Characterization and expression of intracellular lipid-binding protein genes in zebrafish (*Danio rerio*)

by

Mukesh K. Sharma

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: 0-612-89816-4
Our file Notre référence
ISBN: 0-612-89816-4

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 "Characterization and expression of intracellular lipid-binding protein genes in zebrafish (*Danio rerio*)" by Mukesh K. Sharma in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dated: January 30, 2004

External Examiner:

Research Supervisors:

Examining Committee:

Departmental Representative:

DALHOUSIE UNIVERSITY

DATE: 30 January, 2004

AUTHOR: Mukesh K. Sharma

TITLE: Characterization and expression of intracellular lipid-binding protein genes in zebrafish (*Danio rerio*)

DEPARTMENT or SCHOOL: Biology, Dalhousie University

DEGREE: Doctor of Philosophy 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

To my mom and dad.....

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The multigene family of intracellular lipid-binding proteins (iLBP) comprises at least sixteen members that include ten different fatty acid-binding proteins (FABP), four cellular retinol-binding proteins (CRBP) and two cellular retinoic acid-binding proteins (CRABP). All the members of this multigene family are believed to have originated from a single ancestral gene as a result of duplication events. Although the iLBP genes have been intensively investigated in mammals, the physiological functions of their products still remains to be clearly defined and the mechanism of their evolution remains a mystery. I have characterized four members of the iLBP multigene family namely, the intestinal fatty acid-binding protein (I-FABP), basic liver-FABP (Lb-FABP), liver-FABP (L-FABP) and cellular retinoic acid-binding protein II (CRABPII) in zebrafish. The I-FABP, Lb-FABP and CRABPII genes consist of four exons-three introns, similar to that of mammalian iLBPs. The zebrafish L-FABP gene structure is composed of five exons-four introns owing to insertion of an intron in the 5′ UTR. All four genes were mapped to zebrafish linkage groups (LG) by radiation hybrid mapping. The I-FABP gene was mapped to LG 1, the Lb-FABP and CRABPII genes to LG 16 and the L-FABP gene to LG 5. Comparative genomic analysis revealed that zebrafish LG 1 has syntenic relationship to human chromosome 4, zebrafish LG 16 to human chromosome 1 and zebrafish LG 5 to human chromosome 2. The dispersal of the iLBP multigene family members in the zebrafish and the human genome suggests that tandem duplication did not play a major role in evolution of the iLBP multigene family. The pattern of expression of the iLBP genes seems to have diverged as compared to their mammalian orthologs. Whole mount *in situ* hybridization detected I-FABP mRNA in the embryonic zebrafish liver, pancreas and intestine. In adult zebrafish, I-FABP mRNA was detected in the intestine, liver, brain, muscle and testes by RT-PCR analysis. The Lb-FABP mRNA was detected only in the embryonic zebrafish liver but in the adult zebrafish, Lb-FABP mRNA expression was also found in the testes and intestine. The zebrafish L-FABP transcript was not detected during embryonic development and in the adult zebrafish L-FABP mRNA expression was restricted to the intestine. The zebrafish CRABPII mRNA was widely distributed during embryonic development and in adult tissues. The finned-fish lineage is believed to have undergone a whole genome duplication after divergence from the lobe finned-fish lineage. Therefore, zebrafish may have two genes for each of the mammalian orthologs. Database searches at the National Centre for Biotechnology Information and Wellcome Trust Sanger Institute identified putative duplicated copies of L-FABP (L-FABPb) and CRABPII (CRABPIIb) but not of I-FABP or Lb-FABP genes. Preliminary mRNA expression analysis suggests that subfunctionalization may be responsible for the retention of the duplicated genes in the zebrafish genome.

ABBREVIATIONS AND SYMBOLS

ARP-1	Apolipoprotein regulatory protein
B-FABP	Brain fatty acid-binding protein
bp	base pairs
C/EBP	CCAAT/enhancer-binding protein
CIP	Calf intestinal phosphatase
CNS	Central nervous system
cR	centiRay
CRABP	Cellular retinoic acid-binding protein
CRBP	Cellular retinol-binding protein
cRNA	Complementary ribonucleic acid
DHA	Docosahexaenoic acid
dNTP	Deoxy-nucleoside tri-phosphate
DR-1	Direct repeat-1
EST	Expressed sequence tags
FABP	Fatty acid-binding protein
FGB	Fibrinogen, B beta polypeptide
H-FABP	Heart fatty acid-binding protein
HNF	Hepatic nuclear factor
hpf	Hour post-fertilization
IlLBP	Ileal lipid-binding protein
I-FABP	Intestinal fatty acid-binding protein
iLBP	Intracellular lipid-binding protein

kDa	Kilo dalton
Lb-FABP	Basic liver fatty acid-binding protein
LEF1	Lymphoid enhancer binding factor 1
L-FABP	Liver fatty acid-binding protein
LG	Linkage group
LM-PCR	Linker-mediated polymerase chain reaction
LRPAP1	Low-density lipoprotein receptor-related protein associated protein 1
mya	Million years ago
M-FABP	Muscle fatty acid-binding protein
mRNA	messenger ribonucleic acid
MTNR1A	melatonin receptor 1A
mM	Millimolar
NCBI	National Centre for Biotechnology Information
nt	Nucleotide
PCR	Polymerase chain reaction
PPAR	Peroxisome proliferator-activated receptor
PPRE	Peroxisome-proliferator response element
3′ RACE	3′ Rapid amplification of cDNA ends
RACK1	Receptor for activated C kinase 1
RAR	Retinoic acid receptor
5′ RLM-RACE	5′ RNA ligase mediated rapid amplification of cDNA ends
RT-PCR	Reverse transcription-polymerase chain reaction

RXR	Retinoid X-receptor
TAP	Tobacco acid pyrophosphatase
TR	Thyroid hormone receptors
TLBP	Testicular lipid-binding protein
TLL	Tolloid-like
μL	Microliter
μM	Micromolar
UTR	Untranslated region
VDR	Vitamin D_3 receptor
YSL	Yolk syncytial layer

ACKNOWLEDGEMENTS

I am highly indebted to my supervisors Drs. Jonathan M. Wight and Eileen M. Denovan-Wright for their constant encouragement and guidance throughout the period of this study. I express my deep sense of gratitude to Dr. Thomas H. MacRae and Dr. Vett Lloyd, my advisory committee members, for their valuable advice and support. The work involving LM-PCR was performed in collaboration with Dr. Marc Mansour, RLM-RACE with Rong-Zong Liu, phylogenetic analysis with Mary Ellen Boudreau and quantitative RT-PCR with Dr. Eileen M. Denovan-Wright. I wish to thank them for their help. The whole mount *in situ* hybridization was performed by Agnes Degrave, Dr. Christine Thisse and Dr. Bernard Thisse. I also wish to thank my colleagues Steve Mockford, Vishal Saxena and Qian Sun for their informative suggestions and help during the experimental stages of this study. This work would not have been possible without the financial support of a Dalhousie University Graduate Scholarship and Izaak Walton Killam Memorial Scholarship.

