50 promoters (P2 & P3) could also drive EGFP expression in the retina of zebrafish embryos (**C** & **D**).

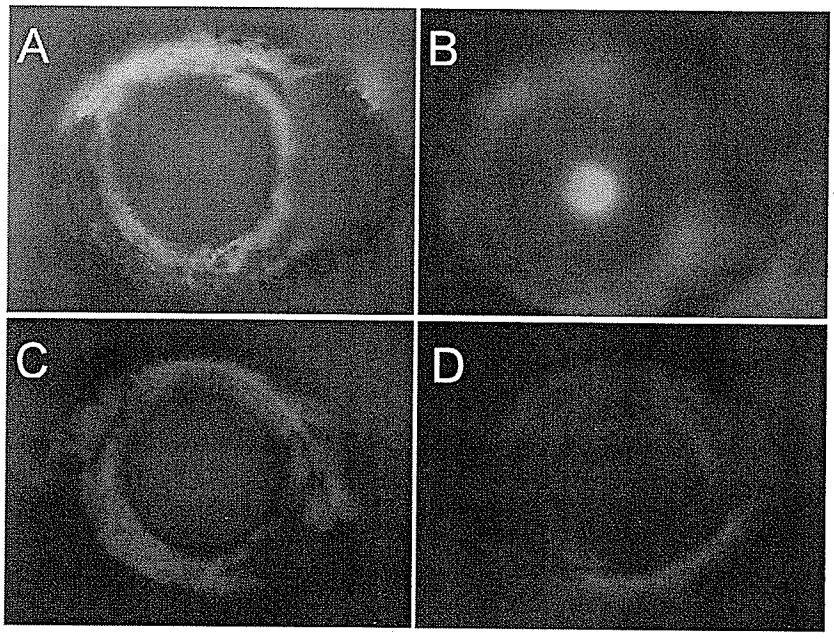


Table 2-1. Primers used for 5'RACE of zebrafish Cx44.1 and mouse Cx50.

Primer Name	Sequences (5' – 3')
GeneRacer™ RNA oligo	CGA CUG GAG CAC GAG GAC ACU GAC AUG GAC UGA AGG AGU AGA AA
GeneRacer™ 5'Primer	CGA CTG GAG CAC GAG GAC ACT GA
GeneRacer™ 5'Nested Primer	GGA CAC TGA CAT GGA CTG AAG GAG TA
Cx44.1 5'RACE Primer	CCA GAC GAA ACT TCT TGC TGC CCT TTG
Cx44.1 5'RACE Nested Primer	TCT GCG GCT GTG CCC AAG ATG AGG A
Cx50 5'RACE Primer	GGC TGC TGG GTG TTG CAT ACA AAA TCA G
Cx50 5'RACE Nested Primer	GAT TGC TCA TCG CCC CAC ACA AAC TC

Table 2-2 Zebrafish Cx44.1 and Mouse Cx50 Real-time PCR Primers

Templates	Sequences of Primers (5' – 3')	Amplicon
Cx44.1_I	(Cx441/q/1/F1) CTCTGGGAGCTCTTACTGTGTG	93 bp
	(Cx441/q/1/R1) CCGTCGAATGCTCATTTACTTCCTC	
Cx44.1_IIa	(Cx441/q/2a/F1) CGACAAGGATTGGAGGCCATAG	128 bp
	(Cx441/q/2a/R1) CCATGGTTGACTGTCAGCACAC	
Cx44.1_IIb	(Cx441/q/2b/F1) GACAAGGATTGGAGGCCATAGTG	94 bp
	(Cx441/q/1/R1) CCGTCGAATGCTCATTTACTTCCTC	
zEF1α	(zEF1α/q/F1) GAGGAGTGATCTCTCAATCTTGAAAC	94 bp
	(zEF1α/q/R1) TTTCCGGAGTCGACGTGGCCAATAAC	
zEF1α	(zEF1α/q/F1) GAGGAGTGATCTCTCAATCTTGAAAC	133 bp
	(zEF1α/q/R2) CCACCGCATTTGTAGATCAGATGG	
Cx50_I	(mCx50/q/1/F1) GTGGGACGCTACTCAGACGTGAGC	144 bp
	(mCx50/q/R1) CTGCTGTCCCGAGGATGAGGATGCGGAAG	
Cx50_II	(mCx50/q/2/F1) GCCTTCTGCTTCCTTGGTAGTGAGC	147 bp
	(mCx50/q/R1) CTGCTGTCCCGAGGATGAGGATGCGGAAG	
Cx50_III	(mCx50/q/3/F1) GGAAGTTGGAGGTGGGCTTTTAGGTG	133 bp
	(mCx50/q/3/R1) CAGTCGCCCATTGCTCACGTCTGAGTAG	
mEF1α1	(mEF1α1/q/F1) GTTTGCCGTCAGAACGCAGGTGTTGT	147 bp
	(mEF1α1/q/R1) CACATTTGTAGATCAGGTGGCCGGTTG	

Table 2-3. PCR primers for ZfCx44.1-GFP and mCx50-GFP constructs. Underlined sequences are primer adapters for eGFP coding sequence.

Primer Name	Sequences (5' – 3')
ZfCx441/PF10	CTC GAG TTG AGT TTG ATC TTA TAG TG
ZfCx44.1/gfp/R1	TCC TCG CCC TTG CTC ACC ATA GAA ACT CCA GTC ACC GAT GGT TG
gfp/ZfCx441/F1	TCG GCA TGG ACG AGC TGT ACA AGT AAT TGA CGG TAT AAG TGA TCC AGT G
mCx50/PF4	CTG CAG AAG CCC CAG CTA AGC
mCx50/PR2	GAG TAG CGT CCC ACG GAG AAA
mCx50/F2	GCA GCA AGA GAG AAA GAC AGC
mCx50/PF7	GCT ATC CCA AAA GTC CCC CGT AC
mCx50/gfp/R1	TCC TCG CCC TTG CTC ACC ATA GAA ACT CCA GTC GCC CAA TGC TC
gfp/F1	ATG GTG AGC AAG GGC GAG GAG
gfp/R1	GCG CGC AAT TAA CCC TCA CTA AAG

CHAPTER THREE

ZEBRAFISH CONNEXIN 45.6

3.1 Introduction

The adult zebrafish heart is a tubular structure consisting of four chambers in a series: the sinus venosus, the atrium, the ventricle, and the bulbous arteriosus (Lohr and Yost 2000; Hu et al. 2000). While the anatomy of the adult zebrafish heart displays some differences relative to the hearts of higher vertebrates, the embryonic development of the zebrafish heart is remarkably similar to that of other vertebrates, including mammals (Lohr and Yost 2000; Opitz and Clark 2000). For example, the zebrafish embryonic heart actually resembles that of a three-week post implantation human embryo (Harvey 1999). Genes that drive vasculogenesis and subsequent angiogenesis in zebrafish heart development appear common to those required in mammals, including for example, Notch, Sonic hedgehog, and vascular endothelial growth factor (Lawson et al. 2002). In zebrafish, the cardiovascular system begins to function by 24 hpf, but the cardiovascular function is not essential for early embryos because the embryo is small enough to obtain oxygen through simple diffusion (Warren and Fishman 1998; Grillitsch et al. 2005). This feature has provided an additional benefit specific to developmental analyses of the cardiovascular system in zebrafish embryos (Lohr and Yost 2000). For example, in morpholino-injected embryos, disruptions of morphogenesis and the abnormal function of the cardiovascular system can be observed for a period of time in the living embryos

without functional circulation. For these reasons, the zebrafish is becoming a favourite model organism for genetic studies of heart development and modeling human cardiovascular diseases (Lohr and Yost 2000; Chico et al. 2008).

In the heart, gap junctions are important for the propagation of action potentials between myocytes, and are therefore essential for the coordinated and rhythmic contractions of the atria and ventricles (Bleeker et al. 1980; Gourdie 1995; Kupperman et al. 2000). The connexins abundantly expressed in mammalian heart include Cx37, Cx40, Cx43 and Cx45. The expression pattern of each connexin is tightly controlled in a cell-type specific manner (Gros and Jongsma 1996; Jongsma and Wilders 2000; Coppen and Severs 2002; van Veen et al. 2006). Cx43, the most predominant connexin in the heart overall, is expressed in all the cardiac components excluding the sinoatrial (SA) and atrioventricular (AV) nodes, the His bundle, and proximal parts of the bundle branches (BBs). Cx45 is predominantly expressed in the AV node and His bundle, and also detected in low levels in the peripheral portion of the ventricular conduction system. Cx40 is found abundantly in the atrial working myocardium, the His bundle, and conducting BBs and Purkinje fibres. Cx37 is expressed in the cardiac endothelial cells (Reed et al. 1993). A similar set of connexins is also found in blood vessels. The endothelial cells express Cx40, Cx43, and Cx37; whereas the smooth muscle cells primarily express Cx43 and sometimes Cx40, Cx45, and/or Cx37 (Christ et al. 1996). Knockout mice have been used to analyze the functional roles of these connexins in cardiovascular function and development. For example, knockout mice demonstrated the

significant role of Cx40 in atrionodal conduction and proximal His-bundle conduction (Simon et al. 1998; Simon and McWhorter 2002; Kirchhoff et al. 1998; Kirchhoff et al. 2000; VanderBrink et al. 2000; Bagwe et al. 2005).

The zebrafish Cx43.4 (ortholog of mammalian Cx45) (Essner et al. 1996) and Cx43 (Chatterjee et al. 2005) have been cloned and characterized in the zebrafish. Previous studies in this lab have isolated zebrafish Cx45.6 gene (the ortholog of mammalian Cx40) from a genomic library, and its mRNA expression patterns have been extensively characterized using Northern blots, RT-PCR and *in situ* hybridization (Christie et al. 2004). These studies showed that ZfCx45.6 and mCx40 show a high level of similarity in amino acid sequences, expression patterns, and biophysical properties (Christie et al. 2004). With the isolation of these cardiovascular connexins in the zebrafish, further studies can be performed in order to elucidate the regulatory mechanisms for the expression and functional roles of these connexin genes during the normal and abnormal development and function of the cardiovascular system.

In the present study, the transcriptional start sites and the 5'UTR sequences of Cx45.6 were identified using 5'RACE techniques. Several types of transcript variants were detected in adult zebrafish heart and liver tissues. These transcript variants have distinct 5'UTR sequences that resulted from multiple promoter usage and alternative splicing. Using real-time RT-PCR, the relative amounts of each of these transcript variants of Cx45.6 were analyzed in early stage zebrafish embryos, as well as in several adult tissues, including the heart. Tissue expression patterns of Cx45.6 total mRNA were

also characterized by real-time RT-PCR method. These studies lay the groundwork for future studies on the regulatory mechanisms for gene expression and functional roles of Cx45.6 in the development and function of the zebrafish cardiovascular system.

3.2 Materials and Methods

3.2.1 Fish Care

See Chapter 2 for details.

3.2.2 RNA Preparation:

RNA samples were isolated using Trizol reagent (Invitrogen). The RNA samples with A260/A280 ratios between 1.8 and 2.0 were accepted for further use. The quality of RNA samples was also tested by RNA gel electrophoresis. For all experiments, fresh RNA samples were used, and no DNase treatments were performed before reverse transcription in order to avoid as much as possible any loss and degradation of RNA.

3.2.3 5'RACE (Rapid Amplification of cDNA 5' Ends)

5' RACE experiments were performed using a GeneRacer^{TM} Kit (Invitrogen) according to the manufacturer's recommended protocols (for details see Chapter 2). RNA samples were isolated from adult zebrafish heart and liver tissues (a mixture of RNA samples). 1 μg of total RNA was used in 5'RACE experiments. The gene-specific reverse

primers used were 5'- GGC CCA TGT AGA TGA GGG AGG GTG -3' for the first round of PCR and 5'- GGT GTC GAT ACG AAG ACG ATC TGC AGC -3' for the nested PCR. The nested PCR products were cloned into the pCR4-TOPO vector (Invitrogen), and the ligation was then transformed into One Shot® TOP10 chemically competent *E. coli* cells (Invitrogen). Independent positive clones were randomly selected and sent out for sequencing (DNA Centre, University of Calgary). Resultant sequences were then aligned against genomic DNA sequences of Cx45.6 to determine their transcriptional start sites and exon-intron structures. Using RNA structure prediction software Sfold (the server Web site: http://sfold.wadsworth.org.) (Ding and Lawrence 2003; Ding et al. 2005), we also analyzed the secondary structure of Cx45.6 5' UTRs.

3.2.4 Real-time RT-PCR

RNA samples for real-time RT-PCR of ZfCx45.6 were isolated from zebrafish embryos (12 hpf, and 1 – 5 dpf stages) and a series of adult tissues. For each sample, DNase pre-treatment was not performed in order to maintain as much as possible the RNA integrity. Approximately 3 μg of total RNA were reverse transcribed by Superscript III (Invitrogen), with a mix of random hexamers and oligo-$dT_{(18)}$ primers (Invitrogen), for 1 hour at 52°C, in a reaction volume of 20μl. The resulting cDNA was then diluted to 100μl with ddH_2O for further use. A reverse transcriptase negative control was also used for monitoring genomic DNA contamination.

Primer pairs for ZfCx45.6 were designed to specifically target their splice

variants. In order to avoid the amplification of potential contaminating genomic DNA, all primer pairs have at least one primer spanning an exon-exon boundary. For this purpose, the primers were manually designed, and their properties were checked by free software OligoAnalyzer 3.0 (Integrated DNA Technologies, Coralville, IA, USA). Zebrafish elongation factor 1α (zEF1α) was used as reference gene. Similarly, in order to restrict the amplification of genomic templates, the forward primer spanned the junction between exon 1 and exon 2, while the reverse primer was located in exon 2. Standard-curve tests were performed in order to validate the specificity and efficiency of all primer pairs. All primer pairs gave acceptable high specificities and efficiencies, thus allowing the comparative cycle threshold (ΔCt) method of analysis to be used. All primer pairs used for real-time PCR and their anticipated amplicon size are summarized in Table 3-1.

PCR reactions were performed on an iQ5 Real-Time PCR detection system (BIO-RAD). cDNA samples (2μl per well) were mixed with 250 nM of each primer and 10μl of SYBR Green Supermix (BIO-RAD, Hercules, CA, USA) reagent in a final volume of 20μl. In addition, a reverse transcriptase negative control and a template-negative control were also set up in parallel. PCR conditions were as follows: 95°C for 3 min followed by 40 cycles at 95°C for 10 sec, 61°C – 62°C for 10 sec, and 72°C for 10 sec. After amplification, melting curve analysis was employed to check PCR specificity, and monitor genomic DNA contamination.

Results were exported into Microsoft Excel for further analysis. Each sample was tested in at least duplicate reactions and all PCR runs were performed twice with cDNAs

prepared from independent batches of RNA samples. Mean Ct values were calculated, and then used for the calculation of ΔCt. The relative level of each transcript variants was first calculated relative to the reference gene and results were then presented as column graphs.

3.3 Results

3.3.1 Mapping transcriptional start sites and identification of 5'UTRs

Using 5'RACE method, we identified the transcription start site and the 5' UTR sequences of zebrafish Cx45.6. As discussed previously, this 5'RACE strategy can selectively capture the full-length of the 5' ends of the transcript, thus making it possible to map the transcriptional start sites. As previous studies demonstrated Cx45.6 expression occurs predominantly in the heart and liver tissues, RNA samples used for the 5'RACE experiment was a mixture of total RNA isolated from zebrafish adult heart and liver. The nested PCR products were cloned into pCR4-TOPO vector (Invitrogen), and 19 positive clones were randomly picked for sequencing. Sequencing results were then aligned against the genomic DNA sequence, resulting in the identification of the transcriptional start sites and the 5'UTR sequences of zebrafish Cx45.6 (Figure 3-1).

Six different 5'UTR sequences were identified and a sequence alignment is shown in Figure 3-1. For convenience of description, the adenine site in ATG codon was assigned as position +1, with position -1 representing its 5' nucleotide and position +2 its

3' nucleotide. Transcriptional start sites (TSSs) separated by less than 50 base pairs are considered to be controlled by the same promoter (also see Chapter 2). According to this criterion, three promoters of Cx45.6 were identified (P1, P2, and P3; Figure 3-2). P1 and P2 are alternate core promoters, whereas P3 is located about 4 kb upstream of both P1 and P2. A search for a TATA-box did not find TATA-like elements for either P1 or P2, while a TATA-like element was found upstream of the TTS driven by P3 (Figure 3-1).

Transcript variants produced by P1 and P2 contain two exons, while transcript variants from P3 contain three exons (Figure 3-2). It appears that these transcript variants only differ in the 5'UTR. As shown in Figure 3-1, the pre-mRNA splicing events follow the conserved "GU-AG" rule. Interestingly, two competing tandem "AGs" at the 3' splicing site were found, which are close to the translational start codon. To date, this is the first study that has demonstrated tandem alternative splicing sites in a connexin gene. As shown in Figure 3-1, alternate transcripts (Ia, IIa, and IIIa) selectively use the intron-distal AG as the splicing acceptor. Therefore, six types of transcript variants of Cx45.6 are generated due to multiple promoter usage and alternative pre-mRNA splicing.

3.3.2 5'UTR Secondary Structures of Zebrafish Cx45.6

Using software Sfold (http://sfold.wadsworth.org.) (Ding and Lawrence 2003; Ding et al. 2005), predictions of secondary structures of Cx45.6 5' UTRs were performed (Figure 3-4 to 3-6). Results show that the 5'UTRs (II/IIa and III/IIIa) are very highly structured. The most thermodynamically stable structures predicted by Sfold are shown,

along with the calculated free energy for each folding variant (e.g., $\Delta G°_{37}$ = -91.50 kcal/mol for Cx45.6-II). It appears that all transcripts (except I/Ia) have very strong secondary structures, with stable stem-loops throughout the 5'UTRs. It's also notable that the Cx45.6 5'UTRs II and III are relatively long 5'UTRs compared to typical eukryotic mRNAs.

A search for upstream AUG codons (uAUGs) revealed a couple of uAUGs within the transcript II and III of Cx45.6 (Figure 3-7). As the alternate 3' splice sites do not shift the open reading frame, only results for transcripts I, II and III are shown in this figure. All these uAUGs are followed by in-frame stop codons prior to the normal AUG codon, thus resulting in a couple of uORFs (upstream open reading frame) of variable length. It has been thought that upstream AUGs can inhibit cap-dependent translation (Morris and Geballe 2000; Iacono et al. 2005). The existence of uAUGs might provide a mechanism for potential control of translational initiation from the correct AUG codon.

3.3.3 Expression Patterns of ZfCx45.6 Transcript Variants

Using real-time RT-PCR, the tissue expression patterns of Cx45.6 mRNA were analyzed in adult zebrafish tissues (Figure 3-8). RNA samples were isolated from 15 different adult tissues including lens, retina, heart, liver, ovary, testis, spleen, gill, intestinal gut, brain, swim bladder, inner ear, skin, fin, and skeletal muscle. The forward primers span a common exon/exon junction, while the reverse primer is located in the coding region, thus allowing all transcript variants to be targeted by the primer pairs (see

below; primer pairs for transcript-I/Ia actually target all variants). The zebrafish EF1α gene was selected as the reference gene, since EF1α has been demonstrated as the most suitable reference gene for zebrafish real-time RT-PCR analyses, both in time-course and tissue based experiments (Tang et al. 2007). Experiments showed that zebrafish Cx45.6 is primarily expressed in adult heart tissue, with only low levels of expression in some other tissues such as the retina, liver, spleen, inner ear, and skeletal muscle (Figure 3-8). These results suggest that zebrafish Cx45.6 could be regarded as a heart-specific connexin.

Using real-time RT-PCR, I further analyzed the composition of alternative transcripts of Cx45.6 in the adult retina, heart, liver, spleen, and intestinal tract, respectively. The primers for zebrafish EF1α (reference gene) are the same as that used in Chapter 2 (also see Table 3-1). In order to prevent the amplification of genomic DNA templates, all primer pairs have at least one primer spanning an exon-exon boundary. As shown in Figure 3-3, primer pairs for transcripts II/IIa and III/IIIa were designed specifically to amplify these variants. Primer pairs for transcript I/Ia actually can target all transcript variants of Cx45.6. Therefore, the results for transcript-I/Ia can be corrected by calculation. The results showed that transcripts I and Ia are predominant among these six types of isoforms. Only trace expression of transcripts III and IIIa were detected among these tissues (Figure 3-9). Interestingly, the intron-proximal AG is preferably selected for 3' splice acceptor site. In order to compare the selection frequency between proximal and distal AGs, the same data were recalculated, in terms of the usage of

alternate 3' splice sites. In all tissues examined, the intron-proximal AG was used about 4 times more than the intron-distal AG (Figure 3-10).

In order to investigate the expression patterns of Cx45.6 transcript variants during embryonic development, analyses were further performed with embryos at early developmental stages (0.5 dpf, 1 – 5 dpf). Results showed that transcript I/Ia and II/IIa are expressed during zebrafish early embryonic development, while only small traces of expression of transcript III or IIIa were detected (Figure 3-11). The mRNA expression of Cx45.6 was readily detected at 1 dpf stage and at later stages, whereas no mRNA expression of Cx45.6 was detected in 0.5 dpf embryos. This observation is consistent with previous studies in this lab (Christie et al. 2004). Significantly, similar to that in adult tissues, transcript I/Ia is the most abundant transcript. It appears that the intron-proximal AG is preferably selected as the 3' splice site, with about 80% transcripts using the proximal AG. Therefore, the proximal AG appears to be selected about 4 times more than the distal AG in the splicing of zebrafish Cx45.6 transcripts both in embryos and adult tissues.

3.4 Discussion

In the present study, the 5'UTR sequences of zebrafish Cx45.6 and its exon-intron genomic organization were identified by 5'RACE techniques. Cx45.6 mRNA has multiple isoforms that are derived from multiple promoter usage and alternative splicing, differing only in the 5'UTR. Real-time RT-PCR demonstrated that most transcripts of

Cx45.6 contain two exons, which is similar to typical connexin genes. Interestingly, however, there are other types of transcripts (III and IIIa) that are transcribed from an alternative distal promoter (about 4kb upstream) and contain three exons. These transcripts have an extra exon that carries 127 nucleotides and share sequences with part of the first exon of the other, shorter transcripts (Figure 3-1 and 3-2). This finding is interesting as connexin genes typically contain only two exons (Sohl and Willecke 2003; Willecke et al. 2002), and to date, only a very few connexin genes have been found to have three exons (reviewed in Chapter 1). By real-time RT- PCR, transcript III was detected only in adult retina tissues - none were not detected in other tissues (Figure 3-9) or at early embryo stages (Figure 3-11). The mouse Cx40 (the ortholog of zebrafish Cx45.6) was found to transcribe a similar extra alternatively spliced exon (called "AS exon") (Anderson et al. 2005). For some transcripts of mouse Cx40, the first exon is spliced to this 57-bp AS exon and then to the second exon; whereas for other transcripts, the first exon is directly spliced to the second exon. Interestingly, similar to the situation of zebrafish Cx45.6 transcript III/IIIa, the transcripts containing the AS exon are nearly absent in mouse embryos and adult heart tissues, whereas the AS exon transcripts appear to be much more abundant in other adult tissues such as esophagus (Anderson et al. 2005). Though future analyses are needed, these observations suggest that the molecular mechanisms for expression of these connexin genes could be evolutionarily conserved.

Different promoter usage and alternative splicing that produce mRNAs of connexins with identical coding regions but alternative 5'UTRs have been previously

reported in recent years, including Cx32 (Sohl et al. 2001), Cx43 (Pfeifer et al. 2004), and five mouse connexin genes (Cx31, Cx40, Cx45, Cx46, and Cx47) (Anderson et al. 2005), and our own data on zebrafish Cx44.1, Cx48.5, and mouse Cx50 (see Chapter 2 and 4). Generally, 5'UTRs play essential roles in the regulation of gene expression by affecting mRNA nuclear export, cytoplasmic localization, translational efficiency and stability (Vilela et al. 1999; Wang et al. 1999; Kindler et al. 2005; Gebauer and Hentze 2004; Pickering and Willis 2005; Mignone et al. 2002). It has been estimated that 10-18% of genes have alternative 5'UTRs arising from multiple promoter usage (Zhang et al. 2004; Trinklein et al. 2003), and at least 13% of genes undergo alternative splicing in their 5'UTRs (Carninci et al. 2005). Multiple promoter usage is strongly linked to alternative splicing of 5'UTRs (Kornblihtt 2005), and gene regulation via alternative 5'UTRs has been shown to be important in a number of mammalian genes (Hughes 2006; Zhang et al. 2004), such as the genes encoding neuronal nitric oxide synthase (Wang et al. 1999; Newton et al. 2003), acetylcholinesterase (Meshorer et al. 2004), and axin2 (Hughes and Brady 2005). Interestingly, the addition of an extra non-coding exon has been demonstrated to play an important role in translational control for the transcripts encoding proinsulin (Shalev et al. 2002) and neuronal nitric oxide synthase (Wang et al. 1999; Newton et al. 2003). It was suggested that mouse Cx40 could use similar regulatory mechanisms for translation due to the existence of such an additional "AS exon" (Anderson et al. 2005). In this study, an extra non-coding exon formed by different promoter usage and alternative splicing in transcript III/IIIa of zebrafish Cx45.6 has been

identified, which implicates similar mechanisms for Cx45.6 translational control. Overall, the complexity of the 5'UTR diversity suggests that the expression of zebrafish Cx45.6 can be tightly regulated at the post-transcriptional level.

In this study, the sequence and secondary structures of ZfCx45.6 were analyzed to help look into whether this gene employed such regulatory mechanisms. 5'RACE demonstrated that these 5'UTR variants, especially transcript II/IIa and III/IIIa, are significantly long 5'UTRs, compared to typical eukaryotic 5'UTRs. A longer 5'UTR usually inhibits the initiation of the cap-dependent translation, because a long 5'UTR is more likely to contain uORFs as well as stable secondary structures that may interfere with the ribosomal scanning process (Kozak 1991). In this study, a search for uAUGs found a couple of uAUGs within the 5'UTRs of Cx45.6, for those transcripts with larger size. Prediction of the RNA secondary structure showed that Cx45.6 transcripts have stable loop-stem structures throughout their 5'UTRs (Figure 3-4, to 3-6). Transcript-I/Ia of Cx45.6 has moderately stable secondary structures, whereas transcripts II/IIa and III/IIIa form very stable structures. It has been suggested that the efficiency of translation can be significantly decreased when very strong structures (e.g., $\Delta G < -50$ kcal/mol) are formed in the 5'UTRs (Mignone et al. 2002). Therefore, the uORFs and stable secondary structures of 5'UTR variants of Cx45.6 may provide a mechanism by which rapid alteration of connexin expression can be achieved without the requirement for new transcription.

It was estimated that about 10% of all mRNAs contain atypically long 5' UTRs,

and their protein products typically regulate vital cellular processes, including cell growth, cell proliferation, apoptosis, etc (Pickering and Willis 2005). In situations where cap-dependent translation is impaired or suppressed, an alternative mechanism can be used by those long 5'UTRs, as translational initiation can be mediated by a secondary structure called an internal ribosome entry site (IRES) (Kaminski et al. 1994; Jackson 2005). Recent estimates suggest that up to 10% of all cellular mRNAs have the potential to initiate translation via this mechanism (Pickering and Willis 2005). Many of the IRESs identified in cellular mRNAs have complex RNA secondary structures. However, thus far, there is no recognizable conservation of secondary structure or sequence among these IRESs. It is hypothesized that many cellular IRESs contain short modules and internal initiation is mediated by the combined effect of these modules (Stoneley and Willis 2004), including a Y-shaped double hairpin structure (Le and Maizel 1997), the sequence GNRA (where N is any nucleotide, and R is a purine) (Hoffman and Palmenberg 1995; Robertson et al. 1999), polypyrimidine tracts, and the sequence AGACA (Huez et al. 1998). IRES-mediated translational initiation has been studied in Cx43 and Cx32 (Schiavi et al. 1999; Hudder and Werner 2000). These examples have interested scientists in the gap junction community to further look into other connexin genes (Werner 2000). 5'UTRs of zebrafish Cx45.6 were predicted to have stem-loops with an extended Y-shaped secondary structures. Closely looking into its 5' UTR primary sequences, we further found "GNRA", "AGACA" semi-conserved sequences and polypyrimidine tracts needed for IRESs. Such short RNA modules in its 5' UTR suggest that zebrafish Cx45.6

might have the potential to use the IRES-mediated mechanism for regulation of translation. In future studies, cistronic reporter assays could be employed to evaluate the functional roles of these motifs.

5'RACE experiments of zebrafish Cx45.6 identified a "CAGCAG" tandem alternative acceptor for pre-mRNA splicing (Figure 3-1). This is interesting as alternative splicing at such acceptor sites produces mRNAs that differ only by 3 nucleotides. To date, the present study is the first one that has reported tandem alternative splice sites in connexin genes. Although alternative splicing at tandem sites was observed a few years after the discovery of pre-mRNA splicing, it is not until recent years that this phenomenon has been well appreciated. Recent studies showed that many alternative splicing acceptor (and donor) sites are very close to each other, and they produce subtle splice variants with only a few nucleotides of difference (Zavolan et al. 2003; Chern et al. 2006; Hiller et al. 2004; Sugnet et al. 2004; Dou et al. 2006). Among such tandem alternative splice sites, acceptors with a "NAGNAG" motif (N = A, C, G, T) are most frequent in plants and animals (Hiller et al. 2004; Dou et al. 2006; Campbell et al. 2006; Tadokoro et al. 2005; Hiller et al. 2007). *In vitro* studies have suggested both a scanning process and competition between AGs that are responsible for the selection of the acceptor AG (Smith et al. 1993; Smith et al. 1989). According to the scanning mechanism, a spliceosomal component starts from the branch point, scans downstream, and generally selects the first AG (intron-proximal AG) as the 3' splice site during catalytic step II of splicing (Smith et al. 1993; Chen et al. 2000; Chua and Reed 2001).

The proximal AG can be skipped if it is too close to the branch point or if there is a distal AG in the adjacent as a competitor. Alternative splicing can result from this competition between the proximal and distal AGs. In general, a shorter distance between both AGs provides higher competition. Studies showed that efficient competition can occur if the distance is less than 6 nucleotides (Dou et al. 2006; Chen et al. 2000; Chua and Reed 2001). It has been suggested that the ratio of splicing at the intron-proximal versus the distal AGs can be highly regulated under specific spatio-temporal conditions or external stimuli (Hiller et al. 2004; Hu et al. 1999; Koenig Merediz et al. 2000; Takeda et al. 1994; Yan et al. 2000; Xu et al. 2004). This is not surprising because alternative splicing is a highly regulated process. It was shown that splicing factors (SF2/ASF and hnRNP A1) and step II factors (hSlu7) can affect the ratio between the proximal and distal sites (Pollard et al. 2002; Chua and Reed 1999; Lev-Maor et al. 2003). However, in some other cases, the tandem splicing ratio is rather constant in different tissues, at different development stages, or as a response to external stimuli (Tadokoro et al. 2005; Hammes et al. 2001; Unoki et al. 2006; Tsai and Lin 2006; Burgar et al. 2002; Li and Howe 2001; Rivers et al. 1999; Makielski et al. 2003; Treacy et al. 1992; Vogan et al. 1996). In these cases, the competition model suggests that the splice site selection can be achieved by stochastically binding to spliceosomal core components (such as U1 snRNA or U2AF35) (Chern et al. 2006). Studies showed that many tandem splice events can lead to the production of functionally different proteins, for example, the zebrafish pou5f1 gene (Takeda et al. 1994), and human SCN5A gene (Makielski et al. 2003). Stochastic splicing

of *Drosophila* DSCAM is essential for proper axon guidance (Graveley 2005). Furthermore, tandem splice sites in the 5' UTR can affect the efficiency of translation with the protein products identical between splice variants (Joyce-Brady et al. 2001). By contrast, some tandem splice variants do not necessarily show functional differences, for example, acceptor variants in the *Arabidopsis* PIMT2 gene (Xu et al. 2004). This type of tandem alternative splicing events is thought to be evolutionary neutral and thus simply tolerated by the cell (Hiller et al. 2004).

Real-time RT-PCR results in the present study suggest that both tandem acceptor AGs of zebrafish Cx45.6 are spliced at a constant ratio in different tissues and at different embryonic stages. This feature of Cx45.6 expression implies that the stochastic mechanism for tandem AG selection might be used. It is not clear at this stage whether or not the tandem alternative splicing of Cx45.6 may have functional importance since tandem acceptors are located in the 5'UTRs (thus producing the identical protein). Nevertheless, the prediction of 5'UTR secondary structure in this study showed that tandem alternative splicing of Cx45.6 can significantly affect the formation of loop-stem structures (Figure 3-4 to 3-6). Though further analyses are needed, this comparison between secondary structures implicates that tandem splice acceptors in zebrafish Cx45.6 could possibly play a role in the regulation of translational efficiency, as was shown in the regulation of γ-glutamyltransferase transcript processing (Joyce-Brady et al. 2001). Considering that this phenomenon appears to be not frequent in the connexin gene family, the tandem splicing of zebrafish Cx45.6 could possibly be a stochastic event and

therefore it may not have any evolutionary functional significance. In any case, *in vitro* reporter assays would provide a practical approach in our future studies to test whether or not these tandem splice acceptors regulate Cx45.6 expression.

Taking advantage of the zebrafish model, morpholino knockdown would provide a powerful strategy to elucidate the functional roles of Cx45.6 (and individual transcript variants) for the development of the cardiovascular system in zebrafish embryos. The identification of the 5'UTRs of Cx45.6 in this study enables us to design effective morpholino oligos for such functional analyses. The results in this study lays groundwork for us to further investigate the transcriptional and post-transcriptional regulation of Cx45.6 expression during embryonic development and in adult tissues. These studies are important to characterize the roles of connexin genes in normal and abnormal vertebrate cardiovascular development and function, and provide insight into the prevention and treatment of certain forms of human cardiovascular diseases.

3.5 Summary

In this chapter, the transcriptional start sites of zebrafish Cx45.6 were mapped and the 5'UTR sequences were identified using 5'RACE techniques. Cx45.6 is transcribed from different promoters and undergoes alternative splicing in the 5'UTRs, which result in six types of transcript variants that differ in the 5'UTRs. Interestingly, two tandem alternative splice acceptors in the "CAGCAG" motif were identified in this study, which produce mRNAs that only differ by 3 nucleotides. This is the first study that

has demonstrated tandem alternative splicing events in the connexin gene family. Using real-time RT-PCR, the tissue expression patterns of Cx45.6 mRNA were characterized among 15 zebrafish adult tissues. Zebrafish Cx45.6 was demonstrated to be a primarily heart-specific connexin gene. In addition, the relative abundance of these transcript variants was determined by real-time RT-PCR in early zebrafish embryonic stages and in adult tissues (such as the heart, retina, liver, etc.). Results showed that transcript I and Ia are the most abundant among these transcript variants, and the ratio between the selection of the intron-proximal and distal AG splice acceptor sites appears to be constant (about 4:1) during embryonic development and in adult tissues. *In silico* analyses of the 5'UTRs showed that transcript II(IIa) and III(IIIa) are highly structured with very stable loop-stem structures throughout the 5'UTRs. Interestingly, the subtle alternative splicing variants each showed different secondary structures relative to their alternatives (I vs Ia, II vs IIa, and III vs IIIa), which suggests that the tandem acceptors of Cx45.6 possibly play a role in the regulation of translational efficiency. Also, the Cx45.6 gene contains a couple of uORFs in the 5'UTRs. These analyses suggest that the regulation of Cx45.6 expression can be complicated at both the transcriptional and post-transcriptional levels. In future studies, *in vitro* assays could be employed to study the regulatory mechanisms for Cx45.6 expression. With the identification of the 5'UTR sequences of Cx45.6 in this study, effective morpholino oligos can be designed, and functional analyses using morpholino knockdown could provide insights into whether the splice variants are functionally important. In summary, the extensive studies presented in this chapter have

laid the groundwork for future studies on the functional roles of connexin genes in the vertebrate cardiovascular system using the favorable zebrafish model.

Figure 3-1. Alignment of ZfCx45.6 transcript variants. The adenine site in ATG codon was assigned as position +1. Splicing events follow the conserved "GU-AG" rule (highlighted with yellow color). Six types of transcripts are derived from different promoter usage and alternative pre-mRNA splcing. Tandem alternative 3' splicing sites were found among these transcripts. Transcript I(Ia) and II(IIa) have two exons each, whereas transcript III (IIIa) are three-exon transcripts.

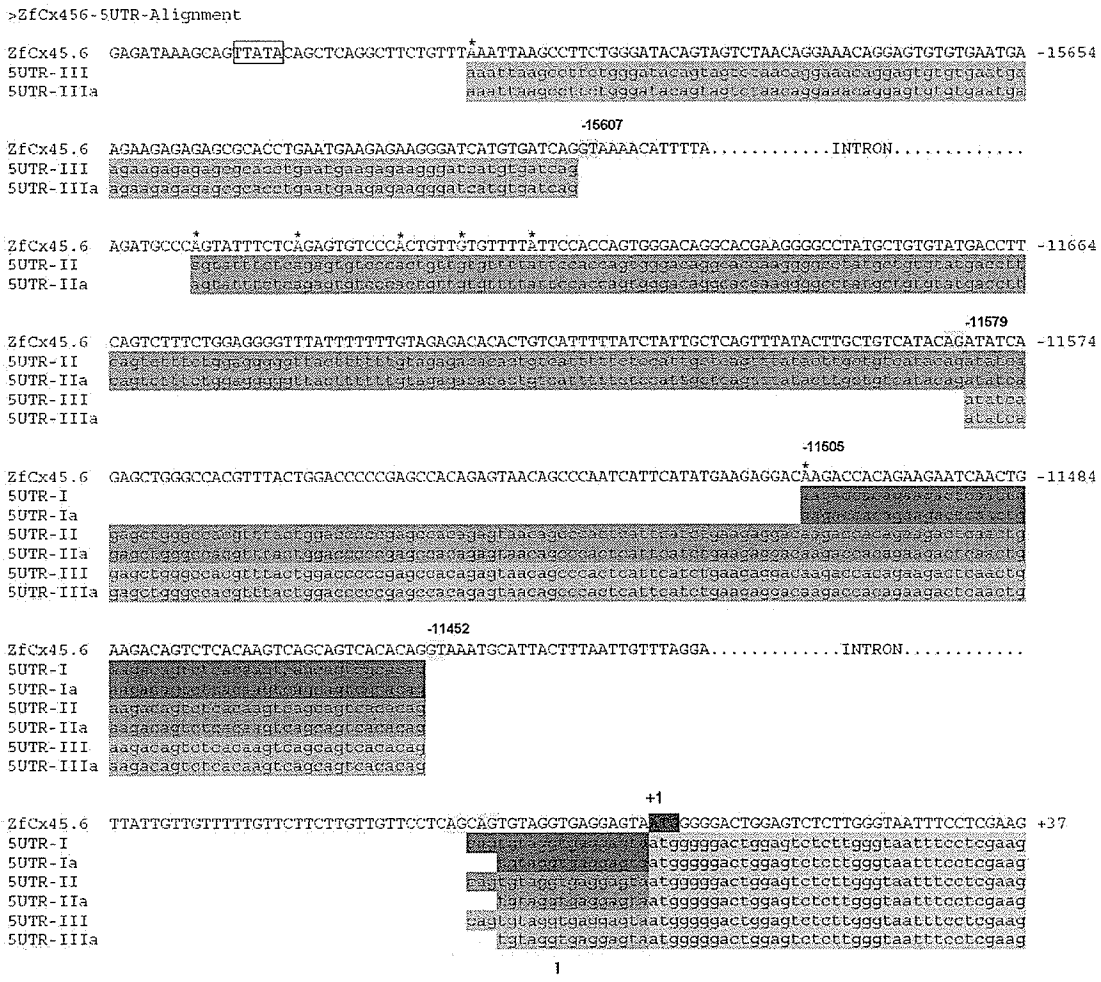


Figure 3-2. Sketch of ZfCx45.6 transcripts. Boxes represent the exons of Cx45.6 transcripts, while bold lines (or blank boxes) are genomic sequences or introns. All three transcripts share the same coding sequence and 3'UTR, whereas alternative first exons are derived from different promoter usage (P1, P2 or P3) and pre-mRNA alternative splicing (tandem alternative 3' sites). 5'UTRs are marked by the same colors as shown in Figure 3-1.

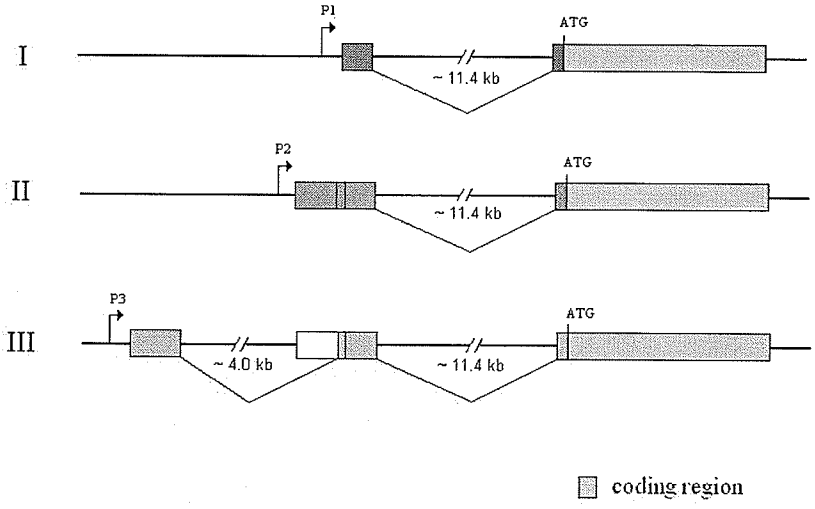


Figure 3-3. Design of ZfCx45.6 real-time PCR primers. Real-time PCR primer pairs were designed for targeting specific transcript variants of Cx45.6 (except that the primer pair for transcript I/Ia also targets two other transcripts). Forward primers are indicated by red arrows, while reverse primers by blue arrows. All primer pairs have at least one primer spanning the exon-exon boundary formed by the tandem alternative 3' sites.

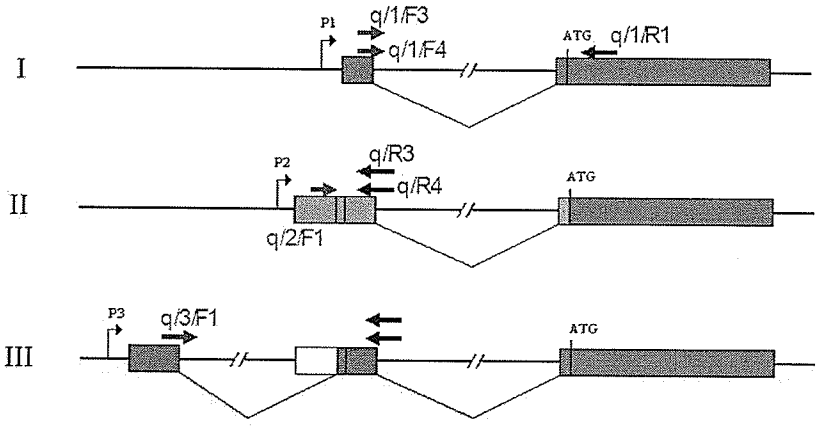


Figure 3-4. ZfCx45.6 transcript-I/Ia 5'UTR secondary structures. Predictions were performed using Sfold software. The most thermodynamically stable structure is shown with calculated free energy. Transcript-I (using the intron-proximal AG) is shown in (**A**) and transcript Ia (using the intron-distal AG) is shown in (**B**) ($\Delta G°_{37}$ = -11.90 kcal/mol and -8.90 kcal/mol, respectively).

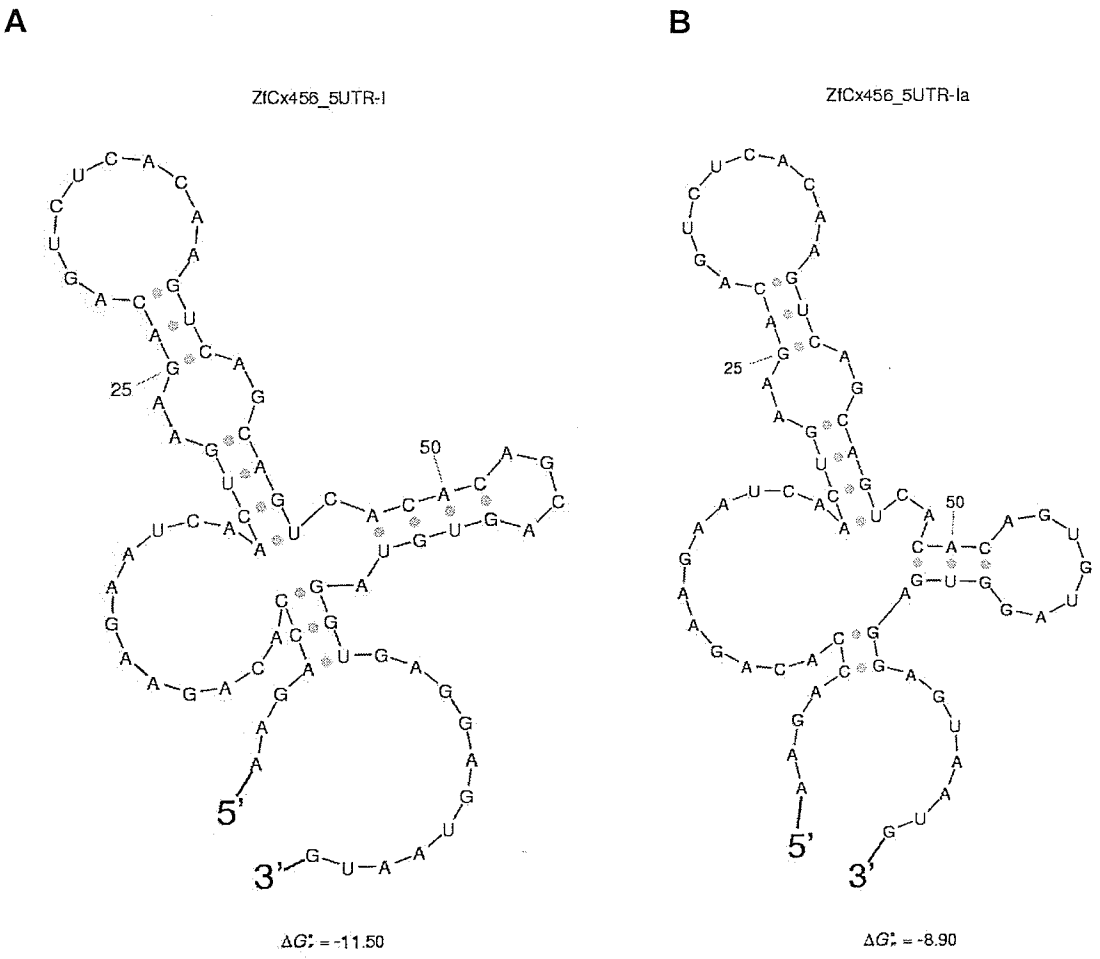


Figure 3-5. ZfCx45.6 transcript-II/IIa 5'UTR secondary structure. Predictions were performed using Sfold software. The most thermodynamically stable structure is shown with calculated free energy. Transcript-II (using the intron-proximal AG) is shown in (**A**) and transcript IIa (using the intron-distal AG) is shown in (**B**) ($\Delta G°_{37}$ = -91.50 kcal/mol and -92.80 kcal/mol, respectively).

