# INTRACELLULAR LIPID-BINDING PROTEIN GENES FROM ZEBRAFISH (*Danio rerio*): DIFFERENTIAL EXPRESSION AND THE FATE OF DUPLICATED GENES

by

Rong-Zong Liu

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
December 2003

*Your file Votre référence*
*ISBN: 0-612-89806-7*
*Our file Notre référence*
*ISBN: 0-612-89806-7*

DALHOUSIE UNIVERSITY

DEPARTMENT OF BIOLOGY

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "Intracellular lipid-binding protein genes from zebrafish (*Danio rerio*): differential expression and the fate of duplicated genes" by Rong-Zong Liu in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dated: December 19, 2003

External Examiner:

Research Supervisors:

Examining Committee:

Departmental Representative:

**DALHOUSIE UNIVERSITY**

DATE: December 19th, 2003

AUTHOR: RONG-ZONG LIU

TITLE: Intracellular lipid-binding protein genes from zebrafish (*Danio rerio*): differential expression and the fate of duplicated genes

DEPARTMENT OR SCHOOL: BIOLOGY

DEGREE: PhD CONVOCATION: May YEAR: 2004

Permission is herewith granted to Dalhousie University to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions.

_____
Signature of Author

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ABSTRACT

Six genes encoding intracellular lipid-binding proteins (iLBPs), including fatty acid-binding proteins 3 (*fabp3*) and 7 (*fabp7a* and *fabp7b*), cellular retinol-binding protein 1 (*rbp1a* and *rbp1b*) and 2 (*rbp2*) were identified in the zebrafish (*Danio rerio*) genome. cDNAs for *fabp3*, *fabp7b*, *rbp1a* and *rbp1b* were cloned by 3' RACE and 5'RLM-RACE and sequenced. The structure of each gene was defined through data mining of the zebrafish genomic DNA database. Transcription start site(s) for each gene were mapped by 5'RLM-RACE. Phylogenetic analysis and synteny with mammalian orthologs confirmed the identity of each gene, and radiation hybrid mapping assigned each gene to a linkage group. Tissue-specific distribution of iLBP gene transcripts in adult zebrafish was investigated by RT-PCR, and by tissue-section *in situ* hybridization and emulsion autoradiography. Spatio-temporal distribution of the iLBP gene transcripts during development was examined using whole mount *in situ* hybridization on zebrafish embryos and larvae at different developmental stages. The results showed that the genomic properties of the zebrafish iLBP genes, including gene structure, coding features and syntenic relationship, are conserved when compared with their mammalian orthologs, while both conserved and diverged expression patterns were observed. Gene duplication is common within the zebrafish iLBP multigene family. Genomic analysis of two duplicated iLBP gene pairs showed that the duplicate iLBP genes might arise from chromosomal, or possibly, whole genome duplication. Comparative analysis of the spatio-temporal distribution of mRNA from sister copies of the duplicated iLBP gene pairs provided new insights into the evolution of gene functions and the mechanism for preservation of duplicate genes in the vertebrate genomes.

# LIST OF ABBREVIATIONS AND SYMBOLS

| | |
|---|---|
| B-FABP | brain fatty acid-binding protein |
| BLBP | brain lipid-binding protein |
| bp | base pair |
| cDNA | complementary deoxyribonucleic acid |
| CIP | calf intestinal phosphatase |
| cM | centimorgan |
| CNS | central nervous system |
| cpm | count per minute |
| CRABP | cellular retinoic acid-binding protein |
| CRBP | cellular retinol-binding protein |
| CRBPI | cellular retinol-binding protein type I |
| CRBPII | cellular retinol-binding protein type II |
| DDC | duplication-degeneration-complementation |
| DHA | docosahexaenoic acid |
| dpf | days post fertilization |
| EST | expressed sequence tag |
| FABP | fatty acid-binding protein |
| FABP3 | heart-type fatty acid-binding protein |
| *fapb3* | gene for the heart-type fatty acid-binding protein |

| | |
|---|---|
| *fabp7* | gene for the brain-type fatty acid-binding protein |
| FFA | free fatty acid |
| hpf | hours post fertilization |
| iLBP | intracellular lipid-binding protein |
| kb | kilobase pairs |
| LFP | lateral floor plate |
| LG | linkage group |
| LM-PCR | linker-mediated polymerase chain reaction |
| MFP | medial floor plate |
| mya | million years ago |
| nt | nucleotide |
| PCR | polymerase chain reaction |
| PBS | phosphate-buffered saline |
| PPAR | peroxisome proliferator-activated receptor |
| qRT-PCR | quantitative reverse transcription-polymerase chain reaction |
| r4 | rhombomere four |
| RA | retinoic acid |
| RACE | rapid amplification of cDNA ends |
| RAR | retinoic acid receptor |
| *rbp1* | cellular retinol-binding protein type I gene |
| *rbp2* | cellular retinol-binding protein type II gene |
| RH | radiation hybrid |

| | |
|---|---|
| RLM-RACE | RNA ligase mediated- rapid amplification of cDNA ends |
| RT-PCR | reverse transcription-polymerase chain reaction |
| RXR | retinoid x receptor |
| TAP | tobacco acid pyrophosphatase |
| UTR | untranslated region |
| VA | vitamin A |

# ACKNOWLEDGEMENTS

I am most indebted to my supervisors, Dr. Jonathan Wright and Dr. Eileen Denovan-Wright, who made my PhD studies possible, enjoyable and, I would say, rewarding. They not only instructed me how to perform scientific research, but also taught me how to become a qualified scientist. Their teachings were the key to my achievements during the four and half years of my PhD studies, and will be a treasure for my future academic life. I will never forget their kindness and all their help during the whole journey of my studies. I would like to give my heartfelt thanks to Dr. Tom MacRae and Dr. Vett Lloyd, my supervising committee members, for their directions on my project and their encouragement.

The embryo whole mount *in situ* hybridization experiments of this study were done by Dr. Christine Thisse, Dr. Bernard Thisse and Dr. Agnes Degrave in the Institut de Génétique et Biologie Moléculaire et Cellulaire, University of Strasbourg, France. I thank them for their collaboration in this research project. I am grateful to Dr. Zhou Yi from Children's Hospital, Boston, USA, for his assistance in mapping the zebrafish *fabp7b* gene. I thank Miss Kathleen Murphy for her kind help in the tissue section *in situ* hybridization experiment.

I thank my labmates, Mukesh, Steve, Qian and Vishal for their friendship and their help. I enjoyed so much working with them.

I would like to give my sincere thanks to my wife, Yuying, for her understanding, love and support. She left her job and everything that she had achieved and came to Canada to accompany me and help me with my studies. I can never make up what I owe her! I would like to express my greatest gratitude to my mother and my passed father for their consistent and unconditional love.

I thank God, my Lord and Saviour, for His everlasting love and mercy on me. I believe that "where there is faith, there is a way"!

# 1 General Introduction

## 1.1 INTRACELLULAR LIPID-BINDING PROTEINS AND THEIR GENES

Intracellular lipid-binding proteins (iLBPs) are a group of low-molecular-mass (~ 15 kDa) proteins that bind long-chain fatty acids, retinoids or other hydrophobic ligands [reviewed in Ong *et al.*, 1994; Glatz *et al.*, 1996; Bernlohr *et al.*, 1997]. iLBPs which bind long-chain fatty acids are referred to as fatty acid-binding proteins (FABPs) and another sub-group of iLBPs are know as the cellular retinoid-binding proteins; these iLBPs specifically bind to retinol and retinal (CRBPs) or retinoic acids (CRABPs). Following the first report of a cytoplasmic protein binding to fatty acids thirty years ago [Ockner, 1972], iLBPs have been isolated from various species and their structure and physical properties intensively studied, primarily in mammals. iLBPs are single-chain polypeptides that consist of ten anti-parallel β-strands and two short α-helices situated between the first and second β-stands. Each β-stand is hydrogen-bonded to the adjacent strand to form two five-stranded β-sheets, which fold together to form a tertiary structure known as the so-called "β-barrel" [Ong *et al.*, 1994; Glatz *et al.*, 1996; Bernlohr *et al.*, 1997]. Although the amino acid sequence of different paralogous members of the iLBP family varies from 20% to 70%, the "β-barrel" tertiary structure is strikingly similar for all members of the iLBP family. Hydrophobic ligands such as fatty acids and retinoids are accommodated in the central cavity of the iLBP β-barrel, which, therefore, dramatically enhances the solubility of the ligands within the aqueous cytoplasm and

helps them move to the target sites within the cell to exert their biological actions [reviewed in Glatz *et al.*, 1996; Bernlohr *et al.*, 1997].

In vertebrates, at least sixteen paralogous iLBPs, including ten fatty acid-binding proteins and six cellular retinoid-binding proteins, have been identified, each of which shows specific temporal and spatial distribution patterns during mammalian development and in adulthood [Glatz *et al.*, 1996; Ong *et al.*, 1994]. The tissue-specific distribution of orthologous iLBPs is basically conserved among species. For instance, B-FABP is usually present in the brain tissues and CRBPII is always detected in the intestine. However, this is not true for all iLBP members. CRBPI, for example, is abundant in the rat kidney and liver, while in humans CRBPI levels are highest in the ovary and CRBPI levels are relatively low in the kidney and liver [reviewed by Ong *et al.*, 1994]. A basic liver type FABP (Lb-FABP) was identified in birds and fishes [Di Pietro *et al.*, 1997; Denovan-Wright *et al.*, 2000; Vasile *et al.*, 2003] but has not been identified in mammals. In contrast, mammals but not other vertebrates have the unique CRBPIII and CRBPIV [Folli *et al.*, 2001; 2002; Vogel *et al.*, 2001].

Although several iLBP gene-knockout experiments in mice have tried to provide direct evidence for the biological functions of FABPs [Binas *et al.*, 1999; Schaap *et al.*, 1999; Ribarik *et al.*, 1999], the precise physiological role(s) played by each iLBP have not been well defined. The proposed functions of iLBPs include cellular uptake and transport of long-chain fatty acids and retinoids, interaction with other transport and enzyme systems, signal transduction, and cell protection. [Glatz *et al.*, 1996; Ong *et al.*, 1994; Napoli, 1999; Stewart, 2000].

iLBPs are encoded by a multigene family, the members of which are dispersed throughout the vertebrate genome. The gene structure of the iLBP family has been conserved among all paralogous members from all vertebrate species [Ong *et al.*, 1994; Bernlohr *et al.*, 1997]. Each gene contains four exons with similar coding capacity and the exons are separated by three introns of varying sizes [Ong *et al.*, 1994; Bernlohr *et al.*, 1997; Schaap *et al.*, 2002]. In mammals, each iLBP gene only exists in a single functional copy in the genome, although nontranscribed pseudogenes have been identified for H-FABP [Treuner *et al.*, 1994]. It has been proposed that the iLBP multigene family has undergone at least 14 gene duplications. The liver/intestinal/ileal FABP clade emerged from the heart/adipose/myelin P2 FABP lineage some 700 million years ago (mya), prior to the vertebrate/invertebrate divergence [Schleicher *et al.*, 1995]. The CRBP and CRABP members seem to have diverged from the liver/intestinal FABP clade about 500 mya. The mammalian CRBPI and CRBPII are thought to have arisen by gene duplication after the split with the amphibia, as *Xenopus* has a single CRBP gene [Matarese *et al.*, 1989].

## 1.2 THE ZEBRAFISH (*Danio rerio*) SYSTEM

Although some iLBPs have been identified in fishes, the largest and most evolutionary diverse group of vertebrates, and other lower (or earlier) vertebrates [reviewed by Londraville, 1996], our knowledge of the characteristics of iLBPs and theirs genes are mainly from studies in mammals. However, lower vertebrates, such as fish, differ from mammals dramatically in lipid metabolism and homeostasis. Comparative studies of iLBPs and their genes in these non-mammalian systems may, therefore,

provide us with greater knowledge of iLBP functions, and offer insight into the mechanisms governing their function. Zebrafish (*Danio rerio*) has become an increasingly popular model organism of vertebrate in developmental and genetic studies. Zebrafish have advantages over mammalian systems to study development that include rapid development, high reproductivity, transparent embryos, and low cost and ease of maintenance. Most importantly, genomic, molecular and cellular methodologies are advancing rapidly for study of zebrafish. Several high-resolution linkage mapping panels are available for rapid genomic assignment of cloned genes [Kwok *et al.*, 1998; Hukriede *et al.*, 1999; Kelly *et al.*, 2000; Knapik *et al.*, 2000], large-scale zebrafish mutant screens have been achieved [Driever, et al., 1996; Pattonand Zon, 2001], and sequencing of the zebrafish genome is almost complete. Other techniques such as whole mount *in situ* hybridization [Westerfield, 1995; Jowett, 1999], green fluorescent protein reporter systems [Amsterdam *et al.*, 1995; 1996; Moss *et al.*, 1996] and ectopic gene expression techniques [Hyatt *et al.*, 1999] are now routinely used in zebrafish. Undoubtedly, studies using the zebrafish will make great contributions to our understanding of vertebrate gene function and development.

## 1.3 GENE DUPLICATION AND EVOLUTION OF VERTEBRATE GENOMES

For over three decades it has been argued that duplication of individual genes, chromosomal segments or whole genomes plays a crucial role in genome evolution, and likely contributes to morphological diversification of vertebrates [Ohno, 1970; Holland *et al.*, 1994; Sidow, 1996; Robinson-Rechavi *et al.*, 2001]. One prominent example is the *hox* gene family. In the amphioxus genomes, only one *hox* cluster has been defined

[Spruyt, 1998]. Human and mouse have four paralogous *hox* gene clusters in their genomes [Holland *et al.*, 1994; Holland, 1999], In contrast, zebrafish has seven paralogous hox gene clusters, presumably arising from either chromosomal or genome duplication as genes closely linked to the *hox* clusters are also duplicated in zebrafish [Amores *et al.*, 1998; Prince *et al.*, 1998; Woods *et al.*, 2000]. Recently, the number of teleost ESTs, mainly zebrafish (*Danio rerio*) and pufferfish (*Takifugu rubripes*), has increased dramatically. Analysis of these ESTs has indicated that teleost fishes have numerous duplicated gene loci and phylogenetic analysis suggests a whole genome duplication may have occurred in the euteleost fish lineage [Taloy *et al.*, 2003].

Investigation of phylogeny, syntenic relationship and spatio-temporal expression patterns of large multigene families would provide insight into the genesis and fate of duplicated genes, and the evolution of the vertebrate genome. Gene duplication events, however, have only been investigated in a few multigene families, such as the *hox* family [Amores *et al.*, 1998], the *sox* family [Chiang *et al.*, 2001a], the *jak* family [Oates *et al.*, 1999], the nuclear receptor gene family [Robinson-Rechavi et a., 2001], the P450 aromatase gene family [Chiang *et al.*, 2001b], *h/E(spl)/hey* family [Gajewski and Voolstra, 2002] and the *anx* gene family [Farber *et al.*, 2003]. The iLBP multigene family is dispersed in the whole genome, rather than forming clusters in the vertebrate genome like *hox* genes, and, therefore, can provide information from more gene loci. Because the genomic properties and expression patterns are known for many mammalian iLBP genes, the determination of iLBP gene structure and expression, and comparative analysis of iLBPs in mammals and zebrafish may provide insight into the vertebrate gene and genome evolution.

## 1.4 OBJECTIVES OF THIS STUDY

The objectives of this project were: (1) to clone and characterize several iLBP genes from zebrafish, (2) to determine their structure and patterns of expression and compare them to mammalian orthologs, and (3) to investigate the function and fate of duplicated genes of the zebrafish iLBP multigene family by analyzing the spatio-temporal distribution of their transcripts during development and in adulthood.

# 2 Materials and Methods

## 2.1 Zebrafish culture and breeding

Zebrafish were purchased from a local aquarium store and cultured in filtered, aerated water at 28.5 °C in 35 L aquaria. Fish were maintained on a 24 h cycle of 14 h light and 10 h darkness. Fish were fed with a dry fish feed, TetraMin Flakes (TetraWerke, Melle, Germany), in the morning and hatched brine shrimp (*Artemia* cysts from INVE, Grantsville, UT, USA) in the afternoon. Fish breeding and embryo manipulation were conducted according to established protocols [Westerfield, 1995].

## 2.2 3' rapid amplification of cDNA ends (3' RACE)

3′ Rapid Amplification of cDNA Ends (3′ RACE) was employed to clone the cDNA sequences for the zebrafish iLBPs using primers based on the corresponding zebrafish ESTs in GenBank or gene sequences from the zebrafish genomic DNA sequence database (The Wellcome Trust Sanger Institute, Cambridge, UK). The gene-specific primers used for 3'RACE are summarized in Table 2-1.

First strand cDNA was synthesized from total RNA of adult zebrafish with a 3′ adaptor primer (5′-GGCCACGCGTCGACTAGTAC$T_{17}$-3′) and reverse transcriptase Superscript II (GibcoBRL, Burlington, Ontario, Canada). Using the reverse-transcription reaction as template, the cDNA sequence of each gene was amplified by PCR using the "antisense" primer complementary to the 3′ adaptor and nested gene-specific 5′ "sense"

primers (Table 2-1). The PCR products were separated by gel-electrophoresis in low melting point agarose and purified from the gel using the Qiaquick gel extraction kit (Qiagen, Mississauga, Ontario, Canada). The purified DNA fragments were cloned into the plasmid vector, pGEM-T (Promega, Madison, WI, USA), and three well-isolated positive clones were sequenced (University of Toronto Sequencing Facility, Toronto, Ontario, Canada) for each clone from both directions. The deduced amino acid sequence of the iLBP was determined using the algorithm in Gene Runner V. 3.05 (Hastings Software, Inc.). Nucleotide and amino acid sequences were aligned using CLUSTALW [Thompson *et al.*, 1994]. cDNA sequences for the zebrafish *fabp7a* [Denovan-Wright *et al.*, 2001] and *rbp2a* [Cameron *et al.*, 2002] were previously determined.

**Table 2-1. Gene-specific primers used for 3'RACE**

| Gene | Primers |
|---|---|
| *fabp7b* | Outer: 5′-TGATTCAGAAAATGGTGGATG-3′<br>Inner: 5′-GGTCAGCAGTGACAACTTTG-3′ |
| *fabp3* | Primer: 5′-TCAGCTCAAACATGGCAGAC-3′ |
| *rbp1a* | Outer: 5′-CACGGAGTAAACTCTCGG-3′<br>Inner: 5′-CAGAGGAACAGTGAGCACAC-3′ |
| *rbp1b* | Outer: 5′-CGACACGGCATTCAGGTCTG-3′<br>Inner: 5′-AATCGCGTTGCCTGTAGTTC-3′ |

### 2.3 Southern blot and hybridization

Genomic DNA was isolated from adult zebrafish by standard methodology [Sambrook *et al.*, 1989]. Eight μg aliquots of genomic DNA were digested individually with *Hinc*II, *Hae*III, *Mbo*I or *Rsa*I. The digested DNA was size-fractionated in a 0.8%

(w/v) agarose gel and transferred to a nylon membrane in an alkaline transfer solution containing 0.4 N NaOH and 1 M NaCl. A hybridization probe specific for the coding region of the *fabp3* gene was generated by RT-PCR and radiolabelled in a subsequent PCR with [$\alpha^{32}$P]-dATP (Amersham Pharmacia Biotech, Baie d'Urfé, Quebec, Canada). The membrane was pre-hybridized at 68°C for 2 h in a solution containing 5xSSPE (1xSSPE: 180 mM NaCl, 10 mM $NaH_2PO_4$, 1 mM EDTA, pH 7.4), 5x Denhardt's solution (1X Denhardt's: 0.1% w/v polyvinylpyrollidone, 0.1% w/v bovine serum albumin, 0.1% w/v Ficoll), 100 μg/mL yeast tRNA and 0.5% w/v sodium dodecyl sulphate (SDS). The hybridization probe was added to the solution at $5x10^5$ cpm/mL and hybridization was allowed to proceed for approximately 16 h. The blot was washed twice at room temperature for 5 min in 2xSSC (1X SSC: 0.15 M NaCl, 0.015 sodium citrate, pH 7.0) and 0.1% SDS, twice at 68°C for 15 min in 0.2xSSC/0.1%SDS, and the membrane was exposed to X-ray film at –70°C for 48 h.

## 2.4 Promoter cloning for iLBP promoters

Linker-mediated polymerase chain reaction (LM-PCR) was employed to clone the core promoter and upstream regulatory elements of the zebrafish *fabp7a* and *fabp3* gene. Genomic DNA was isolated from adult zebrafish and purified according to a standard protocol [Sambrook *et al.*, 1989]. Two μg of genomic DNA was digested with the restriction enzyme, *Bam*HI, and 0.5 μg of the digest was ligated to the double-stranded DNA linker,

```
5′-GTACATATTGTCGTTAGAACGCGTAATACGACTCACTATAGGGA-3′
3′-CATGTATAACAGCAATCTTGCGCATTATGCTGAGTGATATCCCTCTAG-5′
```

using T4 DNA ligase (Promega). Following precipitation, the DNA was resuspended in

15 μL of sterile, distilled water. Two partially overlapping sense primers (C1, C2) were synthesized based on the linker sequence (C1: 5′-GTACATATTGTCGTTAGAACGCG TAATACGACTCA -3′; C2: 5′- CGTTAGAACGCGTAATACGACTCACTATAGG GAGA-3′). The sequences of the gene-specific antisense primers are shown in Table 2-2.

**Table 2-2. Gene-specific antisense primers used for promoter cloning**

| Gene | Nucleotide sequences of primers |
|---|---|
| *fabp7a* | Outer: 5'-CTCGTCGAAGTTCTGGCTGTC-3'<br>Inner: 5'-GATGATGAAACACACAGTGGTC-3' |
| *fabp3* | Outer: 5′-TGCTCTCCTTCAAGTTCCACG-3′<br>Inner: 5′-AATGAGAGCGAGAGCAGATGG-3′ |

First round PCR was performed using primer C1 and an external gene-specific antisense outer primer that would anneal to a sequence within the first exon of the zebrafish *fabp7a* or *fabp3* gene. The 50 μL reaction contained 1× PCR buffer, 1.25 U of *Taq* DNA polymerase (IBM Fermentas), 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.2 μM of each primer and 1 μL of linker-ligated genomic DNA. Following an initial denaturation step at 94°C for 2 min, the reaction was subjected to 35 cycles of amplification at 94°C for 30 s, 55°C for 40 s, 72°C for 2.5 min, and a final extension for 5 min. One μL of the primary PCR product was used as template for a second round of PCR (nested PCR) with primers C2 and an inner gene-specific antisense primer. The conditions for the secondary PCR were similar to those of the primary PCR with the following modifications: 94°C for 1 min, 24 cycles of amplification at 94°C for30 s, 57°C

for 40 s, 72°C for 2.5 min. The products from the secondary PCR was fractionated by 1% agarose gel-electrophoresis and purified using Qiaquick gel extract kit (Qiagen). The purified DNA fragments were cloned into the plasmid, pGEM-T (Promega), and a single clone for *fabp3* and *fabp7* were sequenced in its entirety from both directions. To identify potential *cis*-acting regulatory elements in the *fabp7a* and *fabp3* promoter sequences, Computer-assisted analysis using MatInspector professional [Quandt *et al.*, 1995] was performed.

**2.5 5′ RNA ligase mediated rapid amplification of cDNA ends (5′ RLM-RACE)**

5′ RLM-RACE was employed to determine the initiation sites for transcription of the zebrafish iLBP genes. Total RNA was extracted from adult zebrafish using Trizol (Gibco BRL). cDNA for 5′ RLM-RACE was prepared using the Ambion RLM-RACE kit following the supplier's instructions. Briefly, 10 μg of total RNA was treated with calf intestinal phosphatase (CIP) and divided into two aliquots. One aliquot was then treated with tobacco acid pyrophosphatase (TAP) to remove the 5′ $^{7}$methyl guanine cap of intact, mature mRNA molecules. RNA molecules that had 5′ phosphate groups including degraded or unprocessed mRNAs lacking a 5′ cap, structural RNAs and traces of contaminating genomic DNA, were dephosphorylated by CIP treatment and, therefore, unable to be ligated to the adapter primer sequence. The two preparations of RNA (plus and minus TAP treatment) were incubated with a 45 base RNA adapter (5′-GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUG AAA-3′) and T4 RNA ligase. A random-primed reverse transcription reaction was performed to synthesize cDNA. A nested PCR was performed to amplify the 5′ cDNA ends using two nested

sense primers corresponding to the RNA adapter sequence (outer: 5′-GCTGATGGCGA TGATGAACACTG-3′; inner: 5′-CGCGGATCCGAACACTGCGTTTGCTGGCTTT GATG-3′) and two nested gene-specific antisense primers (Table 2-3).

**Table 2-3. Gene-specific antisense primers used for 5' RLM-RACE**

| Gene | Nucleotide sequences of primers |
|---|---|
| *fabp7a* | Outer: 5′-CACCACCATCCATCATTGAC-3′<br>Inner: 5′-CTCGTCGAAGTTCTGGCTGTC-3′ |
| *fabp7b* | Outer: 5′-GAAAGAAGAACAGCGACACC-3′<br>Inner: 5′-ATGACCACCTTTTCTCCCTC-3′ |
| *Fabp3* | Outer: 5′-TTGATGAGAGCGGATTGAGG-3′<br>Inner: 5′-ATTGGCAACTTGACGCGTG-3′ |
| *rbp1a* | Outer: 5′-TCCTAAAGTGAAGTCCATGTCG-3′<br>Inner: 5′- ACGCCAGTGTAATTAGGTTTG-3′ |
| *rbp1b* | Outer: 5′-AACTCCTGACCAACTACAAAG-3′<br>Inner: 5′-CGTAGTTCTTAAATGTGCTGAC-3′ |
| *rbp2a* | Outer: 5′-GCCTGTTTGGAGCAATAG-3′<br>Inner: 5′-AAACTCCTCTCCAATGACG-3′ |

The ten μL reaction of the first round of PCR contained 1× PCR buffer, 0.75 U of *Taq* DNA polymerase (IBM Fermentas), 1.5 mM $MgCl_2$, 0.25 mM of each dNTP, 0.5 μM of each outer primer and 0.5 μL of cDNA from the reverse transcription reaction. The PCR conditions were 94°C for 1 min followed by 30 cycles of 94°C for 30 s, 57°C for 30 s, 72°C for 40 s, and a final extension at 72°C for 10 min. 0.5 μL of the primary PCR product from the TAP+ and TAP- reactions were used as template for the secondary PCR containing 1× PCR buffer, 1 U of *Taq* DNA polymerase (IBM Fermentas), 1.5 mM $MgCl_2$, 0.25 mM of each dNTP and 0.25 μM of each inner primer. The thermal cycle conditions were the same as the primary PCR except that the annealing temperature was

increased to 60°C and the number of cycles were increased to 35. The PCR products in the TAP+ reaction were size-fractionated by agarose gel-electrophoresis and purified using the Qiaquick gel extraction kit (Qiagen), cloned and sequenced. The transcription start site of each gene was mapped by aligning the 5′ RLM-RACE sequence with the corresponding gene sequence.

### 2.6 Linkage analysis by radiation hybrid mapping

Radiation hybrids of the LN54 panel [Hukriede *et al.*, 1999] were used to map each iLBP gene to a specific zebrafish linkage group by PCR. DNA (100 ng) from each of the 93 mouse-zebrafish cell hybrids was amplified using a pair of zebrafish gene-specific primers (Table 2-4). The PCR reactions contained 1x PCR buffer (MBI Fermentas), 1.5mM $MgCl_2$, 0.25μM each forward and reverse primer, 0.2 mM each dNTP, 1U of *Taq* DNA polymerase. The PCR templates for the three controls were 100ng of DNA from zebrafish (cell line AB9), mouse (cell line B78) and 1:10 mixture of zebrafish/mouse DNA (AB9/B78), respectively. Following an initial denaturation at 94 °C for 4 min, the PCR was subjected to 32 cycles of amplification at 94 °C for 30s, 55 °C for 30s, 72 °C for 30s and a final extension at 72 °C for 7 min. Fifteen μL of the PCR reaction was fractionated by gel-electrophoresis in 2 % agarose. The radiation hybrid panel was scored based on the absence (0) or presence (1) of the DNA fragment of expected size, or an ambiguous result (2) to generate the RH vector and analyzed according to the directions at http://mgchd1.nichd.nih.gov:8000/zfrh/beta.cgi [Hukriede *et al.*, 1999]. The *fabp7b* gene was mapped by Dr. Zhou Yi using the T51 radiation hybrid panel.

**Table 2-4. Gene-specific primers used for radiation hybrid mapping**

| Gene | Mapping panel | Nucleotide sequence of primers |
|---|---|---|
| *fabp7a* | LN54 | Forward: 5′-TGCGCACATACGAGAAGGC-3′<br>Reverse: 5′-CACCACCATCCATCATTGAC-3′ |
| *fabp7b* | T51 | Forward: 5′-CCTATTTTACACTGCATTCCC-3′<br>Reverse: 5′-GAAAGAAGAACAGCGACACC-3′ |
| *fabp3* | LN54 | Forward: 5′-ACTTGGCGACATCGTCTCC-3′<br>Reverse: 5′-TCTGGAGGTTTGGAAGTTGG 3′ |
| *rbp1a* | LN54 | Forward: 5′-ATGCACTGAGGAAGGTAGTCTG-3′<br>Reverse: 5′- TCCTAAAGTGAAGTCCATGTCG-3′ |
| *rbp1b* | LN54 | Forward: 5′-GTGGGCAAACAGATTTTCAAG-3′<br>Reverse: 5′-AACATCACATCATCTCTCACC-3′ |

## 2.7 RT-PCR assay of iLBP mRNA expression

Reverse transcription-polymerase chain reaction (RT-PCR) was used to determine the tissue distribution of mRNA for each gene in adult zebrafish. Total RNA was extracted from adult zebrafish tissues and embryos at various stages of development using Trizol reagent and the protocol recommended by the supplier (GibcoBRL). One μg of total RNA from each sample was used as template for the synthesis of first strand cDNA by reverse transcriptase (SuperScript II). For PCR amplification, oligonucleotide primers were synthesized based on the cDNA sequence of each gene (Table 2-5). PCR reactions contained 1x PCR buffer, 1.25 U of *Taq* DNA polymerase, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.4 μM of each primer, and 1 μL from the reverse transcription reaction. Following an initial denaturation step at 94°C for 2 min, the reaction was subjected to 30 cycles of amplification at 94°C for 30 s, 57°C for 30 s, 72°C for 1 min, and a final extension at 72 °C for 5 min. Fifteen μL of each PCR reaction was size-

fractionated by 1% agarose gel-electrophoresis. The gel was stained with ethidium bromide and photographed under UV light. As a positive control in RT-PCR experiments, the constitutively expressed mRNA for Receptor for Activated C Kinase (RACK) [Hamilton and Wright, 1999] was RT-PCR amplified in tandem with experimental samples from all RNA samples assayed using forward (5′-ATCCAA CTCCATCCACCTTC-3′) and reverse (5′-ATCAGGTTGTC AGTGTAGCC-3′) primers. The RT-PCR conditions employed for detection of RACK mRNA were the same as RT-PCR of each iLBP mRNA (see above). As a negative control, reactions contained all RT-PCR components and gene-specific primers, but lacked the cDNA templates.

**Table 2-5. Gene-specific primers for RT-PCR**

| Gene | Nucleotide sequences of primers |
|---|---|
| *fabp7a* | Forward: 5′-TTGACAGCCAGAACTTCGAC-3′<br>Reverse: 5′-CACCACCATCCATCATTGAC-3′ |
| *fabp7b* | Forward: 5′-GGTCAGCAGTGACAACTTTG-3′<br>Reverse: 5′-GGTGTCGCTGTTCTTCTTTC-3′ |
| *fabp3* | Forward: 5′-TCAGCTCAAACATGGCAGAC-3′<br>Reverse: 5′-TTGATGAGGACGGATTGAGG -3′ |
| *rbp1a* | Forward: 5′-ATGCACTGAGGAAGGTAGTCTG-3′<br>Reverse: 5′-ACTGCAAATTAATCACGACTTG-3′ |
| *rbp1b* | Forward: 5′-GTGGGCAAACAGATTTTCAAG-3′<br>Reverse: 5′-AACATCACATCATCTCTCACC-3′ |
| *rbp2a* | Forward: 5′-CCAGCACATCCAGCTTC-3′<br>Reverse: 5′- GCCTGTTTGGAGCAATAG -3′ |

## 2.8 Quantitative PCR

Quantitative PCR for *fabp7a* and β-actin cDNA was performed using the LightCycler thermal cycler system (Roche Diagnostics) according to the manufacturer's instructions. The *fabp7a*-specific primers used for qualitative PCR were also used for quantitative PCR. β-actin cDNA was amplified using forward (5′- AAGCAGGA GTACGATGAGTCTG-3′, GenBank Accession number NM_131031) and reverse (5′- GGTAAACGCTTCTGGAATGAC 3′, GenBank Accession number NM_131031). Serial dilutions of bacteriophage lambda DNA and gel-purified *fabp7a* and β-actin RT-PCR products were allowed to bind SYBR® Green dye and the amount of bound SYBR® Green I was determined by fluorimetry. The concentration of *fabp7a* and β-actin RT-PCR gel-purified products were determined by extrapolation from the standard curve of concentration-dependent bacteriophage lambda DNA fluorescence and the copy number per µl was calculated. Five dilutions of the *fabp7a* and β-actin product ranging from 8 x $10^5$ to 8 x $10^1$ copies per reaction were used in individual quantitative PCR reactions to determine the standard curve of the crossing points for the amplification of *fabp7a* and β-actin from tissue-specific cDNA samples. Melting curve analysis of each standard and experimental sample following PCR demonstrated that only one product was generated in these reactions (data not shown). The ratio of *fabp7a* / β-actin PCR product for each experimental sample was calculated. The PCR reaction to amplify *fabp7a* contained 1 µL of cDNA, 0.2 µM sense and antisense primers, 3 mM $MgCl_2$, and 1 x LightCycler-DNA FastStart SYBR® Green I Mix containing nucleotides, buffer, and hot start *Taq* DNA polymerase. The PCR conditions for β-actin differed from those used for the B-FABP cDNA in that 0.25 µM sense and antisense primers and 5 mM $MgCl_2$ were used.

Multiple cDNA samples were simultaneously analyzed in parallel reactions. The PCR conditions were as follows: 15 min at 95°C to activate *Taq* DNA polymerase, 45 cycles of denaturation (15 s at 95°C), annealing (5 s at 54°C), and enzymatic chain extension (10 s at 72°C). Fluorescent signal was measured at the end of each extension phase. Melting curve analysis of the PCR products was performed after the 45 cycles by continuously measuring the total fluorescent signal in each PCR reaction while slowly heating the samples from 65°C to 95°C. For negative controls, cDNA was omitted.

**2.9 Northern blot and hybridization**

Fifteen μg of total RNA from adult zebrafish was size-fractionated by 2% (w/v) agarose gel-electrophoresis in a MOPS buffer (40 mM 3-[N-morpholino]propanesulfonic acid, 10 mM sodium acetate, 1 mM EDTA, pH 7.2) and 0.2M formaldehyde. The resolved RNA was transferred to Hybond-N+ membrane (Amersham Pharmacia Biotech, Baie d'Urfé, Quebec, Canada) according to standard methods [Sambrook *et al.*, 1989]. The *fabp3* cDNA generated by RT-PCR from total RNA of adult zebrafish was labeled with [$\alpha^{32}$P]-dATP during the reaction and used as a hybridization probe. The membrane was pre-hybridized in 5 mL of ExpressHyb™ (BD Biosciences Clontech, Franklin Lakes, New Jersey, USA) solution at 68°C for 30 min, and then hybridized to the denatured probe in the same solution at 68°C for 1 h. The blot was rinsed three times at room temperature with 2×SSC and 0.05% SDS (w/v), washed once at room temperature for 20 min with the same solution, and then washed at 50°C for 1 h in 0.1×SSC/0.1% (w/v) SDS. The membrane was then exposed to X-ray film at –70°C for 48 h.

## 2.10 *In situ* hybridization to tissue sections and emulsion autoradiography

Gene-specific antisense oligonucleotides corresponding to the cDNA sequence of each gene (Table 2-6) were synthesized and used as probes in tissue section *in situ* hybridization and emulsion autoradiography to localize the mRNA of each gene at the tissue and cellular level in adult zebrafish. *In situ* hybridization was performed as described by Denovan-Wright *et al.* [1998]. Briefly, 20 μm cryostat sections obtained from fresh-frozen adult zebrafish were mounted onto Superfrost Plus glass slides (Fisher Scientific, Nepean, Ontario, Canada), fixed in 4% paraformaldehyde for five min, rinsed in 1×PBS twice for three min, and then equilibrated in 1×SSC for 20 min. The fixed tissue sections were immersed in a hybridization buffer containing 50% formamide, 5×SSC, 1× Denhardt's solution, 20 mM sodium phosphate (pH 6.8), 0.2% w/v SDS, 5mM EDTA, 10 μg/mL poly $(A)_n$, 10% dextran sulfate, and $5\times10^6$ cpm/mL of [$\alpha^{33}$P]-dATP 5' end-labelled oligonucleotide probe. Hybridization was performed at 42°C for 16 h. The slides were washed sequentially in 1×SSC (four times for 15 min), 0.5×SSC (four times for 15 min) and 0.25×SSC (twice for 15 min) at 55°C, and then once in 0.25× SSC for 60 min at room temperature. The sections were exposed to Kodak Biomax single emulsion film at room temperature for five days, and then dipped in Kodak NTB2 nuclear track emulsion and exposed for 14 days at 4°C. The tissue sections were then developed, counter-stained with cresyl violet and viewed under bright-field and dark-field illumination [Denovan-Wright *et al.*, 1998]. Hybridization using a sense oligonucleotide probe corresponding to the specific iLBP cDNA sequence (Table 2-6) was employed as a negative control.

**Table 2-6. Gene-specific oligonucleotide probes used for tissue section *in situ* hybridization**

| Gene | Sequence of oligonucleotide probes |
|---|---|
| *fabp3* | Antisense: 5′-TACATAACCATTGATGAGGACGGATTGAGGC-3′<br>Sense: 5′-CACTACGTGAAAGCAGAATAAAGGGAAAGC-3′ |
| *fabp7b* | Antisense: 5′-GAAAGAAGAACAGCGACACCTGATAGTATTCCC-3′ |
| *rbp1a* | Antisense: 5′-CAGGCAGCTCTATCTCATTGTATGATGGTT-3′ |
| *rbp1b* | Antisense: 5′-TTGTTTGTTTCTTCTCCAGCAGAGGGCACCAC-3′ |

## 2.11 Whole mount *in situ* hybridization to embryos

Whole mount *in situ* hybridization to zebrafish embryos was performed to reveal the spatio-temporal expression patterns of each gene during zebrafish development. The experiment was done by Christine Thisse, Bernard Thisse and Agnes Degrave in the Institut de Génétique et Biologie Moléculaire et Cellulaire, University of Strasbourg, France. Antisense RNA probes for whole mount *in situ* hybridization were synthesized from the corresponding cDNA clone for each gene (Table 2-7).

**Table 2-7. cDNA templates for antisense RNA probe synthesis**

| Gene | Clone ID | GenBank accession number | Vector |
|---|---|---|---|
| *fabp7a* | fb62f07.y1 | AI497146 | pSPORT1 |
| *fabp7b* | - | AY380814 | pGEM®-T |
| *fabp3* | - | AF448057 | pGEM®-T |
| *rbp1a* | STR00496 | BU670746 | pSPORT1 |
| *rbp1b* | - | AY395732 | pGEM®-T |
| *rbp2a* | fb69e02.y1 | AI544932 | pSPORT1 |

# 3 Structure, Linkage Mapping and Expression of the Heart-Type Fatty Acid-Binding Protein Gene *(fabp3)* from Zebrafish (*Danio rerio*)

## 3.1 INTRODUCTION

Intracellular lipid-binding proteins (iLBPs) are low molecular mass (~15 kDa) polypeptides encoded by a multigene family [reviewed by Stewart, 2000; Hertzel and Bernlohr, 2000; Schaap *et al.*, 2002]. Sixteen members of this multigene family have been identified in mammals and named according to the initial site of isolation, e.g., adipocyte fatty acid-binding protein (A-FABP), brain (B-FABP), epidermal (E-FABP), heart (H-FABP), intestinal (I-FABP), liver (L-FABP), *etc.* Nomenclature based on the initial site of isolation or patterns of tissue-specific expression has given rise to multiple names for the same proteins, which is, on occasion, confusing. For instance, the H-FABP was named mammary-derived growth inhibitor (MDGI) and muscle-type FABP (M-FABP) due to its presence in mammary gland and skeletal muscle [Treuner *et al.*, 1994; Prinsen and Veerkamp, 1996]. In addition, tissue-distribution or function, or both, of proteins encoded by orthologous genes from different species may have different physiological functions. Hertzel and Bernlohr [2000] have suggested, therefore, an alternative nomenclature for iLBPs , which I have followed in this report.

Lipid storage and utilization differ in various taxa. For example, mammals store lipids subcutaneously and in adipose tissue, whereas fish deposit and store lipids in several tissues including mesenteric fat, liver, dark muscle [Vayda *et al.*, 1998] and in oocytes [Selman *et al.*, 1993 and references therein]. Tissue-specific patterns of iLBP gene expression in non-mammalian species may, therefore, differ from that observed in

mammals. Few studies have focused on the tissue-specific patterns of iLBP expression in fishes, the largest and most evolutionary diverse group of vertebrates. FABPs have been detected in the white heart muscle of ocean pout and sea raven, the liver of nurse shark, elephant fish, lamprey and catfish, and the aerobic muscle of striped bass. More recently, cDNAs have been characterized for a H-FABP from rainbow trout and two distinct isoforms of H-FABP from the ventricle of four Antarctic fishes [reviewed by Stewart, 2000]. Recently, the nucleotide sequence of cDNAs encoding three FABPs (I-, B-, and Lb-type FABP) and a related CRBPII, and their tissue-specific expression in adult zebrafish have been reported [Denovan-Wright *et al.*, 2000; Pierce *et al.*, 2000; Denovan-Wright *et al.*, 2000; Cameron *et al.*, 2002]. Here I describe a H-FABP (*fabp3*) gene from zebrafish. While high amino acid sequence similarity, identical gene structure and coding capacity, and conserved syntenic relationship to the mammalian H-FABP suggest that it is an orthologous H-FABP gene (or *fabp3*), expression patterns of the zebrafish and mammalian *fabp3* genes, as well as their 5′ *cis* regulatory elements were strikingly different. This zebrafish FABP mRNA was detected in abundance in primary oocytes and in the liver by both tissue section *in situ* hybridization and RT-PCR, but the mRNA was detected in the heart and other tissues of adult zebrafish only by RT-PCR. Messenger RNA levels for the zebrafish *fabp3* decreased during the first phase of oogenesis, the growth phase, to the point that it was undetectable in mature oocytes. I propose that differences in tissue expression patterns of the *fabp3* genes from mammals and zebrafish are the result of evolutional divergence of transcriptional regulation.

## 3.2 RESULTS

### 3.2.1 Nucleotide sequence of cDNAs coding for the zebrafish FABP3

Three cDNA clones derived from the major product and one clone from the minor product of 3′ RACE (Fig. 3-1) for the zebrafish FABP3 were 636 bp and 528 bp in length, respectively, not including the poly(A) tail. Both sequences contain the same open reading frame for a polypeptide of 133 amino acids (Fig. 3-2). The sequence also contains a 11 bp 5′ untranslated region, a 223 bp (major transcript) or 115 bp (minor transcript) of the 3′ untranslated region with an alternative polyadenylation signal of AATAAA at nucleotides 616-621(major transcript) or 509-514 (minor transcript) of the cDNA sequence (Fig. 3-2). A single nucleotide difference, a C to a T transition in one of the 3′RACE cDNA clones was seen at position 104 in the coding region. This nucleotide difference, however, did not change the encoded amino acid. The nucleotide difference may be due to either an artefact during RT-PCR cloning of the FABP3 cDNA, or, more likely, the nucleotide difference represents an allelic variant of the zebrafish *fabp3* gene. The latter conclusion is supported by the finding that independently cloned EST sequences (GenBank accession numbers AI617818, AW077983, AW281103, AW778251, BI672083, BI673099, BM082353, BQ481071, BQ481317) contain C rather than T at this position, while other ESTs (accession numbers BQ480714, BM186245, BM005090, BM025130, BI867082, BM186677, BQ260001, BI868173) contain T rather than C.