Chapter 1: General Introduction

1.1 Intracellular lipid-binding proteins

Long-chain fatty acids and retinoic acid contribute to a variety of biological processes. Long-chain fatty acids serve as metabolic energy sources, are constituents of biological membranes, play essential roles in certain signal transduction pathways and are thought to mediate gene expression (reviewed in McArthur *et al.*, 1999). Retinoic acid is important for normal vertebrate development, cell growth and differentiation (reviewed in Lohnes *et al.*, 1992; Lampron *et al.*, 1995). The intracellular transport of the highly hydrophobic fatty acids and retinoic acid is facilitated by a group of low molecular mass (~15 kDa) proteins encoded by a multigene family of fatty acid (FABP), retinol (CRBP) and retinoic acid (CRABP) binding proteins, collectively referred to as intracellular lipid-binding proteins (iLBP) (Schaap *et al.*, 2002).

At least sixteen members of the iLBP multigene family, believed to have formed from a single ancestral gene by multiple duplication events (Schleicher *et al.*, 1995), have been described (Schaap *et al.*, 2002). Schaap and co-workers (2002) estimated that the first iLBP gene duplication occurred approximately 930 million years ago (mya). Subsequent duplications gave rise to the sixteen distinct iLBP genes currently described. The liver, intestinal, ileal clade of FABPs emerged from the brain, heart, myelin P2 FABP lineage approximately 700 mya. The CRBPs and CRABPs appear to have diverged from the liver/intestinal FABP clade approximately 500 mya (reviewed in Stewart, 2000). In humans, most iLBP genes are dispersed on different chromosomes. Only the genes encoding adipocyte-FABP, myelin P2 FABP, testicular lipid-binding protein (TLBP), and epidermal-FABP located on chromosome 8 (Schaap *et al.*, 2002;

Zimmerman and Veerkamp, 2002) and CRBP I and CRBPII located on chromosome 3 are tightly linked. These two groups of iLBP genes may have arisen by tandem duplication events (Schaap *et al.*, 2002).

The genomic organization of the iLBP genes has been highly conserved over hundreds of millions of years. Each iLBP gene contains four exons interrupted by three introns (Bernlohr *et al.*, 1997). The exceptions to this genomic organization are the desert locust muscle-type FABP gene and a putative muscle-type FABP gene from *Drosophila*, both of which lack intron 2 (Wu *et al.*, 2001).

The proteins encoded by the iLBP multigene family are all small, ranging in size from 127 to 142 amino acid residues (Zimmerman and Veerkamp, 2002; Sharma *et al.*, 2003). The amino acid sequence similarities between iLBPs vary from 20-70% but, based on x-ray crystallography and protein modelling studies, all iLBPs have a similar tertiary structure. The tertiary structure of iLBPs consists of two short α-helices and ten antiparallel β-strands. The β-strands are organized into two orthogonal β-sheets, giving the protein an overall shape of a clam shell (Bernlohr *et al.*, 1997). The hydrophobic ligand is bound within the cavity of the clam shell-like structure.

Different iLBPs exhibit distinct ligand binding preferences. For example, brain-FABP (B-FABP) binds docosahexaenoic acid (DHA) with higher affinity than other fatty acids (Xu *et al.*, 1996). Liver-FABP (L-FABP) binds a broad range of ligands including heme, steroids, acyl CoA, leukotrienes, peroxisome proliferators, prostaglandins, and long-chain fatty acids (Raza *et al.*, 1989; Sorof, 1994; Miller and Cistola, 1993; Richieri *et al.*, 1994) and can bind two ligands at the same time (Thompson *et al.*, 1997b; Haunerland *et al.*, 1984). CRABP type I and II only bind all-*trans*-retinoic acid (Ong and

Chytil, 1978; Kleywegt *et al.*, 1994), and CRBP type I and II bind all-*trans*-retinol or retinal (Levin *et al.*, 1988; Inagami and Ong, 1992). It has been suggested that the ancestral iLBP gene encoded a universal hydrophobic ligand-binding protein that acquired specialized binding affinity during evolution (Schaap *et al.*, 2002).

Some iLBPs are expressed in only one tissue while others are expressed in several tissues. Adipocyte-FABP is expressed only in adipose tissue, TLBP is expressed only in testis, L-FABP is expressed in liver, intestine and kidney and the heart-FABP (H-FABP) is expressed in heart, skeletal muscle, brain, kidney, stomach, mammary gland, lung, placenta and ovary (reviewed in Glatz and van der Vusse, 1996; Hertzel and Bernlohr, 2000). The iLBP family members differ with respect to transcriptional regulation and ligand binding preference, suggesting that individual iLBPs have distinct functions.

Despite more than thirty years of research, the precise biological functions of these proteins remains uncertain. Proposed generic functions for FABPs can be summarised as follows: regulation of fatty acid metabolism by transporting long-chain fatty acids through the aqueous cytoplasm to the sites of β-oxidation or triglyceride synthesis, modulation of activity of enzymes involved in fatty acid metabolism, protection of enzymes and membranes from detrimental effects of excess fatty acids, and regulation of the expression of fatty acid-responsive genes (Börchers and Spener, 1994; Coe and Bernlohr, 1998; Hertzel and Bernlohr, 2000; Storch and Thumser, 2000; Besnard *et al.*, 2002). The functions proposed for CRBPs and CRABPs include transport of retinol/retinoic acid within the cell and to the nucleus, delivery of retinoic acid to the appropriate nuclear receptor, involvement in retinoic acid metabolism and retinol esterification (Wolf, 1991, Fiorella and Napoli, 1991, Ghyselinck *et al.*, 1999).

Most of the work describing iLBP structure and function has been performed on mammalian iLBPs. Few studies have investigated iLBPs in fishes, a large and evolutionary diverse group of vertebrates. The present study was focused on characterization of iLBPs in zebrafish, a model system for vertebrate genetics and development. Studies of zebrafish iLBPs will provide comparative data between the mammalian (rodent and human) and non-mammalian iLBP homologs and may shed light on the function of iLBPs. These studies will also provide important clues for understanding the evolution of the iLBP multigene family and may enrich our understanding of the fundamental processes for the genesis and fate of multigene families that are believed to be the result of duplication events during vertebrate evolution.