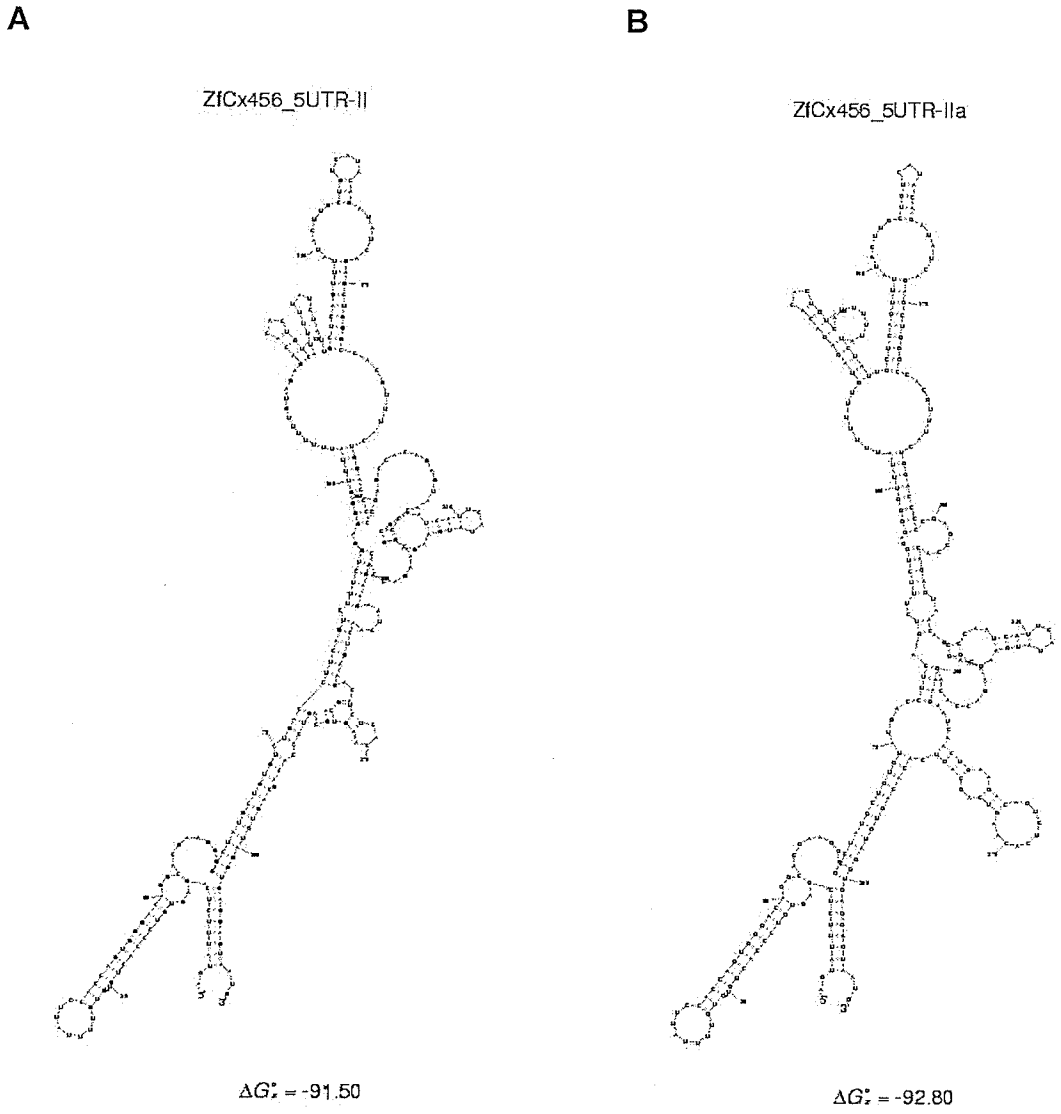


Figure 3-6. ZfCx45.6 transcript-III/IIIc 5'UTR secondary structure. Predictions were performed using Sfold software. The most thermodynamically stable structure is shown with calculated free energy. Transcript-III (using the intron-proximal AG) is shown in (**A**) and transcript IIIa (using the intron-distal AG) is shown in (**B**) ($\Delta G°_{37}$ = -50.20 kcal/mol and -50.10 kcal/mol, respectively).

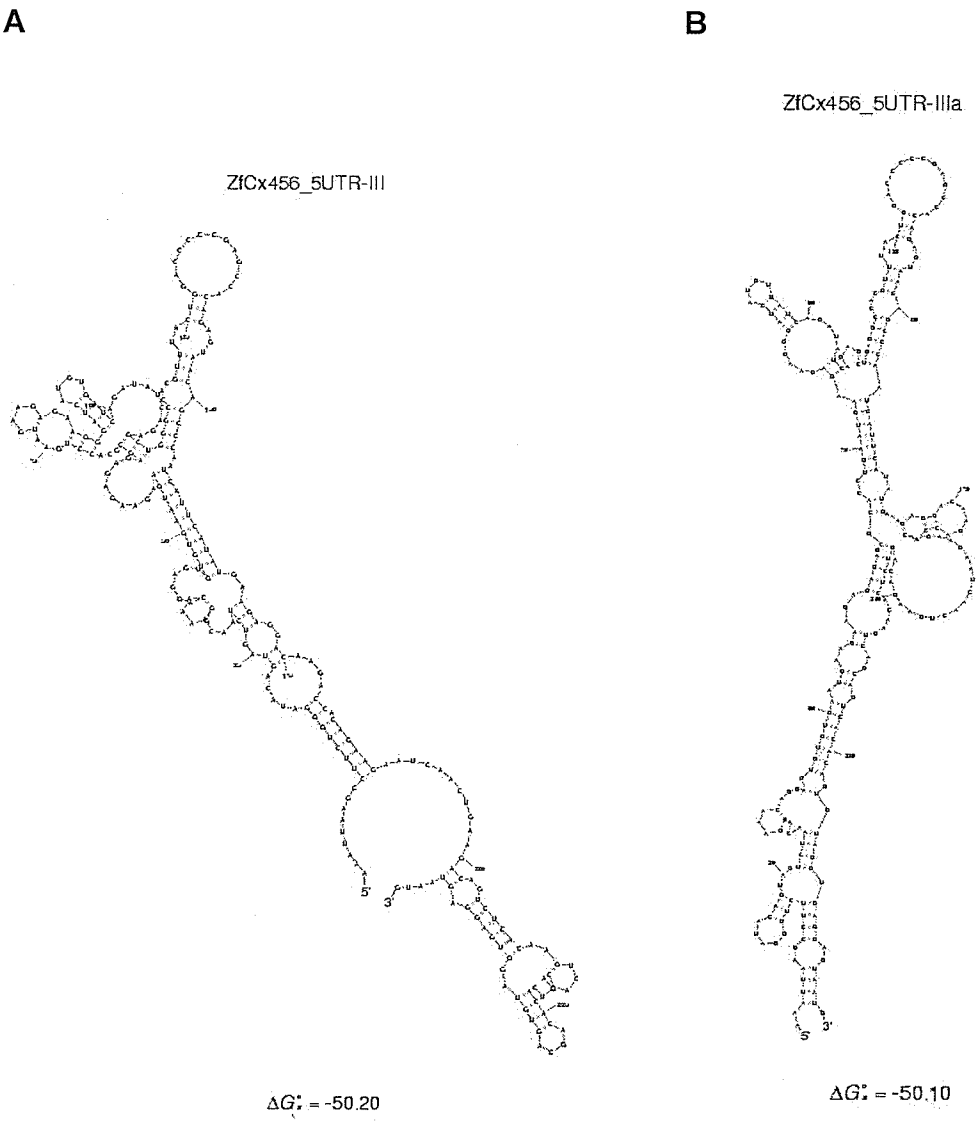


Figure 3-7. uORFs in 5'UTRs of ZfCx45.6. The normal translational start codons (AUG) are indicated by green solid triangles. uORFs (upstream open reading frames) of variable length are indicated by red arrows. As the alternate 3' splice sites do not shift the open reading frame, only results for transcripts I, II and III are shown in this figure.

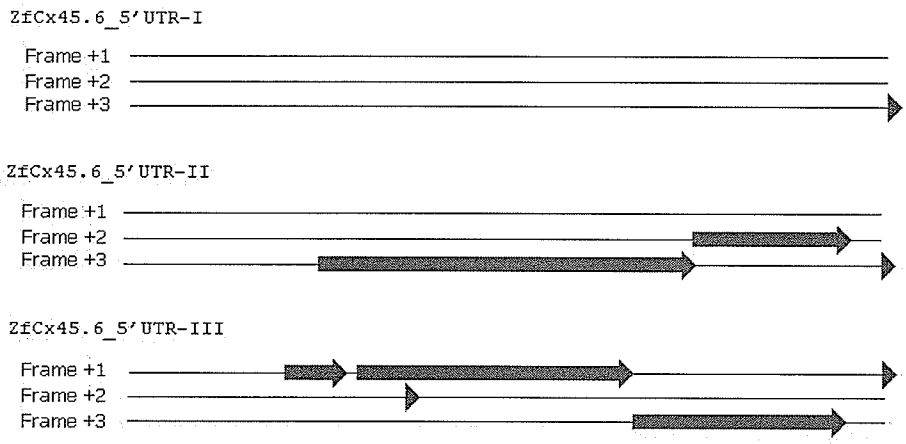


Figure 3-8. Tissue distribution of ZfCx45.6 total mRNA. Expression of zebrafish Cx45.6 mRNA (all transcripts combined) is tissue-specific. The expression ratios were calculated by calculating ΔCt between mean Ct values of test and EF1α genes in two replicate experiments, and the results were presented as a column graph. Total mRNA of Cx45.6 is predominantly expressed in the heart, while very low level expression was also detected in the retina, liver, spleen, inner ear, and skeletal muscle. ZfCx45.6 could be regarded as a heart-specific connexin.

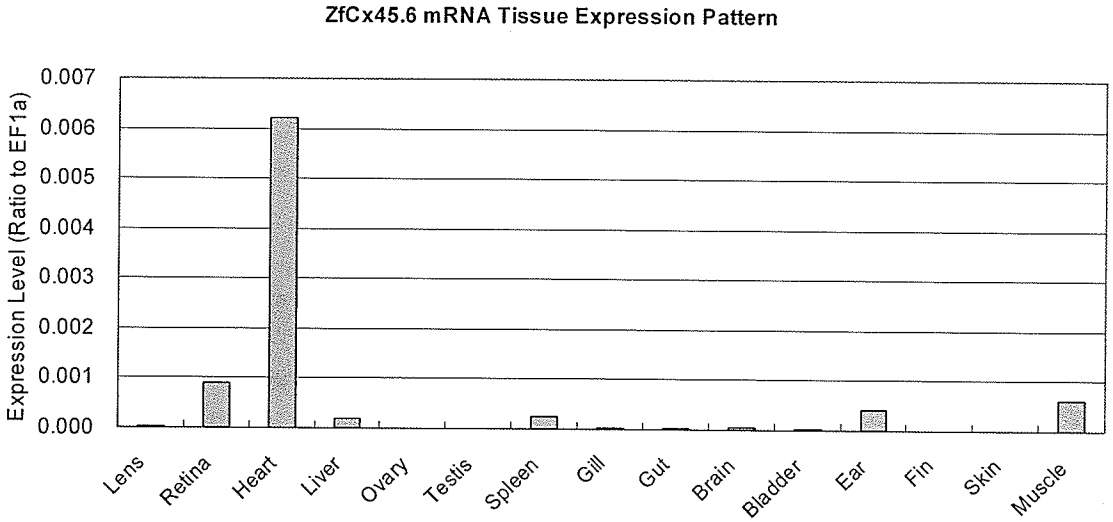


Figure 3-9. ZfCx45.6 transcripts in adult tissues. The expression ratios were calculated by calculating ΔCt between mean Ct values of test and EF1α genes in two replicate experiments, and the results were presented as a column graph.

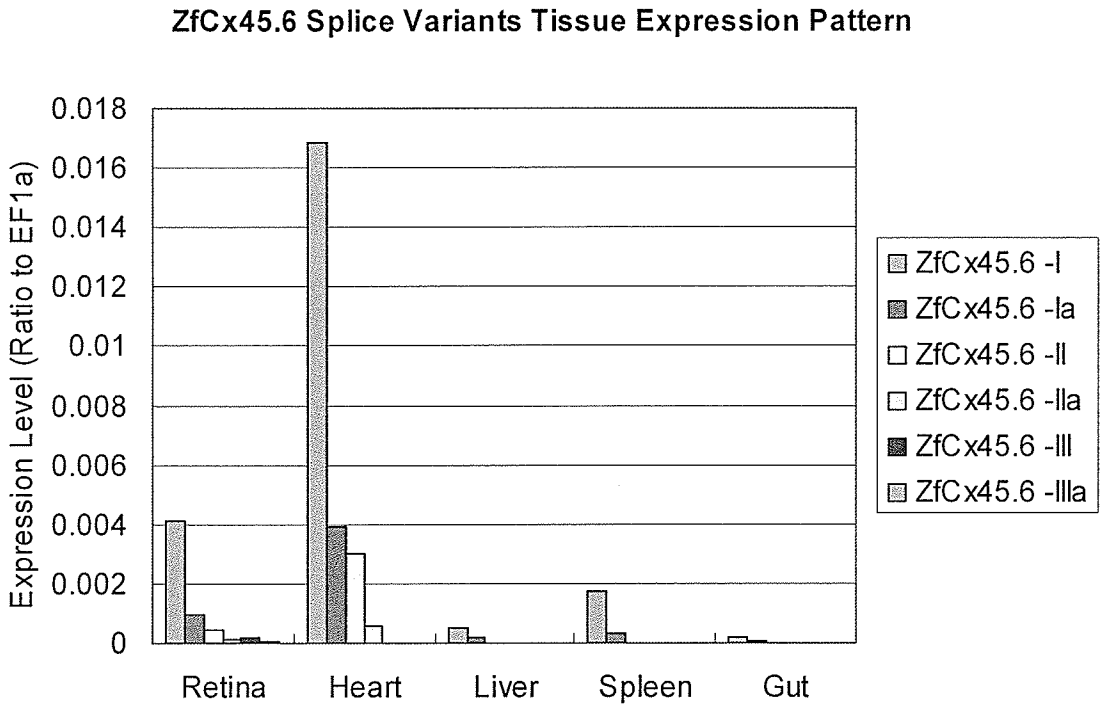


Figure 3-10. Distal- or proximal-AGs selection of ZfCx45.6 among tissues. The same data (as Figure 3-9) were re-calculated according to the selection of either intron-proximal or intron-distal alternative 3' splicing sites. The expression ratios were calculated by calculating ΔCt between mean Ct values of test and EF1α genes in two replicate experiments, and the results were presented as a column graph.

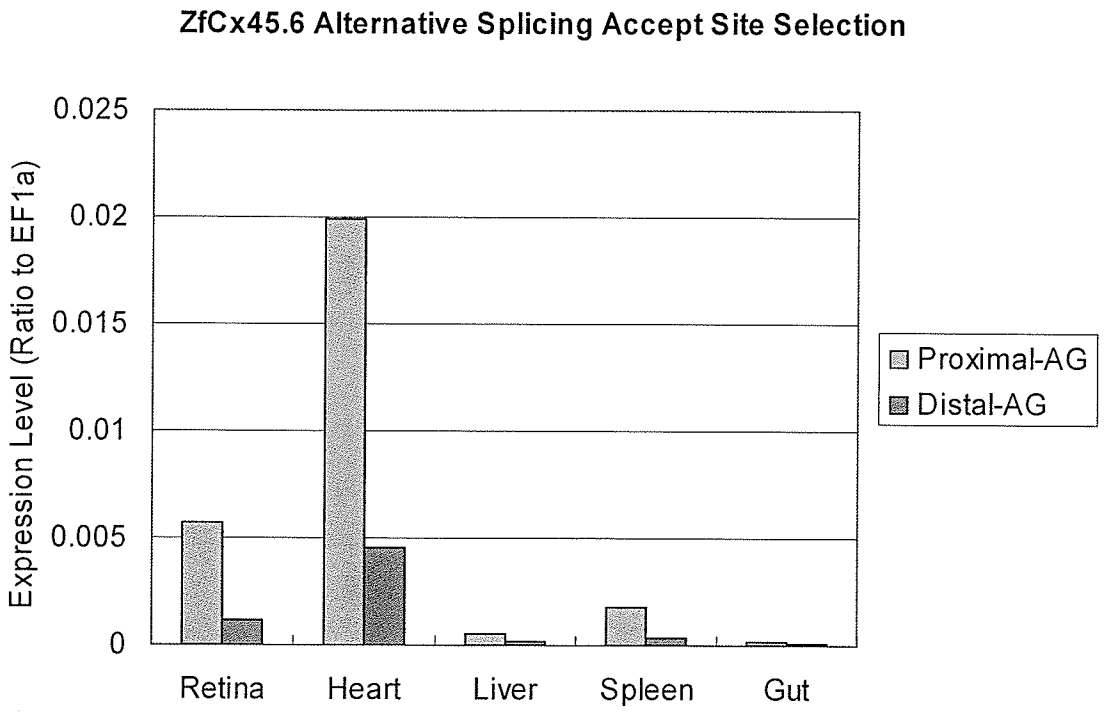


Figure 3-11. Expression pattern of ZfCx45.6 transcripts in embryos. Expression of zebrafish Cx45.6 transcript variants were detected before 1dpf (day post fertilization). The expression ratios were calculated by calculating ΔCt between mean Ct values of test and EF1α genes in two replicate experiments, and the results were presented as a column graph. Only small traces of transcript III were detected during late stage embryos. mRNA levels are expressed by ratios to that of zebrafish EF1α.

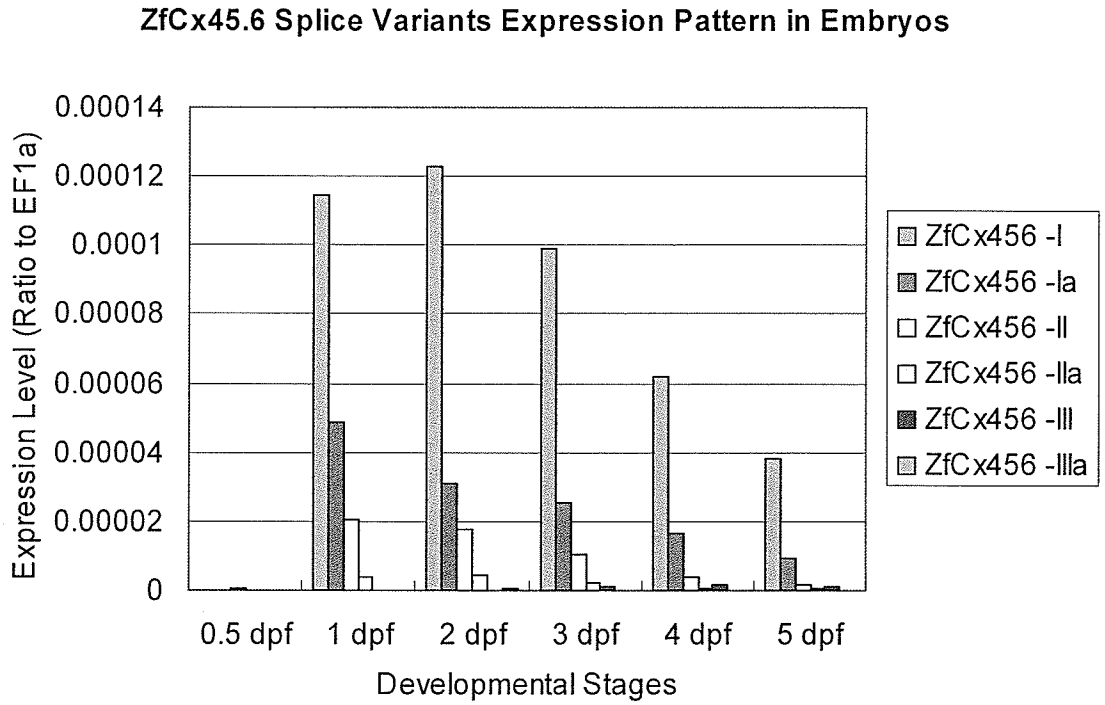


Table 3-1 ZfCx45.6 real-time PCR primers

Templates	Sequences of Primers (5' – 3')	Amplicon
ZfCx45.6 -I	(F) CACAAGTCAGCAGTCACACAGCA	173 bp
	(R) CATGATGACTCTGCCGCTGTGCCTAAC	
ZfCx45.6 -Ia	(F) CTCACAAGTCAGCAGTCACACAGTG	172 bp
	(R) CATGATGACTCTGCCGCTGTGCCTAAC	
ZfCx45.6 -II	(F) GCTCAGTTTATACTTGCTGTCATACAG	175 bp
	(R) CATTACTCCTCACCTACACTGCTGT	
ZfCx45.6 -IIa	(F) GCTCAGTTTATACTTGCTGTCATACAG	174 bp
	(R) CCCATTACTCCTCACCTACACTGTG	
ZfCx45.6 -III	(F) GAGAAGGGATCATGTGATCAGATATCAG	169 bp
	(R) CATTACTCCTCACCTACACTGCTGT	
ZfCx45.6 -IIIa	(F) GAGAAGGGATCATGTGATCAGATATCAG	168 bp
	(R) CCCATTACTCCTCACCTACACTGTG	
zEF1α	(F) GAGGAGTGATCTCTCAATCTTGAAAC	171 bp
	(R) CTTCCTTCTCGAACTTCTCGATGGTTC	

CHAPTER FOUR

ZEBRAFISH CONNEXIN 48.5

4.1 Introduction

In addition to zebrafish Cx44.1 and Cx43 (Chatterjee et al. 2005; Cason et al. 2001), this lab has previously isolated another zebrafish lens connexin gene, Cx48.5, by screening a zebrafish PAC library with a rat Cx46 cDNA probe. Amino acid sequence and biophysical similarities suggested that Cx48.5 is the zebrafish ortholog of mammalian Cx46 (Cheng et al. 2004). Using *in situ* hybridization, mRNA expression of ZfCx48.5 was detected in the otic vesicles in 1 dpf embryos (but not at later stages), and in the lens at 1.5 dpf and at later stages. RT-PCR showed that ZfCx48.5 is expressed in adult tissues, such as the lens and heart (Cheng et al. 2003; Cheng et al. 2004). Functional analyses using morpholino knock-down demonstrated that Cx48.5 has an essential role in the maintenance of lens homeostasis. Disruption of Cx48.5 expression by morpholino oligos resulted in the development of small lenses and eyes, as well as severe cardiovascular abnormalities in zebrafish embryos (Cheng et al. 2004).

Previous studies have isolated the zebrafish Cx48.5 gene, and extensively characterized the expression patterns and functional roles of Cx48.5 during zebrafish embryonic development, however, the mRNA sequence as well as molecular mechanisms for the expression and function of ZfCx48.5 remained unknown. As previously mentioned in Chapter 1, in the connexin research community, our knowledge

about gene structure and regulatory mechanisms for connexin gene expression and function is still rather limited. We believe that the zebrafish Cx48.5 gene could be an excellent model for studies on regulatory mechanisms for connexin gene expression and function. First, a number of molecular and genetic approaches are available in the powerful zebrafish model, including transient and stable expression of transgenes, and reverse genetic methods such as morpholino knockdown (reviewed in Chapter 1). Taking advantage of the favourite zebrafish model, a comprehensive investigation on the regulatory mechanisms for the expression and the functional roles of Cx48.5 can be conducted. Second, ZfCx48.5 has been demonstrated to be expressed in the lens during early embryonic development, as well as multiple adult tissues including the lens and the heart. The avascular lens is an excellent model for connexin research due to the essential roles of gap junctional communication in lens homeostasis. There are only two types of cells in the lens tissue, and there are only a few connexins known to be expressed in the lens (there are only three in mammalian lenses). These features indeed facilitate *in vivo* analyses of the expression and function of connexin genes. In addition, zebrafish is an excellent model system for studies on the development of cardiovascular system (Lohr and Yost 2000). In zebrafish, even though the cardiovascular system is functional by 24 hpf (hours post fertilization), it is not essential in the early embryonic stages, because the embryo is small enough to obtain oxygen through simple diffusion (Warren and Fishman 1998). This feature of zebrafish embryos provides an additional benefit specific to developmental analyses of the cardiovascular system by morpholino knockdown,

because disruptions of the morphogenesis and functions of cardiovascular system can be observed for a period of time even in the absence of functional vascular circulation.

The present study has aimed to address some questions regarding the regulatory mechanisms for the expression and function of ZfCx48.5. First, I have mapped the transcriptional start sites and identified the full-length cDNA sequences of zebrafish Cx48.5 by using RACE (Rapid Amplification of cDNA Ends) technology. Results showed that Cx48.5 contains three exons, which is among the unique, and very few known three-exon connexins. More interestingly, expression of Cx48.5 uses alternative first exons which arise from different promoter usage and alternative pre-mRNA splicing. Second, I have performed real-time RT-PCR to analyze the tissue distribution of Cx48.5 mRNA expression, as well as to determine the relative amount of these transcript variants during embryonic development and in several adult tissues. Third, I performed promoter analyses of Cx48.5 by *in vivo* GFP transient expression assays. Fourth, I have proposed and designed experiments for functional analyses of individual transcript variants of Cx48.5 by morpholino knockdown using stable transgenic zebrafish embryos that express GFP-labelled Cx48.5 (for details see Appendix II). In this chapter, I present the identification of cDNA sequences of zebrafish Cx48.5 and its mouse ortholog Cx46, the tissue distribution of mRNA expression of Cx48.5 as well as the expression patterns of transcript variants of Cx48.5 during embryonic development and in several adult tissues, and the GFP transient expression driven by a Cx48.5 promoter.

4.2 Methods and Materials

4.2.1 Fish care

See Chapter 2 for details.

4.2.2 RNA preparation

RNA samples used in the present study were isolated using Trizol reagent (Invitrogen). The concentration and quality of RNA preparations were monitored by UV spectrometry. The RNA samples with A260/A280 ratios between 1.8 and 2.0 were accepted for further use. The quality of RNA samples was also tested by RNA gel electrophoresis. For all experiments, freshly isolated RNAs were used, and no DNase treatments were performed before reverse transcription so as to avoid as much as possible any loss and degradation of RNA.

4.2.3 Mapping transcriptional start sites and the identification of the 5'UTRs

5' RACE (Rapid Amplication of cDNA 5' Ends) experiments were performed using a GeneRacer^{TM} Kit (Invitrogen) according to the manufacturer's recommended protocols. In brief, RNA samples for 5'RACE were isolated from adult zebrafish tissues (the lens, retina, and heart) and adult mouse lens tissues respectively. Each RNA sample (~1 μg) was treated with calf intestinal phosphatase (CIP) in order to de-phosphorylate the 5'end of truncated mRNA and non-mRNA molecules. The de-phosphorylated RNA

was then treated with tobacco acid pyrophospatase (TIP) to remove the 5' cap structure from mature full-length mRNA, leaving a 5' phosphate group, to which a universal RNA oligonucleotide (GeneRacer™ RNA oligo, 5'-CGA CUG GAG CAC GAG GAC ACU GAC AUG GAC UGA AGG AGU AGA AA-3') was then ligated by T4 RNA ligase. After these procedures, the full-length mRNAs were selectively added at their 5' ends with a priming site for the following PCR reactions. First strand cDNA was then reverse transcribed at 52°C for 1 hour using SuperScript III reverse transcriptase with a mix of random hexamers. Primers used for the first round and nested PCRs are listed in Table 4-1.

For each 5'RACE experiment, the nested PCR products were cloned into the pCR4-TOPO vector (Invitrogen), and the ligation was then transformed into One Shot® TOP10 chemically competent *E. coli* cells (Invitrogen). Positive colonies were randomly picked and grown in LB medium at 37°C with shaking (250 rpm) overnight. Independent plasmids were extracted and then sent out for sequencing (DNA Centre, University of Calgary). The resulting 5'UTR sequences were then aligned against genomic DNA sequences to determine their transcriptional start sites and exon-intron structures. Using RNA structure prediction software Sfold (the server Web site: http://sfold.wadsworth.org.) (Ding and Lawrence 2003; Ding et al. 2005), I also analyzed the secondary structure of 5' UTR of zebrafish Cx48.5 and mouse Cx46.

4.2.4 Identification of the 3'UTR sequence of ZfCx48.5

3' RACE (Rapid Amplification of cDNA Ends) experiments were performed using the GeneRacer™ Kit (Invitrogen). RNA sample was freshly isolated from the zebrafish lens, and its integrity/quality was confirmed by RNA gel tests. Without DNase treatments, ~1 μg of total RNA sample was reverse transcribed at 52°C for 1 hour using SuperScript III reverse transcriptase (Invitrogen) with the GeneRacer™ oligo dT primer (5'-GCT GTC AAC GAT ACG CTA CGT AAC GGC ATG ACA GTG $(T)_{24}$-3'). The resultant first cDNA (containing universal priming sites at the 5' end) was then used as template for the following first round and nested PCRs. All PCR primers are listed in Table 4-1. Similar to 5' RACE experiments described above, the nested PCR product was cloned into the pCR4-TOPO vector (Invitrogen), and 3 independent positive clones were sequenced (DNA Centre, University of Calgary). Using RNA structure prediction software mFOLD (Zuker 2003), its secondary structure was predicted.

4.2.5 Real-time RT-PCR

RNA samples for real-time RT-PCR of ZfCx48.5 were isolated from zebrafish embryos (0.5 dpf, and 1 – 5 dpf stages) and a series of adult tissues. For mouse Cx46, RNA samples were isolated from adult lens and retina tissues. For each RNA sample, DNase pre-treatment was not performed in order to maintain as much as possible of RNA integrity. Approximately 3 μg of total RNA were reverse transcribed by Superscript III (Invitrogen), with oligo-$dT_{(18)}$ (Invitrogen), for 1 hour at 52°C, in a reaction volume of 20μl. The resulting cDNA was then diluted to 100μl with ddH_2O for further use. Also, a

reverse transcriptase negative control was also used for monitoring genomic DNA contamination.

Primer pairs for ZfCx48.5 and mCx46 were designed to specifically target their splice variants (Figure 4-3). In order to avoid the amplification of potential contaminating genomic DNA, all primer pairs have at least one primer spanning an exon-exon boundary. For this purpose, the primers were manually designed, and their properties were checked by free software OligoAnalyzer 3.0 (Integrated DNA Technologies, Coralville, IA, USA). Zebrafish elongation factor 1α (zEF1α) and mouse elongation factor 1α-1 (mEF1α1) were used as reference genes for analyses of Cx48.5 and Cx46 respectively. Similarly, in order to restrict the amplification of genomic templates, both primer pairs of the reference genes were designed to span their first introns respectively. Standard-curve tests were performed in order to validate the specificity and efficiency of all primer pairs. All primer pairs gave acceptable high specificities and efficiencies, thus allowing the comparative cycle threshold (ΔCt) method of analysis to be used. All primer pairs used for real-time PCRs and their anticipated amplicon size are summarized in Table 4-2.

PCR reactions were performed on an iQ5 Real-Time PCR detection system (BIO-RAD). cDNA samples (2μl per well) were mixed with 250 nM of each primer and 10μl of SYBR Green Supermix (BIO-RAD, Hercules, CA, USA) reagent in a final volume of 20μl. In addition, a reverse transcriptase negative control and a template-negative control were also set up in parallel. PCR conditions were as follows:

95°C for 3 min followed by 40 cycles at 95°C for 10 sec, 61°C – 62°C for 10 sec, and 72°C for 10 sec. After amplification, melting curve analysis was employed to check PCR specificity, and monitor genomic DNA contamination.

Results were exported into Microsoft Excel for further analysis. Each sample was tested in at least duplicate reactions and all PCR runs were performed twice with cDNAs prepared from independent batches of RNA samples. Mean Ct values were calculated, and then used for the calculation of ΔCt. The results were expressed as a ratio of mRNA expression level to that of the respective reference gene. The ratios were calculated by comparing ΔCt between mean Cts of test and reference genes.

4.2.6 ZfCx48.5-EGFP transient expression assays in zebrafish embryos

By the screening of a zebrafish PAC library (RZDP Deutsches Ressourcen-Zentrum für Genomforschung GMbH, Berlin), PAC clones positive for Cx48.5 were previously isolated in this lab (Cheng et al. 2004). Cx48.5 positive PAC clones were digested with EcoRI and the resulting fragments were sub-cloned into a pBluescript II SK +/- vector (Stratagene, La Jolla, CA). A ^{32}P-labelled oligo probe was designed to hybridize to the second exon of Cx48.5, and used for a screening of the EcoRI subclones. One of the positive subclones containing the first two exons of Cx48.5 and a ~12 kb upstream fragment was selected for constructing the Cx48.5-EGFP expression plasmid (Figure 4-1, A). The subclone was digested with HindIII and then re-ligated to remove the upstream EcoRI site, thus leaving only one remaining EcoRI site

(situated in the second intron of Cx48.5) for further cloning use.

The EGFP coding sequence and the SV40 poly-A adenylation sequence (carrying NotI site at the 3' end) were retrieved by high fidelity PCR from the RARE3E(-SP6) EGFP expression vector (a generous gift from Dr. Robert Davis, Harvard Medical School). Using a forward primer situated on the second exon of Cx48.5, a genomic fragment immediately upstream of the translational start codon of Cx48.5 was retrieved from zebrafish genomic DNA by PCR. In order to fuse the EGFP coding sequence in frame to the 5'UTR leading sequence of Cx48.5, a specific fusion primer was designed (ZfCx485/gfp/R1, Table 4-3), and the above two PCR fragments were linked by a PCR strategy (see Appendix I). All PCR primers are listed in Table 4-3. All PCR reactions described above used Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen) to ensure the sequence accuracy of PCR products. The resulting "Cx48.5_eGFP_SV40-Poly-A" recombinant DNA fragment was then double digested with EcoRI and NotI restriction enzymes, and then cloned into the above modified Cx48.5_pBlueScript subclone; thus a plasmid with Cx48.5 ~12kb promoter driving EGFP expression was successfully constructed (Figure 4-1, B). The open reading frame of EGFP, and the sequence surrounding the specific recombinant sites were verified by DNA sequencing.

The Cx48.5_eGFP expression plasmid DNA was linearized with NotI restriction enzyme. The digestion products were purified with the GENECLEAN® Spin Kit (BIO-101 Inc., Vista, CA,USA), and then the concentration of DNA was adjusted to

about 100 ng/μl in 100 mM KCl solution with 0.05% of phenol red (as the injection indicator). DNA constructs were then used to inject into 1-cell stage zebrafish embryos. Injected and non-injected control embryos were separately incubated in egg water in a water-bath incubator with the temperature set at 28.5°C. Egg water was regularly changed every 4 – 8 hours, and at about 1 dpf stage, 0.003% of PTU (1-phenyl-2-thiourea) was added into the egg water in order to repress the pigment formation. Embryos were de-chorionated and relocated to 24-well cell-culturing plates, with 1.5 - 2 ml of fresh egg water, and one embryo in each well. Each injected embryo was then assigned with a numbers to be identified. Fluorescence expression was observed and recorded from this stage and later on, until day 5. During observation, live embryos were temporarily anesthetised in 4% Tricaine (Sigma). Images were captured and recorded using a Zeiss Axioskop FS microscope equipped with a Sony DXC-950 CCD camera and Northern Eclipse software (Empix, Mississauga, ON).

4.3 Results

4.3.1 Multiple transcriptional start sites and 5'UTR variants of ZfCx48.5 and mCx46

Using 5'RACE, I identified the transcription start site and 5' UTR sequence of zebrafish Cx48.5 and its mouse ortholog Cx46. As discussed previously, this 5'RACE strategy can selectively capture the full-length of the 5' ends of the transcript, thus

making it possible to map the transcriptional start sites. This method has been demonstrated to be comparable and, in many cases, more accurate than other approaches, for example, S1 nuclease protection, RNase protection, and primer extension (Liu and Gorovsky 1993; Schaefer 1995; Gum et al. 2003). For Cx48.5, the RNA samples were isolated from zebrafish adult lens, retina, and heart, respectively, and 5' RACE experiments were performed respectively. For each tissue, the nested PCR products were cloned into pCR4-TOPO vector (Invitrogen), and positive clones were randomly picked for sequencing. Sequencing results were then aligned against the genomic DNA sequence, resulting in the identification of the transcriptional start sites and the intron-exon structures of zebrafish Cx48.5 and mouse Cx46.

Using RNA samples from zebrafish adult tissues (the lens, retina, and heart), 5' RACE experiments identified multiple transcriptional start sites (TSSs) and the 5'UTR sequences of zebrafish Cx48.5. The alignments of 5'UTR sequences against the genomic DNA of ZfCx48.5 are presented in the Figure 4-2. For convenience of description, the adenine site in the ATG codon was assigned as position +1, with position -1 representing the 5' nucleotide and position +2 for its 3' nucleotide. Since the core promoter is typically situated within about 100 bp around the transcriptional start site (TSS), TSSs separated by less than 50 base pairs are considered to be controlled by the same core promoter. According to this criterion, two promoters of Cx48.5 were identified (P1 and P2) in the present study. As shown as stars in Figure 4-2, transcription of Cx48.5 can be initiated from multiple TSSs, for example, TTSs at positions -2044, -2039, and -2035 for

Cx48.5_P1. A search for a TATA-box did not find any TATA-like element for both P1 and P2.

The exon-intron organization of Cx48.5 gene is presented in Figure 4-3. Four types of transcripts were detected among three adult tissues (the lens, retina, and heart). All transcript variants contain three exons. Different types of transcripts differ in their first exons, which arise from multiple promoter usage and pre-mRNA alternative splicing. As shown in Figure 4-3, transcript I is transcribed from P1, whereas transcripts IIa, IIb and IIc from P2. The second exon of Cx48.5 is common among all transcripts, and it only contains a fragment of the 5' UTR. The coding region and the 3'UTR of Cx48.5 are situated in the third exon. The expression patterns of Cx48.5 transcript variants appear to be tissue-specific among these three adult tissues. For the 5'RACE experiment in the lens, transcript IIc was not detected; and in the retina, both transcripts I and IIb were not detected; while in the heart, transcript I was not detected. Although their tissue distribution patterns need to be further confirmed by other techniques, the 5' RACE experiments suggest that the promoter usage and the pre-mRNA splicing of Cx48.5 might be regulated in a tissue-specific manner.

Pre-mRNA splicing follows the conserved "GU-AG" rule, and undergoes alternative splicing. The first intron can be spliced out from alternative 5' splicing donor sites while sharing the same 3' splice accept site (nucleotide -35). Both transcripts I and IIc use the same 5' splicing site at nucleotide -1974, generating an intron of 874 bp in length (intron 1). Transcript IIa is spliced from positions -2254, generating an intron of

1154 bp in size (intron 1a). Transcript IIb uses the 5' splicing site at position -2126, splicing out an intron of 1026 bp in size (intron 1b). (Figure 4-3). Therefore, it is the usage of alternative 5' splicing sites that leads to the generation of three types of transcripts that are transcribed from the P2 promoter of Cx48.5 (i.e., IIa, IIb, and IIc).

In order to look into the possibility of evolutionary conservation of mRNA expression between zebrafish Cx48.5 and its mouse ortholog Cx46, I also performed 5'RACE analyses for Cx46 using RNA samples isolated from adult mouse lenses. It should be noted that at the time when our analyses were conducted (in 2003), another research group reported similar studies in which two Cx46 transcript isoforms were detected in mouse embryos (Anderson et al. 2005). However, using RNA samples from the adult mouse lens tissues, only transcript I was detected by our own 5'RACE experiments. Although it is not clear to us whether there are significant differences in the expression level of Cx46 transcript II during mouse embryonic development or in adult lens tissues, it appears that the expression level of transcript II could be very low in the lens tissues. In fact, this has been demonstrated by further analyses using real-time RT-PCR (see below). Here, I have presented both isoforms of mCx46 mRNA by aligning them against the genomic DNA sequence (Figure 4-4). Similar to zebrafish Cx48.5, it is clear that mouse Cx46 also has alternative promoters which result in alternative first exons. For P1 in the lens tissues, transcription of Cx46 can be initiated from multiple sites (at positions -20936, -20953, and -20955, respectively). Like Cx48.5, no TATA-like elements were found for both P1 and P2 of mouse Cx46. These preliminary results thus

far suggest some degree of evolutionary conservation regarding alternative promoter usage and mRNA expression between ZfCx48.5 and mCx46.

Interestingly, 5'RACE experiments of mCx46 unexpectedly detected a "trans-splicing" event between mouse Cx50 and Cx46 – the first exon of Cx50 was connected to the second exon of Cx46 in the adult mouse lenses. Real-time RT-PCR showed that the trans-splicing events of Cx46 are very rare in the adult mouse lenses (but not in the retina, see below). Further studies could be conducted in order to address the biological significance of the trans-splicing between these two lens-specific connexins.

4.3.2 5'UTRs of ZfCx48.5 and mCx46 are highly structured

The secondary structure of 5' UTRs of zebrafish Cx48.5 and mouse Cx46 were analyzed by using software Sfold (the server Web site: http://sfold.wadsworth.org.) (Ding and Lawrence 2003; Ding et al. 2005). Figure 4-6 to 4-11 show that their 5'UTRs are very highly structured. The most thermodynamically stable structures predicted by Sfold are shown with calculated free energy indicated in the figures (e.g., ΔG°_{37} = -34.30 kcal/mol for Cx48.5-I). The 5'UTRs of Cx48.5 are relatively long 5'UTRs compared to typical eukaryotic mRNAs, and the lengths vary substantially from each other (from 164 to 442 nucleotides). It appears that all transcripts have very strong secondary structures, with stable stem-loops throughout the 5'UTRs. Similar to zebrafish Cx48.5, both 5'UTRs of mouse Cx46 are also relatively long 5'UTRs, and their length are significantly different from each other. Prediction of the secondary structures of mouse

Cx46 transcripts also showed remarkably strong loop-stem structures (Figure 4-10, 4-11).

A search for upstream AUG codons (uAUGs) revealed a couple of uAUGs within the 5'UTR variants of zebrafish Cx48.5 (Figure 4-12, A), and one uAUG for each transcript isoform of mouse Cx46 (Figure 4-12, B). All these uAUGs are followed by in-frame stop codons prior to the normal AUG codon, resulting in a couple of uORFs (upstream open reading frame) of variable length (Figure 4-12, red arrows). It has been thought that upstream AUGs can inhibit cap-dependent translation (Morris and Geballe 2000; Iacono et al. 2005). The existence of uAUGs might provide a mechanism that allows fine control of translational initiation from the correct AUG codon.

4.3.3 Sequence analyses of 3'UTR of zebrafish Cx48.5

The 3'UTR sequence of Cx48.5 was identified by 3'RACE method. Zebrafish Cx48.5 has a relatively long 3'UTR with 2118 bp in size. Using mFOLD software (Zuker 2003), the prediction of secondary structure showed very strong loop-stem structures throughout the 3'UTR of Cx48.5. The most thermodynamically stable structure is shown with a calculated free energy (ΔG°_{37} = -445.22 kcal/mol) indicated in Figure 4-13.

4.3.4 Expression patterns of ZfCx48.5 and mCx46 transcript variants

Using real-time RT-PCR methods, I have performed analyses for the distribution of alternative transcripts of Cx48.5 in the adult lens, retina, and heart, respectively. The zebrafish EF1α gene was selected as a reference gene, since EF1α has been demonstrated

as the most suitable reference gene for zebrafish quantitative RT- PCR analyses, both in time-course and tissue based experiments (Tang et al. 2007). The primers for zebrafish EF1α are the same as that used in Chapter 2 (also see Table 4-2). As mentioned, primer pairs for ZfCx48.5 were designed to specifically target their splice variants (Figure 4-3). In order to prevent amplifying genomic DNA templates, all primer pairs have at least one primer spanning an exon-exon boundary. As shown in Figure 4-3, transcript I, IIa, and IIb share the reverse primer that spans the boundary of exon 2 and 3. Forward primers for transcript I, IIa, IIb were designed to span the boundaries between their first and second exons respectively, thus allowing the PCR fragments from different transcripts to have very similar length and nucleotide compositions with minimal differences only in their 5' ends. These considerations of primer design are valuable for my real-time RT- PCR experiments because the relative amounts of different transcripts are of interest. In addition, an additional primer pair was designed specifically for transcript IIc, since the primer pair targeting transcript-I also targets transcript IIc (Figure 4-3). The reverse primer for transcript IIc spans the boundary between exon 1 and exon 2, while the forward primer is situated within a unique 5'UTR region of transcript IIc. The results for transcript-I can therefore be corrected by calculation. Figure 4-14 shows the composition of Cx48.5 transcript variants in the adult lens, retina, and heart tissues. Cx48.5 mRNA expression is primarily expressed in the lens, whereas in the retina and heart tissues, only very low levels of Cx48.5 transcripts were detected. More significantly, the expression of transcript variants of Cx48.5 in different tissues appears to be in a tissue-specific manner.

For example, transcript I, IIa, and IIb are primarily expressed in the lens, whereas transcript IIc is mostly expressed in the heart tissues. Transcript IIb is the most abundant transcript in the lens tissues, whereas transcript IIc is the most abundant Cx48.5 transcript in the heart tissues (Figure 4-14).

Expression pattern analyses for the distribution of alternative transcripts of mCx46 were also performed with RNA samples from the mouse adult lens and retina respectively (Figure 4-15). Although the mRNA expression of Cx46 was detected in both tissues, it is primarily expressed in the lens. Both lens and retina express transcript I. Both transcript variants of mouse Cx46 were detected in the adult lens tissues, but the level of expression of transcript II and the trans-spliced transcript are very low. Their expression was not detected in the retina.

In order to investigate the expression patterns of the transcript variants of Cx48.5 during embryonic development, I further performed real-time RT-PCR analyses with embryos at early developmental stages (12 hpf, 1 – 5 dpf). Results showed that all of these four transcript variants are indeed expressed during zebrafish early embryonic development, although the relative amount of Cx48.5 transcripts in the embryo is very low when compared with that of EF1α (less than 0.02% of EF1α mRNA expression, Figure 4-16). Small amounts of transcripts (IIb) were detected at 12 hpf, and the expression of Cx48.5 transcripts were readily detected at 1 dpf and later stages. Similar to adult lens tissues, transcript IIb is the most abundant transcript; but unlike adult lens tissues, expression of transcript IIc is also very strong in the embryos. The expression

patterns of Cx48.5 transcript variants might be tissue-specific in the early embryos, but due to the small size of the embryos, it was too hard to obtain tissue samples from early embryos for RT-PCR analyses. According to the results obtained from adult tissues, transcript IIc might be a "heart-specific" transcript (see above). This may also be true in early embryos. However, it is also possible that expression of transcript IIc could be stronger in the developing lens compared to adult lens. In any case, it will be interesting to further investigate the dynamic spatio-temporal expression patterns of these transcript variants in early embryos. For example, i*n situ* hybridization could be employed using probes that specifically target different transcript variants. However, taking advantage of other molecular biology tools available in zebrafish, stable transgenesis might be a more powerful strategy to mimic the dynamic spatio-temporal expression patterns of these transcript variants in living embryos.

The expression pattern of zebrafish Cx48.5 mRNA was previously studied in this lab. Using a probe spanning from 612 nucleotides downstream of the start codon to 81 nucleotides downstream of the stop codon, mRNA expression of ZfCx48.5 was detected by *in situ* hybridization in the otic vesicles in 1 dpf embryos (but not at later stages), and in the lens at 1.5 dpf and at later stages. RT-PCR showed that ZfCx48.5 is expressed in adult tissues, such as the lens and heart (Cheng et al. 2003; Cheng et al. 2004). In the present study, using real-time RT-PCR, I further studied the tissue expression patterns of Cx48.5 mRNA expression in adult zebrafish (Figure 4-17). RNA samples were isolated from 15 different adult tissues including lens, retina, heart, liver, ovary, testis, spleen, gill,

gut, brain, swim bladder, inner ear, skin, fin, and skeletal muscle. The primers used in these analyses are the same as those used for specific transcript variants (I, IIa, IIb) of Cx48.5. The common reverse primer, together with a mixture of these three forward primers, was used as a "primer pair", which can amplify all four types of transcripts of Cx48.5 and produce almost the same PCR products in a single reaction. RT-PCR experiments showed that zebrafish Cx48.5 is clearly expressed in lens, while its expression is also detected at 10-fold lower levels or less in some other tissues such as the heart, skeletal muscle, retina, spleen, and inner ear (Figure 4-17).

4.3.5 GFP transient expression driven by a Cx48.5 ~12 kb promoter in the zebrafish embryos

Promoter activity of zebrafish Cx48.5 was tested by EGFP transient expression assays in micro-injected zebrafish embryos. As described above, the Cx48.5-eGFP expression construct used in the present study contains a ~12 kb promoter fragment of Cx48.5 originating from a zebrafish PAC subclone. In this expression construct, the EGFP coding sequence (plus SV40 poly-A fragment) is fused in-frame to the 5'UTR leading sequence of Cx48.5 without modifying the exon-intron genomic structure, and thus only the coding sequence and 3'UTR of Cx48.5 are replaced (Figure 4-1, B). Therefore, normal mechanisms for pre-mRNA splicing of endogenous Cx48.5 were expected to regulate the EGFP transient expression *in vivo* in zebrafish embryos. I believe that this design can effectively control potential artefacts since the pre-mRNA

splicing also appears to be a significant aspect of the regulation of Cx48.5 expression.

As shown in Figure 4-19, very strong EGFP expression was observed in the lens during embryonic development. EGFP expression in the retina was also detected although it was relatively weak, and did not accompany a strong EGFP expression in the lens. Unexpectedly, EGFP expression in the heart was not detected in the injected embryos.