The deduced amino acid sequence of the zebrafish FABP3 was aligned with FABP sequences from zebrafish and six other species using CLUSTALW [21] (Fig. 3-3). The zebrafish FABP3 exhibits highest sequence identity with rainbow trout H-FABP

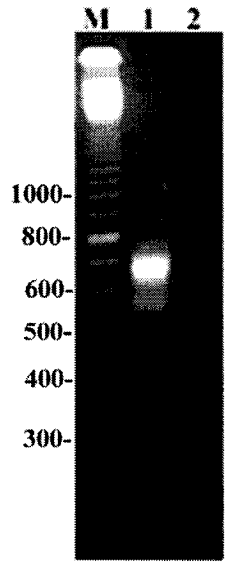



**Figure 3-1. Cloning of the zebrafish FABP3 cDNA by 3′ RACE.** Agarose gel-electrophoretic separation of 3′ RACE products for zebrafish FABP3 cDNA cloning. Both the major band of ~650 bp and the minor band of ~570 bp were excised, cloned and sequenced. M: 100 bp DNA ladder (with molecular sizes in bp shown on the left of the panel). Lane 1: 3′ RACE product of FABP3 cDNA; Lane 2: Negative control without template.

```
                                                          5′ tcagctcaaac  11

ATG GCA GAC GCT TTT ATC GGC ACG TGG AAC TTG AAG GAG AGC AAG AAT TTC  62
Met Ala Asp Ala Phe Ile Gly Thr Trp Asn Leu Lys Glu Ser Lys Asn Phe

      RsaI                                                C
GAT GAG TAC ATG AAA GGA ATT GGT GTT GGC TTT GCC ACG CGT CAA GTT GCC  113
Asp Glu Tyr Met Lys Gly Ile Gly Val Gly Phe Ala Thr Arg Gln Val Ala

AAT ATG ACC AAA CCC ACA ACC ATC ATC TCC AAG GAG GGC GAC GTT TTC ACA  164
Asn Met Thr Lys Pro Thr Thr Ile Ile Ser Lys Glu Gly Asp Val Phe Thr

CTT AAA ACC GTC AGC ACT TTC AAA TCC ACA GAA ATC AAT TTC AAA CTG GGA  215
Leu Lys Thr Val Ser Thr Phe Lys Ser Thr Glu Ile Asn Phe Lys Leu Gly

GAG GAG TTC GAT GAG ACC ACT GCA GAT GAC CGA AAA GTC AAG TCT GTT ATA  266
Glu Glu Phe Asp Glu Thr Thr Ala Asp Asp Arg Lys Val Lys Ser Val Ile

ACT TTA GAT GGA GGG AAG TTG TTG CAT GTT CAG AAA TGG GAC GGT AAA GAG  317
Thr Leu Asp Gly Gly Lys Leu Leu His Val Gln Lys Trp Asp Gly Lys Glu

                                                          HincII
ACG ACC CTG CTG CGA GAA GTC AGT GAC AAT AAC CTC ACT CTG ACG TTG ACA  368
Thr Thr Leu Leu Arg Glu Val Ser Asp Asn Asn Leu Thr Leu Thr Leu Thr

CTT GGC GAC ATC GTC TCC ACA CGA CAC TAC GTG AAA GCA GAA TAA agggaaa  420
Leu Gly Asp Ile Val Ser Thr Arg His Tyr Val Lys Ala Glu  #

                                    RsaI
Gctga[gcctcaatccgtcctcatcaatggttatgtacatgtttttctgtggtgttcatgctctcctt  487
                                                          RsaI
Tgcaccgtttcccaaacactgaataaattccaacttccaaacctccagatttctgtatatgatcttg  554
          MobI                          *
Tatttctgggtgtacgatatgcccggcggtcagtgttattataaatacgcatttggtctacaataaa  621

gatgacttttgctcc(A)n                                                   636
```

**Figure 3-2. Nucleotide and deduced amino acid sequences of the zebrafish FABP3 cDNA.** The coding nucleotides are shown in uppercase letters and the 5′ and 3′ UTRs are in lowercase letters. Two polyadenylation signals are double underlined and an alternative polyadenylation site is marked with a star. Underlined sequences correspond to primers used in PCR and RT-PCR experiments (see methods). A single nucleotide difference in one of the cDNA clones is shown above the nucleotide sequence at position 104. The nucleotide sequence complementary to the antisense probe used for tissue section *in situ* hybridization and emulsion autoradiography is boxed. Diagnostic restriction endonuclease sites used in Southern blot analysis are indicated above the recognition nucleotide sequences and in bold type. The GenBank accession number of the zebrafish FABP3 cDNA is AF448057.

```
ZF FABP3     MADAFIGTWNLKESKNFDEYMKGIGVGFATRQVANMTKPTTIISKEGDV
TR H-FABP    ..E..A......D.........AL.........GG........EVA..T
RT H-FABP    .....V...K.VD.....D...SL.........S........E.N..T
HM H-FABP    .....V...K.VD.....D...SL.........S........E.N..T
MS H-FABP    .....V...K.VD.....D...SL.........GS.......E.N..T
PG H-FABP    .V...A...K.VD.....D...S.....................EVN..T
CW H-FABP    .V...V...K.VD.....D...SL.........G.........EVN..T
ZF B-FABP    .V...CA..K.VD.Q.......SL.........G.V....IV..H...K
ZF I-FABP    .T--.N...KVDRNE.YEKF.EQM..NMVK.KL.AHDNLKITLEQT..K
ZF Lb-FABP   ..--.S...QVYAQE.YE.FLRA.SLPEEVIKL.KDV..V.E.QQN.SD

ZF FABP3     FTLKTVSTFKSTEINFKLGEEFDETTADDRKVKSVITLDGG--KLLHVQK
TR H-FABP    V.........N...S....A.............L..I...--.MV....
RT H-FABP    I.I..H....N...S.Q..V....V..........V.....--..V....
HM H-FABP    I.I.......N.....Q..I....V.........LV.....--..I....
MS H-FABP    I.I.......N.....Q..I....V.........LV.....--..I....
PG H-FABP    III...........S....V.............IV.....--..V.L..
CW H-FABP    VII.......N...S....V.............IV.....--..V....
ZF B-FABP    VVI..L....N...S..............H...TVS.E.D--N.VQ..R
ZF I-FABP    .NV.EV...RTL....T..VT..YSL..GTELTGSWVIE.DTL.GTFTR.
ZF Lb-FABP   ..ITSKTPG.TVTNS.T..K.AEI..M.GK.L.CIVK....--..VCR--

ZF FABP3     WDGKETTLLREVSDNNLTLTLTLGDIVSTRHYVKAE      100.00%
TR H-FABP    ........V....G.A.E.......V....S.....       80.45%
RT H-FABP    ...Q....T..L..GK.I....H.NV....T.E.EA       74.44%
HM H-FABP    ...Q....T..LV.GK.I....H.SV....T.E.EA       73.68%
MS H-FABP    ...Q....T..LV.GK.I....H.SV....T.E.EA       72.93%
PG H-FABP    .N.Q....V..LV.GK.I....H.SA.C..T.E.EA       72.18%
CW H-FABP    .N.Q..S.V..MV.GK.I....H.TA.C..T.E.QA       69.93%
ZF B-FABP    ......KFV..IK.GKMVM...FEGVQAV.T.E..-       62.12%
ZF I-FABP    DN..VL.TV.TIVNGE.VQSYSYDGVEAK.IFKR.-       28.03%
ZF Lb-FABP   --TDRFSHIQ.IKAGEMVE...V.GTTMI.KSK.I-       27.78%
```

**Figure 3-3. Comparison of the zebrafish FABP3 to H-FABPs from six different species and paralogs of other zebrafish FABPs.** The deduced amino acid sequence of the zebrafish *fabp3* (ZF FABP3; Swiss-Prot and TrEMBL accession number: Q8UVG7) was compared to the sequences of H-FABPs from rainbow trout (TR H-FABP; O13008), rat (RT H-FABP; P07483), human (HM H-FABP; P05413), mouse (MS H-FABP; P11404), pig (PG H-FABP; O02772), cow (CW H-FABP; P10790) and the zebrafish FABP paralogs, brain-type (ZF B-FABP; Q9I8N9), intestinal-type (ZF I-FABP; Q9PRH9) and the basic liver-type (ZF Lb-FABP; Q9I8L5). Dots indicate amino acid identity. Dashes represent gaps in the alignment. The percentage amino acid sequence similarities between the zebrafish FABP3 and other FABPs are shown at the end of the sequences.

(80%) and mammalian H-FABPs (72-74%). The zebrafish FABP3, like mammalian H-FABPs, had high sequence identity to B-FABPs from zebrafish (Fig. 3-3) and mammals (data not shown). Amino acid identity between the zebrafish FABP3 and two paralogous zebrafish FABPs, I-FABP and liver basic-type FABP (Lb-FABP), was 28% and 26%, respectively.

### 3.2.2 Southern analysis of the zebrafish *fabp3* gene

Using a pair of primers flanking the entire coding region and a portion of the 3′ UTR for the zebrafish FABP3 cDNA sequence (see Fig. 3-1), a radiolabeled hybridization probe from adult zebrafish total RNA was generated by RT-PCR for Southern blot and hybridization analysis. The cDNA probe hybridized to restriction fragments of zebrafish genomic DNA of 4.7 kb and 7.2 kb in *Hinc*II-digested DNA, 2.6 kb and 5.1 kb in *Hae*III-digested DNA, 2.4 kb and 5.0 kb in *Mbo*I-digested DNA, and 1.7 kb and 5.0 kb in *Rsa*I-digested DNA (Fig.3-4). Of the four restriction endonucleases used in the Southern blot, *Rsa*I and *Hinc*II have two and one recognition sites, respectively, while *Hae*III and *Mbo*I have no recognition sites within the FABP3 cDNA sequence. The most parsimonious explanation for the simple hybridization seen in the Southern blot is that the *fabp3* gene exists as a single copy in the zebrafish genome. I was unable to detect the predicted 0.1 kb *Rsa*I restriction fragment in the Southern blot, presumably due to the low hybridization signal from such a small DNA fragment or it may have migrated off the end of the agarose gel during electrophoresis. Since both the *Hae*III and *Mbo*I-digested DNA samples generated two fragments in the Southern blot-hybridization, and neither site is present in the cDNA sequence, I predict that single recognition site for each

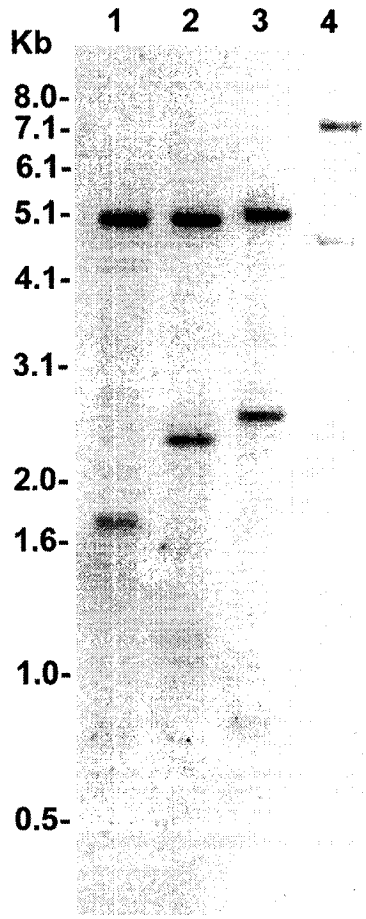



**Figure 3-4. Southern blot analysis of zebrafish genomic DNA using the *fabp3* cDNA as a hybridization probe.** Eight μg aliquots of zebrafish genomic DNA were digested with one of the following restriction endonucleases: *Rsa*I (lane 1), *Mbo*I (lane 2), *Hae*III (lane 3) and *Hinc*II (lane 4). The size-fractionated DNA was transferred to nylon membranes and hybridized to *fabp3* cDNA. Molecular weight markers in kb are shown on the left of the panel.

of these restriction endonucleases is present in one of the introns of the zebrafish *fabp3* gene.

### 3.2.3 DNA sequence and structure of the zebrafish *fabp3* gene

Using the zebrafish FABP3 cDNA sequence to search the zebrafish genome sequence database of the Wellcome Trust Sanger Institute, Four DNA traces containing the coding sequence of exon 1 (zfishG-a1962b06.q1c), exon 2 (Z35725-a5890c08.q1c), exon 3 (Z35724-a1164g06.q1c) and exon 4 (Z35725-a1576b09.p1c) were retrieved. The zebrafish FABP3 cDNA sequence was identical to the DNA trace sequences, except for a T to C substitution in the sequence of DNA trace Z35724-a1164g06.q1c in exon 3, which corresponds to the sequence of some FABP3 ESTs reported in GenBank. The zebrafish *fabp3* gene consists of four exons encoding 24, 58, 34 and 17 amino acids, respectively, interrupted by three introns (Fig. 3-5A). All exon/intron splice junctions of the zebrafish *fabp3* gene conform to the GT-AG rule [Breathnach and Chambon, 1981]. Comparison of the structure of the zebrafish *fabp3* gene with that of the orthologous human and mouse genes revealed an identical exon/intron organization, conserved splice junction sequence, and coding capacity for each exon (Fig. 3-5B).

### 3.2.4 Multiple transcription start sites for the zebrafish *fabp3* gene

Using 5′ RLM-RACE, I determined the 5' end of the capped and complete zebrafish *fabp3* gene transcripts. A major band of approximately 200 bp, and two minor bands of 250 bp and 180 bp, were observed after agarose gel-electrophoresis of nested PCR products from the CIP/TAP treated RNA. No specific band was detected for the negative control using a RNA sample that was not treated with TAP (Fig. 3-6). Thus,

**Figure 3-5. Nucleotide sequence of the 5′ upstream region of zebrafish *fabp3* gene and comparison of the structure of zebrafish *fabp3* gene with its mammalian orthologs.** (**A**) The coding sequence of the first exon is shown in uppercase letters and underlined and the deduced amino acid sequence indicated below. 5′ flanking sequence and the 5′ UTR of the first exon are shown in lowercase letters. The multiple transcription start sites, mapped by 5′ RLM-RACE, are marked by stars and the major transcription start site is numbered as +1. The core sequence of a potential TATA box upstream of the transcription initiation sites and two GC boxes are boxed. The external and internal antisense primer sequences used for promoter cloning are in bold. GenBank accession number: AY246558. (**B**) Exon/intron organization and nucleotide sequences of exon/intron junctions of the zebrafish *fabp3* gene are compared with the orthologous human and mouse *fabp3* genes. Exons are boxed and the numbers of amino acids of each exon are indicated. The initiation codon in exon 1 and stop codon in exon 4 are also indicated. The nucleotide sequence of the 5′ splice donor and 3′ splice acceptor for each exon/intron junction are shown in uppercase and lowercase letters, respectively. The intron residues (gt/ag) adjoining the splice junctions are in bold font. The mouse and human *fabp3* gene structures were identified based on human *fabp3* gene sequence (accession number: NT_034397) and the mouse *fabp3* gene sequences (accession number : NT_039265) obtained from GenBank .

**A**

```
                                                tctaattcagtcttggcaccaaaatcagct   -1191
tttaatcatgcactcatgcaaacatgctttccttttttgggtgtgtgcggcatgctaagtgtttcttaa   -1121
agtagtgttgagtgattggatgattgcacggggaaaatgcattgataataggaacattcttagataaact   -1051
aatgtaaaatttggtcttggcaccaaaatcagctttgatgttttgtgttttttaaagtagtgttcagcga   -981
ttggatgattgcatggagattgtataattgagtatttttaaatcatcactccttttagggtagtttttat   -911
ttgattcacatttattatcattaatttatgcgccacttggcatttctgtatggagtttgcatgttcgccc   -841
cgtgttggcatgggtttcctctgggtgttccctcacagtccgaacacatacactataggtgaataaacta   -771
aatggccatatgtgtgtgaatgagtgtgtatggatgtttcccagtacaaggttgtaccgggaagggcatc   -701
cgctgtgtgaaacatgtgctggataagtgggcggttcattccactgtggcgacccctgataaattaaggg   -631
actaagccatgtccatggggttgacgatcagtccatacagtatattgattatgtggcagacctacatata   -561
tcatataaaagtacttgtatagcctgtctcctgcgtataagtatcttatagtatacgttagtaacggaca   -491
gggtttttaatcgctcgctgatcggtgttatgtttgtatttaggattcatttgtgtagcacagttgtcc   -421
tgcgagacacaacatttcgttcctctgtatgttcacacatgtagtggaatgacaatgaagcttgacttga   -351
cttaatacttagataaaagaatgtaaaatttacatcagccgttaattagttttttaaattaaatgtatat   -281
tcaaattctaaacaaattctattataagtttttttgcattgacattgcacattcccttgacatcctctg   -211
tttattctattatatttaattcaatccccatctctctctctctctttctttctctctctctcccccccc   -141
tgctgtgtgtcatctcatagct ccgcccc cat ccgcccc cccaggcctacactgatgacatcacaaacag   -71
cagtta taaa tcagccagtaggatcgcagaagcagaagctctgtggtggtgctgtcttacctctctctct   -1
                *
ctttctcttgcgtccctccatctgctctcgctctcattctctgctgcgcttcagctcaaacATGGCAGAC   70
*                               *                                  M  A  D
GCTTTTATCGGCACGTGGAACTTGAAGGAGAGCA
 A  F  I  G  T  W  N  L  K  E  S
```

**B**

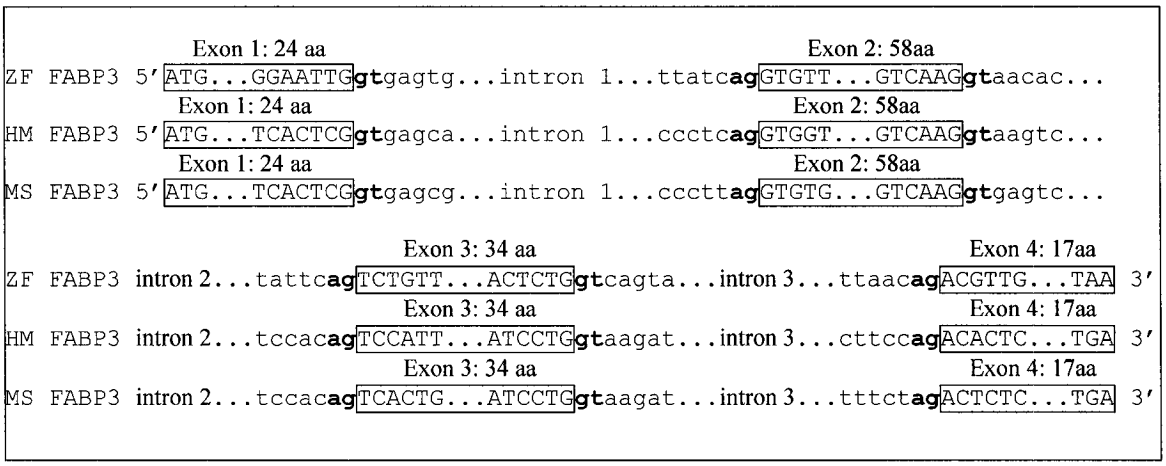



**Figure 3-5. Nucleotide sequence of the 5′ upstream region of zebrafish *fabp3* gene and comparison of the structure of zebrafish *fabp3* gene with its mammalian orthologs.**

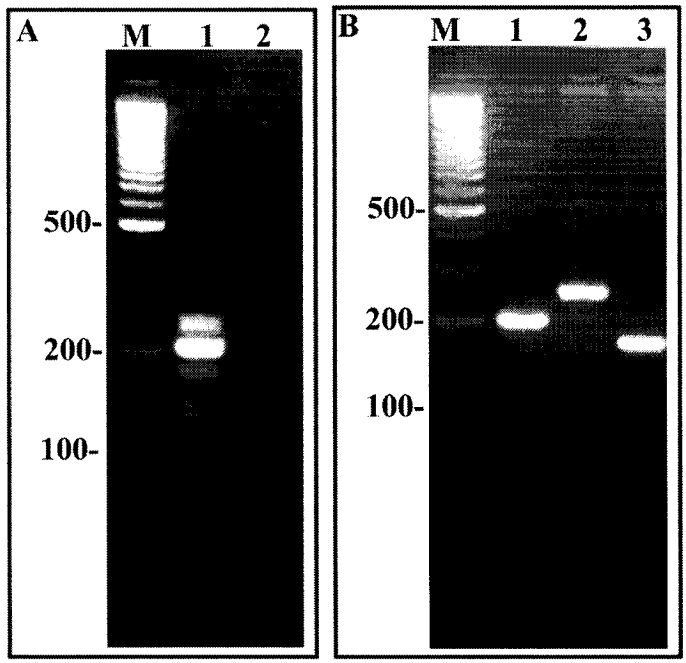



**Figure 3-6. Agarose gel electrophoresis of 5′ RLM-RACE product and PCR identification of the corresponding clones.** (A) Total RNA from whole adult zebrafish was sequentially treated with calf intestinal alkaline phosphatase (CIP), tobacco acid pyrophosphatase (TAP) and then ligated to a designated RNA adapter. Following two rounds of nested PCR, one major PCR-amplified product of approximately 200 bp and two minor products of 180bp and 250 bp were size-fractionated by gel-electrophoresis through 2.5% agarose (lane 1). RNA treated to the same experimental regime, but with TAP digestion omitted, did not generate a product (lane 2). A ladder of 100 bp molecular weight markers (MBI Fermentas) is shown in lane M with the 200 bp marker indicated to the left of the panel. (B) Positive colonies from transformants of the three 5′ RLM-RACE reactions were of 200 bp (lane 1), 250 bp (lane 2) and 180 bp (lane 3). The correct size of DNA insert was confirmed by colony PCR. M: 100 bp DNA ladder with molecular sizes shown on the left of the panel.

these RACE products most likely represent the 5′ ends of the mature *fabp3* gene transcripts. By aligning the sequence of the 5′ RLM-RACE products with the *fabp3* gene sequence, the transcription start sites of zebrafish *fabp3* were mapped to 29 bp, 61 bp and 116 bp upstream of the initiation codon (Fig. 3-5A). Multiple transcription start sites have been reported for mammalian *fabp3* genes [Treuner *et al.*, 1994; Qian *et al.*, 1999].

### 3.2.5 Assignment of the zebrafish *fabp3* gene to linkage group 19

A PCR product of predicted size (229 bp) was generated for positive mouse-zebrafish hybrid cell lines of the LN54 panel [Hukriede *et al.*, 1999] using primers specific to the fourth exon of the zebrafish *fabp3* gene (Fig. 3-7). A PCR product of the predicted size was also generated in the two positive controls, a 1:10 mixture of zebrafish and mouse genomic DNA, and zebrafish genomic DNA. No band was derived from the negative control containing mouse genomic DNA only. Online analysis of the radiation hybrid mapping data assigned the zebrafish *fabp3* gene to linkage group 19 at 365.69 cR (LN54 panel) or 67.51 cM (ZMAP panel) in the zebrafish genome with a LOD score of 14.6. Comparison of the syntenic relationship of *fabp3* gene on zebrafish linkage group 19 with that of the human and mouse orthologous gene revealed conserved syntenies (Table 3-1). Five mapped genes syntenic to the zebrafish *fabp3* gene on LG 19 are located on human chromosome 1, and four of them were syntenic to the mouse *fabp3* gene on mouse chromosome 4. Four other genes syntenic to the *fabp3* gene in zebrafish, however, reside on human chromosome 6 at 6p21.3 and mouse chromosome 17. The syntenic genes in the zebrafish LG 19 are distributed at two distinct chromosomal locations in the human and mouse genomes suggesting that this region has undergone inter-chromosomal rearrangements after the divergence of fish and mammals.

**Table 3-1. Conserved syntenic relationship of the zebrafish *fabp3* gene in human and mouse genome**

| Zebrafish[a] | | Human[b] | | Mouse[b] | |
|---|---|---|---|---|---|
| Gene symbol | location | Gene symbol | location | Gene symbol | location |
| fabp3 | 19 67.51 cM | FABP3 | 1p33-p32 | Fabp3 | 4 61.0 cM |
| oprd1 | 19 8.42 cM | OPRD1 | 1p36.1-p34.3 | Oprd1 | 4 64.8 cM |
| rpl11 | 19 70.21 cM | RPL11 | 1p36.1-35 | Rpl11 | 4 D3 |
| fuca1 | 19 48.02 cM | FUCA1 | 1p34 | Fuca1 | 4 65.7 cM |
| mycl1 | 19 50.99 cM (T51) | MYCL1 | 1p34.3 | Mycl1 | 4 65.7 cM |
| ifl2 | 19 56.41 cM (LN54) | IFL2 | 1q21.1 | Ifl2 | 3 F2 |
| bing1 | 19 38.6 cM | BING1 | 6p21.3 | Bing1 | 17 B1 |
| daxx | 19 38.6 cM | DAXX | 6p21.3 | Daxx | 17 B1/17.0 cM |
| psmb9 | 19 42.56 cM | PSMB9 | 6p21.3 | Psmb9 | 17 18.59 cM |
| rxre | 19 43.76 cM (LN54) | RXRB | 6p21.3 | Rxrb | 17 18.49 cM |

[a]ZMAP (http://zfin.org/cgi-bin/view_zmapplet.cgi );
[b]LocusLink(http://www.ncbi.nlm.nih.gov/LocusLink/list.cgi), NCBI.

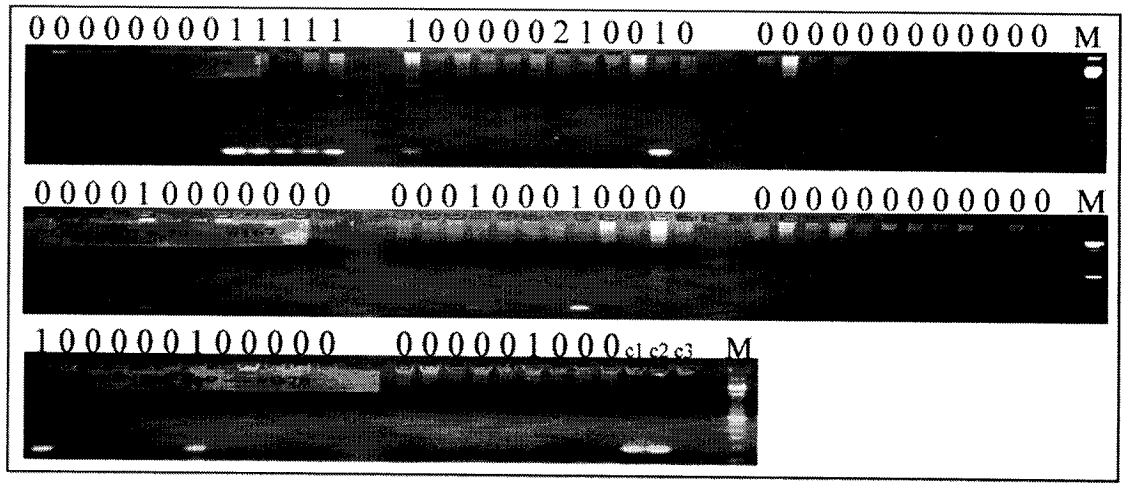



**Figure 3-7. Linkage group assignment by radiation hybrid mapping of the zebrafish *fabp3* gene.** DNA from 93 mouse-zebrafish radiation cell hybrids of the LN54 panel was subjected to PCR using *fabp3* gene-primers and the products subjected to size fractionation by gel-electrophoresis in 2% agarose (arrayed 1-93 from left to right, and top to bottom in the panels). The absence or presence of the expected 203 bp DNA fragment, or an ambiguous result, was scored 0, 1 or 2, respectively, to generate the RH vector (shown above each lane corresponding to the PCR reaction from a DNA of a given radiation cell hybrid) and submitted to http://mgchd1.nichd.nih.gov:8000/zfrh/beta.cgi for linkage group assignment. Genomic DNA from zebrafish (c1), a 1:10 mixture of zebrafish:mouse genomic DNAs (c2) and mouse genomic DNA (c3) were used as controls for the PCR. Molecular weight markers are shown at the right and left of each panel (M).

### 3.2.6 Tissue-specific expression of the *fabp3* gene in zebrafish

Hybridization of a [$\alpha^{32}$P]-dATP labelled zebrafish FABP3 cDNA probe to a northern blot containing total RNA from adult zebrafish revealed a broad RNA band of ~780 nucleotides (Fig. 3-8A). Detection of *fabp3* gene transcripts with two different 3′ polyadenylation sites and three different 5' transcription start sites for the zebrafish *fabp3* mRNA could give rise to six potential transcripts of 741 nt, 686 nt, 654 nt, 654 nt, 620 nt, 575 nt, 543 nt in length, not including the polyA tail. The relative abundance of each of these transcripts in different tissues has not been determined.

An oligonucleotide antisense probe was used to detect the distribution of the *fabp3* gene transcript in adult tissue sections by *in situ* hybridization. The zebrafish *fabp3* mRNA was exclusively localized to liver (Fig. 3-8B1) and ovary (Fig. 3-8C1). No hybridization to adult tissue sections was observed with the oligonucleotide "sense" probe. The signal emanating from a region of the eye and skin (Fig. 3-7B2 and III-7C2) has been observed in our lab with many "sense" probes indicating a non-specific interaction between the DNA probe and components of these tissues. It was concluded, therefore, that the *fabp3* gene transcript detected by *in situ* hybridization is localized to the liver and ovary of adult zebrafish.

In order to further resolve the tissue and cellular localization of the *fabp3* mRNA in adult zebrafish, emulsion autoradiography was performed on tissue sections. The hybridization signal, corresponding to the location of the zebrafish *fabp3* gene-transcript, was most intense and uniform in primary growth stage (stage I) oocytes, including both ooplasm and germinal vesicle (Fig. 3-9A; white granules in the emulsion). Stages of oocyte maturation in zebrafish are described in [Selman *et al.*, 1993]. The hybridization

signal was less intense and restricted to the ooplasm of cortical alveolus stage (stage II) oocytes. In stage III oocytes, hybridization to the *fabp3* mRNA was almost undetectable under the conditions employed here. No hybridization signal was seen in follicular cells surrounding the oocytes in the ovary. A moderate, but uniform hybridization signal was seen over the hepatocytes of the liver (Fig. 3-9B-D). No hybridization signal was evident in adult zebrafish heart, muscle (Fig. 3-9B and C) or brain (data not shown), tissues where mammalian and other fish H-FABPs are known to be expressed [reviewed in Veerkamp and Maatman, 1995; Vayda *et al.*, 1998]. The *in situ* hybridization and emulsion autoradiography were repeated three times using two different antisense oligonucleotide probes to the *fabp3* gene transcript with the same result (data not shown).

To verify the tissue-specific distribution of the zebrafish *fabp3* gene transcript, RT-PCR was performed on total RNA extracted from ovary, liver, heart, muscle, brain, intestine, skin and testis (Fig. 3-10). As a positive control, RT-PCR was employed to amplify the constitutively expressed zebrafish receptor for activated C kinase 1 (RACK1) [Hamilton and Wright, 1999]. RT-PCR product of the correct size for the *fabp3* mRNA was detected in all the tissues examined (Fig. 3-10), indicating a wide tissue-distribution for the zebrafish *fabp3* transcript similar to that reported for mRNA expression of the *fabp3* gene in mammals [Veerkamp and Maatman, 1995; Heuckeroth *et al.*, 1987]. Although the zebrafish *fabp3* mRNA was detected in a wide range of tissues by RT-PCR, the mRNA must be of such low abundance in most of these tissues, with the exception of primary oocytes and liver, that it is below the sensitivity of detection by the technique of tissue section *in situ* hybridization and emulsion autoradiography employed here.

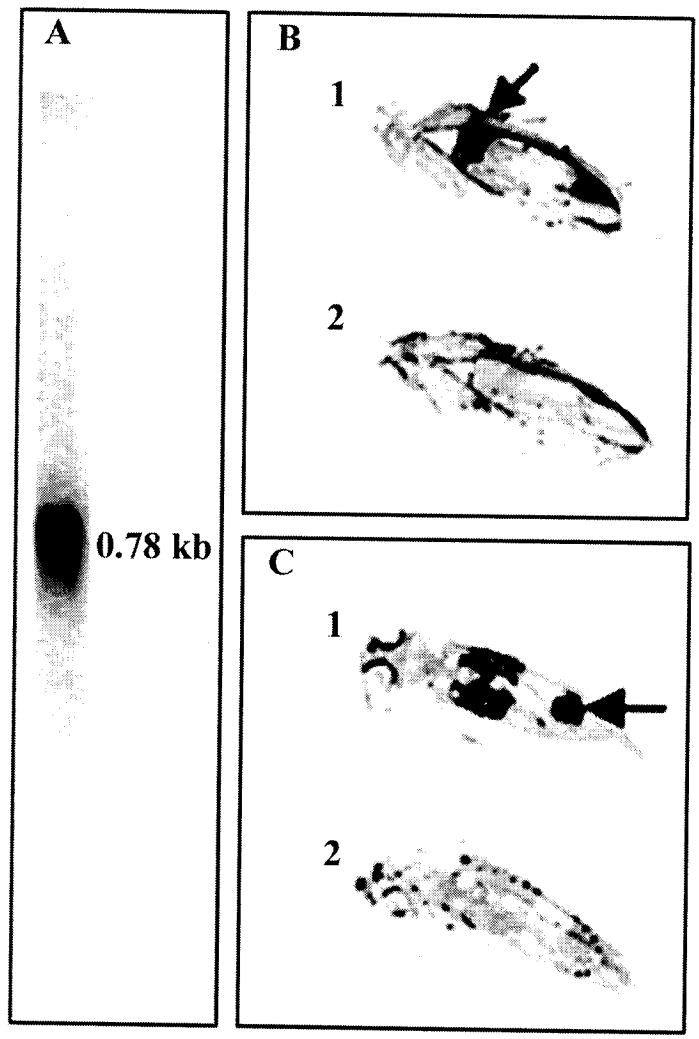



**Figure 3-8. *fabp3* gene expression in adult zebrafish.** (**A**) Northern blot analysis of total RNA isolated from whole adult zebrafish using the *fabp3* cDNA as a hybridization probe detected a single transcript of 780 bases. (**B** and **C**) *In situ* hybridization of sense and antisense *fabp3* specific oligonucleotides to mRNA in transverse tissue sections of adult zebrafish. The arrows indicate specific hybridization of the antisense probe to the *fabp3* mRNA in liver (B1) and ovary (C1). No specific hybridization was evident when the sense oligonucleotide was used as a hybridization probe (B2 and C2).

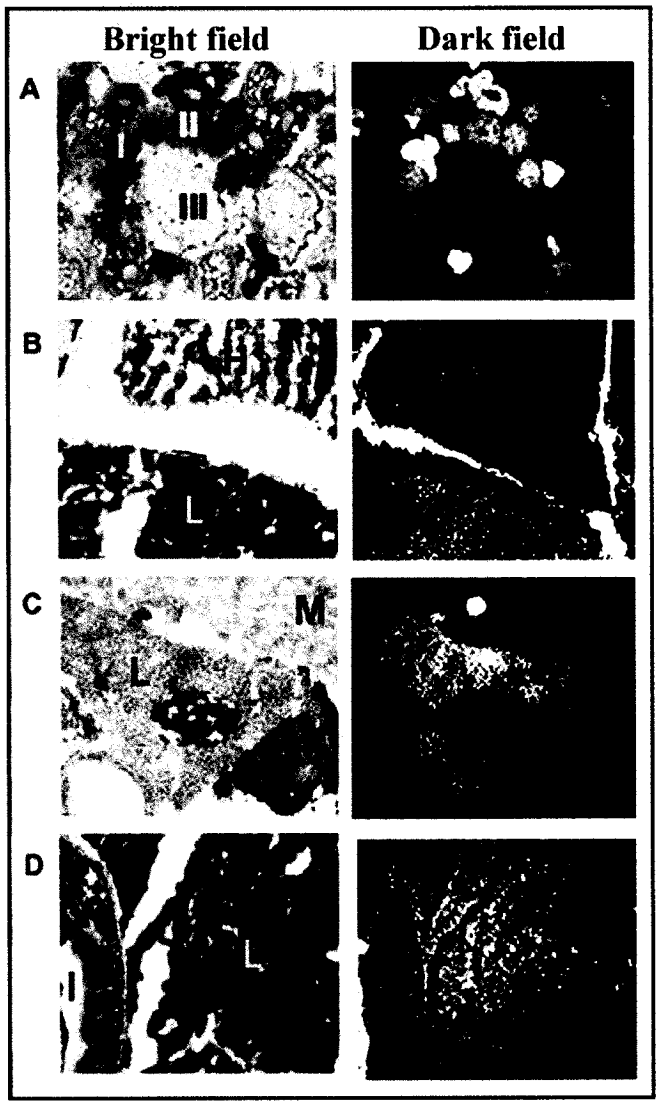



**Figure 3-9. Autoradiographic emulsion of zebrafish sections hybridized to the *fabp3* antisense probe.** The zebrafish sections that hybridized to the *fabp3* probe were exposed to autoradiographic emulsion, cresyl violet counter-stained and viewed under bright field (panel on the left) and dark field (panel on the right) illumination. (A) Silver grains corresponding to the *fabp3* mRNA were visualized by dark field illumination in different stages of zebrafish oocytes. Abundant silver grains were observed throughout stage I oocytes (I in bright field). In stage II oocytes (II in bright field) the density of silver grains diminished relative to stage I oocytes and was restricted to the ooplasm. No silver grains were observed in stage III (III in bright field) and matured oocytes. (B-D) Silver grains were detected over hepatocytes of the liver (L), but not in heart (H, panel B), muscle (M, panel C), or intestine (I, panel D).

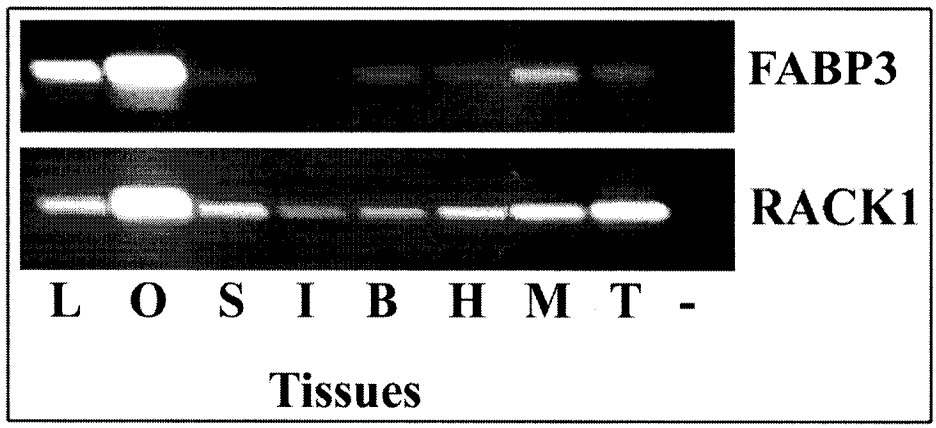



**Figure 3-10. Zebrafish *fabp3* mRNA in adult tissues detected by reverse-transcription polymerase chain reaction.** Zebrafish *fabp3* mRNA-specific primers amplified a product from total RNA extracted from ovary (O), liver (liver), skin (S), intestine (I), brain (B), heart (H), testis (T) and muscle (M). A RT-PCR product was generated from RNA in all samples for the constitutively expressed, Receptor for Activated C Kinase (RACK1), as a positive control and in RT-PCR using total RNA from adult zebrafish (A). A negative control (-) lacking RNA template generated no RT-PCR products.

### 3.2.7 Potential 5′ *cis* regulatory elements of the zebrafish *fabp3* gene

Inspection of the 5′ upstream sequence of the zebrafish *fabp3* gene revealed a typical cellular and viral TATA box element, with a matrix sequence of 5′-ttaTAAAtcagccag-3′. The core sequence (TAAA) of this TATA box is located 60 bp upstream of the major transcription start site. This TATA box in the zebrafish *fabp3* gene differs from the common TATA box element (TTTAAA) found in the *fabp3* gene promoter sequence of mouse [Treuner *et al.*, 1994], rat [Zhang *et al.*, 1999] and pigs [Gerbens *et al.*, 1997]. Two adjacent GC boxes (-102, -112) are located further upstream in the proximal promoter of the zebrafish *fabp3* gene (Fig. 3-5A).

Numerous transcription factor responsive elements were predicted by computer analysis of the 1220 bp 5′ flanking sequence of *fabp3* gene, and some of these may be associated with the tissue-specific expression of this gene in zebrafish (Table 3-2). A Yin Yang 1 (YY1) transcription element was identified in the 5′ flanking sequence that may be associated with oocyte-specific expression of zebrafish *fabp3* gene. Studies in tissue culture suggest that YYI may play a role in controlling expression of developmentally-regulated genes. Recently it has been reported that YY1 is abundant in the oocytes of mouse [Donohoe *et al.*, 1999] and *Xenopus* [Ficzycz and Ovsenek, 2002]. *Fabp3* might be one of the targeted genes of YY1 in zebrafish oocytes. Two additional types of *cis* elements for the transcription factors E2F and GATA-2 were found in the 5′ upstream region of the zebrafish *fabp3* gene. These transcription factors are expressed during *Drosophila* and *Xenopus* oogenesis [Myster *et al.*, 2000; Partington *et al.*, 1997].

The presence of the hepatocyte nuclear factor 1 (HNF1) elements in the 5′ upstream region of the zebrafish *fabp3* gene may be relevant to the expression of this

gene in the zebrafish liver. In a recent report, the expression of L-FABP was markedly reduced in HNF1α-null mice. HNF1α elements were found in the 5′ flanking sequence of the mouse L-FABP gene and HNF1α is required for transactivation of the L-FABP promoter [Akiyama *et al.*, 2000]. These findings indicated an important role of HNF1α in control of expression of the L-FABP gene in mouse liver. Conceivably, expression of the zebrafish *fabp3* gene in liver may be regulated by HNF1α.

**Table 3-2. Potential 5′ cis regulatory elements of the zebrafish *fabp3* gene**

| Name of family/matrix | Further information | Position | Strand | Core sim | Matrix sim. | Sequence |
|---|---|---|---|---|---|---|
| TBPF/TATA.01 | TATA box elements | -53 | (+) | 1.000 | 0.925 | ttaTAAAtca gccag |
| AP2F/AP2.01 | activator protein 2 | -94 | (+) | 0.976 | 0.924 | ccCCCCcag gcc |
| AP1F/AP1.01 | AP1 binding site | -902 | (-) | 0.881 | 0.954 | gtgaATCAa |
| SP1F/SP1.01 | stimulating protein 1 (SP1) | -100 | (-) | 1.000 | 0.896 | ggggGGCG gatgg |
| HNF1/HNF1.01 | hepatic nuclear factor 1 | -297 | (+) | 1.000 | 0.830 | cGTTAatta gttttt |
| HNF1/HNF1.02 | Hepatic nuclear factor 1 | -891 | (-) | 0.806 | 0.774 | tGATAataa atgtgaat |
| GATA/GATA1.05 | GATA-binding factor 1 | -333 | (+) | 1.000 | 0.966 | ttaGATAaa a |
| GATA/GATA1.03 | GATA-binding factor 1 | -635 | (+) | 1.000 | 0.949 | ccctGATAa atta |
| GATA/GATA1.02 | GATA-binding factor 1 | -671 | (+) | 1.000 | 0.966 | tgctgGATA agtgg |
| GATA/GATA2.01 | GATA-binding factor 2 | -889 | (-) | 1.000 | 0.945 | aatGATAat a |
| SORY/SOX5.01 | Sox-5 | -363 | (+) | 1.000 | 0.862 | atgaCAATg a |
| SORY/SOX5.01 | Sox-5 | -955 | (-) | 1.000 | 0.861 | tataCAATct |
| YY1F/YY1.01 | Yin and Yang 1 | -760 | (-) | 1.000 | 0.839 | atatggCCA Tttagtttatt |
| ECAT/NFY.02 | nuclear factor Y | -974 | (-) | 1.000 | 0.928 | catCCAAtc gc |
| ECAT/NFY.02 | nuclear factor Y | -1099 | (-) | 1.000 | 0.946 | catCCAAtca c |
| E2FF/E2F.01 | E2F, involved in cell cycle regulation | -1082 | (+) | 0.750 | 0.777 | tgcacggGG AAaatg |
| E2FF/E2F.02 | E2F, involved in cell cycle regulation | -1198 | (+) | 1.000 | 0.849 | gcacCAAA |

There are differences in the number and location of *cis* regulatory elements in the 5′ upstream sequences of the *fabp3* gene between zebrafish and mammals. The widespread E-box elements in rodent *fabp3* genes [Treuner *et al.*, 1994; Zhang *et al.*, 1999] were not found in the 5′ flanking sequence of the zebrafish *fabp3* gene. Moreover, the DR-1 element, a binding site for the retinoic acid receptor, retinoid X receptor and peroxisome proliferator-activated receptor (PPAR), identified in the 5′ upstream region of rodent *fabp3* genes [Treuner *et al.*, 1994; Zhang *et al.*, 1999], is absent in the 5′ upstream sequence of the zebrafish *fabp3* gene. In contrast, the abundant POU elements distributed throughout the 5′ upstream sequence of the zebrafish *fabp3* gene (data not shown) were not reported for the mammalian *fabp3* genes. However, elements for transcription factors, AP1 and NF1, are present in the 5′ upstream region of both zebrafish and mammalian *fabp3* genes [Treuner *et al.*, 1994; Zhang *et al.*, 1999].