In this thesis, I summarize my work on the intestinal fatty-acid binding protein (I-FABP or *fabp2*), basic liver fatty acid-binding protein (Lb-FABP or *fabp 10*), liver fatty acid-binding protein (L-FABP or *fabp1*) and the cellular retinoic acid-binding protein type II (CRABPII or *crabp2*) genes in zebrafish. The determination of developmental and tissue-specific expression patterns, exon/intron organization, 5′ upstream sequence, transcription start sites, potential *cis*-regulatory elements and the linkage analysis of these iLBP genes was performed in zebrafish to gain insight into the evolution of the multigene family of iLBPs.

Chapter 2: Sequence, expression and linkage mapping of the intestinal-type fatty acid-binding protein gene from zebrafish (*Danio rerio*).

2.1 Introduction

The intestinal fatty acid-binding protein (I-FABP) was first isolated from the intestinal mucosa (Ockner *et al.*, 1972) and its expression was thought to be restricted to the intestine and stomach in mammals (Veerkamp and Maatman, 1995). It has been proposed that I-FABP is involved in uptake of dietary fatty acids (Levy *et al.*, 2001) and intracellular fatty acid transport (Baier *et al*, 1996). An alanine-to-threonine polymorphism of codon 54 (A54T) initially found in the I-FABP/*FABP2* gene of Pima Indians was linked to a predisposition to type 2 diabetes (Tataranni *et al.*, 1996). The mutated threonine^{54} I-FABP has two-fold greater affinity for long-chain fatty acids than the alanine^{54} I-FABP (Baier *et al.*, 1996). The A54T polymorphism is associated with a high plasma triglyceride level, insulin resistance (Baier *et al.*, 1996), and an increase in the body mass index (Hegele *et al.*, 1996), but not diabetes mellitus in Japanese men (Yamada *et al.*, 1997), aboriginal Canadians (Hegele *et al.*, 1996) and the Guadeloupe Indian population (Boullu-Sanchis *et al.*, 1999). I-FABP gene knockout experiments showed that I-FABP deficiency does not decrease the capacity of mice to absorb dietary fat (Vassileva *et al.*, 2000). André *et al.* (2000) correlated the expression of I-FABP in the intestine to the intracellular storage of lipid droplets and synthesis of very low-density lipoproteins in zebrafish larvae. Although the I-FABP gene has been intensively investigated, its physiological function remains to be clearly defined.

In this study, I present the exon/intron structure, linkage analysis, 5′ upstream sequence, transcription start site and the potential *cis*-regulatory elements of the zebrafish

I-FABP (*fabp2*) gene. The mRNA expression in a number of adult zebrafish tissues using RT-PCR and quantification of I-FABP specific message using quantitative RT-PCR is presented. Onset of I-FABP gene expression during zebrafish embryogenesis is analyzed by both RT-PCR and whole-mount *in situ* hybridization.

2.2 Materials and methods

2.2.1 Fish Stocks

Zebrafish (*Danio rerio*) were purchased from a local aquarium store and maintained at 28.5 °C in 35 litre aquaria. Fish were fed with dry fish feed, TetraMin Flakes (TetraWerke, Melle, Germany) in the morning and hatched brine shrimp (*Artemia* cysts from INVE, Grantsvillle, UT, USA) in the afternoon. Fish were maintained on a 14 h light and 10 h dark cycle. Fish breeding and embryo manipulation were conducted according to Westerfield (1995).

2.2.2 Sequence of the I-FABP gene

The sequence of a zebrafish I-FABP cDNA (Pierce *et al.*, 2000) was used to search the zebrafish genome sequence database at The Wellcome Trust Sanger Institute (http://www.ensembl.org/Danio_rerio/). Traces containing different exons and introns of the zebrafish I-FABP gene were retrieved and extended using overlapping traces. The complete sequence of intron 2, not present in the zebrafish genome sequence database, was amplified by PCR with a sense primer specific to exon 2 (5′ CACCTTTGACTATT CTCTGGCA 3’; nt 311 to 332, Fig. 1) and an antisense primer specific to exon 3 (5′ CA CCATTAACGATAGTCCTGAC 3′; nt 2598 to 2577, Fig. 1) of the zebrafish I-FABP gene. The PCR product was size fractionated by 1% agarose gel-electrophoresis and the

single band was excised from the gel. The excised band was purified using QIAquick Gel Extraction kit (QIAGEN Inc.), cloned into the plasmid vector using TOPO®XL PCR Cloning Kit (Invitrogen Corporation) and sequenced. The sequencing was performed using LI-COR® 4000 sequencer at National Research Centre, Halifax, Canada.

2.2.3 I-FABP promoter cloning

To clone the core promoter and the 5′ upstream sequence of the zebrafish I-FABP gene, linker-mediated polymerase chain reaction (LM-PCR) was employed. Genomic DNA was isolated from adult zebrafish and purified according to standard protocols (Sambrook *et al.*, 1989). Two μg samples of genomic DNA were digested separately with the restriction enzymes, *Bgl* II and *Eco* RI, and T4 DNA ligase (Promega) was used to ligate 0.5 μg *Bgl* II digested genomic DNA to the double-standard DNA linker,

5′ GTACATATTGTCGTTAGAACGCGTAATACGACTCACTATAGGGA 3′
3′ CATGTATAACAGCAATCTTGCGCATTATGCTGAGTGATATCCCTCTAG 5′,

and 0.5 μg of *Eco* RI digested genomic DNA to the double-standard DNA linker,

5′ GTACATATTGTCGTTAGAACGCGTAATACGACTCACTATAGGGAGAG 3′
3′ CATGTATAACAGCAATCTTGCGCATTATGCTGAGTGATATCCCTCTCTTAA 5′.

Following ethanol precipitation, the DNA was re-suspended in 15 μL of sterile, double distilled water. Two partially overlapping sense primers (C_1, C_2) were synthesized based on the linker sequence (C_1: 5′ GTACATATTGTCGTTAGAACGCGTAATACGAC TCA 3′; C_2: 5′ CGTTAGAACGCGTAATACGACTCACTATAGGGAGA 3′). The first round of PCR was performed using primer C_1 and an external I-FABP gene-specific antisense primer (5′ CTTCTCGTAGTTCTCATTGCGG 3′; nucleotides 84 to 63, GenBank accession number AY266452) that would anneal to a sequence within the first

exon of the zebrafish I-FABP gene. The 50 μL reaction contained 1×PCR buffer, 1.25 U of *Taq* DNA polymerase (MBI 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 at 94°C for 2 min, the reaction was subjected to 35 cycles of amplification at 94°C for 30s, 63°C for 40s, and 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 C_2 and an internal I-FABP gene-specific antisense primer (5′ CCAGGTCCCGTTGAAGGTCA 3′; nucleotides 54 to 35, GenBank accession number AY266452). The conditions for the secondary PCR were similar to those of the primary PCR with the exception that the reaction was subjected to 24 cycles of amplification. The product from the secondary PCR was size fractionated by 1% agarose gel-electrophoresis. A single band of ~1000 bp was obtained from genomic DNA digested with *Bgl* II and a single band of ~450 bp was obtained from genomic DNA digested with *Eco* RI. These bands were excised from the agarose gel and purified according to standard protocol (Sambrook *et al.*, 1989). The purified DNA fragments were cloned into plasmid using pGEM-T Easy vector system (Promega), and a single clone was sequenced from both directions for each promoter fragment. Potential *cis*-acting regulatory elements for the I-FABP promoter were identified by computer-assisted analysis performed using MatInspector professional at http://transfac.gbf.de/cgi-bin/matSearch/matsearch.pl (Quandt *et al.*, 1995).