4.4 Discussion

In the present study, the mRNA sequence of zebrafish Cx48.5 and its exon-intron genomic organization have been identified by both 5' and 3'RACE techniques. Zebrafish Cx48.5 contains three exons. Exon 1 and exon 2 only contain the 5'UTR, whereas exon 3 contains the remaining 5'UTR, the coding region and the 3'UTR. The genomic organization of Cx48.5 is quite different from typical connexin genes. As reviewed in Chapter 1, connexin genes commonly have two exons with their 5'UTRs interrupted by an intron of variable length (Sohl and Willecke 2003; Willecke et al. 2002). To date, only a very few connexin genes have been found to contain three exons, for example, human Cx30 (Essenfelder et al. 2005), mouse Cx40 (Anderson et al. 2005) and Cx45 (Jacob and Beyer 2001). Similar to typical connexins, however, it appears to be a common rule that the uninterrupted coding region and 3'UTR are located in the last exon. 5'RACE results also showed multiple transcriptional start sites and alternative pre-mRNA splicing of ZfCx48.5 mRNA. Four types of transcript variants have been identified in this study.

These transcript variants share both exon 2 and exon 3, whereas they differ in their first exons, which arise from alternative splicing and promoter usage (Figure 4-2, and 4-3). Also in this study and others (Anderson et al. 2005), multiple variants of exon 1 are also evident for the mouse Cx46 (ortholog of zebrafish Cx48.5). These discoveries, together with previous data regarding ZfCx44.1 and mCx50, indicate that mRNA diversity at the 5'end (the first exon) is a common feature of these connexin genes, and significantly, this feature of connexin expression appears to be evolutionarily conserved across species.

UTRs play essential roles in the regulation of gene expression by affecting mRNA nuclear export, cytoplasmic localization, translational efficiency and stability (Vilela et al. 1999; Wang et al. 1999; Kindler et al. 2005; Gebauer and Hentze 2004; Pickering and Willis 2005; Mignone et al. 2002). Several mechanisms for translational regulation by 5'UTRs have been proposed. These include (but are not limited to) the following examples. First, 5'UTRs provide potential binding sites for regulatory proteins which can modulate many aspects of mRNA function (Derrigo et al. 2000). Secondly, primary structure of the 5'UTR can affect translational initiation, for example, by upstream open reading frames (uORFs) (Morris and Geballe 2000; Iacono et al. 2005). Thirdly, strong secondary structures in 5'UTRs repress cap-dependent translation by inhibiting binding or scanning of the translational machinery, whereas some structural motifs act as internal ribosome entry sites inducing cap-independent translation (Pickering and Willis 2005). Alternations in the 5'UTR have the potential to introduce post-transcriptional regulatory elements.

The complicated 5'UTR diversity suggests that the expression of these connexin genes can be tightly regulated at the post-transcriptional level. In this study, the sequence and secondary structures of ZfCx48.5 and mCx46 were analyzed to examine the potential to employ such regulatory mechanisms. 5'RACE demonstrated that all 5'UTR variants of Cx48.5 and Cx46 are relatively long, compared to typical eukaryotic 5'UTRs, and their lengths are significantly different. Notably, most eukaryotic mRNAs contain short 5'UTRs (frequently fewer than 50 nucleotides), and thus cap-mediated translation of such mRNAs is efficient. A longer 5'UTR usually increases the efficiency of translation by facilitating the recruitment of more pre-initiation complexes, but further increases in 5'UTR length usually down-regulate the cap-dependent translation, because a long 5'UTR usually increases the propensity for the formation of uORFs as well as stable secondary structures that may inhibit translation initiation by interfering with the ribosomal scanning process (Kozak 1991). In this study, a search for uAUGs found a number of uAUGs within the 5'UTRs of Cx48.5, especially for those transcripts with larger size. Prediction of the RNA secondary structure showed that Cx48.5 and Cx46 have very stable loop-stem structures throughout their 5'UTR variants. Transcript I of Cx48.5 has moderately stable secondary structures, whereas transcripts IIa, IIb, and IIc form very stable structures. It has been suggested that the efficiency of translation can be significantly decreased when very strong structures (e.g., $\Delta G < -50$ kcal/mol) are formed in the 5'UTRs (Mignone et al. 2002). Therefore, the uORFs and stable secondary structures of 5'UTR variants of Cx48.5 may provide a switch mechanism by which rapid

alteration of connexin expression can be achieved without the requirement for new transcription.

It was estimated that about 10% of all mRNAs contain atypically long 5' UTRs, and their protein products typically regulate vital cellular processes, including cell growth, cell proliferation, apoptosis, etc (Pickering and Willis 2005). In situations where cap-dependent translation is impaired or suppressed, an alternative mechanism can be used by such long 5'UTRs, as translational initiation can be mediated by a secondary structure called an internal ribosome entry site (IRES) (Kaminski et al. 1994; Jackson 2005). Recent estimates suggest that up to 10% of all cellular mRNAs have the potential to initiate translation via this mechanism (Pickering and Willis 2005). Many of the IRESs identified in cellular mRNAs have complex RNA secondary structures. However, thus far there is no recognizable conservation of secondary structure or sequence among these IRESs. It is hypothesized that many cellular IRESs contain short modules and internal initiation is mediated by the combined effect of these modules (Stoneley and Willis 2004), including a Y-shaped double hairpin structure (Le and Maizel 1997), the sequence GNRA (where N is any nucleotide, and R is a purine) (Hoffman and Palmenberg 1995; Robertson et al. 1999), polypyrimidine tracts, and the sequence AGACA (Huez et al. 1998). IRES-mediated translational initiation has been studied in Cx43 and Cx32 (Schiavi et al. 1999; Hudder and Werner 2000). These examples have interested scientists in the gap junction community to further look into other connexin genes (Werner 2000). As shown in Figure 4-6 to 4-9 (A), 5'UTRs of zebrafish Cx48.5 were predicted to have

stem-loops with extended Y-shaped secondary structures. Looking closely at its 5' UTR primary sequences, we further found "GNRA", "AGACA" semi-conserved sequences and polypyrimidine tracts needed for IRESs. Such short RNA modules in its 5' UTR suggest that zebrafish Cx48.5 might have the potential to use the IRES-mediated mechanism for regulation of translation. In addition, it appears that the second exon of Cx48.5 is essential for the formation of such a Y-shaped structure, as the removal of this exon from the 5'UTRs predicted structures without such a specific structure (Figure 4-6 to 4-9, B). The addition of an extra non-coding exon has been demonstrated to play an important role in translational control for the transcripts encoding proinsulin (Shalev et al. 2002)and neuronal nitric oxide synthase (Wang et al. 1999; Newton et al. 2003). The analysis in this study suggests that the common second exon of Cx48.5 may be important for the potential IRES-mediated translation. In future studies, cistronic reporter assays could be performed to test these hypotheses.

5'RACE experiments detected different sets of transcript variants of Cx48.5 in the lens, retina, and the heart, respectively (Figure 4-3). Real-time RT-PCR also confirmed that different transcript variants of Cx48.5 are indeed expressed in a tissue-specific manner (Figure 4-14). A combination of differential promoter activation and alternative splicing may provide a powerful means for the flexibility to generate complex patterns of gene expression. A number of studies suggested that the translational regulation mediated by the 5' UTR plays a key role in tissue- and developmental stage-specific expression of the gene (Pickering and Willis 2005; van der Velden and

Thomas 1999). For example, different 5'UTR sequences of the NOS1 mRNAs have been implicated in the regulation of their translation (Wang et al. 1999; Newton et al. 2003b). As demonstrated in the present study, alternative promoter usage for both ZfCx48.5 and mCx46 results in variable first exons, and the selection of different 5' splice sites of variable exons also contribute to the specific patterns of connexin gene expression. This plasticity of 5'UTRs may play an important role in the regulation of connexin gene expression in response to a wide variety of intrinsic and extrinsic signals.

Using real-time RT-PCR, expression pattern analyses has shown that zebrafish Cx48.5 is predominantly expressed in the adult lens, and relatively low level expression was also detected in the retina, heart, inner ear, and skeletal muscle. The present study also demonstrated the mRNA expression of mouse Cx46 in the adult lens and retina tissues. Cx46 mRNA or protein has been detected in the adult rat, rabbit, and human heart (Paul et al. 1991; Verheule et al. 2001; Davis et al. 1995). These data suggest that tissue expression patterns of both connexin genes appear to be conserved. Cx46 plays essential roles in lens development and adult lens tissue homeostasis, as the ablation of Cx46 in knock-out mice resulted in the formation of cataracts (Gong et al. 1997). The previous morpholino knock-down of Cx48.5 in this lab reproduced the cataract phenotype seen in the Cx46 knock-out mice. Cx48.5 knock-down also resulted in small retinas as well as severe cardiovascular abnormalities in morpholino-injected zebrafish embryos (Cheng et al. 2004). These two morpholino oligos used in the previous study from our lab were designed using only the genomic sequence of Cx48.5, and in the

present study, the identification of the mRNA sequences of Cx48.5 further confirmed that those morpholino oligos were correctly targeting the Cx48.5 transcripts. The discovery of four types of transcript isoforms of Cx48.5 in the present study now raises questions, such as, why different transcripts are needed, which suggests to us that we should investigate the functional roles of each individual transcript variant during zebrafish embryonic development and in adult tissues. Morpholino knock-down could provide a very powerful strategy for such functional analyses. I have designed a set of morpholino oligos to target the specific individual transcript isoforms, but functional analyses must await the establishment of stable transgenic zebrafish lines that express EGFP-tagged Cx48.5 (for details, see Appendix II).

We tested the Cx48.5 ~12 kb promoter capacity using EGFP transient expression assays in living zebrafish embryos. EGFP expression driven by the Cx48.5 promoter was detected in the lens, retina, and skeletal muscle in early embryos; however, its expression was not seen in the heart. As transient expression of a transgene in the embryos is highly mosaic, it is not surprising that EGFP transient assays cannot reliably reproduce the expression patterns of endogenous genes (Udvadia and Linney 2003). However, these observations certainly suggest that this Cx48.5-eGFP construct contains sufficient regulatory elements needed for driving EGFP expression in the developing lens and retina. Another possible explanation for the failure in detecting EGFP expression in the developing heart may be the weakness of signal. As demonstrated by real-time PCR, the expression of Cx48.5 in the heart is much lower than that seen in the lens. It is also

possible that some key regulatory elements essential for heart-specific expression, such as distal enhancer sequences, might have been missed in this promoter construct. Researchers have demonstrated that endogenous expression patterns could be more perfectly mimicked if using much larger promoter fragments, such as for example, artificial chromosomal vectors (BAC or PAC) (Jessen et al. 1998). Larger constructs could be valuable to search distal enhancers or silencers, however, it is not helpful for mapping regulatory elements within the proximal promoter. Although there are some examples in the literature that *in vivo* GFP transient assays were used to rapidly analyze the *cis*-acting promoter/enhancer elements of specific genes (reviewed in Chapter 1), the intrinsic drawback of this approach (highly mosaic and ectopic expression) makes it very challenging to pool useful expression information from such a mosaic background. In the present study, the original objective of GFP transient expression was to analyze the regulatory elements within the proximal promoter of Cx48.5, however, the analyses here suggested that this approach appears to be much less practical. In our future studies, specific *in vitro* approaches could help address this objective, by using for example, luciferase assays in appropriate cultured cell lines.

4.5 Summary

In this chapter, the transcriptional start sites of both zebrafish Cx48.5 and mouse Cx46 were mapped using 5'RACE techniques. The identification of the full-length mRNA sequences revealed that Cx48.5 has three exons, with an uninterrupted coding

region and 3'UTR located in the third exon. Four types of Cx48.5 transcripts were detected, and they only differ in their first exons, which arise from alternative promoter usage and pre-mRNA splicing. Using real-time RT-PCR, the relative amount of transcript variants of zebrafish Cx48.5 were compared in adult tissues (the lens, retina, and heart), as well as at different early embryonic stages (0.5 dpf, 1 – 5 dpf). Results showed that four isoforms of Cx48.5 mRNA are distributed in adult tissues (the lens, retina, and heart) in a tissue-specific manner. Tissue expression pattern of Cx48.5 was also determined using RT-PCR with a set of adult tissues. Results showed that Cx48.5 is predominantly expressed in the zebrafish lens, while mRNA expression was also detected in other tissues including the heart, inner ear, retina, spleen, and skeletal muscle. Analyses of the sequence and secondary structures of Cx48.5 5'UTR variants showed a number of uORFs, stable loop-stem structures, and potential motifs for IRES-mediated translation throughout the 5'UTR variants. In order to look into the possibility of an evolutionary conservation of connexin gene expression across species, I also performed analyses for mouse Cx46 in parallel. Two isoforms of Cx46 transcripts were detected in adult mouse lenses, whereas only transcript-I of Cx46 was detected in the retina. Transcript variants of Cx46 also differ only in the first exons, which also arise from different promoters and pre-mRNA splicing. A combination of different promoter usage and alternative selection of 5' splice sites are common features of connexin gene expression across species. These results suggest that post-transcriptional regulation could play important roles in connexin gene expression. The promoter activity of zebrafish Cx48.5 was tested using *in vivo* GFP

transient expression assays. Results suggest that a ~12kb upstream fragment can effectively drive eGFP expression in the lens and retina tissues in living embryos. Further mapping the regulatory elements within the Cx48.5 promoter awaits the availability of appropriate *in vitro* approaches. Transcript variant-specific morpholino knock-down analyses have also been proposed for future studies to elucidate the functional roles of individual transcript variants of Cx48.5.

Figure 4-1. Construction of ZfCx48.5-eGFP plasmid. (A) An EcoRI fragment, digested from a Cx48.5 PAC clone, was sub-cloned and screened in a pBluescript II SK +/- vector. (**B**) The above plasmid was modified to remove one extra EcoRI site, into which a NotI-EcoRI fragment produced by PCR from both Cx48.5 genomic DNA and an EGFP expression vector was then cloned. Also see the text for details.

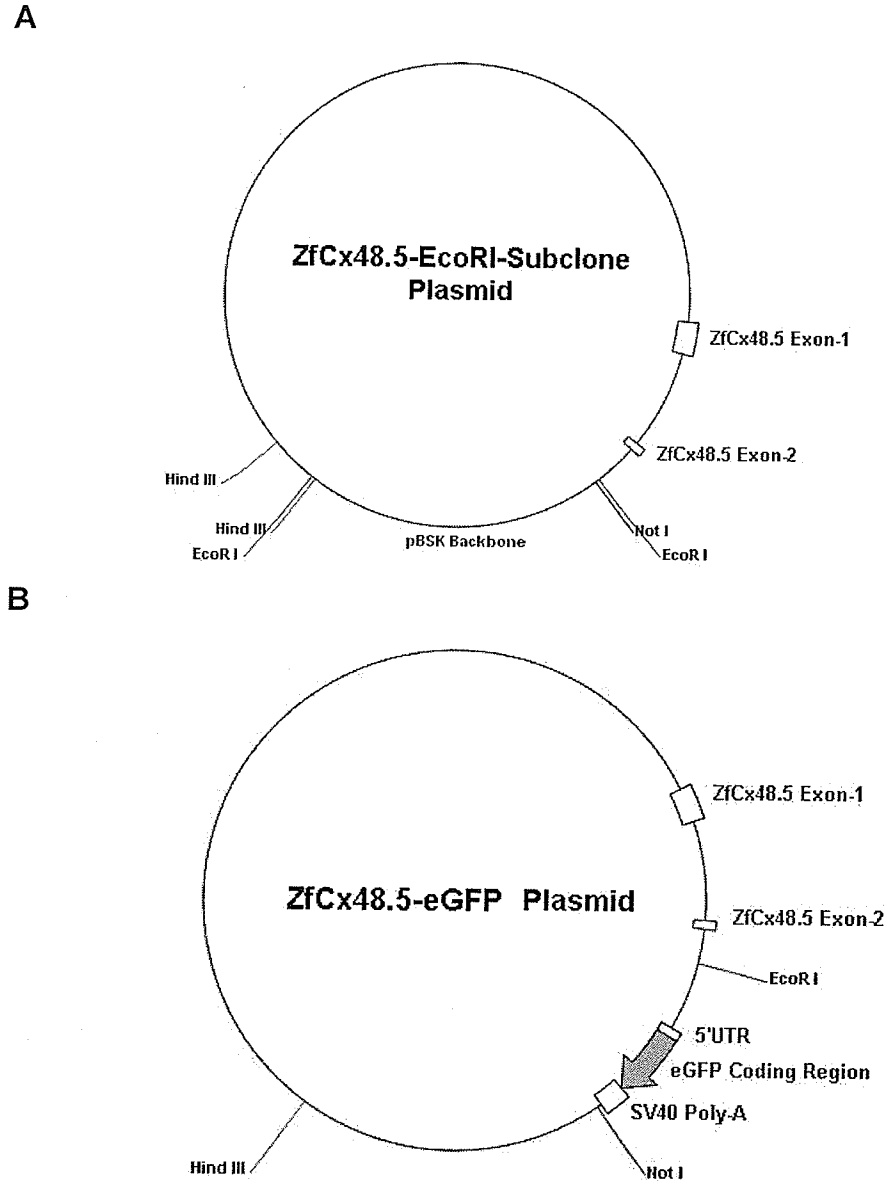


Figure 4-2. Alignment of ZfCx48.5 transcripts against genomic DNA. Cx48.5 is a three-exon connexin gene. Four transcript variants share the same exon 2, exon 3 (mostly coding sequence and 3'UTR), while their first exons are derived from different promoter usage and alternative pre-mRNA splicing. Transcript I is transcribed from P1, whereas IIa, IIb and IIc are from P2. Transcription can start from multiple transcriptional start sites (TSSs, indicated by * above the specific nucleotides). Transcript I and IIc share the same splicing sites, but differ from that of other transcripts (IIa and IIb).

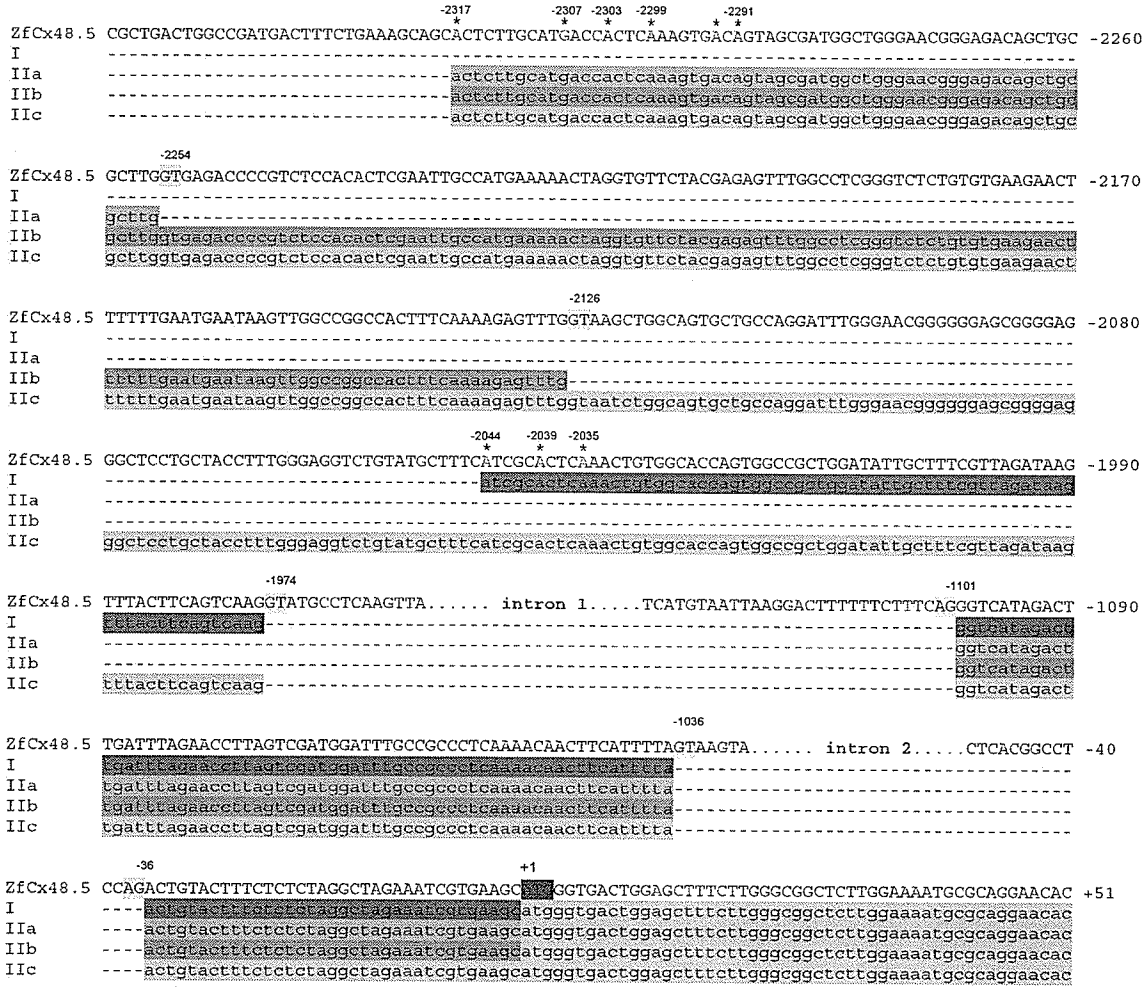


Figure 4-3. Sketch of ZfCx48.5 transcripts and real-time PCR primers. Boxes represent the exons of Cx48.5 transcripts, while bold lines (or blank boxes) are genomic sequences or introns. All four transcripts share the same coding sequence and 3'UTR, whereas alternative first exons are derived from different promoter usage (P1 or P2) and pre-mRNA alternative splicing (alternative 5' sites). 5'UTRs are marked by the same colors as shown in Figure 4-2. The expression of these transcript variants is in a tissue-specific manner among the adult lens, retina and heart. "+" indicates that the expression was detected by 5'RACE method. Real-time PCR primer pairs were designed for targeting specific transcript variants of Cx48.5. Forward primers are indicated by red arrows, while reverse primers by blue arrows. All primer pairs have at least one primer spanning an exon-exon boundary.

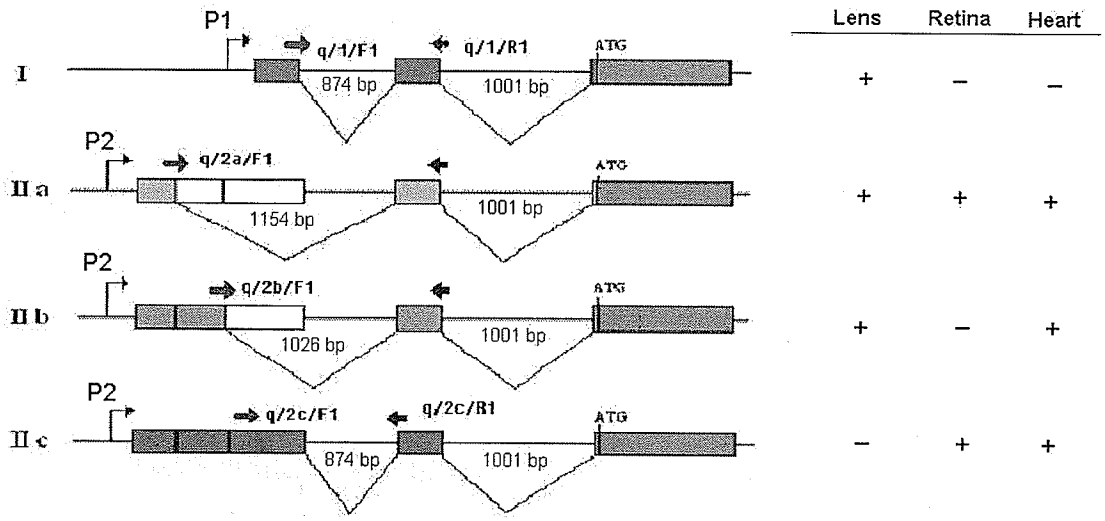


Figure 4-4. Alignment of mCx46 transcripts against genomic DNA. Transcript I was identified in the adult mouse lenses in the present study, while transcript II was not detected by 5'RACE in the lenses and its sequence was retrieved from the GenBank (accession no. AY390398). See the text for details. Transcript variants arise from different promoter usage (P1 and P2) and pre-mRNA splicing (alternative 5' splicing sites). Transcript I can be initiated from multiple transcriptional start sites (TSSs), which are indicated by * above the alignment. Splicing events follow the conserved "GU-AG" rule (highlighted with yellow color). The adenine site in ATG codon was assigned as position +1.

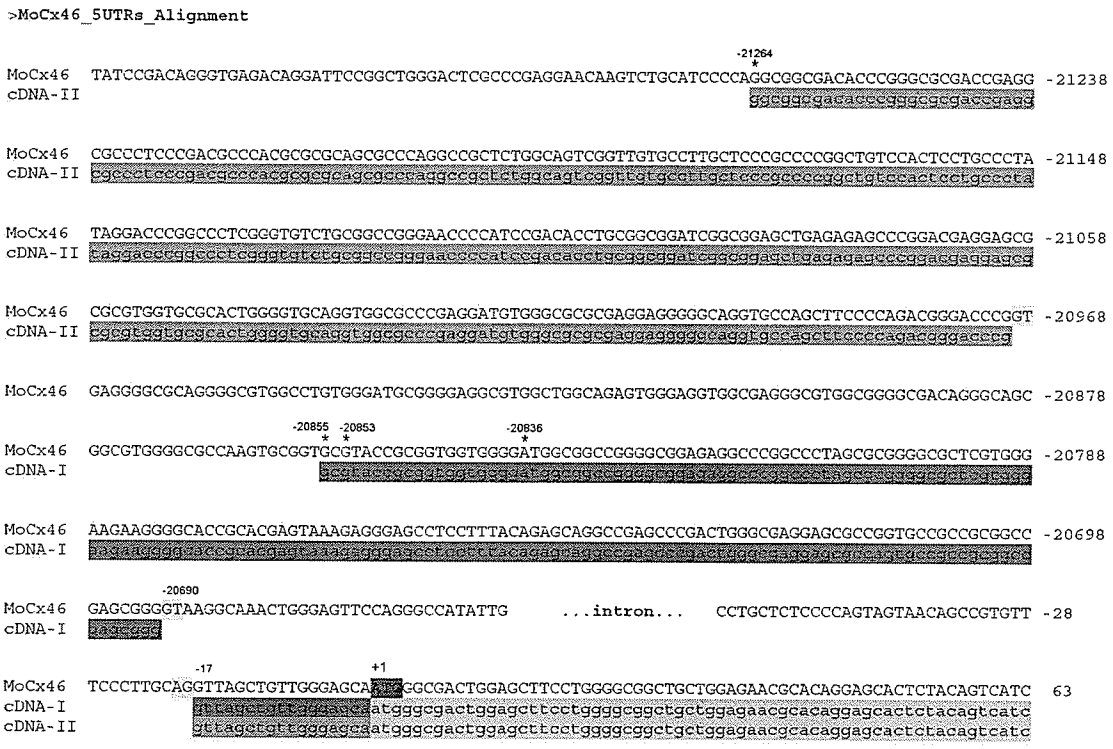


Figure 4-5. Sketch of mCx46 transcripts and real-time PCR primers. Boxes represent the exons of Cx46 transcripts, while bold lines (or blank boxes) are genomic sequences or introns. Both transcripts (I and II) share the same coding sequence and 3'UTR, and alternative 5'UTR are derived from different promoter usage (P1 and P2) and pre-mRNA splicing. 5'UTRs are marked by the same colors as shown in Figure 4-4. Real-time RT-PCR primer pairs were designed for targeting specific transcript variants of Cx46. Forward primers are indicated by red arrows, while reverse primers by blue arrows. Both primer pairs share the reverse primer and the forward primers span the exon-exon boundary.

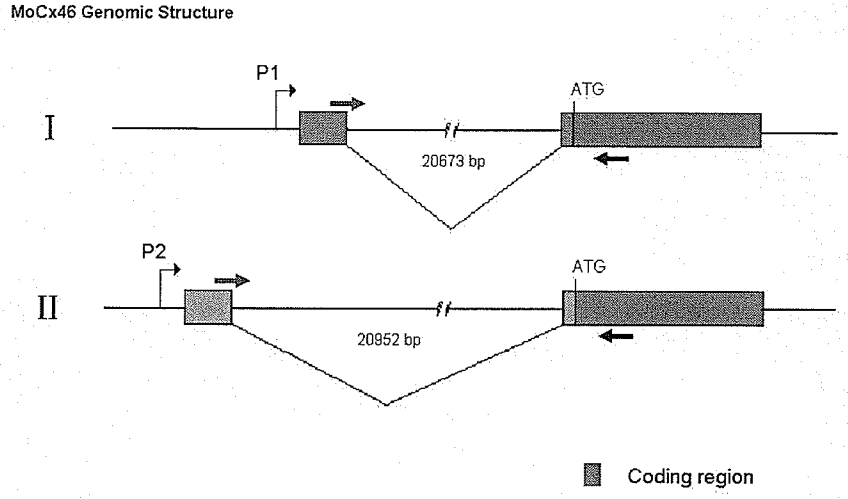
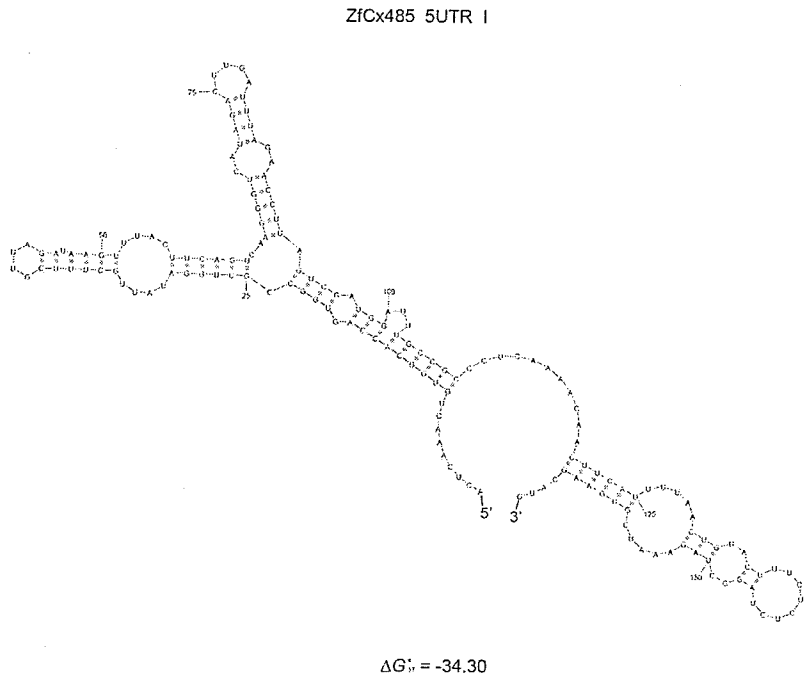


Figure 4-6. ZfCx48.5 transcript-I 5'UTR secondary structure. 5'UTR sequence including exon-2 (**A**) and sequence excluding exon-2 (**B**) were both predicted using Sfold software. The most thermodynamically stable structure is shown with the calculated free energies ($\Delta G°_{37}$ = -34.40 kcal/mol and -20.20 kcal/mol, respectively).

A



B

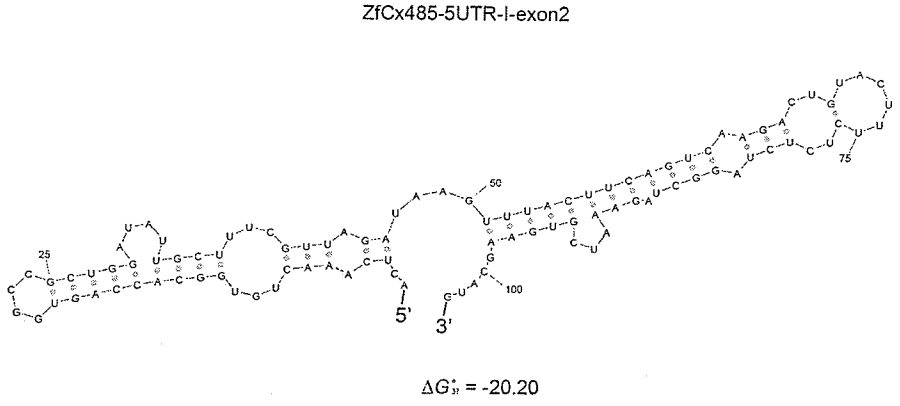


Figure 4-7. ZfCx48.5 transcript-IIa 5'UTR secondary structure. 5'UTR sequence including exon-2 (**A**) and sequence excluding exon-2 (**B**) were both predicted using Sfold software. The most thermodynamically stable structure is shown with the calculated free energy ($\Delta G°_{37}$ = -33.10 kcal/mol and -26.00 kcal/mol, respectively).

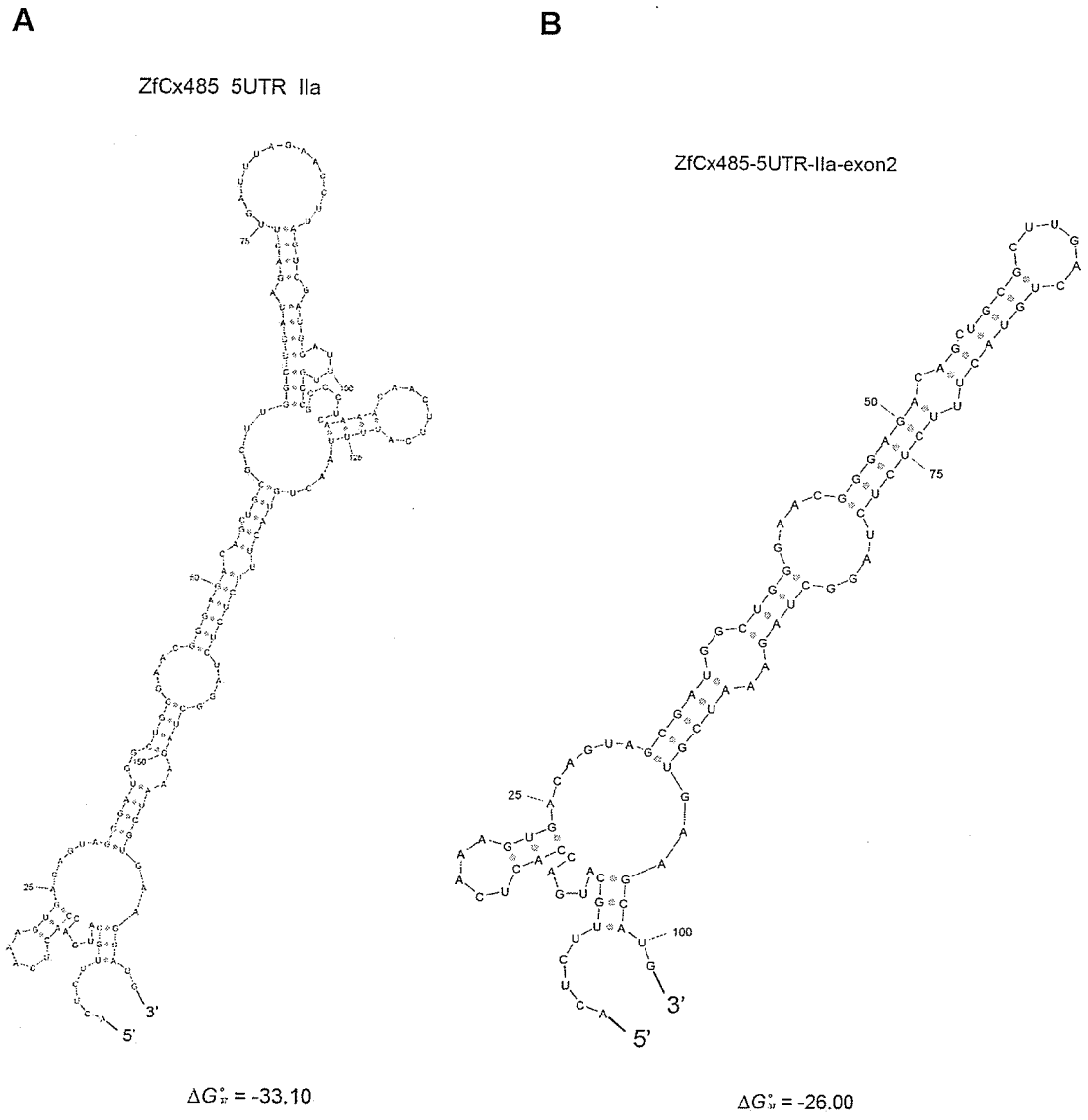


Figure 4-8. ZfCx48.5 transcript-IIb 5'UTR secondary structure. 5'UTR sequence including exon-2 (**A**) and sequence excluding exon-2 (**B**) were both predicted using Sfold software. The most thermodynamically stable structure is shown with the calculated free energy ($\Delta G°_{37}$ = -75.80 kcal/mol and -65.80 kcal/mol, respectively).

A

B

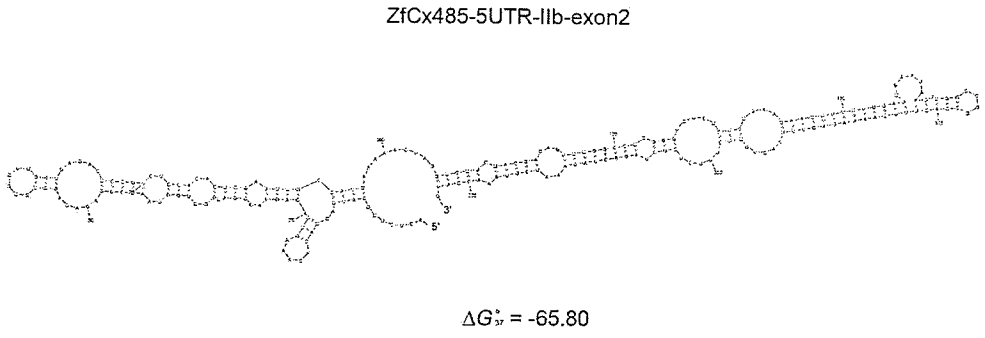
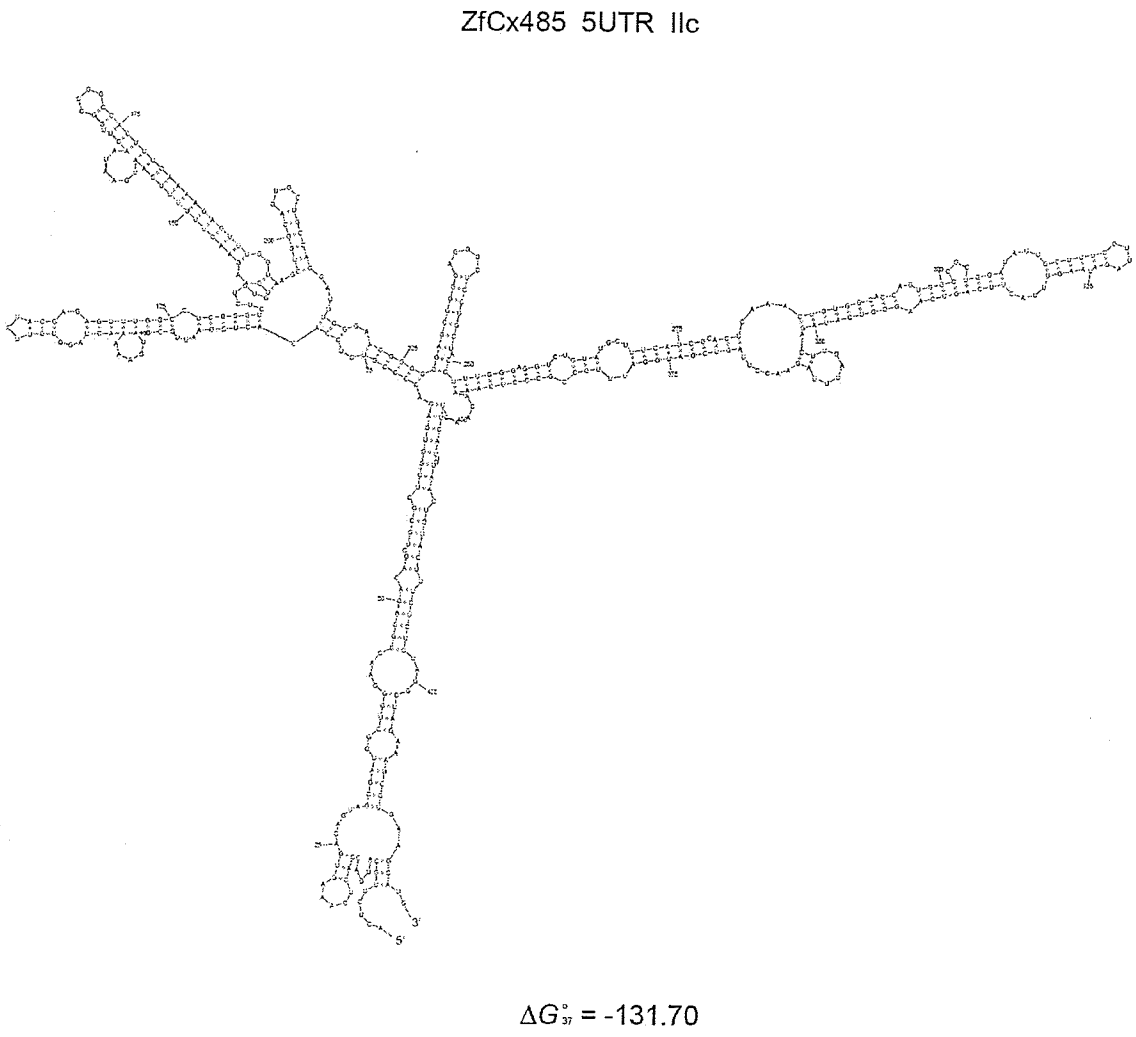


Figure 4-9. ZfCx48.5 transcript IIc 5'UTR secondary structure. 5'UTR sequence including exon-2 (**A**) and sequence excluding exon-2 (**B**) were both predicted using Sfold software. The most thermodynamically stable structure is shown with the calculated free energy ($\Delta G°_{37}$ = -131.70 kcal/mol, and -122.80 kcal/mol, respectively).

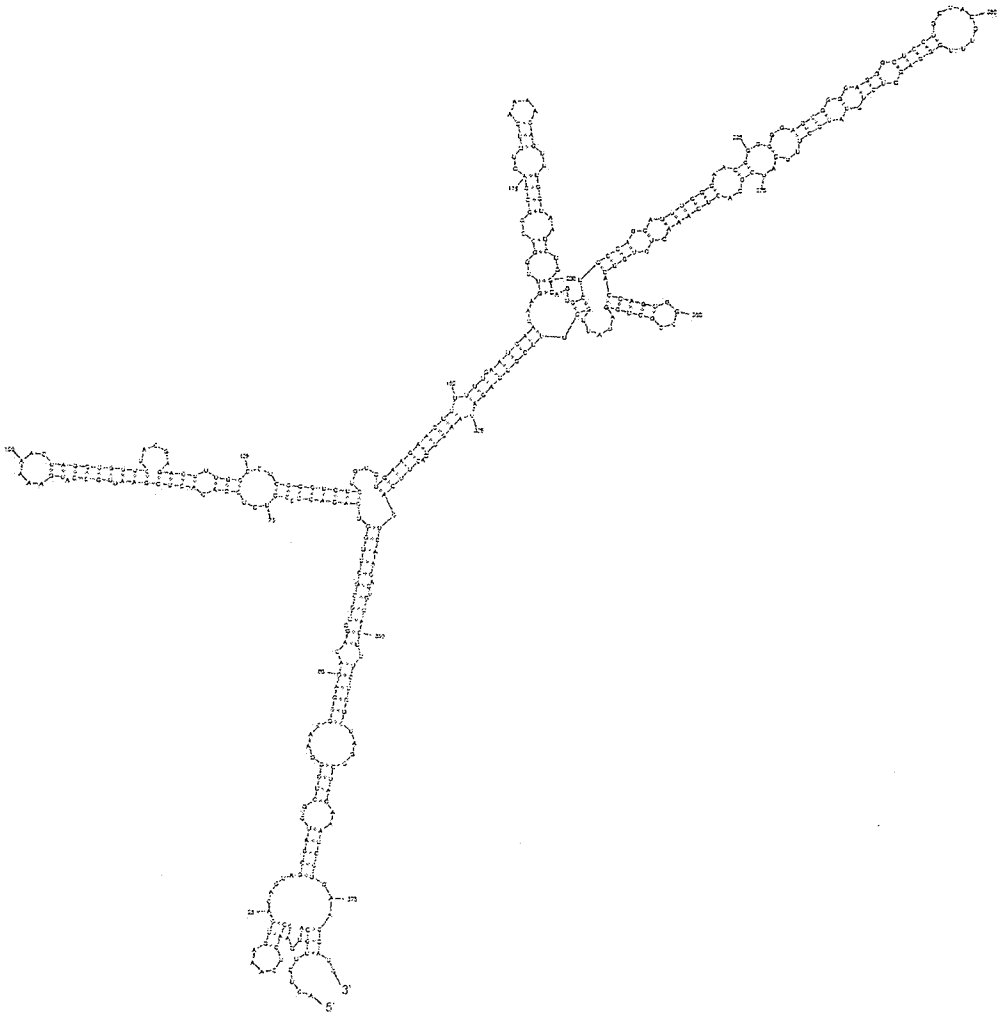


(to be continued on next page)

(continued)

B

ZfCx485-5UTR-IIc-exon2



ΔG°_{37} = -122.80

Figure 4-10. mCx46 transcript-I 5'UTR secondary structure. 5'UTR sequence was predicted using Sfold software. The most thermodynamically stable structure is shown with its calculated free energy ΔG°_{37} = -92.60 kcal/mol.

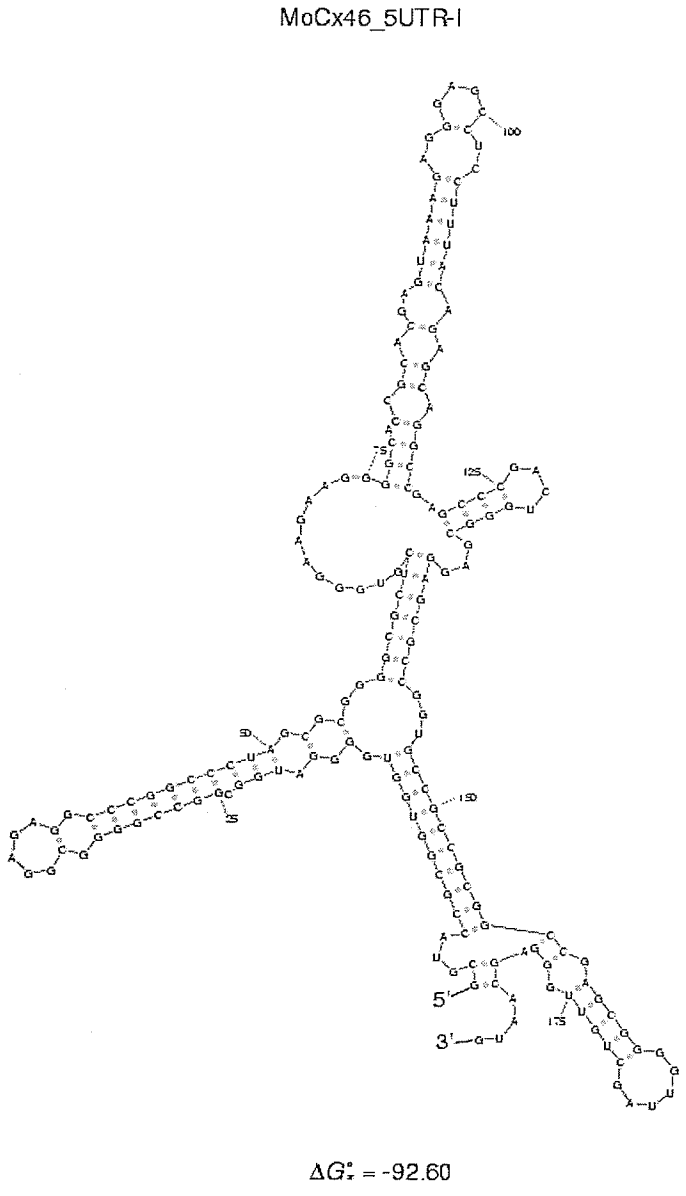


Figure 4-11. mCx46 transcript-II 5'UTR secondary structure. 5'UTR sequence was predicted using Sfold software. The most thermodynamically stable structure is shown with its calculated free energy $\Delta G°_{37}$ = -159.70 kcal/mol.