## 3.3 DISCUSSION

Among the members of the iLBP multigene family, H-FABP shows the widest range of tissue distribution. Mammalian H-FABP is found in heart, skeletal and smooth muscle, specific regions of the brain, distal tubule cells of the kidney, stomach parietal cells, lactating mammary gland, ovary, testis and placenta [reviewed in Veerkam and Maatman, 1995; Heuckeroth *et al.*, 1987], but is absent in the liver, white fat and intestine [Heuckeroth *et al.*, 1987; Chrisman *et al.*, 1987; Claffey *et al.*, 1987; Miller *et al.*, 1988; Bass and Manning, 1986; Hittel and Storey, 2000; Kurtz *et al.*, 1994; Paulussen *et al.*, 1990]. The zebrafish *fabp3* gene, described in the present study, showed highest amino

acid sequence similarity, identical gene structure and coding capacity, and conserved genomic syntenies to mammalian H-FABPs. However, the zebrafish *fabp3* gene displayed a different pattern of tissue distribution to that of the orthologous mammalian H-FABPs. Steady-state mRNA level of the H-FABP in adult tissues of five different fish species has been analyzed by northern blot hybridization. In the four Antarctic teleost fish species, both H-FABP isoforms exhibited similar expression patterns to the mammalian H-FABP, i.e. high mRNA level in the heart, muscle and brain, but absent in the liver [Vayda *et al.*, 1998]. In the mumichog (*Fundulus heteroclitus*), the H-FABP mRNA level was most abundant in the male liver, gills and gonads (female tissues were not examined), which is more similar to the tissue-distribution pattern of the zebrafish FABP3 reported here [Bain, 2002]. By *in situ* hybridization and emulsion autoradiography, I detected the *fabp3* gene transcript in ovary and liver (Fig. 3-8, 9). Only by the highly sensitive technique of RT-PCR was I able to detect the *fabp3* gene transcript in the other tissues such as brain, heart, intestine, muscle, skin and testis of adult zebrafish (Fig. 3-10). Comparison of the 5′ upstream sequence of the zebrafish *fabp3* gene and the orthologous mammalian genes revealed numerous differences in *cis* elements. Together, the differences between zebrafish and mammals in *cis* elements and the expression pattern of the *fabp3* gene suggests that, while primary amino acid, gene structure and syntenic relationships have been conserved, *cis* elements in the 5′ upstream regions of these genes have apparently evolved following divergence of fish and mammals leading to altered patterns of gene expression. This is not the case for the zebrafish B-FABP (*fabp7*) gene, which shows conservation in both expression pattern (brain-specific) and regulatory elements with its mammalian orthologs [Liu *et al.*, 2003].

Although at least sixteen paralogous members of the FABP multigene family have been characterized in various species, the precise *in vivo* physiological functions of each protein are not well understood. In mammalian cardiac and skeletal muscle, H-FABP is thought to participate in fatty acid β-oxidation and energy production [reviewed by Hertzel and Bernlohr, 2000; Glatz and Vusse, 1996]. Studies using H-FABP knockout mice demonstrated that H-FABP is required for efficient uptake, intracellular transportation and utilization of fatty acids in cardiac muscle [Binas *et al.*, 1999; Schaap *et al.*, 1999]. However, H-FABP is also abundant in tissues that do not use fatty acids as an energy source such as mammary gland and developing brain [Treuner *et al.*, 1994, Sellner *et al.*, 1995]. Similarly, the detection of high steady-state levels of the *fabp3* gene transcript in zebrafish oocytes and liver, tissues which do not use fatty acids primarily as energy sources, indicates that FABP3 participates in β-oxidation in muscle and lipogenesis in other tissues in fish and mammals. FABP3 (H-FABP) may play a general and fundamental role in fatty acid transportation in tissues exhibiting anabolic and catabolic lipid metabolism.

During development of the animal oocyte, large quantities of mRNAs, rRNAs and tRNAs are synthesized, some mRNAs are translated immediately into protein, while others are stored in an inactive form, ribosomes and mitochondria accumulate, and quantities of polysaccharides and lipids are synthesized. In addition to the metabolic activity of the oocyte, proteins, lipids and carbohydrates enter the cytoplasm from outside the cell [Wolfe, 1993]. During stage III of oocyte development or vitellogenesis, much of the stored lipid and protein is packaged into yolk granules that accumulate in all animal oocytes except mammals [Selman *et al.*, 1993; Wolfe, 1993]. Oocyte development in

oviparous species (birds, fish, amphibians and reptiles) is, indeed, dependent on the uptake of nutrients and their storage as yolk, whose constituents are subsequently used by the embryo during early stages of development. The yolk granules often comprise 95% of the cytoplasmic volume that accounts for the relatively large size of eggs of oviparous species compared to eggs of mammalian species [Wolfe *et al.*, 1993]. The abundance of the *fabp3* mRNA prior to the vitellogenic (III) stage seen in the zebrafish (Fig. 3-9A) correlates in time with accumulation of fatty acids within the oocyte. Fatty acids may be sequestered by FABP to prevent cytotoxicity to the cell, and transported by FABP to sites of triglyceride synthesis and/or storage within the cytoplasm. As antibodies to the zebrafish *fabp3* are not currently available, I am unable to assess the stage of oocyte maturation at which the *fabp3* gene transcript is translated into protein. I suspect, however, that the *fabp3* mRNA is most abundant immediately prior to and during the vitellogenic (III) stage of oocyte development.

In Antarctic teleost fishes, the mRNA of two distinct heart-type FABP isoforms has been detected in cardiac tissue [Vayda *et al.*, 1998]. It is likely that there is a duplicated *fabp3* gene in zebrafish, which may play a role in muscular β-oxidation. Based on preliminary analysis, I have cloned a cDNA and identified the corresponding gene in the zebrafish genome sequence database (Wellcome Trust Sanger Institute) that may be expressed in the zebrafish heart. Characterization of this newly discovered gene might provide clues to the expression and function of FABPs in zebrafish cardiac tissue.

# 4 Structure, mRNA Expression and Linkage Mapping of the Brain-Type Fatty Acid-Binding Protein Gene (*fabp7*) from Zebrafish (*Danio rerio*)

## 4.1 INTRODUCTION

Long-chain polyunsaturated fatty acids are highly concentrated in brain and play vital roles in visual and brain development [reviewed by Uauy *et al.*, 2001; salem *et al.*, 2001]. Fatty acids are a basic component of biological membranes and their overall quantity and composition affect membrane biophysical properties and function [Lee *et al.*, 1986; Stubbs and Smith, 1990]. In the central nervous system (CNS), fatty acids serve as regulators of gene expression [reviewed by Uauy *et al.*, 2001; Ntambi and Bene, 2001]. Intra-cellular uptake, transport and metabolism of fatty acids are thought to be mediated by fatty acid-binding proteins (FABPs), a group of low molecular weight (14-16 kDa) proteins encoded by a multigene family [reviewed by Glatz and van de Vusse *et al.*, 1996; Coe and Bernlohr, 1998; Hertzel and Bernlohr, 2000]. Brain-type fatty acid-binding protein (B-FABP) was first isolated from rat brain [Bass *et al.*, 1984; Senjo *et al.*, 1985] and was later found to be a brain-specific member of the FABP family with high expression levels in the developing CNS [Kurtz *et al.*, 1994; Feng *et al.*, 1994; Shimizu *et al.*, 1997]. Ligand binding experiments have shown that docosahexaenoic acid (DHA) is the likely physiological ligand for B-FABP as affinity of B-FABP for DHA ($K_d \sim 10$ nM) is the highest ever reported for a FABP/ligand interaction [Xu *et al.*, 1996]. The essential roles of DHA in CNS development [Uauy *et al.*, 2001; salem *et al.*, 2001], the spatial and temporal expression pattern of the B-FABP gene [Kurtz *et al.*, 1994; Feng *et al.*, 1994;

Shimizu *et al.*, 1997], and the ligand specificity of B-FABP for DHA [Xu *et al.*, 1996] suggest an important role for B-FABP in the CNS development through mediation of DHA utilization. How the expression of the B-FABP gene is regulated *in vivo* remains unclear.

Identification of *cis*-acting regulatory elements and the transcription factors that bind to them in the B-FABP gene is an initial step to determine the regulatory mechanisms that govern the tissue-specific and developmental expression of the B-FABP gene. Feng and Heintz [Feng and Heintz, 1995] have identified *cis*-acting elements in the 5′ upstream region of the mouse B-FABP gene involved in regulation of its transcription in radial glia cells. Later, Josephson *et al.* [Josephson *et al.*, 1998] found that expression of the rat B-FABP gene depends on interaction of POU with POU domain binding sites in its promoter region for general CNS expression, while a hormone response element is additionally required for its expression in the anterior CNS.

In a previous paper, Denovan-Wright *et al.* [2000] reported the sequence of cDNA clones coding for a B-FABP in zebrafish and showed by *in situ* hybridization that the B-FABP mRNA is predominantly expressed in the periventricular gray zone of the optic tectum of the adult zebrafish brain. As both mammalian and zebrafish B-FABP genes were found to be expressed predominantly in the brain, I wished to determine whether the zebrafish and mammalian B-FABP genes share common *cis*-acting regulatory elements in their 5′ upstream regions that confer brain-specific expression. In addition, I intended to determine whether the structure and syntenic relationship of B-FABP gene is conserved between the zebrafish and mammalian genomes as the FABP multigene family is thought to have originated by a series of duplications of a common

ancestral gene, with most duplications occurring before the divergence of invertebrates and vertebrates [Schleicher *et al.*, 1995]. Here I report the gene structure, tissue-specific and temporal expression, potential *cis*-acting regulatory elements of the promoter and gene linkage mapping of the B-FABP gene from zebrafish (*Danio rerio*).

## 4.2 RESULTS AND DISCUSSION

### 4.2.1 Sequence and structure of the zebrafish B-FABP gene

DNA traces showing sequence identity to the B-FABP cDNA clone, fb62f07.y1 [Denovan-Wright *et al.*, 2000], were retrieved from the zebrafish genome sequence database of the Wellcome Trust Sanger Institute. One trace (zfishC-a1872h08.q1c) contained the sequence of exon 1, intron 1 and exon 2, while a second trace (z35723-a1961g12.p1c) contained the sequence for exon 2, intron 2 and exon 3. A third trace (zfish43795-71b04.p1c) contained the entire sequence of exon 4. Intron 3, a portion of which was missing from trace z35723-a1961g12.p1c, was PCR amplified and sequenced. In addition, a 1249 bp fragment upstream of exon 1 of the B-FABP gene was obtained by linker-mediated PCR and cloned and sequenced. The exon/intron organization of the zebrafish B-FABP gene (Fig. 4-1), which consists of 4 exons (nucleotides 1-163, 290-462, 616-717 and 2081-2370, respectively) separated by 3 introns (nucleotides 164-289, 463-615 and 718-2080, respectively), is the same as for all the FABP genes and other members of this multigene family reported thus far [Bernlohr *et al.*, 1997], with the exception of the desert locust muscle-type FABP which lacks intron 2 [Wu *et al.*, 2001]. The coding sequence of the zebrafish B-FABP gene was identical to that previously reported for the zebrafish B-FABP cDNA sequence of clone fb62f07.y1 [Denovan-

Wright *et al.*, 2000]. The coding capacity of the 4 exons (encoding 24, 58, 34 and 16 amino acids, respectively) is identical to that of the human and mouse B-FABP genes, whereas the size of introns 1-3 varies among human, mouse and zebrafish (Fig. 4-2A). An interesting note is the increasing size of each of the three introns, i.e. intron 1<intron 2<intron 3 (Fig. 4-2A), is maintained between fishes and mammals. All intron/exon splice junctions of the zebrafish B-FABP gene conform to the GT-AG di-nucleotide rule [Breathnach and Chambon, 1981].

The four exons of the zebrafish B-FABP gene contain 708 nucleotides. Northern blot and hybridization using a zebrafish B-FABP-specific cDNA probe detected an mRNA transcript of approximately 850 nucleotides [Denovan-Wright *et al.*, 2000]. Considering the average size of the poly(A) tail of eukaryotic mRNAs (150-200 nucleotides), the predicted and observed sizes of zebrafish B-FABP mRNA are in close agreement.

The amino acid sequence of the zebrafish B-FABP was deduced from each of the individual exons of the B-FABP gene and aligned with the same peptide sequence from the human, mouse and pufferfish orthologous B-FABP genes (Fig. 4-2B). The percentage amino acid identity between zebrafish and human, mouse and pufferfish B-FABP is 83%, 76%, and 83%, respectively. The percentage amino acid identity between zebrafish and human and mouse is higher in the exons 1 and 2 than it is in the exons 3 and 4 coding for B-FABP. This result is consistent with previous observations for the human and rat I-FABP, and other members of the FABP family, that the $NH_2$-terminal halves of these proteins are more highly conserved than their COOH-terminal halves [Sweetser *et al.*, 1987 and references therein].

**Figure 4-1. Nucleotide sequence of the zebrafish B-FABP gene and its 5′ upstream region.** Exons are shown in uppercase letters with the coding sequences of each exon underlined and the deduced amino acid sequence indicated below. The initiation site for transcription, mapped by 5′ RLM-RACE, is numbered at +1, and a putative polyadenylation signal is highlighted in bold type. A potential TATA box 19 bp upstream of the transcription initiation site, a GC box and a CAAT box are boxed. The GenBank accession number for the sequence of the zebrafish B-FABP gene is AY145893.

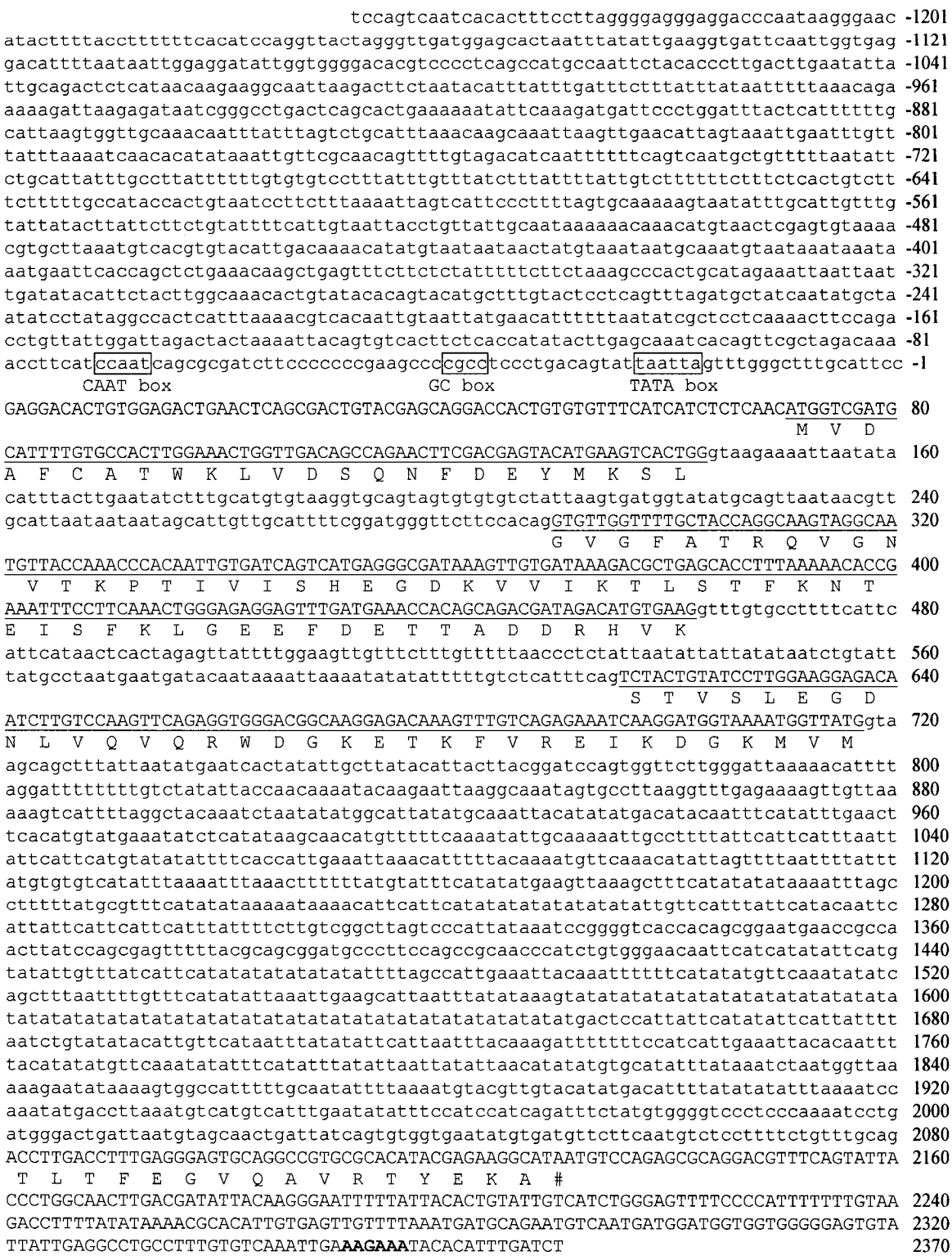
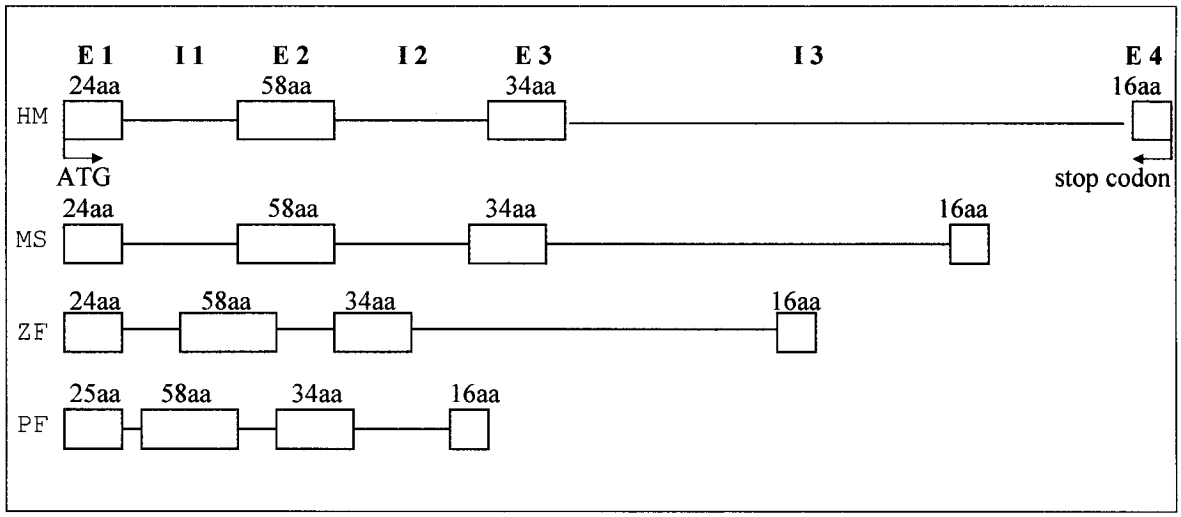



**Figure 4-1. Nucleotide sequence of the zebrafish B-FABP gene and its 5′ upstream region.**

**Figure 4-2. Structure of B-FABP genes from fishes and mammals.** (**A**) Comparison of the exon/intron organization of the zebrafish B-FABP gene (ZF) with the orthologous genes from human (HM), mouse (MS) and pufferfish (PF). Exons (E1-E4) are shown as boxes and introns (I1-I3) as solid lines. The length of the boxes and lines represent the approximate size of the exons and introns, respectively, with the number of amino acids encoded by each exon shown above the boxes. The human and mouse B-FABP gene sequences were obtained from GenBank (accession numbers NT_033944 and U04827). The sequence of the pufferfish B-FABP gene was retrieved from scaffold 3785 by searching the Fugu (pufferfish) genome project database (V1.0) at http://www.jgi.doe.gov /fugu. (Wellcome Trust Sanger Institute). (**B**) The deduced amino acid sequence encoded by each exon of the zebrafish B-FABP gene (ZFb-FABP) was aligned with the amino acid sequence encoded by each exon from the human (HMb-FABP), mouse (MSb-FABP) and pufferfish (PFb-FABP) B-FABP genes using CLUSTALW [Thompson *et al.*, 1994]. Dots indicate amino acid identity and dashes a deletion/insertion. The percentage amino acid sequence identity for the peptides encoded by each exon of the B-FABP gene between zebrafish and human, mouse and pufferfish is shown at the right of each exon.

**A**




**B**

**Exon 1:**

```
ZFbFABP MVDAFCATWKLVDSQNFDEYMKSL-  100%
HMbFABP ..E........TN.........A.-   83%
MSbFABP ..........T.........A.-     91%
PFbFABP ............E.E...D.T.A.G   79%
```

**Exon 2:**

```
ZFbFABP GVGFATRQVGNVTKPTIVISHEGDKVVIKTLSTFKNTEISFKLGEEFDETTADDRHVK 100%
HMbFABP ................VI..Q.......R............Q.............NC.  88%
MSbFABP ................VI..Q..G....R.QC.......N.Q.....E..SI...NC.  75%
PFbFABP ................VI..L......VR.Q.........................NC.  84%
```

**Exon 3:**

```
ZFbFABP STVSLEGDNLVQVQRWDGKETKFVREIKDGKMVM  100%
HMbFABP .V...D..K..HI.K......N.............  79%
MSbFABP .V.R.D..K.IH..K......NCT.........V  67%
PFbFABP ........K..H..K......T.............  88%
```

**Exon 4:**

```
ZFbFABP TLTFEGVQAVRTYEKA 100%
HMbFABP ....GD.V...H....  75%
MSbFABP ....GDIV...C....  68%
PFbFABP D....D.H........  81%
```

**Figure 4-2. Structure of B-FABP genes from fishes and mammals.**

### 4.2.2 Mapping of the initiation site of transcription for the zebrafish B-FABP gene

In order to map the initiation site of transcription for the zebrafish B-FABP gene, I performed 5′ RLM-RACE and obtained the 5′ cDNA end from the capped and complete mRNA sequence. A single band was detected from the CIP/TAP treated RNA after nested PCR amplification, but no product was observed from the RNA sample that was not treated with TAP, which served as negative control (Fig. 4-3). Thus, this single RACE product most likely represents the 5′ end of the mature B-FABP mRNA. The 5′ RACE product contained a 166 bp sequence corresponding to a portion of exon 1 including the 5′ UTR of the zebrafish B-FABP mRNA. The potential transcription start site of zebrafish B-FABP was mapped to 70 bp upstream of the initiation codon by aligning the 5′ RLM RACE sequence with the B-FABP gene sequence. The sequence of the 5' RACE product was identical to its corresponding genomic sequence. In contrast to several mammalian FABP genes, which possess two or more transcription start sites [Sweetser *et al.*, 1987; Qian *et al.*, 1999], only a single transcription start site was found in the zebrafish B-FABP gene. A putative TATA box is present 19 bp upstream from the transcription start site. A GC box (-38) and a CAAT box (-68) are located further upstream in the proximal promoter of the zebrafish B-FABP gene (Fig. 4-1). These elements are general features of many eukaryotic core promoters.

### 4.2.3 Putative 5′ *cis* regulatory elements of the zebrafish fabp7 gene

Neuronal cell differentiation is generally thought to be regulated by a cascade of transcription factors. Analysis of the sequence 5′ upstream of exon 1 of the B-FABP gene revealed a number of potential *cis*-acting regulatory elements, which may provide

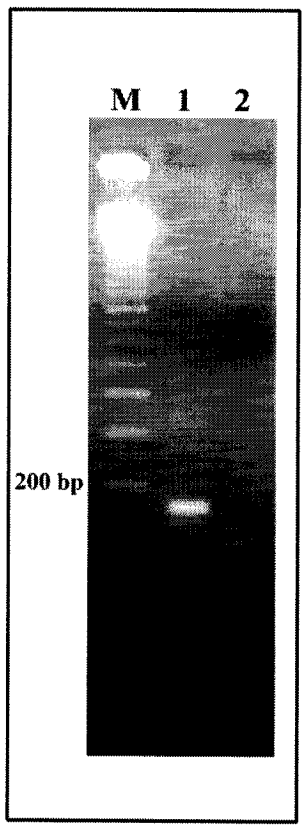



**Figure 4-3. Product of 5′ RLM-RACE derived from the 5′ end of the mature zebrafish B-FABP mRNA.** Total RNA from whole adult zebrafish was sequentially treated with calf intestinal alkaline phosphatase (CIP), tobacco acid pyrophosphatase (TAP) and then ligated to a designated RNA adapter. Following two rounds of nested PCR, a single, PCR-amplified product of approximately 170 bp was size-fractionated by gel-electrophoresis through 1% agarose (lane 1). RNA treated to the same experimental regime, but with TAP digestion omitted, did not generate a product (lane 2). A ladder of 100 bp molecular weight markers (MBI Fermentas) is shown in lane M with the 200 bp marker indicated to the left of the panel.

clues to the spatial and temporal expression patterns of the B-FABP gene in zebrafish (Table 4-1). POU domain recognition elements were the most abundant transcription factor binding sites identified within the 1249 bp 5′ upstream sequence. The nine POU elements dispersed throughout the 5′ upstream sequence of the zebrafish B-FABP gene included three Octamer-binding factor-1 (Oct-1), one Brain-3 (Brn-3), two Brain-2 (Brn-2), two Testis-1 (Tst1) and one GHF-1 pituitary specific POU domain transcription factor (Pit1) elements. POU-domain genes were first identified in mammals, encoding three transcription factors, Pit-1 [Ingraham *et al.*, 1988], Oct-1 [Sturm *et al.*, 1988] and Oct-2 [Clerc *et al.*, 1988]. He *et al.* [1989] reported a large number of POU-domain regulatory genes, which are widely expressed in the developing mammalian neural tube, and exhibit differential, overlapping patterns of expression in the adult mammalian brain. Several CNS-specific genes, including the B-FABP gene, contain POU-domain binding sites, which drive their expression throughout the developing mammalian CNS [Josephson *et al.*, 1998]. Investigation of POU-domain genes in zebrafish has revealed their specific patterns of expression in developing neural tissues [Masuzaki *et al.*, 1992] and in the adult brain [Sampath and Stuart, 1996]. B-FABP is specifically expressed in the mammalian and zebrafish brain [Kurtz *et al.*, 1994; Shimizu *et al.*, 1997; Feng and Heintz, 1995; Denovan-Wright *et al.*, 2000], and its expression correlates temporally to mammalian neuronal and glial differentiation during development [Feng and Heintz, 1995].

Some mammalian POU-domain binding proteins are co-expressed with homeodomain proteins in the brain [He *et al.*, 1989 and references therein] and at least some of the homeobox genes or homeodomain proteins are required for neuronal

development [Doe *et al.*, 1988; Halter *et al.*, 1995]. In a recent morphological and molecular study on the medaka optic tectum, the expression of two homeobox genes, *paired-related-homeobox3* (*Ol-Prx3*) and *genetic-screen-homeobox1* (*O1-Gsh1*), correlated with proliferative events in the developing tectum [Nguyen *et al.*, 1999]. It has previously been shown that the zebrafish B-FABP mRNA is localized to the adult optic tectum [Denovan-Wright *et al.*, 2000]. Neurogenesis is ongoing in the optic tectum of teleost fishes [Birse *et al.*, 1980] and specific brain nuclei in birds [Rousselot *et al.*, 1997]. Significantly, in the 5′ upstream region of the zebrafish B-FABP gene, I identified a number of potential homeodomain binding elements in addition to the abundant POU-domain elements (data not shown).

In the 1249 bp 5′ upstream sequence of the zebrafish B-FABP gene, four copies of nuclear factor Y (NF-Y) binding element are present. NF-Y is a transcription factor that recognizes the consensus sequence 5′ YYRRCCAATCAG 3′ present in the promoter region of many constitutive, inducible and cell-cycle-dependent eukaryotic genes [Maity and de Crombrugghe, 1998]. It has been suggested that NF-Y may interact with other transcription factors or nuclear proteins to regulate genes harboring NF-Y elements [Matuoka and Chen, 1999]. Activation of the neuronal aromatic L-amino acid decarboxylase gene promoter requires a direct interaction between the NF-Y factor and a POU-domain protein Brn-2 [Dugast and Weber, 2001]. Polyunsaturated fatty acids are thought to up-regulate the expression of fatty acid oxidation-related genes by activating peroxisome proliferator-activated receptors α (PPAR-α), but also down-regulate lipogenic genes through their suppressive effect on another group of transcription factors, including NF-Y [Clarke, 2001]. I did not find any PPAR response elements in the 5′

upstream sequence of the zebrafish B-FABP gene, but did find a number of potential NF-Y binding elements. Considering the spatial expression of B-FABP, the physiological function of zebrafish B-FABP may be limited primarily to lipogenic processes rather than lipid oxidation.

Several other distinct transcription factor binding motifs were identified in the 5′ upstream sequence of the zebrafish B-FABP gene, including elements for activator protein-1 (AP-1), SRY-related HMG box-5 (SOX-5), cAMP responsive element binding protein (CREB), GATA-1 and GATA-2. A number of these elements are the target for transcription factors known to play a role in neuronal development or survival and plasticity of neurons in adult mammalian brain. For example, although the precise physiological function for AP-1 is not known, it is generally considered that AP-1 may regulate a wide range of cellular processes including cell proliferation, survival, differentiation and death [Shaulian Karin, 2002]. In the adult mammalian brain, AP-1 is also thought to play a role in neuroprotection and neurodegeneration [Hergegen and Waetzig, 2001]. In humans, the *SOX5* gene is expressed in fetal brain and adult testis [Wunderle *et al.*, 1996]. A large number of potential SOX binding sites have been found in the promoter region of the brain-specific *cyp19* genes in a teleost fish [Tchoudakova *et al.*, 2001]. Among the large SOX family, only the SOX5 binding site is present in the promoter sequence of the zebrafish B-FABP gene. The cAMP-CREB cascade plays an important role in neuronal survival and plasticity, and regulates adult neurogenesis [Nakagawa *et al.*, 2002]. A recent study has shown that disruption of CREB function in brain results in neurodegeneration [Mantamadiotis *et al.*, 2002]. GATA-1 (previously termed as Eryf1, NF-E1, or GF-1) is a transcription factor that recognizes *cis*-elements

widely distributed throughout the promoters of erythroid-specific genes. GATA-1 is also widely expressed in brain [Yamamoto *et al.*, 1990], although little is known about its physiological function in this tissue. Identification of the target genes specifically expressed in brain could be a useful approach to elucidate the function of this transcription factor. GATA-2 was recently found to be required for the generation of V2 interneurons in transgenic mice [Zhou *et al.*, 2000]. Moreover, GATA2 gene expression in the CNS, as assayed by microinjection of the GATA-2 promoter fused to the green fluorescent protein reporter gene into single cell embryos, proceeds the onset of B-FABP mRNA expression during zebrafish embryogenesis reported here. In this cascade of transcription factors, the GATA-2 gene itself is regulated by a neuronal-specific cis-acting element, CCCTCCT, in the GATA-2 gene promoter, that presumably binds a neuronal-specific transcription factor [Meng *et al.*, 1997]. Both GATA-1 and GATA-2 binding elements were found in the 5′ upstream sequence of the zebrafish B-FABP gene, again suggesting their potential function in neuronal development or growth.

The presence of several classes of transcription factor binding elements in the 5′ upstream region of the zebrafish B-FABP gene, elements known to participate in signaling pathways that influence neural growth, differentiation or plasticity, suggests that the zebrafish B-FABP gene plays a role in neurogenesis. Confirmation that these putative transcription factor binding elements in the zebrafish B-FABP gene direct its expression will require detailed functional analysis of the promoter region and DNA gel-shift and DNA footprinting assays using nuclear protein extracts.

**Table 4-1. Potential 5′ *cis* regulatory elements of zebrafish B-FABP gene**

| Name of family/matrix | Further Information | Position | Strand | Core sim. | Matrix sim. | Sequence |
|---|---|---|---|---|---|---|
| SP1F/GC.01 | GC box elements | -34 | (-) | 1.000 | 0.929 | gggaGGCGgggctt |
| PCAT/CAAT.01 | cellular and viral CCAAT box | -66 | (+) | 1.000 | 0.957 | ttcatCCAAtca |
| OCTB/TST1.01 | POU-factor Tst-1/Oct-6 | -126 | (+) | 1.000 | 0.874 | ctaaAATTacagtgt |
| OCTP/OCT1P.01 | POU-specific domain /Oct1 | -238 | (+) | 1.000 | 0.912 | atcaatATGCtaata |
| BRNF/BRN2.01 | POU factor Brn-2 | -435 | (+) | 1.000 | 0.952 | aacatatgTAATaata |
| OCTB/TST1.01 | POU-factor Tst-1/Oct-6 | -522 | (-) | 1.000 | 0.905 | aggtAATTacaatga |
| BRNF/BRN2.01 | POU factor Brn-2 | -788 | (-) | 1.000 | 0.925 | ttgattttAAATaaac |
| BRNF/BRN3.01 | POU transcription factor Brn-3 | -963 | (+) | 1.000 | 0.809 | ATAAtttttaaaca |
| OCT1/OCT1.02 | POU octamer-binding factor 1 | -877 | (-) | 1.000 | 0.941 | aATGCaaaaa |
| PIT1/PIT1.01 | POU factor Pit1 | -911 | (+) | 1.000 | 0.891 | aaatATTCaa |
| OCT1/OCT1.02 | POU octamer-binding factor 1 | -1064 | (+) | 1.000 | 0.869 | cATGCcaatt |
| ECAT/NFY.02 | nuclear factor Y | -147 | (-) | 1.000 | 0.925 | aatCCAAtaac |
| ECAT/NFY.02 | nuclear factor Y | -1091 | (-) | 1.000 | 0.906 | ccaCCAAtatc |
| ECAT/NFY.02 | nuclear factor Y | -1122 | (-) | 1.000 | 0.915 | tcaCCAAttga |
| ECAT/NFY.01 | nuclear factor Y | -1203 | (+) | 1.000 | 0.937 | aggacCCAAtaaggga |
| GATA/GATA2.02 | GATA-binding factor 2 | -177 | (-) | 1.000 | 0.912 | agcGATAtta |
| GATA/GATA1.03 | GATA-binding factor 1 | -672 | (-) | 1.000 | 0.954 | taaaGATAaacaa |
| GATA/GATA1.02 | GATA-binding factor 1 | -940 | (+) | 1.000 | 0.965 | taagaGATAatcgg |
| SORY/SOX5.01 | Sox-5 | -200 | (-) | 1.000 | 0.860 | attaCAATtg |
| SORY/SOX5.01 | Sox-5 | -561 | (-) | 1.000 | 0.989 | caaaCAATgc |
| SORY/SOX5.01 | Sox-5 | -660 | (-) | 1.000 | 0.868 | aagaCAATaa |
| SORY/SOX5.01 | Sox-5 | -771 | (-) | 1.000 | 0.980 | cgaaCAATtt |
| SORY/SOX5.01 | Sox-5 | -858 | (+) | 1.000 | 0.984 | caaaCAATtt |
| CREB/CREB.01 | cAMP-responsive element binding protein | -210 | (-) | 1.000 | 0.934 | TGACgttt |
| AP1F/AP1.03 | activator protein 1 | -597 | (-) | 1.000 | 0.966 | aaTGACtaatt |
| AP1F/AP1.03 | activator protein 1 | -736 | (-) | 1.000 | 0.927 | atTGACtgaaa |
| AP1F/AP1.01 | activator protein 1 | -929 | (-) | 1.000 | 0.995 | ctgaGTCAg |

### 4.2.4 Tissue-specific and temporal distribution of B-FABP mRNA

Previously, I examined B-FABP expression in adult zebrafish by *in situ* hybridization to whole mount sections [Denovan-Wright *et al.*, 2000]. I performed RT-PCR analysis, a more sensitive technique than *in situ* hybridization, to determine B-FABP mRNA distribution in adult tissues and during embryogenesis. RT-PCR products were generated from brain RNA using zebrafish B-FABP cDNA-specific primers. RT-PCR-amplified products were also generated from RNA of liver, testes and intestine, but not in skin, heart, muscle and ovary (Fig. 4-4A). No RT-PCR product was detected in the negative control in which no cDNA template was added. Positive control RT-PCR reactions for each cDNA sample were performed for mRNA of the constitutively expressed zebrafish RACK1 gene. To confirm the tissue-distribution of B-FABP mRNA in adult zebrafish revealed by the conventional RT-PCR, quantitative RT-PCR (qRT-PCR) of B-FABP mRNA from the same tissues was performed using another constitutively expressed gene, the β-actin gene, as a positive control. Levels of B-FABP mRNA in each cDNA sample ranged between undetectable to $3.5 \times 10^2$ copies per μL of cDNA. β-actin RT-PCR products were amplified from every cDNA sample and ranged from $1.5 \times 10^2$ to $3.5 \times 10^5$ copies per μL. The ratio of B-FABP/ β-actin PCR product for each experimental sample was calculated (Fig. 4-4B). This analysis demonstrated that the levels of B-FABP mRNA are ~ 7 times higher in brain than in testes and between 50 and 160 times higher in brain than in muscle, intestine and heart. No product was generated by qRT-PCR from liver, ovary, skin and kidney RNA. Both conventional RT-PCR and qRT-PCR using different controls, i.e., RACK1 and β-actin mRNA, showed similar

tissue-distribution where the zebrafish B-FABP mRNA was abundant, but not in some tissues where the levels of B-FABP mRNA were low.

In a previous report, using tissue section *in situ* hybridization, I detected the B-FABP mRNA in the zebrafish periventricular zone of the optic tectum, but not in any other tissues [Denovan-Wright *et al.*, 2000]. As suggested by the results of conventional RT-PCR and qRT-PCR, the amount of zebrafish B-FABP mRNA in liver, testis, heart, muscle and intestine may be too low to be detected by *in situ* hybridization, but its presence in these tissues was revealed by the more sensitive method of RT-PCR. Using Northern blot and hybridization, B-FABP mRNA was detected in the liver of rat [Bennett *et al.*, 1994], but absent in the liver of mouse [Kurtz *et al.*, 1994]. In rat, however, the hybridization signal for B-FABP mRNA in liver was much weaker than that seen for brain RNA [Bennett *et al.*, 1994]. It is likely, therefore, that the low levels of B-FABP mRNA may not be detected by methods such as Northern blot and hybridization and *in situ* hybridization that are less sensitive than RT-PCR.

RT-PCR of RNA extracted from zebrafish embryos at different times post-fertilization (PF) revealed the temporal expression of the B-FABP gene during embryogenesis. No product was detected for the RNA from embryos at 1 and 12 h PF or in the negative control reactions (Fig. 4-4C). B-FABP-specific RT-PCR product was detected at 24 h PF and thereafter throughout zebrafish embryonic development. During zebrafish embryonic development, a premature central nervous system can be identified at approximately 12 h PF, the forebrain, midbrain and hindbrain can be distinguished at16 h PF, and brain ventricles are present and interneurons developed after 19 h PF (for embryonic zebrafish staging, see http://www.ana.ed. ac.uk/anatomy/database/zebrafish_

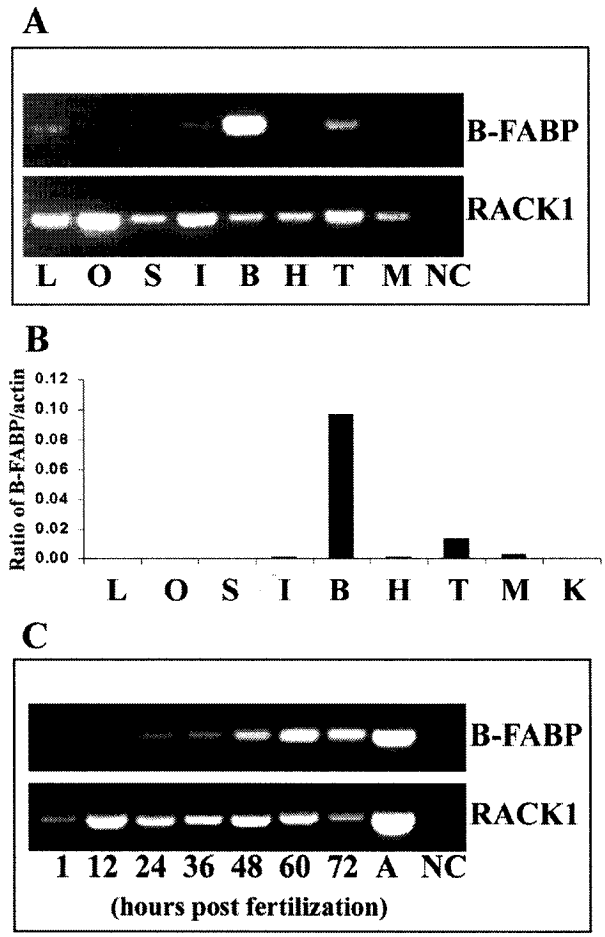



**Figure 4-4. B-FABP mRNA in adult tissues and developing embryos of zebrafish detected by RT-PCR. (A)** Zebrafish B-FABP cDNA-specific primers amplified by qualitative RT-PCR an abundant product in RNA extracted from adult zebrafish brain (B), and detectable product extracted from RNA from adult liver (L), intestine (I) and testis (T), but not from RNA extracted from ovary (O), skin (S), heart (H) or muscle (M). As a negative control (NC), RNA template was omitted from the RT-PCR reaction (upper panel). RT-PCR detected a product for the constitutively expressed RACK1 mRNA using cDNA-specific primers in RNA extracted from all tissues assayed (lower panel). **(B)** Quantitative RT-PCR was performed to determine the levels of zebrafish B-FABP and β-actin mRNAs in adult tissues. The histogram shows the ratio of B-FABP mRNA to β-actin mRNA in various tissues with abundant expression of the B-FABP mRNA seen in RNA extracted from adult brain (B), much lower B-FABP mRNA levels in testis (T), muscle (M), intestine (I), and heart (H), and undetectable levels in liver (L), ovary (O), skin (S) and kidney (K). **(C)** Qualitative RT-PCR did not generate a B-FABP mRNA-specific product from total RNA extracted from embryos, 1 and 12 h post-fertilization, but did generate a product from total RNA extracted from embryos, 24 h post-fertilization and developmental stages thereafter, and from RNA extracted from whole adult zebrafish (A). No product was detected in the negative control (NC) lacking RNA template in the RT-PCR (upper panel). At all stages of embryogenesis, a product specific for RACK1 mRNA was detected (lower panel).

embryo_stages_0-24hrs.pdf, Dr. Jonathan Bard, Anatomy Department, Edinburgh University, UK; see also [Westerfield, 1995]). By 24 h PF and at all later stages examined, B-FABP mRNA was detected. The temporal expression of the zebrafish B-FABP gene seen here correlates well with early development of the zebrafish brain. Similarly, in humans and other mammals, it has been shown that B-FABP is expressed at high levels in the developing CNS. The expression is also spatially and temporally correlated with neuronal migration and differentiation in radial glia, which support the differentiation and migration of developing neurons [Kurtz *et al.*, 1994; Feng *et al.*, 1994]. As stated previously, the expression of B-FABP in the brain of adult canary [Rousselot *et al.*, 1997] and fish [Denovan-Wright *et al.*, 2000] suggests a role for this protein in the neuronal migration and synaptic reorganization of adult avian and fish brain. The temporal expression of the B-FABP gene reported here (Fig. 4-4C) and our previous report of its expression in the periventricular grey zone of the optic tectum of adult zebrafish brain, a site of neurogenesis [Denovan-Wright *et al.*, 2000], further implicates B-FABP as playing a role in embryonic and adult neurogenesis.

### 4.2.5 Radiation hybrid mapping of the B-FABP to LG17

Using radiation hybrids, LN54 panel [Hukriede *et al.*, 1999], I mapped the zebrafish B-FABP (*fabp7*) gene to linkake group 17 (LG 17) at 21.11 cR or 1.05 cM (merged ZMAP panel) in the zebrafish genome with a LOD score of 16.2. (Fig. 4-5). The B-FABP gene is closely linked to the expressed sequence tag (EST) for myristoylated alanine-rich protein kinase C substrate (MACS) in the zebrafish linkage map. This linkage relationship is well conserved among zebrafish, mouse and human (Table 4-2). In the human cytogenetic map, the B-FABP gene (q22-q23) and MACS (q22.2) are also

closely linked (Table 4-2). Some of the other genes or ESTs that are syntenic with the B-FABP gene in zebrafish LG17 also have conserved syntenies in the human and mouse genomes. The genes for B-FABP, MACS and GNMT on zebrafish LG17 have conserved syntenies on human chromosome 6, but they are located on two linkage groups (LG10 and LG17) in the mouse genome, suggesting an inter-chromosome rearrangement of the surrounding region of B-FABP in the mouse genome after the divergence of fishes and mammals, and following the human-mouse divergence (Table 4-2). Interestingly, a similar syntenic relationship and its conservation among zebrafish, human and mouse has also been observed for another intracellular lipid-binding protein gene, CRBPII [Cameron *et al.*, 2002].

**Table 4-2. Conserved syntenic relationship of zebrafish B-FABP gene**

| **zebrafish**[a] | | **human**[c] | | **mouse**[c] | |
|---|---|---|---|---|---|
| Gene symbol | location[b] | Gene symbol | location | Gene symbol | location |
| Fabp7 | 17 40.9 cM | FABP7 | 6 q22-q23 | Fabp7 | 10 |
| Macs | 17 40.9 cM | MACS | 6 q22.2 | Macs | 10 22 cM |
| Gnmt | 17 57.2-62.1 cM | GNMT | 6 p12 | Gnmt | 17 |
| Pax9 | 17 37.2-49.8 | PAX9 | 14 q12-q13 | Pax9 | 12 26 cM |
| Foxa1 | 17 37.2-49.8 cM | FAXA1 | 14 q12-q13 | Faxa1 | 12 26 cM |
| Otx2 | 17 54.3-56.2 cM | OTX2 | 14 q21-q22 | Otx2 | 14 19 cM |
| Bmp4 | 17 67.7 cM | BMP4 | 14 q22-q23 | Bmp4 | 14 15 cM |
| Snap25b | 17 73.7 cM | SNAP25 | 20 p12-p11.2 | Snap25 | 2 78.2 cM |
| Bmp2a | 17 18.4 cM | BMP2 | 20 p12 | Bmp2 | 2 76.1 cM |

[a] Mapped ESTs by Woods *et al* [2000]; [b]ZMAP (http://zfin.org/cgi-bin/view_zmapplet.cgi), Zebrafish Information Network (ZFIN), the Zebrafish International Resource Center, University of Oregon, Eugene, USA; [c]LocusLink (http://www.ncbi.nlm.nih.gov/LocusLink/list.cgi), National Center for Biotechnology Information, U.S. National Library of Medicine, Bethesda, USA.