2.2.4 Linkage analysis by radiation hybrid mapping

LN54 radiation hybrid panel (zebrafish DNA in a mouse background) (Hukriede *et al.*, 1999) was used to localize the I-FABP gene to a zebrafish linkage group. Sense

(5′ GCTGCCCATGACAACCTGAAG 3′; nucleotides 224 to 204, GenBank accession number AY266452) and antisense (5′ AAGGTGACGCCCAGAGTAAAG 3′; nucleotides 316 to 296, GenBank accession number AY266452) primers were used to PCR-amplify exon 2 of the zebrafish I-FABP gene. Each 25 μL reaction contained 1×PCR buffer (MBI Fermentas), 1.5 mM $MgCl_2$, 0.4 μM each of sense and antisense primers, 0.2 mM dNTP, 1.25 U of *Taq* DNA polymerase and 100 ng of hybrid-cell DNA. The three control reactions contained 100 ng of DNA from each of the parental cell lines (AB9 or B78 DNA and a 1:10 mixture of zebrafish AB9 DNA and mouse B78 DNA). PCR conditions were: 94°C for 4 min, 32 cycles of 94°C for 30s, 59°C for 30s and 72°C for 30s followed by 72°C for 7 min. Fifteen μL of the PCR product was subjected to 2% agarose gel electrophoresis. The radiation hybrid panel was scored based on the absence (0) or presence (1) of the expected 307 bp DNA fragment, 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).

2.2.5 Mapping the transcription start site of the zebrafish I-FABP gene

To map the transcription start site of the zebrafish I-FABP gene, 5′ RNA ligase-mediated rapid amplification of cDNA ends (5′ RLM-RACE) was employed. Total RNA was extracted from adult zebrafish using Trizol® (Gibco BRL). The RLM-RACE kit (Ambion) was used to prepare cDNA following the manufacturer's instructions. Briefly, total RNA was treated with calf intestinal phosphatase (CIP) to remove the 5′-phosphate groups from unprocessed or degraded mRNA, structural RNAs or contaminating genomic DNA. In this treatment, the ^{7}methyl G cap structure (cap) on the intact 5′ ends

of mRNA is not cleaved by CIP. The RNA was divided into two aliquots. One aliquot was treated with tobacco acid pyrophosphatase (TAP) to remove the cap structure from full-length mRNA and the other aliquot was not treated with TAP. The two preparations of RNA (plus and minus TAP treatment) were incubated with a 45 base RNA adapter (5′ GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUGAAA 3′) and T4 RNA ligase. The RNA adapter was ligated to the full-length mRNA and not to the dephosphorylated RNA that lacked 5′ phosphate groups necessary for ligation. A random primed reverse transcription was performed to synthesize cDNA from both plus and minus TAP RNA preparations. Nested PCR using two sense adapter primers (outer: 5′ GCTGATGGCGATGAATGAACACTG 3′ and inner: 5′ CGCGGATCCGAACACTGC GTTTGCTGGCTTTGATG 3′) and two antisense primers specific to I-FABP mRNA (outer: 5′ AAGGTGACGCCCAGAGTAAAG 3′; nucleotides 316 to 296, GenBank accession number AY266452 and inner: 5′ CTTCTCGTAGTTCTCATTGCGG 3′; nucleotides 84 to 63, GenBank accession number AY266452) was performed to amplify the 5′ end of the I-FABP mRNA. The 10 μL reaction of the first round of PCR contained 1×PCR buffer (ClonTech), 1.25×cDNA polymerase mix (ClonTech), 0.5 mM of each dNTP, 0.5 μM of each outer primer and 0.5 μL of cDNA from reverse transcription reactions. Following an initial denaturation at 94°C for 1 min, the reaction was subjected to 35 cycles of amplification at 94°C for 30 s, 60°C for 30 s, 72°C for 30s, and a final extension at 72°C for 10 min. Half a μL of the primary PCR product from the plus and minus TAP reactions was used as template for a second round of PCR (nested PCR) with 1×PCR buffer (ClonTech), 0.625×cDNA polymerase mix (ClonTech), 0.5 mM of each dNTP, 0.25 μM of each inner primer in a total volume of 20 μL. The conditions for the

secondary PCR were identical to those of the primary PCR. The product from the secondary PCR was size fractionated by 1% agarose gel-electrophoresis and a single band representing abundant transcript specific for the I-FABP gene was excised from the gel. The excised band was purified using QIAquick Gel Extraction Kit (QIAGEN), cloned into plasmid using pGEM Easy vector system (Promega) and sequenced to determine the I-FABP transcription start site.

2.2.6 Developmental and tissue-specific expression of I-FABP mRNA in zebrafish

To determine the tissue-specific expression of the zebrafish I-FABP, total RNA was extracted from adult zebrafish heart, brain, muscle, testis, liver, ovary, skin, intestine and brain. For studies of the temporal expression of the I-FABP gene during embryogenesis, total RNA was extracted from 1, 3, 12, 24, 60 hour post-fertilization (hpf) embryos and adult zebrafish. All RNA extractions were performed using Trizol® (Gibco BRL). First strand cDNA was synthesized by reverse transcriptase-Superscript II (Gibco BRL) using 1 μg of total RNA from each sample as template. The cDNA generated was used as template in PCR with sense (5′ AAGTCGACCGCAATGAGAAC 3′; nucleotides 56 to 75, GenBank accession number AY266452) and antisense (5′ GTTTGACATTGGGAGTGCAG 3′; nucleotides 2771 to 2752, GenBank accession number AY266452) primers specific to I-FABP mRNA. The constitutively expressed mRNA for the receptor for activated C kinase (RACK1) (Hamilton and Wright, 1999) was employed as a positive control in RT-PCR. The RACK1 cDNA was amplified in tandem with experimental samples from all tissue-derived cDNAs assayed using sense (5′ ATCCAACTCCATCCACCTTC 3′; nucleotides 14 to 33, GenBank accession number

AF025330) and antisense (5′ ATCAGGTTGTCAGTGTAGCC 3′; nucleotides 977 to 958, GenBank accession number AF025330) primers. PCR reactions without cDNA template were used as negative controls in each RT-PCR experiment. The 25 μL PCR reactions contained 1×PCR buffer, 1.25 U of *Taq* DNA polymerase (MBI Fermentas), 1.5 mM $MgCl_2$, 0.2 μM each dNTP, 0.4 μM each of sense and antisense primers, and 1 μL of cDNA. PCR conditions involved initial denaturation at 94°C for 2 min, followed by 30 cycles of 94°C for 30s, 53°C for 30s, 72°C for 1 min and final extension at 72°C for 5 min. Fifteen μL of each PCR product was size fractionated by 1% (w/v) agarose gel-electrophoresis.