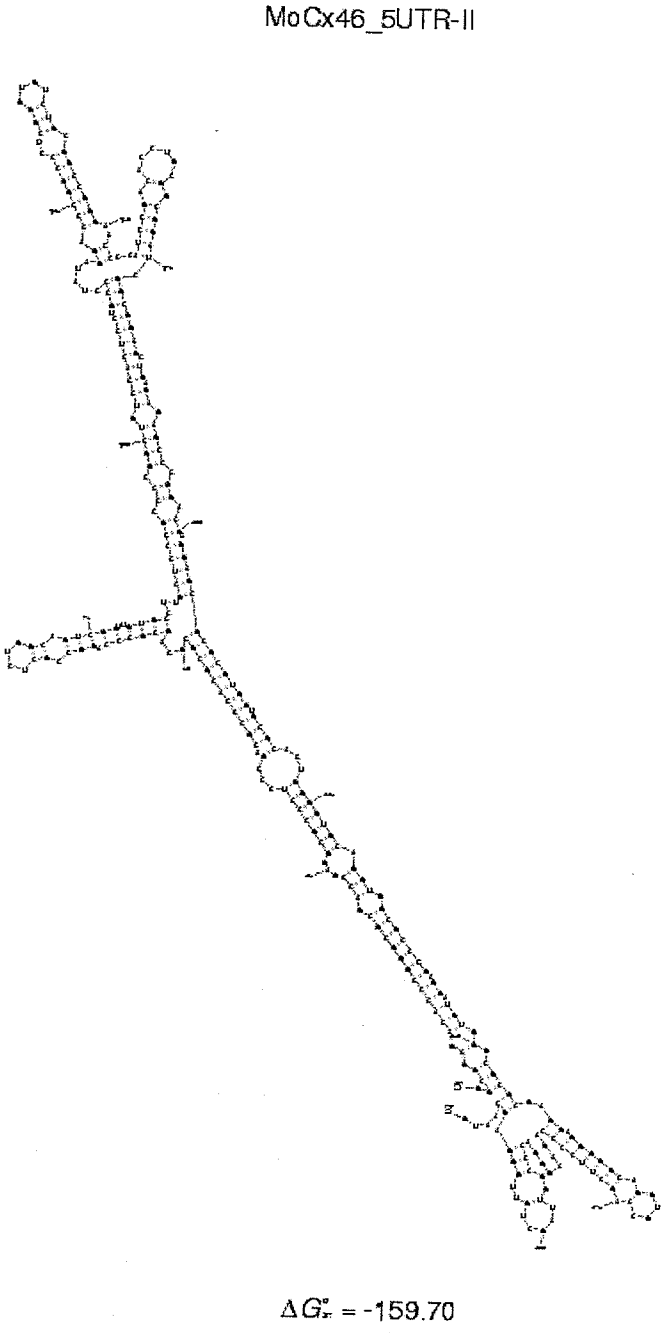


Figure 4-12. uORFs of ZfCx48.5 and mCx46 5'UTRs. The normal translational start codons (AUG) are indicated by green solid triangles. uORFs (upstream open reading frames) of variable length are indicated by red arrows.

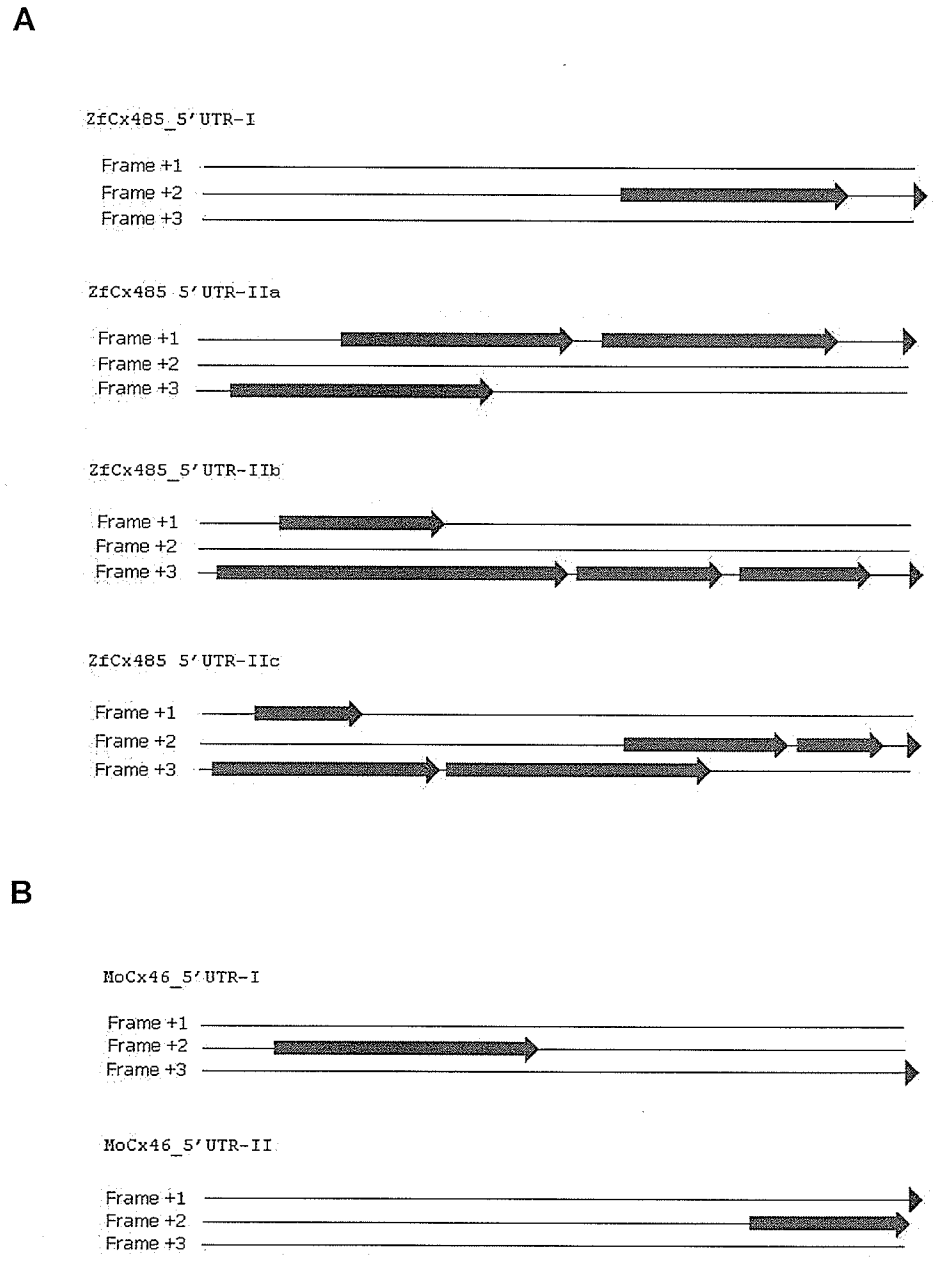


Figure 4-13. ZfCx48.5 3'UTR secondary structure. The most thermodynamically stable structures predicted by mFOLD program are shown with its calculated free energy ($\Delta G°_{37}$ = -445.22 kcal/mol).

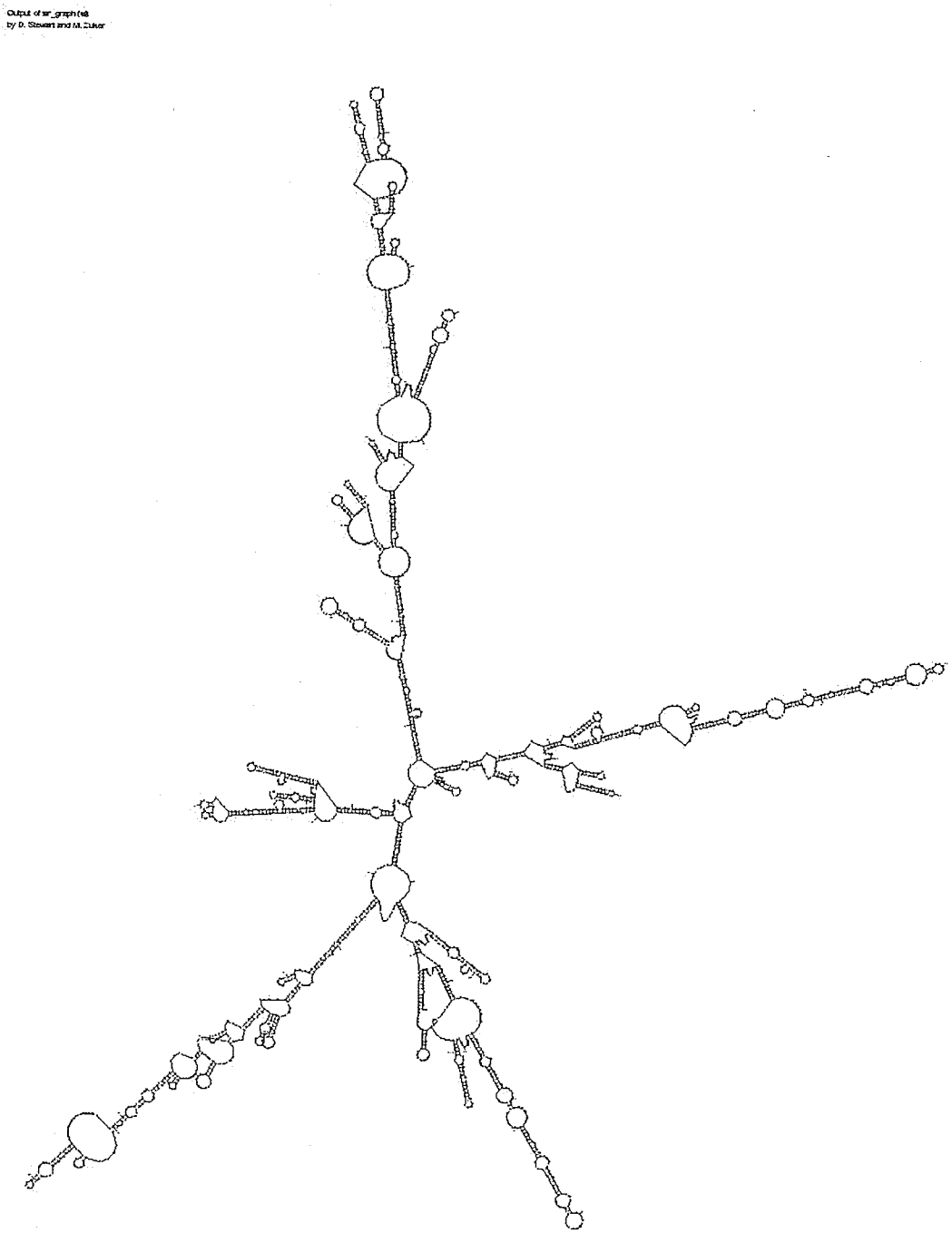


Figure 4-14. ZfCx48.5 transcripts in the lens, retina and heart. Expression of transcript variants is tissue-specific. Three transcript variants are predominantly expressed in the lens, while very low levels of expression for some of these transcript variants were also detected in the retina and heart. Transcript IIb is the most abundantly expressed in the lens, while in the heart, expression of transcript IIc is the most abundant. The expression ratios were calculated by comparing ΔCt between mean Cts of ZfCx48.5 transcripts and zebrafish EF1α mRNA in two replicate experiments, and presented as a column graph.

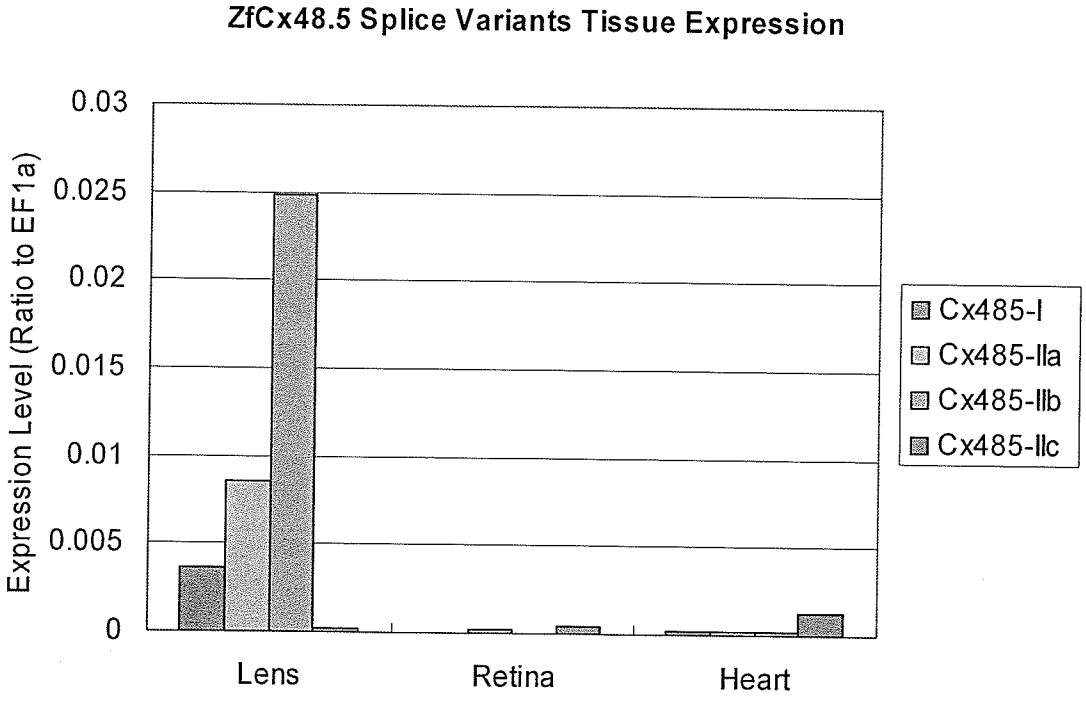


Figure 4-15. mCx46 transcripts in adult lens and retina. Expression of transcript variants is tissue-specific. Transcript I is predominantly expressed in the lens, while lower level of expression was detected in the retina. Transcript II and a mCx50 trans-splicing transcript were not detected in the retina, but with very low level in the lens. The expression ratios were calculated by comparing ΔCt between mean Cts of test and reference genes in two replicate experiments, and presented by a column graph.

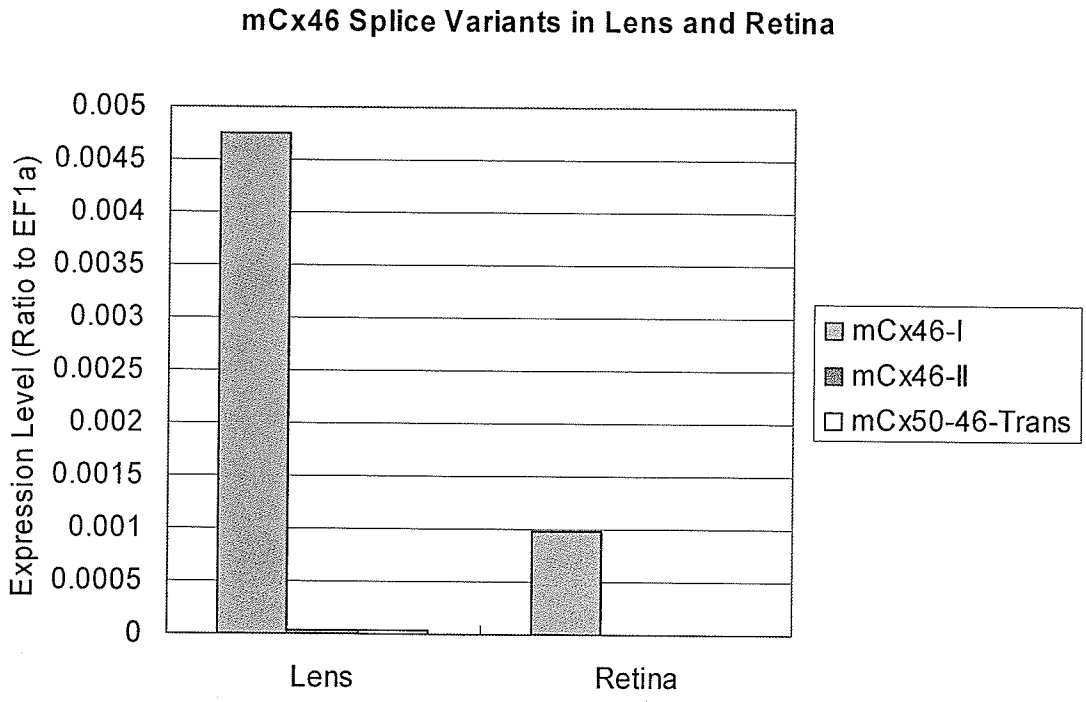


Figure 4-16. ZfCx48.5 transcript variants during development. Analyses were performed using about 50 embryos at different developmental stages indicated by dpf (days post fertilization). Expression of Cx48.5 was not detected before 0.5 dpf, but strong expression was detected from day 1 and later on. The expression ratios were calculated by comparing ΔCt between mean Cts of ZfCx48.5 transcripts and zebrafish EF1α mRNA in two replicate experiments, and presented as a column graph.

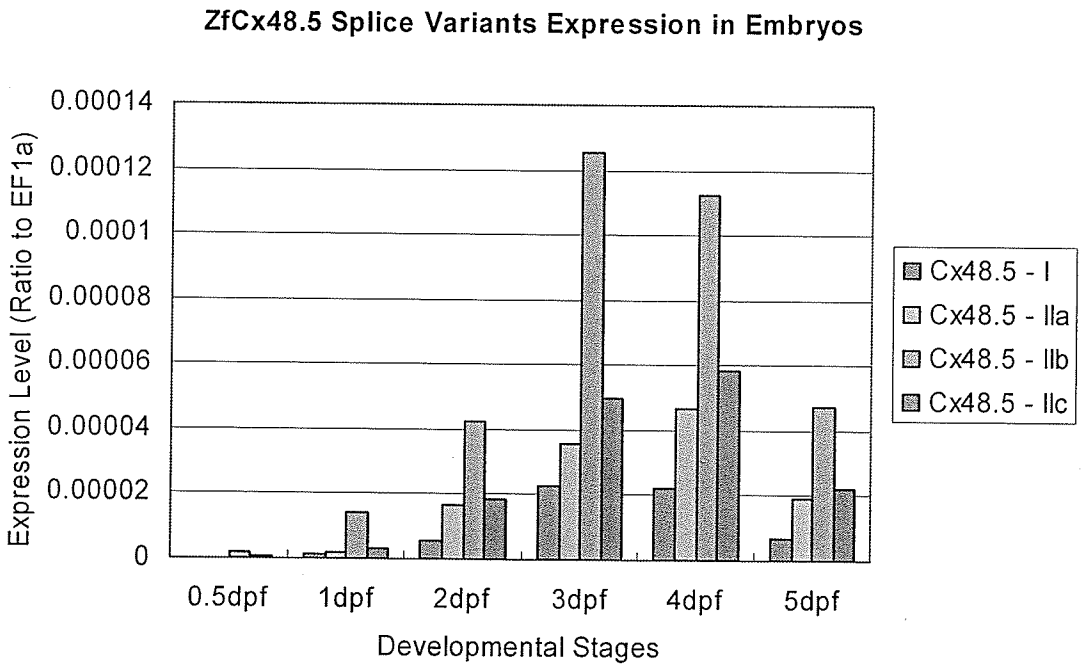


Figure 4-17. Tissue distribution of ZfCx48.5 total mRNAs. Cx48.5 total mRNA (all transcripts combined) is predominantly expressed in the adult lens tissues, while a relatively low level of Cx48.5 expression was also detected in the retina, heart, spleen, inner ear and skeleton muscles. The expression ratios were calculated by comparing ΔCt between mean Cts ofZfCx48.5 and zebrafish EF1α genes in two replicate experiments, and presented as a column graph.

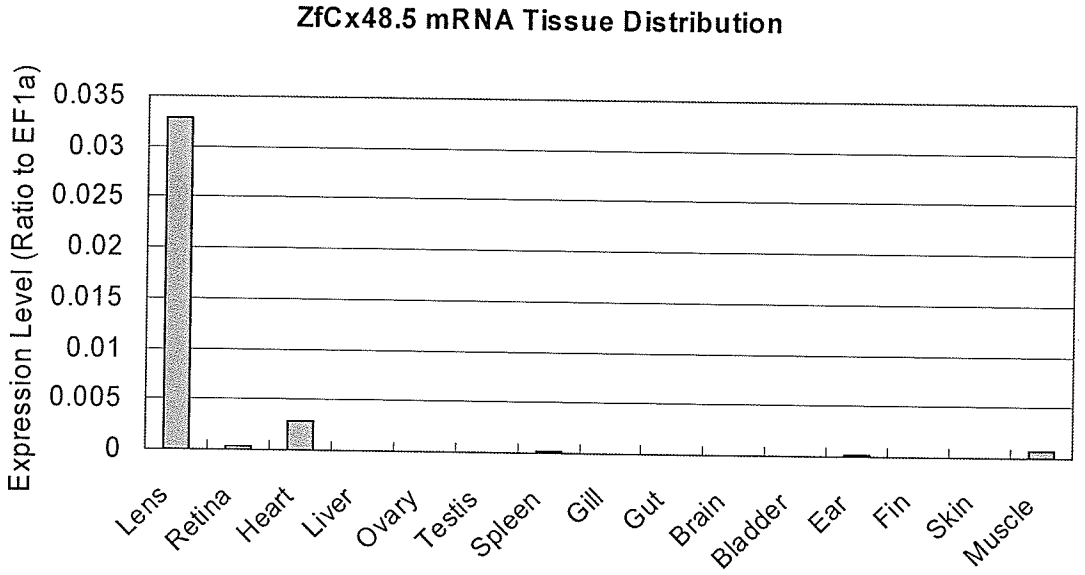


Figure 4-18. ZfCx48.5-eGFP transient expression in zebrafish embryos. A ~ 12 kb promoter fragment immediately upstream of the AUG codon of zebrafish Cx48.5 could drive EGFP transient expression in the lens and retina of fish embryos (**B, D, F**).

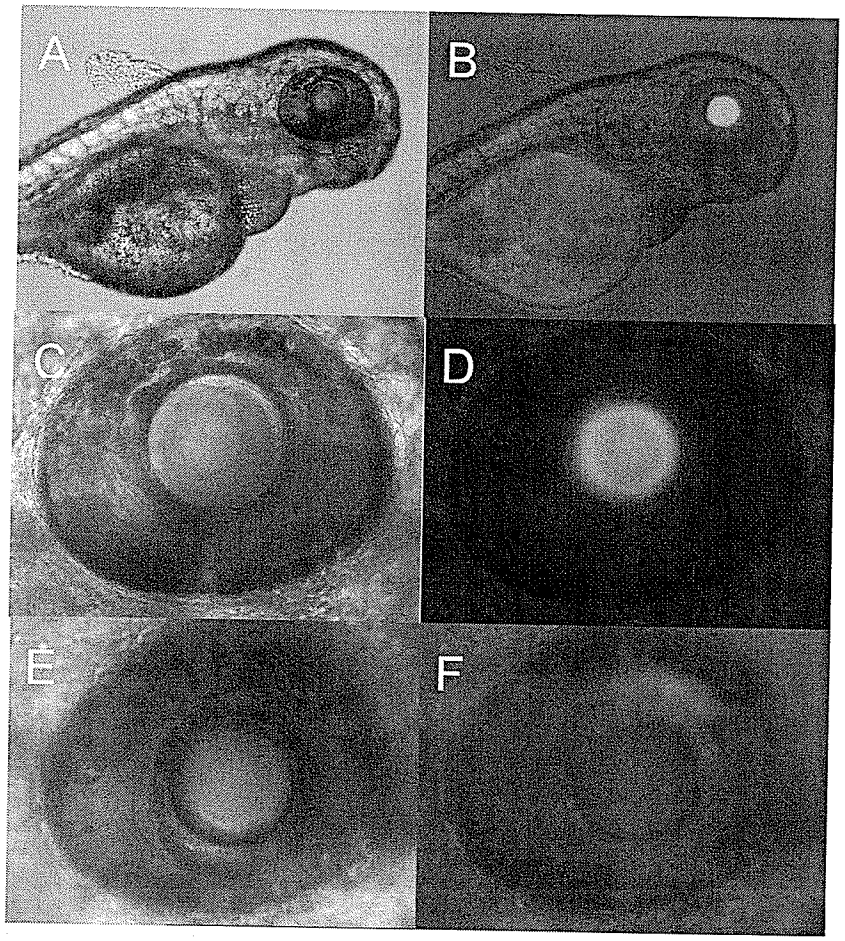


Table 4-1. Primers used for 5' and 3' RACE of ZfCx48.5 and mCx46.

Primer Name	Sequences (5' – 3')
GeneRacer™ RNA oligo	CGA CUG GAG CAC GAG GAC ACU GAC AUG GAC UGA AGG AGU AGA AA
GeneRacer™ 5'Primer	CGA CTG GAG CAC GAG GAC ACT GA
GeneRacer™ 5'Nested Primer	GGA CAC TGA CAT GGA CTG AAG GAG TA
GeneRacer™ 3'Primer	GCTGTCAACGATACGCTACGTAACG
GeneRacer™ 3'Nested Primer	CGCTACGTAACGGCATGACAGTG
Cx48.5 5'RACE Primer	GCG GCA TGTGGGAGAATGGGGAAGGC
Cx48.5 5'RACE Nested Primer	GATCTGGAGCACCCAGAAGCGGATGTG
Cx48.5 3' RACE Primer	CCT TAG TCG ATG G AT TTG CC
Cx48.5 3' RACE Nested Primer	GTG AGA GCG AGG CCC AAT GAT CTG G
Cx46 5'RACE Primer	GGT CGT AGC AGA CGT TCT CAC AGC CTG
Cx46 5'RACE Nested Primer	GGC TGC TGT GTG TTG CAG GTG AAG TC

Table 4-2 ZfCx48.5 and mCx46 real-time PCR primers

Templates	Sequences of Primers (5' – 3')	Amplicon
ZfCx48.5 -I	(F) GAT AAG TTT ACT TCA GTC AAG GCT CAT AG	105 bp
	(R) GCC TAG AGA GAA AGT ACA GTT AAA ATG AAG	
ZfCx48.5 -IIa	(F) AAC GGG AGA CAG CTG CGC TTG GCT CAT AG	105 bp
	(R) GCC TAG AGA GAA AGT ACA GTT AAA ATG AAG	
ZfCx48.5 -IIb	(F) GGC CAC TTT CAA AAG AGT TTG GCT CAT AG	105 bp
	(R) GCC TAG AGA GAA AGT ACA GTT AAA ATG AAG	
ZfCx48.5 -IIc	(F) CTG CTA CCT TTG GGA GGT CTG TAT GC	131 bp
	(R) GGT TCT AAA TCA AGT CTA TGA CCC TTG AC	
zEF1α cDNA	(F) GAG GAG TGA TCT CTC AAT CTT GAA AC	94 bp
	(R) TTT CCG GAG TCG ACG TGG CCA ATA AC	
zEF1α cDNA	(F) GAG GAG TGA TCT CTC AAT CTT GAA AC	133 bp
	(R) CCA CCG CAT TTG TAG ATC AGA TGG	
mCx46 cDNA-I	(F) CGC GGC CGA GCG GGG TTA GCT G	143 bp
	(R) CTA ACA CCA GAA TGC GGA AGA TGA ACA G	
mCx46 cDNA-II	(F) TTC CCC AGA CGG GAC CCG GTT AGC	147 bp
	(R) CTA ACA CCA GAA TGC GGA AGA TGA ACA G	
mCx50-mCx46 Trans-splicing cDNA	(F) GTG GGA CGC TAC TCA GAC GTG AGC	148 bp
	(R) CTA ACA CCA GAA TGC GGA AGA TGA ACA G	
mEF1α1 cDNA	(F) GTT TGC CGT CAG AAC GCA GGT GTT GT	147 bp
	(R) CAC ATT TGT AGA TCA GGT GGC CGG TTG	

Table 4-3. PCR primers used for ZfCx48.5-eGFP construct.

Primer Name	Sequences (5' – 3')
ZfCx485/PF9	CCT TAG TCG ATG GAT TTG CC
ZfCx485/gfp/R1	TCC TCG CCC TTG CTC ACC ATA GCT TCA CGA TTT CTA GCC TAG AG
gfp/F1	TGG CCT CCT CCG AGG ACG TCA TC
gfp/R1	TAG ATG AAG GAG CCG TCT TGC AGG G

CHAPTER FIVE

ZEBRAFISH CONNEXIN 30.3

5.1 Introduction

This lab has previously cloned and characterized several zebrafish connexin genes with an aim to unravel the molecular mechanisms regulating the expression and function of connexin genes in early embryonic development (Cason et al. 2001; McLachlan et al. 2003; Christie et al. 2004; Cheng et al. 2004). As a part of this effort, the isolation of zebrafish Cx48.5 and its functional analysis using morpholino knockdown were previously reported (Cheng et al. 2004). As described in Chapter 4, I subsequently analysed its 5' and 3'UTR sequences, as well as its transcriptional start sites further. Cx48.5 is a three-exon connexin gene with the 5' UTR located in alternative first exons and a common second exon, while the entire coding and 3' UTR sequences are located in the third exon. Notably, differential promoter usage and alternative splicing result in variable first exons of Cx48.5, which suggests tight regulatory mechanisms for the expression of Cx48.5 mRNAs, both at the transcriptional and post-transcriptional levels (Kornblihtt 2005). In order to analyze the regulatory elements essential for the transcription of Cx48.5, further upstream sequence information was required. Since the zebrafish genome sequence project had not been completed at that stage, I further obtained positive EcoRI sub-clones by screening the PAC clones previously isolated in this lab (Cheng et al. 2004). One of the positive subclones containing the first two exons

of Cx48.5 and a ~12 kb upstream fragment was selected for promoter studies (see Chapter 4, Figure 4-1, A). Interestingly, sequencing the upstream promoter region of Cx48.5 resulted in the discovery of a new zebrafish connexin gene, and we named it Cx30.3, according to the common Cx molecular weight nomenclature (Beyer et al. 1990). Sequence analyses suggested that zebrafish Cx30.3 is in fact the same gene that was previously deposited in the GenBank as Cx33.8 (Accession numbers: AY135446, BC095350, BC165699, NM212825). Phylogenetic analyses suggested that Cx30.3 is the zebrafish ortholog of mammalian Cx26.

Although Cx26 was first identified in mouse liver (Janssen-Timmen et al. 1983), most reports suggest that it is predominantly expressed and functions in the inner ear and the skin, as mutations in Cx26 cause defects of hearing and disorders in the skin (Gerido and White 2004; Lefebvre and Van De Water 2000; Richard 2000; Rabionet et al. 2002). My hypothesis is that zebrafish Cx30.3 has similar expression patterns and functional roles in zebrafish embryonic development as that of Cx26 in mammals. In this chapter, comprehensive analyses of the sequences and expression of Cx30.3 mRNA will be presented. My results suggest that zebrafish Cx30.3 and mammalian Cx26 have similar expression patterns, and that they may act in similar functions during embryonic development and in adult inner ear and skin tissues. Taking advantage of the excellent zebrafish model system, I suggest that zebrafish Cx30.3 can be employed to model the role of Cx26 in the embryonic development of human inner ear and skin, as well as the mechanisms for Cx26-related human diseases.

5.2 Materials and Methods

5.2.1 Fish care.

Wild type zebrafish (*Danio rerio*) were purchased from a local pet store and maintained in an aquarium facility at 28.5 °C on a daily photoperiod of 14 h light and 10 h dark. The embryos were incubated in egg water (60 μg/ml Instant Ocean sea salts) in a 28.5 °C water bath, and staged using time elapsed since fertilization. Embryos for EGFP expression studies and for *in situ* hybridization studies were treated with 0.003% 1-phenyl-2-thiourea (Sigma, Mississauga, ON, Canada) in egg water beginning at ~24 hpf (hours past fertilization) to prevent pigment formation. Zebrafish maintenance, breeding and embryo collection followed standard protocols (Westerfield, 1995). The use of experimental animals was approved by the University of Manitoba Institutional Animal Care Committee.

5.2.2 Cloning and sequence analysis

By screening a Zebrafish PAC library (RZDP Deutsches Ressourcen-Zentrum für Genomforschung GMbH, Berlin), PAC clones positive for Cx48.5 were previously isolated in this lab (Cheng et al. 2004). Cx48.5 positive PAC clones were digested with EcoRI and PstI, respectively, and the resulting fragments were subcloned into a pBluescript II SK +/- vector (Stratagene, La Jolla, CA). The subclones were screened

with a ^{32}P-labelled Cx48.5 probe (located in exon 2) and several positive subclones were obtained. One of these EcoRI subclones, carrying a ~12 kb fragment upstream of the translational start site, was selected for further analysis. This subclone was sent to the DNA Center at the University of Calgary for sequencing. The assembled sequence was used in a BLAST search of GenBank and resulted in the discovery of a new zebrafish connexin – Cx30.3. The predicted amino acid sequence was analyzed by tools on the ExPASy server (http://www.expasy.org/tools/protparam.html) (Gasteiger et al. 2003), and transmembrane domains were predicted by software on the "DAS Transmembrane Prediction" server (http://www.sbc.su.se/~miklos/DAS/) (Cserzo et al. 1997). Amino acid sequence similarity analysis was performed by aligning the predicted Cx30.3 amino acid sequence against human and mouse Cx26, using the Kalign program (Kalign web server http://www.ebi.ac.uk/Tools/kalign/) (Lassmann and Sonnhammer 2005).

5.2.3 Phylogenetic analyses and multiple sequence alignment

To assess which sequences were most closely related to Cx30.3, a phylogenetic analysis was performed on protein sequences aligned using Kalign with default parameters. Connexin sequences were obtained from GenBank (http://www.ncbi.nlm.nih.gov/). A neighbour-joining tree was created to determine which connexin sequences most closely matched the Cx30.3 sequence.

5.2.4 RNA extraction

RNA was isolated from zebrafish embryos and adult fish tissues using Trizol reagent (Invitrogen Corp., Carlsbad, CA) according to the manufacturer's instructions. RNA concentration was determined by UV spectrometry. The A260/A280 ratios were between 1.8 and 2.0 for all samples. The quality of RNA samples was also tested by RNA gel electrophoresis. For RT-PCR, a DNase I treatment was performed to remove residual genomic DNA contamination, whereas for real time PCR and RACE-PCR experiments, no DNase treatments were performed in order to avoid as much as possible any loss and degradation of RNA.

5.2.5 RACE-PCR and sequencing of UTRs

5' and 3'RACE (rapid amplification of cDNA ends) were performed using a GeneRacer Kit (Invitrogen) according to the manufacturer's recommended protocols. In brief, total RNA samples were treated with calf intestinal phosphatase (CIP) in order to de-phosphorylate the 5'ends of truncated mRNA and non-mRNA molecules. The dephosphorylated RNA samples were then treated with tobacco acid pyrophosphatase (TIP) to remove the 5' cap structure from mature full-length mRNA, leaving a 5' phosphate group available. Using T4 RNA ligase, a universal RNA oligonucleotide (provided by the manufacturer, 5'- CGA CUG GAG CAC GAG GAC ACU GAC AUG GAC UGA AGG AGU AGA AA -3') was then ligated to the de-capped mRNA, providing a priming site for the following PCR reactions. Total RNA samples for both 5' and 3' RACE were prepared from adult zebrafish retina tissue. Reverse transcription was

carried out with SuperScript III reverse transcriptase and the GeneRacer oligo dT primer (5'- GCT GTC AAC GAT ACG CTA CGT AAC GGC ATG ACA GTG $(T)_{24}$ -3'). For 5'RACE, the gene-specific reverse primer used was 5'-TCC CAT GAT GGC GGA GTT GAA G-3' for the first round of PCR and 5'-GCT GCT ATG ACC AGG ATG CAA ATG-3' for the nested PCR. For 3'RACE, the gene-specific forward primers used was 5'-ATC TCA CTC TGA AGA CCT CAG-3' for the first round of PCR and 5'-GAG AAA ATG TCC ACA GAA AAG GCG-3' for the nested PCR.

The nested PCR products were cloned into the pCR4-TOPO vector (Invitrogen). The ligation mixture was then used to transform One Shot® TOP10 chemically competent *E. coli* cells (Invitrogen). Several positive colonies were randomly picked up and grown in LB medium at 37°C with shaking (250 rpm) overnight. Plasmids were extracted from 3 clones for each of the 5' and 3'RACE, and sent out for sequencing (DNA Centre, University of Calgary).

5.2.6 RT-PCR

Two separate sets of total RNA samples from embryos at different developmental stages (0.5, 1, 2, 3, 4, 5 dpf) were isolated for RT-PCR experiments. For each sample, 3 µg of total RNA was treated with DNase I and then reverse transcribed using Superscript II reverse transcriptase (Invitrogen) with oligo-$dT_{(18)}$ primers (Invitrogen) for 1 hr at 52°C in a reaction of 20µl. The resulting cDNA was diluted to 50µl with ddH_2O, and 1µl of cDNA was used in each 25µl PCR reaction. PCR primers were designed to span the

intron of Cx30.3 in order to discriminate RT-PCR products from genomic contamination. A forward primer (5'-ATC TCA CTC TGA AGA CCT CAG-3') that anneals in the first exon and a reverse primer (5'-GCT GCT ATG ACC AGG ATG CAA ATG-3') that anneals in the second exon amplified a 252 bp fragment from the cDNA template. The gene encoding elongation factor 1α (EF1α) was also amplified as a positive control to assess the presence and integrity of the cDNA in the samples. Accordingly, EF1α primers (forward: 5'-CAA GGG CTC CTT CAA GTA CGC CTG-3', and reverse: 5'-GGC AGA ATG GCA TCA AGG GCA-3') were designed to span an intron, and this primer pair amplifies a 569 bp fragment from a cDNA template or a 733 bp fragment from genomic DNA (Christie et al. 2004; Cheng et al. 2004).

5.2.7 Real-time RT-PCR

The level of Cx30.3 expression at different developmental stages was assessed using real-time RT-PCR, using an iQ5 Real-Time PCR detection system (Bio-Rad). Without DNase pre-treatments, 3 μg of total RNA were reverse transcribed by Superscript III (Invitrogen), with a mix of random hexamers and oligo-$dT_{(18)}$, for 1 hr at 52°C, in a reaction volume of 20μl. The resulting cDNA was then diluted to 100μl with ddH_2O. The Cx30.3 primers used for real-time PCR were the same as those for RT-PCR (described above, the predicted amplicon is 252 bp in size). Standard-curve tests were performed according to general real-time PCR guidelines to validate the specificity and efficiency of this primer pair. The zebrafish elongation factor 1α (EF1α) was used as a

reference cDNA, using primers designed to span the first intron splicing boundary of EF1α gene, so as to avoid amplifying genomic sequences in real-time PCR reactions. The forward primer was 5'-GAG GAG TGA TCT CTC AAT CTT GAA AC-3', and the reverse primer was 5'-CCA CCG CAT TTG TAG ATC AGA TGG-3' and the amplicon was 133 bp in length. The specificity and efficiency of the Cx30.3 and EF1α primers were validated by standard-curve methods according to the manufacturer's specifications (Bio-Rad).

cDNA samples (1μl per well) were mixed with 250 nM of each primer and 10μl of SYBR Green Supermix (Bio-Rad) reagent in a final volume of 20μl. PCR conditions were as follows: 95°C for 3 min followed by 40 cycles at 95°C for 10 sec, 62°C for 10 sec, and 72°C for 10 sec. After amplification, melting curve analysis was employed to check PCR specificity, and/or monitor genomic DNA contamination. In addition, a reverse transcriptase negative control and a template-negative control were also tested in parallel. Each sample was tested in at least duplicate reactions and all PCR runs were performed twice with cDNAs prepared from independent batches of RNA samples. The relative amounts of Cx30.3 gene transcripts were normalized to the EF1α reference gene, using a modified comparative C_t method (Pfaffl 2001).

5.2.8 *In situ* hybridization

Digoxigenin-labelled anti-sense and sense RNA probes for Cx30.3 were synthesized by *in vitro* transcription with DIG-11 UTP (Roche, Mississauga, ON). The

DNA template was from one of the 3'RACE products of Cx30.3 and contained the entire 3'UTR region plus a 94 bp C-terminal fragment. The resulting anti-sense and sense RNA probes were about 1.4 kb in size. The *in situ* hybridization (ISH) was carried out with various developmental stages of embryos between 1 and 5 dpf (days post fertilization), following Thisse's protocol with slight modifications (Christie et al. 2004; Thisse et al. 1993). Embryos were mounted in glycerol for whole mount observations. To obtain plastic sections, the embryos were embedded in JB-4 resin (Polysciences; Warrington, PA) and 5 μm sections were cut with a microtome. The embryos were observed, and all images were recorded using a Zeiss Axioskop FS microscope equipped with a Sony DXC-950 CCD camera and Northern Eclipse software (Empix, Mississauga, ON).

5.2.9 Functional analysis / Electrophysiology

Construction of Cx30.3 expression plasmid

The Cx30.3 coding sequence was PCR amplified using Platinum Taq DNA polymerase High Fidelity (Invitrogen) with primers carrying BamHI and EcoRI restriction sites (forward primer 5'-GGT TAA TTG GAT CCG ACA GGA TGA GTT GGG GAG CAC-3', and reverse primer 5'-CCA TGA TGG CGG AAT TCA AGA TCA TC-3'). The PCR products were cut with BamHI and EcoRI and cloned into a pCS2+ expression vector (Turner and Weintraub 1994). The accuracy of the DNA primary sequence and its open reading frame (ORF) were confirmed by sequencing.

Electrophysiological analyses

Analysis performed by Dr. T. W. White. For some of the results see Appendix-IV.

5.3 Results

5.3.1 Identification of the genomic and amino acid sequences of Cx30.3

DNA sequencing of a 12 kb DNA fragment upstream of the Cx48.5 gene revealed the presence of another gene. A BLAST search indicated that this new gene matched a putatively- identified connexin gene, referred to as Cx33.8. (accession No. AY135446; Kremer, M. and Dermietzel, R., 2002, unpublished). As this Cx33.8 gene had not been fully described, I chose to examine my novel clone further by performing 5' and 3' RACE to identify the full-length gene. Based on these analyses, which I describe more thoroughly below, it appears that the predicted translation start site of the Cx33.8 gene was misidentified, and instead, this connexin gene contains a coding region of 804 nucleotides, coding for a protein of 267 amino acids (Figure 5-1). The connexin therefore has a predicted molecular weight of 30,323 Da (protein analysis tools on the ExPASy server http://www.expasy.org/tools/protparam.html) (Gasteiger et al. 2003). Hereafter, I will refer to the gene and its protein as Cx30.3, according to the conventional Cx molecular weight nomenclature (Beyer et al. 1990).

Using transmembrane domain prediction software ("DAS" Transmembrane Prediction server http://www.sbc.su.se/~miklos/DAS/)(Cserzo et al. 1997), four potential transmembrane regions were predicted (Figure 5-2). The amino acid sequence of Cx30.3 has a similar structure to that of other connexins – with four transmembrane domains,

two extracellular loops, one intracellular loop, and cytoplasmic N- and C- terminals. Each of the two extracellular loops contains three conserved cysteines (Figure 5-2). These features of the amino acid sequence of Cx30.3 are consistent with that of a typical connexin (Unger et al. 1999; Yeager et al. 1998; Rahman and Evans 1991; John and Revel 1991; Dahl et al. 1992). Further analysis of the amino acid sequence revealed that Cx30.3 has a relatively high isoelectric point (pI = 9.54) (ExPASy protein analysis tools) (Gasteiger et al. 2003).

Multiple DNA sequence comparisons suggested that Cx30.3 is most closely related to mammalian Cx26 (data not shown), as has been shown by other researchers (Eastman et al. 2006; Cruciani and Mikalsen 2006; Cruciani and Mikalsen 2007). A comparison of Cx30.3's amino acid sequence with its most similar connexins reveals a high degree of identity with human and mouse Cx26 proteins (Figure 5-2). We therefore suggest that Cx30.3 is the zebrafish ortholog of mammalian Cx26.

5.3.2 Identification of the transcriptional start sites and UTRs of Cx30.3

Using the 5'RACE method, we identified the transcription start site and 5' UTR sequence of zebrafish Cx30.3. This is a RNA ligasemediated (RLM)- 5'RACE strategy that ensures that only cDNAs produced from mature capped mRNA transcripts are selected for the following PCR amplification, and thus only the full-length of the 5' ends can be amplified and sequenced. This feature of the method makes it possible to identify the transcriptional start sites when the full-length 5' UTR sequence is aligned against its

genomic sequence. The 5' RACE nested PCR generated a single clear band of about 250 bp in size. For 3' RACE, the nested PCR produced a single clear band of about 1.4 kb in size. The 5' and 3' RACE-PCR fragments were then cloned into pCR4-TOPO vector (Invitrogen, CA) and three independent positive clones were sequenced for each of the 5' and 3' UTRs. Sequencing results were then aligned against the genomic DNA sequence, resulting in the identification of the transcriptional start site and the intro-exon genomic structure of zebrafish Cx30.3 (Figure 5-5).

The 5' RACE results identified a single transcriptional start site for ZfCx30.3, which is located at position -1760 nts upstream of the ATG translation start codon. For convenience of description, the adenine site in the ATG codon was assigned as position +1. The 5' UTR of Cx30.3 is interrupted by an intron of 1627 nucleotides in size, which starts from position -1561 nt to position -24 nt upstream of the ATG codon. Pre-mRNA splicing uses the conserved "GU-AG" rule (Figure 5-1). The 3' UTR is of 1308 nucleotides in size, which starts from position +805 nt, and ends at position +2113 nt downstream of the ATG codon, a couple of nucleotides following the conserved hexamer "AAUAAA" poly (A) signal (Figure 5-1) (Nag et al. 2006). In these 5' and 3' RACE experiments, mRNA samples were from zebrafish adult retina tissue. The 5' and 3' RACE data suggest that Cx30.3 gene contains two exons, with an intron interrupting its 5' UTR. The complete coding region and 3' UTR is located in exon 2.

Using mFOLD software (Zuker 2003), we predicted the secondary structure of the 5' UTR and 3'UTR of zebrafish Cx30.3 (Figure 5-6, and 5-7). The most

thermodynamically stable structures predicted by mFOLD are shown with calculated free energy indicated in the figures ($\Delta G°_{37}$ = -36.4 kcal/mol for 5'UTR, and $\Delta G°_{37}$ = -344.3 kcal/mol for 3'UTR). Both 5'and 3' UTR of Cx30.3 have very strong secondary structures, with very stable stem-loops throughout the UTRs. In addition to their stable secondary structures, the 5' UTR and 3' UTR of Cx30.3 are relatively long (125 nts and 1308 nts in size, respectively).

Searching for upstream AUG codons (uAUGs) revealed two uAUGs within the first exon. Both uAUGs are followed by in-frame stop codons prior to the normal AUG codon, resulting in two uORFs (upstream open reading frame), each of which contains 5 codons (Figure 5-6).

5.3.3 Analyses of Zebrafish Cx30.3 mRNA Expression

Using RT-PCR (reverse transcriptase polymerase chain reactions), we studied the mRNA expression of Cx30.3 in embryos at stages from 0.5, 1 to 5 dpf. RNA samples were isolated from two independently collected batches of embryos, and reverse transcribed to obtain independent cDNA samples. Results from these parallel RT-PCR experiments were consistent with each other. The zebrafish EF1α gene was used as a positive control. The primers for EF1α were designed so as to anneal to different exons, thus amplification from cDNA templates will result in shorter PCR fragments than amplification from genomic DNA templates (569 versus 733 bp). Therefore, the parallel amplification of EF1α also serves as an indicator for the integrity and

non-genomic-contamination of the cDNA samples. As shown in Figure 5-8, Cx30.3 transcripts were not detected in 0.5 dpf embryos, but were readily detected at 1 dpf and later stages.

Whole mount *in situ* hybridization (WMISH) was performed with embryos at different embryonic stages from 1-5 dpf. The anti-sense and sense probes were about 1.4 kb in size, and were designed to target the 3'UTR plus a very short fragment from the 3' end of its coding region. These regions of connexin gene sequences are usually less conserved than other regions, thus providing higher specificity for mRNA hybridization than probes from coding regions. Cx30.3-specific *in situ* hybridization signal was detected in 1 dpf stage embryos, in the otic vesicle (Figure 5-9, A) and dorsal fins (Figure 5-9, B). At 2 and 3 dpf, the Cx30.3-specific mRNA signal was readily detected in the inner ear (Figure 5-9, D), the lateral line (Figure 5-9, E & G), the fins, and the skin (Figure 5-9, C). This pattern remained unchanged at 4 and 5 dpf (data not shown). In order to obtain a more detailed picture regarding the internal expression pattern, we embedded embryos in plastic and obtained 5 μm sections with a microtome. Images captured from 2 dpf stage embryos are presented (Figure 5-10, A – C), and the staining pattern in later stage embryos was very similar. As seen in these images, Cx30.3 mRNA signals were strongly detected in the epithelial cells in the skin of the whole embryo, as well as cells in the inner ear. Signals were not detected in the heart or other embryonic tissues, although the more sensitive technique of real-time RT-PCR, suggested that Cx30.3 mRNA expression occurs in a number of adult tissues (see below, tissue

distribution analysis by real-time RT-PCR). Either Cx30.3 is not expressed during 1-5 dpf of development in these tissues, or the transcript levels are simply too low to be detected by our *in situ* hybridization protocol.