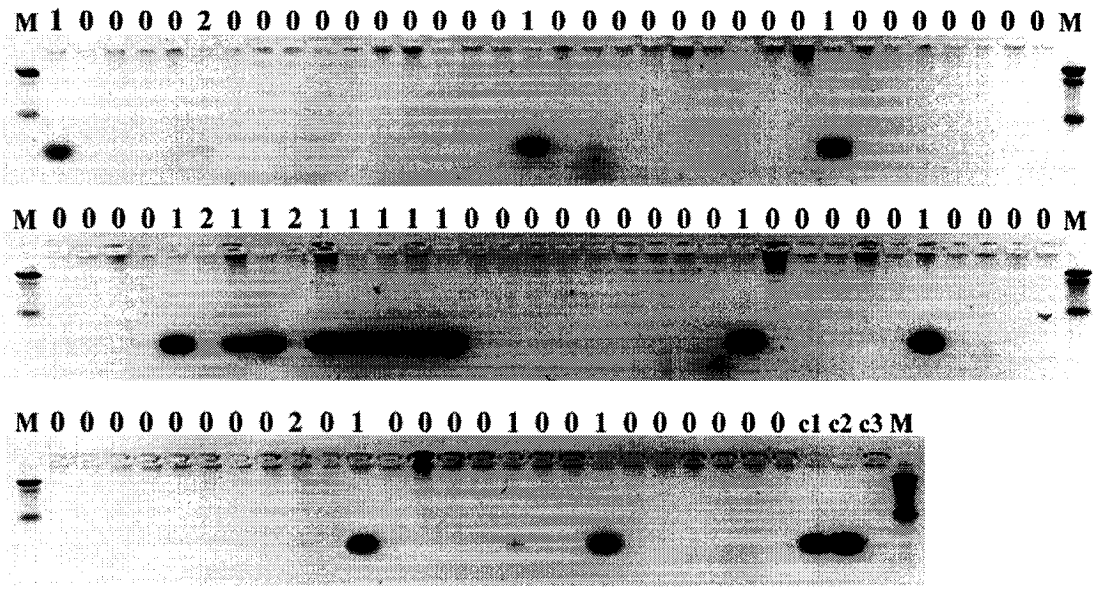



**Figure 4-5. Linkage group assignment by radiation hybrid mapping of the zebrafish B-FABP gene.** DNA from 93 mouse-zebrafish radiation cell hybrids of the LN54 panel was subjected to PCR using primers specific to the coding and 3′ UTR sequence of the fourth exon of the zebrafish B-FABP gene and the products subjected to size fractionation by gel-electrophoresis in 2% agarose (arrayed 1-93 from left to right, and top to bottom in the panels). The absence or presence of the expected 203 bp DNA fragment, or an ambiguous result, was scored 0, 1 or 2, respectively, to generate the RH vector (shown above each lane corresponding to the PCR reaction from a DNA of a given radiation cell hybrid) and submitted to http://mgchd1.nichd. nih.gov:8000/zfrh/beta.cgi for linkage group assignment. Genomic DNA from zebrafish (c1), a 1:10 mixture of zebrafish:mouse genomic DNAs (c2) and mouse genomic DNA (c3) were used as controls for the PCR. Molecular weight markers are shown at the right and left of each panel (M).

# 5 Identification of a Duplicated Gene for Zebrafish Brain-Type Fatty Acid-Binding Protein (*fabp7b*) and Differential Expression Patterns of *fabp7a* and *fabp7b* during Development

## 5.1 INTRODUCTION

Polyunsaturated long chain fatty acids, especially docosahexaenoic acid (DHA), play important roles in the development of vertebrate brain and visual system [Clandinin, 1999; Uauy *et al.*, 2001; Jeffrey *et al.*, 2001]. The distribution of DHA in the mammalian brain was determined early in 1960s, and thereafter the effects of DHA and other polyunsaturated long chain fatty acids on the development and function of the mammalian brain and visual system were studied extensively [reviewed by Salem *et al.*, 2001; Uauy *et al.*, 2001]. The precise sites, critical developmental periods and mechanism of DHA action in the developing brain and retina are, however, not well understood. It has been known that DHA, as a component of membrane lipid, has a role in regulation of the membrane ion channels, which is important for the cellular signal transduction processes of the neural transmitters and retinal photoreceptors [Salem *et al.*, 2001; Jeffrey *et al.*, 2001]. The effects of DHA on the biophysical properties of membranes can not fully explain its involvement in developmental programming and patterning of the neural and visual system.

In addition to being incorporated in membrane phospholipids, free fatty acids (FFA) are present in the cytoplasm. Over the past 10 years, intracellular long chain fatty acids have been recognized as crucial signaling molecules in gene regulation through binding and activating nuclear receptors, such as the peroxisome proliferator-activated receptor (PPAR) and retinoid x receptors (RXR) [reviewed in Duplus *et al.*, 2000; Clarke,

2001; de Urquiza *et al.*, 2000]. Recently, DNA microarray analysis of gene expression in human fetal retinal explants showed that DHA treatment altered the expression of various genes encoding proteins involved in neurogenesis, neurotransmission and refinement of neural connectivity [Rojas et a., 2003].

Intracellular fatty acid-binding proteins (FABPs) bind to FFAs and facilitate their uptake and transport within the aqueous cytoplasm. The biological functions of the intracellular FFAs are, therefore, mediated by FABPs [Glatz *et al.*, 1996; Bernlohr *et al.*, 1997]. Brain type fatty acid-binding protein (B-FABP) belongs to the large intracellular lipid-binding protein (iLBP) family, which includes FABPs and cellular retinoid-binding proteins [Feng *et al.*, 1994; Kurtz *et al.*, 1994]. The mammalian B-FABP shows highest affinity for DHA [Xu *et al.*, 1996]. B-FABP is abundant in the developing mouse brain and its distribution is spatially and temporally correlated with neuronal differentiation during the development of the central nervous system (CNS). B-FABP is present in both the cytoplasm and the nucleus [Feng *et al.*, 1994]. The levels of B-FABP mRNA are elevated in undifferentiated chick retina, suggesting that this protein is involved in early retinal development [Godbout, 1993].

The cDNA sequence, gene structure, linkage relationship and adult tissue distribution of the transcript encoding B-FABP from zebrafish (*Danio rerio*) were reported previously [Denovan-Wright *et al.*, 2000; Liu *et al.*, 2003a]. Here, I report the discovery of a duplicated B-FABP gene (*fabp7b*) from the zebrafish genome. The deduced amino acid sequence, gene structure and phylogenetic analysis suggested that I have identified a duplicated gene for B-FABP in zebrafish. While *fabp7b*, the duplicate copy of the B-FABP gene, was assigned to a separate linkage group (LG) from *fabp7a*

[Liu *et al.*, 2003a], both genes possess conserved syntenies to their mammalian orthologs. To explore the potential functions of the two B-FABP genes (*fabp7a* and *fabp7b*) during zebrafish embryogenesis and early development, I determined the spatio-temporal distribution of the *fabp7a* and *fabp7b* transcripts using embryo whole mount *in situ* hybridization. I detected abundant transcripts of the *fabp7a* gene in distinct regions of the early developing CNS and retina, indicating potential sites of DHA action that likely contribute to neurogenesis and patterning of the developing neural and visual system. The transcripts of the *fabp7b* gene were less abundant than those of *fabp7a* but in distinct sites in the developing brain and retina. Distribution of *fabp7b* mRNA in the developing swim bladder, which is a fish-specific organ, implies that neofunctionalization of the duplicate *fabp7b* gene has occurred in the fish lineage.

## 5.2 RESULTS

### 5.2.1 Identification of a duplicated gene for B-FABP (*fabp7b*) from the zebrafish genome

Using a genome search strategy, I identified a sequence with high sequence identity to the previously described zebrafish B-FABP gene, referred to hereafter as *fabp7a* [Denovan-Wright *et al.*, 2001; Liu *et al.*, 2003a]. I named the new B-FABP gene *fabp7b*, referring to a duplicate gene for the zebrafish B-FABP after analysis of its gene structure, linkage relationship, phylogeny and expression patterns (discussed below). The zebrafish *fabp7b* is a small gene spanning about 1.5 kb. The *fabp7b* gene consisted of four exons separated by three introns (Fig. 5-1), a gene structure shared by all of the iLBP multigene family in vertebrates [Bernlohr *et al.*, 1997]. The four exons of the zebrafish

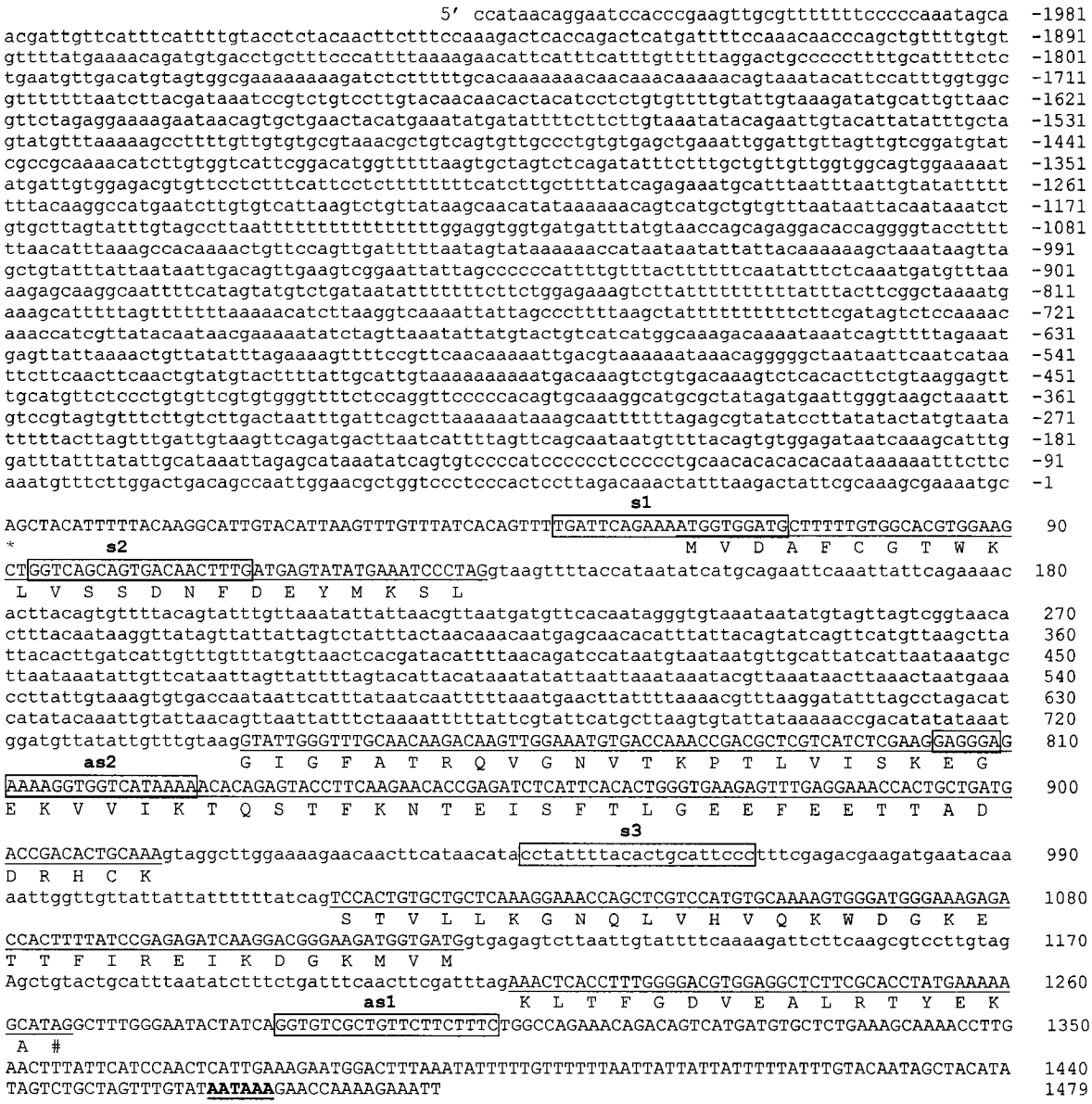



**Figure 5-1. Nucleotide sequence of the zebrafish *fabp7b* gene and its 5′ upstream region.** Exons are shown in uppercase letters with the coding sequences of each exon underlined and the deduced amino acid sequence indicated below. Numbers on the right indicate nucleotide positions in the gene sequence. The initiation site for transcription is marked by an asterisk and numbered at +1, and a putative polyadenylation signal is highlighted in bold font. PCR primer sequences (s1 and s2 for 3′ RACE; as1 and as2 for 5′ RACE; s2 and as1 for RT-PCR; s3 and as1 for radiation hybrid mapping for linkage group assignment) used in this study are boxed. The zebrafish *fabp7b* gene and its 5′ upstream sequence were identified from a DNA sequence assembly of clone CH211-199I15 (GenBank accession number: AL845421, Welcome Trust Sanger Institute).

*fabp7b* gene encoded 24, 58, 34 and 16 amino acids, respectively, which is identical to the coding capacity of each exon of the zebrafish *fabp7a* gene [Liu *et al.*, 2003a]. The size for each of the three introns of the zebrafish *fabp7b* gene was 608 bp, 105 bp and 94 bp, respectively. The introns in *fabp7b* were considerably smaller than those in *fabp7a* [Liu *et al.*, 2003a]. The introns were in the same position and obey the GT/AG, intron/exon splice junction rule in both *fabp7a* and *fabp7b* [Breathnach and Chambon, 1981].

A single product of the intact 5′ cDNA end with the $^{7}$methyl G cap for the *fabp7b* transcript was obtained using 5′ RLM-RACE (Fig. 5-2, lane 1). A negative control with tobacco acid pyrophosphatase (TAP) treatment omitted yielded no RACE product (Fig. 5-2, lane 2). Alignment of the 5′ RLM-RACE sequence with the zebrafish *fabp7b* gene sequence mapped a single transcription start site at the nucleotide position 60 bp upstream of the initiation codon (Fig. 5-1). A single transcription start site was previously determined for the zebrafish *fabp7a* gene [Liu *et al.*, 2003a], and the human [Bisgrove *et al.*, 2000] and mouse [Kurtz *et al.*, 1994] B-FABP genes. The size of the 5′ UTR of the zebrafish *fabp7b* gene (60 bp) was similar to that of the zebrafish *fabp7a* (69 bp) [Liu *et al.*, 2003a] and the human B-FABP gene (81 bp) [Bisgrove *et al.*, 2000].

In contrast to the zebrafish *fabp7a* gene, no TATA box was observed in the core promoter region of the *fabp7b* gene (Fig. 5-1). Inspection of the 2 kb 5′ upstream flanking sequence of the *fabp7b* gene with MatInspector professional 6.2 [Quandt *et al.*, 1995] revealed that several *cis* elements for the POU-domain transcription factors were present in this region (data not shown). POU-domain elements are essential for general CNS expression of the mouse B-FABP gene [Josephson *et al.*, 1998]. POU-

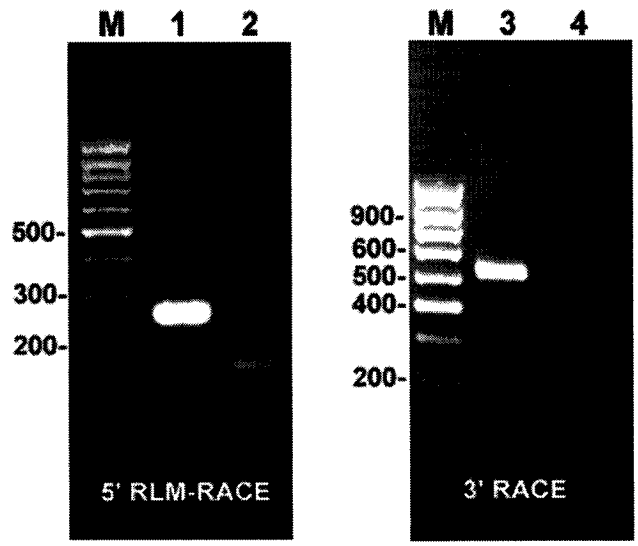



**Figure 5-2. Products of cDNA ends generated by 3′ RACE and 5′ RLM-RACE for cDNA cloning and transcription start site mapping**. Lane 1, product of 5′ RLM-RACE; Lane 2, negative control of 5′ RLM-RACE (RNA without tobacco acid pyrophos-phatase treatment); Lane 3, product of 3′ RACE; Lane 4, negative control of 3′ RACE (without the reverse transcription template). Lanes indicated M are 100 bp molecular weight markers. The size (in bp) of the molecular markers is indicated to the left of each panel.

domain elements are also present in the upstream flanking sequence of the zebrafish *fabp7a* gene [Liu *et al.*, 2003a]. The transcription factor NF1 binding sites in the human B-FABP gene promoter sequence [Bisgrove *et al.*, 2000] were not observed in the 5′ upstream region of the zebrafish *fabp7b* gene.

### 5.2.2 Isolation and characterization of the zebrafish *fabp7b* cDNA

3′ RACE with the zebrafish *fabp7b* gene-specific primers within exon 1 (Fig. 5-1) yielded a single product of the 3′ cDNA end with a short poly(A) tail (Fig. 5-2, lane 3). A negative control, which lacked the reverse transcription template, generated no RACE product (Fig. 5-2, lane 4). After combining the 3′ RACE and the 5′ RLM-RACE sequences, I determined the complete nucleotide sequence of the *fabp7b* transcript (GenBank accession number: AY380814). The size of the entire cDNA sequence was 672 bp, excluding the poly(A) tail. The cDNA encodes a polypeptide of 132 amino acids with an estimated isoelectric point of 6.4 and a molecular mass of 15 kDa. The coding sequence of the zebrafish *fabp7b* cDNA is identical to that of the corresponding *fabp7b* gene sequence. A T/A and a G/T transversion in the 5′ UTR was observed between the *fabp7b* cDNA and the gene sequences, which may represent polymorphisms or sequencing errors in either the 5′ RLM-RACE or the gene sequence. The coding sequence and the deduced amino acid sequence of the zebrafish *fabp7b* cDNA (B-FABPb) shared 73% and 82% identity with that of the zebrafish *fabp7a* (B-FABPa), respectively (Fig. 5-3). Both Arg-106 and Arg-126 are conserved in the amino acid sequences of B-FABPb and B-FABPa (indicated by stars in Fig. 5-3), amino acids known to be involved in fatty acid binding, and conserved among FABPs and CRABPs [Cheng *et al.*, 1991], Mutation of either of these two residues dramatically reduces ligand binding

**A**

```
AGCTACATTTTTACAAGGCTTTGTACAGTAAGTTTGTTTATCACAGTTTTGATTCAGAAAATGGTGGATGCTTTTTGTG
GCACGTGGAAGCTGGTCAGCAGTGACAACTTTGATGAGTATATGAAATCCCTAGGTATTGGGTTTGCAACAAGACAAGT
TGGAAATGTGACCAAACCGACGCTCGTCATCTCGAAGGAGGGAGAAAAGGTGGTCATAAAAACACAGAGTACCTTCAAG
AACACCGAGATCTCATTCACACTGGGTGAAGAGTTTGAGGAAACCACTGCTGATGACCGACACTGCAAATCCACTGTGC
TGCTCAAAGGAAACCAGCTCGTCCATGTGCAAAAGTGGGATGGGAAAGAGACCACTTTTATCCGAGAGATCAAGGACGG
GAAGATGGTGATGAAACTCACCTTTGGGGACGTGGAGGCTCTTCGCACCTATGAAAAAGCATAGGCTTTGGGAATACTA
TCAGGTGTCGCTGTTCTTCTTTCTGGCCAGAAACAGACAGTCATGATGTGCTCTGAAAGCAAAACCTTGAACTTTATTC
ATCCAACTCATTGAAAGAATGGACTTTAAATATTTTTGTTTTTTAATTATTATTATTTTTATTTGTACAATAGCTACAT
ATAGTCTGCTAGTTTGTATAATAAAGAACCAAAAGAAATT(A)n
```

**B**

```
                0                   20                  40                  60
                |                   |                   *                   |
ZF B-FABPb      MVDAFCGTWKLVSSDNFDEYMKSLGIGFATRQVGNVTKPTLVISKEGEKVVIKTQSTFKNTEISFTLG
ZF B-FABPa      ......A.....D.Q..........V.............I...H..D......L..........K..

                          80                  100                 120
                          |                   |       *           |       *
ZF B-FABPb      EEFEETTADDRHCKSTVLLKGNQLVHVQKWDGKETTFIREIKDGKMVMKLTFGDVEALRTYEKA  100%
ZF B-FABPa      ...D........V....S.E.DN..Q..R......K.V..........T...EG.Q.V......   82%
```

**Figure 5-3. Nucleotide sequence of the zebrafish *fabp7b* cDNA and alignment of its deduced amino acid sequence with that of the zebrafish *fabp7a*.** (**A**) Complete nucleotide sequence of the zebrafish *fabp7b* cDNA was determined by 3′ RACE and 5' RLM-RACE (GenBank accession number: AY380814). The initiation codon and stop codon of the cDNA sequence were boxed and the polyadenylation signal sequence was in bold. (**B**) The deduced amino acid sequence of the zebrafish B-FABP was aligned with the amino acid sequence of the zebrafish B-FABPa (GenBank accession number: AAF79948). Dots indicate amino acid identity. The percentage amino acid sequence identity between the zebrafish B-FABPb and B-FABPa is shown at the end of the sequence. Positions of certain amino acids are indicated by numbers above them and the three residues important for the fatty acid ligand binding are indicated by stars.

of the murine BLBP [Xu *et al.*, 1996]. Residue, Gly-33, conserved among the zebrafish and mammalian B-FABPs (Fig. 5-3), is adjacent to the opening of the murine BLBP cavity, and mutations at this position reduce ligand binding presumably by preventing ligand access to the binding pocket of BLBP [Xu *et al.*, 1996].

### 5.2.3 Phylogenetic analysis suggested a fish-specifc duplication of the *fabp7* gene

To reveal the evolutionary relationship of the zebrafish *fabp7b* gene with other identified vertebrate iLBP genes, I conducted phylogenetic analysis (Fig. 5-4). A bootstrap neighbour-joining phylogenetic tree was constructed with ClustalX [Thompson *et al.*, 1997] using the nematode lipid-binding protein sequence as an outgroup to root the tree. The zebrafish B-FABPb clustered with the zebrafish B-FABPa and all the orthologous mammalian B-FABPs in the same clade (bootstrap value1000/1000; Fig. 5-4). The zebrafish B-FABPa and B-FABPb formed a specific sub-clade from the mammalian B-FABP group (Fig. 5-4), suggesting that they are sister duplicated genes that arose from a zebrafish-specific duplication event.

### 5.2.4 The zebrafish *fabp7b* gene was assigned to LG 20

Using radiation hybrid panel, the zebrafish *fabp7b* gene was mapped to LG 20 of the zebrafish genome with a mapping distance of 4 cR from the marker unp1843 (scoring data can be provided on request). In contrast, zebrafish *fabp7a* is on LG 17 [Liu *et al.*, 2003a]. Integrating our mapping data with the existing synteny maps [Woods *et al.*, 2000; Liu *et al.*, 2003a] and data obtained from LocusLink (http://www.ncbi.nlm.nih.gov/locuslink), I found that zebrafish *fabp7b*, like zebrafish *fabp7a* [Liu *et al.*, 2003a], had conserved syntenic gene loci with the human *FABP7* on human chromosome 6 and the

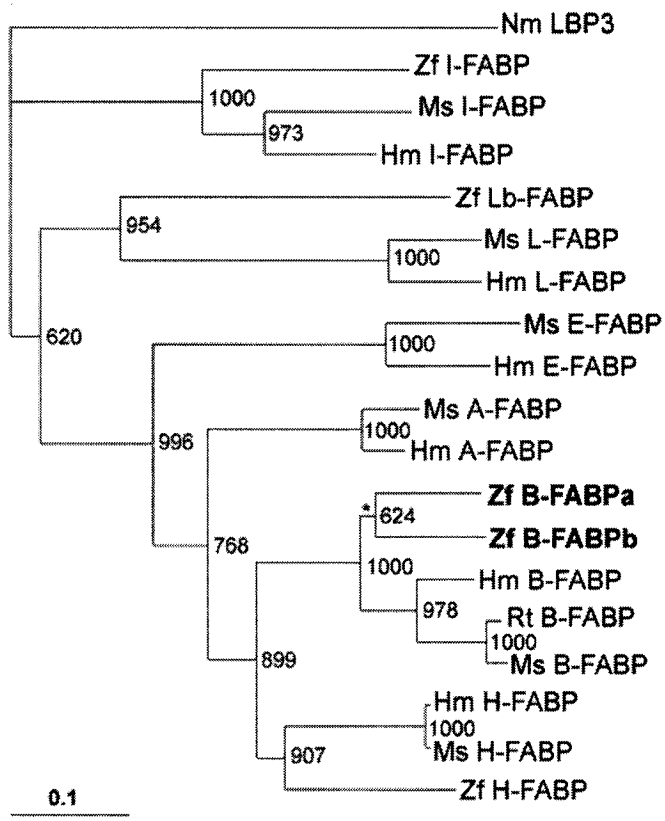



**Figure 5-4. Phylogenetic tree of the vertebrate iLBPs.** The Bootstrap neighbour-joining phylogenetic tree was constructed with ClustalX [Thompson *et al.*, 1997] using the nematode lipid-binding protein 3 (GenBank accession number NP_508556) as an outgroup. The bootstrap values (based on number per 1000 duplicates) are indicated on the nodes. Sequences used in this analysis in addition to zebrafish B-FABP7a (Zf B-FABP7a) and B-FABP7b (Zf FABP7b) include the amino acid sequences of the zebrafish intestinal-type FABP (Zf I-FABP, GenBank accession number: CAB64945), liver-basic type FABP (Zf Lb-FABP, NP_694492), heart type FABP (Zf H-FABP); human (Hm) I-FABP (Hm I-FABP, NP_000125), L-FABP (Hm L-FABP, AAA52419), epidermal type FABP (E-FABP, AAH19385), adipocyte FABP (Hm A-FABP, AAH03672), B-FABP (Hm B-FABP, NM_001446), H-FABP (Hm H-FABP, NP_004093); mouse (Ms) I-FABP (Ms I-FABP, P55050), L-FABP (Ms L-FABP, NP_059095), E-FABP (Ms E-FABP, NP_034764), A-FABP (Ms A-FABP, P04117), B-FABP (Ms B-FABP, NM_021272), H-FABP (Ms H-FABP, NP_034304) and rat B-FABP (Rt B-FABP, NM_030832). Scale bar = 0.1 substitutions per site.

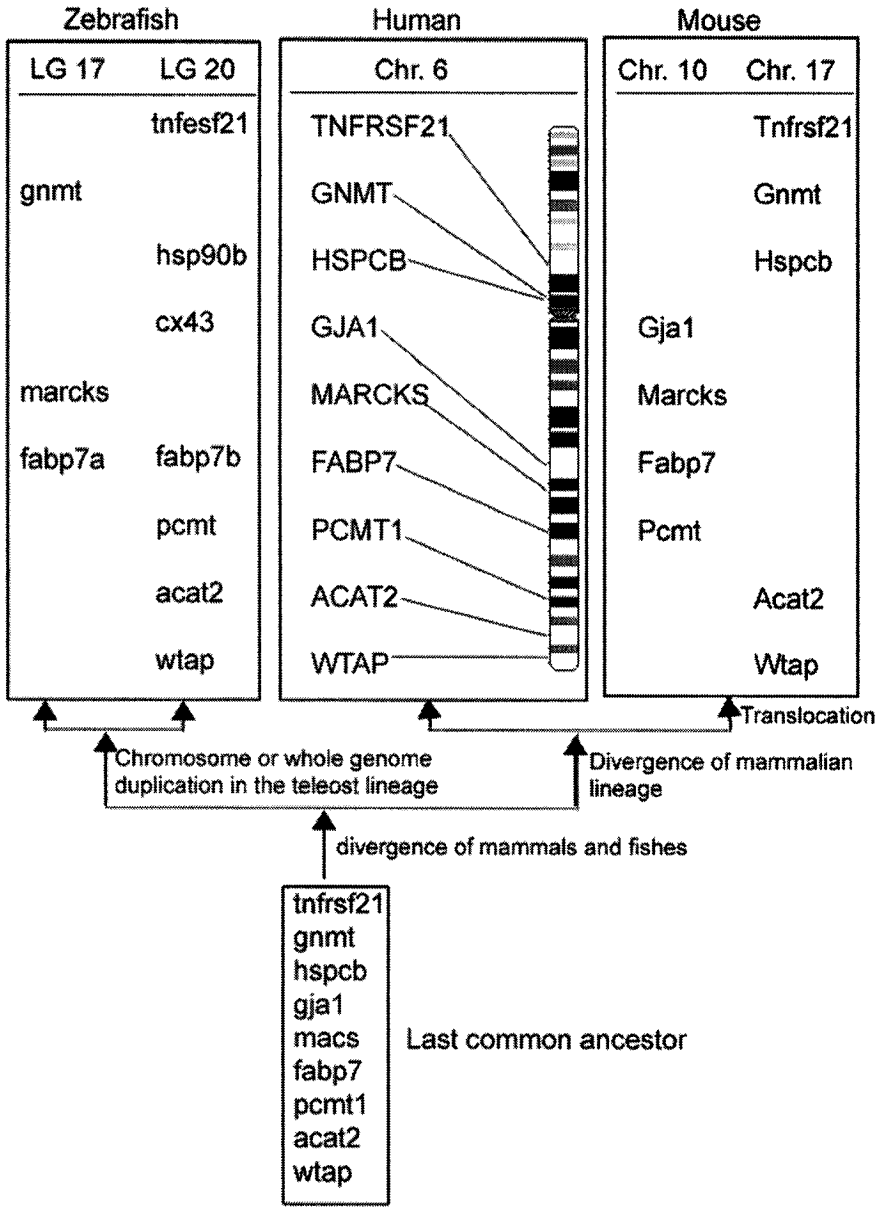



**Figure 5-5. Comparative syntenies with *fabp7* genes from zebrafish, human and mouse**. Both the zebrafish *fabp7a* gene on LG 17 and *fabp7b* gene on LG 20 have conserved syntenies (left box) with the human *FABP7* gene on chromosome 6 (middle box) and the mouse *Fabp7* gene on chromosome 10 and 17 (right box), which suggests that the zebrafish LG 17 and 20 arose from duplication of a chromosome corresponding to the human chromosome 6. The order of the human FABP7 gene and its syntenic genes was determined based on the cytogenetic mapping data from LocusLink (http://www.ncbi.nlm.nih.gov/locuslink) and the gene loci on the zebrafish and mouse LGs and the last common ancestor's chromosome are listed in the order appearing on the human chromosome 6. A putative evolutionary pathway leading to the syntenic relationships of the zebrafish, human and mouse *Fabp7* genes is indicated by arrows below the panels.

mouse *Fabp7* on LG 10 (Fig. 5-5), indicating they are all orthologous genes arising from the same ancestral gene.

### 5.2.5 Tissue-specific distribution of *fabp7b* transcripts in adult zebrafish

RT-PCR was performed to reveal the distribution pattern of *fabp7b* mRNA in adult zebrafish tissues. RT-PCR products were generated from mRNA extracted from the liver, intestine, brain and testis of adult zebrafish, but no product was observed from mRNA of the ovary, heart, and muscle (Fig. 5-6). The zebrafish *fabp7b* mRNA was detected by RT-PCR in the same set of adult tissues as *fabp7a*, although the relative abundance of the resulting PCR products from each tissue was considerably different (Fig. 5-6). The RT-PCR used was not quantitative, but the very faint RT-PCR product detected with the *fabp7b* cDNA-specific primers in the adult zebrafish brain indicated that low levels of *fabp7b* mRNA are present in this tissue. Tissue section *in situ* hybridization with a *fabp7b* gene-specific oligonucleotide probe was performed on several sections of adult zebrafish, however, no hybridization signal was observed (data not shown), indicating low levels of expression of the zebrafish *fabp7b* gene in adult zebrafish tissue.

### 5.2.6 Specific distribution of the zebrafish *fabp7a* mRNA in the developing CNS

To demonstrate the spatial and temporal distribution of *fabp7a* transcripts during zebrafish embryonic and larval development, whole mount *in situ* hybridization on zebrafish embryos and larvae at different developmental stages was performed (Fig. 5-7). Transcripts of the *fabp7a* gene were first detected in embryos at the late gastrulation and early segmentation stage but the hybridization signals were not spatially restricted (data not shown). At 17 hours post fertilization (hpf), *fabp7a* mRNA was widely distributed in

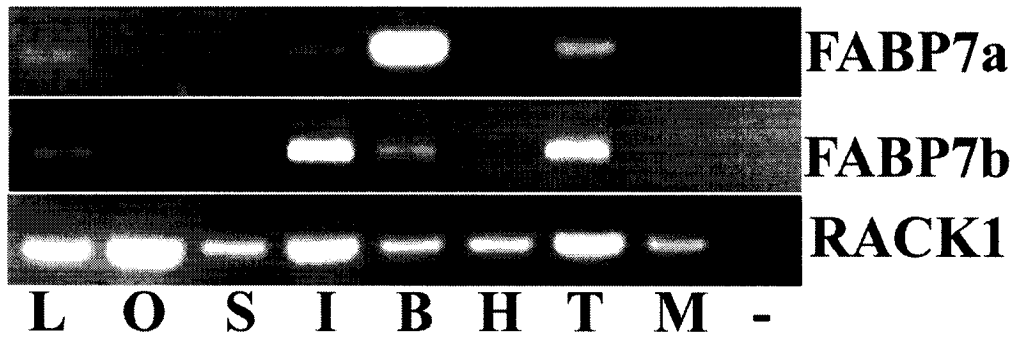



**Figure 5-6. Tissue-specific distribution of the *fabp7a* and *fabp7b* transcripts in adult zebrafish detected by RT-PCR.** The zebrafish *fabp7b* cDNA-specific primers amplified a PCR product from total RNA extracted from the liver (L), skin (S), intestine (I), brain (B) and testis (T), but not in the ovary (O), heart (H) and muscle (M). The zebrafish *fabp7a* cDNA-specific primers detected mRNA in nearly the same set of tissues but the relative abundance of the RT-PCR products in the brain is much higher than that of the *fabp7b* transcripts. A RT-PCR product corresponding to the constitutively expressed, receptor for activated C kinase (RACK1) was generated from RNA in all samples as a positive control. A negative PCR control (-) did not contain cDNA template and did not generate any RT-PCR product.

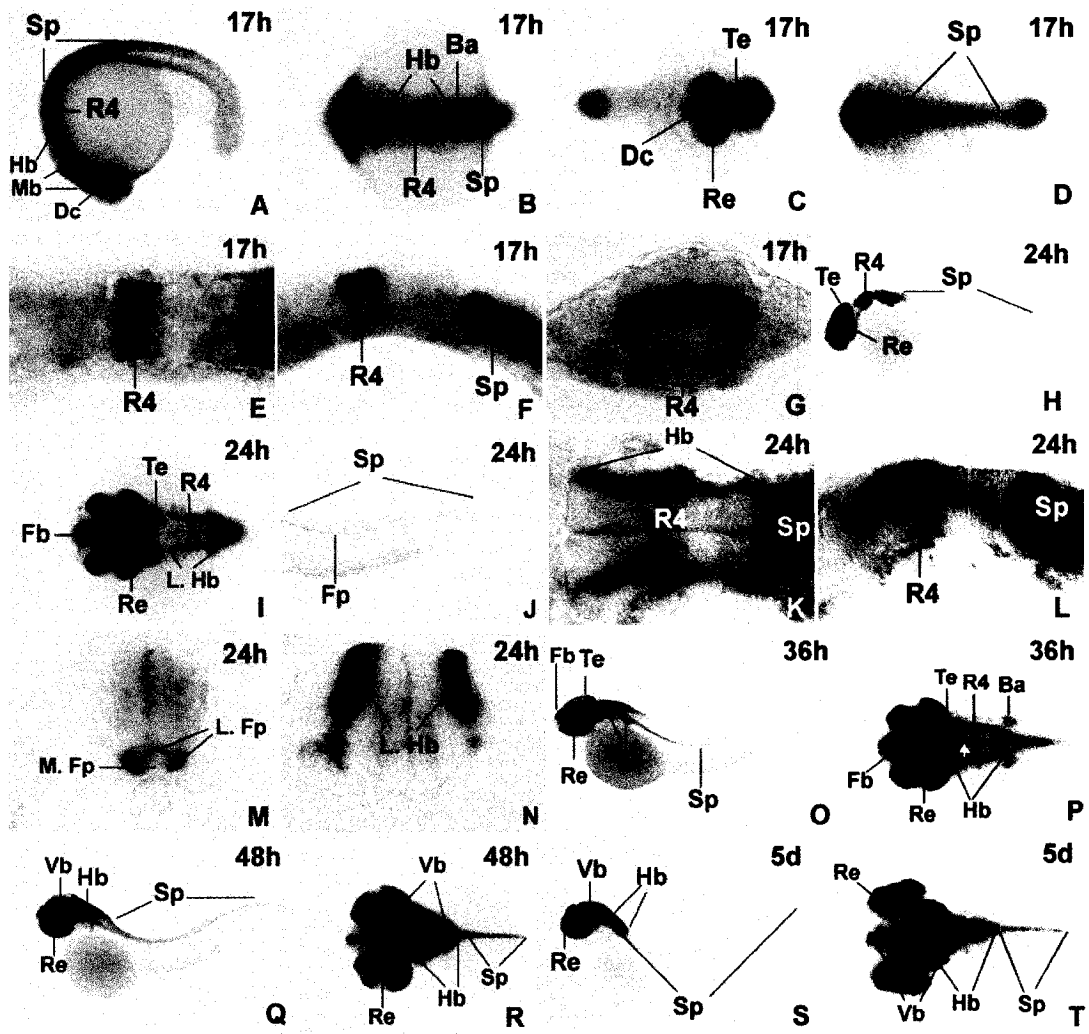



**Figure 5-7. Spatio-temporal mRNA distribution of the zebrafish *fabp7a* gene in the early developing CNS.** Lateral (A, F, H, J, L, O, Q, S), dorsal (B-E, I, K, P, R, T) and cross section (G, M, N) views of embryos and larvae showing the localization of *fabp7a* transcripts at the stages indicated. (A-D) embryos showing *fabp7a* mRNA distribution in the diencephalon (Dc), anterior midbrain (Mb), tectum (Te), anterior hindbrain (Hb), rhomobomere 4 (R4), spinal cord (Sp), tectum (Te), branchial arches (Ba) and retina (Re) at the middle segmentation stage (17 hpf). (E-G, K, L) Zebrafish embryonic hindbrain with restricted *fabp7a* mRNA distribution in R4 during 17 to 24 hpf. From 24 hpf embryos to 5 dpf larvae, *fabp7a* mRNA was detected in the lateral floor plate (L. Fp) of the spinal cord, lateral hindbrain (L. Hb), ventricular zone of brain (Vb, indicated by an arrow in P) and the retina (H-T). Transcripts of *fabp7a* were also observed in the ventricular zone of the forebrain (Fb) at 24 and 36 hpf stages (I, O, P).

the developing CNS, including the ventral retina (Fig. 5-7 A, C), diencephalon (Fig. 5-7 A, C), anterior midbrain (Fig. 5-7 A), optic tectum (Fig. 5-7 C), hindbrain and spinal cord (Fig. 5-7 A, B, D). In the hindbrain, *fabp7a* mRNA was abundant and restricted to the middle region of rhombomere 4 (Fig. 5-7 B, E-G). At 24 hpf, hybridization signals for *fabp7a* transcripts were observed in the retina, forebrain, tectum, hindbrain and spinal cord (Fig. 5-7 H-J). In the hindbrain, the abundance of *fabp7a* mRNA was reduced compared to 17 hpf in rhombomere 4 and predominantly located at the lateral region (Fig. 5-7 I, K, L). *fabp7b* mRNA was distributed throughout the whole lateral hindbrain at 24 hpf stage (Fig. 5-7 I, K, N). A cross section of the spinal cord at 24 hpf revealed a restricted distribution of *fabp7a* mRNA in the lateral floor plate of the developing spinal cord (Fig. 5-7 M). The spatial distribution of the *fabp7a* mRNA observed at 36 hpf (Fig. 5-7 O, P), 48 hpf (Fig. 5-7 Q, R) and 5 dpf (days post fertilization) larvae (Fig. 5-7 S, T) were similar, and the signals retained at all those regions showed *fabp7a* mRNA at 24 hpf. From 36 hpf on, however, the *fabp7a* mRNA was detected in the ventricular zone of the whole brain (Fig. 5-7 P, R, T). *fabp7a* mRNA was present in the ventral forebrain at 24 and 36 hpf (Fig. 5-7 I, O, P).

### 5.2.7 Spatio-temporal distribution of *fabp7b* transcripts during development

*fabp7b* transcripts were detected in the developing brain, but at lower levels and at later stages than *fabp7a* transcripts. *fabp7b* mRNA was first observed in a small domain of the dorsal diencephalon and pretectum of the 24 hpf embryo (Fig. 5-8 A, B). At 36 hpf, the hybridization signal in the dorsal diencephalon was extended and *fabp7b* mRNA was detected in the dorsal lateral hindbrain (Fig. 5-8 C, D). The *fabp7b* mRNA distribution in the dorsal diencephalons and dorsal and lateral hindbrain became more

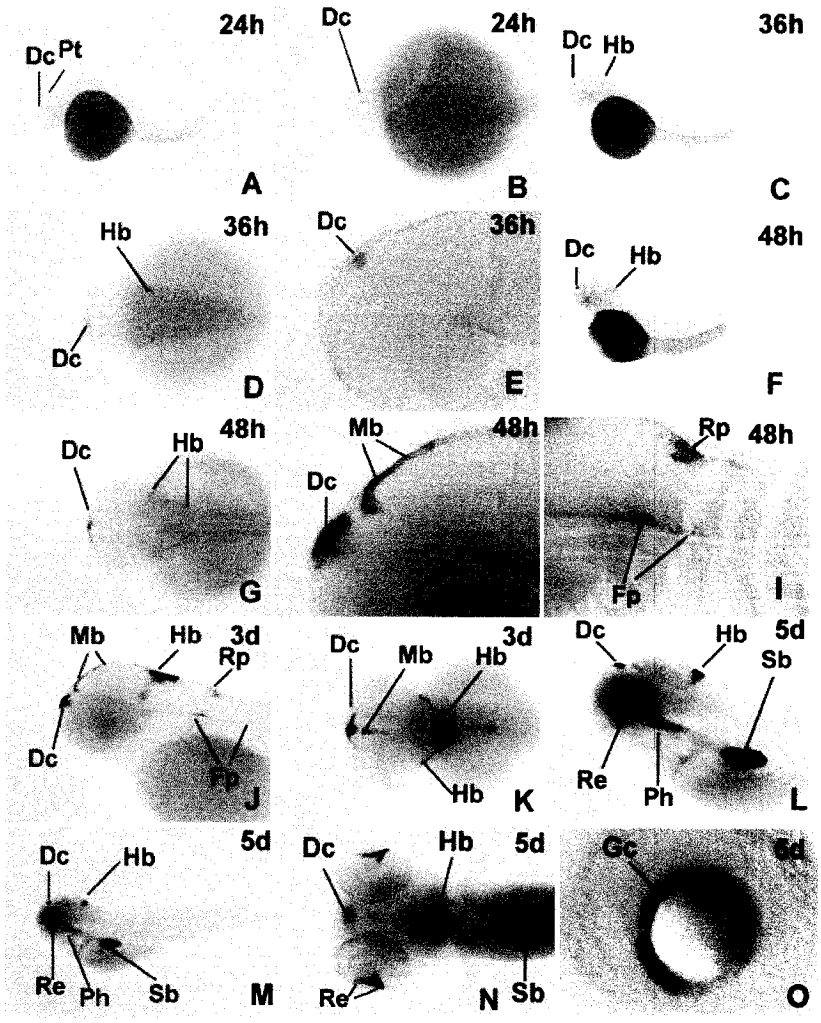



**Figure 5-8. Spatio-temporal distribution of the zebrafish *fabp7b* transcripts during early development.** Lateral (A, C, E, F, H-J) and dorsal (B, D, G, K) views of zebrafish embryos and larvae showing *fabp7b* mRNA distribution in the diencephalons (Dc), pretectum (Pt), hindbrain (Hb), midbrain (Mb) and the anterior spinal roof plate (Rp) and floor plate (Fp) during development of 24 hpf to 3 dpf. Lateral view (L, M) and dorsal view (N) of 5 dpf larvae showing differentiated mRNA expression in the developing swim bladder (Sb) and pharynx (Ph) in addition to the developing retina (Re) and brain. (O) Magnified lateral view of the 5 dpf retina showing specific localization of the zebrafish *fabp7b* mRNA in the ganglion cell layer (Gc).

intense at 48 hpf (Fig. 5-8 F-H) and 3 dpf (Fig. 5-8 J, K). At the same stage, *fabp7b* mRNA was detected in the dorsal midbrain (Fig. 5-8 H, J, K) and anterior roof and floor plate (Fig. 5-8 I, J). At 5 dpf stage, distribution of the *fabp7b* mRNA in the diencephalon and hindbrain was restricted to a small region, and abundant levels were detected in a region associated with the swim bladder and the ganglion cell layer of the retina (Fig. 5-8 L-O). Hybridization signal also appeared in the pharynx at this stage (Fig. 5-8 M).