Quantitative RT-PCR for I-FABP and β-actin, another constitutively expressed gene, was performed using the LightCycler thermal cycler system (Roche Diagnostics) according to the manufacturer's instructions. The I-FABP-specific primers used for conventional RT- PCR were also used for quantitative RT- PCR. The β-actin cDNA was amplified using sense (5′ AAGCAGGAGTACGATGAGTCTG 3′; nucleotides 1128 to 1149, GenBank accession number NM_131031) and antisense (5′ GGTAAACGCTTC TGGAATGAC 3′; nucleotides 1405 to1385, GenBank accession number NM_131031) primers. Serial dilutions of bacteriophage lambda DNA and gel-purified I-FABP and β-actin RT-PCR products were allowed to bind SYBR® Green I dye and the amount of bound SYBR® Green I was determined by fluorimetry. The concentration of I-FABP 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. Dilutions of the I-FABP and β-actin product ranging from 5×10^6 to 5×10^1 copies per reaction were used in individual quantitative

PCR reactions to determine the standard curve of the crossing points for the amplification of I-FABP and β-actin from tissue-specific cDNA samples. 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. 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 I-FABP/ β-actin PCR product for each experimental sample was calculated. The PCR reaction to amplify I-FABP contained 1 μL of cDNA, 0.2 μM sense and antisense primers, 3 mM $MgCl_2$, and 1×LightCycler-DNA FastStart SYBR® Green I Mix containing nucleotides, buffer, and hot start *Taq* DNA polymerase. Parallel reactions were performed simultaneously to analyze multiple cDNA samples. The PCR conditions involved 15 min at 95°C to activate *Taq* DNA polymerase, 45 cycles of 95°C for 15 s, 55°C for 5 s and 72°C for 16 s. Fluorescent signal was measured at the end of each extension phase. For negative controls, cDNA was omitted from the reaction.

Whole-mount *in situ* hybridization of embryos was performed by A. Degrave, C. Thisse and B. Thisse using riboprobes based on the I-FABP cDNA clone (Pierce *et al.*, 2000) as described by Thisse and Thisse (http://zfin.org/zf_info/zfbook/chapt9/9.82.html).

2.3 Results and discussion

2.3.1 Sequence and organization of the zebrafish I-FABP gene

DNA traces with identity to a zebrafish I-FABP cDNA (Pierce *et al.*, 2000) were retrieved from the Wellcome Trust Sanger Institute (http://trace.ensembl.org/perl/ssahaview). Trace Z35723-a1242f09.p1c contained the sequence for exon 1, intron 1 and

part of exon 2, trace Zfish37251-1064h08.q1c contained the sequence for exon 2, trace Z35725-a3053h05.p1c contained part of intron 2, exon 3, intron 3 and part of exon 4, and a fourth trace zfish44909-2631a10.q1k contained the sequence for part of intron 2, exon 3, intron 3 and exon 4 of the zebrafish I-FABP gene, respectively. The complete sequence of intron 2, not present in the zebrafish genome sequence database, was obtained from genomic DNA by PCR. A 924 bp genomic DNA fragment 5′ upstream of the initiation codon for the I-FABP gene was cloned by linker-mediated PCR following digestion of genomic DNA with the restriction enzyme *Bgl* II and sequenced.

The zebrafish I-FABP gene contains four exons of 100 bp, 173 bp, 108 bp and 243 bp interrupted by three introns of 77 bp, 2151 bp, 83 bp, respectively (Fig. 1). Southern blot and hybidization has previously shown evidence for a single copy of the I-FABP gene in the zebrafish genome (Pierce *et al.*, 2000). Each of the intron/exon splice junctions of the zebrafish I-FABP gene follows the "GT-AG" dinucleotide rule. Similar exon/intron organization consisting of four exons interrupted by three comparably positioned introns has been reported for mammalian FABPs (Green *et al.*, 1992; Sweetser *et al.*, 1987; Bernlohr *et al.*, 1997). The exception to this FABP gene organization is the muscle-type FABP (M-FABP) gene from the desert locust M-FABP and the putative *Drosophila* M-FABP, both of which lack intron 2 (Wu *et al.*, 2001). Data mining of genome sequence banks showed that intron 2 was present in the zebrafish and the pufferfish I-FABP genes at a position comparable to mammalian I-FABPs (Fig. 2A). Thus, the loss of intron 2 in the locust and *Drosophila* M-FABP seems to be restricted to the insect lineage or probably occurred in the invertebrates after the divergence from vertebrates.