Using real-time RT-PCR, we further analyzed the tissue distribution of Cx30.3 mRNA among different tissues in adult zebrafish. We selected the zebrafish EF1α gene as the reference gene in these real-time RT-PCR experiments, as EF1α has been demonstrated as the most suitable reference gene for zebrafish real-time PCR analyses, both in time-course and tissue based experiments (Tang et al. 2007). Cx30.3 expression levels were highest in the skin, the fins (which are, of course, covered in skin), and the inner ear of adult zebrafish (Figure 5-11). Relatively low levels of expression were also detected in the heart, retina, liver, ovary, swim bladder, and the gill.

5.4 Discussion

5.4.1 Cx30.3 is the zebrafish ortholog of mammalian Cx26

In the present study, we identified a novel connexin gene, Cx30.3, in the zebrafish genome when we were looking into the promoter elements of ZfCx48.5 by genomic sequencing. A BLAST search revealed that one similar sequence had previously been deposited in GenBank, and was named Cx33.8 (accession No. AY_135446; Kremer,M. and Dermietzel,R., 2002, unpublished). Using 5' and 3' RACE, we identified the open reading frame (ORF) of this novel connexin gene and following careful comparison of the ORF with that of the previously deposited Cx33.8 in GenBank, we found that an

upstream ATG sequence located in the first intron was mistakenly considered as the translation start codon of Cx33.8, thus giving a larger prediction of its molecular weight and therefore so named. In 2005, mRNA sequence of Cx33.8 was updated as NM_212825.1 (GI:47086382, Apr., 2005), and this version was subsequently replaced by a full-length mRNA sequence with accession number NM_ 212825.2 (GI:66472875, May, 2005). Further, in May 2005, M. Dickson M. and colleagues also deposited in GenBank a full-length mRNA sequence of Cx33.8 with accession No. BC095350 (GI:63101893). Comparing the full-length cDNA sequence of Cx30.3 identified in this study with that of Cx33.8 in GenBank, it is clear that zebrafish Cx30.3 and Cx33.8 are in fact the same gene. We therefore re-named this connexin gene as Cx30.3, following the conventional molecular weight nomenclature of the connexin gene family (Beyer et al. 1990). We think that Cx30.3 is the more appropriate name for this gene, and will continue to use this name in the future. The full-length mRNA sequence of Cx30.3 has been deposited into the GenBank (accession No. GQ469999).

The predicted amino acid sequence of Cx30.3 is consistent with all the features of a typical connexin gene, having for example, four α-helical transmembrane domains and two extracellular loops with three highly conserved cysteine residues in each. By the formation of intramolecular disulfide bonds, these conserved cysteines play essential roles for normal folding and/or channel function (Unger et al. 1999; Yeager et al. 1998; Rahman and Evans 1991; John and Revel 1991; Dahl et al. 1992).

Phylogenetic analyses suggest that zebrafish Cx30.3 is the ortholog of

mammalian Cx26 (Figure 5-3). At least two other groups have also made this suggestion (Eastman et al. 2006; Cruciani and Mikalsen 2006; Cruciani and Mikalsen 2007). Amino acid sequence similarity comparisons showed high degree of identities between ZfCx30.3 and mCx26. Looking into the genomic organization of both zebrafish and mouse connexin genes, we also found that the genomic organization has been maintained during evolution, as a cluster of three connexins. In humans, Cx30, Cx26, and Cx46 are all found physically linked in the same gene orientation, on one chromosome (chromosome 13). In zebrafish, Cx30.3 (ortholog of human Cx26) and Cx48.5 (ortholog of human Cx46) are physically linked on chromosome 9, in a similar organization to that of human Cx26 and Cx46 (Figure 5-4). These similarities (ortholog connexin genes, and similar genomic cluster organizations) suggest that these connexin genes are most likely evolved from a single common ancestor (Eastman et al. 2006).

In such a genomic cluster on zebrafish chromosome 9, about 600 nucleotides downstream of the poly (A) signal of Cx30.3 is one of the transcriptional start sites of the Cx48.5 gene, and also notably, both connexin genes are in the same open reading frames (ORF) orientation. Cx48.5 is primarily expressed and functions in the lens tissue, as well as in the heart. Expression of Cx48.5 was also detected in the otic vesicle and later in the inner ear (Cheng et al. 2004). It appears that both 30.3 and 48.5 connexin genes have distinct and overlapping expression patterns (see below for more details). Given that both connexin genes are so closely linked to each other, it is possible that there are some transcriptional regulatory elements of Cx48.5 embedded within the Cx30.3 gene, and/or

there are potential mechanisms for co-regulation of the expression of both connexin genes. Further studies of the regulatory sequences of these two connexin genes are required to address these questions.

5.4.2 Highly structured untranslated regions of zebrafish Cx30.3

In the present study, we identified the transcription start site, 5'- and 3'- UTRs (untranslated regions) of zebrafish Cx30.3 using RACE methods. The RLM-mediated 5'RACE strategy takes advantage of the 5' cap to selectively capture full-length 5' end of mRNAs, thus making it feasible to map the transcriptional start sites of a gene. This method has been demonstrated to be comparable and, in many cases, more accurate than other approaches, for example, S1 nuclease protection, RNase protection, and primer extension (Liu and Gorovsky 1993; Schaefer 1995; Gum et al. 2003). 5' RACE identified a single transcriptional start site for ZfCx30.3, and a TATA box located 28 bp upstream appears to be responsible for the transcriptional initiation. 3' RACE identified the 3' UTR sequence of Cx30.3. These RACE experiments suggest that Cx30.3 gene contains two exons, with its 5' UTR interrupted by an intron. A two-exon genomic structure module has been described as the common gene structure for most connexin genes (Sohl and Willecke 2003; Sohl and Willecke 2004). It appears that zebrafish Cx30.3 follows this general rule.

In silico analyses of the UTR regions suggest that Cx30.3 has a highly structured 5'UTR and 3' UTR (Figure 5-6, and 5-7). Generally, UTRs are important in the

post-transcriptional regulation of gene expression, including mRNA transportation out of the nucleus, subcellular localization, mRNA stability, and efficiency of translation. These regulations can be mediated by interactions between nucleotide patterns or motifs located in 5' and 3' UTRs with specific trans-acting proteins, or specific non-complementary non-coding RNAs. Unlike DNA-mediated regulatory elements, however, the function of these RNA regulatory motifs relies on a combination of primary and secondary structure (Mignone et al. 2002; Pickering and Willis 2005; van der Velden and Thomas 1999; Jansen 2001; Bashirullah et al. 2001; de Moor et al. 2005). Eukaryotic 5' UTRs are important for regulation of protein synthesis by influencing mRNA stability and translational efficiency (Vilela et al. 1999; Wang et al. 1999; Kindler et al. 2005; Gebauer and Hentze 2004; Pickering and Willis 2005; Mignone et al. 2002). The initiation of mRNA translation begins with the binding of the small ribosomal subunit (40S) to the 5' cap structure. The 40S subunit then scans the 5' UTR until it encounters an AUG codon, where the 60S subunit enters to form the entire translational machinery, and polypeptide synthesis begins. Most cellular mRNAs use this cap-dependent mechanism for protein synthesis, and the 5' UTRs are required to contain more than 12 nts. A longer 5' UTR usually results in increased efficiency of translation by facilitating the recruitment of more pre-initiation complexes, but further increases in 5' UTR length usually down-regulate the translational efficiency due to secondary structure or alternative initiation of translation at upstream AUG codons (Kozak 1991). Most eukaryotic mRNAs contain short 5'UTRs (usually fewer than 50 nts), and thus, cap-mediated translation of

such mRNAs is efficient. About 10% of all mRNAs contain atypically long 5' UTRs, and their protein products typically regulate vital cellular processes, including cell growth, cell proliferation, apoptosis, etc. A long 5' UTR gives the propensity to form stable secondary structures that may inhibit translation initiation by impeding the ribosomal scanning process (Pickering and Willis 2005). In situations where cap-dependent translation is impaired or suppressed, an alternative mechanism can be used by such long 5'UTRs, as translational initiation can be mediated by a secondary structure called the internal ribosome entry site (IRES) (Kaminski et al. 1994; Jackson 2005). Recent estimates suggest that up to 10% of all cellular mRNAs have the potential to initiate translation via this mechanism (Pickering and Willis 2005). Many of the IRESs identified in cellular mRNAs are GC-rich, and therefore most likely have complex RNA secondary structures. However, thus far there is no recognizable conservation of secondary structure or sequence among these IRESs. It is hypothesized that many cellular IRESs contain short modules and internal initiation is mediated by the combined effect of these modules (Stoneley and Willis 2004), such as a Y-shaped double hairpin structure (Le and Maizel 1997), the sequence GNRA (where N is any nucleotide, and R is a purine) (Hoffman and Palmenberg 1995; Robertson et al. 1999), polypyrimidine tracts, and the sequence AGACA (Huez et al. 1998). IRES-mediated translational initiation has been studied in the mammalian Cx43 and Cx32 genes (Schiavi et al. 1999; Hudder and Werner 2000). These examples have prompted scientists in the gap junction research community to further look at other connexin genes (Werner 2000). As shown in Figure 5-6, zebrafish

Cx30.3 has a relatively long 5' UTR (125 nt in size), with a predicted stem-loop with an extended Y-shaped secondary structure. On closer inspection, the 5' UTR primary sequence also has the motif sequences "GNRA" and an "AGACA" found in some IRESs (Figure 5-6). Such short RNA modules in its 5' UTR suggest that the zebrafish Cx30.3 transcript might have potential to use the IRES-mediated mechanism for regulation of translation.

A search for uORFs revealed that the zebrafish Cx30.3 mRNA contains two upstream open reading frames (uORFs) within its 5' UTRs. We suggest that this also might be a common feature for most connexin genes, since a number of connexin genes have been found to contain upstream open reading frames (uORFs) in their 5' UTRs (Jacob and Beyer 2001; Pfeifer et al. 2004; Hudder and Werner 2000; Anderson et al. 2005; Meijer et al. 2000). The presence of uORFs in the 5' UTR usually decreases the efficiency of translational initiation on the downstream main ORF (Morris and Geballe 2000; Iacono et al. 2005). For example, Cx41 of *Xenopus laevis* has three uORFs in its 5' UTR, and each of these uORFs largely down-regulates the translational efficiency (Meijer et al. 2000; Meijer and Thomas 2003). uAUGs and uORFs may also play a role in the regulation of mRNA stability through either the nonsense-mediated mRNA decay (NMD) pathway (Ruiz-Echevarria and Peltz 2000), or NMD-independent mechanisms (Vilela et al. 1999), thus providing an alternative mechanism for the regulation of gene expression by uORFs. Whether these short uORFs of zebrafish Cx30.3 play an important role in the translation of mRNA certainly needs further detailed studies.

Synthesis of proteins involved in vital cellular and developmental processes, such as growth factors, transcription factors or proto-oncogenes, pro-apoptosis factors, all need to be strongly and finely regulated. The mRNAs encoding such proteins frequently have 5'UTRs that are longer than average, with uAUGs or uORFs, and stable secondary structures impeding efficient translation. In addition, these mRNAs often have the potential to utilize IRESs to initiate translation under conditions where cap-dependent translation is reduced or abolished, such as under a patho-physiological cell stress like heat shock or apoptosis (Pickering and Willis 2005). Recent studies suggest that many connexin genes possess such similar features in their mRNAs (Jacob and Beyer 2001; Pfeifer et al. 2004; Schiavi et al. 1999; Hudder and Werner 2000; Werner 2000; Anderson et al. 2005). Clearly, the zebrafish Cx30.3 gene shares many features with other connexin genes, and it will be of interest to determine whether these features play a similar role in regulating expression of this gene.

5.4.3 Expression patterns are conserved between Cx30.3 and Cx26

Using RT-PCR, *in situ* hybridization, and real-time RT-PCR, I have investigated the transcriptional expression of zebrafish Cx30.3 during embryonic development as well as in some adult tissues. These data clearly demonstrated that the expression of Cx30.3 starts at early embryonic stages and is strongly expressed in some adult tissues, including the inner ear, lateral line, the fins and the skin. Interestingly, the mammalian ortholog Cx26, is similarly expressed and functions in the inner ear and the skin (Gerido and

White 2004; Lefebvre and Van De Water 2000; Richard 2000; Rabionet et al. 2002). In the inner ear, human Cx26 is the most prevalent connexin gene in both the supporting cells and hair cells in the cochlea (Kikuchi et al. 1995; Martin et al. 1999; Kikuchi et al. 2000). Cx26 plays an essential role in the recycling of potassium ions during auditory transduction. Mutations of Cx26 gene may disrupt the recycling pathway of K+ ions, leading to cellular dysfunction and ultimately cell death, thus resulting in loss or disorders of hearing (Lefebvre and Van De Water 2000). In the skin, human Cx26 is expressed in hair follicles and eccrine sweat glands and, at lower levels, in the basal keratinocytes of the palmar and plantar epidermis, and its expression is up-regulated in damaged and psoriatic epidermis. Mutations in Cx26 are also associated with skin disease – a number of disorders with varying phenotypes, but similar genetic abnormalities (Richard 2000; Richard 2005). As clearly demonstrated in the present study, in zebrafish, Cx30.3 has similar expression patterns to its mammalian ortholog Cx26. If Cx30.3 plays similar roles in development and homeostasis in zebrafish, these animals could serve to model some human hearing defects or skin diseases, as well as evaluating therapeutic drugs to ameliorate such disorders.

Connexin-expressing hair cells in the inner ear are involved in the sensory perception of sound and motion in mammals. Like other vertebrates, zebrafish also uses hair cells to detect sound and motion. Although the anatomy of the zebrafish inner ear differs significantly from that of the human inner ear, many of the same genetic pathways are utilized for the development and function of the inner ear in all vertebrates (Whitfield

2002). For example, myosin VIIa is required and functions the same way in hair cells, whether in human or zebrafish (Hasson 1999). In addition, fish have a lateral line system of hair cells, which are assembled to form the so-called "neuromasts", the sensory organs of the lateral line. Hair cells in neuromasts are directly exposed to the surrounding water and sense its movement along the surface of the fish head and body, thus play essential roles in a variety of behaviors, including prey detection, predator avoidance, school swimming, and sexual courtship (Ghysen and Dambly-Chaudiere 2004). It is the neuromasts of the lateral line that make zebrafish an excellent model system for studying hair cell development and function, because fish hair cells are readily accessible on the surface of the animal (Ghysen and Dambly-Chaudiere 2004; Dambly-Chaudiere et al. 2003; Goodrich 2005). In mammals, Cx26 is primarily expressed in sensory organs functionally related to hearing and balance, where mechanosensory stimulation of hair cells is essentially involved. In zebrafish, accordingly, Cx30.3 is dominantly expressed in the inner ear, the lateral line, the swim bladder, the fins, and the skin, all of which are functionally related to fish hearing and balance. Although further studies on the function of zebrafish Cx30.3 are absolutely needed, its expression patterns suggest that this connexin gene most likely plays significant roles in zebrafish hearing and balance. Zebrafish is a favorable model organism both genetically and experimentally (Driever et al. 1994; Lieschke and Currie 2007). In particular, microinjection techniques are well-developed for zebrafish embryos, and the delivery of anti-sense morpholino oligos has facilitated numerous studies of gene function during zebrafish embryonic

development. An obvious next step in our studies of Cx30.3 will be the use of morpholinos to assess the function of Cx30.3 in hearing and skin development and function in zebrafish, and perhaps the development of zebrafish as models for understanding the genetic basis of human diseases of hearing or disorders in the skin. I have designed morpholino oligos for these future studies (for details, see Appendix II).

In addition, mRNA expression pattern of Cx30.3 in the cardiovascular system and in the retina tissue is overlapping with that of Cx48.5 (Cheng et al. 2004). This expression pattern suggests that Cx30.3 might also play potentially significant roles in the development and function of these other tissues. Overlapping expression and functional diversity are general features of connexin genes. Whether or not both connexin genes have similar functions in the retina and cardiovascular systems awaits to be determined by detailed functional studies of zebrafish Cx30.3 in the future.

5.5 Summary

In the present sub-project of my thesis research, a new zebrafish connexin gene – Cx30.3 – has been identified and characterized. Consistent with most connexin genes, Cx30.3 also contains two exons, with its 5'UTR interrupted by an intron. Its mRNA contains a relatively long 5' UTR, with two uORFs upstream of the predicted AUG start codon. RNA secondary structure predictions showed very highly-structured 5' and 3' UTRs of Cx30.3 mRNA, suggesting potential complicated mechanisms for post-transcriptional regulation of expression. RT-PCR and ISH experiments with Cx30.3

demonstrated that its mRNA expression in zebrafish embryos starts from about 1 dpf of age. Real-time RT-PCR showed that Cx30.3 is expressed in a spectrum of tissues in adult fish, including the skin, fins, inner ear, heart, and the retina, etc. Its mammalian ortholog gene, Cx26, has been well established to be expressed and function in the inner ear and the skin of mammals, including humans. Expression patterns of the zebrafish Cx30.3 gene suggest that this protein may play significant roles in the development and normal function of hearing and balance in zebrafish. If these expression patterns can be confirmed, the zebrafish would serve as an excellent model to study the development and function of hair cells, as well as to model human diseases of hearing and disorders in the skin.

Figure 5-1. Genomic sequence of zebrafish Cx30.3. Underlined sequence indicates the first exon of Cx30.3. The box in the first line indicates a TATA-like element of the core promoter. The parentheses sequence is the intron of Cx30.3, with bold nucleotides indicating the 5' and 3' end of the intron. Pre-mRNA splicing follows the "GU-AG" rule. Sequence between the bold and highlighted nucleotides (ATG and TAA) are coding sequence. The box in the last line indicates a conserved "AAUAAA" signal.

TCAGGTGCAC CTCTCCCCTA GAAAATAAGT GCTCACTGTC TTCAGTCAAT CACTTCATCT Exon 1
CACTCTGAAG ACCTCAGAGG CACCACACAA GTGATTTGCT CAACTTAGCT CACACTGTCG 5'UTR
GATATATGTC ATGGATTTGA TCTGAATCAT CGTGTAGCCT AACAG{*GTAAA CAGAAGTCT*
……(intron, ~1.6 kb)……
CTGTTTGCT GCGTTGGTGC TGTATCAATT TCCTCTCTTT CTTTTCAAAC AG}GAGCTGAG Exon 2
AGCACCAGTA GACAGGATGA GTTGGGGAGC ACTTTATGCT CAGCTGGGAG GAGTGAATAA Coding Region
ACACTCCACC AGCTTGGGGA AGATCTGGCT GTCTGTCCTC TTCATCTTCC GCATTTGCAT
CCTGGTCATA GCAGCAGAGA CGGTCTGGGG AGACGAACAG TCAGACTTCA CCTGCAACAC
ACAACAGCCT GGTTGCAAAA ACGTCTGCTA TGACCACTTC TTTCCAGTCT CGCACATACG
TTTCTGGTGT CTGCAGCTCA TCTTTGTGTC CACACCGGCT TTACTGGTAG CTATGCATGT
GGCATATCGC AAGCGCAACA TGAAAAAGAA AAGCATTTTA GCCAAGCGTG GTGGTAATGG
TAAAGGAGAT GACCTGGAGA GCTTGAAGAA CCGGCGTCTA CCCATCACTG GGCCACTGTG
GTGGACCTAC ACATCCAGCC TGTTCTTCAG ACTTCTTTTC GAGGCCGGAT TCATGTATGC
TCTCTATTAC GTCTATGATG GCTTTCAGAT GGCACGCCTT GTGAAGTGTG AGCAATGGCC
TTGTCCCAAT AAAGTTGACT GTTTCATCTC AAGGCCGACA GAGAAGACGG TCTTCACCAT
CTTTATGGTG GGATCTTCTG CTATCTGCAT TGTGCTCAAT GTGGCTGAAC TGGCCTATCT
GATTGTCAAA GCATTGCTCA GGTGCTCAGC CAGAGCCAAA GGGAGGCGCT CATTTGTACA
CCAAGAGAAA ATGTCCACAG AAAAGGCGCA CCTACAGAAT GAAAAAAATG CAAGGTTGCT
GTCATCAGCT TCGGACTCAT CGAGCAATAA GACTGTTTAA GATGATCTTC AACTCCGCCA
TCATGGGAGA TGATAACAGT AGATTTAACA GTATGTGCTT TGTAGCTCTT TTTGACAAGT 3'UTR
AGCCCTGAGA ACACATCAAA AGACACACAC TTGAAAGCAG GTTTCAGCGT GGCTCTTTCC
GAAGCTGAAA TTCTTTCCCC TAGCCGAAAG ATGATTCAGA GTTATAACAC TGCATGGTCA
ATGAAGAAGA TCTGAGGGTT GTTGAGATAT ATTTGTAATT TTTTTACTTG CTAACTGATT
ACAGTACTTA GGAGTGTCCA ACTTTGGTCC TGGAGGACCG GTGTCCTGGA GAGTTTAGCT
CCAACCCTAA TCAAACACAC CTGAACCAGC TAATTAAGCT CTTACTAGAT ATACTAGAAA
ATCTCTGGCA GGTGTGTTGA AGCAAGTTGG AGCTAAACTA TGCAGGACAC CGGCCGTCCA
GAACCGAGTT TGCACACCCC TGGGCATTAG AAAAGCGAAA GAGTGACTCC TAGTGGTCAA
ACAACACACA CGTTCAAAGG CCGCCTTTGC ACAGAATTTG TGTCCTGTTT TTGAAGGATC
GTTTTTATGT TGTACAGAAA TGCATCTCTA TATTTGGACA CTTTGCTATG TTTACTTAAT
TTTTAAATGC ATTTTAGTCT TTCAGTTTCC TACCACCAAT TTTAGCATGC AAATGTTTTT
TGTATGTGTG ATTTGTTTTT GGGTTTTTAA TATTCTTTTG GTAAACGATT CAAATGTTTC
ACAACTCACA AGGTCAGTCT TACATTCAAC CGTACCTGGT ATTTCATAGC AGCTTAAAGC
AATGTGCCTT GTTACATGAA TCTGCAGTAT TTAATGAGCA ATAATAGATA TTTCAAACAA
AGTTTTAGCC TTATAGTCAA TCATTAAAGC AAGCCATTAA GTATTTCTAT TTGGACTACT
TGGATGATGT TTGCTGAAGA GTAAGCACTA CATGACATAT CCCAAAAGAG TAGTGGCTCT
GGAGTATGTT ATTGTAATAC AGTCTTATTT ATTAGCCGTT AGAGTTGCAA ACAGGCTGCA
ATCGCGTGAG AATGTTTACT TTTTTACTAG TCGCGCTTCT CTTTGACATC TAGTCAGGAT
ACCTGTTGAC AGATGAGAGC GCATTCAGTG ACAGAAGTTG GGAATGATGC GTACCTAAAT
GTCCACATGG TATTGGTTGA ATCTTGTTTC CCAAAGCAAG CTGTGATCTT TTTTATTGTC
AACTGATTGC CAGCGTAATA CATGTTCTTT GCGATTTTGG CTAACAACTA ATATTTTATA
TGAAGGTCTT AATAAAAAAT CTCTAAAAA

Figure 5-2. Amino acid sequence multi-alignment of ZfCx30.3. Zebrafish Cx30.3, and its two most closely related sequences, mouse Cx26 (mCx26), and human Cx26 (HuCx26) are aligned using Kalign program. Solid lines under the sequences indicate the four predicted transmembrane domains (M1, M2, M3, M4). The X symbols indicate the six conserved cysteine residues within the two predicted extracellular loops of ZfCx30.3.

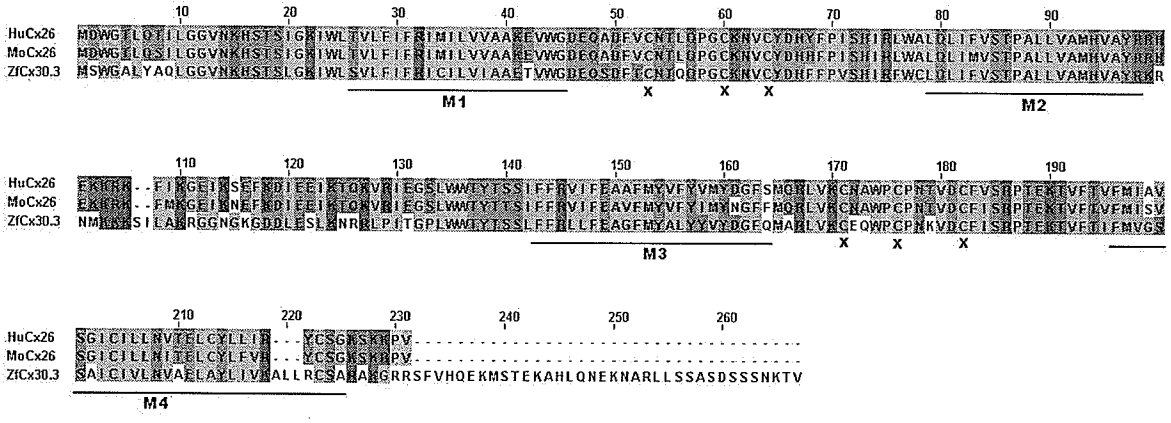


Figure 5-3. Phylogenetic analysis of zebrafish and mammalian connexins. Neighbour-joining tree shows that zebrafish Cx30.3 is most closely related to human Cx26. Figure kindly provided by Dr. Gunnar Valdimarsson.

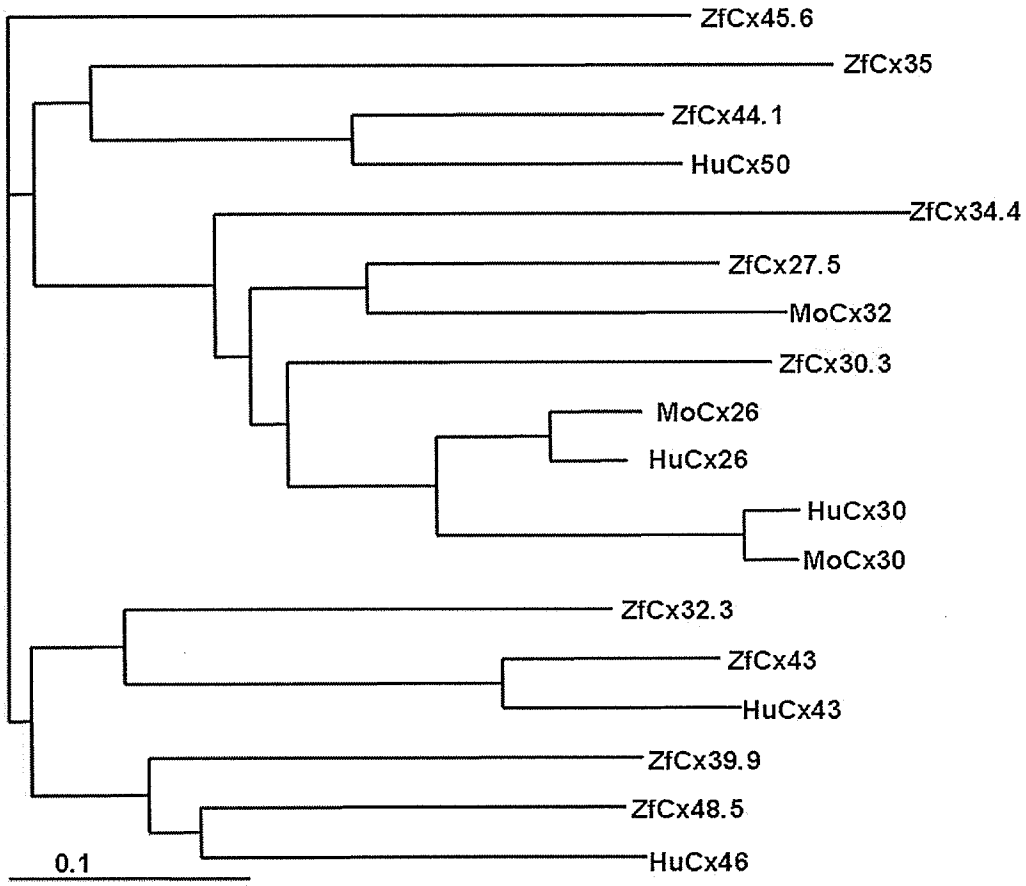


Figure 5-4. Genomic organization of connexin gene clusters. Connexin genes are represented by arrows. Similar to a cluster of three human connexins (Cx30, Cx26, Cx46) on human chromosome 13, ZfCx48.5, the zebrafish ortholog of Cx46, is closely linked to ZfCx30.3, the zebrafish ortholog of human Cx26. Gene distance is not drawn in scale.

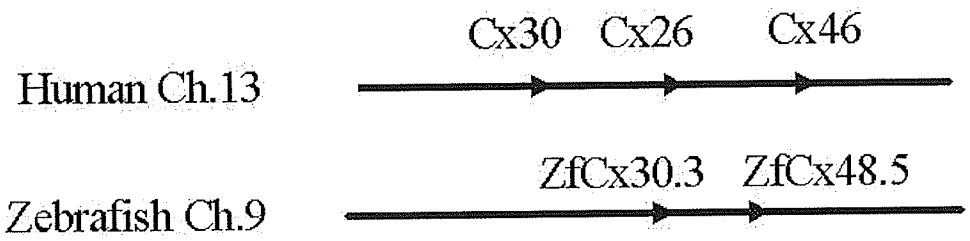


Figure 5-5. Gene structure of zebrafish Cx30.3. ZfCx30.3 contains two exons separated by an intron of about 1.6 kb in length. The complete coding region and 3'UTR are located in the second exon. The first exon is 110 bp in size, and the coding sequence is 804 bp, while the 3'UTR is about 1.3 kb in size.

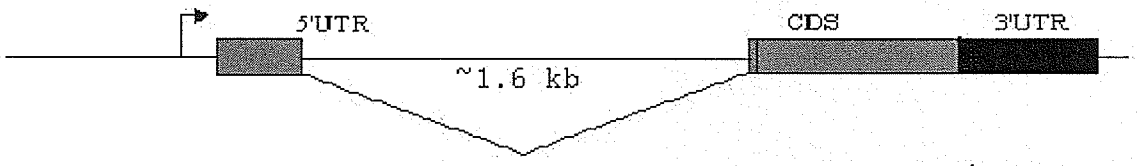
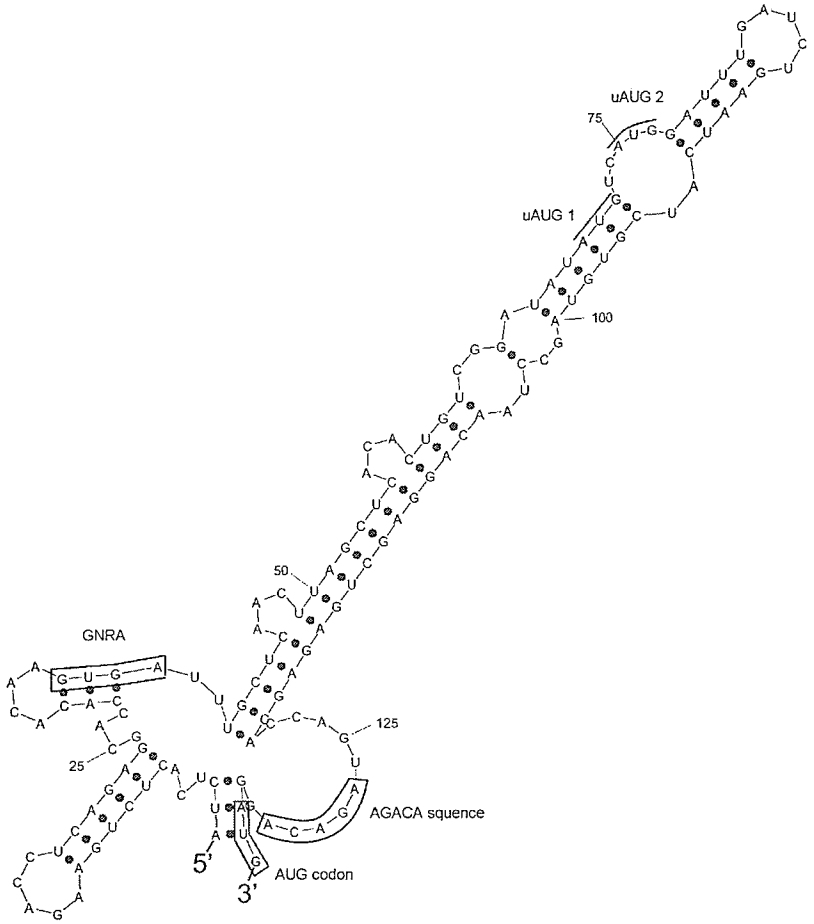


Figure 5-6. ZfCx30.3 5'UTR secondary structure. The most thermodynamically stable structures are shown with free energy at 37 °C (ΔG = -36.4 kcal/mol). The 5' UTR of zebrafish Cx30.3 forms a semi-conserved Y-shaped structure with extended stable stem loop. IRES *cis*-acting elements, such as the GNRA (nt 35-38) and the AGACA (nt 127-131), are marked by blue boxes. Normal translational start codon is marked by the red box, while two upstream AUGs (uAUGs) are indicated by red lines (nt 70 and 75, respectively). Both upstream open reading frames (uORFs) contain 5 codons, and both terminated by a UGA codon.



$\Delta G^{\circ}_{37} = -36.40$

Figure 5-7. ZfCx30.3 3'UTR secondary structure. The most thermodynamically stable structures are shown with free energy at 37 °C (ΔG = -344.3 kcal/mol).

Figure 5-8. RT-PCR analyses of ZfCx30.3 mRNA expression. Expression of Cx30.3 was detected at 1 dpf (day past fertilization) stage, and thereafter maintained throughout embryonic development. Zebrafish EF1α gene was used as positive control to demonstrate the integrity and genomic-DNA-contamination-free status of the cDNA samples.

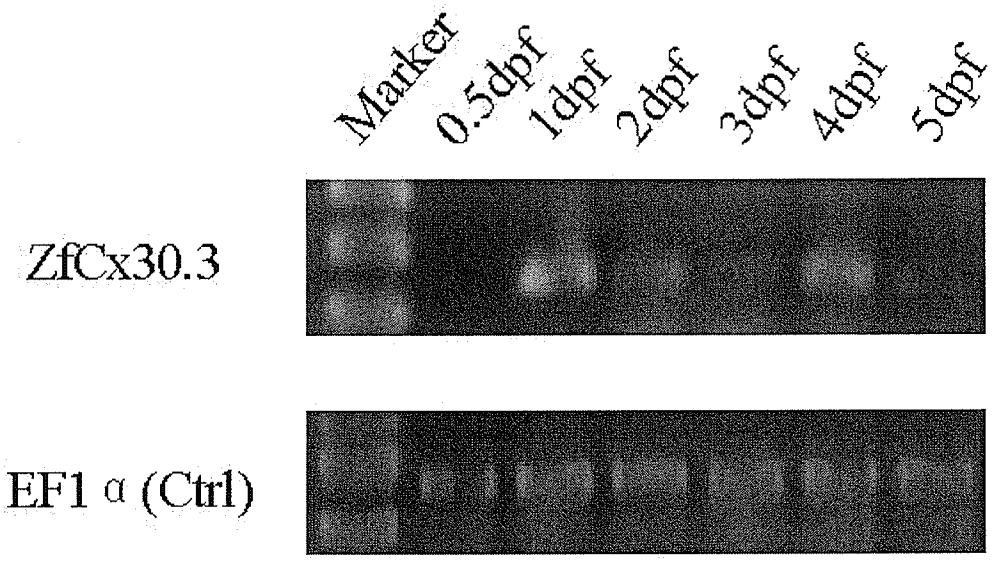


Figure 5-9. ***In situ*** **hybridization of ZfCx30.3 mRNA in embryos.** Embryos were stained by whole-mount in situ hybridization with a Cx30.3 anti-sense riboprobe. Signal was detected (and highlighted with arrows) in 1 dpf stage embryos, in the otic vesicle (**A**) and dorsal fins (**B**). At 2 dpf (**C**, **G**, **H**) or 3 dpf (**D**, **E**, **F**), Cx30.3 mRNA signal was readily detected in the skin (**C**), the inner ear (**D**), the lateral line (**E**, **G**), and the fins (**C**). Sense-probe (negative) controls are shown in (**F**) and (**H**).

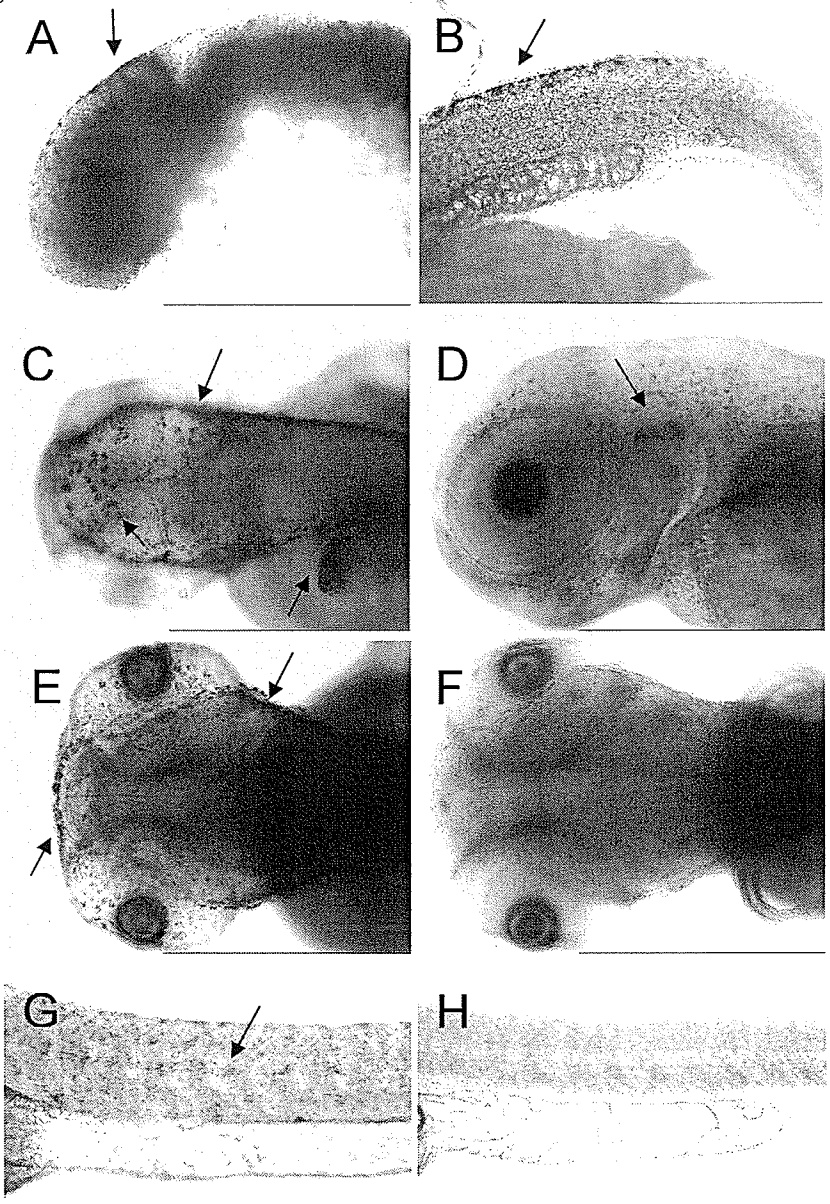


Figure 5-10. Sections of ZfCx30.3 ISH stained embryos. Embryos were stained by whole-mount in situ hybridization with Cx30.3 anti-senses riboprobe, then embedded in plastic. 5 μm sections were cut with a microtome. Only sections from embyos at 2 dpf (days post fertilization) are shown. The staining pattern in later stage embryos is similar. (**A**) transverse cross-section of embryonic eyes, showing (lack of accumulation) of probe; (**B**) shows developing inner ear with probe labeling the otic vesicle (arrow); and (**C**) shows a tranverse section of the embryonic fish tail and fin, with probe accumulating around the peripheral skin and on the fin (far right).

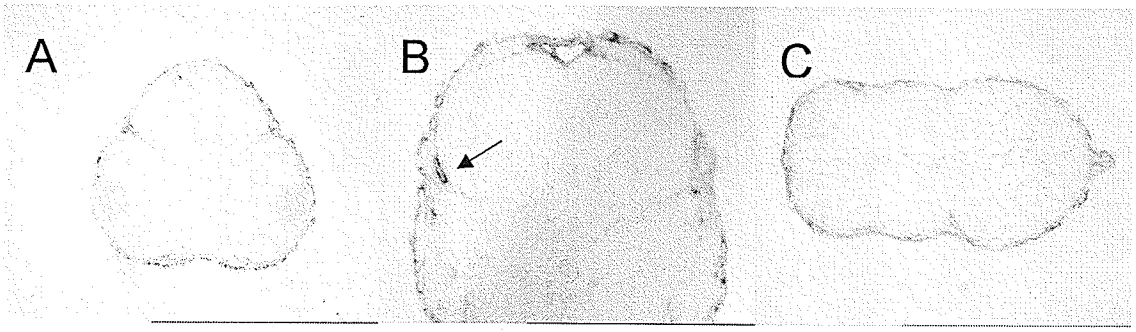
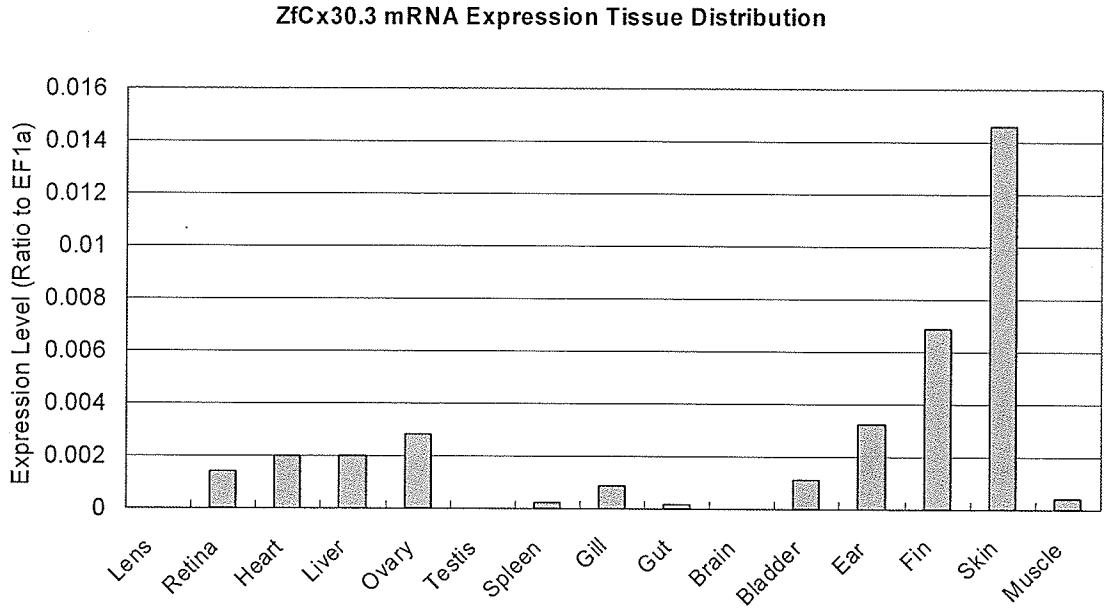


Figure 5-11. Tissue distribution of ZfCx30.3 mRNAs. Cx30.3 mRNA is predominantly expressed in the adult skin, fin, inner ear, ovary, liver, heart, retina, while a relatively low level of Cx30.3 expression was also detected in the gill, swim bladder, skeletal muscle, spleen and gut. mRNA levels are expressed by ratios to that of zebrafish EF1α.. The expression ratios were calculated by comparing ΔCt between mean Cts of ZfCx30.3 and zebrafish EF1α mRNA in two replicate experiments, and presented as a column graph.



CHAPTER SIX

GENERAL CONCLUSIONS

This thesis has presented a series of studies that were aimed at exploring the molecular genetic mechanisms that regulate the expression and function of connexin genes using zebrafish as the model system.

Using 5' RACE techniques, this thesis has mapped the transcriptional start sites and identified the 5' UTR sequences of several zebrafish connexin genes (including Cx43, Cx44.1, Cx45.6, Cx48.5, and Cx30.3), as well as mouse Cx46 (the ortholog of zebrafish Cx44.1) and Cx50 (the ortholog of zebrafish Cx48.5). The 3' UTR sequences of zebrafish Cx48.5 and Cx30.3 were also identified using 3' RACE method. Sequence analyses showed that most of these connexin genes (except zebrafish Cx43 and Cx30.3) have alternative 5' UTR variants (differing in their first exons) that are generated by different promoter usage and alternative pre-mRNA splicing. This could be a more widespread feature of connexin gene expression, as it was also shown in some other connexins (Anderson et al. 2005). Also significantly, this feature of connexin gene expression appears to be evolutionarily conserved between species, as shown in this thesis by comparative analyses of some zebrafish and mouse connexin genes (ZfCx44.1 vs mCx50, ZfCx45.6 vs mCx40, and ZfCx48.5 vs mCx46).

As UTRs are generally important for post-transcriptional regulation, the complicated 5' UTR diversity demonstrated in this thesis suggests that the expression of connexin genes are tightly regulated at both transcriptional and post-transcriptional levels.

In this thesis, I performed both *in silico* and *in vivo* comparisons between promoter activities of zebrafish Cx44.1 and mouse Cx50. Preliminary results suggest that transcription of homologous connexin genes might share regulatory networks between species, though further analyses will be needed to dissect the *cis*-elements and *trans*-factors for both connexin genes (Chapter 2). I also performed GFP transient expression experiments to test the capacity of a ~12 kb upstream promoter fragment of zebrafish Cx48.5 (Chapter 4). The promoter analyses in this thesis have laid the groundwork for further studies that map the regulatory elements in these connexin promoters, as well as provided useful information on which regulatory sequences may be useful in future genetic studies using transgenic zebrafish lines. *In silico* analyses of 5' UTRs of these connexin genes revealed relatively long primary sequences and very strong loop-stem secondary structures for all connexin genes in this thesis (Chapter 2 to 5). 5' UTRs of these connexins are also rich in uORFs, which are generally thought to regulate the translation efficiency by inhibiting the initiation of translation. These analyses suggest that 5' UTRs could play an important role in the translational regulation of connexin genes, and some mechanisms for translational regulation, such as IRESs, could be used by the 5' sequences of the connexin genes examined in this study. Based on these analyses, *in vitro* reporter assays have been suggested for future studies, to investigate the functional roles of the 5' UTR variants in the translational regulation of connexin gene expression.

While alternative transcripts of these connexin genes have their variable first

exons, a significant feature is that all alternative transcripts have common coding sequences, thus translating the same protein product. This feature of gene organization raises interesting questions, for example, why alternative transcripts are needed or what the functional role of the individual transcript variants can be. We hypothesized that transcript variants may have distinct (and overlapping) functional roles during zebrafish embryonic development and in adult tissues. For future studies, functional analyses have been proposed to use "transcript variant-specific morpholino knock-down" approaches with fluorescent protein tagged transgenic zebrafish embryos (see Appendix II).

Interestingly, identification 5' UTR variants of zebrafish Cx45.6 in this thesis has revealed two tandem alternative splice acceptors in the "CAGCAG" motif, which produce alternative transcripts that only differ by 3 nucleotides. This is the first study that has demonstrated tandem alternative splicing events in the connexin gene family. Real time RT-PCR analyses showed that the ratio between the selection of the intron-proximal and distal AG splice acceptor sites appears to be constant (about 4:1) during embryonic development and in adult tissues. In future studies, it would be interesting to investigate whether or not the tandem alternative splicing of zebrafish Cx45.6 have specific biological significance.

The sequencing of an upstream promoter fragment of Cx48.5 led to the identification of a new connexin gene, Cx30.3, in the zebrafish genome (Chapter 5). Sequence similarity and biophysical analyses suggest that zebrafish Cx30.3 could be the ortholog of mammalian Cx26, which has been well established to be expressed and

function in the inner ear and the skin of mammals including humans (Lefebvre and Van De Water 2000; Richard 2000; Rabionet et al. 2002; Gerido and White 2004). In this thesis, the transcriptional start sites and full-length cDNA sequence of Cx30.3 have been identified. Its expression patterns were extensively studied using RT-PCR, real-time RT-PCR, and ISH. Cx30.3 is expressed in early embryos, and in adult tissues including the skin, inner ear, fin, heart, and retina. The expression pattern suggests that zebrafish Cx30.3 might play significant roles in the development and function of hearing and balance in zebrafish, as does Cx26 in mammals. Taking advantage of the zebrafish model, Cx30.3 potentially provides an excellent model to study the development and function of hair cells, as well as to model human diseases of hearing and disorders in the skin. For this purpose, I have designed two non-overlapping morpholino oligos that can target either the 5' UTR sequence or the 3' splicing site, and proposed to use fluorescent protein labelled stable transgenic zebrafish embryos to perform the morpholino analyses (see Appendix II). The data presented in this thesis have provided groundwork for future functional studies.