## 5.3 DISCUSSION

### 5.3.1 B-FABP and vertebrate brain development

The development of the vertebrate CNS follows a specific spatial and temporal order. It has been proposed that several organizing centers in the early developing vertebrate CNS are involved in local patterning of axis formation in CNS, including the midbrain-hindbrain boundary [reviewed by Rhinn and Brand, 2001; Wurst and Bally-Cuif, 2001], the forebrain organizer [Shimamura and Rubenstein, 1997; Houart *et al.*, 1998] and the hindbrain-spinal cord border [Maden, 2002]. However, the molecular mechanisms involved in signaling neuronal and glial differentiation from organizing centers are not well understood. Isolation of genes with restricted patterns of expression in these regions followed by mutant screening is a powerful approach toward understanding the signaling pathways governing axis patterning in vertebrate CNS. A number of genes, most prominently the *Hox* family, play a crucial role in specifying the anterior-posterior (A-P) identity in the neural plate [Krumlauf, 1994].

B-FABP was isolated from the mammalian brain nearly a decade ago [Kuhar *et al.*, 1993; Kurtz *et al.*, 1994], and a study of the spatio-temporal distribution of the

mammalian B-FABP suggests that it might represent a signaling system in the developing mammalian CNS [Feng *et al.*, 1994]. A detailed expression profile of the gene encoding B-FABP and direct evidence of its function in neurogenesis and patterning in the vertebrate CNS is, however, still not available. The zebrafish embryo provides a model for gene expression studies in the developing CNS because its yolk granules are sequestrated in a single, large yolk cell, making the whole embryo optically transparent [Westerfield, 1995]. Using whole mount *in situ* hybridization of zebrafish embryos, I obtained distinct and detailed mRNA distribution profiles of the two duplicated *fabp7* genes at different stages of early developmental, and I revealed some important expression features in the developing CNS which were not evident from expression studies of the mammalian B-FABPs [Feng *et al.*, 1994; Kurtz *et al.*, 1994]. In the zebrafish embryonic hindbrain, I detected *fabp7a* mRNA in a restricted region of rhombomere 4 (r4) during the middle and late segmentation phase, suggesting an important role of the *fabp7a* product in patterning the hindbrain development. The vertebrate hindbrain consists of seven subdivided segments, called rhombomeres, each of which has distinct cellular and molecular characteristics [see review by Lumsden and Krumlauf, 1996]. During the development of the hindbrain in zebrafish and other vertebrates, r4 is the first rhombomere to form and acts as a signaling center in hindbrain patterning [Graham *et al.*, 1993; Graham and Lumsden, 1996; Maves *et al.*, 2002]. Transplantation studies in chick embryos have shown that r4 regulates neural crest cell death in r3 and r5 [Graham *et al.*, 1993]. In a recent study in zebrafish, transplantation of r4 cells at bud/1-somite stage induces expression of r5/r6 markers and the signaling activity is mediated by the fibroblast growth factor (FGF) [Maves *et al.*, 2002]. I found in

the present study that abundant *fabp7a* transcripts were transiently expressed in the early stages of r4 development, implying that B-FABPa might be involved in r4 signaling activity. As B-FABP is present in the nucleus as well as the cytoplasm [Feng *et al.*, 1994] and DHA, its potential ligand *in vivo* [Xu et a., 1996], is an activator of the nuclear receptor, RXR [ed Urquiza *et al.*, 2000], the *fabp7a* gene product might mediate expression of genes essential for r4 patterning activity through the DHA-RXR-target gene pathway. It would be interesting to define the interrelationships among r4-specific genes (e.g. *fgf*, *hox*, *fabp7*, etc.) in the whole cascade of r4 signaling through gene function analysis.

### 5.3.2 *fabp7a* gene expression in the spinal cord

The vertebrate spinal cord arises from the neural plate at the gastrula stage of embryogenesis. Differentiation of the ventral region of the spinal cord and the floor plate requires signals from the underlying axial mesoderm [Lumsden and Krumlauf, 1996]. The differentiated floor plate itself acts as a major signaling center, guiding the growth of axons along and across the midline of the spinal cord [reviewed by Colammarino and Tessier-Lavigne, 1995; Stoeckli and Landmesser, 1998]. In the developing spinal cord of zebrafish, the floor plate consists of a single row of medial floor plate (MFP) cells flanked by lateral floor plate (LFP) cells [Bernhardt *et al.*, 1992; Odenthal *et al*, 2000]. Similar to the MFP cells, the LFP cells play a role in axon guidance [Kuwada etv al., 1990]. Although these two cell populations share expression of many genes, they also have unique patterns of gene expression [Odenthal *et al.*, 2000 and references therein]. I found that the zebrafish *fabp7a* transcripts were specifically located in the LFP region, but not in the MFP region of the developing spinal cord, suggesting that B-FABP

mediates LFP signaling activity. Again, the B-FABP ligand, DHA, might be a potential signaling molecule in the LFP region. B-FABP could serve as a specific molecular marker for LFP cells in the developing vertebrate neural tube.

### 5.3.3 B-FABP, DHA and vertebrate retinal development

The polyunsaturated fatty acid, DHA, is abundant in the retina and is essential for development of the vertebrate visual system [Fliesler and Anderson, 1983; Uauy *et al.*, 2001]. The effects of DHA deficiency on retinal development and function have been studied intensively [reviewed by Uauy *et al.*, 2001; Jeffrey *et al.*, 2001; Salem *et al.*, 2001]. Animal experiments showed that inadequate supply of DHA at a certain developmental stage might result in permanent dysfunction, indicating existence of critical periods of DHA accretion for normal retinal development and function in animals [reviewed by Jeffrey *et al.*, 2001]. However, the mechanisms of DHA action in developing visual system are still not known. Spatio-temporal expression studies of the B-FABP gene during embryogenesis and later development might be a fruitful approach toward identification of the sites and timing of DHA action in the developing visual system. The mRNA levels of B-FABP increase in the undifferentiated chick retina but decreases upon retinal maturation, providing evidence of a role for B-FABP in retinal development [Godbout, 1993]. In the present study, I found abundant transcripts of the zebrafish *fabp7a* gene in the developing retina from early embryonic stages to 5 dpf larvae, while *fabp7b* mRNA was detected specifically in the ganglion cell layer of the 5 dpf larval retina. *fabp7a* mRNA was also observed in the early developing optic tectum. To our knowledge, our data provide the first detailed spatio-temporal expression of the *fabp7* genes in the developing vertebrate visual system, which indicates the time and sites

of DHA action in the developing visual system. Relating our data to the results of B-FABP mRNA expression in the undifferentiated chick retina [Godbout, 1993], DHA effects on retinal development and function from animal studies [reviewed by Uauy *et al.*, 2001; Jeffrey *et al.*, 2001; Salem *et al.*, 2001] and the evidence of DHA involvement in regulation of gene expression [ed Urquiza *et al.*, 2000; Rojas *et al.*, 2003], I propose that vertebrate retinal development is mediated by B-FABP through DHA action on regulation of crucial genes associated with retinal development and function.

### 5.3.4 Tandem gene duplication or whole genome duplication

With the identification of genes and mapping of zebrafish and pufferfish genomes, it appears that many fish genes and gene clusters have been duplicated [Gates *et al.*, 1999; Postlethwait *et al.*, 2000; Barbazuk *et al.*, 2000; Woods *et al.*, 2000; Taylor *et al.*, 2003]. It has been hypothesized that a whole genome duplication specific to the ray-finned fish lineage occurred following the divergence of ray-finned and lobe-finned fishes [Taylor *et al.*, 2003 and references therein]. The vertebrate iLBPs are encoded by a multigene family and their gene loci are dispersed throughout the vertebrate genome [reviewed by Bernlohr *et al.*, 1997]. I identified here a duplicated *fabp7* gene in the zebrafish genome, and based on phylogenetic analysis, the duplication appears to have occurred after divergence of mammals and fishes. Each of the duplicate *fabp7* genes was mapped to different linkage groups, and each had conserved syntenies with human and mouse orthologs. As such, the data support the fish-specific genome duplication theory rather than tandem gene duplication [Taylor *et al.*, 2001; 2003].

### 5.3.5 Function and fate of duplicated genes

The neofunctionalization theory of duplicated genes states that new patterns of expression or novel favorable functions might be a major mechanism for the preservation of duplicated genes in the host genome [Cooke *et al.*, 1997]. The duplication-degeneration-complementation (DDC) model was recently proposed for the mechanism of duplicate gene preservation. The DDC model predicts that after gene duplication, mutations in the upstream *cis* regulatory elements of both duplicates result in partitioning of the ancestral functions into each of the duplicate genes (subfunctionalization theory) [Force *et al.*, 1999]. Although a number of recent studies on expression patterns of duplicate genes support the DDC model [Force *et al.*, 1999; de Matino *et al.*, 2000; Serluca *et al.*, 2001], Lynch *et al.* [2001] suggest that neofunctionalization may be a more important mechanism for preservation of duplicate genes than subfunctionization in populations consisting of large number of individuals. Our present study demonstrated that *fabp7a* was expressed during development and adulthood similar to that observed for its mammalian orthologs, whereas the *fabp7b* gene showed only low levels of expression in the developing dorsal brain. Abundant transcripts of *fabp7b* were, however, detected in the developing swim bladder, an organ unique to fishes. During 3 to 5 dpf, zebrafish undergo a critical embryonic to larval transition, marked by major morphological and physiological changes including the inflation of the swim bladder, and structural and functional maturation of the mouth and gastrointestinal tract [Kimmel *et al.*, 1995; Pack *et al.*, 1996]. It appears that *fabp7b* acquired a novel function in the developing swim bladder and pharynx that affects this particular transitory stage of development. The expression patterns of the zebrafish *fabp7* gene duplicates suggest that quantitative

subfunctionalization and neofunctionalization account for retention of the duplicate *fabp7* genes in the zebrafish genome.

# 6 Characterization of a Cellular Retinol-Binding Protein Gene (*rbp1*) From Zebrafish (*Danio rerio*): Distinct Distribution of *rbp1* and *rbp2* Gene Transcripts during Development and Adulthood

## 6.1 INTRODUCTION

Retinoids, derived from retinol or vitamin A (VA), play crucial roles in controlling vertebrate growth, development and reproduction. Retinoid exerts its biological action mainly through its active metabolite, retinoic acid, which binds and activates two types of nuclear receptors, the retinoic acid receptors (RAR) and retinoid x receptors (RXR) [Mangelsdorf, 1994]. RAR and RXR are widely distributed in both mammalian and fish tissues. RAR is activated by all-*trans* and 9-*cis* retinoic acids, while RXR is only activated by 9-*cis* retinoic acid [Heyman *et al.*, 1992; Levin *et al.*, 1992]. These activated nuclear receptors in turn serve as transcription factors that regulate the expression of target genes essential to the development and function of the nervous, reproductive and immune system as well as skin and bone throughout the life of vertebrates [Mangelsdorf, 1994; Giguere, 1994; Napoli, 1996].

Hormones, biosynthetic and catabolic enzymes, and plasma and intracellular binding proteins have been implicated in the retinoid signaling pathway [Napoli, 1996; 1999; Noy, 2000]. The cellular retinol-binding proteins (CRBPs) and cellular retinoic acid-binding proteins (CRABPs) are members of the intracellular lipid-binding protein (iLBP) multigene family. CRBPs and CRABPs are small molecular mass (~15 kDa) proteins that specifically bind retinol and retinal (CRBPs) and retinoic acid (CRABPs) [Ong *et al.*, 1994]. CRBPs and CRABPs play important roles in intestinal VA absorption

and cellular retinoid transport. Cellular retinoid-binding proteins also modulate the metabolism and homeostasis of retinoids through interaction with various metabolic enzymes [Noy, 2000; Lohnes *et al.*, 1992].

The mammalian CRBPs and CRABPs and their genes have been studied intensively, although their functions *in vivo*, and regulation of their gene expression are not well defined. In addition to the well studied mammalian CRBPI and CRBPII, two novel types of CRBP, CRBPIII and CRBPIV, have recently been described in human and mouse. Each CRBP is encoded by a distinct gene and shows differential patterns of expression in mammals [Folli *et al.*, 2001; 2002; Vogel *et al.*, 2001].

Knowledge of CRBPs and their genes in fishes, the largest and most evolutionary diverse group of vertebrates, is limited. Thus far, the structure of recombinant protein of CRBPII from zebrafish has been reported [Calderone *et al.*, 2002]. Cameron *et al.*, [2002] previously determined the cDNA sequence, linkage localization and adult mRNA tissue-specific distribution of CRBPII in zebrafish. Here, I describe another CRBP from zebrafish (*Danio rerio*). Conservation of primary amino acid sequence, gene structure and syntenic relationship between the homologous genes in zebrafish and mammals, and phylogenetic analysis suggest this zebrafish CRBP is the ortholog of the mammalian CRBPI gene (*rbp1*). The tissue distribution of CRBPI mRNA revealed a distinct expression profile that correlates to RA distribution and action in the developing central nervous system of zebrafish. I also report the gene structure and spatio-temporal mRNA distribution of the zebrafish *rbp2* gene during development. Whole mount *in situ* hybridization showed that *rbp2* transcripts were abundant in the developing intestinal

bulb during embryogenesis and later the developing intestine, suggesting a novel function of this gene in intestinal organogenesis and development.

## 6.2 RESULTS

### 6.2.1 Cloning and analysis of the zebrafish CRBPI cDNA

After the nested PCR amplification of the primary 3′ RACE product with an internal "sense" primer, a single PCR product of ~1.1 kb was observed following agarose gel-electrophoresis (Fig. 6-1). The cDNA sequence for the CRBPI, determined by sequencing of three randomly selected 3′ RACE clones, was 1052 bp, excluding the poly (A) tail. The cDNA contained an ORF encoding a peptide of 138 amino acids with a theoretical isoelectric point of 6.74 and a molecular mass of 15.87 kDa. A putative polyadenylation signal was located at 522 bp downstream of the stop codon and 13 bp upstream of the poly(A) sequence (Fig. 6-2A). The deduced amino acid sequence of *rbp1*, aligned with several other cellular retinoid-binding proteins from mammals and fishes, exhibited the highest amino acid sequence identity (58-59%) with mammalian CRBPI and less than 50% of amino acid sequence identity with the other CRBPs and CRABPs (Fig. 6-2B). Based on the alignment, the encoded protein has three amino acid insertions at amino acid positions 3, 50 and 51 compared with other CRBPs. A phylogenetic analysis using ClustalX [Thompson *et al.*, 1997] grouped the zebrafish CRBP with the mammalian CRBPI clade (Fig. 6-3).

### 6.2.2 Structure and transcription start sites of the zebrafish *rbp1* and *rbp2* genes

Using the cDNA sequence of the putative zebrafish *rbp1* to search against the zebrafish genomic sequence database (Wellcome Trust Sanger Institute), I identified a

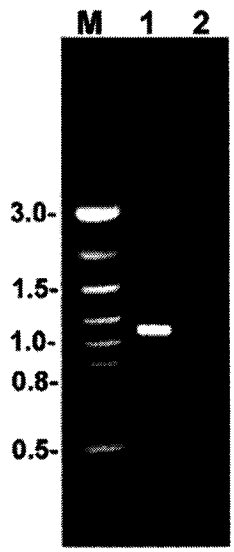



**Figure 6-1. Cloning of the zebrafish CRBPI cDNA by 3′ RACE.** Zebrafish CRBPI cDNA 3′ RACE products were separated by agarose gel-electrophoresis. The ~1.1 kb RACE product was excised, cloned and sequenced. M: 100 bp DNA ladder, molecular sizes of the molecular weight markers are indicated on the left of the panel in kb. Lane 1: 3′ RACE product of CRBPI cDNA; Lane 2: Negative control.

**Figure 6-2. Nucleotide and deduced amino acid sequences of the zebrafish *rbp1* cDNA.** **(A)** cDNA clones for the zebrafish CRBPI were generated by 3′ RACE. Both strands of the cDNA clones were sequenced, aligned, and the amino acid sequence for the zebrafish CRBPI was deduced. The coding nucleotides are shown in uppercase letters and the 5′ and 3′ UTRs are in lowercase letters. The potential polyadenylation signal is double underlined. Underlined sequences correspond to primers used in RT-PCR experiments (s2 and as2), 3′ RACE (internal, s1) and radiation hybrid mapping (s2 and as 1). Numbers on the right correspond to the nucleotides in the cDNA sequence. GenBank accession number: AY309074. **(B)** The deduced amino acid sequence of the zebrafish CRBPI (ZF CRBPI, accession number: AAQ54326) was aligned with the amino acid sequence of the human CRBP1 ( HM CRBPI, P09455), mouse CRBPI (MS CRBPI, Q00915), human CRBPIII (HM CRBPIII, P82980), zebrafish CRBPII (ZF CRBPII, Q8UVG6), human CRBPII (HM CRBPII, P50120), human CRBPIV (HM CRBPIV, Q96R05), pufferfish CRABPI (PF CRABPI, O42386) and human CRABPII (HM CRABPII, P29373). The intron positions relative to the amino acid sequence are indicated by "▼". Dots indicate amino acid identity and dashes represent gaps. The percentage amino acid sequence identities between the zebrafish CRBPI and other CRBPs or CRABPs are shown at the end of the sequences.

**A**

```
   S1                                                                    5' cagaggaa   8

cagtgagcacactaagccctatatcacacaaagtttagctgcattcgtgtcgattcgaaacaaaaggaaaatagaaagaaagagaaa   95


ATG TCC AAA CCT AAT TAC ACT GGC GTT TAC AAC ATG ATT TCC CAG GAT AAT TTT GAA GAA TAT CTG   161
Met Ser Lys Pro Asn Tyr Thr Gly Val Tyr Asn Met Ile Ser Gln Asp Asn Phe Glu Glu Tyr Leu
                                        S2
GCG GCT TTG GAT GTT AAT TAT GCA CTG AGG AAG GTA GTC TGT ATT CTG AAA CCC AGC AAA CAC ATT   227
Ala Ala Leu Asp Val Asn Tyr Ala Leu Arg Lys Val Val Cys Ile Leu Lys Pro Ser Lys His Ile

GAA CAC GAT GTG AAC ACT GGC AGG ATG AAG ATA AAG ACC GTC ACC ACT TTT AAG AAC TTC GAC ATG   293
Glu His Asp Val Asn Thr Gly Arg Met Lys Ile Lys Thr Val Thr Thr Phe Lys Asn Phe Asp Met
       as1
GAC TTC ACT TTA GGA CAG GAG TTC ACT GAG GAC CTG GGG CCA GTC GAT GGG CGA AAG TGC CAA ACT   259
Asp Phe Thr Leu Gly Gln Glu Phe Thr Glu Asp Leu Gly Pro Val Asp Gly Arg Lys Cys Gln Thr

ACG GTG GAC TGG GAC GGT GAT AAA TTG ATC TGC GTG CAG CGA GGA GAG AAA GAG GGC AGA GGC TGG   425
Thr Val Asp Trp Asp Gly Asp Lys Leu Ile Cys Val Gln Arg Gly Glu Lys Glu Gly Arg Gly Trp

ACA CAC TGG TTA GAA GGA AAC ATC CTC CAT TTG GAG CTG AGG GCT CAG GGT GTC ACT GCC AAA CAA   491
Thr His Trp Leu Glu Gly Asn Ile Leu His Leu Glu Leu Arg Ala Gln Gly Val Thr Ala Lys Gln

GTC TTT AAG AAG GTT GAG TGA                                                            512
Val Phe Lys Lys Val Glu  #


ggatgaagagagagagagaatcagaggggccattgctgggggggtttgagctgagtttatttctttgagcccatggctcagaaaa   599
aggctggattgccacagttttttttgtccctgattcaaaacaaatcgtatttaaatgaagaccataaatgaagaaattggcacccgc   686
tctatgatgtttcagttttttaccccccggagttgtgattgaactaaaacaatagcagcaccctgtgtgaacatccatcaatggtact   773
attgtaaaatgtgtatctgaaaaatggcacctgaaaattgttttctgcagatatgtttacataagttgaagtcatcaaaacagacag   860
                                                                                  as2
tggttgttcaatccaaagttgtatttaatgtctagtaatctgctactggatttgtgcaaatgttatgaaataatcaagtcgtgatta   947
atttgcagttgtcaaacattgtctgaaagtgtcaaaatttgtaagctctgaaaagaatgacttgtcattttatttgtcaactctaaa  1034
atacaaacgtttatattc(a)n                                                                  1052
```

**B**

```
                    10          Intron 1            40                          70
ZF CRBPI    MSKPNYTGVYNMISQDNFEEYLAAL▼DVNYALRKVVCILK--PSKHIEHDVNTGRMKIKTVTTFKNFDMDFTLGQEF
HM CRBPI    .P-VDF..YWK.LVNE......R..  ...V....IANL..--...E.VQ.GD--H.I.R.LS..R.YI...QV.K..
MS CRBPI    .P-VDFN.YWK.L.NE......R..  ...V....IANL..--...E.VQ.GD--H.I.R.LS..R.YI...QV.K..
HM CRBPIII  .P-..L..YYRFV..K.M.D..Q..  NISL.V..IALL..--...E...QG.--HMTVR.LS..R.YTVQ.DV.V..
ZF CRBPII   .P-ADFN.TWE.L.N....DVMK..  .IDF.T..IAVH..--QT.V.VQNGD--KFET..LS..R.YEVN.VI.E..
HM CRBPII   .T-RDQN.TWE.E.NE...G.MK..  .IDF.TP.IAVR.T--QT.V.DQ.GD--NF.T..TS..R.Y.V...V.V..
HM CRBPIV   .P-ADLS.TWTLL.S....G.ML..  GIDF.T..IAKL..--.Q.V..QNGD--SFT.H.NSSLR.YFVK.KV.E..
PF CRABPI   .P--.FA.TWK.K.SE..D.L.K..  G.NTM....AVAAASN.HVE.RQ.GE--KFY...S..VRTTEINFHI.E..
HM CRABPII  .P--.FS.NWKIIRSE....L.KV.  G.NVM...IAVAAASK.AVE.KQEGD--TFY...S..VRTTEINFKV.E..

                    Intron 2        100                     Intron 3    130
ZF CRBPI    TEDLGPVDGRKCQ▼TTVDWDGD-KLICVQ---RGEKEGRGWTHWLEGN-ILHL▼ELRAQGVTAKQVFKKVE  100%
HM CRBPI    E...TGI.D...M  ...S....-..Q...---K.........Q.I..D-E... .M.VE..VC.......Q   58%
MS CRBPI    E...TGI.D...M  ...S....-..Q...---K.........Q.I..D-E... .M..E..IC.......H   59%
HM CRBPIII  E...RSV......  .I.T.EEE-H.V...---K..VPN...R.....E-M.Y. ..T.RDAVCE...R..R   50%
ZF CRBPII   D.QTKGL.N.TVK  .L.K....-..V...---K....N...KQ.I..D-L... .IHC.DKVCH.....KN   43%
HM CRBPII   D.YTKSL.N.HVK  AL.T.E..-V.V...---K....N...KQ.I..D-K.Y. ..TCGDQVCR.....K-   42%
HM CRBPIV   D..NRGL.N...K  SL.I..N.-R.T.I.---K...KN......I..D-K... .MFCE.QVC..T.QRA-   41%
PF CRABPI   D.ET--......K  SLPT.ESEN.IR.K.TLVE.DGPKTF..RE.N.D-E.T. VFG.DD.VCTRIYVRE-   31%
HM CRABPII  E.QT--....P.K  SL.K.ESEN.MV.E.KLLK..GPKTS..RE.TNDGE.I. TMT.DD.VCTR.YVRE-   27%
```

**Figure 6-2. Nucleotide and deduced amino acid sequences of the zebrafish *rbp1* cDNA**

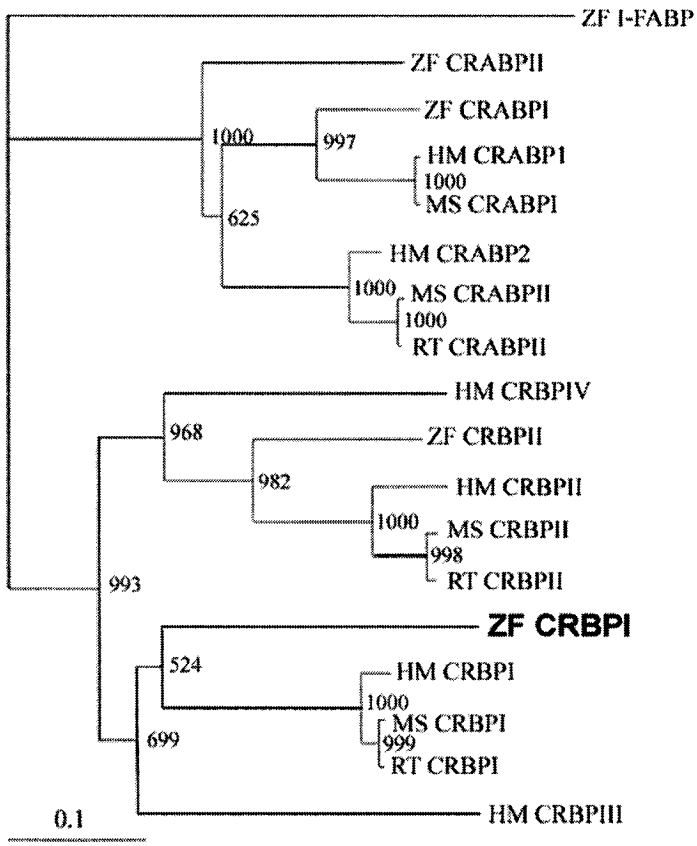



**Figure 6-3. Phylogenetic tree of the cellular retinoid-binding proteins.** The Bootstrap neighbour-joining phylogenetic tree was constructed with ClustalX [20] using the zebrafish intestinal type fatty acid-binding protein (I-FABP) amino acid sequence (Swiss-Prot and TrEMBL accession number Q9PRH9) as an outgroup. The bootstrap values (based on number per 1000 duplicates) are indicated on the nodes. Sequences used in this analysis in addition to those in figure 6-1 include the amino acid sequences for rat (RT) CRBPI (P02696), mouse (MS) CRBPII (Q08652), rat CRBPII (P06768), zebrafish (ZF) CRABPI (AAP44333), human (HM) CRABPI (P29762), mouse CRABPI (P02695), zebrafish CRABPII (NP_878279), mouse CRABPII (P22935), rat CRABPII (P51673). Scale bar = 0.1 substitutions per site.

contig sequence (ctg11808.1) harboring the entire coding sequences of exon 1 (76 bp), exon 2 (185 bp), exon 3 (102 bp), the genomic sequence of intron 1 (2723 bp), intron 2 (166 bp) and a genomic sequence of 2973 bp upstream of the initiation codon. Genomic trace sequence (235723-a1145g05.p1c) contains the entire coding sequence for exon 4 and a partial sequence of intron 3 and 3′ untranslated region (UTR). The coding sequence of the putative zebrafish *rbp1* cDNA is identical to the corresponding genomic nucleotide sequence from the database except for a G/A transition at the position of 247 in the cDNA sequence (Fig. 6-2A), which results in an amino acid change from glycine to aspartic acid. However, the three isolated 3′ RACE clones and all the zebrafish ESTs corresponding to CRBPI reported in GenBank have a "G" rather than an "A" at this position, therefore the nucleotide A in the genomic sequence is likely a sequencing error. The zebrafish *rbp1* gene consists of four exons interrupted by three introns, identical to all iLBP genes identified thus far. All exon/intron splice junctions of the zebrafish *rbp1* conform to the GT-AG rule [Breathnach and Chambon, 1981]. Exon/intron positions in the zebrafish CRBPI amino acid sequence are shown in Fig.6-2B.

A search of the zebrafish genomic sequence database with the zebrafish CRBPII cDNA sequence [Cameron *et al.*, 2002] found four assembly or trace sequences, z06s122453, z35725-a2618a07.q1c, zfishB-a2744c03.p1c and z06s007814, containing the coding sequences for exon 1, exon 2, exon 3 and exon 4 of the zebrafish CRBPII gene (*rbp2*), respectively. Assembly z06s122453 also contains a genomic sequence of 2365 bp upstream of the initiation codon of the zebrafish CRBPII cDNA. Two nucleotide variations between the cDNA and genomic sequences were observed that changed the $10^{th}$ codon of exon 1 from GAG to TAA. This variation most likely resulted from a

sequencing error in the assembly genomic sequence because the published cDNA sequence, the 5′ RACE sequence reported here (see below), corresponding ESTs from GenBank and another trace sequence (zfish35935-775b01.p1c) containing exon 1 have GAG in codon 10 rather then TAA. The zebrafish *rbp2* consists of four exons separated by three introns and the exon/intron splice junctions of the zebrafish *rbp2*, like those of *rbp1*, conform to the GT-AG rule (Fig. 6-1B) [Breathnach and Chambon, 1981].

The structures of zebrafish *rbp1* and *rbp2* genes were compared with those of the othologous genes from humans and mouse (data not shown). The exon/intron structure, coding capacity and the boundary sequences around the splicing junctions are conserved among mammals and zebrafish.

To map the transcription start site(s) of the zebrafish *rbp1* and *rbp2* genes, the $^{7}$methyl G capped ends of the zebrafish *rbp1* and *rbp2* mRNA transcripts were amplified from total adult zebrafish RNA using 5′ RLM-RACE (Fig. 6-4). After the nested PCR amplification, two distinct bands each of ~350 bp for *rbp1* and one single band of ~300 bp for *rbp2* were isolated and purified from the electrophoresis gel (Fig. 6-4A). Sequencing of the cloned 5′ RLM-RACE products (Fig. 6-4B) demonstrated that the cDNA ends were 352 and 336 bp in length for *rbp1*, and 289 bp for *rbp2*, including 38 bp of 5′ RLM-RACE adaptor primer sequence in each clone. Alignment of the two *rbp1* 5′ RLM-RACE sequences with the 5′ genomic sequence of the zebrafish *rbp1* revealed that this gene has two transcription start sites located at positions 288 bp and 272 bp, respectively, upstream of the initiation codon. In the same way, the transcription start site of the zebrafish *rbp2* was mapped to the position 39 bp upstream of its initiation codon.

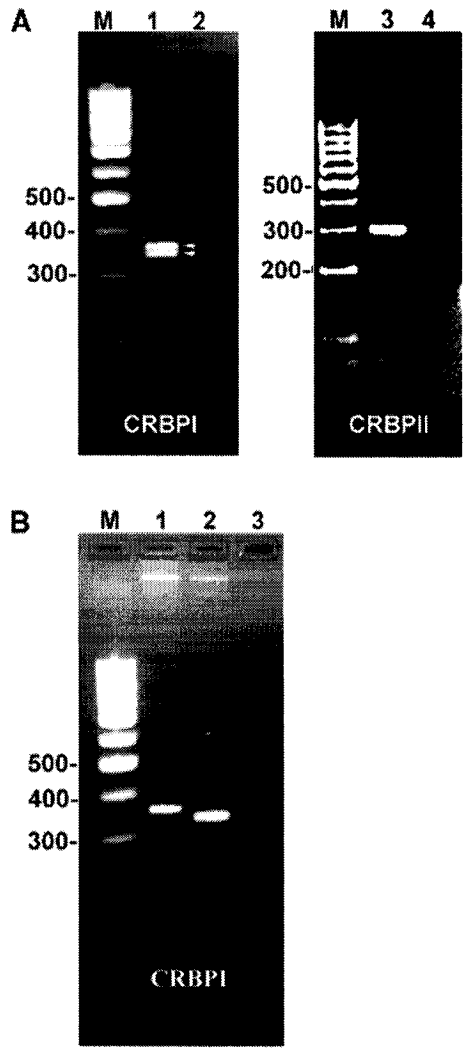



**Figure 6-4. Agarose gel electrophoresis of 5′ RLM-RACE products.** (**A**) Total RNA from whole adult zebrafish was sequentially treated with calf intestinal alkaline phosphatase (CIP), tobacco acid pyrophosphatase (TAP) and then ligated to a designated RNA adapter. Following the nested PCR amplification, two PCR-amplified products of approximately 350 bp for *rbp1* (Panel A, lane 1) and one product of 300 bp for *rbp2* (Panel A, lane 3) were size-fractionated by gel-electrophoresis through 2.5% agarose. RNA treated to the same experimental regime, but with TAP digestion omitted, did not generate a product (Panel A, lanes 2 and 4 for rbp1 and rbp2, respectively). (**B**) The inserts of two 5'RLM-RACE *rbp1* clones were 352 bp (lane 1) and 336 bp (lane 2) in size. Lanes indicated M are 100 bp molecular weight marker. The sizes (in bp) of the molecular markers are indicated to the left of each panel.

The 5′ upstream sequences of *rbp1* and *rbp2* were analyzed using MatInspector professional 6.2 [Quandt *et al.*, 1995] to reveal potential 5′ *cis* regulatory elements of the zebrafish *rbp1* and *rbp2* genes (Fig. 6-5 A, B). As shown for mammalian CRBPI gene promoters [Ong *et al.*, 1994], no TATA box was present in the zebrafish *rbp1* promoter. Several functional *cis* elements, including the NF1-binding element, the AP-2 and Sp1 elements and the RA-responsive element (RARE) observed in mammalian CRBPI genes [Ong et al., 1994], were not found in the 2.7 kb 5′ upstream sequence of the zebrafish *rbp1* gene. However, a number of POU domain elements, including the response elements for Brn-2, Brn-3 and Tst-1 (Oct-6), are present in the 5′ upstream region of the zebrafish *rbp1* gene (Fig. 6-5 A). POU-domain transcription factors are widely expressed in the developing mammalian brain and involved in regulation of gene expression during CNS development [He *et al.*, 1989].

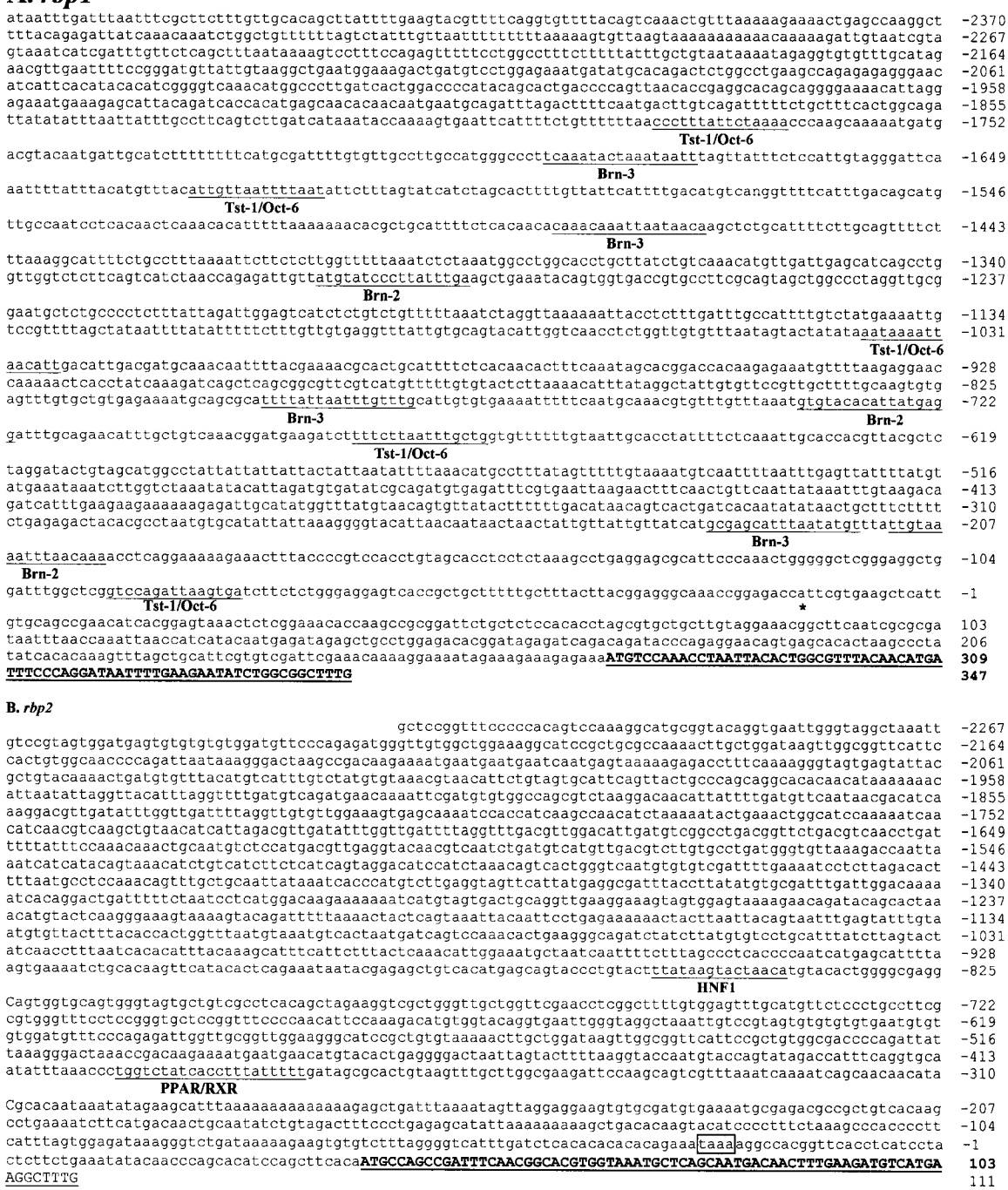
Inspection of the 2326 bp 5′ upstream sequence of the zebrafish *rbp2* disclosed a putative TATA box at position -27 of the zebrafish *rbp2* gene sequence (Fig. 6-5B). A TATA box was also found in the promoter of mammalian genes for CRBPII [8]. A PPAR/RXR-specific *cis* element was observed in the proximal promoter region of the zebrafish *rbp2* gene (Fig. 6-5B). The PPAR/RXR heterodimer is an important transcription factor involved in regulation of the mammalian and avian *rbp2* genes [Takase *et al.*, 2000].

### 6.2.3 Linkage group assignment and syntenic relationship of the zebrafish *rbp1*

I used the zebrafish LN54 radiation hybrid panel [Hukriede *et al.*, 1999] to assign the zebrafish *rbp1* gene to linkage group (LG) 16 at a distance 1.02 centiRays (cR) from the marker fb30f09 on the LN54 panel map and 47.48 cM on the merged ZMAP. The

**Figure 6-5. Nucleotide sequences of the 5′ upstream region of zebrafish *rbp1* (A) and *rbp2* (B) gene.** The coding sequence of the first exon is shown in uppercase letters and underlined. 5′ flanking sequence and the 5′ UTR of the first exon are shown in lowercase letters. The transcription start sites are numbered as +1 (the downstream one for *rbp1*) and the second transcription start site for *rbp1* is indicated by a star under the nucleotide. The core sequence of a potential TATA box upstream of the transcription initiation site of the *rbp2* gene is boxed. The matrix sequences of the selected potential *cis* regulatory elements in the *rbp1* (A) and *rbp2* (B) genes identified by MatInspector [Quandt et al., 1995] are underlined and indicated.

**A. *rbp1***




**B. *rbp2***

**Figure 6-5. Nucleotide sequences of the 5′ upstream region of zebrafish *rbp1* (A) and *rbp2* (B) gene.**

zebrafish *rbp1* gene is syntenic to a mapped EST similar to the human huntington interacting protein B (*hypb*, 43.40 cM of LG 16), and the gene for beta-catenin, (*ctnnb*, 20.70 cM of LG 16), glycogen synthase kinase 3 (*gsk3b*, 32.0 cM of LG 16) and caspase (*caspa*, 48.9 cM of LG 16) in the zebrafish genome (Fig. 6-6). In humans, the *RBP1* gene (chromosomal location 3q23) is also syntenic to *HYPB* (3p21.31), *CTNNB* (3p21) and *GSK3a* (3q13.3). In the mouse genome, the *Rbp1* gene (52.0 cM of LG 9), *Ctnnb* (72.0 cM of LG 9) and *Casp1* (1.0 cM of LG 9) are located on LG 9 (Fig. 6-6). This sytenic relationship has been conserved after the divergence of mammals and fishes for approximately 450 mya. In the human and mouse genome, however, the genes for CRBP1 and CRBP2 are closely linked [Demmer *et al.*, 1987], while in the zebrafish genome, *rbp1* and *rbp2* genes were assigned to LG16 and LG15 [Cameron *et al.*, 2002], respectively. Both the zebrafish *rbp1* and *rbp2* genes exhibit conserved syntenies with their mammalian orthologs. The assignment of zebrafish *rbp1* and *rbp2* genes on different linkage groups compared to their mammalian orthologs, which are closely linked on a single chromosome in human and mouse, suggests that inter-chromosomal rearrangement, possibly by translocation, has occurred in the fish lineage after divergence of the euteleost fishes from the mammals.

### 6.2.4 Tissue-specific distribution of the *rbp1* and *rbp2* gene transcripts in adult zebrafish

RT-PCR was performed with mRNA extracted from adult zebrafish tissues to determine the tissue-specific distribution of the mRNA encoded by the zebrafish *rbp1* and *rbp2* genes. After 35 cycles of PCR amplification, the *rbp1*-specific product was detected in the reactions with cDNA template from the intestine, liver, ovary, brain and testis (Fig.

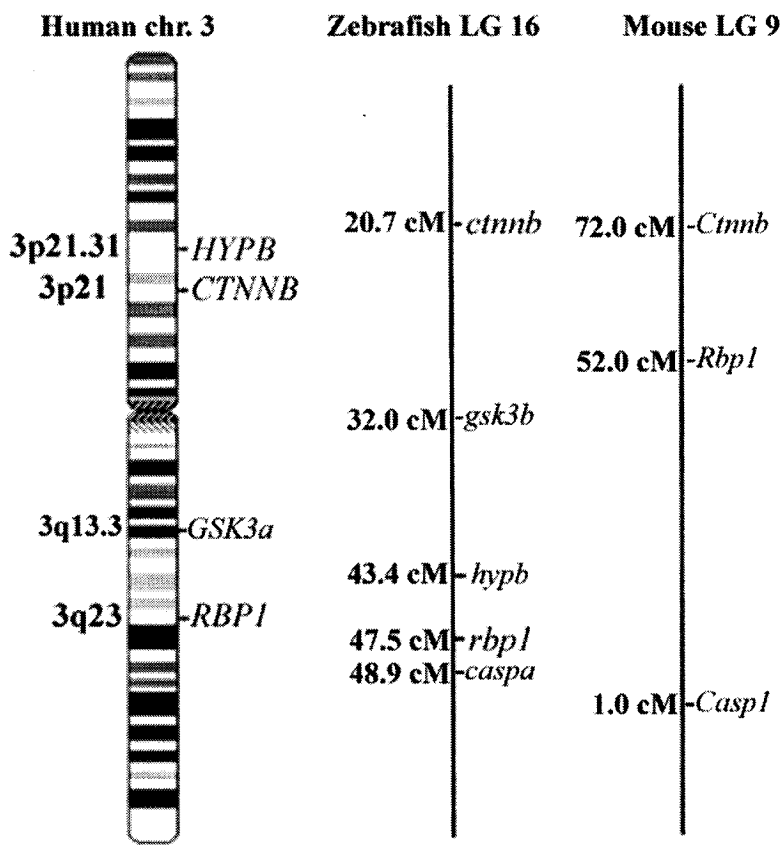



**Figure 6-6. Comparison of syntenic relationships of the CRBPI genes on human chromosome 3 (Chr. 3), zebrafish linkage group 16 (LG 16) and mouse linkage group 9 (LG 9).** The human and mouse gene map were constructed based on gene loci from LocusLink (http://www. ncbi.nlm.nih.gov/LocusLink/list.cgi), NCBI. The zebrafish linkage map was constructed based on the gene loci shown on the ZMAP at ZFIN website (http://zfin.org). *ctnnb*: gene for beta-catenin; *hypb*: gene for huntingtin interacting protein B.

6-7). No PCR products were observed using *rbp1*-specific primers and cDNA derived from the skin, heart and muscle (Fig. 6-7). *In situ* hybridization of the zebrafish CRBPI cDNA to tissue sections of adult zebrafish did not yield a hybridization signal (data not shown) possibly owing to the low levels of mRNA in adult tissues.

RT-PCR with the zebrafish *rbp2* cDNA-specific primers generated abundant amplification product in the reactions with cDNA template from the intestine and liver, but no PCR product was amplified when cDNA from other tissues was used as template (Fig. 6-7). The tissue-specific mRNA distribution pattern of the zebrafish *rbp2* gene in adult zebrafish observed using RT-PCR was in agreement with the expression pattern demonstrated using tissue section *in situ* hybridization [Cameron *et al.*, 2002].

### 6.2.5 Distribution of the *rbp1* mRNA in the developing zebrafish

I examined the distribution of both the zebrafish *rbp1* and *rbp2* gene transcripts during zebrafish embryonic and larval development by whole mount *in situ* hybridization. Images of zebrafish *rbp1* gene expression patterns are deposited in the ZFIN database at http://zfin.org [Thisse *et al.*, 2001]. Zebrafish *rbp1* gene transcripts were first detected at the early segmentation stage (11 hpf), showing mRNA distributed along the lateral neural plate (Fig. 6-8 a). During the middle segmentation stage (17 hpf), *rbp1* gene transcripts were detected in the anterior spinal cord, midbrain, epiphysis and diencephalon. A strong hybridization signal was observed in somites. The yolk syncytial layer (YSL) also showed *rbp1* mRNA distribution at the 17 hpf stage (Fig. 6-8 b). Later, during the early pharyngula period (24 hpf), abundant *rbp1* transcripts were found in the anterior spinal cord, the developing ear and branchial arches (Fig. 6-8 c, d). In the eye, the *rbp1* transcripts were restricted in a small domain in the retina and the lens (Fig. 6-8 c, e).