Figure 1. Nucleotide sequence of the zebrafish I-FABP gene and its 5′ upstream sequence. Exons are shown in capital letters with the deduced amino acid sequence for each exon indicated below. Introns are shown in lower case and are italicized. The arrow indicates the transcription initiation site mapped using 5′ RLM-RACE. The sequence 5′ upstream of the transcription start site is in lower case and putative TATA and GC boxes are outlined. A polyadenylation signal sequence, AATAAA, is italicized and in bold font. A variation between the genomic sequence and the cDNA sequence is shown in bold font with the variation indicated above.

```
5′tctatgaaataatattggcaggtgcaaattttcattttaataatgtactagtactgtaaatatttgagttagcttcagaactggctttg -803
aagtatattcactggtagctgtctgttaagtagaactgcacaaaaaaaaggtacgaggaagaacgaagcattgccatatactgtattctca -712
attttagctgcaatagctgcacttttgtagtaaaaagttcttcaaagacccttttttatatagtgtgaagaacttttcaataatcttaaaac -621
atttttctacataaataaaatttttgtgcaatgaaatatttcaaaaagtactttataaaaccaaccattcatgtcaacacagagccttaaa -529
ataattgttaccagtatttctcaatagccaatgaggcatgcagcaagattgtaaactaatttaactgtaaccattaaatctaaagttttga -439
cctttgacctttcaccacaattgtttaaaaatcaatacatgtggctagaattcatagctttcaaacactcttcatagtccctcagtagtta -348
ctgcaacacaaggcacttactagtataacagacacaccctacacagtttaaacagcaaacaattcaaacaaaagcaggactttgttatgag -257
aacaacagagcgtgtatgtttgaccatcgcagtacaagtgagtgcttgttctcaggacttgaaatctccgctctttgcacaatcaatgaat -166
aagcaaggcatgctgggatgtgtgtaacatatagcctgtt[gggcgg]gtgagatttatacttggtgagctttactcggccacatcagcatga -75
                                          GC BOX
agataatactcagataaggcaacgcttcgccactcgcacagg[tataaaa]gagtggctcggggtaaagttaggccACTGTCAGGATCACACA  17
                                           TATA Box
ACAGTCTGTCATCATC    ATG ACC TTC AAC GGG ACC TGG AAA GTC GAC CGC AAT GAG AAC TAC GAG AAG TTC  87
                    Met Thr Phe Asn Gly Thr Trp Lys Val Asp Arg Asn Glu Asn Tyr Glu Lys Phe

ATG GAA CAA ATG G gtaagtgatgaataataatcgatgcttgcttgagtgagttgattgattaagtgagatttactgcattccttta  173
Met Glu Gln Met

acag GC GTC AAC ATG GTG AAA AGG AAA CTG GCT GCC CAT GAC AAC CTG AAG ATC ACC CTG GAG CAG ACC  242
    Gly Val Asn Met Val Lys Arg Lys Leu Ala Ala His Asp Asn Leu Lys Ile Thr Leu Glu Gln Thr

GGA GAC AAG TTC AAC GTG AAG GAA GTC AGC ACT TTC CGC ACA CTG GAA ATT AAC TTT ACT CTG GGC GTC  311
Gly Asp Lys Phe Asn Val Lys Glu Val Ser Thr Phe Arg Thr Leu Glu Ile Asn Phe Thr Leu Gly Val

ACC TTT GAC TAT TCT CTG GCA GAC GGC ACT GAG CTC ACA gtaagcacagtcattgctgatttcagatgcataagcgtg  389
Thr Phe Asp Tyr Ser Leu Ala Asp Gly Thr Glu Leu Thr

gttccatagtaagcccataacattcaatctgcacaaagggttctttatagtgtgctgactgcacttttttttaagaatttttatttttaaac  480
attgttatacaaattgacaaaggtttaatttttcagtttagcaaaaaataaaaatagtaaattaaataaattaaaagaagttaatacaaaa  571
aacatcaagaataatatagattcaaaataacttaacacactttgagagaaaaaaaataaaataacaatattaatgttaaataatacagcaat  662
aacatatttgtctaatctatatactttcaatgtcagtttaagaaaagagtcaccatcaccatttctccaaaaagccatagaagattttcca  753
aatagaccaaaataaatgaattttgttttccaaaaaataggtttgctgtctgcaccttactactagtttctttaaatgtttccagttcat  844
agtctttaacattctctggtgcacataatgtgctgtgaaagaacaaaaaatgtactaatttataaacactaaagactttatttttacttaaa  935
aaaaaaacacacatttttgaaaaacctgcagcaaatgttgaccacagctggtctctacataggaaaaaaaatggttatttgattatttcaa 1026
attttctgcacactgggttcttataagaactgttcactgaaaggatttttggagaaccaaaaaaaaggggtaaaaaaatcaagagcaaagg 1117
ttgagaggtattgaatggatgatgtaaccatatcacatagaaatctgacatagtcctttgacaatatatgcaaattgcttgtctgagacac 1208
aagttcatatttgtatttgtaaatataaaatataagtcatatatgcagcatccactaccagtcaaaagtttggggtcagtgcgatttta 1299
aatgttttaaaataagcttcttctgctcaccaaggctgcatttatttaaacaaaaatacagtacaaatttgtaaagttgtatataacttttc 1390
aaaagtagtttatcatttaatttaatcatttattccactgattttaatggtgaatttttagcattactccagtcttcagagtcacaggatc 1481
cttcagacatccctcttatattaaatattattgttattattattattattgttattattattattattattgttgttgtaataataataat 1572
aataataataataataataattattattattattattattattattattattattattattattgataatagtaaaaaaaaacaataattactg 1663
aagtattaatttcatttaaaactacatttaaatttttaataattgtttttttaattacatttaataaatgctaccttgatgagcagtataat 1754
aacaaatgaaaatacctgaccccaaaatttttggtatattttaaatacatatacagtatatttcagcctaaacgaaattctaaatatgtaaa 1845
tatatattacaacatattgcttatttatatatacttattacttatataacatacatatataatatacttatataacatttatgcatattta 1936
tctaattccaatggtggaaaaaatatacagtatgtatgtggccatatatttcaaaatatgttgcatatatgacatatattgtaggtgaaat 2027
ccaaatatgaacttgtatctcatatttgcaatttgcacatcctgccatgaatcagatttgtaaagaggataccatgaaatagtgtaaaaaa 2118
aaaaaaaactattccacacaatttatgcatgttatacttacacaaatcgatttatttactgtaagtggattaaagataaacaattaagttt 2209
tccaaaaaatctttagaattgtgttgttgcaactcattttaaacaagtagattaaacaaacagcaaaaaacattttctgagtgtcatgttg 2300
tacaatggcttactgactatttttgctgcatcattgtgctttattgataagcaacaaagttcaggctttgagtatagcagagaagaataaa 2391
ttaattacataacacatgtttacaagctttctggagagaaaatcactaaccggtcaactagatctggcgccgcttgaaaatgagcctgacc 2482
tttgactttctgattgcag GGA TCC TGG GTC ATA GAG GGA GAC ACG CTC AAG GGG ACT TTC ACA CGC AAG GAC 2555
                    Gly Ser Trp Val Ile Glu Gly Asp Thr Leu Lys Gly Thr Phe Thr Arg Lys Asp

AAC GGA AAG GTA CTA ACA ACA GTC AGG ACT ATC GTT AAT GGT GAA CTT GTA CAG gtgaggagttgtttcttat 2628
Asn Gly Lys Val Leu Thr Thr Val Arg Thr Ile Val Asn Gly Glu Leu Val Gln

atattatgacgaatgtttggatttcctgtaaagtagtgttctcacttttgcatttgcttacag    AGC TAT AGC TAT GAT GGA 2710
                                                                    Ser Tyr Ser Tyr Asp Gly

GTC GAG GCC AAG AGG ATT TTC AAG AGG GCT TAA ACTGTTATCTGCACTCCCAATGTCAAACTTACAACTGAAGTGGGACA 2790
Val Glu Ala Lys Arg Ile Phe Lys Arg Ala  *

ATATGAAACTTTATAGTGTTTGAATATTTAACCTGAAAATAGCTTTAATGATTGGCAGCCCTAAAGTTTACAAAACATTTATGTAAATCGT 2881
                                                t
GTCATTTGCACTGTGAGTTTTGATATTTATGAT[AATAAA]GCTATCTAAAGGCCC 3′                                  2953
```

Figure 1. Nucleotide sequence of the zebrafish I-FABP gene and its 5′ upstream sequence.