Most of the zebrafish connexin genes studied in this thesis were previously isolated from a genomic library and their mRNA expression patterns were extensively characterized by former researchers in this lab (excluding Cx30.3). In this thesis, however, tissue mRNA distribution of all these connexin genes was further analyzed using real-time RT-PCR (Chapters 2 to 5). These results confirmed their distinct but overlapping tissue-specific expression patterns in adult zebrafish. For example, Cx44.1

and Cx48.5 are both lens-specific connexins; Cx45.6 is a heart-specific connexion; and Cx30.3 is a skin-specific connexin. Expression of specific genes in specific tissues frequently implicates certain functional importance in the development and function of the correspondent tissues. With the identification of the 5' UTR sequences of these genes, morpholino oligos have been designed. Targeted gene knockdown could provide a very powerful approach for future functional studies of these connexion genes.

APPENDIX-I EGFP TRANSIENT EXPRESSION ASSAY

Introduction

Transgenesis has been widely used to study the expression and function of specific genes *in vivo* as well as endow the host organism with properties of interest. Taking advantage of the accessibility and optical clarity of zebrafish embryos, fluorescent reporter genes can be assayed in live and developing embryos, making it possible to visualize and track the details of dynamic gene expression as well as cellular processes during embryonic development. One of these applications is to analyze the function of putative transcriptional regulatory sequences by *in vivo* EGFP transient expression assay. The capacity of a specific regulatory DNA sequence to drive EGFP expression can be rapidly tested in developing embryos by micro-injecting recombinant DNA construct into early embryos.

The conventional method for preparation of DNA injection includes multiple steps with constructing DNA plasmids as the necessary first step. Making recombinant DNA plasmids is usually a laborious and time-consuming procedure; and in some cases, the expected constructs are not easily attainable. During my PhD research, I developed a simple method to prepare DNA constructs for micro-injection into zebrafish embryos. Since linear DNA molecules are recommended for micro-injection, it could be much more efficient if constructs are direct products of PCR reactions. Based on this idea, I made my EGFP expression constructs taking advantage of the powerful PCR strategy. In an example described below, specific constructs can be easily prepared only by two

rounds of PCR reactions; and purified PCR products can directly be used for micro-injection.

Method

The zebrafish αA-crystallin promoter has been well studied using *in vivo* EGFP transient expression assays by another group: a promoter fragment greater than 564 bp has been used to drive EGFP expression in the lens in developing embryos (Kurita et al. 2003). I selected this promoter as a positive control to test the feasibility of my method. The preparation of the recombinant DNA construct includes the following four steps. For all PCR reactions described here, the Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen) was used to ensure the sequence accuracy of PCR products.

(A) Zebrafish αA-crystallin promoter fragment

A ~1 kb promoter fragment of zebrafish αA-crystallin was PCR amplified from genomic DNA. The forward primer used was ZfCrs/aA/F1 (5'- GAG TAT TTG GGT GTT TCC CAG -3'), and the reverse primer was ZfCrs/aA/gfp/R1 (5'- TCC TCG CCC TTG CTC ACC ATA GTT GGA TCG CAA TAT CCT TAA TG -3'). It should be noted that the underlined part of the reverse primer is an adapter sequence that can be annealed against the N-terminal of EGFP coding sequence from the ATG codon (Fig AI-1).

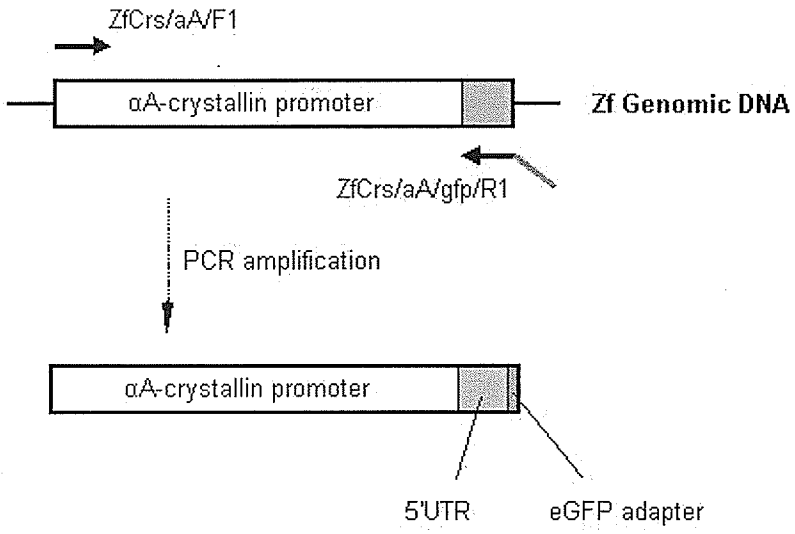


Figure A1-1. Zebrafish αA-crystallin ~1kb promoter fragment.

(B) EGFP-SV40polyA fragment

The EGFP-SV40polyA fragment was amplified by PCR from RARE3E(-SP6), a well-established EGFP expression vector (generous gift from Dr. Robert Davis, Harvard Medical School). The fragment includes the EGFP coding sequence (starting at the ATG codon and ending at the TAA stop codon), and the SV40 poly-A tail. The forward primer is gfp/F1 (5'- TGG CCT CCT CCG AGG ACG TCA TC -3'), while the reverse primer is gfp/R1 (5'- TAG ATG AAG GAG CCG TCT TGC AGG G -3'). It should be noted that the first nucleotides were selected to be a "T" in order to ensure there was no 3' hanging "A" to be added on the ends of PCR products (Fig AI-2).

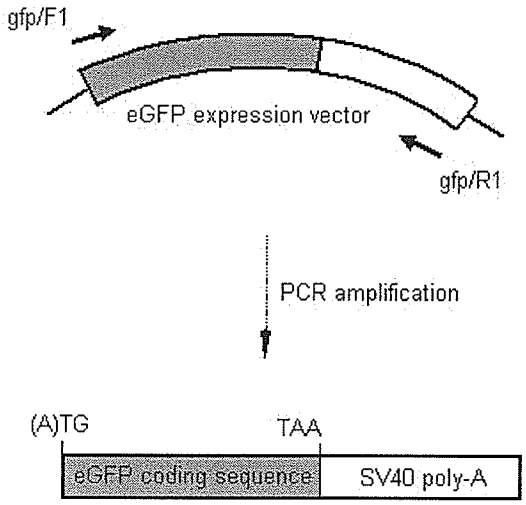


Figure A1-2. Amplification of the EGFP-SV40polyA fragment.

(C) Final recombinant DNA molecule

The above PCR products (**A**) and (**B**) were mixed at a ratio of about 1:1, and used as the template for the following round of PCR. The forward primer is ZfCrs/aA/F1, and the reverse primer is gfp/R1 (Fig. A1-3)

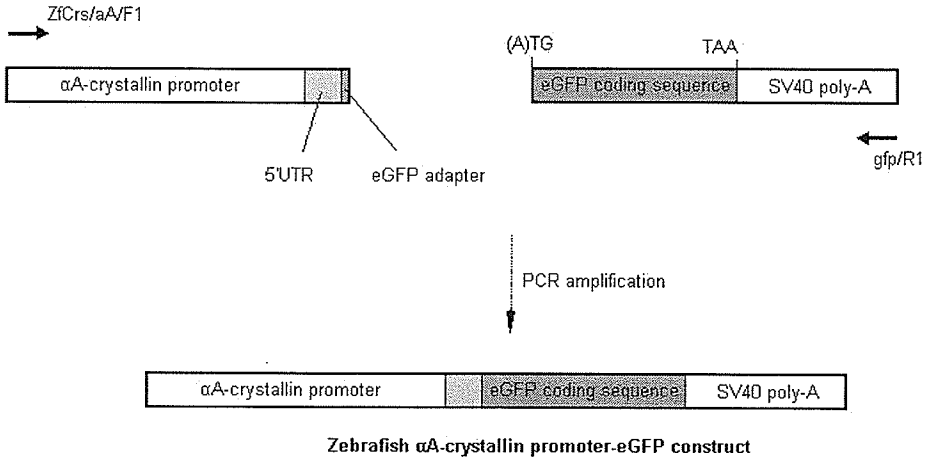


Figure A1-3. Amplification of ZfCrs-eGFP construct.

(D) Purification of final PCR products and micro-injection

The final PCR products (i.e., the recombinant DNA constructs) were purified using the GENECLEAN® Spin Kit (BIO-101 Inc., Vista, CA,USA), and then the concentration was adjusted to about 100 ng/μl in 100 mM KCl solution with 0.05% of phenol red (as the injection indicator). DNA constructs were injected into 1-cell stage zebrafish embryos.

Result

Strong lens-specific EGFP expression can be driven by zebrafish αA-crystallin ~1kb promoter.

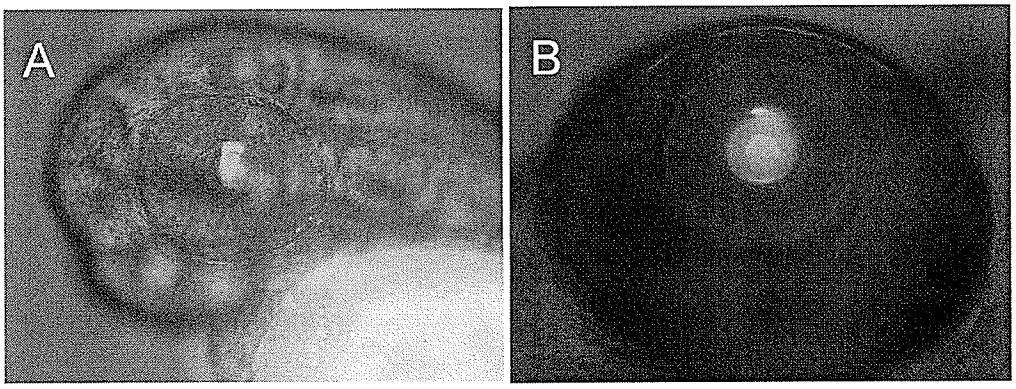


Figure A1-4. Lens-specificZfCrs-GFP expression. (**A**) 2dpf. Some autofluorescence is observed in the yolk sac (lower right), but fluorescence is obvious in the eye lens (**B**) 7dpf.

Discussion

As described in previous chapters and Appendix II, all of my recombinant DNA

constructs for zebrafish micro-injection (either linear or plasmid DNA molecules) were built using such a PCR strategy. I found this is a very powerful and flexible method to make recombinant DNA constructs.

The EGFP transient expression strategy described herein has many advantages. First, it is very efficient. The conventional method for linear DNA constructs preparation usually includes making plasmid DNA constructs, transformation to bacteria, isolating plasmid, restriction cutting, purification etc. This is a very laborious and time-consuming procedure. In contrast, in our method, it takes only two rounds of PCR followed by purification of PCR products. That is, it takes only several hours to get the DNA constructs ready for injection. This efficient method makes it possible for the researcher to work with several genes and lots of promoter deletion analyses in parallel. Second, in conventional methods for plasmid DNA constructs, the flexibility of selecting cloning sites are frequently limited by specific restriction enzyme sites. In contrast, our method provides much more flexibility, because the site where the two fragments are linked only depends on the design of PCR primers. Third, this strategy can also be readily used to make protein fusion constructs, as well as site-directed mutations. In addition, besides saving time, it also saves money.

APPENDIX-II

CX30.3-RFP/CX48.5-EGFP DOUBLE FLUORESCENT TRANSGENIC FISH AND POTENTIAL TRANSCRIPT VARIANT-SPECIFIC KNOCK-DOWN ANALYSES OF CX48.5

Introduction

As discussed in Chapter 4, the discovery of four types of transcript isoforms of Cx48.5 in the present study may raise questions, such as, why different transcripts are needed, or what are the functional roles of each individual transcript variant during zebrafish embryonic development and in adult tissues. Similar questions also may be asked about zebrafish Cx44.1. To address these questions, morpholino knock-down could provide a very powerful strategy, and we have proposed that "transcript variant-specific morpholino knock-down" could be employed for such functional analyses. Generally proteins determine phenotypes. However, as these transcript variants of Cx48.5 (and Cx44.1) only differ in their 5' UTRs and actually they lead to the production of the same protein, it is very challenging to link each individual transcript to potential fish embryonic phenotypes. Obviously there is a "black box" blocking this linkage (Figure A2-1, A). To address this question, I proposed to use fluorescently- labelled stable transgenic zebrafish embryos to perform such "transcript variant-specific knock-down" analyses. Fluorescent reporter proteins can be assayed in live and developing embryos, making it possible to visualize and track the details of dynamic gene expression as well as cellular processes during embryonic development. By this strategy, the "black box"

that blocks the linkage between transcript variants and potential phenotypes hopefully becomes transparent (Figure A2-1, B).

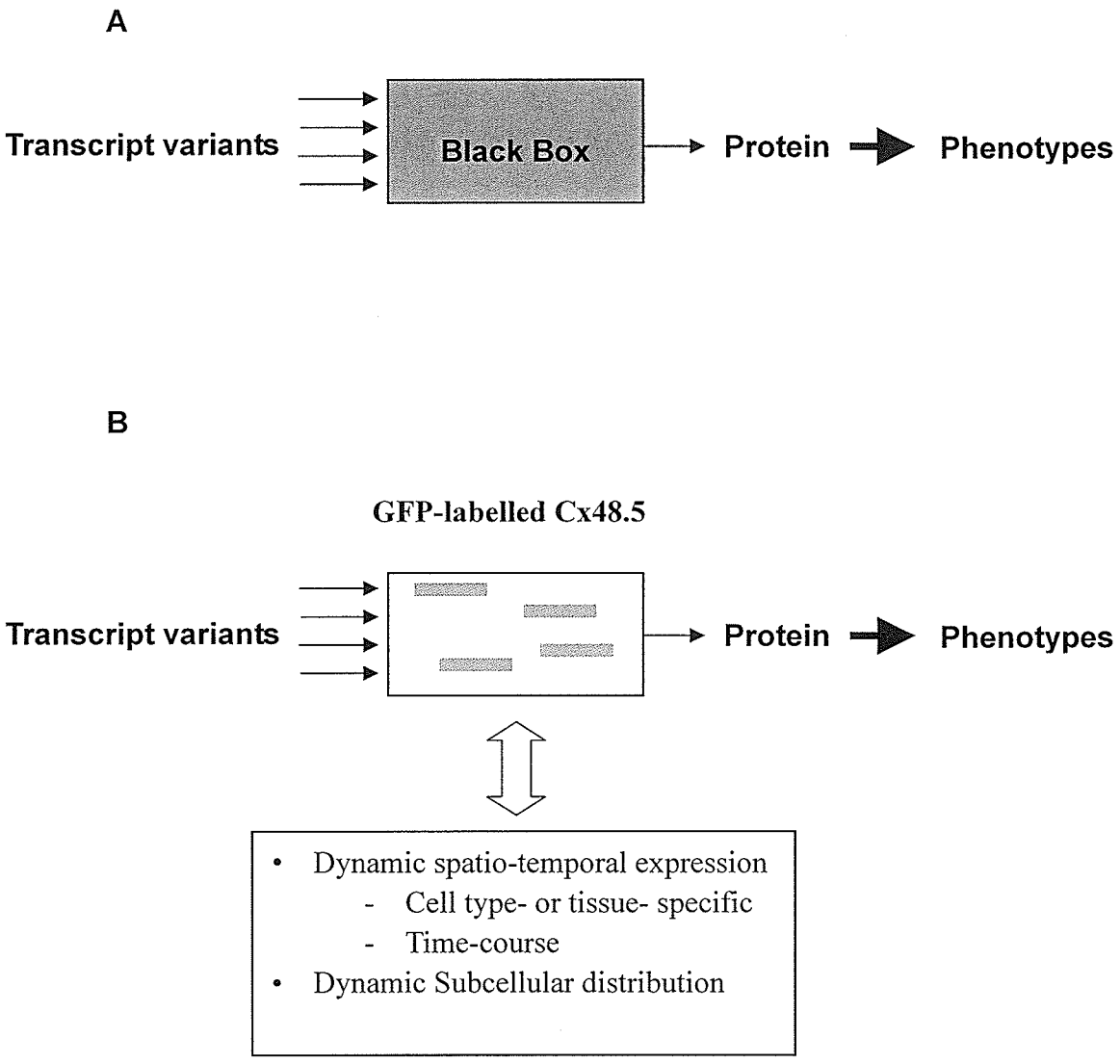


Figure A2-1. Relationship between transcripts and potential phenotypes.

As discussed in Chapter 5, the genomic organization of a cluster of three connexins, Cx30, Cx26, and Cx46, has also been found physically linked, all with the same gene orientation, on human chromosome 13; and in zebrafish, Cx30.3 (ortholog of

human Cx26) and Cx48.5 (ortholog of human Cx46) are physically linked on chromosome 9, in a similar organization to that of human Cx26 and Cx46 (Figure 5-4). Based on these similarities (ortholog connexin genes, and similar genomic cluster organizations), I hypothesized that the Cx30.3 and Cx48.5 gene cluster could be an excellent model to investigate regulatory mechanisms for the diversity of expression and function between connexin genes.

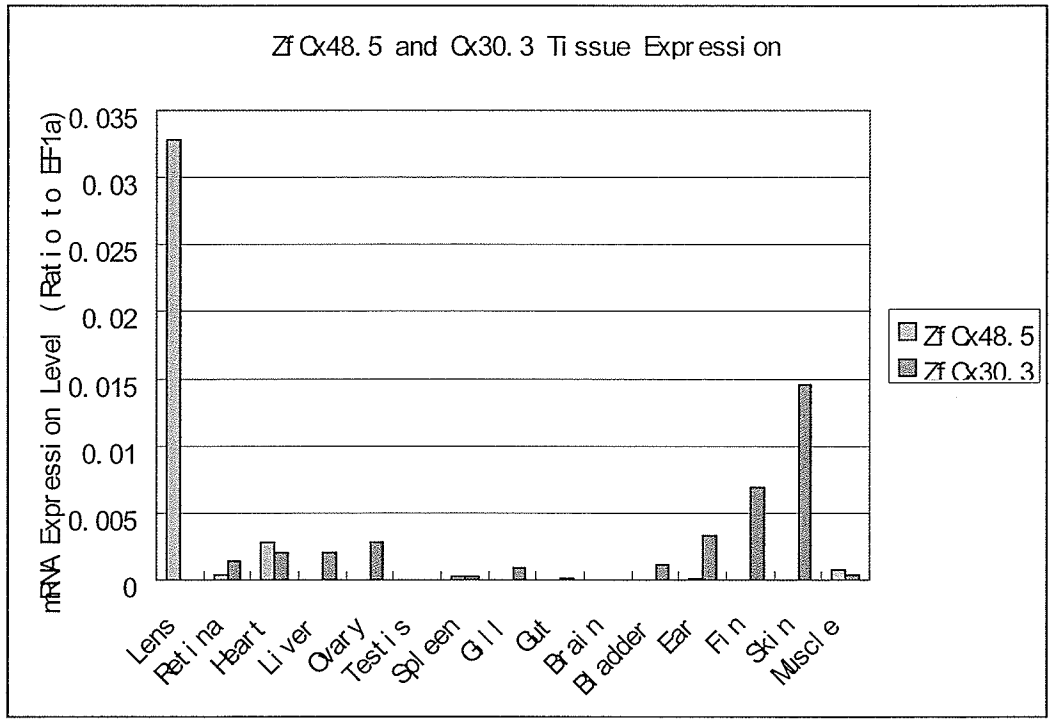


Figure A2-2. Tissue distribution of ZfCx48.5 and ZfCx30.3 mRNAs. The expression ratios were calculated by comparing ΔCt between mean Cts of test transcripts and zebrafish EF1α mRNA in two replicate experiments, and presented as a column graph.

The distance from the poly (A) signal of Cx30.3 to the transcriptional start sites of Cx48.5 is only about 600 nucleotides, and notably, both connexin genes are in the same orientation. Cx48.5 is predominantly expressed and functions in the zebrafish lens as well as in the heart during development and in adult tissues. Expression of Cx48.5 was also detected in the otic vesicle and later in the inner ear during embryonic development (Cheng et al. 2004). In adult tissues, it appears that both connexin genes have distinct and partially overlapping expression patterns (Figure A2-2). Considering the above, we hypothesize that there might be some mechanisms for co-regulation of the expression of both connexin genes, or some sequence elements of both genes could function as regulatory sequences for each other. To address such questions will require further detailed studies of the Cx30.3-Cx48.5 gene cluster.

For these studies, I proposed to establish Cx30.3-Cx48.5 double fluorescent protein labelled stable transgenic zebrafish lines. In addition, the expected stable transgenic fish can also be used for functional analyses of ZfCx30.3 using morpholino knock-down.

ZfCx30.3-RFP/Cx48.5-eGFP Expression Plasmid

I designed and constructed a recombinant DNA plasmid that contains a ~15kb fragment of this genome cluster. In this plasmid, the coding sequences of the Cx30.3 and Cx48.5 genes are replaced by Cx30.3-DsRFP and Cx48.5-GFP fusion coding sequences, respectively, so that protein products of these two genes are labelled by different

fluorescent markers, but without modifying any regulatory components (Figure A2-3).

I previously isolated and characterized an EcoRI- pBlueScript sub-clone from a Cx48.5 positive genomic PAC clone (see Chapter 4 & 5). This sub-clone contains a ~12kb genomic fragment of the gene cluster of interest, but is missing the Cx48.5 coding region and 3'UTR. This sub-clone was digested with HindIII and then re-ligated to remove the 5' upstream EcoRI site, and then digested by PstI to remove part of Cx30.3 sequence. This intermediate plasmid was used as a "vector plasmid" for the following cloning steps. Briefly, RFP and EGFP coding sequences were retrieved by PCR from a pCMV-DsRed express vector (gift from Dr. Steve Whyard, University of Manitoba) and RARE3E(-SP6) express vector (gift from Dr. Robert Davis, Harvard Medical School), respectively. Fusion coding sequences (Cx30.3-RFP and Cx48.5-eGFP) were generated by PCR, and then linked to their 3'UTR sequences by one more round of PCR, respectively. The resulting recombinant DNA fragments were then digested with suitable restriction enzymes and cloned into the above "vector plasmid". For Cx30.3-RFP, PstI sites were used, while for Cx48.5-eGFP, EcoR1-NotI sites were used (Figure A2-3). Sequences surrounding all cloning sites were confirmed by sequencing. In the final plasmid construct, all endogenous components of the Cx30.3-Cx48.5 gene cluster remain intact, except that their coding sequences were tagged by fluorescent proteins.

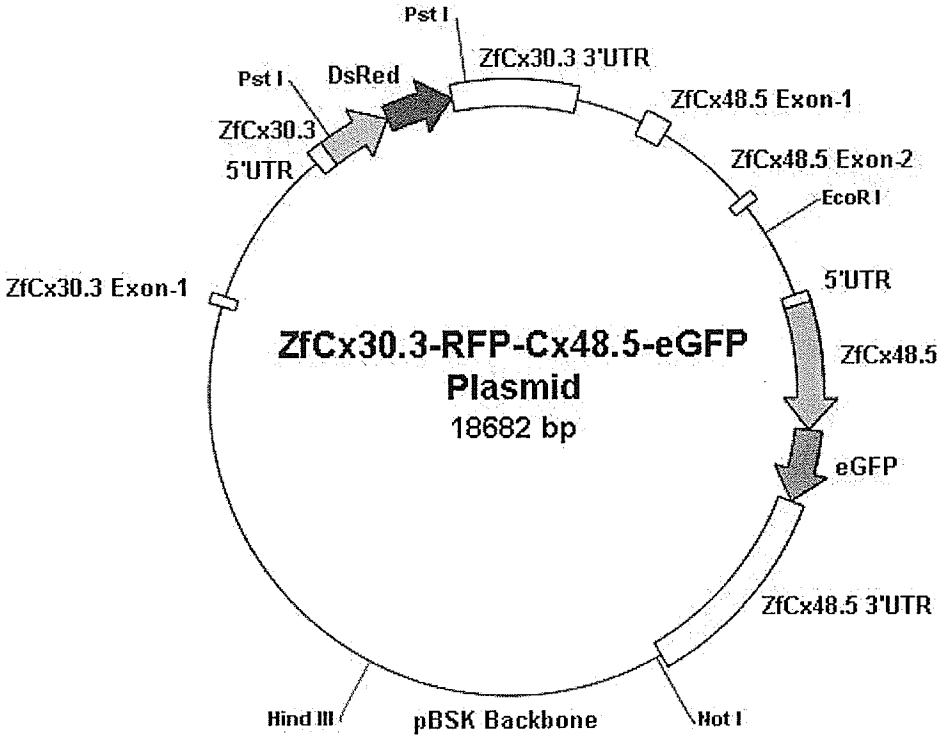


Figure A2-3. ZfCx30.3-RFP/Cx48.5-eGFP plasmid

Design of Transcript Variant-specific Morpholino Oligos for ZfCx48.5

I have designed a set of morpholino oligos (MO 3-8) that specifically target individual transcript isoforms of zebrafish Cx48.5, as shown in Figure A2-4 (red or blue arrows). Sequences of these six morpholino oligos are listed in Table A2-1.

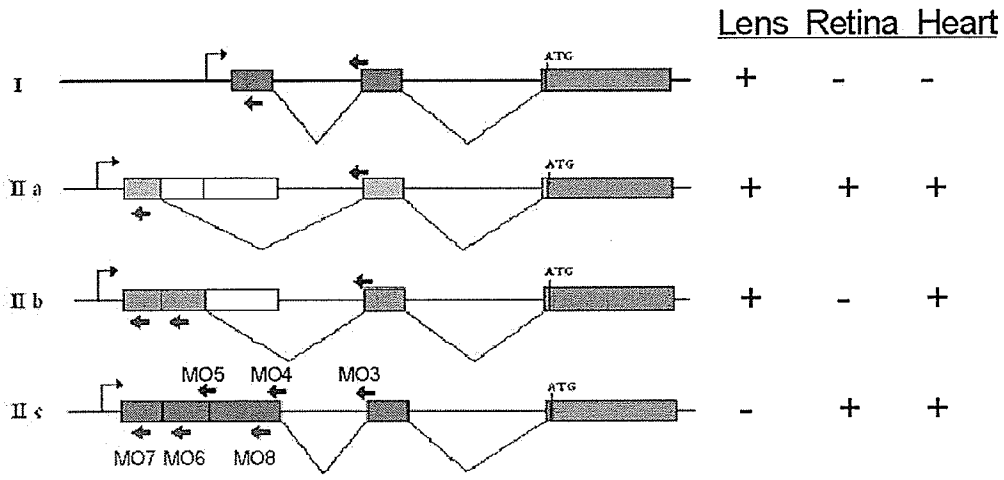


Figure A2-4. Design of ZfCx48.5 morpholino oligos. The left panel shows the 4 transcript variants and their associated morpholinos, while the right panel indicates the tissues in which the variants have been detected by 5' RACE.

Morpholino oligos (MO3, 4, 5) can modify pre-mRNA splicing events, as they were designed to span specific exon-exon boundaries. Morpholino oligos MO6, MO7, and MO8 can each specifically target a fraction of these transcript variants. This design makes it possible to use one (or a subset of) morpholinos to modify the ratio (or composition) of Cx48.5 transcripts in early fish embryos. By using Cx48.5-EGFP stable transgenic zebrafish embryos, the disruption (or interference) of expression of Cx48.5 transcripts will be traced in real time, and potential phenotypes will be assayed during early embryonic development.

Design of Morpholino Oligos for ZfCx30.3

In order to investigate the role of zebrafish Cx30.3 in embryonic development, we proposed to perform morpholino analyses (as discussed in Chapter 5). I designed two morpholino oligos for this purpose (Figure A2-5). Sequences of both morpholino oligos are listed in Table A2-1. MO1 was designed to span the 3' splicing site (boundary of the intron and the second exon), and thus it will interfere with the pre-mRNA splicing and result in non-functional transcripts. MO2 was designed to target the first exon, and thus interfere with the initiation of translation of Cx30.3 mature mRNAs. Co-injection of both morpholino oligos into embryos is expected to have synergetic effects, thus ensuring sufficient specificity. Functional analyses of ZfCx30.3 await the establishment of stable transgenic fish lines described in this Appendix.

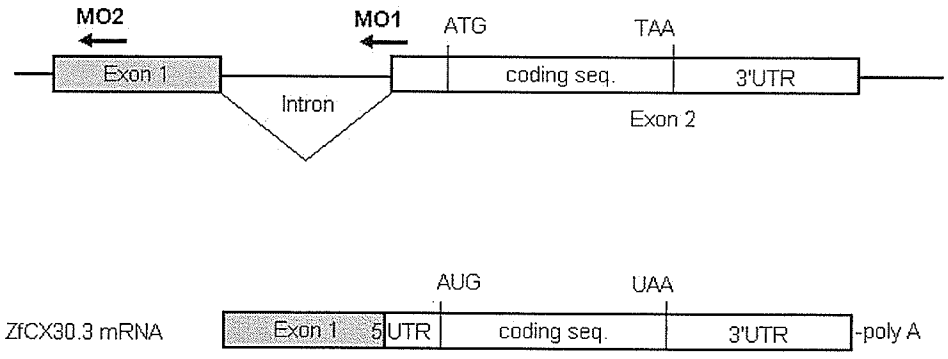


Figure A2-5. Design of ZfCx30.3 morpholino oligos.

Transient Expression of ZfCx30.3-RFP/Cx48.5-eGFP Plasmid

Transient expression experiments were performed to test the capability of the

Cx30.3-RFP/Cx48.5-eGFP plasmid to drive fluorescent protein-tagged expression of both connexin genes (Figure A2-6). Cx30.3 and Cx48.5 show distinct but overlapping expressions.

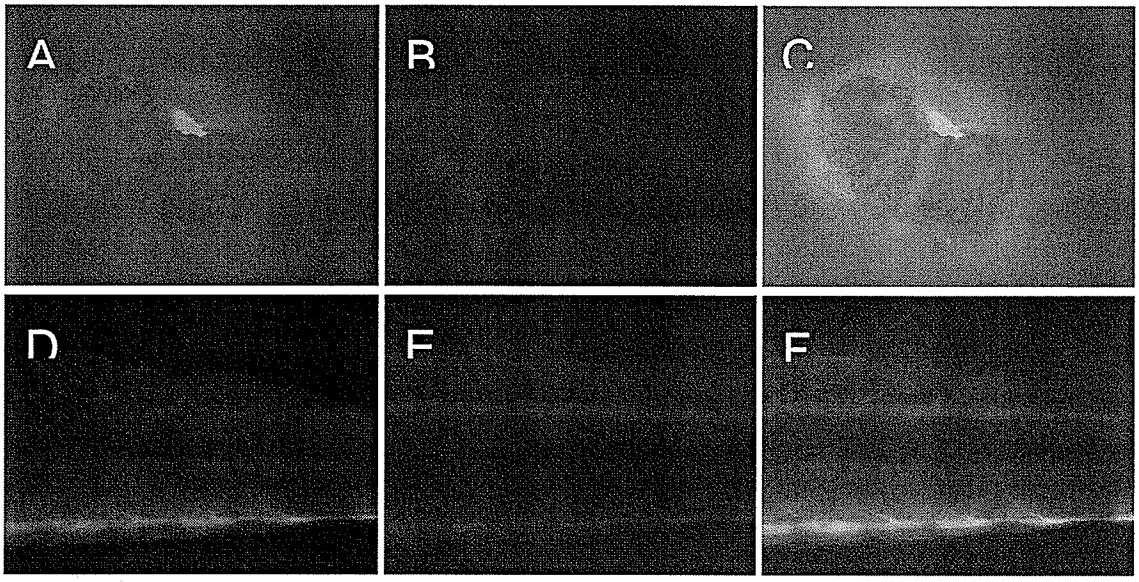


Figure A2-6. Transient expression of Cx30.3-RFP/Cx48.5-eGFP construct. (**A**) Cx48.5-eGFP expression in the lens. (**B**) Cx30.3-RFP expression in the retina. (**C**) overlay image. (**D**) EGFP expression in the vascular system. (**E**) DsRed expression in vascular system. (**F**) overlay image.

Table A2-1. Morpholino oligos for zebrafish Cx48.5 and Cx30.3.

Name	Sequence (5' – 3')
Cx48.5 – MO3	AAG UCC UUA AUU ACA UGA UAA UAA A
Cx48.5 – MO4	AAC UUG AGG CAU ACC UUG ACU GAA G
Cx48.5 – MO5	GCA GCA CUG CCA GCU UAC CAA ACU C
Cx48.5 – MO6	AAG UUC UUC ACA CAG AGA CCC GAG G
Cx48.5 – MO7	CAA GCG CAG CUG UCU CCC GUU CCC A
Cx48.5 – MO8	AUC UAA GCA AAG CAA UAU CCA GCG G
Cx30.3 – MO1	UGU CUACUG GUG CUC UCA GCU CCU G
Cx30.3 – MO2	CAA AUC ACU UGU GUG GUG CCU CUG AG

APPENDIX-III

STUDIES ON ZEBRAFISH CX43

Transcriptional start sites and the 5' UTR sequence of zebrafish Cx43 were identified by 5' RACE technique using mRNA samples from adult lens tissue. Results were published in 2005. For details, please see the below article.

Bishwanath Chatterjee*, Alvin J. Chin*, Gunnar Valdimarsson*, Jennifer M. Sonntag, Carla Finis, Linda Choi, Liang Tao, Krithika Balasubramanian, Carolyn Bell, Alison krufka, David J. Kozlowski, Ross G. Johnson, and Cecilia W. Lo.. (2005) Evolutionary conservation of structure and developmental regulation of the zebrafish connexin43 gene. *Developmental Dynamics*. 233(3): 890-906. (* these authors contribute equally)

APPENDIX-IV

ELECTRO-PHYSIOLOGICAL PROPERTIES OF ZFCX30.3

This work was conducted in collaboration with Dr. T. W. White (Department of Physiology and Biophysics, SUNY, Stony Brook) in order to examine the physiological properties of the gap junction channels composed of Cx30.3 using patch clump techniques with the *Xenopus* oocytes expression system.

Table A4-1. Boltzmann parameters for zebrafish Cx30.3 gap junctions.

	V_j	V_0	G_{jmin}	A
Cx30.3	+	91	0.51	0.05
Cx30.3	-	-93	0.43	0.04

G_{jmin} represents the minimum conductance value, V_0 indicates the voltage measured midway through the G_j decline, and A denotes the cooperativity constant, reflecting the number of charges moving through the transjunctional field. Signs + and – for V_j indicate transjunctional membrane potential polarity.

REFERENCES

Adam A, Bartfai R, Lele Z, Krone PH, and Orban L. (2000)"Heat-Inducible Expression of a Reporter Gene Detected by Transient Assay in Zebrafish." *Experimental cell research* 256, 1: 282-90.

Ai Z, Fischer A, Spray DC, Brown AMC, and Fishman GI. (2000)"Wnt-1 Regulation of connexin43 in Cardiac Myocytes." *Journal of Clinical Investigation* 105, 2: 161-71.

Amatruda JF, Shepard JL, Stern HM, and Zon LI. (2002)"Zebrafish as a Cancer Model System." *Cancer cell* 1, 3: 229-31.

Amsterdam A, and Becker TS. (2005)"Transgenes as Screening Tools to Probe and Manipulate the Zebrafish Genome." *Developmental dynamics : an official publication of the American Association of Anatomists* 234, 2: 255-68.

Amsterdam A, Burgess S, Golling G, Chen W, Sun Z, Townsend K, Farrington S, Haldi M, and Hopkins N. (1999)"A Large-Scale Insertional Mutagenesis Screen in Zebrafish." *Genes & development* 13, 20: 2713-24.

Amsterdam A, and Hopkins N. (2004)"Retroviral-Mediated Insertional Mutagenesis in Zebrafish." *Methods in Cell Biology* 2004, 77: 3-20.

Amsterdam A, and Hopkins N. (1999)"Retrovirus-Mediated Insertional Mutagenesis in Zebrafish." *Methods in cell biology* 60, : 87-98.

Amsterdam A, Lin S, and Hopkins N. (1995)"The Aequorea Victoria Green Fluorescent Protein can be used as a Reporter in Live Zebrafish Embryos." *Developmental biology* 171, 1: 123-9.

Amsterdam A. (2006)"Insertional Mutagenesis in Zebrafish: Genes for Development, Genes for Disease." *Briefings in functional genomics & proteomics* 5, 1: 19-23.

Anderson CL, Zundel MA, and Werner R. (2005)"Variable Promoter Usage and Alternative Splicing in Five Mouse Connexin Genes." *Genomics* 85, 2: 238-44.

Andolfatto P. (2005)"Adaptive Evolution of Non-Coding DNA in Drosophila." *Nature* 437, 7062: 1149-52.

Arora V, Knapp DC, Smith BL, Statdfield ML, Stein DA, Reddy MT, Weller DD, and Iversen PL. (2000)"C-Myc Antisense Limits Rat Liver Regeneration and Indicates

Role for c-Myc in Regulating Cytochrome P-450 3A Activity." *The Journal of pharmacology and experimental therapeutics* 292, 3: 921-8.

Austin CP, Battey JF, Bradley A, Bucan M, Capecchi M, Collins FS, Dove WF, Duyk G, Dymecki S, Eppig JT, et al. (2004)"The Knockout Mouse Project." *Nature genetics* 36, 9: 921-4.

Bagwe S, Berenfeld O, Vaidya D, Morley GE, and Jalife J. (2005)"Altered Right Atrial Excitation and Propagation in connexin40 Knockout Mice." *Circulation* 112, 15: 2245-53.

Bajoghli B, Aghaallaei N, Heimbucher T, and Czerny T. (2004)"An Artificial Promoter Construct for Heat-Inducible Misexpression during Fish Embryogenesis." *Developmental biology* 271, 2: 416-30.

Balciunas D, Davidson AE, Sivasubbu S, Hermanson SB, Welle Z, and Ekker SC. (2004)"Enhancer Trapping in Zebrafish using the Sleeping Beauty Transposon." *BMC Genomics* 5, : .

Baranova A, Ivanov D, Petrash N, Pestova A, Skoblov M, Kelmanson I, Shagin D, Nazarenko S, Geraymovych E, Litvin O, et al. (2004)"The Mammalian Pannexin Family is Homologous to the Invertebrate Innexin Gap Junction Proteins." *Genomics* 83, 4: 706-16.

Barbe MT, Monyer H, and Bruzzone R. (2006)"Cell-Cell Communication Beyond Connexins: The Pannexin Channels." *Physiology (Bethesda, Md.)* 21, : 103-14.

Bashirullah A, Cooperstock RL, and Lipshitz HD. (2001)"Spatial and Temporal Control of RNA Stability." *Proceedings of the National Academy of Sciences of the United States of America* 98, 13: 7025-8.

Bassett DI, and Currie PD. (2003)"The Zebrafish as a Model for Muscular Dystrophy and Congenital Myopathy." *Human molecular genetics* 12 Spec No 2, : R265-70.

Bennett MV, Contreras JE, Bukauskas FF, and Saez JC. (2003)"New Roles for Astrocytes: Gap Junction Hemichannels have Something to Communicate." *Trends in neurosciences* 26, 11: 610-7.

Bennett MVL, and Goodenough DA. (1978)"Gap Junctions, Electrotonic Coupling, and Intercellular Communication." *Neurosciences Research Program Bulletin* 16, 3: 373-486.

Berghmans S, Jette C, Langenau D, Hsu K, Stewart R, Look T, and Kanki JP. (2005)"Making Waves in Cancer Research: New Models in the Zebrafish." *BioTechniques* 39, 2: 227-37.

Berthoud VM, Bassnett S, and Beyer EC. (1999)"Cultured Chicken Embryo Lens Cells Resemble Differentiating Fiber Cells in Vivo and Contain Two Kinetic Pools of connexin56." *Experimental eye research* 68, 4: 475-84.

Berthoud VM, and Beyer EC. (2008)"Oxidative Stress, Lens Gap Junctions, and Cataracts." *Antioxidants & redox signaling* , : .

Beyer EC, Kistler J, Paul DL, and Goodenough DA. (1989)"Antisera Directed Against connexin43 Peptides React with a 43-kD Protein Localized to Gap Junctions in Myocardium and Other Tissues." *The Journal of cell biology* 108, 2: 595-605.

Beyer EC, Paul DL, and Goodenough DA. (1990)"Connexin Family of Gap Junction Proteins." *The Journal of membrane biology* 116, 3: 187-94.

Black DL. (2003)"Mechanisms of Alternative Pre-Messenger RNA Splicing." *Annual Review of Biochemistry* 72, : 291-336.

Bleeker WK, Mackaay AJ, Masson-Pevet M, Bouman LN, and Becker AE. (1980)"Functional and Morphological Organization of the Rabbit Sinus Node." *Circulation research* 46, 1: 11-22.

Blencowe BJ. (2000)"Exonic Splicing Enhancers: Mechanism of Action, Diversity and Role in Human Genetic Diseases." *Trends in biochemical sciences* 25, 3: 106-10.

Bondurand N, Girard M, Pingault V, Lemort N, Dubourg O, and Goossens M. (2001)"Human Connexin 32, a Gap Junction Protein Altered in the X-Linked Form of Charcot-Marie-Tooth Disease, is Directly Regulated by the Transcription Factor SOX10." *Human Molecular Genetics* 10, 24: 2783-95.

Brinster RL, Allen JM, Behringer RR, Gelinas RE, and Palmiter RD. (1988)"Introns Increase Transcriptional Efficiency in Transgenic Mice." *Proceedings of the National Academy of Sciences of the United States of America* 85, 3: 836-40.

Bruneau BG, Nemer G, Schmitt JP, Charron F, Robitaille L, Caron S, Conner DA, Gessler M, Nemer M, Seidman CE, et al. (2001)"A Murine Model of Holt-Oram Syndrome Defines Roles of the T-Box Transcription Factor Tbx5 in Cardiogenesis and

Disease." *Cell* 106, 6: 709-21.

Bruzzone R, White TW, and Goodenough DA. (1996a)"The Cellular Internet: On-Line with Connexins." *BioEssays: news and reviews in molecular, cellular and developmental biology* 18, 9: 709-18.

Bruzzone R, White TW, and Paul DL. (1996b)"Connections with Connexins: The Molecular Basis of Direct Intercellular Signaling." *European journal of biochemistry / FEBS* 238, 1: 1-27.

Bukauskas FF, Elfgang C, Willecke K, and Weingart R. (1995)"Heterotypic Gap Junction Channels (connexin26 - connexin32) Violate the Paradigm of Unitary Conductance." *Pflugers Archiv European Journal of Physiology* 429, 6: 870-2.

Buratti E, and Baralle FE. (2004)"Influence of RNA Secondary Structure on the Pre-mRNA Splicing Process." *Molecular and cellular biology* 24, 24: 10505-14.

Burgar HR, Burns HD, Elsden JL, Lalioti MD, and Heath JK. (2002)"Association of the Signaling Adaptor FRS2 with Fibroblast Growth Factor Receptor 1 (FGFR1) is Mediated by Alternative Splicing of the Juxtamembrane Domain." *Journal of Biological Chemistry* 277, 6: 4018-23.

Burge CB, Tuschl T, Sharp PA. (1999). In *The RNA World* (2nd ed), ed. RF Gesteland, TR Cech, JF Atkins, pp. 525 – 60. Cold Spring Harbor, NY: Cold Spring Harb. Lab.

Caceres JF, and Kornblihtt AR. (2002)"Alternative Splicing: Multiple Control Mechanisms and Involvement in Human Disease." *Trends in genetics : TIG* 18, 4: 186-93.

Campbell MA, Haas BJ, Hamilton JP, Mount SM, and Robin CR. (2006)"Comprehensive Analysis of Alternative Splicing in Rice and Comparative Analyses with Arabidopsis." *BMC Genomics* 7, : .

Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C, et al. (2005)"The Transcriptional Landscape of the Mammalian Genome." *Science (New York, N.Y.)* 309, 5740: 1559-63.

Carpio Y, and Estrada MP. (2006)"Zebrafish as a Genetic Model Organism." *Biotecnologia Aplicada* 23, 4: 265-70.

Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M,

Bayerlein M, and Werner T. (2005)"MatInspector and Beyond: Promoter Analysis Based on Transcription Factor Binding Sites." *Bioinformatics (Oxford, England)* 21, 13: 2933-42.

Carystinos GD, Kandouz M, Alaoui-Jamali MA, and Batist G. (2003)"Unexpected Induction of the Human Connexin 43 Promoter by the Ras Signaling Pathway is Mediated by a Novel Putative Promoter Sequence." *Molecular Pharmacology* 63, 4: 821-31.

Cason N, White TW, Cheng S, Goodenough DA, and Valdimarsson G. (2001)"Molecular Cloning, Expression Analysis, and Functional Characterization of connexin44.1: A Zebrafish Lens Gap Junction Protein." *Developmental dynamics : an official publication of the American Association of Anatomists* 221, 2: 238-47.

Chatterjee B, Chin AJ, Valdimarsson G, Finis C, Sonntag JM, Choi BY, Tao L, Balasubramanian K, Bell C, Krufka A, et al. (2005)"Developmental Regulation and Expression of the Zebrafish connexin43 Gene." *Developmental dynamics: an official publication of the American Association of Anatomists* 233, 3: 890-906.

Chen J-, Cheng Y-, Chou M-, Sen-Lin T, Lai W-, Ho WL, and Lee H. (2003)"The Correlation between Aberrant Connexin 43 mRNA Expression Induced by Promoter Methylation and Nodal Micrometastasis in Non-Small Cell Lung Cancer." *Clinical Cancer Research* 9, 11: 4200-4.

Chen J-. (1998)"Isolation and Characterization of Tilapia (Oreochromis Mossambicus) Insulin-Like Growth Factors Gene and Proximal Promoter Region." *DNA and Cell Biology* 17, 4: 359-76.

Chen S, Anderson K, and Moore MJ. (2000)"Evidence for a Linear Search in Bimolecular 3′ Splice Site AG Selection." *Proceedings of the National Academy of Sciences of the United States of America* 97, 2: 593-8.

Chen W, Burgess S, Golling G, Amsterdam A, and Hopkins N. (2002)"High-Throughput Selection of Retrovirus Producer Cell Lines Leads to Markedly Improved Efficiency of Germ Line-Transmissible Insertions in Zebra Fish." *Journal of Virology* 76, 5: 2192-8.

Chen Y-, Lee W-, Liu C-, and Tsai H-. (2001)"Molecular Structure, Dynamic Expression, and Promoter Analysis of Zebrafish (Danio Rerio) Myf-5 Gene." *Genesis* 29, 1: 22-35.

Chen Z-, Lefebvre D, Bai X-, Reaume A, Rossant J, and Lye SJ. (1995)"Identification of Two Regulatory Elements within the Promoter Region of the Mouse Connexin 43 Gene." *Journal of Biological Chemistry* 270, 8: 3863-8.

Cheng S, Christie T, and Valdimarsson G. (2003)"Expression of connexin48.5, connexin44.1, and connexin43 during Zebrafish (Danio Rerio) Lens Development." *Developmental dynamics : an official publication of the American Association of Anatomists* 228, 4: 709-15.

Cheng S, Shakespeare T, Mui R, White TW, and Valdimarsson G. (2004)"Connexin 48.5 is Required for Normal Cardiovascular Function and Lens Development in Zebrafish Embryos." *The Journal of biological chemistry* 279, 35: 36993-7003.

Chern TM, van Nimwegen E, Kai C, Kawai J, Carninci P, Hayashizaki Y, and Zavolan M. (2006)"A Simple Physical Model Predicts Small Exon Length Variations." *PLoS genetics.* 2, 4: .

Chico TJ, Ingham PW, and Crossman DC. (2008)"Modeling Cardiovascular Disease in the Zebrafish." *Trends in cardiovascular medicine* 18, 4: 150-5.

Chow RL, and Lang RA. (2001)"Early Eye Development in Vertebrates." *Annual Review of Cell and Developmental Biology* 17, : 255-96.

Christ GJ, Spray DC, el-Sabban M, Moore LK, and Brink PR. (1996)"Gap Junctions in Vascular Tissues. Evaluating the Role of Intercellular Communication in the Modulation of Vasomotor Tone." *Circulation research* 79, 4: 631-46.

Christie TL, Mui R, White TW, and Valdimarsson G. (2004)"Molecular Cloning, Functional Analysis, and RNA Expression Analysis of connexin45.6: A Zebrafish Cardiovascular Connexin." *American journal of physiology.Heart and circulatory physiology* 286, 5: H1623-32.

Chua K, and Reed R. (1999)"The RNA Splicing Factor hSlu7 is Required for Correct 3' Splice-Site Choice." *Nature* 402, 6758: 207-10.