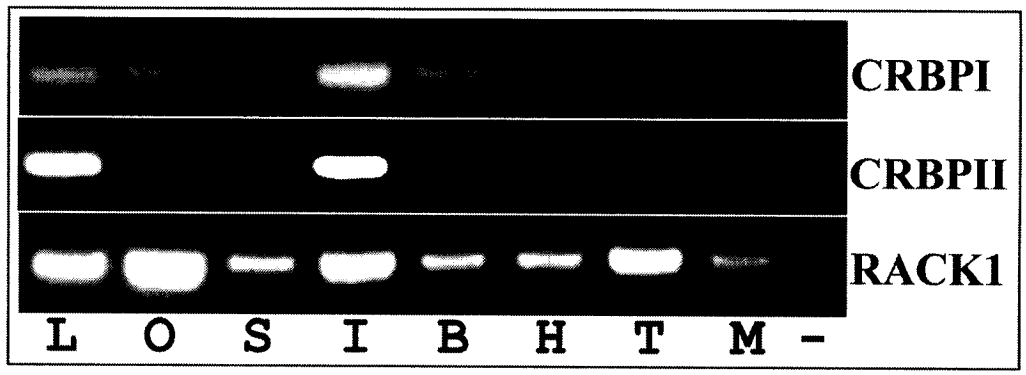



**Figure 6-7. Zebrafish *rbp1* and *rbp2* mRNA in adult tissues.** The zebrafish *rbp1* mRNA-specific primers amplified a PCR product from total RNA extracted from the liver (L), ovary (O), intestine (I), brain (B) and testis (T), but not in skin (S), heart (H) and muscle (M). The zebrafish *rbp2* mRNA-specific primers amplified a product from the liver and intestine. A RT-PCR product corresponding to the constitutively expressed, receptor for activated C kinase (RACK1) was generated from RNA in all samples. A negative PCR control (-) did not contain cDNA template and did not generated any RT-PCR product.

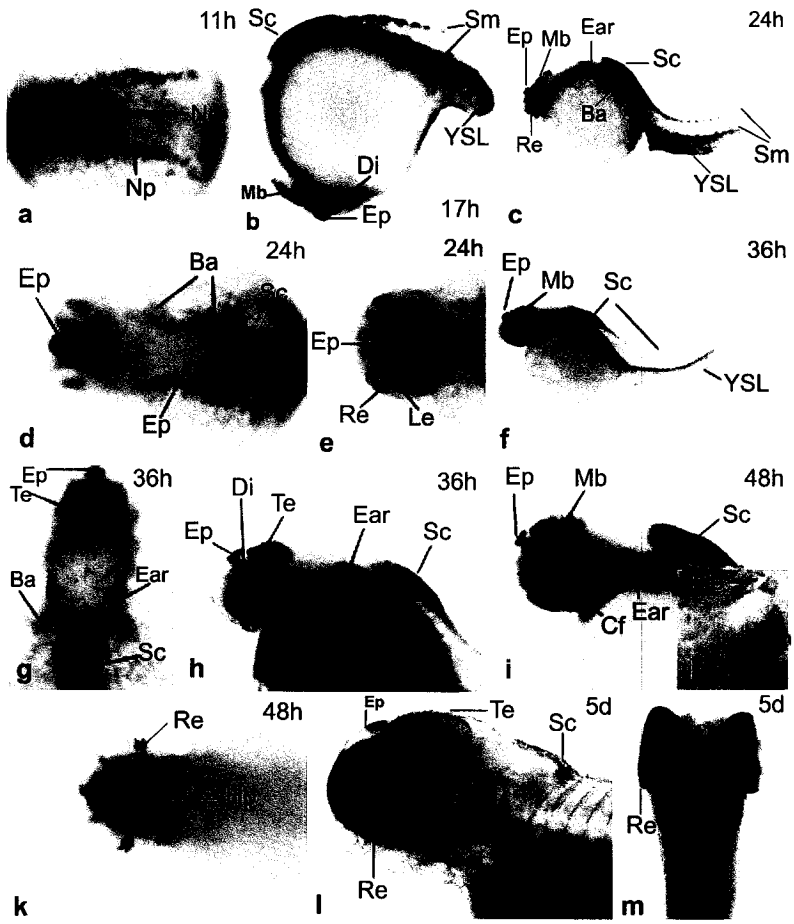



**Figure 6-8. Zebrafish *rbp1* mRNA expression during embryonic and larval development.** Whole mount *in situ* hybridization showing *rbp1* mRNA expression in the zebrafish embryo and larva during early and middle segmentation stages at 11 hours post fertilization (hpf, a) and 17 hpf (b), late segmentation stage at 24 hpf (c-e), then at 36 hpf (f-h), 48 hpf (i-k) and 5 days (l-m). b, c, f, h, I, l: lateral view, anterior is to the left; a, d, k: dorsal view, anterior is to the left; c: dorsal view showing the eyes; g, m: dorsal view, anterior is to the top; j: retina showing an area near the choroid fissure (Cf) at 48 hpf with a high level of *rbp1* expression. mRNA expression in the anterior spinal cord (Sc), retina (Re), branchial arches (Ba), midbrain (Mb), epiphysis (Ep), yolk syncytial layer (YSL), lens (Le), tectum (Te), diencephalon (Di) and hypophysis (Hp) is indicated.

The *rbp1* mRNA in the epiphysis, midbrain, somites and YSL is retained at the 24 hpf stage (Fig. 6-8 c, d). At 36 hpf stage, the *rbp1* expression of mRNA is down-regulated in the somites and YSL, but retained in the spinal cord (Fig. 6-8 f). *rbp1* mRNA was localized to a discrete region of the optic tectum at the 36 hpf stage (Fig. 6-8 g, h) and in the developing ear and branchial arches that can be readily distinguished at this stage of development (Fig. 6-8 g). At 48 hpf, *rbp1* mRNA was detected in the ventricular zone of the midbrain, epiphysis, and anterior spinal cord (Fig. 6-8 i). A weak hybridization signal of *rbp1* mRNA was located at the dorsal retina and the ventral side of the eye close to the choroid fissure (Fig. 6-8 j, k). In 5 day-old larvae, levels of *rbp1* transcripts in the spinal cord were reduced relative to that observed in the 48 hpf larvae, based on the intensity of the hybridization signal, while *rbp2* mRNA was distributed in the retina, restricted to the external layer of the pigmented epithelium (Fig. 6-8 l, m).

### 6.2.6 Distribution of the *rbp2* mRNA in the developing intestine and retina of zebrafish

Transcripts of zebrafish *rbp2* gene were first detected at much later developmental stages than transcripts of the *rbp1*gene. The presence of *rbp2* mRNA was first observed in larvae at post segmentation stage (36 hpf) in the developing eye with a distinct region of hybridization in the posterior part of the lens (Fig. 6-9 a, b, c). *rbp2* mRNA was then absent in the lens at 48 hpf stage, but was present in the photoreceptor cell layer of the retina (Fig. 6-9 d, f, g). *rbp2* mRNA persisted in the retina and extended to the dorsal and posterior retina in 5 day-old larvae (Fig. 6-9 h, i, j, k). At 48 hpf stage, *rbp2* mRNA was detected in the intestinal bulb, and based on the hybridization signal,

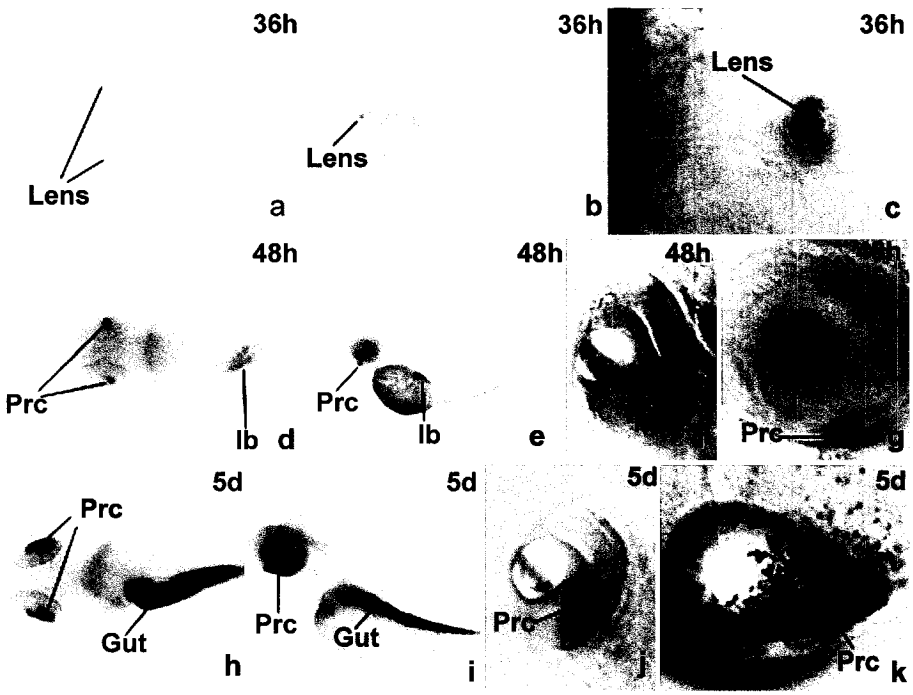



**Figure 6-9. Zebrafish *rbp2* mRNA expression during embryonic and larval development.** Whole mount *in situ* hybridization revealed mRNA expression of the zebrafish *rbp2* gene at 36 hpf in the developing lens (a-c), 48 hpf in part of the photoreceptor layer (Prc, d-g) and the intestinal bulb (Ib, d-e). The arrow indicates the optic nerve in the retina (f). At 5 days post fertilization, *rbp2* mRNA is abundantly distributed along the gut (h-i) and distributed in the photoreceptor cells in the retina (h-k). a, d, h: dorsal view, head to the left; b, e, I: lateral view, head to the left; c: lateral view of the lens; g: lateral view of the eye; f, k: dorsal view of the eye; j: ventral view of the eye.

*rbp2* mRNA levels were elevated in the whole of the developing gut in 5 day-old larvae (Fig. 6-9 d, e, h, i).

## 6.3 DISUSSION

I cloned the cDNA encoding a CRBP from zebrafish (*Danio rerio*). Based on conserved gene structure, syntenic relationship and phylogenetic analysis, I identified the zebrafish ortholog of mammalian CRBPI genes. To our knowledge, this is the first report of a CRBPI gene and its developmental and tissue-specific expression in fishes.

The tissue distribution of CRBPI protein and mRNA in adult mammals is strikingly different among species. For instance, in rats, CRBPI is abundantly distributed in the liver and kidney, and, to much lesser density, in the lung, testis, spleen, eye, ovary, uterus and intestine [Ong *et al.*, 1994], whereas in human, the ovary contains the highest amount of CRBPI, and the kidney and lung have much lower levels of CRBPI [Fex and Johannesson, 1984]. The tissue distribution of human CRBPI was later confirmed by radioimmuno assay [Ong and Page, 1986] and mRNA dot-blot analysis [Folli *et al.*, 2001]. In the present study, I found the zebrafish *rbp1* mRNA was present in the liver, intestine, brain, ovary and testis, but could only be detected by RT-PCR (Fig. 6-7) and not by adult tissue section *in situ* hybridization (data not shown), indicating that its expression is at a relatively low level. The abundant and restricted presence of CRBPII in the adult small intestine is conserved in mammals and birds. Similarly the zebrafish *rbp2* mRNA was also detected in abundance in the adult intestine by RT-PCR (Fig. 6-7) and by adult tissue section *in situ* hybridization [Cameron *et al.*, 2002]. However, the zebrafish *rbp2* mRNA was also observed in the adult liver, which is not the case for any

vertebrates investigated to date, suggesting a novel function for the zebrafish *rbp2* gene product in retinol metabolism in the adult fish liver, or a different metabolic pathway for retinol in the fishes. To invoke a novel function for the zebrafish *rbp2* gene is speculative. Confirmation of such a novel function for *rbp2* will require additional studies including localization of the CRBPII protein in the liver by immunocytochemistry, and detailed ligand binding assays *in vitro* using recombinant zebrafish CRBPII. The differential expression of the zebrafish *rbp2* in the adult liver may therefore be controlled by a specific *cis* regulatory sequence not found in mammalian CRBPII genes. In zebrafish and mammals, the *rbp2* genes contain PPAR/RXR heterodimer responsive elements [Takase *et al.*, 2000]. In zebrafish, the *rbp2* gene also has a potential *cis* element for the hepatic nuclear factor 1 (HNF1). HNF1 is a required transcription factor driving the liver-type fatty acid-binding protein gene expression in the mouse liver [Akiyama *et al.*, 2000].

During development, the mammalian CRBPI protein and mRNA are abundant in the mammalian central nervous system (CNS). At the early somite stage of the mouse embryonic development, CRBPI is widely distributed from the posterior neural tube up to the hindbrain, but very low levels were seen in the forebrain [Ruberte *et al.*, 1991]. CRBPI was also detected in the mouse and rat olfactory bulb, spinal cord and hippocampus using immunohistochemical assay and tissue section *in situ* hybridization [Perez-Castro *et al.*, 1989; Zetterstrom *et al.*, 1994; Romand *et al.*, 2000]. The mRNA expression of the *rbp1* gene in developing zebrafish determined here is similar to that of its mammalian orthologs.

Retinoic acid (RA), the biologically active derivative of retinol, acts as an embryonic hormonal signal in controlling patterns of gene activity in several systems including the CNS [Maden, 2001; 2002], the limb [Stratford *et al.*, 1996; 1999] and the eye [McCaffery *et al.*, 1999]. In the vertebrate CNS, RA is essential for the anterior-posterior axis (AP) patterning [Guidato *et al.*, 2003; Grandel *et al.*, 2002; Begemann and Meyer, 2001; Dupe and Lumsden, 2001]. RA is produced from retinol through two oxidation steps. In the first step, retinol is oxidized to retinal by retinol dehydrogenase [Ang *et al.*, 1996]. Subsequently, retinal is oxidized to RA by the action of retinaldehyde dehydrogenases. To date, three isozymes of retinaldehyde dehydrogenase, type 1 (Raldh1), type 2 (Raldh2) and type 3 (Raldh3), have been identified in the CNS during vertebrate embryonic development [reviewed in Maden, 2001]. In the developing CNS, Raldh2 is mainly expressed in the neural tube and Raldh1 and Radhl3 are expressed in the eye and other distinct sites in the CNS [Maden, 2001]. Upon synthesis, RA exerts its action on gene regulation in the cells of the target tissues through binding and activating the nuclear transcription factors, the RA receptors (RARα, β, γ) and retinoid x receptors (RXRα, β, γ) [Napoli, 1996]. CRBPI binds retinol and retinal with high affinity and then the holo-CRBPI serves as the substrate for the enzymes in the retinol and retinal oxidation pathways, thus mediating RA synthesis [Noy, 2000; Lohnes *et al.*, 1992]. CRBPI distribution, therefore, may serve to identify centers for RA synthesis and action in the developing embryo. For example, RA levels are high in the spinal cord but extremely low or undetectable in the fore-, mid- and hindbrain in mouse and chicken [reviewed in Maden, 2001; 2002]. RA levels are high in the developing mouse retina [Maden, 2001]. The distribution of *rbp1* mRNA in the zebrafish embryo found in this

study (Fig. 6-8) is concordant with distribution of RA in mammalian embryos. In the zebrafish embryos, RA is generated directly from the abundant retinal deposited in the yolk or retinal produced from another natural precursor, β-carotene, by the action of β,β–carotene-15,15′-oxygenase [Lampert *et al.*, 2003]. The latter pathway for RA production has also been implicated in pattern formation and differentiation during zebrafish embryogenesis [45].

Recently, patterns of expression of Raldh2 were revealed in developing zebrafish by whole mount *in situ* hybridization [Grandel *et al.*, 2002]. The distribution of Raldh2 mRNA in retina and somites coincides with the sites of CRBPI mRNA expression in zebrafish, but in contrast to CRBPI, Raldh2 is not expressed in the spinal cord. Instead, Raldh2 mRNA is found in the mesenchyme adjacent to the spinal cord, but shows a sharp on-off expression border at the spinal cord-hindbrain junction at the same level of the AP axis as CRBPI [Grandel *et al.*, 2002]. This difference in the RA and Raldh2 localization was evident in developing mammals and chickens [Maden, 2001]. This observation is noteworthy as RA at this particular site of the anterior spinal cord is thought to act as a localized signal to pattern the posterior hindbrain [Maden, 2002]. Whether RA is synthesized in the mesoderm and then transferred into the spinal cord or if the spinal cord itself has the ability to synthesize RA remains a controversial issue. On the one hand, none of the identified RA-synthesizing Raldh enzymes are detected in the early embryonic spinal cord, suggesting RA must be synthesized in the adjacent mesoderm and then transferred into the spinal cord [Maden, 2002]. On the other hand, the developing spinal cord of the chicken embryo is able to synthesize RA from retinol [Maden 2001; Maden *et al.*, 1998]. Recently, activity of an RA-responsive reporter transgene (*RARE-*

*hsp68-lacZ*) was detected in the medial-ventral cells of the spinal cord and the posterior hindbrain of the Raldh2$^{-/-}$ mutant mouse, indicating RA is synthesized in the spinal cord by RA synthesizing enzymes other than Raldh2 [Niederreither *et al.*, 2002; Mic *et al.*, 2002]. In the present study, I found that CRBPI, which is involved in RA synthesis, is present abundantly in the anterior spinal cord, but not in the adjacent paraxial mesenchyme, supporting the idea that the early developing anterior spinal cord is a site of RA synthesis using an as yet unidentified enzyme.

The expression of the zebrafish *rbp2* gene, as assayed by mRNA distribution, showed a distinctly different spatio-temporal pattern during embryonic and larval development compared to that of the *rbp1* gene. Zebrafish *rbp2* mRNA was abundant in the developing intestine being detected at a very early development state in the intestinal bulb (Fig. 6-9). CRBPII and its mRNA were detected in the intestine during development in mammals, but the function of CRBPII was not defined, although it was suggested that it is involved in the transportation and/or metabolism of the retinoids absorbed by the developing intestine [Crow and Ong, 1985; Li *et al.*, 1986; Leven *et al.*, 1987]. I found in zebrafish the presence of *rbp2* mRNA in the very early stage of intestinal development prior to the establishment of absorptive function, implying a role in intestinal organogenesis and development possibly through mediation of RA signaling. In rat, CRBPII mRNA is transiently expressed in the developing liver during the perinatal period [Li *et al.*, 1986; Leven *et al.*, 1987], whereas in zebrafish, CRBPII mRNA was absent in the developing liver up to 5 days post fertilization (Fig. 6-9). Clearly, there are differences between mammals and fishes during development and in adulthood regarding the expression of the *rbp2* gene in the liver. During early developmental stages, teleost

fishes obtain nutrients stored in yolk. CRBPII may, therefore, not be necessary in the developing fish liver to process β-carotene. In contrast, CRBPII may be needed in the liver of perinatal mammals for metabolism of hepatic β-carotene transferred from the immature intestine [Cameron *et al.*, 2002].

Both zebrafish *rbp1* and *rbp2* mRNA were distributed at distinct cell layers of the developing retina, an important tissue of RA distribution and action [Blanner and Olson, 1994]. CRBPI and CRBPII may mediate retinoid metabolism and RA production for development of the retina and later for maintenance of its function at the sites where they are distinctively distributed. I detected *rbp1* mRNA in the primitive developing ear and branchial arches, implying RA may also be involved in organogenesis and development of these structures. Radlh2 has been detected in the branchial arches [Grandel *et al.*, 2002] during the same zebrafish embryonic stages when *rbp1* is expressed, again suggesting the involvement of CRBPI in RA synthesis in this region.

In summary, I cloned the cDNA and characterized a gene encoding CRBPI in zebrafish. The zebrafish *rbp1* and *rbp2* genes showed differential patterns of expression in developing embryos and larvae. The zebrafish *rbp1* was expressed in the developing CNS at sites of RA synthesis and action, implying a possible role for CRBPI in the CNS patterning and organogenesis during the early development of the zebrafish embryo. A novel function of CRBPII in intestinal organogenesis and development was implicated by the presence of the mRNA of the zebrafish *rbp2* gene in the intestinal bulb during embryogenesis.

# 7 Identification of a Duplicated Gene Encoding the Zebrafish Cellular Retinol-Binding Protein Type I (*rbp1b*) and its Specialized Expression in the Primary Oocytes and the Developing Gall Bladder

## 7.1 INTRODUCTION

Retinoids play crucial roles in organogenesis during embryonic development and, thereafter, in the maintenance of a wide range of biological processes, including vision, reproduction, growth, immunity. [Napoli, 1996; Ross *et al.*, 2000]. It has long been recognized that retinoids are essential micronutrients for normal reproductive function of animals (Thompson *et al.*, 1964). Some recent animal experiments show that retinoids may serve as important regulators in oogenesis and oocytes maturation [Livera *et al.*, 2000; Morita and Tilly, 1999; Whaley *et al.*, 2000].

Retinoids are hydrophobic molecules and their transportation in the cytoplasm is facilitated by a group of intracellular carrier proteins which specifically bind different isomeric retinoids and target them to various intracellular enzymes or receptors for retinoid metabolism and eventual actualization of their physiological activities. During the last decade, increasing evidence suggests the involvement of intracellular retinoid-binding proteins in retinoid biological action and metabolism [reviewed by Napoli, 1999; Noy, 2000; Adida and Spener, 2002]. In mammals, cellular retinoid-binding proteins and their genes have been intensively studied, although their functions *in vivo*, precise mechanism for the regulation of their genes and the genomic evolution across various taxa await more detailed investigation.

Cellular retinoid-binding proteins belong to the large family of low molecular mass (~15 kDa) intracellular lipid-binding proteins (iLBP) that bind fatty acids, retinoids or steroids [reviewed by Ong *et al.*, 1994; Napoli, 1999; Noy, 2000]. In mammals, two different groups of cellular retinoid-binding proteins with distinctive ligand-binding properties have been identified, the cellular retinol-binding proteins (CRBPs) and the cellular retinoic acid-binding proteins (CRABPs). Four types of CRBP (namely CRBPI, CRBPII, CRBPIII and CRBPIV) encoded by distinct genes have been identified in mammals [Bashor *et al.*, 1973; Ong *et al*, 1984; Folli *et al.*, 2001; Vogel *et al.*, 2001; Folli *et al.*, 2002]. Although the ligand binding affinity (measured by $K_d$ values) varies among different CRBP types, all CRBPs show the greatest affinity for retinol and retinal, but do not bind retinoic acid or retinyl esters [Ong, 1994; Folli *et al.*, 2001; Vogel *et al.*, 2001; Folli *et al.*, 2002].

In previous studies, the gene structure, cDNA sequence, linkage relationship and expression of the genes coding for cellular retinol-binding protein type I (hereafter referred to as *rbp1a*) and II (*rbp2*) from zebrafish have been determined [Cameron *et al.*, 2003; Liu *et al.*, 2003, in press]. Here, I report the discovery of a duplicated *rbp1* gene (*rbp1b*) from the zebrafish genome. Phylogenetic analysis, conserved gene structure and syntenies suggest that the zebrafish *rbp1b* gene is another ortholog of the mammalian *Rbp1* gene. Transcripts of the zebrafish *rbp1b* gene were specifically distributed in the primary oocytes and the developing gall bladder, implying a specialized function of this gene in both the developing and adult zebrafish, and that *rbp1a* and *rbp1b* genes have evolved distinct tissue-specific patterns of expression after gene duplication and divergence.

## 7.2 RESULTS

### 7.2.1 Identification and structure of the zebrafish *rbp1* gene

A zebrafish genomic DNA sequence record (GenBank accession number: AL953896) was identified using tblastN in the zebrafish genomic DNA sequence database that has potential amino acid sequence similarity to the zebrafish CRBP1a protein amino acid sequence. The identified genomic DNA sequence contained a segment corresponding to a potential duplicated copy of the *rbp1* gene and was designated *rbp1b*. The gene size and structure (Fig. 7-1) of *rbp1b* was further defined after aligning the cDNA sequence (discussed later) with the genomic DNA sequence. The gene spans 8267 bp and consisted of four exons separated by three introns. The four exons were 159 bp, 179 bp, 102 bp and 527 bp in length and coded for 22, 60, 34 and 16 amino acids, respectively (Fig. 7-1). The three introns of the zebrafish *rbp1b* gene were 3018 bp, 988 bp and 3226 bp in length, respectively, accounting for 89% of the whole gene sequence. The nucleotides at the splice site of each exon/intron boundaries conformed to GT-AG rule [Breathnach and Chambon, 1981].

The transcription start site of the zebrafish *rbp1b* gene was defined by 5′ RLM-RACE, which generated a single product of the intact 5′ cDNA end with the 7-methyl G cap (Fig. 7-2, lane 1). A negative control with tobacco acid pyrophosphatase (TAP) treatment omitted yielded no RACE product (Fig. 7-2, lane 2). The 5′ RLM-RACE product was cloned and sequenced, and the transcription start site was mapped to a nucleotide position 92 bp upstream of the first ATG within the cDNA identified as the *rbp1b* initiation codon.

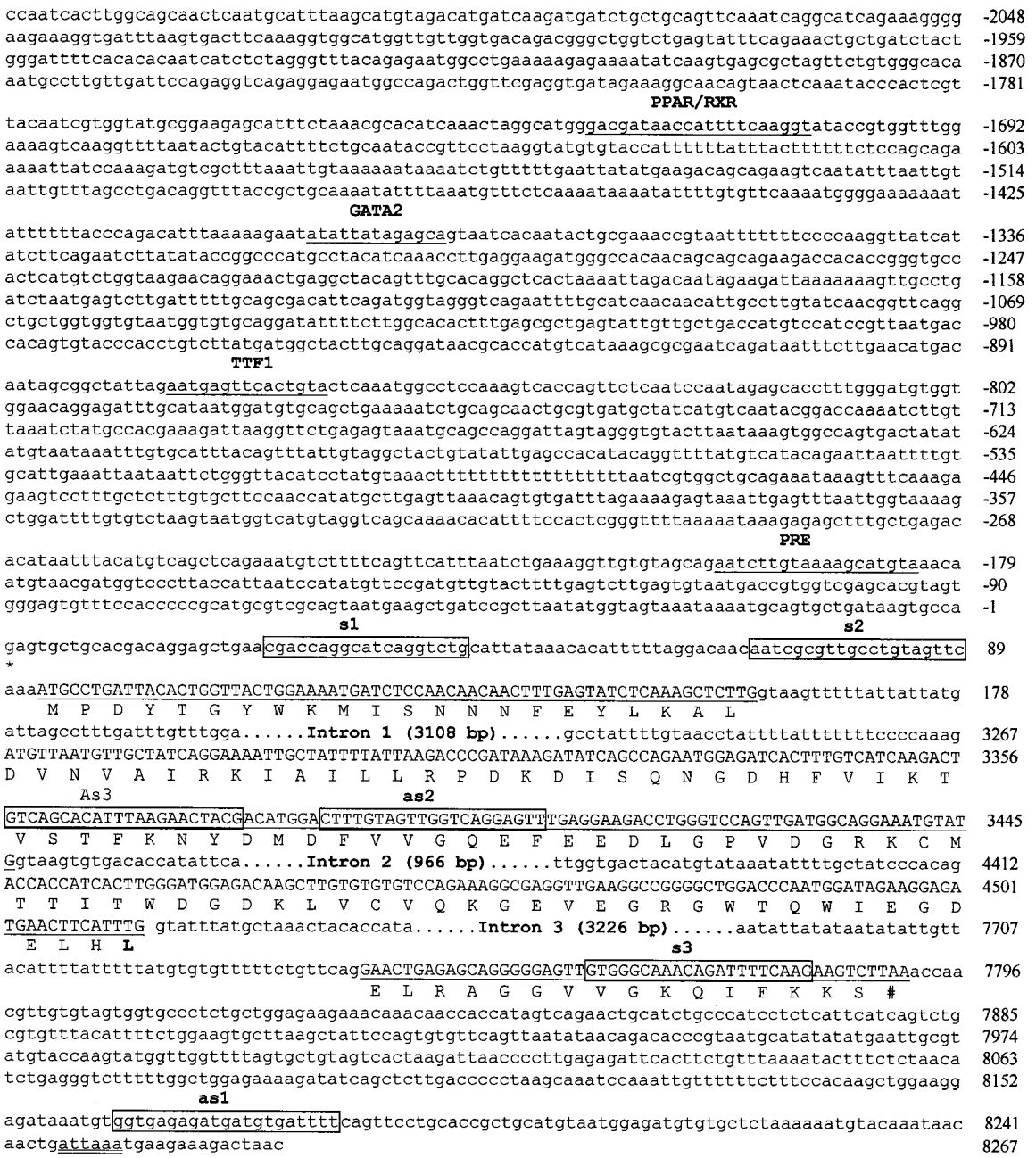



**Figure 7-1. Nucleotide sequence of the zebrafish *rbp1b* gene and its 5′ upstream region.** Exons are shown in uppercase letters with the coding sequences of each exon underlined and the deduced amino acid sequence indicated below. Numbers on the right indicate nucleotide positions in the gene sequence. The initiation site for transcription is marked by a star and numbered +1, and a putative polyadenylation signal is double underlined. The introns and their sizes are indicated. Potential *cis* elements in the 5′ upstream region are underlined. PCR primer and tissue section in situ hybridization probe sequences used in this study are boxed. The zebrafish *rbp1b* gene and its 5′ upstream sequence were identified from a zebrafish genomic DNA sequence assembly of clone CH211-11908 (GenBank accession number: AL953896).

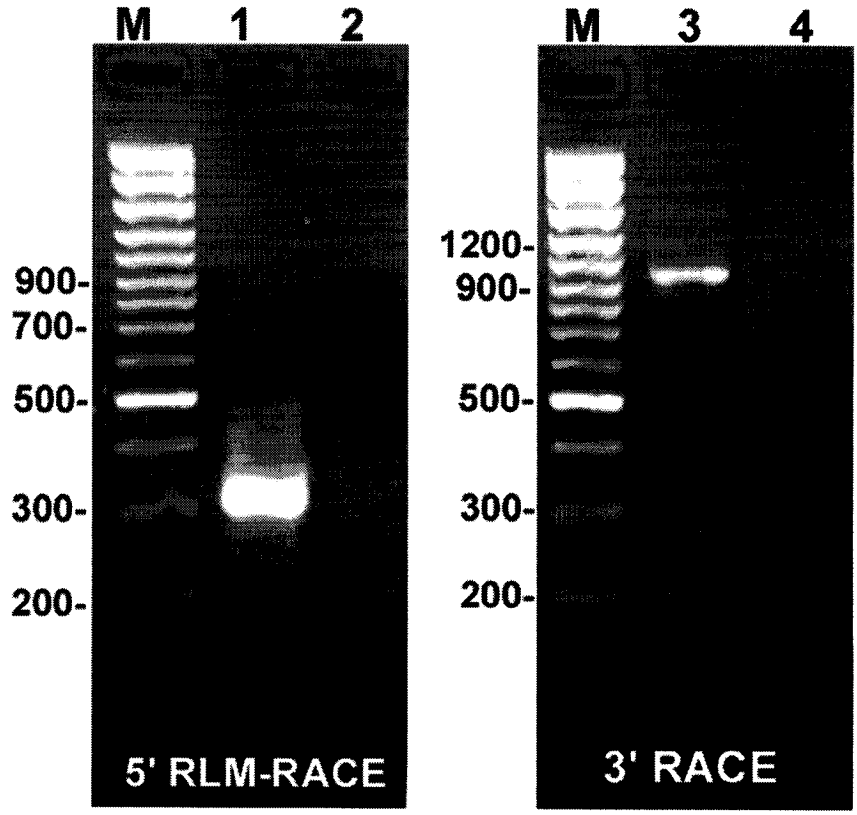



**Figure 7-2. Products of zebrafish *rbp1b* cDNA ends generated by 3′ RACE and 5′ RLM-RACE for cDNA cloning and transcription start site mapping.** Lane 1, product of 5′ RLM-RACE; Lane 2, negative control of 5′ RLM-RACE (RNA without tobacco acid pyrophosphatase treatment); Lane 3, product of 3′ RACE; Lane 4, negative control of 3′ RACE without the reverse transcription template. Lanes indicated M are 100 bp molecular weight marker. The sizes (in bp) of the molecular markers are indicated to the left of each panel.

The 5′ upstream DNA sequence of the zebrafish *rbp1b* gene was A/T rich. A/T nucleotides accounted for over 60% of the total nucleotides in that region. Inspection of the 5′ upstream sequence of the zebrafish *rbp1b* gene with MatInspector professional 6.2 [Quandt *et al.*, 1995] revealed no TATA box within the proximal promoter region, a feature shared by the zebrafish *rbp1a* gene as well as *rbp1* gene from other vertebrate species [Ong *et al.*, 1994; Liu *et al.*, 2003, in press]. Several *cis* regulatory elements, including a binding element for the peroxisome proliferator-activated receptor/retinoid x receptor heterodimer (PPAR/RXR), a progesterone receptor element (PRE), a thyroid transcription factor-1 (TTF1) binding site and a GATA2 element, were observed in upstream region of the zebrafish *rbp1b* gene (Fig. 7-1). Peroxisome proliferator-responsive element (PPRE) and/or retinoic acid-responsive element (RARE) has been found in the 5′ upstream regulatory sequence of several other iLBP genes [reviewed by Ong *et al.*, 1994; Bernlohr *et al.*, 1997]. Progesterone receptor (PR) and GATA2 are supposed to be transcription factors involved in mediation of gene expression in *Xenopus* oocytes [Partington, *et al.*, 1997; Maller, 2001]. Thyroid hormones and the timely expression of its receptors are important for embryonic to larval transitory phase in zebrafish [Liu and Chan, 2002].

### 7.2.2 Isolation and analysis of the zebrafish *rbp1b* cDNA

To further confirm that the potential zebrafish *rbp1b* gene was expressed and to characterize the spatial and temporal distribution patterns of its transcript, I isolated *rbp1b* cDNA. In addition to the 5′ RLM-RACE, I performed 3′ RACE using a gene-specific sense primer corresponding to a sequence in the 5′ UTR (Fig. 7-3A). A single

**A**

```
                                                    5′ gagtgctgcacgacaggagct  21
gaacgaccaggcatcaggtctgcattataaacacatttttaggacaacaatcgcgttgcctgtagttcaaa  92
ATG CCT GAT TAC ACT GGT TAC TGG AAA ATG ATC TCC AAC AAC AAC TTT GAG GAG  146
Met Pro Asp Tyr Thr Gly Tyr Trp Lys Met Ile Ser Asn Asn Asn Phe Glu Glu

TAT CTC AAA GCT CTT GAT GTT AAT GTT GCT ATC AGG AAA ATT GCT ATT TTA TTA  200
Tyr Leu Lys Ala Leu Asp Val Asn Val Ala Ile Arg Lys Ile Ala Ile Leu Leu

AGA CCC GAT AAA GAT ATC AGC CAG AAT GGA GAT CAC TTT GTC ATC AAG ACT GTC  254
Arg Pro Asp Lys Asp Ile Ser Gln Asn Gly Asp His Phe Val Ile Lys Thr Val

AGC ACA TTT AAG AAC TAC GAC ATG GAC TTT GTA GTT GGT CAG GAG TTT GAG GAA  308
Ser Thr Phe Lys Asn Tyr Asp Met Asp Phe Val Val Gly Gln Glu Phe Glu Glu

GAC CTG GGT CCA GTT GAT GGC AGG AAA TGT ATG ACC ACC ATC ACT TGG GAT GGA  362
Asp Leu Gly Pro Val Asp Gly Arg Lys Cys Met Thr Thr Ile Thr Trp Asp Gly

GAC AAG CTT GTG TGT GTC CAG AAA GGC GAG GTT GAA GGC CGG GGC TGG ACC CAA  416
Asp Lys Leu Val Cys Val Gln Lys Gly Glu Val Glu Gly Arg Gly Trp Thr Gln

TGG ATA GAA GGA GAT GAA CTT CAT TTG GAA CTG AGA GCA GGG GGA GTT GTG GGC  470
Trp Ile Glu Gly Asp Glu Leu His Leu Glu Leu Arg Ala Gly Gly Val Val Gly

AAA CAG ATT TTC AAG AAG TCT TAA accaacgttgtgtagtggtgccctctgctggagaagaaa  533
Lys Gln Ile Phe Lys Lys Ser  #
caaacaaccaccatagtcagaactgcatctgcccatccactcattcatcagtccgcgtgtttacattttct  604
ggaagtgcctaagctattccagtgtgtttagttaatataacagacacctgtaatgcatatatatgaattgc  675
gtatgtaccaagtatggttggttttagtgctgtagtcactaagattaacccttgagagattgacttctgt  746
ttaaaatactttctctaacatctgagggtctttttggctggaggaaaaatatcagctcttgaccatctaag  817
caaatccaaatagtttttttctttttcacaagctggaaggagataaatgtggtgagagatgatgtgatgttc  888
agttcctgcaccgctgcatgtaatggagatgtgtgctctaaaaaatgtacaaataacaactg attaaa atga  959
agaaagactaac(a)n                                                          971
```

**B**

```
CRBP1b  M--PDYTGYWKMISNNNFEEYLKALDVNVAIRKIAILLRPDKDISQ--NGDHFVIKTVSTFKNYDMDFVVGQ
CRBP1a  MSK.N...VYN...QD......A.....Y.L..VVCI.K.S.H.EHDV.TGRMK....T....F....TL..

CRBP1b  EFEEDLGPVDGRKCMTTITWDGDKLVCVQKGEVEGRGWTQWIEGDELHLELRAGGVVGKQIFKKS-
CRBP1a  ..T..........Q..VD......I...R..K......H.L..NI.......Q..TA..V...VE   66%
```

**Figure 7-3. Nucleotide sequence of the zebrafish *rbp1b* cDNA and alignment of its deduced amino acid sequence with that of the zebrafish *rbp1a*.** (**A**) Complete nucleotide sequence of the zebrafish *rbp1b* cDNA was determined by 3′ RACE and 5′ RLM-RACE (GenBank accession number: AY395732). The coding nucleotides are shown in uppercase letters and the 5′ and 3′ UTRs are in lowercase letters. Numbers at right correspond to the nucleotides in the cDNA sequence. A potential polyadenylation signal is in bold. The insertional variation of a codon in the cloned cDNA sequence is boxed. The antisense probe used for tissue section *in situ* hybridization is double underlined. (**B**) The deduced amino acid sequence of the zebrafish CRBP1b was aligned with the amino acid sequence of the zebrafish CRBP1a (GenBank accession number: AY309074). Dots indicate amino acid identity. The percentage amino acid sequence identity between the zebrafish CRBP1b and CRBP1a is shown at the end of the sequence.

3′ RACE product of ~ 950 bp was produced, cloned and sequenced (Fig. 7-2, lane 3). No RACE product was generated in the negative control without reverse transcription templates (Fig. 7-2, lane 4). The complete nucleotide sequence of the zebrafish *rbp1b* cDNA was determined by combining the sequencing data of both 5′ RLM-RACE and 3′ RACE. Alignment of the zebrafish *rbp1b* cDNA sequence with the gene sequence revealed a GAG insertion in the cDNA sequence at position 144-146 (Fig. 7-3A). The extra GAG in the cDNA sequence resulted in an insertion of an amino acid of glutamic acid in exon 1. This discrepancy between the cDNA and genomic sequence is unlikely a PCR artifact, because independent clones from separate 3′ RACE and 5′ RLM-RACE reactions contained this extra GAG while the *rbp1b* gene sequence and an expressed sequence tag (EST) record from GenBank (accession number: BI886241) did not contain these three nucleotides. At the present time, I do not know whether this variation has any functional implications. Except for the difference of the above GAG existence, the coding sequence of the zebrafish *rbp1b* gene and the cloned cDNA were identical.

The deduced amino acid sequence from the zebrafish *rbp1b* cDNA consisted of 133 amino acids with a theoretical isoelectric point of 4.94 and molecular mass of 15.2 kDa. The zebrafish CRBP1b sequence was aligned with a number of mammalian CRBPs and CRABPs, and it showed highest sequence identity with that of the mammalian CRBP1 (~ 74%) and the zebrafish CRBPIa (66%, Fig. 7-3B)), followed by mammalian and zebrafish CRBP2s (~56%), human CRBP3 (54%), human CRBP4 (51%), mammalian CRABP1 (~33%) and mammalian CRABP2 (29-33%).

### 7.2.3 Gene phylogeny of the zebrafish *rbp1*

Phylogenetic analysis was performed to reveal the evolutionary relationship among the newly identified zebrafish *rbp1b* gene and other cellular retinoid-binding protein genes from mammals and fishes (Fig. 7-4). The results showed that all mammalian and fish CRBP1 genes clustered together on the same clade, indicating that they are orthologous genes (Fig. 7-4). The zebrafish CRBP1b and a pufferfish EST (GenBank accession number: CA341008) that is similar to CRBP1 formed a sub-branch within the CRBP1 clade, suggesting a closer relationship between these two CRBP1s.

### 7.2.4 Linkage mapping and conserved syntenies of the zebrafish *rbp1b* gene

To reveal the linkage relationship of the zebrafish *rbp1b* gene, I performed radiation hybrid mapping analysis using the LN54 panel [Hukriede *et al.*, 1999]. The zebrafish *rbp1b* gene was assigned to the zebrafish linkage group (LG) 2 with a LOD score of 13.4 to the overall markers on LG 2 and a distance of 5.23 cR from the marker Z211023 (mapping data can be provided upon request). By examining an existing comparative map of the zebrafish LG 2 [Woods *et al.*, 2000] and data from LocusLink (http://www.ncbi.nlm.nih.gov/locuslink), I found that both *rbp1a* [Liu *et al.*, 2003, in press] and *rbp1b* have conserved syntenic genes with the human *RBP1* gene on chromosome 3 and mouse *Rbp1* on chromosome 9 although the zebrafish *rbp1s* are located on different LGs (Fig. 7-5). This observation indicated that the zebrafish *rbp1* genes may have been derived from a duplicated chromosome rather than duplication of only the *rbp1* gene (Fig. 7-5).

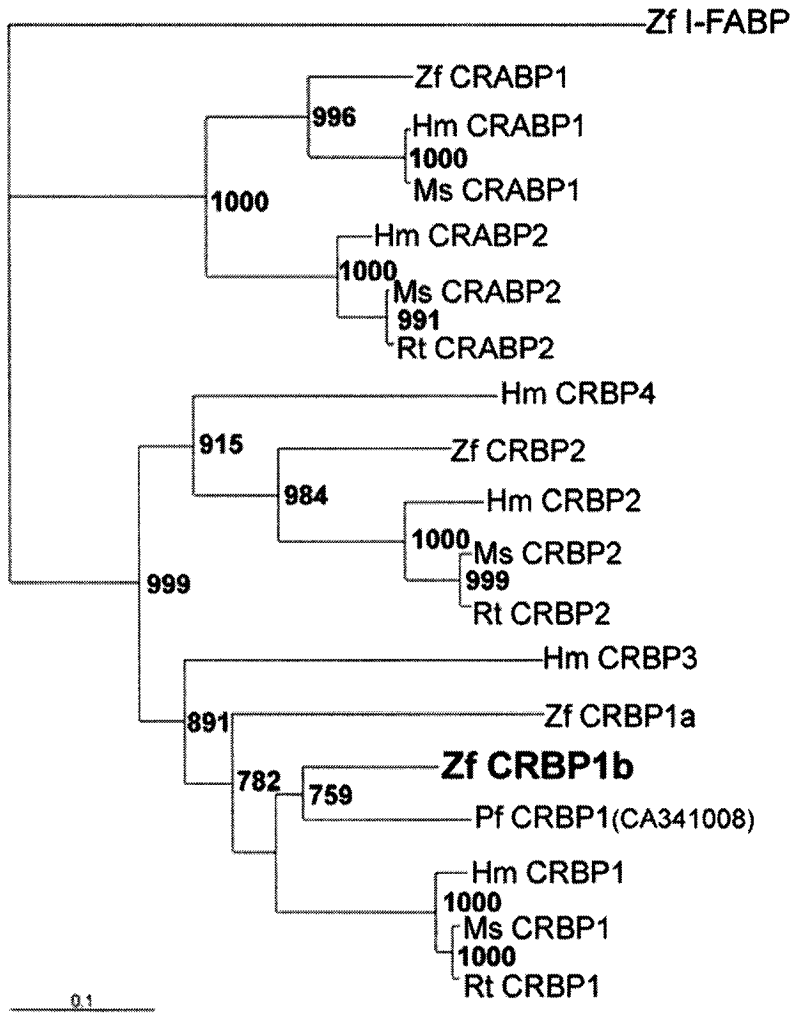



**Figure 7-4. Phylogenetic tree of the cellular retinoid-binding proteins.** The bootstrap neighbour-joining phylogenetic tree was constructed with ClustalX [20] using the zebrafish (Zf) intestinal type fatty acid-binding protein (I-FABP) amino acid sequence (Swiss-Prot and TrEMBL accession number Q9PRH9) as an outgroup. The bootstrap values (based on number per 1000 duplicates) are indicated on the nodes. Sequences used in this analysis include the amino acid sequences for human (Hm) CRBPI (P094555), mouse (Ms) CRBPI (Q00915), rat (Rt) CRBPI (P02696), human CRBPII (P50120), mouse CRBPII (Q08652), rat CRBPII (P06768), human CRBPIII (P82980), human CRBPIV (Q96R05), zebrafish CRABPI (Q7T0F4), human CRABPI (P29762), mouse CRABPI (P02695), human CRABPII (P29373), mouse CRABPII (P22935), rat CRABPII (P51673). Scale bar = 0.1 substitutions per site.