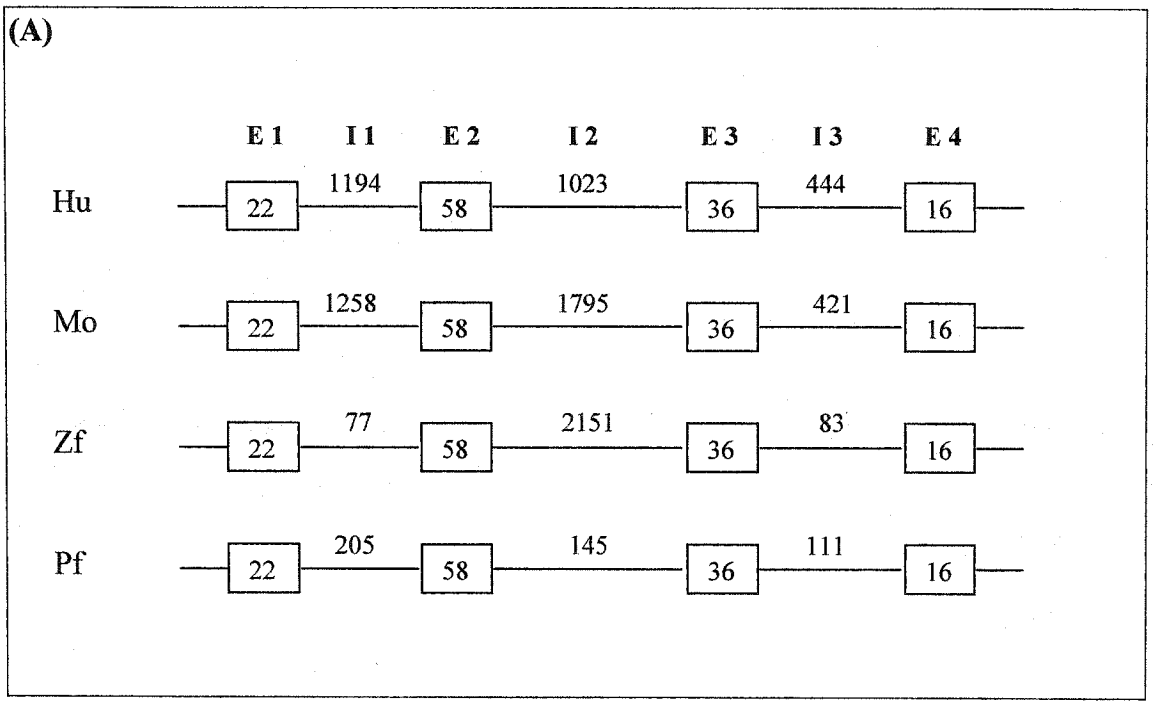
Exons 1- 4 of the zebrafish I-FABP gene encodes 22, 58, 36 and 16 amino acids, respectively, as reported for the human and mouse I-FABP genes (Green *et al.*, 1992; Sweetser *et al.*, 1987), but there are large differences in the size of the introns 1-3 among human, mouse, pufferfish and zebrafish (Fig. 2A). The amino acid sequence for each of the individual exons of the orthologous I-FABP gene from mouse and human were aligned with the deduced amino acid sequence of the pufferfish and the zebrafish I-FABP gene (Fig. 2B). The amino acid sequence encoded by exon 3 was the least conserved with only 55% identity between zebrafish and human, zebrafish and mouse and 63% identity between zebrafish and pufferfish. This observation is in agreement with earlier reports that the NH_2-terminal halves of the FABP family members have been more highly conserved through the course of evolution than their COOH-terminal halves (Sweetser *et al.*, 1987 and references therein).

2.3.2 Radiation hybrid mapping of the zebrafish I-FABP gene

To localize the I-FABP gene to a zebrafish linkage group, radiation hybrid mapping using the LN 54 panel of radiation hybrids was performed (Hukriede *et al.*, 1999). The zebrafish I-FABP gene was assigned to linkage group 1 at a distance of 30.79 centiRay (cR) from the marker Z1463 with a LOD of 16.7.

Comparison of the mapped zebrafish, human and rat I-FABP genes revealed that zebrafish linkage group 1 has a syntenic relationship with human chromosome 4 and rat chromosome 2. The genes for homeo box (H6 family) 1 (HMX1), low-density lipoprotein receptor-related protein associated protein 1 (LRPAP1), lymphoid enhancer binding factor 1 (LEF1), fatty acid-binding protein 2, intestinal (I-FABP/FABP2), fibrinogen, B beta polypeptide (FGB), tolloid-like 1, human (TLL1), and melatonin

Figure 2. Organization of I-FABP genes in fishes and mammals. (A) Comparison of the exon/intron organization of the zebrafish I-FABP gene (Zf) with the orthologous genes from human (Hu), mouse (Mo) and pufferfish (Pf). Coding region of the genes is shown in boxes as E1-E4 with the number of amino acids encoded by each exon indicated within the box. Introns (I1-I3) are shown as horizontal lines connecting adjacent exons and the size of each intron in bp is shown above the line. The sequences for the I-FABP genes were obtained from: Sweetser *et al.* (1987) (human, GenBank accession number M18079); Green *et al.* (1992) (mouse, GenBank accession number M65033); the pufferfish sequence was retrieved from scaffold 537 by searching the *Fugu rubripes* (pufferfish) genome database v.2.0 at http://genome.jgipsf.org/fugu3/fugu3.home.html. (B) The deduced amino acid sequence encoded by each exon of zebrafish I-FABP gene (ZfExon 1-4) was aligned with the amino acid sequence encoded by respective exons from human (HuExon 1-4), mouse (MoExon1-4) and pufferfish (PfExon 1-4) I-FABP genes using CLUSTALW (Thompson *et al.*, 1994). The percentage identity for the amino acids encoded by each exon of the zebrafish I-FABP gene with each exon of the human, mouse and pufferfish I-FABP gene is shown on the right of each exon. The dots in the alignment indicate amino acid identity.