Chua K, and Reed R. (2001)"An Upstream AG Determines Whether a Downstream AG is Selected during Catalytic Step II of Splicing." *Molecular and Cellular Biology* 21, 5: 1509-14.

Civitelli R, Ziambaras K, Warlow PM, Lecanda F, Nelson T, Harley J, Atal N, Beyer EC, and Steinberg TH. (1997)"Regulation of connexin43 Expression and Function by

Prostaglandin E2 (PGE2) and Parathyroid Hormone (PTH) in Osteoblastic Cells." *Journal of Cellular Biochemistry* 68, 1: 8-21.

Clairmont A, and Sies H. (1997)"Evidence for a Posttranscriptional Effect of Retinoic Acid on connexin43 Gene Expression Via the 3'-Untranslated Region." *FEBS Letters* 419, 2-3: 268-70.

Clairmont A, Tessmann D, and Sies H. (1996)"Analysis of connexin43 Gene Expression Induced by Retinoic Acid in F9 Teratocarcinoma Cells." *FEBS Letters* 397, 1: 22-4.

Clough DW, Lepinske HM, Davidson RL, and Storti RV. (1985)"Drosophila P Element-Enhanced Transfection in Mammalian Cells." *Molecular and cellular biology* 5, 4: 898-901.

Collas P, and Aleströ̈m P. (1998)"Nuclear Localization Signals Enhance Germline Transmission of a Transgene in Zebrafish." *Transgenic Research* 7, 4: 303-9.

Collas P, and Alestrom P. (1997)"Nuclear Localization Signals: A Driving Force for Nuclear Transport of Plasmid DNA in Zebrafish." *Biochemistry and cell biology = Biochimie et biologie cellulaire* 75, 5: 633-40.

Condorelli DF, Parenti R, Spinella F, Salinaro AT, Belluardo N, Cardile V, and Cicirata F. (1998)"Cloning of a New Gap Junction Gene (Cx36) Highly Expressed in Mammalian Brain Neurons." *European Journal of Neuroscience* 10, 3: 1202-8.

Coppen SR, and Severs NJ. (2002)"Diversity of Connexin Expression Patterns in the Atrioventricular Node: Vestigial Consequence Or Functional Specialization?" *Journal of cardiovascular electrophysiology* 13, 6: 625-6.

Corey DR, and Abrams JM. (2001)"Morpholino Antisense Oligonucleotides: Tools for Investigating Vertebrate Development." *Genome biology* 2, 5: REVIEWS1015.

Cretekos CJ, and Grunwald DJ. (1999)"Alyron, an Insertional Mutation Affecting Early Neural Crest Development in Zebrafish." *Developmental Biology* 210, 2: 322-38.

Crosier PS, Bardsley A, Horsfield JA, Krassowska AK, Lavallie ER, Collins-Racie LA, Postlethwait JH, Yan YL, McCoy JM, and Crosier KE. (2001)"In Situ Hybridization Screen in Zebrafish for the Selection of Genes Encoding Secreted Proteins." *Developmental dynamics : an official publication of the American Association of Anatomists* 222, 4: 637-44.

Cruciani V, and Mikalsen SO. (2007)"Evolutionary Selection Pressure and Family Relationships among Connexin Genes." *Biological chemistry* 388, 3: 253-64.

Cruciani V, and Mikalsen SO. (2006)"The Vertebrate Connexin Family." *Cellular and molecular life sciences : CMLS* 63, 10: 1125-40.

Cserzo M, Wallin E, Simon I, von Heijne G, and Elofsson A. (1997)"Prediction of Transmembrane Alpha-Helices in Prokaryotic Membrane Proteins: The Dense Alignment Surface Method." *Protein engineering* 10, 6: 673-6.

Culp P, Nusslein-Volhard C, and Hopkins N. (1991)"High-Frequency Germ-Line Transmission of Plasmid DNA Sequences Injected into Fertilized Zebrafish Eggs." *Proceedings of the National Academy of Sciences of the United States of America* 88, 18: 7953-7.

Cvekl A, and Piatigorsky J. (1996)"Lens Development and Crystallin Gene Expression: Many Roles for Pax-6." *BioEssays : news and reviews in molecular, cellular and developmental biology* 18, 8: 621-30.

Dahl G, Werner R, Levine E, and Rabadan-Diehl C. (1992)"Mutational Analysis of Gap Junction Formation." *Biophysical journal* 62, 1: 172,80; discussion 180-2.

Dahm R, and Geisler R. (2006)"Learning from Small Fry: The Zebrafish as a Genetic Model Organism for Aquaculture Fish Species." *Marine biotechnology (New York, N.Y.)* 8, 4: 329-45.

Dahm R, van Marle J, Prescott AR, and Quinlan RA. (1999)"Gap Junctions Containing alpha8-Connexin (MP70) in the Adult Mammalian Lens Epithelium Suggests a Re-Evaluation of its Role in the Lens." *Experimental eye research* 69, 1: 45-56.

Dambly-Chaudiere C, Sapede D, Soubiran F, Decorde K, Gompel N, and Ghysen A. (2003)"The Lateral Line of Zebrafish: A Model System for the Analysis of Morphogenesis and Neural Development in Vertebrates." *Biology of the cell / under the auspices of the European Cell Biology Organization* 95, 9: 579-87.

Davidson AE, Balciunas D, Mohn D, Shaffer J, Hermanson S, Sivasubbu S, Cliff MP, Hackett PB, and Ekker SC. (2003)"Efficient Gene Delivery and Gene Expression in Zebrafish using the Sleeping Beauty Transposon." *Developmental Biology* 263, 2: 191-202.

Davidson EH, 2001, *Genomic Regulatory Systems: Development and Evolution,*

Academic, San Diego

Davidson EH, 2006, *The Regulatory Genome: Gene Regulatory Network in Development and Evolution.*(Academic, New York.

Davis LM, Rodefeld ME, Green K, Beyer EC, and Saffitz JE. (1995)"Gap Junction Protein Phenotypes of the Human Heart and Conduction System." *Journal of cardiovascular electrophysiology* 6, 10 Pt 1: 813-22.

Davison JM, Akitake CM, Goll MG, Rhee JM, Gosse N, Baier H, Halpern ME, Leach SD, and Parsons MJ. (2007)"Transactivation from Gal4-VP16 Transgenic Insertions for Tissue-Specific Cell Labeling and Ablation in Zebrafish." *Developmental biology* 304, 2: 811-24.

de Moor CH, Meijer H, and Lissenden S. (2005)"Mechanisms of Translational Control by the 3' UTR in Development and Differentiation." *Seminars in cell & developmental biology* 16, 1: 49-58.

den Hertog J. (2005)"Chemical Genetics: Drug Screens in Zebrafish." *Bioscience reports* 25, 5-6: 289-97.

Deng TL, Li DW, Jenh CH, and Johnson LF. (1986)"Structure of the Gene for Mouse Thymidylate Synthase. Locations of Introns and Multiple Transcriptional Start Sites." *The Journal of biological chemistry* 261, 34: 16000-5.

Derrigo M, Cestelli A, Savettieri G, and Di Liegro I. (2000)"RNA-Protein Interactions in the Control of Stability and Localization of Messenger RNA (Review)." *International journal of molecular medicine* 5, 2: 111-23.

Ding Y, Chan CY, and Lawrence CE. (2005)"RNA Secondary Structure Prediction by Centroids in a Boltzmann Weighted Ensemble." *RNA (New York, N.Y.)* 11, 8: 1157-66.

Ding Y, and Lawrence CE. (2003)"A Statistical Sampling Algorithm for RNA Secondary Structure Prediction." *Nucleic acids research* 31, 24: 7280-301.

Dobrowolski R and Willecke K. (2009) "Connexin-Caused Genetic Diseases and Corresponding Mouse Models" *Antioxidants & Redox Signaling* 11, 2: 283-95.

Dou Y, Fox-Walsh KL, Baldi PF, and Hertel KJ. (2006)"Genomic Splice-Site Analysis Reveals Frequent Alternative Splicing Close to the Dominant Splice Site." *RNA* 12,

12: 2047-56.

Draper BW, McCallum CM, Stout JL, Slade AJ, and Moens CB. (2004)"A High-Throughput Method for Identifying N-Ethyl-N-Nitrosourea (ENU)-Induced Point Mutations in Zebrafish." *Methods in cell biology* 77, : 91-112.

Draper BW, Morcos PA, and Kimmel CB. (2001)"Inhibition of Zebrafish fgf8 Pre-mRNA Splicing with Morpholino Oligos: A Quantifiable Method for Gene Knockdown." *Genesis (New York, N.Y.: 2000)* 30, 3: 154-6.

Dreyfuss G, Kim VN, and Kataoka N. (2002)"Messenger-RNA-Binding Proteins and the Messages they Carry." *Nature reviews.Molecular cell biology* 3, 3: 195-205.

Driever W, Solnica-Krezel L, Schier AF, Neuhauss SCF, Malicki J, Stemple DL, Stainier DYR, Zwartkruis F, Abdelilah S, Rangini Z, et al. (1996)"A Genetic Screen for Mutations Affecting Embryogenesis in Zebrafish." *Development* 123, : 37-46.

Driever W, Stemple D, Schier A, and Solnica-Krezel L. (1994)"Zebrafish: Genetic Tools for Studying Vertebrate Development." *Trends in genetics : TIG* 10, 5: 152-9.

Drummond IA. (2005)"Kidney Development and Disease in the Zebrafish." *Journal of the American Society of Nephrology : JASN* 16, 2: 299-304.

Du SJ, and Dienhart M. (2001)"Zebrafish Tiggy-Winkle Hedgehog Promoter Directs Notochord and Floor Plate Green Fluorescence Protein Expression in Transgenic Zebrafish Embryos." *Developmental Dynamics* 222, 4: 655-66.

Duga S, Asselta R, Giacco LD, Malcovati M, Ronchi S, Tenchini ML, and Simonic T. (1999)"A New Exon in the 5' Untranslated Region of the connexin32 Gene." *European Journal of Biochemistry* 259, 1-2: 188-96.

Dupays L, Jarry-Guichard T, Mazurais D, Calmels T, Izumo S, Gros D, and Thveniau-Ruissy M. (2005)"Dysregulation of Connexins and Inactivation of NFATc1 in the Cardiovascular System of Nkx2-5 Null Mutants." *Journal of Molecular and Cellular Cardiology* 38, 5: 787-98.

Dupays L, Mazurais D, Rücker-Martin C, Calmels T, Bernot D, Cronier L, Malassiné A, Gros D, and Théveniau-Ruissy M. (2003)"Genomic Organization and Alternative Transcripts of the Human Connexin40 Gene." *Gene* 305, 1: 79-90.

Eastman SD, Chen TH, Falk MM, Mendelson TC, and Iovine MK. (2006)"Phylogenetic

Analysis of Three Complete Gap Junction Gene Families Reveals Lineage-Specific Duplications and Highly Supported Gene Classes." *Genomics* 87, 2: 265-74.

Ebihara L, Xu X, Oberti C, Beyer EC, and Berthoud VM. (1999)"Co-Expression of Lens Fiber Connexins Modifies Hemi-Gap-Junctional Channel Behavior." *Biophysical journal* 76, 1 Pt 1: 198-206.

Ekker SC, and Larson JD. (2001)"Morphant Technology in Model Developmental Systems." *Genesis* 30, 3: 89-93.

Ekker SC. (2000)"Morphants: A New Systematic Vertebrate Functional Genomics Approach." *Yeast (Chichester, England)* 17, 4: 302-6.

Elfgang C, Eckert R, Lichtenberg-Frate H, Butterweck A, Traub O, Klein RA, Hulser DF, and Willecke K. (1995)"Specific Permeability and Selective Formation of Gap Junction Channels in Connexin-Transfected HeLa Cells." *Journal of Cell Biology* 129, 3: 805-17.

Elias LA, Wang DD, and Kriegstein AR. (2007)"Gap Junction Adhesion is Necessary for Radial Migration in the Neocortex." *Nature* 448, 7156: 901-7.

Ellingsen S, Laplante MA, Konig M, Kikuta H, Furmanek T, Hoivik EA, and Becker TS. (2005)"Large-Scale Enhancer Detection in the Zebrafish Genome." *Development (Cambridge, England)* 132, 17: 3799-811.

Essenfelder GM, Larderet G, Waksman G, and Lamartine J. (2005)"Gene Structure and Promoter Analysis of the Human GJB6 Gene Encoding connexin30." *Gene* 350, 1: 33-40.

Essner JJ, Laing JG, Beyer EC, Johnson RG, and Hackett PB,Jr. (1996)"Expression of Zebrafish connexin43.4 in the Notochord and Tail Bud of Wild-Type and Mutant no Tail Embryos." *Developmental biology* 177, 2: 449-62.

Fadool JM, Hartl DL, and Dowling JE. (1998)"Transposition of the Mariner Element from Drosophila Mauritiana in Zebrafish." *Proceedings of the National Academy of Sciences of the United States of America* 95, 9: 5182-6.

Fan L, Crodian J, and Collodi P. (2004a)"Production of Zebrafish Germline Chimeras by using Cultured Embryonic Stem (ES) Cells." *Methods in cell biology* 77, : 113-9.

Fan L, Crodian J, Liu X, Alestrom A, Alestrom P, and Collodi P. (2004b)"Zebrafish

Embryo Cells Remain Pluripotent and Germ-Line Competent for Multiple Passages in Culture." *Zebrafish* 1, 1: 21-6.

Fan L, Moon J, Crodian J, and Collodi P. (2006)"Homologous Recombination in Zebrafish ES Cells." *Transgenic research* 15, 1: 21-30.

Feng H, Langenau DM, Madge JA, Quinkertz A, Gutierrez A, Neuberg DS, Kanki JP, and Look AT. (2007)"Heat-Shock Induction of T-Cell lymphoma/leukaemia in Conditional Cre/lox-Regulated Transgenic Zebrafish." *British journal of haematology* 138, 2: 169-75.

Fessele S, Maier H, Zischek C, Nelson PJ, and Werner T. (2002)"Regulatory Context is a Crucial Part of Gene Function." *Trends in genetics : TIG* 18, 2: 60-3.

Fijnvandraat AC, Lekanne Deprez RH, Christoffels VM, Ruijter JM, and Moorman AFM. (2003)"TBX5 Overexpression Stimulates Differentiation of Chamber Myocardium in P19C16 Embryonic Carcinoma Cells." *Journal of Muscle Research and Cell Motility* 24, 2-3: 211-8.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, and Mello CC. (1998)"Potent and Specific Genetic Interference by Double-Stranded RNA in Caenorhabditis Elegans." *Nature* 391, 6669: 806-11.

Fishman MC. (2001)"Genomics. Zebrafish--the Canonical Vertebrate." *Science (New York, N.Y.)* 294, 5545: 1290-1.

Fisher S, Grice EA, Vinton RM, Bessling SL, Urasaki A, Kawakami K, and McCallion AS. (2006) "Evaluating the biological relevance of putative enhancers using Tol2 transposon-mediated transgenesis in zebrafish." *Nature protocols*, 1, 3: 1297-305.

Francis PJ, Berry V, Moore AT, and Bhattacharya S. (1999)"Lens Biology: Development and Human Cataractogenesis." *Trends in genetics : TIG* 15, 5: 191-6.

Frohnhoefer HG,. (2002) "Table of zebrafish mutations." In: *Zebrafish: A Practical Approach*, Oxford, Oxford University Press, 2002, pp 237-292.

Gabriel H-, Strobl B, Hellmann P, Buettner R, and Winterhager E. (2001)"Organization and Regulation of the Rat Cx31 Gene: Implication for a Crucial Role of the Intron Region." *European Journal of Biochemistry* 268, 6: 1749-59.

Gaiano N, Allende M, Amsterdam A, Kawakami K, and Hopkins N. (1996a)"Highly

Efficient Germ-Line Transmission of Proviral Insertions in Zebrafish." *Proceedings of the National Academy of Sciences of the United States of America* 93, 15: 7777-82.

Gaiano N, Amsterdam A, Kawakami K, Allende M, Becker T, and Hopkins N. (1996b)"Insertional Mutagenesis and Rapid Cloning of Essential Genes in Zebrafish." *Nature* 383, 6603: 829-32.

Gan L, Xiang M, Zhou L, Wagner DS, Klein WH, and Nathans J. (1996)"POU Domain Factor Brn-3b is Required for the Development of a Large Set of Retinal Ganglion Cells." *Proceedings of the National Academy of Sciences of the United States of America* 93, 9: 3920-5.

Gao Y, and Spray DC. (1998)"Structural Changes in Lenses of Mice Lacking the Gap Junction Protein connexin43." *Investigative ophthalmology & visual science* 39, 7: 1198-209.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, and Bairoch A. (2003)"ExPASy: The Proteomics Server for in-Depth Protein Knowledge and Analysis." *Nucleic acids research* 31, 13: 3784-8.

Gebauer F, and Hentze MW. (2004)"Molecular Mechanisms of Translational Control." *Nature reviews.Molecular cell biology* 5, 10: 827-35.

Gehring WJ, and Ikeo K. (1999)"Pax 6: Mastering Eye Morphogenesis and Eye Evolution." *Trends in genetics : TIG* 15, 9: 371-7.

Gehring WJ. (1996)"The Master Control Gene for Morphogenesis and Evolution of the Eye." *Genes to cells : devoted to molecular & cellular mechanisms* 1, 1: 11-5.

Geimonen E, Boylston E, Royek A, and Andersen J. (1998)"Elevated Connexin-43 Expression in Term Human Myometrium Correlates with Elevated c-Jun Expression and is Independent of Myometrial Estrogen Receptors." *Journal of Clinical Endocrinology and Metabolism* 83, 4: 1177-85.

Geimonen E, Jiang W, Ali M, Fishman GI, Garfield RE, and Andersen J. (1996)"Activation of Protein Kinase C in Human Uterine Smooth Muscle Induces Connexin-43 Gene Transcription through an AP-1 Site in the Promoter Sequence." *Journal of Biological Chemistry* 271, 39: 23667-74.

Gerhard GS. (2007)"Small Laboratory Fish as Models for Aging Research." *Ageing*

research reviews 6, 1: 64-72.

Gerido DA, and White TW. (2004)"Connexin Disorders of the Ear, Skin, and Lens." *Biochimica et biophysica acta* 1662, 1-2: 159-70.

Ghysen A, and Dambly-Chaudiere C. (2004)"Development of the Zebrafish Lateral Line." *Current opinion in neurobiology* 14, 1: 67-73.

Giepmans BN.(2004)"Gap Junctions and Connexin-Interacting Proteins." *Cardiovascular research* 62, 2: 233-45.

Glickman NS, and Yelon D. (2002)"Cardiac Development in Zebrafish: Coordination of Form and Function." *Seminars in cell & developmental biology* 13, 6: 507-13.

Goessling W, North TE, and Zon LI. (2007)"New Waves of Discovery: Modeling Cancer in Zebrafish." *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 25, 17: 2473-9.

Golling G, Amsterdam A, Sun Z, Antonelli M, Maldonado E, Chen W, Burgess S, Haldi M, Artzt K, Farrington S, et al. (2002)"Insertional Mutagenesis in Zebrafish Rapidly Identifies Genes Essential for Early Vertebrate Development." *Nature Genetics* 31, 2: 135-40.

Gong X, Cheng C, and Xia CH. (2007)"Connexins in Lens Development and Cataractogenesis." *The Journal of membrane biology* , : .

Gong X, Li E, Klier G, Huang Q, Wu Y, Lei H, Kumar NM, Horwitz J, and Gilula NB. (1997)"Disruption of alpha3 Connexin Gene Leads to Proteolysis and Cataractogenesis in Mice." *Cell* 91, 6: 833-43.

Goodenough DA, Goliger JA, and Paul DL. (1996)"Connexins, Connexons, and Intercellular Communication." *Annual Review of Biochemistry* 65, : 475-502.

Goodenough DA, and Paul DL. (2003)"Beyond the Gap: Functions of Unpaired Connexon Channels." *Nature reviews.Molecular cell biology* 4, 4: 285-94.

Goodenough DA. (1974)"Bulk Isolation of Mouse Hepatocyte Gap Junctions. Characterization of the Principal Protein, Connexin." *Journal of Cell Biology* 61, 2: 557-63.

Goodrich LV. (2005)"Hear, Hear for the Zebrafish." *Neuron* 45, 1: 3-5.

Gourdie RG. (1995)"A Map of the Heart: Gap Junctions, Connexin Diversity and Retroviral Studies of Conduction Myocyte Lineage." *Clinical science (London, England : 1979)* 88, 3: 257-62.

Grabher C, and Wittbrodt J. (2004)"Efficient Activation of Gene Expression using a Heat-Shock Inducible Gal4/Vp16-UAS System in Medaka." *BMC biotechnology* 4, : 26.

Grasso C, Modrek B, Xing Y, and Lee C. (2004)"Genome-Wide Detection of Alternative Splicing in Expressed Sequences using Partial Order Multiple Sequence Alignment Graphs." *Pacific Symposium on Biocomputing.Pacific Symposium on Biocomputing* , : 29-41.

Graveley BR. (2001)"Alternative Splicing: Increasing Diversity in the Proteomic World." *Trends in Genetics* 17, 2: 100-7.

Graveley BR. (2005)"Mutually Exclusive Splicing of the Insect Dscam Pre-mRNA Directed by Competing Intronic RNA Secondary Structures." *Cell* 123, 1: 65-73.

Graveley BR. (2000)"Sorting Out the Complexity of SR Protein Functions." *RNA (New York, N.Y.)* 6, 9: 1197-211.

Grillitsch S, Medgyesy N, Schwerte T, and Pelster B. (2005)"The Influence of Environmental P(O(2)) on Hemoglobin Oxygen Saturation in Developing Zebrafish Danio Rerio." *The Journal of experimental biology* 208, Pt 2: 309-16.

Gros D, Dupays L, Alcoléa S, Meysen S, Miquerol L, and Théveniau-Ruissy M. (2004)"Genetically Modified Mice: Tools to Decode the Functions of Connexins in the Heart - New Models for Cardiovascular Research." *Cardiovascular Research* 62, 2: 299-308.

Gros DB, and Jongsma HJ. (1996)"Connexins in Mammalian Heart Function." *BioEssays : news and reviews in molecular, cellular and developmental biology* 18, 9: 719-30.

Gum JR,Jr, Hicks JW, Crawley SC, Dahl CM, Yang SC, Roberton AM, and Kim YS. (2003)"Initiation of Transcription of the MUC3A Human Intestinal Mucin from a TATA-Less Promoter and Comparison with the MUC3B Amino Terminus." *The Journal of biological chemistry* 278, 49: 49600-9.

Guyon JR, Steffen LS, Howell MH, Pusack TJ, Lawrence C, and Kunkel LM. (2007)"Modeling Human Muscle Disease in Zebrafish." *Biochimica et biophysica acta* 1772, 2: 205-15.

Haffter P, Granato M, Brand M, Mullins MC, Hammerschmidt M, Kane DA, Odenthal J, van Eeden FJ, Jiang YJ, Heisenberg CP, et al. (1996)"The Identification of Genes with Unique and Essential Functions in the Development of the Zebrafish, Danio Rerio." *Development (Cambridge, England)* 123, : 1-36.

Halloran MC, Sato-Maeda M, Warren Jr. JT, Su F, Lele Z, Krone PH, Kuwada JY, and Shoji W. (2000)"Laser-Induced Gene Expression in Specific Cells of Transgenic Zebrafish." *Development* 127, 9: 1953-60.

Hammes A, Guo J-, Lutsch G, Leheste J-, Landrock D, Ziegler U, Gubler M-, and Schedl A. (2001)"Two Splice Variants of the Wilms' Tumor 1 Gene have Distinct Functions during Sex Determination and Nephron Formation." *Cell* 106, 3: 319-29.

Hannon GJ. (2002)"RNA Interference." *Nature* 418, 6894: 244-51.

Hardy ME, Ross LV, and Chien CB. (2007)"Focal Gene Misexpression in Zebrafish Embryos Induced by Local Heat Shock using a Modified Soldering Iron." *Developmental dynamics : an official publication of the American Association of Anatomists* 236, 11: 3071-6.

Harris AL. (2001)"Emerging Issues of Connexin Channels: Biophysics Fills the Gap." *Quarterly reviews of biophysics* 34, 3: 325-472.

Harvey RP. (1999)"Seeking a Regulatory Roadmap for Heart Morphogenesis." *Seminars in cell & developmental biology* 10, 1: 99-107.

Hasson T. (1999)"Molecular Motors: Sensing a Function for Myosin-VIIa." *Current biology : CB* 9, 22: R838-41.

Heasman J, Kofron M, and Wylie C. (2000)"Beta-Catenin Signaling Activity Dissected in the Early Xenopus Embryo: A Novel Antisense Approach." *Developmental biology* 222, 1: 124-34.

Heasman J. (2002)"Morpholino Oligos: Making Sense of Antisense?" *Developmental biology* 243, 2: 209-14.

Hennemann H, Kozjek G, Dahl E, Nicholson B, and Willecke K. (1992a)"Molecular

Cloning of Mouse connexins26 and -32: Similar Genomic Organization but Distinct Promoter Sequences of Two Gap Junction Genes." *European Journal of Cell Biology* 58, 1: 81-9.

Hennemann H, Schwarz H-, and Willecke K. (1992b)"Characterization of Gap Junction Genes Expressed in F9 Embryonic Carcinoma Cells: Molecular Cloning of Mouse connexin31 and -45 cDNAs." *European Journal of Cell Biology* 57, 1: 51-8.

Herve JC, Bourmeyster N, Sarrouilhe D, and Duffy HS. (2007a)"Gap Junctional Complexes: From Partners to Functions." *Progress in biophysics and molecular biology* 94, 1-2: 29-65.

Herve JC, Derangeon M, Bahbouhi B, Mesnil M, and Sarrouilhe D. (2007b)"The Connexin Turnover, an Important Modulating Factor of the Level of Cell-to-Cell Junctional Communication: Comparison with Other Integral Membrane Proteins." *The Journal of membrane biology* 217, 1-3: 21-33.

Herve JC, Phelan P, Bruzzone R, and White TW. (2005)"Connexins, Innexins and Pannexins: Bridging the Communication Gap." *Biochimica et biophysica acta* 1719, 1-2: 3-5.

Heynkes R, Kozjek G, Traub O, and Willecke K. (1986)"Identification of a Rat Liver cDNA and mRNA Coding for the 28 kDa Gap Junction Protein." *FEBS letters* 205, 1: 56-60.

Hieber V, Dai X, Foreman M, and Goldman D. (1998)"Induction of α1-Tubulin Gene Expression during Development and Regeneration of the Fish Central Nervous System." *Journal of Neurobiology* 37, 3: 429-40.

Higashijima S-, Hotta Y, and Okamoto H. (2000)"Visualization of Cranial Motor Neurons in Live Transgenic Zebrafish Expressing Green Fluorescent Protein Under the Control of the Islet-1 promoter/enhancer." *Journal of Neuroscience* 20, 1: 206-18.

Hiller M, Huse K, Szafranski K, Jahn N, Hampe J, Schreiber S, Backofen R, and Platzer M. (2004)"Widespread Occurrence of Alternative Splicing at NAGNAG Acceptors Contributes to Proteome Plasticity." *Nature Genetics* 36, 12: 1255-7.

Hiller M, Nikolajewa S, Huse K, Szafranski K, Rosenstiel P, Schuster S, Backofen R, and Platzer M. (2007)"TassDB: A Database of Alternative Tandem Splice Sites." *Nucleic Acids Research* 35, SUPPL. 1: .

Hinegardner R, and Rosen DE. (1972)"Cellular DNA Content and the Evolution of Teleostean Fishes." *The American Naturalist* 106, 951: 621.

Hirai A, Yano T, Nishikawa K, Suzuki K, Asano R, Satoh H, Hagiwara K, and Yamasaki H. (2003)"Down-Regulation of Connexin 32 Gene Expression through DNA Methylation in a Human Renal Cell Carcinoma Cell." *American Journal of Nephrology* 23, 3: 172-7.

Hoffman MA, and Palmenberg AC. (1995)"Mutational Analysis of the J-K Stem-Loop Region of the Encephalomyocarditis Virus IRES." *Journal of virology* 69, 7: 4399-406.

Hombach S, Janssen-Bienhold U, Söhl G, Schubert T, Büssow H, Ott T, Weiler R, and Willecke K. (2004)"Functional Expression of connexin57 in Horizontal Cells of the Mouse Retina." *European Journal of Neuroscience* 19, 10: 2633-40.

Hsiao C-, Hsieh F-, and Tsai H-. (2001)"Enhanced Expression and Stable Transmission of Transgenes Flanked by Inverted Terminal Repeats from Adeno-Associated Virus in Zebrafish." *Developmental Dynamics* 220, 4: 323-36.

Hu C-A, Lin W-, Obie C, and Valle D. (1999)"Molecular Enzymology of Mammalian Δ1-Pyrroline-5-Carboxylate Synthase. Alternative Splice Donor Utilization Generates Isoforms with Different Sensitivity to Ornithine Inhibition." *Journal of Biological Chemistry* 274, 10: 6754-62.

Hu N, Sedmera D, Yost HJ, and Clark EB. (2000)"Structure and Function of the Developing Zebrafish Heart." *The Anatomical Record* 260, 2: 148-57.

Huang CJ, Jou TS, Ho YL, Lee WH, Jeng YT, Hsieh FJ, and Tsai HJ. (2005)"Conditional Expression of a Myocardium-Specific Transgene in Zebrafish Transgenic Lines." *Developmental dynamics : an official publication of the American Association of Anatomists* 233, 4: 1294-303.

Huang H, Vogel SS, Liu N, Melton DA, and Lin S. (2001)"Analysis of Pancreatic Development in Living Transgenic Zebrafish Embryos." *Molecular and Cellular Endocrinology* 177, 1-2: 117-24.

Hudder A, and Werner R. (2000)"Analysis of a Charcot-Marie-Tooth Disease Mutation Reveals an Essential Internal Ribosome Entry Site Element in the Connexin-32 Gene." *The Journal of biological chemistry* 275, 44: 34586-91.

Huez I, Creancier L, Audigier S, Gensac MC, Prats AC, and Prats H. (1998)"Two Independent Internal Ribosome Entry Sites are Involved in Translation Initiation of Vascular Endothelial Growth Factor mRNA." *Molecular and cellular biology* 18, 11: 6178-90.

Hughes TA, and Brady HJ. (2005)"Expression of axin2 is Regulated by the Alternative 5'-Untranslated Regions of its mRNA." *The Journal of biological chemistry* 280, 9: 8581-8.

Hughes TA. (2006)"Regulation of Gene Expression by Alternative Untranslated Regions." *Trends in genetics : TIG* 22, 3: 119-22.

Iacono M, Mignone F, and Pesole G. (2005)"UAUG and uORFs in Human and Rodent 5'Untranslated mRNAs." *Gene* 349, : 97-105.

Iovine MK, Gumpert AM, Falk MM, and Mendelson TC. (2008)"Cx23, a Connexin with Only Four Extracellular-Loop Cysteines, Forms Functional Gap Junction Channels and Hemichannels." *FEBS Letters* 582, 2: 165-70.

Istrail S, and Davidson EH. (2005)"Logic Functions of the Genomic Cis-Regulatory Code." *Proceedings of the National Academy of Sciences of the United States of America* 102, 14: 4954-9.

Jackson RJ. (2005)"Alternative Mechanisms of Initiating Translation of Mammalian mRNAs." *Biochemical Society transactions* 33, Pt 6: 1231-41.

Jacob A, and Beyer EC. (2001)"Mouse Connexin 45: Genomic Cloning and Exon Usage." *DNA and cell biology* 20, 1: 11-9.

Jansen RP. (2001)"MRNA Localization: Message on the Move." *Nature reviews.Molecular cell biology* 2, 4: 247-56.

Janssen-Timmen U, Dermietzel R, Frixen U, Leibstein A, Traub O, and Willecke K. (1983)"Immunocytochemical Localization of the Gap Junction 26 K Protein in Mouse Liver Plasma Membranes." *The EMBO journal* 2, 3: 295-302.

Jay PY, Harris BS, Maguire CT, Buerger A, Wakimoto H, Tanaka M, Kupershmidt S, Roden DM, Schultheiss TM, O'Brien TX, et al. (2004)"Nkx2-5 Mutation Causes Anatomic Hypoplasia of the Cardiac Conduction System." *Journal of Clinical Investigation* 113, 8: 1130-7.

Jessen JR, Meng A, Mcfarlane RJ, Paw BH, Zon LI, Smith GR, and Lin S. (1998)"Modification of Bacterial Artificial Chromosomes through Chi-Stimulated Homologous Recombination and its Application in Zebrafish Transgenesis." *Proceedings of the National Academy of Sciences of the United States of America* 95, 9: 5121-6.

Jessen JR, Willett CE, and Lin S. (1999)"Artificial Chromosome Transgenesis Reveals Long-Distance Negative Regulation of rag1 in Zebrafish [1]." *Nature Genetics* 23, 1: 15-6.

Jesuthasan S, and Subburaju S. (2002)"Gene Transfer into Zebrafish by Sperm Nuclear Transplantation." *Developmental biology* 242, 2: 88-95.

Jiang JX, and Goodenough DA. (1996)"Heteromeric Connexons in Lens Gap Junction Channels." *Proceedings of the National Academy of Sciences of the United States of America* 93, 3: 1287-91.

Jiang JX, White TW, Goodenough DA, and Paul DL. (1994)"Molecular Cloning and Functional Characterization of Chick Lens Fiber Connexin 45.6." *Molecular Biology of the Cell* 5, 3: 363-73.

John SA, and Revel JP. (1991)"Connexon Integrity is Maintained by Non-Covalent Bonds: Intramolecular Disulfide Bonds Link the Extracellular Domains in Rat Connexin-43." *Biochemical and biophysical research communications* 178, 3: 1312-8.

Jongsma HJ, and Wilders R. (2000)"Gap Junctions in Cardiovascular Disease." *Circulation research* 86, 12: 1193-7.

Joyce-Brady M, Jean J-, and Hughey RP. (2001)"Γ-Glutamyltransferase and its Isoform Mediate an Endoplasmic Reticulum Stress Response." *Journal of Biological Chemistry* 276, 12: 9468-77.

Ju B, Xu Y, He J, Liao J, Yan T, Hew CL, Lam TJ, and Gong Z. (1999)"Faithful Expression of Green Fluorescent Protein (GFP) in Transgenic Zebrafish Embryos Under Control of Zebrafish Gene Promoters." *Developmental Genetics* 25, 2: 158-67.

Kaminski A, Hunt SL, Gibbs CL, and Jackson RJ. (1994)"Internal Initiation of mRNA Translation in Eukaryotes." *Genetic engineering* 16, : 115-55.

Kari G, Rodeck U, and Dicker AP. (2007)"Zebrafish: An Emerging Model System for Human Disease and Drug Discovery." *Clinical pharmacology and therapeutics* 82, 1: 70-80.

Kasahara H, Ueyama T, Wakimoto H, Liu MK, Maguire CT, Converso KL, Kang PM, Manning WJ, Lawitts J, Paul DL, et al. (2003)"Nkx2.5 Homeoprotein Regulates Expression of Gap Junction Protein Connexin 43 and Sarcomere Organization in Postnatal Cardiomyocytes." *Journal of Molecular and Cellular Cardiology* 35, 3: 243-56.

Kasahara H, Wakimoto H, Liu M, Maguire CT, Converso KL, Shioi T, Huang W-, Manning WJ, Paul D, Lawitts J, et al. (2001)"Progressive Atrioventricular Conduction Defects and Heart Failure in Mice Expressing a Mutant Csx/Nkx2.5 Homeoprotein." *Journal of Clinical Investigation* 108, 2: 189-201.

Kawakami K, Amsterdam A, Shimoda N, Becker T, Mugg J, Shima A, and Hopkins N. (2000)"Proviral Insertions in the Zebrafish Hagoromo Gene, Encoding an F-box/WD40-Repeat Protein, Cause Stripe Pattern Anomalies." *Current Biology* 10, 8: 463-6.

Kawakami K, Koga A, Hori H, and Shima A. (1998)"Excision of the tol2 Transposable Element of the Medaka Fish, Oryzias Latipes, in Zebrafish, Danio Rerio." *Gene* 225, 1-2: 17-22.

Kawakami K, Shima A, and Kawakami N. (2000)"Identification of a Functional Transposase of the Tol2 Element, an Ac-Like Element from the Japanese Medaka Fish, and its Transposition in the Zebrafish Germ Lineage." *Proceedings of the National Academy of Sciences of the United States of America* 97, 21: 11403-8.

Kawakami K, and Shima A. (1999)"Identification of the Tol2 Transposase of the Medaka Fish Oryzias Latipes that Catalyzes Excision of a Nonautonomous Tol2 Element in Zebrafish Danio Rerio." *Gene* 240, 1: 239-44.

Kawakami K, Takeda H, Kawakami N, Kobayashi M, Matsuda N, and Mishina M. (2004)"A Transposon-Mediated Gene Trap Approach Identifies Developmentally Regulated Genes in Zebrafish." *Developmental cell* 7, 1: 133-44.

Kawakami K. (2005)"Transposon Tools and Methods in Zebrafish." *Developmental dynamics : an official publication of the American Association of Anatomists* 234, 2: 244-54.

Keating MT. (2004)"Genetic Approaches to Disease and Regeneration." *Philosophical transactions of the Royal Society of London.Series B, Biological sciences* 359, 1445: 795-8.

Keller ET, and Murtha JM. (2004)"The use of Mature Zebrafish (Danio Rerio) as a Model for Human Aging and Disease." *Comparative biochemistry and physiology.Toxicology & pharmacology : CBP* 138, 3: 335-41.

Kelsell DP, Dunlop J, and Hodgins MB. (2001)"Human Diseases: Clues to Cracking the Connexin Code?" *Trends in cell biology* 11, 1: 2-6.

Kikuchi T, Kimura RS, Paul DL, and Adams JC. (1995)"Gap Junctions in the Rat Cochlea: Immunohistochemical and Ultrastructural Analysis." *Anatomy and Embryology* 191, 2: 101-18.

Kikuchi T, Kimura RS, Paul DL, Takasaka T, and Adams JC. (2000)"Gap Junction Systems in the Mammalian Cochlea." *Brain research.Brain research reviews* 32, 1: 163-6.

Kikuta H, Fredman D, Rinkwitz S, Lenhard B, and Becker TS. (2007)"Retroviral Enhancer Detection Insertions in Zebrafish Combined with Comparative Genomics Reveal Genomic Regulatory Blocks - a Fundamental Feature of Vertebrate Genomes." *Genome biology* 8 Suppl 1, : S4.

Kim C-, Taira E, Kuo C-, Li B-, Okamoto H, Nakahira K, Ikenaka K, Higuchi H, and Miki N. (1996)"Neuron-Specific Expression of a Chicken Gicerin cDNA in Transient Transgenic Zebrafish." *Neurochemical Research* 21, 2: 231-7.

Kindler S, Wang H, Richter D, and Tiedge H. (2005)"RNA Transport and Local Control of Translation." *Annual Review of Cell and Developmental Biology* 21, : 223-45.

Kirchhoff S, Kim JS, Hagendorff A, Thonnissen E, Kruger O, Lamers WH, and Willecke K. (2000)"Abnormal Cardiac Conduction and Morphogenesis in connexin40 and connexin43 Double-Deficient Mice." *Circulation research* 87, 5: 399-405.

Kirchhoff S, Nelles E, Hagendorff A, Kruger O, Traub O, and Willecke K. (1998)"Reduced Cardiac Conduction Velocity and Predisposition to Arrhythmias in connexin40-Deficient Mice." *Current biology : CB* 8, 5: 299-302.

Kistler J, Lin JS, Bond J, Green C, Eckert R, Merriman R, Tunstall M, and Donaldson P.

(1999)"Connexins in the Lens: Are they to Blame in Diabetic Cataractogenesis?" *Novartis Foundation symposium* 219, : 97,108; discussion 108-12.

Klingenhoff A, Frech K, Quandt K, and Werner T. (1999)"Functional Promoter Modules can be Detected by Formal Models Independent of overall Nucleotide Sequence Similarity." *Bioinformatics* 15, 3: 180-6.

Kloosterman WP, Lagendijk AK, Ketting RF, Moulton JD, and Plasterk RH. (2007)"Targeted Inhibition of miRNA Maturation with Morpholinos Reveals a Role for miR-375 in Pancreatic Islet Development." *PLoS biology* 5, 8: e203.

Koenig Merediz SA, Schmidt M, Hoppe GJ, Alfken J, Meraro D, Levi B-, Neubauer A, and Wittig B. (2000)"Cloning of an Interferon Regulatory Factor 2 Isoform with Different Regulatory Ability." *Nucleic Acids Research* 28, 21: 4219-24.

Koffler LD, Fernstrom MJ, Akiyama TE, Gonzalez FJ, and Ruch RJ. (2002)"Positive Regulation of connexin32 Transcription by Hepatocyte Nuclear Factor-1α." *Archives of Biochemistry and Biophysics* 407, 2: 160-7.

Kornblihtt AR. (2006)"Chromatin, Transcript Elongation and Alternative Splicing." *Nature structural & molecular biology* 13, 1: 5-7.

Kornblihtt AR. (2005)"Promoter Usage and Alternative Splicing." *Current opinion in cell biology* 17, 3: 262-8.

Korzh V. (2007)"Transposons as Tools for Enhancer Trap Screens in Vertebrates." *Genome biology* 8 Suppl 1, : S8.

Koster RW, and Fraser SE. (2001)"Tracing Transgene Expression in Living Zebrafish Embryos." *Developmental biology* 233, 2: 329-46.

Kozak M. (1991)"An Analysis of Vertebrate mRNA Sequences: Intimations of Translational Control." *The Journal of cell biology* 115, 4: 887-903.

Kozmik Z. (2008)"The Role of Pax Genes in Eye Evolution." *Brain research bulletin* 75, 2-4: 335-9.

Krecic AM, and Swanson MS. (1999)"HnRNP Complexes: Composition, Structure, and Function." *Current opinion in cell biology* 11, 3: 363-71.

Kumar NM, and Gilula NB. (1986)"Cloning and Characterization of Human and Rat

Liver cDNAs Coding for a Gap Junction Protein." *The Journal of cell biology* 103, 3: 767-76.

Kumar NM, and Gilula NB. (1996)"The Gap Junction Communication Channel." *Cell* 84, 3: 381-8.

Kupperman E, An S, Osborne N, Waldron S, and Stainier DY. (2000)"A Sphingosine-1-Phosphate Receptor Regulates Cell Migration during Vertebrate Heart Development." *Nature* 406, 6792: 192-5.

Kurita R, Sagara H, Aoki Y, Link BA, Arai K, and Watanabe S. (2003)"Suppression of Lens Growth by alphaA-Crystallin Promoter-Driven Expression of Diphtheria Toxin Results in Disruption of Retinal Cell Organization in Zebrafish." *Developmental biology* 255, 1: 113-27.

Lacerra G, Sierakowska H, Carestia C, Fucharoen S, Summerton J, Weller D, and Kole R. (2000)"Restoration of Hemoglobin A Synthesis in Erythroid Cells from Peripheral Blood of Thalassemic Patients." *Proceedings of the National Academy of Sciences of the United States of America* 97, 17: 9591-6.

Laird DW. (2006)"Life Cycle of Connexins in Health and Disease." *The Biochemical journal* 394, Pt 3: 527-43.

Lam SH, and Gong Z. (2006)"Modeling Liver Cancer using Zebrafish: A Comparative Oncogenomics Approach." *Cell cycle (Georgetown, Tex.)* 5, 6: 573-7.

Langenau DM, Feng H, Berghmans S, Kanki JP, Kutok JL, and Look AT. (2005)"Cre/lox-Regulated Transgenic Zebrafish Model with Conditional Myc-Induced T Cell Acute Lymphoblastic Leukemia." *Proceedings of the National Academy of Sciences of the United States of America* 102, 17: 6068-73.

Lareau LF, Green RE, Bhatnagar RS, and Brenner SE. (2004)"The Evolving Roles of Alternative Splicing." *Current opinion in structural biology* 14, 3: 273-82.

Lassmann T, and Sonnhammer EL. (2005)"Kalign--an Accurate and Fast Multiple Sequence Alignment Algorithm." *BMC bioinformatics* 6, : 298.

Lawson ND, Vogel AM, and Weinstein BM. (2002)"Sonic Hedgehog and Vascular Endothelial Growth Factor Act Upstream of the Notch Pathway during Arterial Endothelial Differentiation." *Developmental cell* 3, 1: 127-36.

Le SY, and Maizel JV,Jr. (1997)"A Common RNA Structural Motif Involved in the Internal Initiation of Translation of Cellular mRNAs." *Nucleic acids research* 25, 2: 362-9.

Lefebvre PP, and Van De Water TR. (2000)"Connexins, Hearing and Deafness: Clinical Aspects of Mutations in the Connexin 26 Gene." *Brain research.Brain research reviews* 32, 1: 159-62.

Lev-Maor G, Sorek R, Shomron N, and Ast G. (2003)"The Birth of an Alternatively Spliced Exon: 3′ Splice-Site Selection in Alu Exons." *Science* 300, 5623: 1288-91.

Li L, and Howe GA. (2001)"Alternative Splicing of Prosystemin Pre-mRNA Produces Two Isoforms that are Active as Signals in the Wound Response Pathway." *Plant Molecular Biology* 46, 4: 409-19.

Li Y-, Farrell MJ, Liu R, Mohanty N, and Kirby ML. (2000)"Double-Stranded RNA Injection Produces Null Phenotypes in Zebrafish." *Developmental Biology* 217, 2: 394-405.

Liang MR, Alestrom P, and Collas P. (2000)"Glowing Zebrafish: Integration, Transmission, and Expression of a Single Luciferase Transgene Promoted by Noncovalent DNA-Nuclear Transport Peptide Complexes." *Molecular reproduction and development* 55, 1: 8-13.

Lieschke GJ, and Currie PD. (2007)"Animal Models of Human Disease: Zebrafish Swim into View." *Nature reviews.Genetics* 8, 5: 353-67.

Lin JH, Takano T, Cotrina ML, Arcuino G, Kang J, Liu S, Gao Q, Jiang L, Li F, Lichtenberg-Frate H, et al. (2002)"Connexin 43 Enhances the Adhesivity and Mediates the Invasion of Malignant Glioma Cells." *The Journal of neuroscience : the official journal of the Society for Neuroscience* 22, 11: 4302-11.

Lin S, Gaiano N, Culp P, Burns JC, Friedmann T, Yee JK, and Hopkins N. (1994)"Integration and Germ-Line Transmission of a Pseudotyped Retroviral Vector in Zebrafish." *Science (New York, N.Y.)* 265, 5172: 666-9.

Linhares VLF, Almeida NAS, Menezes DC, Elliott DA, Lai D, Beyer EC, Campos De Carvalho AC, and Costa MW. (2004)"Transcriptional Regulation of the Murine Connexin40 Promoter by Cardiac Factors Nkx2-5, GATA4 and Tbx5." *Cardiovascular Research* 64, 3: 402-11.

Linney E, Hardison NL, Lonze BE, Lyons S, and DiNapoli L. (1999)"Transgene Expression in Zebrafish: A Comparison of Retroviral-Vector and DNA-Injection Approaches." *Developmental biology* 213, 1: 207-16.

Liu X, and Gorovsky MA. (1993)"Mapping the 5' and 3' Ends of Tetrahymena Thermophila mRNAs using RNA Ligase Mediated Amplification of cDNA Ends (RLM-RACE)." *Nucleic acids research* 21, 21: 4954-60.

Lo CW. (1996)"The Role of Gap Junction Membrane Channels in Development." *Journal of Bioenergetics and Biomembranes* 28, 4: 379-85.

Lohr JL, and Yost HJ. (2000)"Vertebrate Model Systems in the Study of Early Heart Development: Xenopus and Zebrafish." *American Journal of Medical Genetics* 97, 4: 248-57.

Loncarek J, Yamasaki H, Levillain P, Milinkevitch S, and Mesnil M. (2003)"The Expression of the Tumor Suppressor Gene Connexin 26 is Not Mediated by Methylation in Human Esophageal Cancer Cells." *Molecular Carcinogenesis* 36, 2: 74-81.

Long Q, Meng A, Wang M, Jessen JR, Farrell MJ, and Lin S. (1997)"GATA-1 Expression Pattern can be Recapitulated in Living Transgenic Zebrafish using GFP Reporter Gene." *Development* 124, 20: 4105-11.

Makielski JC, Ye B, Valdivia CR, Pagel MD, Pu J, Tester DJ, and Ackerman MJ. (2003)"A Ubiquitous Splice Variant and a Common Polymorphism Affect Heterologous Expression of Recombinant Human SCN5A Heart Sodium Channels." *Circulation Research* 93, 9: 821-8.

Maniatis T, and Reed R. (2002)"An Extensive Network of Coupling among Gene Expression Machines." *Nature* 416, 6880: 499-506.