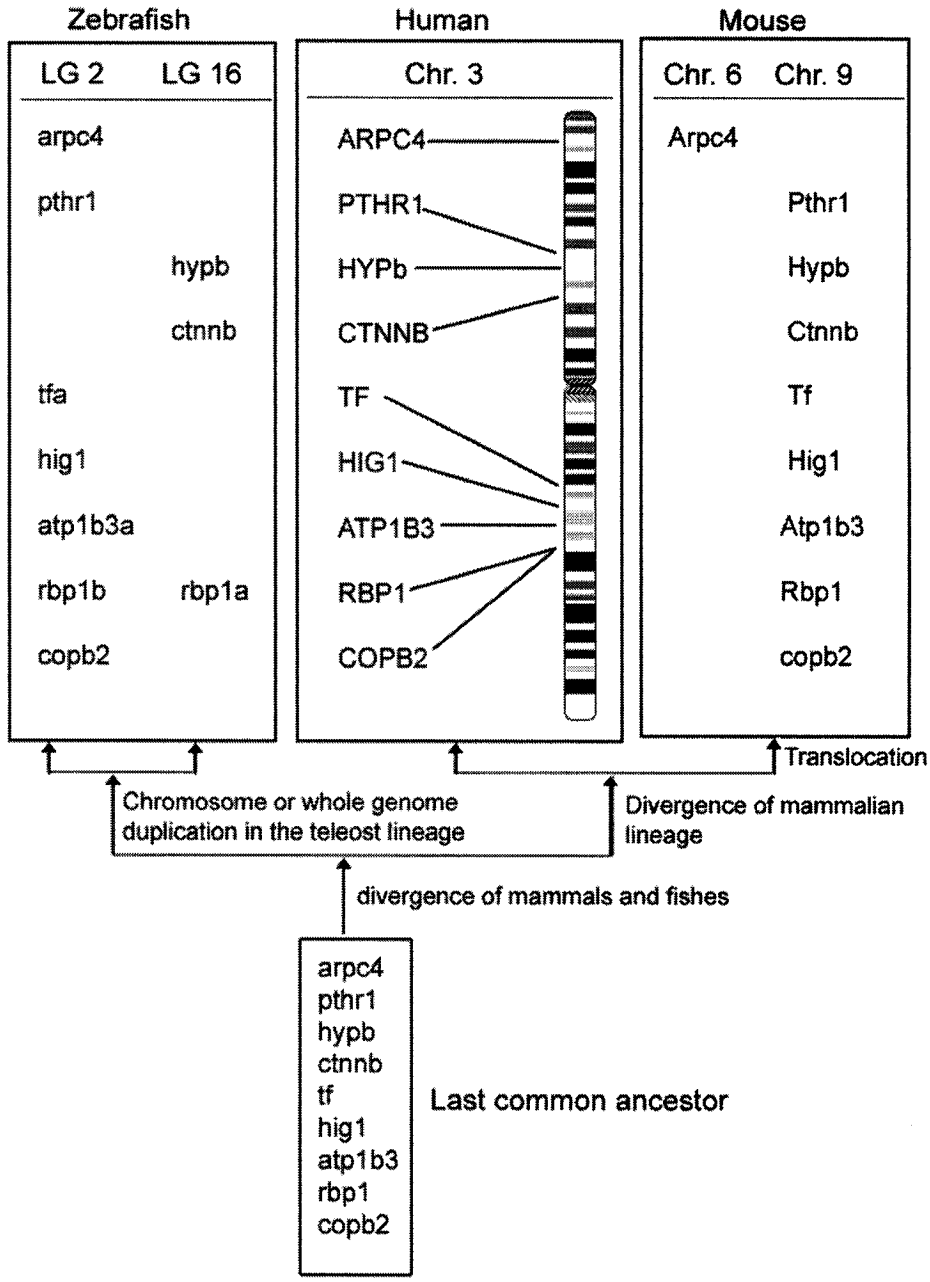



**Figure 7-5. Comparison of syntenies around *rbp1* genes from zebrafish, human and mouse.** Both the zebrafish *rbp1a* gene on LG 16 and *rbp1b* gene on LG 2 have conserved syntenies (left box) with the human *RBP1* gene on chromosome 3 (middle box) and the mouse *Rbp1* gene on chromosome 6 and 9 (right box). The order of the human *RBP1* gene and its syntenic genes was determined based on the cytogenetic mapping data from LocusLink (http://www.ncbi.nlm.nih.gov/locuslink) and the gene loci on the zebrafish and mouse LGs and the last common ancestor's chromosome are listed in the order appearing on the human chromosome 6. Hypothesized evolutionary events leading to the syntenic relationships of the zebrafish, human and mouse *rbp1* genes are indicated at the nodes of the evolutionary tree.

### 7.2.5 Specialized distribution of *rbp1b* transcripts in the ovary of adult zebrafish

The distribution of *rbp1b* mRNA in adult tissues was determined by RT-PCR using *rbp1b* gene-specific primers (Fig. 7-6). RT-PCR products of the expected size were generated from RNA of the ovary, but not from RNA of any other tissues examined including the liver, skin, intestine, brain, heart, muscle and testis (Fig. 7-6). RT-PCR products were generated from RNA of all the tissues examined with specific primers corresponding to the cDNA sequence of the constitutively expressed receptor for activated C kinase (RACK1) gene [Hamilton and Wright, 1999]. No RT-PCR products were detected from the negative controls that lacked the reverse-transcribed cDNA templates (Fig. 7-6).

To confirm the ovary-specific expression of the zebrafish *rbp1b* gene, I performed tissue section *in situ* hybridization with an antisense oligonucleotide probe corresponding to the 3′ UTR of the zebrafish *rbp1b* cDNA (Fig. 7-7 A). Strong hybridization signals were specifically detected in the zebrafish ovary, but not in any other tissues on the sections (Fig. 7-7 A1). The negative control with a sense probe showed no hybridization signal (Fig. 7-7 A2).

To further demonstrate the distribution of the zebrafish *rbp1b* mRNA in the ovary at the cellular level, I conducted emulsion autoradiography of the *in situ* hybridization tissue sections (Fig. 7-7 B). Silver grains corresponding to *rbp1b*-specific hybridization were densely clustered over the primary (stage I) oocytes (Fig. 7-7 B1-3), but not in the stage II or III oocytes (Fig. 7-7 B2). *rbp1b*-specific hybridization was not observed in any other tissues including muscle (Fig. 7-7 B1, 3) or liver (Fig. 7-7 B3).

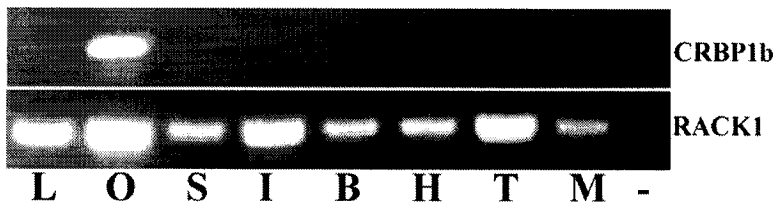



**Figure 7-6. Tissue-specific expression patterns of the *rbp1b* gene in adult zebrafish detected by RT-PCR.** The zebrafish *rbp1b* cDNA-specific primers amplified a PCR product from total RNA extracted from the ovary (O), but not from the liver (L), skin (S), intestine (I), brain (B) and testis (T), heart (H) and muscle (M). A RT-PCR product corresponding to the constitutively expressed, receptor for activated C kinase (RACK1) was generated from RNA in all samples. A negative PCR control (-) did not contain cDNA template and did not generate any RT-PCR product.

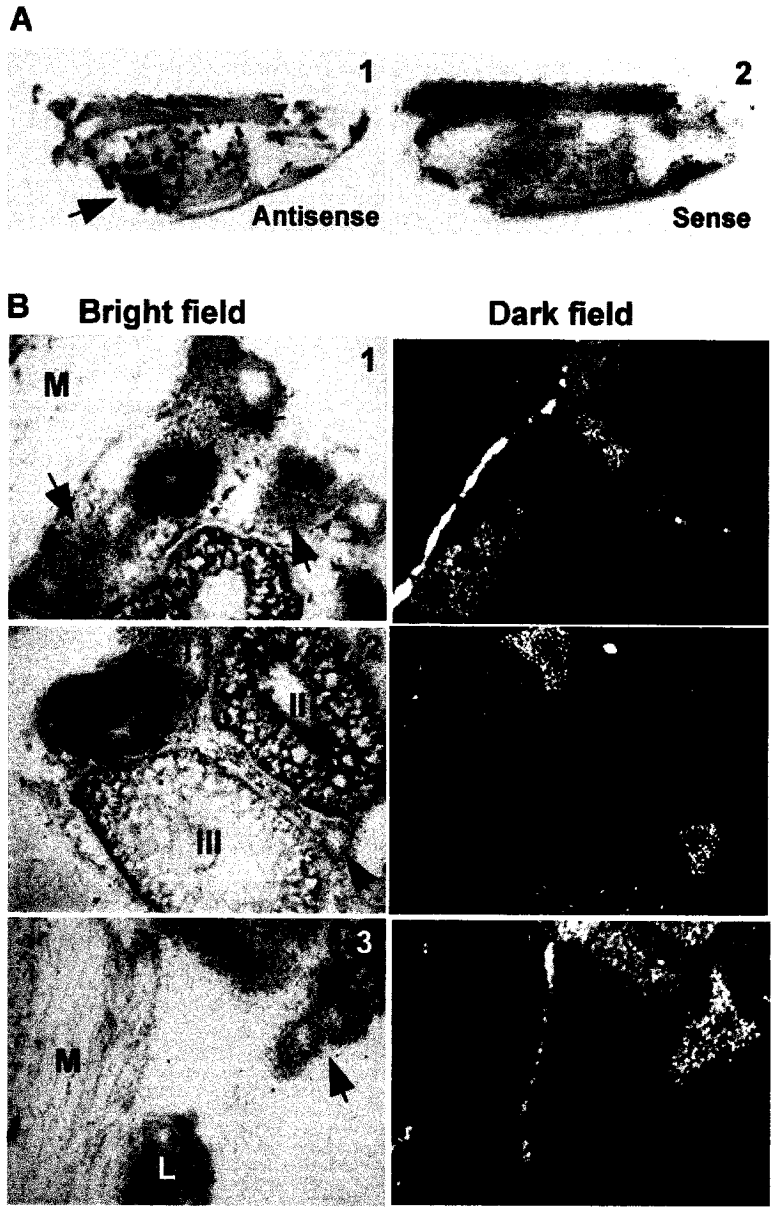



**Figure 7-7. Specific distribution of zebrafish *rbp1* mRNA in the primary oocytes detected by tissue section *in situ* hybridization and autoradiography.** (A) *in situ* hybridization detected *rbp1b* mRNA in the ovary (arrow) on the tissue section hybridized with a *rbp1b* cDNA-specific antisense oligonucleotide probe (A1), but not with the sense probe (A2). (B) Bright and dark field showed the presence of *rbp1b* mRNA (silver grains) at the cellular level by emulsion autoradiography. Silver grains were observed in the stage I primary oocytes (arrows, B1-3), but not in stage II, stage III oocytes (B2), nor in the muscle (M) and liver (L).

### 7.2.6 The zebrafish *rbp1b* mRNA was detected in the developing gall bladder

I examined the spatial and temporal distribution of the zebrafish *rbp1b* transcripts during embryonic and larval stages using whole mount *in situ* hybridization with an antisense RNA probe synthesized from the zebrafish *rbp1b* cDNA plasmid (Fig. 7-8). The zebrafish *rbp1b* mRNA was not detected in the early developing zebrafish until 5 days post fertilization (dpf). In the 5 dpf larvae, *rbp1b* transcripts were restricted to the developing gall bladder (Fig. 7-8 A-D). Hybridization signals were not observed in any other embryonic or larval tissues.

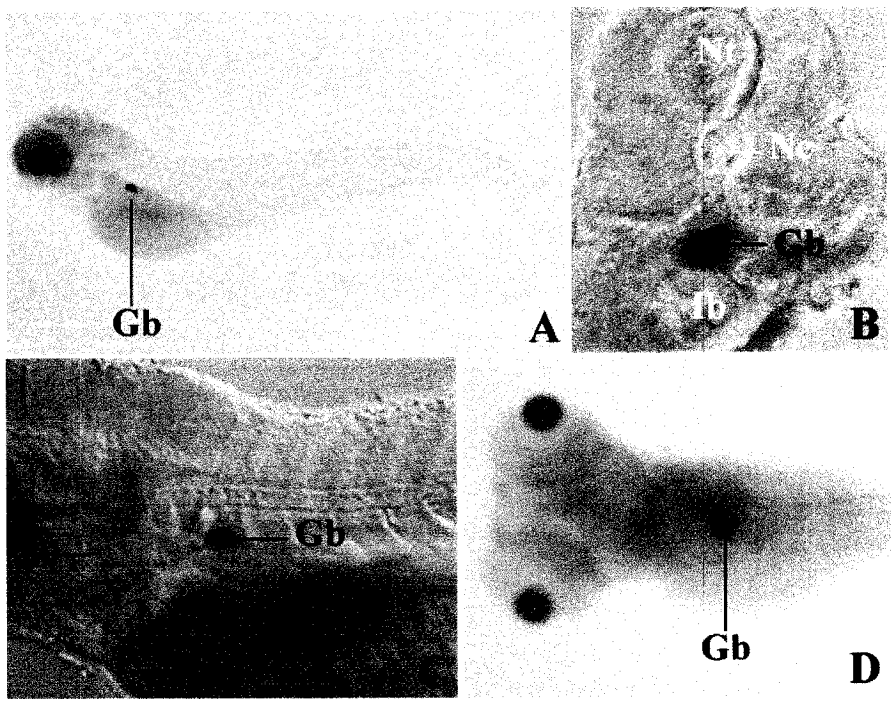



**Figure 7-8. Whole mount *in situ* hybridization of 5 day old zebrafish larvae.** *rbp1b*-specific antisense RNA probe detected *rbp1b* mRNA in the developing gall bladder (Gb) of the larvae at 5 days post fertilization (5dpf). (A) Lateral view, head to the left. (B) Cross section through the gall bladder. (C) sagittal section through the gall bladder. (D) dorsal view, head to the left. Ib, intestinal bulb; Nc, notochord; Nt, neural tube.

## 7.3 DISCUSSION

I have described a duplicate gene encoding cellular retinol-binding protein 1 from zebrafish and the tissue-specific expression of *rbp1b* during development and adulthood. Therefore, zebrafish, unlike mammals, have two *rbp1* genes, *rbp1a* [Liu *et al.*, in press] and *rbp1b*.

Phylogenetic analysis showed that the vertebrate iLBP multigene family may have arisen from a common ancestral gene through at least 14 gene duplications before the divergence of vertebrates and invertebrates [Schleicher *et al.*, 1995]. A number of iLBP genes have been identified in fishes including those encoding cellular retinoid-binding proteins [Kleinjan *et al.*, 1988; Cameron *et al.*, 2002; Sharma *et al.*, 2003; Liu *et al.*, in press]. In the present study, the duplicated *rbp1* gene, *rbp1b*, encodes a polypeptide sharing high sequence similarity to the mammalian and fish CRBP1. The zebrafish *rbp1a* [Liu *et al.*, in press] and *rbp1b* genes are located on different LGs but both have conserved syntenies with the orthologous mammalian *rbp1* genes, indicating that zebrafish *rbp1* genes arose by chromosomal or, possibly, whole genome duplication rather than duplication of the ancestral *rbp1* gene.

There are currently three existing theories regarding the evolutionary fate of duplicate genes: First, the nonfunctionalization theory states that one of the gene duplicates become silenced due to degenerative mutations and it is eventually lost from the genome or become a pseudogene. Second, the neofunctionalization theory states that one of gene duplicates acquires a new and beneficial function while the other duplicate retains the function of the ancestral gene, such that both genes are preserved. Lastly the subfunctionalization theory states that, after duplication, degenerative mutations in the

*cis*-regulatory elements result in partitioning of the original functions into each duplicate, such that the ancestral function is divided between the duplicates and duplicates functionally complement each other and are retained. This last theory is also known as the duplication-degeneration-complementation (DDC) model [Force *et al.*, 1999 and references therein].

The mammalian *Rbp1* genes are expressed in several adult tissues. In rats, CRBPI is abundantly distributed in the liver and kidney, and, to a lesser density, in the lung, testis, spleen, eye, ovary, uterus and intestine [reviewed by Ong *et al.*, 1994], while in humans, the ovary contains the highest amount of *RbpI* gene transcripts and protein, whereas the levels of RBP1 are relatively low in human kidney and lung [Fex and Johannesson, 1984; Ong and Page, 1986; Folli *et al.*, 2001]. Therefore, the relative expression levels of the *RBP1* gene in mammalian adult tissues vary among species. *RbpI* mRNA and protein levels remain in the rat uterus and ovary during the estrous cycle, and *Rbp1* is specifically expressed in oocytes [Wardlaw *et al.*, 1997]. Expression of *RbpI* in these reproductive organs suggests a role for the protein in the metabolism and storage of vitamin A (retinol) for the normal uterine function and ovarian follicular development. During development, CRBPI protein and mRNA are abundant in the mammalian central nervous system (CNS). At the early somite stage of the mouse embryo, CRBPI is widely distributed from the posterior neural tube up to the hindbrain, but very low levels are seen in the forebrain [Ruberte *et al.*, 1991]. CRBPI is also detected in the mouse and rat olfactory bulb, spinal cord and hippocampus [Perez-Castro *et al.*, 1989; Gustafson *et al.*, 1999; Zetterstrom *et al.*, 1994; 1999; Romand *et al.*, 2000].

In zebrafish, of the two duplicated *rbp1* genes, the transcripts of *rbp1a* were detected at low levels in several adult tissues and were abundant in the developing CNS and retina, showing a similar distribution during development to that of the mammalian orthologs [Liu *et al.*, 2003, in press]. In contrast, *rbp1b* was expressed in the adult ovary but not in the developing embryonic CNS. It appears that *rbp1b*, but not *rbp1a* and mammalian *rbp1* was expressed in the developing gall bladder at 5 dpf. Neither zebrafish *rbp1a* nor *rbp1b*, however, was expressed in the adult liver, one of the major tissues showing mammalian CRBP1 distribution in mammals [Ong. *et al.*, 1994] and retinoid metabolism [Blaner and Olson, 1994; Napoli, 1996].

Considering the expression data of the mammalian and the zebrafish *rbp1* gene, it may be assumed that the common ancestral *rbp1* gene be expressed in multiple adult tissues including ovary, liver and in the early developing CNS and retina. It is clear that the duplicated zebrafish *rbp1* genes complimented each other to retain some of the ancestral subfunctions, as was evident by the specialized distribution of *rbp1b* transcripts in the oocytes and *rbp1a* transcripts in the developing CNS and retina [Liu *et al.*, 2003c, in press]. A new feature of the zebrafish *rbp1* duplicated genes is that they were subfunctionalized both spatially and temporally, which is a development of the DDC model of duplicate gene retention. I propose that temporal subfunctionalization might be a common feature for developmentally-regulated gene duplicates. Neither *rbp1a* nor *rbp1b* is expressed in the adult liver and has not, therefore, retained the function of the ancestral *rbp1* gene. This suggests that in addition to non-, sub- and neofunctionalization, some gene duplicates may have undergone "partial nonfunctionalization". In contrast, both the zebrafish *rbp2a* [Cameron *et al.*, 2001; Liu *et al.*, in press] and *rbp2b* [Liu *et al.*,

unpublished data] gained novel expression specificity after duplication in the adult liver as their mRNAs were abundantly detected in this tissue and the mammalian *rbp2* is expressed in the intestine but not the liver. Since CRBP1 and CRBP2 share similar ligand binding specificity and other physiological properties [Ong *et al.*, 1994], it is possible that the partial nonfunctionalization of *rbp1* duplicates has been complemented by the duplicates of the closely related paralogous gene, *rbp2*. In addition, a new function has evolved for the zebrafish *rbp1b*, as its mRNA was specifically detected in the developing gall bladder. Taken together, I conclude that preservation of the zebrafish *rbp1* duplicated genes may be a result of combined subfunctionalization and neofunctionalization, and that partial nonfunctionalization of the duplicate genes with multi-functions (e.g. zebrafish *rbp1a* and *b*) might lead to the retention of duplicates from uni-functional genes such as the zebrafish *rbp2*.

Inspection and comparison of the 5′ regulatory elements of the duplicated zebrafish *rbp1* genes showed that the *rbp1a* gene contains no elements for RAR and PPAR but does contain several POU-domain elements which are associated with gene regulation in the developing mammalian brain [He *et al.*, 1989]. *rbp1b* has elements involved in gene regulation in oocytes and a PPAR/RXR binding site. The divergent evolution of the 5′ *cis* elements of the zebrafish *rbp1a* and *rbp1b* may have led to the complementary differential expression patterns of these two gene duplicates. Functional analysis is needed to demonstrate which putative *cis* elements in the 5' upstream regions of the zebrafish *rbp1* duplicates control gene expression.

Egg yokes from oviparous animals are one of the richest sources of vitamin A. In fish and other oviparous animals, retinol must be taken up from the serum by the rapidly

growing oocytes, and then stored and metabolized prior to functioning during embryonic development. Retinoids have been detected in high levels in unfertilized zebrafish eggs and during embryogenesis [Costaridis *et al.*, 1996]. Holo-CRBPI (retinol-CRBPI complex) is the substrate for lecithin:retinol acyltransferase (LRAT), an enzyme catalyzing the esterification of retinol with long chain fatty acids to form retinyl ester for storage. Apo-CRBPI strongly inhibits the esterification reaction [Herr and Ong, 1992]. Due to the significance and distinct features of the retinoid mobilization and metabolism in the fish ovary, it is reasonable to speculate that *rbp1b* has evolved and subfunctionalized to play a particular role in retinol uptake, metabolism and storage in the fish ovary.

# 8 Conclusion

## 8.1 EXPRESSION OF ILBP GENES IN VERTEBRATES

Distribution of the iLBP gene transcripts and/or proteins is generally conserved among mammalian species [reviewed by Ong *et al.*, 1994; Glataz *et al.*, 1996; Bernlohr *et al.*, 1997]. However, A few exceptions have been found. The gene for CRBPI, for example, shows striking difference in expression patterns between rat and human as discussed in Chapter 6 and 7 [reviewed by Ong *et al.*, 1994], suggesting that the expression patterns of iLBP genes might have diverged during the evolution of vertebrates to fulfill the differentiated physiological requirements of each species. Zebrafish are markedly different in some physiological processes from that of mammals (e.g. reproductive systems), but similar for others (e.g. central nervous systems). In the present study, I found that the differences in gene expression patterns of the iLBP genes between mammals and zebrafish reflected these physiological differences. Specificity of gene expression in the central nervous system (*fabp7a*, *rbp1a*) is well conserved between zebrafish and mammals, while novel expression specificities are evident for some zebrafish iLBP genes, which might reflect their specialized physiological functions (e.g. *fabp3* in the oocytes, Chapter 3; *fabp7b* in swim bladder, Chapter 5).

It is important to know the mechanisms underlying the conservation and divergence of the spatio-temporal expression patterns of iLBP genes between zebrafish and mammals. Although identities of the deduced amino acid sequence for zebrafish iLBPs with their mammalian orthologs varied among paralogous genes, there was no

relationship between the sequence identity levels and expression similarities of orthologous genes. The same is true for the mammalian iLBP genes. Taking, again, the mammalian *Rbp1* gene for example, human and rat CRBPI share 96% identity, but show very different tissue distribution [reviewed by Ong *et al.*, 1994]. Although further functional analysis is needed to determine the *cis* regulatory elements governing the spatio-temporal expression the zebrafish iLBP genes, computer-assisted analysis of the 5' upstream sequence suggested that particular distribution of *cis* elements in the 5' upstream regulatory region of the zebrafish iLBP genes might account for their specific expression and function.

## 8.2 GENE DUPLICATION IN THE ILBP MULTIGENE FAMILY

The hypothesis of "large scale gene duplication in fishes" [Gates *et al.*, 1999; Postlethwait *et al.*, 2000; Woods *et al.*, 2000; Robinson-Rechavi *et al.*, 2001; Taylor *et al.*, 2001; 2003] prompted a search for gene duplicates in the zebrafish genome sequence database (The Wellcome Trust Sanger Institute, Cambridge, UK) for each of the paralogous iLBP members identified in these studies, duplicates for the *fabp3*, *fabp7*, *rbp1*, *rbp2* and *crabp1* genes. Each of these genes has a duplicate copy in the zebrafish genome. In this thesis, the duplicate pairs of *rbp1* and *fabp7* were described in addition to one duplicated copy of *fabp3* and *rbp2* gene from zebrafish. It is apparent from these results that gene duplication is common in the zebrafish iLBP multigene family. In the mammalian genomes, only one copy of each of the above five paralogous gene has been identified [reviewed in Glatz *et al.*, 1996; Bernlohr *et al.*, 1997; Ong *et al.*, 1994]. Although duplicated *rbp1* and *rbp2* genes were identified in the zebrafish genome, we

could not find the orthologous genes for the recently reported mammalian genes for CRBPIII and CRBPIV [Folli *et al.*, 2001; 2002; Vogel *et al.*, 2000]. In addition, *fabp3* psudogenes have been found in the mammalian genome, although only one functional *fabp3* gene has been identified [Heuckeroth *et al.*, 1987]. Therefore, based on the available information, at least the number of gene loci for CRBP and H-FABP in the zebrafish genome is not necessarily more than that in the mammalian genomes.

## 8.3 INDEPENDENT GENE DUPLICATION VERSUS WHOLE GENOME DUPLICATION IN THE FISH LINEAGE

Whether independent gene duplication or ancient whole genome duplication account for the "more fish genes" phenomenon in the teleost fish lineage is a controversial issue. Since Amores *et al.* [1998] proposed a whole genome duplication in the teleost fish lineage based on the finding of duplicate *hox* complexes in the zebrafish genome, several other investigations revealed a large number of fish-specific duplicate orthologs through gene mapping and phylogenetic analysis, supporting the "ancient fish-specific genome duplication" hypothesis [Gates *et al.*, 1999; Postlethwait *et al.*, 2000; Woods *et al.*, 2000; Taylor *et al.*, 2001; 2003]. On the contrary, Robinson-Rechavi *et al.* [2001] noted that the redundant fish genes may result from frequent independent duplications rather than an ancestral genome duplication after examining the phylogenies of several gene families. I analyzed the phylogenies and syntenic relationships of two pairs of duplicate iLBP genes. The results favor a whole genome duplication hypothesis. First, duplication is common for the iLBP genes on various linkage groups (100% of the genes examined, data not shown), indicating a large scale gene duplication may have

shaped the zebrafish genome. Second, each paralogous duplicate resides on a different linkage group, excluding the possibility of independent tandem gene duplication. Third, each pair of duplicate genes has conserved syntenies to the same mammalian orthologs, indicating that each duplicate gene might derive from a chromosomal duplication.

## 8.4 MECHANISM FOR PRESERVATION OF DUPLICATED GENES

There is no doubt that large scale gene duplication has significantly shaped the genomes of vertebrate species. Lynch and Conery [2000] estimate that in eukaryotic genomes duplicate genes arise at a rate of 0.01 per gene per generation on average, suggesting that gene duplication is a frequent event. In theory, the majority of the duplicated genes will be silenced and eventually lost due to accumulation of deleterious mutations in a relatively short time [Bailey *et al.*, 1978; Li, 1980; Lynch and Conery, 2000]. A large fraction of duplicate genes, however, are not silenced in certain lineages for a considerable period of time span [reviewed by Force *et al.*, 1999]. A large number of duplicate genes are retained in the ray-finned fish, presumably resulting from ancient fish-specific genome duplication more than 300 mya [Taylor *et al.*, 2001; 2003]. The fate of duplicate genes and the mechanism for preservation of abundant ancient duplicate genes remains unclear and controversial. Conventional theory suggests that the fate of duplicate genes depends on the occurrence of null mutations, which lead to gene silencing or "nonfunctionalization", or beneficial mutations, which result in "neofunctionalization" or retention of gene duplicates [reviewed by Force *et al.*, 1999; Lynch and Conery, 2000]. The recent "subfunctionalization" theory or "duplication-degeneration-complementation (DDC)" model predicts that gene duplicates are preserved

owing to accumulation of degenerative mutations in regulatory elements and, consequently, each copy of the duplicate pair shares a subset of the ancestral functions [Force *et al.*, 1999; Lynch and Force, 2000].

By analyzing the tissue-specific distribution of gene transcripts of the duplication pair of zebrafish *fabp7* and *rbp1* using RT-PCR, tissue section and whole mount *in situ* hybridization, a complex pattern of functionalization of duplicate genes was found. Assuming that the duplication of the zebrafish *fabp7* and *rbp1* genes is fish lineage-specific and the functions (tissue-specific expression) of mammalian orthologs represent the ancestral functions, subfunctionalization was observed for *rbp1a* and *rbp1b* (Chapter 6, 7), whereas neofunctionalization is suggested for *fabp7a* and *fabp7b* (Chapter 4 and 5). For *rbp1*, both subfunctionalization (in the adult ovary) and neofunctionalization (in the developing gall bladder) occurred simultaneously (Chapter 7). In addition, it seemed that none of the *rbp1* duplicates retained expression and function in the liver (Chapter 6 and 7), which is one of the major tissues in which mammalian *rbp1* is expressed [reviewed in Ong *et al.*, 1994]. This suggests a partial nonfunctionalization of the zebrafish *rbp1* duplicates (Fig. 8-1). Based on the data obtained from this study, I propose that preservation of duplicate genes may not be simply "neofunctionalization" or "subfunctionalization" and may be a combination of both processes (Fig. 8-2). Nonfunctionalization may occur in part of the ancestral subfunctions and, therefore, may not necessarily silence the corresponding duplicate gene. Instead, partial neofunctionalization of duplicates from a multi-functional gene, such as the zebrafish

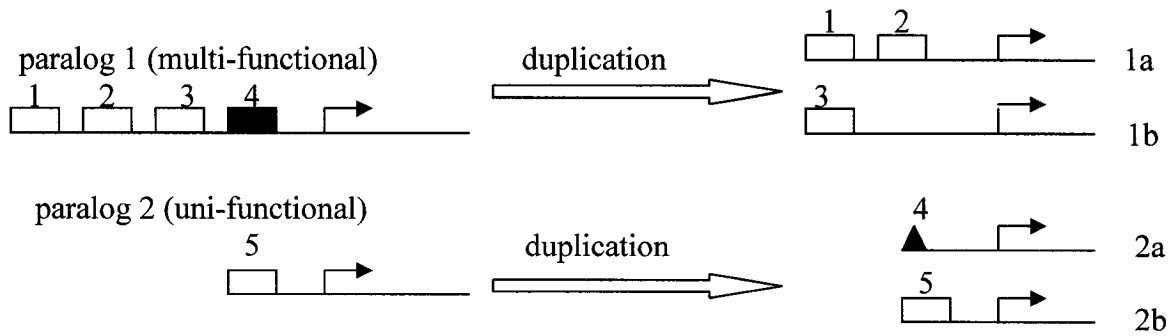



**Figure 8-1. A model for preservation of duplicates of uni-functional genes through "shuffling" subfunctions between closely related paraologous genes.** Mutation at *cis* element 4 results in partial nonfunctioalization of paralog 1 duplicates (1a, 1b). One of the duplicates (2a) of paralog 2, however, gained the *cis* element 4 and developed a novel function, which helped the duplicate gene 2a to remain in the genome after duplication. Boxes numbered 1 to 5 denote regulatory *cis* elements controlling different expressional or functional specificities. The solid box numbered 4 indicates that the subfunction of paralog 1 duplicates is complemented by a paralog 2 duplicate gene (2a). Arrows after *cis* elements denote transcription start sites of genes.

*rbp1*, might be a mechanism for preservation of the duplicates of a uni-functional gene, such as the zebrafish *rbp2*, by "shuffling" their subfunction(s) (Fig. 8-1).

## 8.5 FEATURES OF SUBFUNCTIONS

### 8.5.1 Spatial and temporal

iLBP genes are developmentally regulated and, therefore, each paralogous member exhibits a specific spatio-temporal pattern of expression [reviewed in Glatz *et al.*, 1996; Bernlohr *et al.*, 1997]. I propose that subfunctionalization in the iLBP duplicate genes occurred at two different levels: spatial and temporal (Table 8-1 and Fig. 8-2). Distribution of the transcripts of *rbp1a* and *rbp1b* genes, for example, is not only complemented in the tissues (*rbp1b* expression is specialized in the ovary), but also at the developmental stages (*rbp1a* gene transcripts were detected earlier during development than *rbp1b* gene transcripts).

**Table 8-1. Developmental regulation of iLBP duplicate genes**

| Gene | Development stages | | | | | | | Adulthood | |
|---|---|---|---|---|---|---|---|---|---|
| | G[a] | S[b] | 24h[c] | 36h | 48h | 3d[d] | 5d | RT-PCR | *In situ* |
| *fabp7a* | + | + | + | + | + | + | + | + | + |
| *fabp7b* | - | - | + | + | + | + | + | + | - |
| *rbp1a* | - | + | + | + | + | + | + | + | - |
| *rbp1b* | - | - | - | - | - | - | + | + | + |

a, gastrula; s, somitogenesis; c, hours post fertilization; d, day post fertilization;
+, presence of gene transcripts; -, absence of gene transcripts.

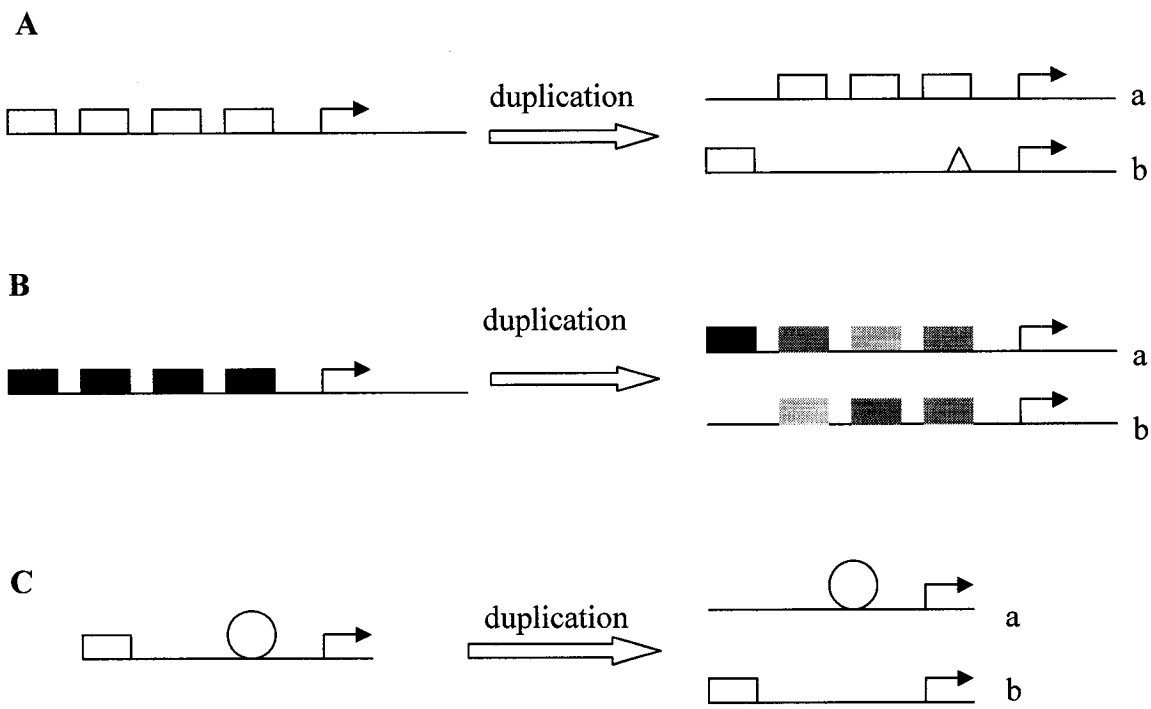



**Figure 8-2. A model for features of subfunctionalization of duplicated genes.** (A) Combination of subfunctionalization (boxes) and neofunctionalization (triangle) in a single duplicated pair. (B) Qualitative and quantitative subfunctionalization. *cis* elements controlling quantity of expression is indicated by boxes with different densities. (C) Degenerative mutations destroyed the *cis* element controlling spatial expression (box) in duplicate gene "a" and that controlling temporal expression during development (circle) in duplicate "b" result in subfunctionalization with spatial and temporal fashion.

### 8.5.2 Qualitative and quantitative

Based on the DDC model, subfunctionalization occurs by either a qualitative or a quantitative route [Force *et al.*, 1999]. Qualitative subfunctionalization results in absolute loss of a curtain subfunction for one duplicate gene copy and silencing of a different subfunction for the other gene copy, while in the case of quantitative subfunctionalization, both duplicated gene copies retain a certain function (indicated by their co-expression in a certain tissue), but the level of expression for each copy is reduced to a point that fulfills the requirement of the host organism [Force *et al.*, 1999]. Quantitative subfunctionalization, therefore, would be the common process for the zebrafish iLBP duplicate genes reported in this study (Fig. 8-2). For both duplicated copies of the zebrafish *fabp7* gene, transcripts were detected in the same set of adult tissues by RT-PCR, but their relative levels varied considerably (Chapter 5). This is also true for the duplicated *rbp1* genes as the transcripts of both genes were detected in the adult ovary, but at very low levels for *rbp1a* (*rbp1a* gene transcripts were only detected with 35 cycles of PCR and not detected by tissue section *in situ* hybridization, Chapter 6) and abundantly for *rbp1b* (revealed by both RT-PCR and tissue section *in situ* hybridization, Chapter 7).

### 8.5.3 Redundancy or complementation

Following duplication, functions of the duplicated genes are presumably redundant. The DDC model predicts that duplicate genes will lose different subfunctions due to degenerative mutations in their regulatory elements, but they must complement each other by jointly retaining the full set of ancestral functions [Force *et al.*, 1999]. Subfunctions of each copy of an iLBP duplicate pair examined in this study are, in

general, complementary. For example, although the mRNA for both *fabp7a* and *fabp7b* was detected in the developing brain and retina, they were distributed in distinct regions. Abundant co-expression of both duplicated genes at the same cellular level was not seen, although putative quantitative subfunctionalization, as mentioned above, was seen.

In one exception as mentioned above, both duplicated copies of the zebrafish *rbp1* gene seemed to have lost their expression capacity in the adult liver. As such, genes for *rbp1* did not complement each other to retain this ancestral subfunction following duplication. On the contrary, both zebrafish *rbp2a* (Cameron *et al.*, 2001; Chapter 6) and *rbp2b* (R-Z Liu, unpublished data) gained expression capacity in the adult liver despite none of the mammalian *rbp2* genes is expressed in this tissue. As such, I can infer from these data that subfunction in the liver for *rbp1* duplicates was complemented by the duplicates of a closely related paralogous gene, *rbp2*.

## 8.6 FINAL SUMMARY

While gene structure, deduced amino acid sequence and syntenic relationship of the zebrafish iLBP genes is conserved with their mammalian orthologs, their expression patterns are either conserved or diverged, presumably as a result of evolutionary divergence of transcriptional regulation. As it has been evident for other multigene families in ray-finned fish genomes, gene duplication is common in the zebrafish iLBP family, and the duplicated genes might arise from chromosomal or, possibly, whole genome duplications. In supplementation to the DDC model for the mechanism of preservation of duplicate genes, gene duplicates may be preserved by a combination of subfunctionalization and neofunctionalization, rather than a simple "sub-" or "neo-"

mechanism, and that subfunctionalization can be both spatial and temporal for developmentally-regulated duplicate genes. Sister copies of a duplicated pair complement each other to jointly retain the full set of ancestral subfunctions, and subfunction(s) of one duplicated pair, however, might be complementarily retained by duplicates of a closely related gene. Shuffling of subfunctions among duplicates of related paralogous members might be a mechanism for retention of duplicates from unifunctional genes. Although it is speculative, divergence in distribution of *cis* regulatory elements in the 5' upstream region may be a major factor contributing to sub- and/or neofunctionalization of the zebrafish duplicate iLBP genes.

# REFERENCES


Adida, A. & Spener, F. (2002) Intracellular lipid binding proteins and nuclear receptors involved in branched-chain fatty acid signaling. Prostaglandins Leukot. Essent. Fatty Acids. 67, 91-98.

Akiyama, T.E., Ward, J.M. & Gonzalez, F.J. (2000) Regulation of the liver fatty acid-binding protein gene by hepatocyte nuclear factor 1alpha (HNF1alpha). Alterations in fatty acid homeostasis in HNF1alpha-deficient mice. J. Biol. Chem. 275, 27117-27122.

Amores, A., Force, A., Yan, Y.L., Joly, L., Amemiya, C., Fritz, A., Ho, R.K., Langeland, J., Prince, V., Wang, Y.L., Westerfield, M., Ekker, M. & Postlethwait, J.H. (1998) Zebrafish *hox* clusters and vertebrate genome evolution. Science. 282, 1711-1714.

Amsterdam, A., Lin, S. & Hopkins, N. (1995) The Aequorea victoria green fluorescent protein can be used as a reporter in live zebrafish embryos. Dev. Biol. 171, 123-129.

Amsterdam, A., Lin, S., Moss, L.G. & Hopkins, N. (1996) Requirements for green fluorescent protein detection in transgenic zebrafish embryos. Gene. 173, 99-103.

Anchan, R.M., Drake, D.P., Haines, C.F., Gerwe, E.A. & LaMantia, A.S. (1997) Disruption of local retinoid-mediated gene expression accompanies abnormal development in the mammalian olfactory pathway. J. Comp. Neurol. 379, 171-184.

Ang, H.L., Deltour, L., Zgombic-Knight, M., Wagner, M.A. & Duester, G. (1996) Expression patterns of class I and class IV alcohol dehydrogenase genes in developing epithelia suggest a role for alcohol dehydrogenase in local retinoic acid synthesis. Alcohol Clin. Exp. Res. 20, 1050-1064.

Bailey, G.S., Poulter, R.T. & Stockwell, P.A. (1978) Gene duplication in tetraploid fish: model for gene silencing at unlinked duplicated loci. Proc. Natl. Acad. Sci. U S A. 75, 5575-5579.

Bain, L.J. (2002) cDNA cloning, sequencing, and differential expression of a heart-type fatty acid-binding protein in the mummichog (Fundulus heteroclitus). Mar. Environ. Res. 54, 379-383.

Barbazuk, W.B., Korf, I., Kadavi, C., Heyen, J., Tate, S., Wun, E., Bedell, J.A., McPherson, J.D. & Johnson, S.L. (2000) The syntenic relationship of the zebrafish and human genomes. Genome Res. 10, 1351-1358.

Bashor, M.M., Toft, D.O. & Chytil, F. (1973) In vitro binding of retinol to rat-tissue components. Proc. Natl. Acad. Sci. U S A. 70, 3483-3487.

Bass, N.M. & Manning, J.A. (1986) Tissue expression of three structurally different fatty acid binding proteins from rat heart muscle, liver, and intestine. Biochem. Biophys. Res. Commun. 137, 929-935.

Bass, N.M., Raghupathy, E., Rhoads, D.E., Manning, J.A. & Ockner, R.K. (1984) Partial purification of molecular weight 12 000 fatty acid-binding proteins from rat brain and their effect on synaptosomal Na+-dependent amino acid uptake. Biochemistry. 23, 6539-6544.

Begemann, G. & Meyer, A. (2001) Hindbrain patterning revisited: timing and effects of retinoic acid signalling. Bioessays. 23, 981-986.

Bennett, E., Stenvers, K.L., Lund, P.K. & Popko, B. (1994) Cloning and characterization of a cDNA encoding a novel fatty acid binding protein from rat brain. J. Neurochem. 63, 1616-1624.

Bernhardt, R.R., Nguyen, N. & Kuwada, J.Y. (1992) Growth cone guidance by floor plate cells in the spinal cord of zebrafish embryos. Neuron. 8, 869-882.

Bernlohr, D.A., Simpson, M.A., Hertzel, A.V. & Banaszak, L.J. (1997) Intracellular lipid-binding proteins and their genes. Annu. Rev. Nutr.17, 277-303.

Binas, B., Danneberg, H., McWhir, J., Mullins, L. & Clark, A.J. (1999) Requirement for the heart-type fatty acid binding protein in cardiac fatty acid utilization. FASEB J. 13, 805-812.

Birse, S.C., Leonard, R.B. & Coggeshall, R.E. (1980) Neuronal increase in various areas of the nervous system of the guppy, Lebistes. J. Comp. Neurol. 194, 291-301.

Bisgrove, D.A., Monckton, E.A., Packer, M. & Godbout, R. (2000) Regulation of brain fatty acid-binding protein expression by differential phosphorylation of nuclear factor I in malignant glioma cell lines. J. Biol. Chem. 275, 30668-30676.

Blaner, W.S. & Olson, J.A. (1994) Retinal and retinoic acid metabolism. In: The Retinoids: Biology, Chemistry and Medicine. In The Retinoids: Biology, Chemistry and Medicine (Sporn, M. B., Roberts, A. B. & Goodman, D. S., eds), pp229-255. Reven Press, NY.

Börchers, T. & Spenser, F. (1994) Fatty acid binding proteins. Curr. Top. Membranes 40, 261-226.

Breathnach, R. & Chambon, P. (1981) Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. 50, 349-383.

Breathnach, R. & Chambon, P. (1981) Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. 50, 349-383.

Calderone, V., Folli, C., Marchesani, A., Berni, R. & Zanotti, G. (2002) Identification and structural analysis of a zebrafish apo and holo cellular retinol-binding protein. J. Mol. Biol. 321, 527-535.

Cameron, M.C., Denovan-Wright, E.M., Sharma, M.K. & Wright, J.M. (2002) Cellular retinol-binding protein type II (CRBPII) in adult zebrafish (*Danio rerio*). cDNA sequence, tissue-specific expression and gene linkage analysis. Eur. J. Biochem. 269, 4685-4692.

Cheng, L., Qian, S.J., Rothschild, C., d'Avignon, A., Lefkowith, J.B., Gordon, J.I. & Li, E. (1991) Alteration of the binding specificity of cellular retinol-binding protein II by site-directed mutagenesis. J. Biol. Chem. 266, 24404-24412.

Chevrette M, Joly L, Tellis P, Knapik EW, Miles J, Fishman M, Ekker M (2000) Characterization of a zebrafish/mouse somatic cell hybrid panel. Genomics. 64, 119-126.

Chiang, E.F., Pai, C.I., Wyatt, M., Yan, Y.L., Postlethwait, J. & Chung, B. (2001a) Two sox9 genes on duplicated zebrafish chromosomes: expression of similar transcription activators in distinct sites. Dev. Biol. 231, 149-163.

Chiang, E.F., Yan, Y.L., Guiguen, Y., Postlethwait, J., Chung, B.C. (2001b) Two Cyp19 (P450 aromatase) genes on duplicated zebrafish chromosomes are expressed in ovary or brain. Mol. Biol. Evol. 18, 542-550.

Chrisman, T.S., Claffcy, K.P., Saouat, R., Hanspal, J. & Brecher, P. (1987) Measurement of rat heart fatty acid binding protein by ELISA. Tissue distribution, developmental changes and subcellular distribution. J. Mol. Cell. Cardiol. 19, 423-431.

Claffey, K.P., Herrera, V.L., Brecher, P. & Ruiz-Opazo, N. (1987) Cloning and tissue distribution of rat heart fatty acid binding protein mRNA: identical forms in heart and skeletal muscle. Biochemistry 26, 7900-7904.

Clandinin, M.T. (1999) Brain development and assessing the supply of polyunsaturated fatty acid. Lipids. 34, 131-137.

Clarke, S.D. (2001) Polyunsaturated fatty acid regulation of gene transcription: a molecular mechanism to improve the metabolic syndrome. J. Nutr. 13, 1129-1132.

Clerc, R.G., Corcoran, L.M., LeBowitz, J.H., Baltimore, D. & Sharp, P.A. (1988) The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. Genes Dev. 2, 1570-1581.

Coe, N.R. & Bernlohr, D.A. (1998) Physiological properties and functions of intracellular fatty acid-binding proteins. Biochim. Biophys. Acta. 1391, 287-306.

Colamarino, S.A. & Tessier-Lavigne, M. (1995) The role of the floor plate in axon guidance. Annu. Rev. Neurosci.18, 497-529.

Cooke, J., Nowak, M.A., Boerlijst, M. & Maynard-Smith, J. (1997) Evolutionary origins and maintenance of redundant gene expression during metazoan development. Trends Genet. 13, 360-364.

Costaridis P, Horton C, Zeitlinger J, Holder N. & Maden M. (1996) Endogenous retinoids in the zebrafish embryo and adult. Dev. Dyn. 205:41-51.

Crow, J.A. & Ong, D.E. (1985) Cell-specific immunohistochemical localization of a cellular retinol-binding protein (type two) in the small intestine of rat. Proc. Natl. Acad. Sci. U S A. 82, 4707-4711.

de Martino, S., Yan, Y.L., Jowett, T., Postlethwait, J.H., Varga, Z.M., Ashworth, A. & Austin, C.A. (2000) Expression of sox11 gene duplicates in zebrafish suggests the reciprocal loss of ancestral gene expression patterns in development. Dev. Dyn. 217, 279-292.

de Urquiza, A.M., Liu, S., Sjoberg, M., Zetterstrom, R.H., Griffiths. W,, Sjovall, J. & Perlmann, T. (2000) Docosahexaenoic acid, a ligand for the retinoid X receptor in mouse brain. Science. 290, 2140-2144.

Demmer, L.A., Birkenmeier, E.H., Sweetser, D.A., Levin, M.S., Zollman, S., Sparkes, R.S., Mohandas, T., Lusis, A.J. & Gordon, J.I. (1987) The cellular retinol binding protein II gene. Sequence analysis of the rat gene, chromosomal localization in mice and humans, and documentation of its close linkage to the cellular retinol binding protein gene. J. Biol. Chem. 262, 2458-2467.

Denovan-Wright, E.M., Newton, R.A., Armstrong, J.N., Babity, J.M. & Robertson, H.A. (1998) Acute administration of cocaine, but not amphetamine, increases the level of synaptotagmin IV mRNA in the dorsal striatum of rat. Mol. Brain Res. 55, 350-354.

Denovan-Wright, E.M., Pierce, M. & Wright, J.M. (2000) Nucleotide sequence of cDNA clones coding for a brain-type fatty acid binding protein and its tissue-specific expression in adult zebrafish (*Danio rerio*). Biochim. Biophys. Acta. 1492, 221-226.

Denovan-Wright, E.M., Pierce, M., Sharma, M.K. & Wright, J.M. (2000) cDNA sequence and tissue-specific expression of a basic liver-type fatty acid binding protein in adult zebrafish (*Danio rerio*). Biochim. Biophys. Acta 1492, 227-232.

Di Pietro, S.M., Dell'Angelica, E.C., Veerkamp, J.H., Sterin-Speziale, N. & Santome, J.A. (1997) Amino acid sequence, binding properties and evolutionary relationships of the basic liver fatty-acid-binding protein from the catfish *Rhamdia sapo*. Eur. J. Biochem. 249, 510-517.

Doe, C.Q., Hiromi, Y., Gehring, W.J. & Goodman, C.S. (1988) Expression and function of the segmentation gene *fushi tarazu* during *Drosophila* neurogenesis. Science. 239, 170-175.

Dolle P, Ruberte E, Leroy P, Morriss-Kay G, Chambon P. (1990) Retinoic acid receptors and cellular retinoid binding proteins. I. A systematic study of their differential pattern of transcription during mouse organogenesis. Development. 110:1133-1151.

Donohoe, M.E., Zhang, X., McGinnis, L., Biggers, J., Li, E. & Shi, Y. (1999) Targeted disruption of mouse Yin Yang 1 transcription factor results in peri-implantation lethality. Mol. Cell Biol. 19, 7237-7244.

Driever, W., Solnica-Krezel, L., Schier, A.F., Neuhauss, S.C., Malicki, J., Stemple, D.L., Stainier, D.Y., Zwartkruis, F., Abdelilah, S., Rangini, Z., Belak, J. & Boggs, C. (1996) A genetic screen for mutations affecting embryogenesis in zebrafish. Development. 123, 37-46.

Dugast, C. & Weber, M.J. (2001) NF-Y binding is required for transactivation of neuronal aromatic L-amino acid decarboxylase gene promoter by the POU-domain protein Brn-2. Brain Res. Mol. Brain Res. 89, 58-70.

Dupe, V. & Lumsden, A. (2001) Hindbrain patterning involves graded responses to retinoic acid signalling. Development. 128, 2199-2208.

Duplus, E., Glorian, M. & Forest, C. (2000) Fatty acid regulation of gene transcription. J. Biol. Chem. 275, 30749-30752.

Farber, S.A., De Rose, R.A., Olson, E.S. & Halpern, M.E. (2003) The zebrafish annexin gene family. Genome Res. 13, 1082-1096.

Feng, L. & Heintz, N. (1995) Differentiating neurons activate transcription of the brain lipid-binding protein gene in radial glia through a novel regulatory element. Development. 121, 1719-1730.

Feng, L., Hatten, M.E. & Heintz, N. (1994) Brain lipid-binding protein (BLBP): a novel signaling system in the developing mammalian CNS. Neuron. 12, 895-908.

Fex, G. & Johannesson, G. (1984) Radioimmunological determination of cellular retinol-binding protein in human tissue extracts. Cancer Res. 44:3029-3032.

Ficzycz, A. & Ovsenek N. (2202) The Yin Yang 1 transcription factor associates with ribonucleoprotein (mRNP) complexes in the cytoplasm of *Xenopus* oocytes. J. Biol. Chem. 277, 8382-8387.

Fliesler, S.J. & Anderson, R.E. (1983) Chemistry and metabolism of lipids in the vertebrate retina. Prog. Lipid. Res. 22, 79-131.

Folli, C., Calderone, V., Ottonello, S., Bolchi, A., Zanotti, G., Stoppini, M. & Berni, R. (2001) Identification, retinoid binding, and x-ray analysis of a human retinol-binding protein. Proc. Natl. Acad. Sci. U S A. 98, 3710-3715.

Folli, C., Calderone, V., Ramazzina, I., Zanotti, G. & Berni, R. (2002) Ligand binding and structural analysis of a human putative cellular retinol-binding protein. J. Biol. Chem. 277, 41970-41977.

Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L. & Postlethwait, J. (1999) Preservation of duplicate genes by complementary, degenerative mutations. Genetics. 151:1531-1545.

Gajewski, M. & Voolstra, C. (2002) Comparative analysis of somitogenesis related genes of the *hairy*/Enhancer of split class in *Fugu* and zebrafish. BMC Genomics. 3, 21.

Gates, M.A., Kim, L., Egan, E.S., Cardozo, T., Sirotkin, H.I., Dougan, S.T., Lashkari, D., Abagyan, R., Schier, A.F. & Talbot, W.S., (1999). A genetic linkage map for zebrafish: comparative analysis and localization of genes and expressed sequences. Genome Res. 9, 334-347.

Geisler, R., Rauch, G.J., Baier, H., van Bebber, F., Brobeta, L., Dekens, M.P., Finger, K., Fricke, C., Gates, M.A., Geiger, H., Geiger-Rudolph, S., Gilmour, D., Glaser, S., Gnugge, L., Habeck, H., Hingst, K., Holley, S., Keenan, J., Kirn, A., Knaut, H., Lashkari, D., Maderspacher, F., Martyn, U., Neuhauss, S. & Haffter P. (1999) A radiation hybrid map of the zebrafish genome. Nat. Genet. 23:86-89.

Gerbens, F., Rettenberger, G., Lenstra, J.A., Veerkamp, J.H. & te Pas, M.F. (1997) Characterization, chromosomal localization, and genetic variation of the porcine heart fatty acid-binding protein gene. Mamm. Genome. 8, 328-332.

Giguere, V. (1994) Retinoic acid receptors and cellular retinoid binding proteins: complex interplay in retinoid signaling. Endocr. Rev. 15, 61-79.

Glatz, J.F. & van der Vusse, G.J. (1996) Cellular fatty acid-binding proteins: their function and physiological significance. Prog. Lipid Res. 35, 243-282.

Glatz, J.F. & Vusse, G.J. (1996) Cellular fatty acid-binding proteins: Their function and physiological significance. Prog. Lipid Res. 35, 243-282.

Godbout, R., 1993. Identification and characterization of transcripts present at elevated levels in the undifferentiated chick retina. Exp. Eye Res. 56, 95-106.

Gould, A., Itasaki, N. & Krumlauf, R. (1998) Initiation of rhombomeric Hoxb4 expression requires induction by somites and a retinoid pathway. Neuron. 21, 39-51.

Graham, A., Heyman, I., Lumsden, A., 1993. Even-numbered rhombomeres control the apoptotic elimination of neural crest cells from odd-numbered rhombomeres in the chick hindbrain. Development. 119, 233-245.

Graham, A., Lumsden A. 1996. Interactions between rhombomeres modulate Krox-20 and follistatin expression in the chick embryo hindbrain. Development. 122, 473-480.

Grandel, H., Lun, K., Rauch, G.J., Rhinn, M., Piotrowski, T., Houart, C., Sordino, P., Kuchler, A.M., Schulte-Merker, S., Geisler, R., Holder, N., Wilson, S.W. & Brand, M. (2002) Retinoic acid signalling in the zebrafish embryo is necessary during pre-segmentation stages to pattern the anterior-posterior axis of the CNS and to induce a pectoral fin bud. Development. 129, 2851-2865.

Guidato, S., Barrett, C. & Guthrie, S. (2003) Patterning of motor neurons by retinoic acid in the chick embryo hindbrain in vitro. Mol. Cell Neurosci. 23, 81-95.

Gustafson, A.L., Donovan, M., Annerwall, E., Dencker, L. & Eriksson, U. (1996) Nuclear import of cellular retinoic acid-binding protein type I in mouse embryonic cells. Mech. Dev. 58:27-38.

Halter, D.A., Urban, J., Rickert, C., Ner, S.S., Ito, K., Travers, A.A. & Technau, G.M. (1995) The homeobox gene repo is required for the differentiation and maintenance of glia function in the embryonic nervous system of *Drosophila melanogaster*. Development. 121, 317-332.

Hamilton, L.C. & Wright, J.M. (1999) Isolation of complementary DNAs coding for a receptor for activated C kinase (RACK1) from zebrafish (*Danio rerio*) and tilapia (*Oreochromis niloticus*): constitutive developmental and tissue expression. Marine Biotechnol. 1, 279-285.

He, X., Treacy, M.N., Simmons, D.M., Ingraham, H.A., Swanson, L.W. & Rosenfeld, M.G. (1989) Expression of a large family of POU-domain regulatory genes in mammalian brain development. Nature. 340, 35-41.

Herdegen, T. & Waetzig, V. (2001) AP-1 proteins in the adult brain: facts and fiction about effectors of neuroprotection and neurodegeneration. Oncogene. 20, 2424-2437.

Herr, F.M. & Ong, D.E. (1992) Differential interaction of lecithin-retinol acyltransferase with cellular retinol binding proteins. Biochemistry. 31, 6748-6755.

Hertzel, A.V. & Bernlohr, D.A. (2000) The mammalian fatty acid-binding protein multigene family: molecular and genetic insights into function. Trends Endocrinol. Metab. 11, 175-180.

Heuckeroth, R.O., Birkenmeier, E.M., Levin, M.S. & Gordon, J.I. (9187) Analysis of the tissue –specific expression, developmental regulation, and linkage relationships of a rodent gene encoding heart fatty acid binding protein. J. Biol. Chem. 262, 9709-9717.

Heyman, R.A., Manglesdorf, D.J., Kyck, J.A., Stein, R.B., Eichele, G., Evans, R.E & Thaller, C. (1992) 9-*cis* retinoic acid is a high affinity ligand for the retinoid X receptor. Cell. 66, 397-406.

Hittel, D. & Storey, K.B. (2001) Differential expression of adipose- and heart-type fatty acid binding proteins in hibernating ground squirrels. Biochim. Biophys. Acta 1522, 238-243.

Holland, P.W. & Garcia-Fernandez, J. (1999) Hox genes and chordate evolution. Dev. Biol. 173, 382-395.

Holland, P.W., Garcia-Fernandez, J., Williams, N.A. & Sidow, A. (1994) Gene duplications and the origins of vertebrate development. Development (Suppl.) 125-133.

Houart, C., Westerfield, M. & Wilson, S.W. (1998) A small population of anterior cells patterns the forebrain during zebrafish gastrulation. Nature. 391, 788-792.

Hukriede, N., L. Joly, M. Tsang, J. Miles, P. Tellis, J. Epstein, W. Barbazuk, F. Li, B. Paw, J. Postlethwait, T. Hudson, L. Zon, J. McPherson, M. Chevrette, I. Dawid, S. Johnson, and M. Ekker (1999) Radiation hybrid mapping of the zebrafish genome. Proc. Natl. Acad. Sci.,USA 96, 9745-9750.

Hyatt, T.M., & Ekker, S.C. (1999) Vectors and techniques for ectopic gene expression in zebrafish. Methods Cell Biol. 59, 117-126.

Ingraham, H.A., Chen, R.P., Mangalam, H.J., Elsholtz, H.P., Flynn, S.E., Lin, C.R., Simmons, D.M., Swanson, L. & Rosenfeld, M.G. (1988) A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. Cell. 55, 519-529.

Itasaki, N., Sharpe, J., Morrison, A. & Krumlauf, R. (1996) Reprogramming Hox expression in the vertebrate hindbrain: influence of paraxial mesoderm and rhombomere transposition. Neuron. 16, 487-500.

Jeffrey, B.G., Weisinger, H.S., Neuringer, M. & Mitchell, D.C. (2001) The role of docosahexaenoic acid in retinal function. Lipids. 36, 859-871.

Josephson. R., Müller, T., Pickel, J., Okabe, S., Reynolds, K., Turner, P.A., Zimmer, A. & McKay, R.D. (1998) POU transcription factors control expression of CNS stem cell-specific genes. Development. 125, 3087-3100.

Jowett, T. (1999) Analysis of protein and gene expression. Methods Cell Biol. 59, 63-85.

Kelly, P.D., F. Chu, I.G. Woods, P. Ngo-Hazelett, T. Cardozo, H. Huang, F. Kimm, L. Liao, Y.-L. Yan, Y. Zhou, S.L. Johnson, R. Abagyan, A.F. Schier, J.H. Postlethwait, and W.S. Talbot (2000) Genetic linkage mapping of zebrafish genes and ESTs. Genome Res. 10:558-567.

Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B. & Schilling, T.F. (1995) Stages of embryonic development of the zebrafish. Dev. Dyn. 203, 253-310.

Kleinjan, D.A., Dekker, S., Guy, J.A. & Grosveld, F.G. (1998) Cloning and sequencing of the CRABP-I locus from chicken and pufferfish: analysis of the promoter regions in transgenic mice. Transgenic Res. 7, 85-94.

Knapik, E. W., Goodman, A., Atkinson, O. S., Roberts, C. T., Shiozawa, M., Sim, C. U., Weksler-Zangen, S., Trolliet, M. R., Futrell, C., Innes, B. A., Koike, G., McLaughlin, M. G., Pierre, L., Simon, J. S., Vilallonga, E., Roy, M., Chiang, P. W., Fishman, M. C., Driever, W., and Jacob, H. J. (1996). A reference cross DNA panel for zebrafish (*Danio rerio*) anchored with simple sequence length polymorphisms. Development 123:451-460.

Krumlauf, R. (1994) Hox genes in vertebrate development. Cell. 78, 191-201.

Kuhar, S.G., Feng, L., Vidan, S., Ross, M.E., Hatten, M.E. & Heintz, N. (1993) Changing patterns of gene expression define four stages of cerebellar granule neuron differentiation. Development. 117, 97-104.

Kurtz, A., Zimmer, A., Schnütgen, F., Brüning, G., Spener, F. & Müller, T. (1994) The expression pattern of a novel gene encoding brain-fatty acid binding protein correlates with neuronal and glial cell development. Development. 120, 2637-2649.

Kuwada, J.Y., Bernhardt, R.R. & Chitnis, A.B. (1990) Pathfinding by identified growth cones in the spinal cord of zebrafish embryos. J. Neurosci.10, 1299-1308.

Kwok, C., Korn, R.M., Davis, M.E., Burt, D.W., Critcher, R., M cCarthy, L., Paw, B.H., Zon, L.I., Goodfellow, P.N. & and Schmitt, K. (1998) Characterization of whole genome radiation hybrid mapping resources for non-mammalian vertebrates. Nucleic Acids Res. 26:3562 3566.

Lee, A.G., East, J.M. & Froud, R.J. (1986) Are essential fatty acids essential for membrane function? Prog. Lipid Res. 25, 41-46.

Levin, A.A., Sturzenbecker, L.J., Kazmer, S., Bosakowski, T., Huselton, C., Allenby, G., Speck, J., Kratzeisen, C., Rosenberger, M. & Lovey, A. (1992) 9-*cis* retinoic acid stereoisomer binds and activates the nuclear receptor RXR alpha. Nature. 355, 359-361.

Levin, M.S., Li, E., Ong, D.E. & Gordon, J.I. (1987) Comparison of the tissue-specific expression and developmental regulation of two closely linked rodent genes encoding cytosolic retinol-binding proteins. J. Biol. Chem. 262, 7118-7124.

Li, E., Demmer, L.A., Sweetser, D.A., Ong, D.E & Gordon, J.I. (1986) Rat cellular retinol-binding protein II: use of a cloned cDNA to define its primary structure, tissue-specific expression, and developmental regulation. Proc. Natl. Acad. Sci. U S A. 83, 5779-5783.

Li, W.H. (1980) Rate of gene silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes. Genetics. 95, 237-258.

Liu, R.Z., Denovan-Wright, E.M. & Wright, J.M. (2003b) Structure, linkage mapping and expression of the heart-type fatty acid-binding protein gene (*fabp3* ) from zebrafish (*Danio rerio*). Eur. J. Biochem. 270, 3223-3234.

Liu, R.-Z., Denovan-Wright, E.M. & Wright, J.M., (2003a) Structure, mRNA expression and linkage mapping of the brain-type fatty acid-binding protein gene (*fabp7*) from zebrafish (*Danio rerio*). Eur. J. Biochem. 270, 715-725.

Liu, R-Z., Denovan-Wright, E.M., Dgrave, A., Thisse, C., Thisse, T. & Wright, J.M. (2003c) Spatio-temporal distribution of cellular retinol-binding protein (CRBPI and CRBPII) gene transcripts in the developing and adult zebrafish (*Danio rerio*). Eur. J. Biochem. In press.

Liu, Y.-W. & Chan, W.-K. (2002) Thyroid hormones are important for embryonic to larval transitory phase in zebrafish. Differentiation. 70, 36-45.

Livera, G., Rouiller-Fabre., V., Valla, J. & Habert, R. (2000) Effects of retinoids on meiosis in the fetal rat ovary in culture. Mol. Cell Endocrinol. 165, 225-231.

Lohnes, D., Dierich, A., Ghyselinck, N., Kastner, P., Lampron, C., LeMeur, M., Lufkin, T., Mendelsohn, C., Nakshatri, H. & Chambon, P. (1992) Retinoid receptors and binding proteins. J. Cell Sci. Suppl. 16, 69-76.

Londraville, R.L. (1996) Intracellular fatty acid-binding proteins: putting lower vertebrates in perspective. Braz. J. Med. Biol. Res. 29, 707-720.

Lumsden, A. & Krumlauf, R. (1996) Patterning the vertebrate neuraxis. Science. 274, 1109-1115.

Lynch, M. & Conery, J.S. (2000) The evolutionary fate and consequences of duplicate genes. Science. 290, 1151-1155.

Lynch, M., O'Hely, M., Walsh, B. & Force, A. (2001) The probability of preservation of a newly arisen gene duplicate. Genetics. 159, 1789-1804.

Maden, M. (2001) Role and distribution of retinoic acid during CNS development. Int. Rev. Cytol. 209, 1-77.

Maden, M. (2002 ) Retinoid signalling in the development of the central nervous system. Nat. Rev. Neurosci. 3, 843-853.

Maden, M., Sonneveld, E., van der Saag, P.T. & Gale, E. (1998) The distribution of endogenous retinoic acid in the chick embryo: implications for developmental mechanisms. Development. 125, 4133-4144.

Maity, S.N. & de Crombrugghe, B. (1998) Role of the CCAAT-binding protein CBF/NF-Y in transcription. Trends Biochem. Sci. 23, 174-178.

Maller, J.L. (2001) The elusive progesterone receptor in *Xenopus* oocytes. Proc. Natl. Acad. Sci. U S A. 98, 8-10.

Mangelsdorf, D.J. (1994) Vitamin A receptors. Nutr. Rev. 52, S32-44.

Mantamadiotis, T., Lemberge, T., Bleckmann, S.C., Kern, H., Kretz, O., Martin, Villalba, A., Tronche, F., Kellendonk, C., Gau, D., Kapfhammer, J., Otto, C., Schmid, W. & Schutz, G. (2002) Disruption of CREB function in brain leads to neurodegeneration. Nat. Genet. 31, 47-54.

Matarese, V., Stone, R.L., Waggoner, D.W. & Bernlohr, D.A. (1989) Intracellular fatty acid trafficking and the role of cytosolic lipid binding proteins. Prog. Lipid Res. 28, 245-272.

Matsuzaki, T., Amanuma, H. & Takeda, H. (1992) A POU-domain gene of zebrafish, ZFPOU1, specifically expressed in the developing neural tissues. Biochem. Biophys. Res. Commun. 187, 1446-1453.

Matuoka, K. & Chen, K.Y. (1999) Nuclear factor Y (NF-Y) and cellular senescence. Exp. Cell Res. 253, 365-371.

Maves. L, Jackman, W. & Kimmel ,C.B. (2002) FGF3 and FGF8 mediate a rhombomere 4 signaling activity in the zebrafish hindbrain. Development. 129, 3825-3837.

McCaffery P, Drager UC. (1994) Hot spots of retinoic acid synthesis in the developing spinal cord. Proc. Natl. Acad. Sci. U S A. 91:7194-7197.

McCaffery, P., Wagner, E., O'Neil, J., Petkovich, M. & Drager, U.C. (1999) Dorsal and ventral retinoic territories defined by retinoic acid synthesis, break-down and nuclear receptor expression. Mech. Dev. 85, 203-214.

Meng, A., Tang, H., Ong, B.A., Farrell, M.J. & Lin, S. (1997) Promoter analysis in living zebrafish embryos identifies a *cis*-acting motif required for neuronal expression of GATA-2. Proc. Natl. Acad. Sci. U S A. 94, 6267-6272.

Mic, F.A., Haselbeck, R.J., Cuenca, A.E. & Duester, G. (2002) Novel retinoic acid generating activities in the neural tube and heart identified by conditional rescue of *Raldh2* null mutant mice. Development. 129, 2271-2282.

Miller, W.C., Hickson, R.C. & Bass, N.M. (1988) Fatty acid binding proteins in the three types of rat skeletal muscle. Proc. Soc. Exp. Biol Med. 189, 183-188.

Morita, Y. & Tilly, J.L. (1999) Segregation of retinoic acid effects on fetal ovarian germ cell mitosis versus apoptosis by requirement for new micromolecular synthesis. Endocrinology 140, 2696-2703.

Moss, J.B., Price, A.L., Raz, E., Driever, W. & Rosenthal, N. (1996) Green fluorescent protein marks skeletal muscle in murine cell lines and zebrafish. Gene. 173, 89-98.

Myster, D.L., Bonnette, P.C. & Duronio RJ. (2000) A role for the DP subunit of the E2F transcription factor in axis determination during *Drosophila* oogenesis. Development. 127, 3249-3261.

Nakagawa, S., Kim, J.E., Lee, R., Malberg, J.E., Chen, J., Steffen, C., Zhang, Y.J., Nestler, E.J. & Duman, R.S. (2002) Regulation of neurogenesis in adult mouse hippocampus by cAMP and the cAMP response element-binding protein. J. Neurosci. 22, 3673-3682.

Napoli, J.L. (1999) Retinoic acid: its biosynthesis and metabolism. Prog. Nucleic. Acid. Res. Mol. Biol. 63:139-188.

Napoli, J.L. (1996) Biochemical pathways of retinoid transport, metabolism, and signal transduction. Clin. Immunol. Immunopathol. 80, S52-62.

Napoli, J.L. (1999) Interactions of retinoid binding proteins and enzymes in retinoid metabolism. Biochim. Biophys. Acta. 1440, 139-162.

Nguyen, V., Deschet, K., Henrich, T., Godet, E., Joly, J.S., Wittbrodt, J., Chourrout, D. & Bourrat, F. (1999) Morphogenesis of the optic tectum in the medaka (*Oryzias*

*latipes*): a morphological and molecular study, with special emphasis on cell proliferation. J. Comp. Neurol. 413, 385-404.

Niederreither, K., Vermot, J., Schuhbaur, B., Chambon, P. & Dolle, P. (2002) Embryonic retinoic acid synthesis is required for forelimb growth and anteroposterior patterning in the mouse. Development. 129, 3563-3574.

Noy, N. (2000) Retinoid-binding proteins: mediators of retinoid action. Biochem. J. 348, 481-495.

Ntambi, J.M. & Bene, H. (2001) Polyunsaturated fatty acid regulation of gene expression. J. Mol. Neurosci. 16, 273-278.

Oates, A.C., Brownlie, A., Pratt, S.J., Irvine, D.V., Liao, E.C., Paw, B.H., Dorian, K.J, Johnson, S.L., Postlethwait, J.H., Zon, L.I. & Wilks, A.F. (1999) Gene duplication of zebrafish JAK2 homologs is accompanied by divergent embryonic expression patterns: only jak2a is expressed during erythropoiesis. Blood. 94, 2622-2636.

Ockner, R.K., Manning, J.A., Poppenhausen, R.B. & Ho, W.K. (1972) A binding protein for fatty acids in cytosol of intestinal mucosa, liver, myocardium, and other tissues. Science. 177, 56-58.

Odenthal, J., van Eeden, F.J., Haffter ,P., Ingham, P.W. & Nusslein-Volhard, C. (2000) Two distinct cell populations in the floor plate of the zebrafish are induced by different pathways. Dev. Biol. 219, 350-363.

Ohno, S. (1970) Evolution by gene duplication. Spinger-Verlag, New York. Curr. Opin. Genet. Dev. 6, 715-722.

Ong DE. (1984) A novel retinol-binding protein from rat. Purification and partial characterization. J. Biol. Chem. 259, 1476-1482.

Ong DE. (1994) Cellular transport and metabolism of vitamin A: roles of the cellular retinoid-binding proteins. Nutr. Rev. 52, S24-31.

Ong, D.E. & Page, D.L. (1986) Quantitation of cellular retinol-binding protein in human organs. Am. J. Clin. Nutr. 44, 425-430.

Ong, D.E., Newcomer, M.E. & Chytil, F. (1994) Cellular retinoid-binding proteins. In The Retinoids: Biology, Chemistry and Medicine (Sporn, M. B., Roberts, A. B. & Goodman, D. S., eds), pp283-317. Reven Press, NY.

Pack, M., Solnica-Krezel, L., Malicki, J., Neuhauss, S.C., Schier, A.F., Stemple, D.L., Driever, W. & Fishman, M.C. (1996) Mutations affecting development of zebrafish digestive organs. Development. 123, 321-328.

Partington, G.A., Bertwistle, D., Nicolas, R.H., Kee, W.J., Pizzey, J.A. & Patient, R.K. (1997) GATA-2 is a maternal transcription factor present in *Xenopus* oocytes as a nuclear complex which is maintained throughout early development. Dev. Biol. 181, 144-155.

Patton, E.E. & Zon, L.I. (2001) The art and design of genetic screens: zebrafish. Nat. Rev. Genet. 2, 956-966.

Paulussen, R.J.A., van Moerkerk, H.T.B. & Veerkamp, J.H. (1990) Immunochemical quantitation of fatty acid-binding proteins. Tissue distribution of liver and heart FABP types in human and porcine tissues. Int. J. Biochem. 22, 393-398.

Perez-Castro AV, Toth-Rogler LE, Wei LN & Nguyen-Huu MC. (1989) Spatial and temporal pattern of expression of the cellular retinoic acid-binding protein and the cellular retinol-binding protein during mouse embryogenesis. Proc. Natl. Acad. Sci. U S A. 86:8813-8817.

Pierce, M., Wang, Y., Denovan-Wright, E.M. & Wright, J.M. (2000) Nucleotide sequence of a cDNA clone coding for an intestinal-type fatty acid binding protein and its tissue-specific expression in zebrafish (*Danio rerio*). Biochim. Biophys. Acta 1490, 175-183.

Postlethwait, J.H., Woods, I.G., Ngo-Hazelett, P., Yan, Y.-L., Kelly, P.D., Chu, F., Huang, H., Hill Force, A. & Talbot, W.S. (2000) Zebrafish comparative genomics and the origins of vertebrate chromosomes. Genome Res. 10:1890-1902.

Prince, V.E., Joly, L., Ekker, M., Ho & R.K. (1998) Zebrafish hox genes: genomic organization and modified colinear expression patterns in the trunk. Development. 125, 407-420.

Prinsen, C.F. & Veerkamp, J.H. (1996) Fatty acid binding and conformational stability of mutants of human muscle fatty acid-binding protein. Biochem. J. 314, 253-260.

Qian, Q., Kuo, L., Yu, Y.T. & Rottman, J.N. (1999) A concise promoter region of the heart fatty acid-binding protein gene dictates tissue-appropriate expression. Circ. Res. 84, 276-289.

Quandt, K., Frech, K., Karas, H., Wingender, E. & Werner, T. (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic. Acids. Res. 23, 4878-4884.

Rhinn, M. & Brand, M. (2001) The midbrain—hindbrain boundary organizer. Curr. Opin. Neurobiol. 11, 34-42.

Ribarik Coe, N., Simpson, M.A. & Bernlohr, D.A. (1999) Targeted disruption of the adipocyte lipid-binding protein (aP2 protein) gene impairs fat cell lipolysis and increases cellular fatty acid levels. J. Lipid Res. 40, 967-972.

Robinson-Rechavi, M., Marchand, O., Escriva, H., Bardet, P.L., Zelus, D., Hughes, S. & Laudet V. (2001) Euteleost fish genomes are characterized by expansion of gene families. Genome Res. 11, 781-788.

Rojas, C.V., Martinez, J.I., Flores, I., Hoffman, D.R. & Uauy, R. (2003) Gene expression analysis in human fetal retinal explants treated with docosahexaenoic acid. Invest. Ophthalmol. Vis. Sci. 44, 3170-3177.

Romand, R., Sapin, V., Ghyselinck, N.B., Avan, P., Le Calvez, S., Dolle, P., Chambon, P. & Mark, M. (2000) Spatio-temporal distribution of cellular retinoid binding protein gene transcripts in the developing and the adult cochlea. Morphological and functional consequences in CRABP- and CRBPI-null mutant mice. Eur. J. Neurosci. 12, 2793-2804.

Ross, A.C. (2000) Addressing research questions with national survey data-the relation of vitamin A status to infection and inflammation. Am. J.Clin. Nutr. 72, 1069-1070.

Rousselot, P., Heintz, N. & Nottebohm, F. (1997) Expression of brain lipid binding protein in the brain of the adult canary and its implications for adult neurogenesis. J. Comp. Neurol. 385, 415-426.

Ruberte, E., Dolle, P., Chambon, P. & Morriss-Kay G. (1991) Retinoic acid receptors and cellular retinoid binding proteins. II. Their differential pattern of transcription during early morphogenesis in mouse embryos. Development. 111:45-60.

Salem, N. Jr., Litman, B., Kim, H.Y. & Gawrisch, K. (2001) Mechanisms of action of docosahexaenoic acid in the nervous system. Lipids 36, 945-959.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. 2nd Ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA.

Sampath, K. & Stuart, G.W. (1996) Developmental expression of class III and IV POU domain genes in the zebrafish. Biochem. Biophys. Res. Commun. 219, 565-571.

Schaap, F.G., Binas, B., Danneberg, H., van der Vusse, G.J. & Glatz, J.F.C. (1999) Impaired long-chain fatty acid utilization by cardiac myocytes isolated from mice lacking the heart-type fatty acid binding protein gene. Circ. Res. 85, 329-337.

Schaap, F.G., van der Vusse, G.J. & Glatz, J.F. (2002) Evolution of the family of intracellular lipid binding proteins in vertebrates. Mol. Cell Biochem. 239, 69-77.

Schleicher, C.H., Cordoba, O.L., Santomé, J.A. & Dell'Angelica, E.C. (1995) Molecular evolution of the multigene family of intracellular lipid-binding proteins. Biochem. Mol. Biol. Int. 36, 1117-1125.

Sellner, P.A., Chu, W., Glatz, J.F. & Berman, N.E. (1995) Developmental role of fatty acid-binding proteins in mouse brain. Brain Res. Dev. Brain Res. 89, 33-46.

Selman, K., Wallace, R.A., Sarka, A. & Qi, X. (1993) Stages of oocyte development in the zebrafish, *Brachydanio rerio*. J. Morphology 218, 203-224.

Senjo, M., Ishibashi, T., Imai, Y., Takahashi, K. & Ono, T. (1985) Isolation and characterization of fatty acid-binding protein from rat brain. Arch. Biochem. Biophys. 236, 662-668.

Serluca, F.C., Sidow, A., Mably, J.D. & Fishman, M.C. (2001) Partitioning of tissue expression accompanies multiple duplications of the Na+/K+ ATPase alpha subunit gene. Genome Res. 11, 1625-1631.

Sharma, M.K., Denovan-Wright, E.M., Boudreau, M.E. & Wright, J.M. (2003) A cellular retinoic acid-binding protein from zebrafish (*Danio rerio*): cDNA sequence, phylogenetic analysis, mRNA expression, and gene linkage mapping. Gene. 311, 119-128.

Shaulian, E. & Karin, M. (2002) AP-1 as a regulator of cell life and death. Nat. Cell Biol. 4, E131-E136.

Shimamura, K., Rubenstein, J.L., 1997. Inductive interactions direct early regionalization of the mouse forebrain. Development. 124, 2709-2718.

Shimizu, F., Watanabe, T.K., Shinomiya, H., Nakamura, Y. & Fujiwara, T. (1997) Isolation and expression of a cDNA for human brain fatty acid-binding protein (B-FABP). Biochim. Biophys. Acta. 1354, 24-28.

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

Sidow A. (1996) Gen(om)e duplications in the evolution of early vertebrates. Curr. Opin. Genet. Dev. 6, 715-722.

Spruyt, N., Delarbre, C., Gachelin, G. & Laudet, V. (1998) Complete sequence of the amphioxus (*Branchiostoma lanceolatum*) mitochondrial genome: relations to vertebrates. Nucleic Acids Res. 26, 3279-3285.

Stewart, J.M. (2000) The cytoplasmic fatty-acid-binding proteins: thirty years and counting. Cell Mol. Life Sci. 57, 1345-1359.

Stoeckli, E.T. & Landmesser, L.T. (1998) Axon guidance at choice points. Curr. Opin. Neurobiol. 8, 73-79.

Stratford, T., Horton, C. & Maden, M. (1996) Retinoic acid is required for the initiation of outgrowth in the chick limb bud. Curr. Biol. 6, 1124-1133.

Stratford, T., Logan, C., Zile, M. & Maden, M. (1999) Abnormal anteroposterior and dorsoventral patterning of the limb bud in the absence of retinoids. Mech. Dev. 81, 115-125.

Stubbs, C.D. & Smith, A.D. (1990) Essential fatty acids in membrane: physical properties and function. Biochem. Soc. Trans. 18, 779-781.

Sturm, R.A., Das, G. & Herr, W. (1988) The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. Genes Dev. 2, 1582-1599.

Sweetser, D.A., Birkenmeier, E.H., Klisak, I.J., Zollman, S., Sparkes, R.S., Mohandas, T., Lusis, A.J. & Gordon, J.I. (1987) The human and rodent intestinal fatty acid binding protein genes. A comparative analysis of their structure, expression, and linkage relationships. J. Biol. Chem. 262, 16060-16071.

Takase, S., Suruga, K. & Goda, T. (2000) Regulation of vitamin A metabolism-related gene expression. Br. J. Nutr. 84, 217-221.

Taylor, J.S., Braasch, I., Frickey, T., Meyer, A. & Van de Peer, Y. (2003) Genome duplication, a trait shared by 22000 species of ray-finned fish. Genome Res. 382-390.

Taylor, J.S., Van de Peer, Y., Braasch, I. & Meyer, A. (2001) Comparative genomics provides evidence for an ancient genome duplication event in fish. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 356, 1661-1679.

Tchoudakova, A., Kishida, M., Wood, E. & Callard, G.V. (2001) Promoter characteristics of two cyp19 genes differentially expressed in the brain and ovary of teleost fish. J. Steroid Biochem. Mol. Biol. 78, 427-439.

Thisse, C., Thisse, B., Schilling, T. F., and Postlethwait, J. H. (1993). Structure of the zebrafish snail1 gene and its expression in wild-type, spadetail and no tail mutant embryos. Development. 119, 1203-1215.

Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 11, 4673-4680.

Thompson, J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. & Higgins, D.G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876-4882.

Thompson, J.H., Howell J.M & Pit, G.M. (1964) Vitamin A and reproduction in rats. Proc. Natl. Sci. Lond B 159, 510-535.

Treuner. M., Kozak, C.A., Gallahan, D., Grosse, R. & Muller, T.(1994) Cloning and characterization of the mouse gene encoding mammary-derived growth inhibitor/heart-fatty acid-binding protein. Gene. 147, 237-242.

Uauy, R., Hoffman, D.R., Peirano P., Birch, D.G. & Birch, E.E. (2001) Essential fatty acids in visual and brain development. Lipids. 36, 885-895.

van Eeden, F.J., Granato, M., Odenthal, J. & Haffter, P. (1999) Developmental mutant screens in the zebrafish. Methods Cell Biol. 60, 21-41.

Vasile, F., Ragona, L., Catalano, M., Zetta, L., Perduca, M., Monaco, H. & Molinari, H. (2003) Solution structure of chicken liver basic fatty acid binding protein. J. Biomol. NMR. 25, 157-160.

Vayda M.E., Londravelli. R.L., Cashon, R.E., Costello, L. & Sidell, B.D. (1998) Two distinct types of fatty acid-binding protein are expressed in heart ventricle of Antarctic teleost fishes. Biochem. J. 330, 375-382.

Veerkamp, J.H. & Maatman, G.H.J. (1995) Cytoplasmic fatty acid-binding proteins: their structure and genes. Prog. Lipid Res. 34, 17-52.

Vogel, S., Mendelsohn, C.L., Mertz, J.R., Piantedosi, R., Waldburger. C., Gottesman, M.E. & Blaner, W.S. (2001) Characterization of a new member of the fatty acid-binding protein family that binds all-trans-retinol. J. Biol. Chem. 276, 1353-6130.

Wardlaw, S,A,, Bucco, R.A., Zheng, W.L. & Ong, D.E. (1997) Variable expression of cellular retinol- and cellular retinoic acid-binding proteins in the rat uterus and ovary during the estrous cycle. Biol. Reprod. 56:125-132.

Werner, E.A. & Deluca, H.F. (2002) Retinoic acid is detected at relatively high levels in the CNS of adult rats. Am. J. Physiol. Endocrinol. Metab. 282:E672-678.

Westerfield, M. (1995) The zebrafish book: a guide for the laboratory use of zebrafish (*Danio rerio*), 3rd edn. University of Oregon Press, Eugene, USA.

Whaley, S.L., Hedgpeth, V.S., Farin, C.E., Martus, N.S., Jayes, F.C.L. & Britt J.H. (2000) Influence of vitamin A injection on oocytes development, follicular hormones and ovulation in gilts fed high-energy diets. J. Anim. Sci. 78, 1598-1607.

Whitesides, J., Hall, M., Anchan, R. & LaMantia, A.S. (1998) Retinoid signaling distinguishes a subpopulation of olfactory receptor neurons in the developing and adult mouse. J. Comp. Neurol. 394:445-461.

Wolfe, S.L. (1993) Molecular and Cellular Biology. pp. 1097-1098. Wadsworth Publishing Co., Belmont, CA, USA.

Woods, I.G., Kelly, P.D., Chu, F., Ngo-Hazelett, P., Yan, Y.-L., Huang, H., Postlethwait, J.H. & Talbot, W.S. (2000) A comparative map of the zebrafish genome. Genome Res. 10:1903-1914.

Wu, Q., Andolfatto, P. & Haunerland, N.H. (2001) Cloning and sequence of the gene encoding the muscle fatty acid binding protein from the desert locust, *Schistocerca gregaria*. Insect. Biochem. Mol. Biol. 31, 553-562.

Wunderle, V.M., Critcher, R., Ashworth, A. & Goodfellow, P.N. (1996) Cloning and characterization of *SOX5*, a new member of the human SOX gene family. Genomics. 36, 354-358.

Wurst, W., Bally-Cuif, L., 2001. Neural plate patterning: upstream and downstream of the isthmic organizer. Nat. Rev. Neurosci. 2, 99-108.

Xu, L.Z., Sánchez, R., Sali, A. & Heintz, N. (1996) Ligand specificity of brain lipid-binding protein. J. Biol. Chem. 271, 24711-24719.

Yamamoto, M., Ko, L.J., Leonard, M.W., Beug, H., Orkin, S.H. & Engel, J.D. (1990) Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. Genes Dev. 4, 1650-1662.

Zetterstrom, R.H., Simon, A., Giacobini, M.M., Eriksson, U. & Olson, L. (1994) Localization of cellular retinoid-binding proteins suggests specific roles for retinoids in the adult central nervous system. Neuroscience. 62, 899-918.

Zhang, J. & Haunerland, N.H. (1998) Transcriptional regulation of FABP expression in flight muscle of the desert locust, *Schistocerca gregaria*. Insect Biochem. Mol. Biol. 28, 683-691.

Zhang, J., Rickers-Haunerland, J., Dawe, I. & Haunerland, N.H. (1999) Structure and chromosomal location of the rat gene encoding the heart fatty acid-binding protein. Eur. J. Biochem. 266, 347-351.

Zheng, W.L. & Ong, D.E. (1998) Spatial and temporal patterns of expression of cellular retinol-binding protein and cellular retinoic acid-binding proteins in rat uterus during early pregnancy. Biol. Reprod. 58:963-970.

Zhou, Y., Yamamoto, M. & Engel, J.D. (2000) GATA2 is required for the generation of V2 interneurons. Development. 127, 3829-3838.