Maniatis T, and Tasic B. (2002)"Alternative Pre-mRNA Splicing and Proteome Expansion in Metazoans." *Nature* 418, 6894: 236-43.

Marsh-Armstrong N, McCaffery P, Hyatt G, Alonso L, Dowling JE, Gilbert W, and Drager UC. (1995)"Retinoic Acid in the Anteroposterior Patterning of the Zebrafish Trunk." *Roux's Archives of Developmental Biology* 205, 3-4: 103-13.

Martin PE, Coleman SL, Casalotti SO, Forge A, and Evans WH. (1999)"Properties of connexin26 Gap Junctional Proteins Derived from Mutations Associated with

Non-Syndromal Heriditary Deafness." *Human molecular genetics* 8, 13: 2369-76.

Martinez-Wittinghan FJ, Sellitto C, White TW, Mathias RT, Paul D, and Goodenough DA. (2004)"Lens Gap Junctional Coupling is Modulated by Connexin Identity and the Locus of Gene Expression." *Investigative ophthalmology & visual science* 45, 10: 3629-37.

Martins J, Oliva Teles L, and Vasconcelos V. (2007)"Assays with Daphnia Magna and Danio Rerio as Alert Systems in Aquatic Toxicology." *Environment international* 33, 3: 414-25.

Mathias RT, Kistler J, and Donaldson P. (2007)"The Lens Circulation." *The Journal of membrane biology* 216, 1: 1-16.

Matlin AJ, Clark F, and Smith CWJ. (2005)"Understanding Alternative Splicing: Towards a Cellular Code." *Nature Reviews Molecular Cell Biology* 6, 5: 386-98.

Matlin AJ, and Moore MJ. (2007)"Spliceosome Assembly and Composition." *Advances in Experimental Medicine and Biology* 623, : 14-35.

Mattick JS. (2004)"RNA Regulation: A New Genetics?" *Nature reviews.Genetics* 5, 4: 316-23.

McLachlan E, White TW, Ugonabo C, Olson C, Nagy JI, and Valdimarsson G. (2003)"Zebrafish Cx35: Cloning and Characterization of a Gap Junction Gene Highly Expressed in the Retina." *Journal of neuroscience research* 73, 6: 753-64.

Meijer HA, Dictus WJ, Keuning ED, and Thomas AA. (2000)"Translational Control of the Xenopus Laevis Connexin-41 5'-Untranslated Region by Three Upstream Open Reading Frames." *The Journal of biological chemistry* 275, 40: 30787-93.

Meijer HA, and Thomas AA. (2003)"Ribosomes Stalling on uORF1 in the Xenopus Cx41 5' UTR Inhibit Downstream Translation Initiation." *Nucleic acids research* 31, 12: 3174-84.

Mendes Soares LM, and Valcarcel J. (2006)"The Expanding Transcriptome: The Genome as the 'Book of Sand'." *The EMBO journal* 25, 5: 923-31.

Meng A, Tang H, Ong BA, Farrell MJ, and Lin S. (1997)"Promoter Analysis in Living Zebrafish Embryos Identifies a Cis-Acting Motif Required for Neuronal Expression of GATA-2." *Proceedings of the National Academy of Sciences of the United States*

of America 94, 12: 6267-72.

Meng A, Tang H, Yuan B, Ong BA, Long Q, and Lin S. (1999)"Positive and Negative Cis-Acting Elements are Required for Hematopoietic Expression of Zebrafish GATA-1." *Blood* 93, 2: 500-8.

Mercatante DR, Sazani P, and Kole R. (2001)"Modification of Alternative Splicing by Antisense Oligonucleotides as a Potential Chemotherapy for Cancer and Other Diseases." *Current cancer drug targets* 1, 3: 211-30.

Mese G, Richard G, and White TW. (2007)"Gap Junctions: Basic Structure and Function." *The Journal of investigative dermatology* 127, 11: 2516-24.

Meshorer E, Toiber D, Zurel D, Sahly I, Dori A, Cagnano E, Schreiber L, Grisaru D, Tronche F, and Soreq H. (2004)"Combinatorial Complexity of 5' Alternative Acetylcholinesterase Transcripts and Protein Products." *The Journal of biological chemistry* 279, 28: 29740-51.

Mignone F, Gissi C, Liuni S, and Pesole G. (2002)"Untranslated Regions of mRNAs." *Genome biology* 3, 3: REVIEWS0004.

Mitchell JA, Ou C-, Chen Z-, Nishimura T, and Lye SJ. (2001)"Parathyroid Hormone-Induced Up-Regulation of Connexin-43 Messenger Ribonucleic Acid (mRNA) is Mediated by Sequences within both the Promoter and the 3′ Untranslated Region of the mRNA." *Endocrinology* 142, 2: 907-15.

Moav B, Liu Z, Caldovic LD, Gross ML, Faras AJ, and Hackett PB. (1993)"Regulation of Expression of Transgenes in Developing Fish." *Transgenic Research* 2, 3: 153-61.

Modrek B, Resch A, Grasso C, and Lee C. (2001)"Genome-Wide Detection of Alternative Splicing in Expressed Sequences of Human Genes." *Nucleic acids research* 29, 13: 2850-9.

Morris DR, and Geballe AP. (2000)"Upstream Open Reading Frames as Regulators of mRNA Translation." *Molecular and cellular biology* 20, 23: 8635-42.

Moss JB, Price AL, Raz E, Driever W, and Rosenthal N. (1996)"Green Fluorescent Protein Marks Skeletal Muscle in Murine Cell Lines and Zebrafish." *Gene* 173, 1 Spec No: 89-98.

Motoike T, Loughna S, Perens E, Roman BL, Liao W, Chau TC, Richardson CD, Kawate

T, Kuno J, Weinstein BM, et al. (2000)"Universal GFP Reporter for the Study of Vascular Development." *Genesis* 28, 2: 75-81.

Muller F, Blader P, and Strahle U. (2002)"Search for Enhancers: Teleost Models in Comparative Genomic and Transgenic Analysis of Cis Regulatory Elements." *BioEssays : news and reviews in molecular, cellular and developmental biology* 24, 6: 564-72.

Muller F, Chang B, Albert S, Fischer N, Tora L, and Strahle U. (1999)"Intronic Enhancers Control Expression of Zebrafish Sonic Hedgehog in Floor Plate and Notochord." *Development (Cambridge, England)* 126, 10: 2103-16.

Murchison EP, and Hannon GJ. (2004)"MiRNAs on the Move: MiRNA Biogenesis and the RNAi Machinery." *Current opinion in cell biology* 16, 3: 223-9.

Musil LS, Beyer EC, and Goodenough DA. (1990)"Expression of the Gap Junction Protein connexin43 in Embryonic Chick Lens: Molecular Cloning, Ultrastructural Localization, and Post-Translational Phosphorylation." *The Journal of membrane biology* 116, 2: 163-75.

Nag A, Narsinh K, Kazerouninia A, and Martinson HG. (2006)"The Conserved AAUAAA Hexamer of the Poly(A) Signal can Act Alone to Trigger a Stable Decrease in RNA Polymerase II Transcription Velocity." *RNA (New York, N.Y.)* 12, 8: 1534-44.

Nasevicius A, and Ekker SC. (2000)"Effective Targeted Gene 'Knockdown' in Zebrafish." *Nature genetics* 26, 2: 216-20.

Neuhaus IM, Bone L, Wang S, Ionasescu V, and Werner R. (1996a)"The Human connexin32 Gene is Transcribed from Two Tissue-Specific Promoters." *Bioscience reports* 16, 3: 239-48.

Neuhaus IM, Dahl G, and Werner R. (1995)"Use of Alternate Promoters for Tissue-Specific Expression of the Gene Coding for connexin32." *Gene* 158, 2: 257-62.

Newton DC, Bevan SC, Choi S, Robb GB, Millar A, Wang Y, and Marsden PA. (2003)"Translational Regulation of Human Neuronal Nitric-Oxide Synthase by an Alternatively Spliced 5'-Untranslated Region Leader Exon." *The Journal of biological chemistry* 278, 1: 636-44.

Nüsslein-Volhard C, Gilmour DT, Dahm R. (2002) Zebrafish as a system to study development and organogenesis: in Nüsslein-Volhard C, Dahm R (eds): *Zebrafish – A Practical Approach*. Oxford, Oxford University Press.

Oates AC, Bruce AEE, and Ho RK. (2000)"Too Much Interference: Injection of Double-Stranded RNA has Nonspecific Effects in the Zebrafish Embryo." *Developmental Biology* 224, 1: 20-8.

O'Brien J, Bruzzone R, White TW, Al-Ubaidi MR, and Ripps H. (1998)"Cloning and Expression of Two Related Connexins from the Perch Retina Define a Distinct Subgroup of the Connexin Family." *The Journal of neuroscience : the official journal of the Society for Neuroscience* 18, 19: 7625-37.

Ogino H, and Yasuda K. (2000)"Sequential Activation of Transcription Factors in Lens Induction." *Development, growth & differentiation* 42, 5: 437-48.

Oltra E, Pfeifer I, and Werner R. (2003)"Ini, a Small Nuclear Protein that Enhances the Response of the connexin43 Gene to Estrogen." *Endocrinology* 144, 7: 3148-58.

Opitz JM, and Clark EB. (2000)"Heart Development: An Introduction." *American Journal of Medical Genetics* 97, 4: 238-47.

Oyamada M, Oyamada Y, and Takamatsu T. (2005a)"Regulation of Connexin Expression." *Biochimica et biophysica acta* 1719, 1-2: 6-23.

Panchin YV. (2005)"Evolution of Gap Junction Proteins--the Pannexin Alternative." *The Journal of experimental biology* 208, Pt 8: 1415-9.

Park H-, Kim C-, Bae Y-, Yeo S-, Kim S-, Hong S-, Shin J, Yoo K-, Hibi M, Hirano T, et al. (2000)"Analysis of Upstream Elements in the HuC Promoter Leads to the Establishment of Transgenic Zebrafish with Fluorescent Neurons." *Developmental Biology* 227, 2: 279-93.

Parng C. (2005)"In Vivo Zebrafish Assays for Toxicity Testing." *Current opinion in drug discovery & development* 8, 1: 100-6.

Paul DL, Ebihara L, Takemoto LJ, Swenson KI, and Goodenough DA. (1991)"Connexin46, a Novel Lens Gap Junction Protein, Induces Voltage-Gated Currents in Nonjunctional Plasma Membrane of Xenopus Oocytes." *The Journal of cell biology* 115, 4: 1077-89.

Paul DL. (1986)"Molecular Cloning of cDNA for Rat Liver Gap Junction Protein." *The Journal of cell biology* 103, 1: 123-34.

Petrocelli T, and Lye SJ. (1993)"Regulation of Transcripts Encoding the Myometrial Gap Junction Protein, Connexin-43, by Estrogen and Progesterone." *Endocrinology* 133, 1: 284-90.

Pfaffl MW. (2001)"A New Mathematical Model for Relative Quantification in Real-Time RT-PCR." *Nucleic acids research* 29, 9: e45.

Pfeifer I, Anderson C, Werner R, and Oltra E. (2004)"Redefining the Structure of the Mouse connexin43 Gene: Selective Promoter Usage and Alternative Splicing Mechanisms Yield Transcripts with Different Translational Efficiencies." *Nucleic acids research* 32, 15: 4550-62.

Phelan P, Bacon JP, Davies JA, Stebbings LA, and Todman MG. (1998)"Innexins: A Family of Invertebrate Gap-Junction Proteins." *Trends in Genetics* 14, 9: 348-9.

Phelan P, and Starich TA. (2001)"Innexins Get into the Gap." *BioEssays : news and reviews in molecular, cellular and developmental biology* 23, 5: 388-96.

Phelan P. (2005)"Innexins: Members of an Evolutionarily Conserved Family of Gap-Junction Proteins." *Biochimica et biophysica acta* 1711, 2: 225-45.

Picker A, Scholpp S, Bohli H, Takeda H, and Brand M. (2002)"A Novel Positive Transcriptional Feedback Loop in Midbrain-Hindbrain Boundary Development is Revealed through Analysis of the Zebrafish pax2.1 Promoter in Transgenic Lines." *Development* 129, 13: 3227-39.

Pickering BM, and Willis AE. (2005)"The Implications of Structured 5' Untranslated Regions on Translation and Disease." *Seminars in cell & developmental biology* 16, 1: 39-47.

Piechocki MP, Burk RD, and Ruch RJ. (1999)"Regulation of connexin32 and connexin43 Gene Expression by DNA Methylation in Rat Liver Cells." *Carcinogenesis* 20, 3: 401-6.

Piechocki MP, Toti RM, Fernstrom MJ, Burk RD, and Ruch RJ. (2000)"Liver Cell-Specific Transcriptional Regulation of connexin32." *Biochimica et Biophysica Acta - Gene Structure and Expression* 1491, 1-3: 107-22.

Piersanti M, and Lye SJ. (1995)"Increase in Messenger Ribonucleic Acid Encoding the

Myometrial Gap Junction Protein, Connexin-43, Requires Protein Synthesis and is Associated with Increased Expression of the Activator Protein-1, c-Fos." *Endocrinology* 136, 8: 3571-8.

Pollard AJ, Krainer AR, Robson SC, and Nicholas Europe-Finner G. (2002)"Alternative Splicing of the Adenylyl Cyclase Stimulatory G-Protein Gαs is Regulated by SF2/ASF and Heterogeneous Nuclear Ribonucleoprotein A1 (hnRNPA1) and Involves the use of an Unusual TG 3′ -Splice Site." *Journal of Biological Chemistry* 277, 18: 15241-51.

Poss KD. (2007)"Getting to the Heart of Regeneration in Zebrafish." *Seminars in cell & developmental biology* 18, 1: 36-45.

Postlethwait J, Amores A, Force A, and Yan YL. (1999)"The Zebrafish Genome." *Methods in cell biology* 60, : 149-63.

Postlethwait JH, Woods IG, Ngo-Hazelett P, Yan YL, Kelly PD, Chu F, Huang H, Hill-Force A, and Talbot WS. (2000)"Zebrafish Comparative Genomics and the Origins of Vertebrate Chromosomes." *Genome research* 10, 12: 1890-902.

Postlethwait JH, Yan YL, Gates MA, Horne S, Amores A, Brownlie A, Donovan A, Egan ES, Force A, Gong Z, et al. (1998)"Vertebrate Genome Evolution and the Zebrafish Gene Map." *Nature genetics* 18, 4: 345-9.

Pozzoli U, and Sironi M. (2005)"Silencers Regulate both Constitutive and Alternative Splicing Events in Mammals." *Cellular and molecular life sciences : CMLS* 62, 14: 1579-604.

Qin G, Taylor M, Ning YY, Iversen P, and Kobzik L. (2000)"In Vivo Evaluation of a Morpholino Antisense Oligomer Directed Against Tumor Necrosis Factor-Alpha." *Antisense & Nucleic Acid Drug Development* 10, 1: 11-6.

Quandt K, Frech K, Karas H, Wingender E, and Werner T. (1995)"MatInd and MatInspector: New Fast and Versatile Tools for Detection of Consensus Matches in Nucleotide Sequence Data." *Nucleic acids research* 23, 23: 4878-84.

Rabionet R, Lopez-Bigas N, Arbones ML, and Estivill X. (2002)"Connexin Mutations in Hearing Loss, Dermatological and Neurological Disorders." *Trends in molecular medicine* 8, 5: 205-12.

Rae JL, Bartling C, Rae J, and Mathias RT. (1996)"Dye Transfer between Cells of the

Lens." *The Journal of membrane biology* 150, 1: 89-103.

Rahman S, and Evans WH. (1991)"Topography of connexin32 in Rat Liver Gap Junctions. Evidence for an Intramolecular Disulphide Linkage Connecting the Two Extracellular Peptide Loops." *Journal of cell science* 100 (Pt 3), Pt 3: 567-78.

Ramos DM, Kamal F, Wimmer EA, Cartwright AN, and Monteiro A. (2006)"Temporal and Spatial Control of Transgene Expression using Laser Induction of the hsp70 Promoter." *BMC developmental biology* 6, : 55.

Raz E, Van Luenen HGAM, Schaerringer B, Plasterkt RHA, and Driever W. (1998)"Transposition of the Nematode Caenorhabditis Elegans Tc3 Element in the Zebrafish Danio Rerio." *Current Biology* 8, 2: 82-8.

Reed KE, Westphale EM, Larson DM, Wang HZ, Veenstra RD, and Beyer EC. (1993)"Molecular Cloning and Functional Expression of Human connexin37, an Endothelial Cell Gap Junction Protein." *The Journal of clinical investigation* 91, 3: 997-1004.

Reinhard E, Nedivi E, Wegner J, Skene JHP, and Westerfield M. (1994)"Neural Selective Activation and Temporal Regulation of a Mammalian GAP-43 Promoter in Zebrafish." *Development* 120, 7: 1767-75.

Renshaw SA, Loynes CA, Elworthy S, Ingham PW, and Whyte MK. (2007)"Modeling Inflammation in the Zebrafish: How a Fish can Help Us Understand Lung Disease." *Experimental lung research* 33, 10: 549-54.

Revel JP, and Karnovsky MJ. (1967)"Hexagonal Array of Subunits in Intercellular Junctions of the Mouse Heart and Liver." *Journal of Cell Biology* 33, 3: .

Richard G. (2005)"Connexin Disorders of the Skin." *Clinics in dermatology* 23, 1: 23-32.

Richard G. (2000)"Connexins: A Connection with the Skin." *Experimental dermatology* 9, 2: 77-96.

Rivers C, Levy A, Hancock J, Lightman S, and Norman M. (1999)"Insertion of an Amino Acid in the DNA-Binding Domain of the Glucocorticoid Receptor as a Result of Alternative Splicing." *Journal of Clinical Endocrinology and Metabolism* 84, 11: 4283-6.

Robertson JD. (1963)"The Occurrence of a Subunit Pattern in the Unit Membrane of

Club Endings in Mauthner Cell Synapses in Goldfish Brain." *J. Cell Biol.* 19, : 201-21.

Robertson ME, Seamons RA, and Belsham GJ. (1999)"A Selection System for Functional Internal Ribosome Entry Site (IRES) Elements: Analysis of the Requirement for a Conserved GNRA Tetraloop in the Encephalomyocarditis Virus IRES." *RNA (New York, N.Y.)* 5, 9: 1167-79.

Rogers M, Berestecky JM, Hossain MZ, Guo H, Kadle R, Nicholson BJ, and Bertram JS. (1990)"Retinoid-Enhanced Gap Junctional Communication is Achieved by Increased Levels of Connexin 43 mRNA and Protein." *Molecular Carcinogenesis* 3, 6: 335-43.

Rubin GM, and Spradling AC. (1983)"Vectors for P Element-Mediated Gene Transfer in Drosophila." *Nucleic acids research* 11, 18: 6341-51.

Rubinstein AL. (2003)"Zebrafish: From Disease Modeling to Drug Discovery." *Current opinion in drug discovery & development* 6, 2: 218-23.

Ruiz-Echevarria MJ, and Peltz SW. (2000)"The RNA Binding Protein Pub1 Modulates the Stability of Transcripts Containing Upstream Open Reading Frames." *Cell* 101, 7: 741-51.

Rukstalis JM, Kowalik A, Zhu L, Lidington D, Pin CL, and Konieczny SF. (2003)"Exocrine Specific Expression of connexin32 is Dependent on the Basic Helix-Loop-Helix Transcription Factor Mist1." *Journal of Cell Science* 116, 16: 3315-25.

Rup DM, Veenstra RD, Wang HZ, Brink PR, and Beyer EC. (1993)"Chick Connexin-56, a Novel Lens Gap Junction Protein. Molecular Cloning and Functional Expression." *The Journal of biological chemistry* 268, 1: 706-12.

Saez JC, Berthoud VM, Branes MC, Martinez AD, and Beyer EC. (2003)"Plasma Membrane Channels Formed by Connexins: Their Regulation and Functions." *Physiological Reviews* 83, 4: 1359-400.

Salameh A. (2006)"Life Cycle of Connexins: Regulation of Connexin Synthesis and Degradation." *Advances in Cardiology* 42, : 57-70.

Schaefer BC. (1995)"Revolutions in Rapid Amplification of cDNA Ends: New Strategies for Polymerase Chain Reaction Cloning of Full-Length cDNA Ends." *Analytical Biochemistry* 227, 2: 255-73.

Scheer N, and Campos-Ortega JA. (1999)"Use of the Gal4-UAS Technique for Targeted Gene Expression in the Zebrafish." *Mechanisms of development* 80, 2: 153-8.

Scheer N, Riedl I, Warren JT, Kuwada JY, and Campos-Ortega JA. (2002)"A Quantitative Analysis of the Kinetics of Gal4 Activator and Effector Gene Expression in the Zebrafish." *Mechanisms of development* 112, 1-2: 9-14.

Schiavi A, Hudder A, and Werner R. (1999)"Connexin43 mRNA Contains a Functional Internal Ribosome Entry Site." *FEBS letters* 464, 3: 118-22.

Scott EK, Mason L, Arrenberg AB, Ziv L, Gosse NJ, Xiao T, Chi NC, Asakawa K, Kawakami K, and Baier H. (2007)"Targeting Neural Circuitry in Zebrafish using GAL4 Enhancer Trapping." *Nature methods* 4, 4: 323-6.

Segretain D, and Falk MM. (2004)"Regulation of Connexin Biosynthesis, Assembly, Gap Junction Formation, and Removal." *Biochimica et biophysica acta* 1662, 1-2: 3-21.

Sellitto C, Li L, and White TW. (2004)"Connexin50 is Essential for Normal Postnatal Lens Cell Proliferation." *Investigative ophthalmology & visual science* 45, 9: 3196-202.

Shafizadeh E, Huang H, and Lin S. (2002)"Transgenic Zebrafish Expressing Green Fluorescent Protein." *Methods in molecular biology (Clifton, N.J.)* 183, : 225-33.

Shalev A, Blair PJ, Hoffmann SC, Hirshberg B, Peculis BA, and Harlan DM. (2002)"A Proinsulin Gene Splice Variant with Increased Translation Efficiency is Expressed in Human Pancreatic Islets." *Endocrinology* 143, 7: 2541-7.

Sharp PA. (2005)"The Discovery of Split Genes and RNA Splicing." *Trends in biochemical sciences* 30, 6: 279-81.

Sharp PA. (1985)"Splicing of Messenger RNA Precursors." *Harvey lectures* 81, : 1-31.

Sharp PA. (1994)"Split Genes and RNA Splicing." *Cell* 77, 6: 805-15.

Shearer D, Ens W, Standing K, and Valdimarsson G. (2008)"Posttranslational Modifications in Lens Fiber Connexins Identified by Off-Line-HPLC MALDI-Quadrupole Time-of-Flight Mass Spectrometry." *Investigative ophthalmology & visual science* 49, 4: 1553-62.

Shestopalov VI, and Panchin Y. (2008)"Pannexins and Gap Junction Protein Diversity." *Cellular and molecular life sciences : CMLS* 65, 3: 376-94.

Silva MT, do Vale A, and dos Santos NM. (2008)"Fish and Apoptosis: Studies in Disease and Pharmaceutical Design." *Current pharmaceutical design* 14, 2: 170-83.

Simon AM, Goodenough DA, and Paul DL. (1998)"Mice Lacking connexin40 have Cardiac Conduction Abnormalities Characteristic of Atrioventricular Block and Bundle Branch Block." *Current biology : CB* 8, 5: 295-8.

Simon AM, and McWhorter AR. (2002)"Vascular Abnormalities in Mice Lacking the Endothelial Gap Junction Proteins connexin37 and connexin40." *Developmental biology* 251, 2: 206-20.

Singal R, Tu ZJ, Vanwert JM, Ginder GD, and Kiang DT. (2000)"Modulation of the connexin26 Tumor Suppressor Gene Expression through Methylation in Human Mammary Epithelial Cell Lines." *Anticancer Research* 20, 1 A: 59-64.

Sivasubbu S, Balciunas D, Amsterdam A, and Ekker SC. (2007)"Insertional Mutagenesis Strategies in Zebrafish." *Genome biology* 8 Suppl 1, : S9.

Smith CW, Chu TT, and Nadal-Ginard B. (1993)"Scanning and Competition between AGs are Involved in 3' Splice Site Selection in Mammalian Introns." *Molecular and cellular biology* 13, 8: 4939-52.

Smith CW, Porro EB, Patton JG, and Nadal-Ginard B. (1989)"Scanning from an Independently Specified Branch Point Defines the 3' Splice Site of Mammalian Introns." *Nature* 342, 6247: 243-7.

Smith CWJ, Tung Tammy Chu, and Nadal-Ginard B. (1993)"Scanning and Competition between AGs are Involved in 3' Splice Site Selection in Mammalian Introns." *Molecular and Cellular Biology* 13, 8: 4939-52.

Sohl G, Degen J, Teubner B, and Willecke K. (1998)"The Murine Gap Junction Gene connexin36 is Highly Expressed in Mouse Retina and Regulated during Brain Development." *FEBS Letters* 428, 1-2: 27-31.

Sohl G, Eiberger J, Jung YT, Kozak CA, and Willecke K. (2001)"The Mouse Gap Junction Gene connexin29 is Highly Expressed in Sciatic Nerve and Regulated during Brain Development." *Biological chemistry* 382, 6: 973-8.

Sohl G, Gillen C, Bosse F, Gleichmann M, Muller HW, and Willecke K. (1996)"A Second Alternative Transcript of the Gap Junction Gene connexin32 is Expressed in Murine Schwann Cells and Modulated in Injured Sciatic Nerve." *European journal of cell biology* 69, 3: 267-75.

Sohl G, Theis M, Hallas G, Brambach S, Dahl E, Kidder G, and Willecke K. (2001)"A New Alternatively Spliced Transcript of the Mouse connexin32 Gene is Expressed in Embryonic Stem Cells, Oocytes, and Liver." *Experimental cell research* 266, 1: 177-86.

Sohl G, and Willecke K. (2004)"Gap Junctions and the Connexin Protein Family." *Cardiovascular research* 62, 2: 228-32.

Sohl G, and Willecke K. (2003)"An Update on Connexin Genes and their Nomenclature in Mouse and Man." *Cell communication & adhesion* 10, 4-6: 173-80.

Soller M. (2006)"Pre-Messenger RNA Processing and its Regulation: A Genomic Perspective." *Cellular and molecular life sciences : CMLS* 63, 7-8: 796-819.

Stetefeld J, and Ruegg MA. (2005)"Structural and Functional Diversity Generated by Alternative mRNA Splicing." *Trends in biochemical sciences* 30, 9: 515-21.

Stock A, and Sies H. (2000)"Thyroid Hormone Receptors Bind to an Element in the connexin43 Promoter." *Biological Chemistry* 381, 9-10: 973-9.

Stoneley M, and Willis AE. (2004)"Cellular Internal Ribosome Entry Segments: Structures, Trans-Acting Factors and Regulation of Gene Expression." *Oncogene* 23, 18: 3200-7.

Streisinger G, Walker C, Dower N, Knauber D, and Singer F. (1981)"Production of Clones of Homozygous Diploid Zebra Fish (Brachydanio Rerio)." *Nature* 291, 5813: 293-6.

Stuart GW, McMurray JV, and Westerfield M. (1988)"Replication, Integration and Stable Germ-Line Transmission of Foreign Sequences Injected into Early Zebrafish Embryos." *Development (Cambridge, England)* 103, 2: 403-12.

Stuart GW, Vielkind JR, McMurray JV, and Westerfield M. (1990)"Stable Lines of Transgenic Zebrafish Exhibit Reproducible Patterns of Transgene Expression." *Development (Cambridge, England)* 109, 3: 577-84.

Sugnet CW, Kent WJ, Ares Jr. M, and Haussler D. (2004)"Transcriptome and Genome Conservation of Alternative Splicing Events in Humans and Mice." *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing* , : 66-77.

Sumanas S, Kim HJ, Hermanson S, and Ekker SC. (2001)"Zebrafish Frizzled-2 Morphant Displays Defects in Body Axis Elongation." *Genesis (New York, N.Y.: 2000)* 30, 3: 114-8.

Sumanas S, and Larson JD. (2002)"Morpholino Phosphorodiamidate Oligonucleotides in Zebrafish: A Recipe for Functional Genomics?" *Briefings in functional genomics & proteomics* 1, 3: 239-56.

Summerton J, Stein D, Huang SB, Matthews P, Weller D, and Partridge M. (1997)"Morpholino and Phosphorothioate Antisense Oligomers Compared in Cell-Free and in-Cell Systems." *Antisense & Nucleic Acid Drug Development* 7, 2: 63-70.

Summerton J, and Weller D. (1997)"Morpholino Antisense Oligomers: Design, Preparation, and Properties." *Antisense & Nucleic Acid Drug Development* 7, 3: 187-95.

Summerton J. (1999)"Morpholino Antisense Oligomers: The Case for an RNase H-Independent Structural Type." *Biochimica et biophysica acta* 1489, 1: 141-58.

Szeto DP, and Kimelman D. (2004)"Combinatorial Gene Regulation by Bmp and Wnt in Zebrafish Posterior Mesoderm Formation." *Development (Cambridge, England)* 131, 15: 3751-60.

Tadokoro K, Yamazaki-Inoue M, Tachibana M, Fujishiro M, Nagao K, Toyoda M, Ozaki M, Ono M, Miki N, Miyashita T, et al. (2005)"Frequent Occurrence of Protein Isoforms with Or without a Single Amino Acid Residue by Subtle Alternative Splicing: The Case of Gln in DRPLA Affects Subcellular Localization of the Products." *Journal of Human Genetics* 50, 8: 382-94.

Takeda H, Matsuzaki T, Oki T, Miyagawa T, and Amanuma H. (1994)"A Novel POU Domain Gene, Zebrafish pou2: Expression and Roles of Two Alternatively Spliced Twin Products in Early Development." *Genes and Development* 8, 1: 45-59.

Tan L-, Bianco T, and Dobrovic A. (2002)"Variable Promoter Region CpG Island Methylation of the Putative Tumor Suppressor Gene Connexin 26 in Breast Cancer." *Carcinogenesis* 23, 2: 231-6.

Tanaka H, Maeda R, Shoji W, Wada H, Masai I, Shiraki T, Kobayashi M, Nakayama R, and Okamoto H. (2007)"Novel Mutations Affecting Axon Guidance in Zebrafish and a Role for Plexin Signalling in the Guidance of Trigeminal and Facial Nerve Axons." *Development* 134, 18: 3259-69.

Tang R, Dodd A, Lai D, McNabb WC, and Love DR. (2007)"Validation of Zebrafish (Danio Rerio) Reference Genes for Quantitative Real-Time RT-PCR Normalization." *Acta biochimica et biophysica Sinica* 39, 5: 384-90.

Thermes V, Grabher C, Ristoratore F, Bourrat F, Choulika A, Wittbrodt J, and Joly JS. (2002)"I-SceI Meganuclease Mediates Highly Efficient Transgenesis in Fish." *Mechanisms of development* 118, 1-2: 91-8.

Thisse B, Heyer V, Lux A, Alunni V, Degrave A, Seiliez I, Kirchner J, Parkhill JP, and Thisse C. (2004)"Spatial and Temporal Expression of the Zebrafish Genome by Large-Scale in Situ Hybridization Screening." *Methods in cell biology* 77, : 505-19.

Thisse C, Thisse B, Schilling TF, and Postlethwait JH. (1993)"Structure of the Zebrafish snail1 Gene and its Expression in Wild-Type, Spadetail and no Tail Mutant Embryos." *Development (Cambridge, England)* 119, 4: 1203-15.

Till BJ, Colbert T, Codomo C, Enns L, Johnson J, Reynolds SH, Henikoff JG, Greene EA, Steine MN, Comai L, et al. (2006)"High-Throughput TILLING for Arabidopsis." *Methods in molecular biology (Clifton, N.J.)* 323, : 127-35.

Till BJ, Colbert T, Tompa R, Enns LC, Codomo CA, Johnson JE, Reynolds SH, Henikoff JG, Greene EA, Steine MN, et al. (2003)"High-Throughput TILLING for Functional Genomics." *Methods in molecular biology (Clifton, N.J.)* 236, : 205-20.

Traver D, Herbomel P, Patton EE, Murphey RD, Yoder JA, Litman GW, Catic A, Amemiya CT, Zon LI, and Trede NS. (2003)"The Zebrafish as a Model Organism to Study Development of the Immune System." *Advances in Immunology* 81, : 253-330.

Treacy MN, Neilson LI, Turner EE, He X, and Rosenfeld MG. (1992)"Twin of I-POU: A Two Amino Acid Difference in the I-POU Homeodomain Distinguishes an Activator from an Inhibitor of Transcription." *Cell* 68, 3: 491-505.

Trinklein ND, Aldred SJ, Saldanha AJ, and Myers RM. (2003)"Identification and Functional Analysis of Human Transcriptional Promoters." *Genome research* 13, 2:

308-12.

Trosko JE. (2007)"Gap Junctional Intercellular Communication as a Biological "Rosetta Stone" in Understanding, in a Systems Biological Manner, Stem Cell Behavior, Mechanisms of Epigenetic Toxicology, Chemoprevention and Chemotherapy." *The Journal of membrane biology* 218, 1-3: 93-100.

Tsai K-, and Lin W-. (2006)"Quantitative Analysis of Wobble Splicing Indicates that it is Not Tissue Specific." *Genomics* 88, 6: 855-64.

Tsonis PA, and Fuentes EJ. (2006)"Focus on Molecules: Pax-6, the Eye Master." *Experimental eye research* 83, 2: 233-4.

Udvadia AJ, Köster RW, and Skene JHP. (2001)"GAP-43 Promoter Elements in Transgenic Zebrafish Reveal a Difference in Signals for Axon Growth during CNS Development and Regeneration." *Development* 128, 7: 1175-82.

Udvadia AJ, and Linney E. (2003)"Windows into Development: Historic, Current, and Future Perspectives on Transgenic Zebrafish." *Developmental biology* 256, 1: 1-17.

Uemura O, Okada Y, Ando H, Guedj M, Higashijima S-, Shimazaki T, Chino N, Okano H, and Okamoto H. (2005)"Comparative Functional Genomics Revealed Conservation and Diversification of Three Enhancers of the isl1 Gene for Motor and Sensory Neuron-Specific Expression." *Developmental Biology* 278, 2: 587-606.

Unger VM, Kumar NM, Gilula NB, and Yeager M. (1999)"Three-Dimensional Structure of a Recombinant Gap Junction Membrane Channel." *Science (New York, N.Y.)* 283, 5405: 1176-80.

Unoki M, Shen JC, Zheng Z-, and Harris CC. (2006)"Novel Splice Variants of ING4 and their Possible Roles in the Regulation of Cell Growth and Motility." *Journal of Biological Chemistry* 281, 45: 34677-86.

Valcarcel J, and Green MR. (1996)"The SR Protein Family: Pleiotropic Functions in Pre-mRNA Splicing." *Trends in biochemical sciences* 21, 8: 296-301.

Valiunas V, Mui R, McLachlan E, Valdimarsson G, Brink PR, and White TW. (2004)"Biophysical Characterization of Zebrafish connexin35 Hemichannels." *American journal of physiology.Cell physiology* 287, 6: C1596-604.

Van der Heyden MAG, Rook MB, Hermans MMP, Rijksen G, Boonstra J, Defize LHK,

and Destrée OHJ. (1998)"Identification of connexin43 as a Functional Target for Wnt Signalling." *Journal of Cell Science* 111, 12: 1741-9.

van der Sar AM, Appelmelk BJ, Vandenbroucke-Grauls CM, and Bitter W. (2004)"A Star with Stripes: Zebrafish as an Infection Model." *Trends in microbiology* 12, 10: 451-7.

van der Velden AW, and Thomas AA. (1999)"The Role of the 5' Untranslated Region of an mRNA in Translation Regulation during Development." *The international journal of biochemistry & cell biology* 31, 1: 87-106.

van Veen TA, van Rijen HV, and Jongsma HJ. (2006)"Physiology of Cardiovascular Gap Junctions." *Advances in Cardiology* 42, : 18-40.

VanderBrink BA, Sellitto C, Saba S, Link MS, Zhu W, Homoud MK, Estes NA,3rd, Paul DL, and Wang PJ. (2000)"Connexin40-Deficient Mice Exhibit Atrioventricular Nodal and Infra-Hisian Conduction Abnormalities." *Journal of cardiovascular electrophysiology* 11, 11: 1270-6.

Verheule S, van Kempen MJ, Postma S, Rook MB, and Jongsma HJ. (2001)"Gap Junctions in the Rabbit Sinoatrial Node." *American journal of physiology.Heart and circulatory physiology* 280, 5: H2103-15.

Vilela C, Ramirez CV, Linz B, Rodrigues-Pousada C, and McCarthy JE. (1999)"Post-Termination Ribosome Interactions with the 5'UTR Modulate Yeast mRNA Stability." *The EMBO journal* 18, 11: 3139-52.

Vine AL, Leung YM, and Bertram JS. (2005)"Transcriptional Regulation of Connexin 43 Expression by Retinoids and Carotenoids: Similarities and Differences." *Molecular Carcinogenesis* 43, 2: 75-85.

Vinken M, Vanhaecke T, and Rogiers V. (2003)"The Role of Intercellular Communication Via "Gap Junctions" in Disease." *Nederlands tijdschrift voor geneeskunde* 147, 50: 2463-6.

Vogan KJ, Underhill DA, and Gros P. (1996)"An Alternative Splicing Event in the Pax-3 Paired Domain Identifies the Linker Region as a Key Determinant of Paired Domain DNA-Binding Activity." *Molecular and Cellular Biology* 16, 12: 6677-86.

von Maltzahn J, Euwens C, Willecke K, and Söhl G. (2004)"The Novel Mouse connexin39 Gene is Expressed in Developing Striated Muscle Fibers." *Journal of*

Cell Science 117, 22: 5381-92.

Wada H, Tanaka H, Nakayama S, Iwasaki M, and Okamoto H. (2006)"Frizzled3a and Celsr2 Function in the Neuroepithelium to Regulate Migration of Facial Motor Neurons in the Developing Zebrafish Hindbrain." *Development* 133, 23: 4749-59.

Wang Y, Newton DC, Robb GB, Kau CL, Miller TL, Cheung AH, Hall AV, VanDamme S, Wilcox JN, and Marsden PA. (1999)"RNA Diversity has Profound Effects on the Translation of Neuronal Nitric Oxide Synthase." *Proceedings of the National Academy of Sciences of the United States of America* 96, 21: 12150-5.

Wargelius A, Ellingsen S, and Fjose A. (1999)"Double-Stranded RNA Induces Specific Developmental Defects in Zebrafish Embryos." *Biochemical and Biophysical Research Communications* 263, 1: 156-61.

Warnecke C, Willich T, Holzmeister J, Bottari SP, Fleck E, and Regitz-Zagrosek V. (1999)"Efficient Transcription of the Human Angiotensin II Type 2 Receptor Gene Requires Intronic Sequence Elements." *The Biochemical journal* 340 (Pt 1), Pt 1: 17-24.

Warren KS, and Fishman MC. (1998)""Physiological Genomics": Mutant Screens in Zebrafish." *The American Journal of Physiology* 275, 1 Pt 2: H1-7.

Wasserman WW, and Sandelin A. (2004)"Applied Bioinformatics for the Identification of Regulatory Elements." *Nature reviews.Genetics* 5, 4: 276-87.

Wawersik S, Purcell P, and Maas RL. (2000)"Pax6 and the Genetic Control of Early Eye Development." *Results and problems in cell differentiation* 31, : 15-36.

Wei CJ, Xu X, and Lo CW. (2004)"Connexins and Cell Signaling in Development and Disease." *Annual Review of Cell and Developmental Biology* 20, : 811-38.

Weis L, and Reinberg D. (1992)"Transcription by RNA Polymerase II: Initiator-Directed Formation of Transcription-Competent Complexes." *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 6, 14: 3300-9.

Werner R. (2000)"IRES Elements in Connexin Genes: A Hypothesis Explaining the Need for Connexins to be Regulated at the Translational Level." *IUBMB life* 50, 3: 173-6.

Werner T. (1999)"Models for Prediction and Recognition of Eukaryotic Promoters."

Mammalian Genome 10, 2: 168-75.

Westerfield M. (1995) The Zebrafish Book. University of Oregon Press, Eugene.

Westerfield M, Wegner J, Jegalian BG, DeRobertis EM, and Puschel AW. (1992)"Specific Activation of Mammalian Hox Promoters in Mosaic Transgenic Zebrafish." *Genes and Development* 6, 4: 591-8.

White TW, Bruzzone R, Goodenough DA, and Paul DL. (1992)"Mouse Cx50, a Functional Member of the Connexin Family of Gap Junction Proteins, is the Lens Fiber Protein MP70." *Molecular Biology of the Cell* 3, 7: 711-20.

White TW, and Bruzzone R. (1996)"Multiple Connexin Proteins in Single Intercellular Channels: Connexin Compatibility and Functional Consequences." *Journal of Bioenergetics and Biomembranes* 28, 4: 339-50.

White TW, Goodenough DA, and Paul DL. (1998)"Targeted Ablation of connexin50 in Mice Results in Microphthalmia and Zonular Pulverulent Cataracts." *Journal of Cell Biology* 143, 3: 815-25.

White TW, Paul DL, Goodenough DA, and Bruzzone R. (1995)"Functional Analysis of Selective Interactions among Rodent Connexins." *Molecular Biology of the Cell* 6, 4: 459-70.

White TW, and Paul DL. (1999)"Genetic Diseases and Gene Knockouts Reveal Diverse Connexin Functions." *Annual Review of Physiology* 61, : 283-310.

White TW, Sellitto C, Paul DL, and Goodenough DA. (2001)"Prenatal Lens Development in connexin43 and connexin50 Double Knockout Mice." *Investigative ophthalmology & visual science* 42, 12: 2916-23.

White TW. (2003)"Nonredundant Gap Junction Functions." *News in physiological sciences : an international journal of physiology produced jointly by the International Union of Physiological Sciences and the American Physiological Society* 18, : 95-9.

White TW. (2002)"Unique and Redundant Connexin Contributions to Lens Development." *Science (New York, N.Y.)* 295, 5553: 319-20.

Whitelaw CB, Archibald AL, Harris S, McClenaghan M, Simons JP, and Clark AJ. (1991)"Targeting Expression to the Mammary Gland: Intronic Sequences can

Enhance the Efficiency of Gene Expression in Transgenic Mice." *Transgenic research* 1, 1: 3-13.

Whitfield TT. (2002)"Zebrafish as a Model for Hearing and Deafness." *Journal of neurobiology* 53, 2: 157-71.

Wienholds E, and Plasterk RH. (2004)"Target-Selected Gene Inactivation in Zebrafish." *Methods in cell biology* 77, : 69-90.

Wienholds E, Schulte-Merker S, Walderich B, and Plasterk RH. (2002)"Target-Selected Inactivation of the Zebrafish rag1 Gene." *Science (New York, N.Y.)* 297, 5578: 99-102.

Wienholds E, van Eeden F, Kosters M, Mudde J, Plasterk RHA, and Cuppen E. (2003)"Efficient Target-Selected Mutagenesis in Zebrafish." *Genome Research* 13, 12: 2700-7.

Willecke K, Eiberger J, Degen J, Eckardt D, Romualdi A, Guldenagel M, Deutsch U, and Sohl G. (2002)"Structural and Functional Diversity of Connexin Genes in the Mouse and Human Genome." *Biological chemistry* 383, 5: 725-37.

Woltering JM, and Durston AJ. (2008)"MiR-10 Represses HoxB1a and HoxB3a in Zebrafish." *PLoS ONE* 3, 1: e1396.

Woodley L, and Valcarcel J. (2002)"Regulation of Alternative Pre-mRNA Splicing." *Briefings in functional genomics & proteomics* 1, 3: 266-77.

Wu YL, Pan X, Mudumana SP, Wang H, Kee PW, and Gong Z. (2008)"Development of a Heat Shock Inducible Gfp Transgenic Zebrafish Line by using the Zebrafish hsp27 Promoter." *Gene* 408, 1-2: 85-94.

Xie J, and Black DL. (2001)"A CaMK IV Responsive RNA Element Mediates Depolarization-Induced Alternative Splicing of Ion Channels." *Nature* 410, 6831: 936-9.

Xie J. (2008)"Control of Alternative Pre-mRNA Splicing by Ca(++) Signals." *Biochimica et biophysica acta* 1779, 8: 438-52.

Xu Q, Belcastro MP, Villa ST, Dinkins RD, Clarke SG, and Downie AB. (2004)"A Second Protein L-Isoaspartyl Methyltransferase Gene in Arabidopsis Produces Two Transcripts Whose Products are Sequestered in the Nucleus." *Plant Physiology* 136,

1: 2652-64.

Xu X, Yang D, Ding JH, Wang W, Chu PH, Dalton ND, Wang HY, Bermingham JR,Jr, Ye Z, Liu F, et al. (2005)"ASF/SF2-Regulated CaMKIIdelta Alternative Splicing Temporally Reprograms Excitation-Contraction Coupling in Cardiac Muscle." *Cell* 120, 1: 59-72.

Yan M, Wang L-, Hymowitz SG, Schilbach S, Lee J, Goddard A, De Vos AM, Gao W-, and Dixit VM. (2000)"Two-Amino Acid Molecular Switch in an Epithelial Morphogen that Regulates Binding to Two Distinct Receptors." *Science* 290, 5491: 523-7.

Yancey SB, Biswal S, and Revel JP. (1992)"Spatial and Temporal Patterns of Distribution of the Gap Junction Protein connexin43 during Mouse Gastrulation and Organogenesis." *Development (Cambridge, England)* 114, 1: 203-12.

Yano T, Ito F, Kobayashi K, Yonezawa Y, Suzuki K, Asano R, Hagiwara K, Nakazawa H, Toma H, and Yamasaki H. (2004)"Hypermethylation of the CpG Island of Connexin 32, a Candiate Tumor Suppressor Gene in Renal Cell Carcinomas from Hemodialysis Patients." *Cancer Letters* 208, 2: 137-42.

Yeager M, Unger VM, and Falk MM. (1998)"Synthesis, Assembly and Structure of Gap Junction Intercellular Channels." *Current opinion in structural biology* 8, 4: 517-24.

Yen MR, and Saier MH,Jr. (2007)"Gap Junctional Proteins of Animals: The innexin/pannexin Superfamily." *Progress in biophysics and molecular biology* 94, 1-2: 5-14.

Yeo S-, Little MH, Yamada T, Miyashita T, Halloran MC, Kuwada JY, Huh T-, and Okamoto H. (2001)"Overexpression of a Slit Homologue Impairs Convergent Extension of the Mesoderm and Causes Cyclopia in Embryonic Zebrafish." *Developmental Biology* 230, 1: 1-17.

Yoder JA, Nielsen ME, Amemiya CT, and Litman GW. (2002)"Zebrafish as an Immunological Model System." *Microbes and infection / Institut Pasteur* 4, 14: 1469-78.

Yu W, Dahl G, and Werner R. (1994)"The connexin43 Gene is Responsive to Oestrogen." *Proceedings of the Royal Society B: Biological Sciences* 255, 1343: 125-32.

Zavolan M, Kondo S, Schönbach C, Adachi J, Hume DA, Arakawa T, Carninci P, Kawai

J, Hayashizaki Y, and Gaasterland T. (2003)"Impact of Alternative Initiation, Splicing, and Termination on the Diversity of the mRNA Transcripts Encoded by the Mouse Transcriptome." *Genome Research* 13, 6 B: 1290-300.

Zavolan M, and van Nimwegen E. (2006)"The Types and Prevalence of Alternative Splice Forms." *Current opinion in structural biology* 16, 3: 362-7.

Zhang JT, and Nicholson BJ. (1989)"Sequence and Tissue Distribution of a Second Protein of Hepatic Gap Junctions, Cx26, as Deduced from its cDNA." *The Journal of cell biology* 109, 6 Pt 2: 3391-401.

Zhang T, Haws P, and Wu Q. (2004)"Multiple Variable First Exons: A Mechanism for Cell- and Tissue-Specific Gene Regulation." *Genome research* 14, 1: 79-89.

Zhao Z, Cao Y, Li M, and Meng A. (2001)"Double-Stranded RNA Injection Produces Nonspecific Defects in Zebrafish." *Developmental Biology* 229, 1: 215-23.

Zon LI, and Peterson RT. (2005)"In Vivo Drug Discovery in the Zebrafish." *Nature reviews.Drug discovery* 4, 1: 35-44.

Zuber ME, Gestri G, Viczian AS, Barsacchi G, and Harris WA. (2003)"Specification of the Vertebrate Eye by a Network of Eye Field Transcription Factors." *Development (Cambridge, England)* 130, 21: 5155-67.

Zuker M. (2003)"Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction." *Nucleic acids research* 31, 13: 3406-15.