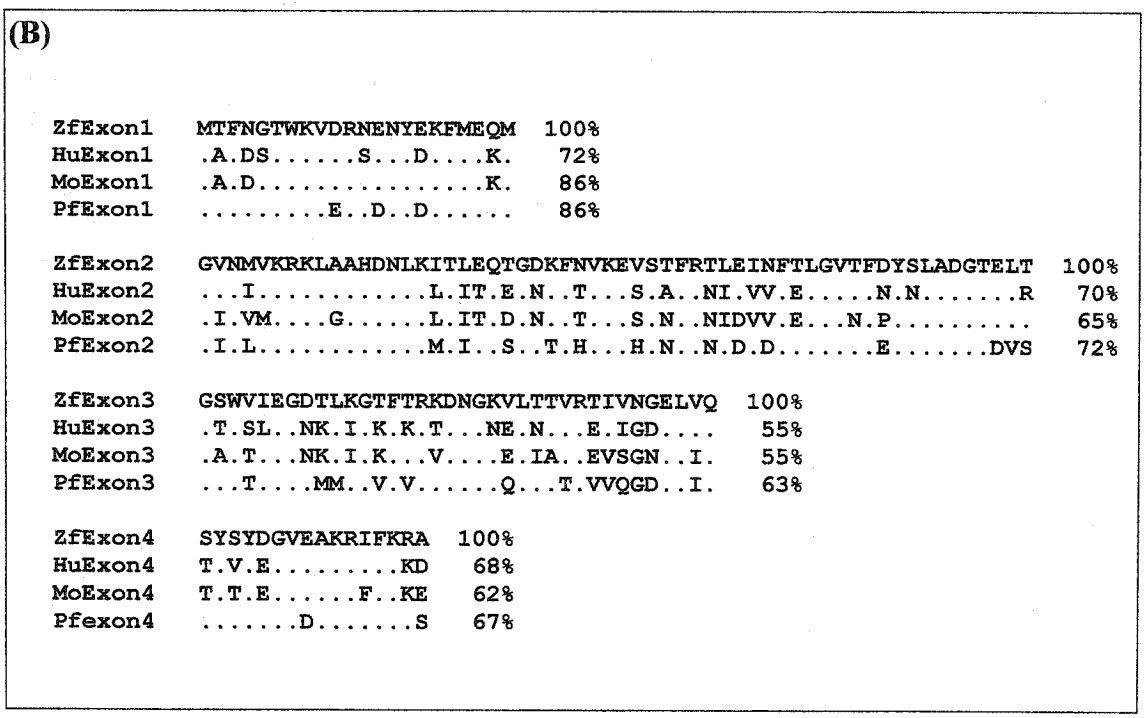


Figure 2. Organization of I-FABP genes in fishes and mammals.

receptor 1A (MTNR1A) are located on chromosome 4 in humans and linkage group 1 in zebrafish. In the rat genome, the genes for I-fabp/Fabp2, Lef1 and Fgb are located on chromosome 2 (Table 1). This synteny suggests that a common linkage group was inherited from the ancestor of fishes and mammals. In mouse, however, the synteny has not been maintained as the genes for I-fabp, Lef1 and Fgb are located on chromosome 3, Hmx1and Lrpap1 on chromosome 5, and Tll and Mtnr1a on chromosome 8 (Table 1). The location of the same set of genes on one linkage group in zebrafish and a single chromosome in humans, but on three different chromosomes in mouse suggests that rearrangements or translocations have taken place in this region of the mouse genome after the divergence of fishes and mammals.

2.3.3 Transcription start site of the zebrafish I-FABP gene

To map the transcription start site of the zebrafish I-FABP gene, we obtained the 5′ cDNA end from the capped and complete mRNA using 5′ RLM-RACE. One abundant transcript of approximately 120 bp was detected in the reaction with calf intestinal phosphatase/tobacco acid pyrophosphatase (TAP) treated RNA amplified in nested PCR with the primer specific to I-FABP mRNA and the adapter primer (Fig. 3). This band, which represents the 5′ end of the mature I-FABP mRNA, contained a 62 bp sequence corresponding to a portion of exon 1 including the 5′ untranslated region (UTR) of the zebrafish I-FABP mRNA. Alignment of the 5′ RLM-RACE sequence with the zebrafish I-FABP gene sequence localized the potential transcription start site 33 bp upstream of the initiation codon (Fig. 1). A putative TATA box and a GC box are located 25 bp and 120 bp, respectively, upstream of the transcription start site. A putative CCAAT box, generally found 80 bp upstream of the transcription start site in

Table 1. Zebrafish-mammals conserved syntenies defined by the I-FABP gene.

Zebrafish[1]			Human[2]		Mouse[2]		Rat[2]	
LG	**Locus zebrafish**	**Accession Number**	**Gene**	**Chromosomal Position**	**Gene**	**Chromosomal Position**	**Gene**	**Chromosomal Position**
1	hmx1	AI658291	HMX1	4p16.1	Hmx1	5 18.0 cM	Hmx1	-
1	lrpap1	AW154084	LRPAP1	4p16.3	Lrpap1	5 20.0 cM	Lrpap1	-
1	lef1	AF136454	LEF1	4q23-q25	Lef1	3 61.6 cM	Lef1	2q34-q45
1	**i-fabp/ fabp2**	**AF180921**	**I-FABP/ FABP2**	**4q28-q31**	**Fabp2**	**3 55.0 cM**	**Fabp2**	**2**
1	fgb	AA658651	FGB	4q28	Fgb	3 48.2 cM	Fgb	2q31-q34
1	tolloid	AF027596	TLL1	4q32-q33	Tll	8 32.4 cM	Tll	-
1	mel1ar	U31822	MTNR1A	4q35.1	Mtnr1a	8 25.0 cM	Mtnr1a	-

[1]Woods *et al.*, 2000.

[2] LocusLink (http://www.ncbi.nlm.nih.gov/LocusLink/list.cgi), National Centre for Biotechnology Information.

Figure 3. Determination of the transcription start site of zebrafish I-FABP gene by 5′ RLM-RACE. Total RNA from whole zebrafish was treated with calf intestinal phosphatase. One aliquot was treated with tobacco acid pyrophosphatase and ligated to an RNA adapter (+TAP, lane 1). A second aliquot was not treated with tobacco acid pyrophosphatase prior to adapter ligation (–TAP, lane 2). Nested PCR with TAP-treated and TAP-untreated total RNA was performed with zebrafish I-FABP mRNA-specific primers and the adapter primer. An abundant transcript of approximately 120 bp representing the 5′ end of the mature zebrafish I-FABP mRNA was amplified in the +TAP reaction but not in the –TAP reaction. The size marker (L) was the 100bp DNA ladder from MBI Fermentas.

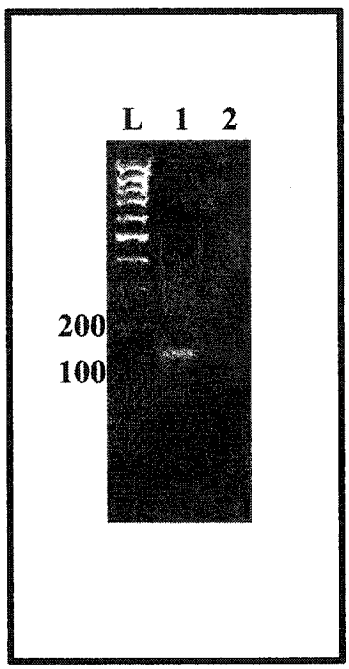


Figure 3. Determination of the transcription start site of zebrafish I-FABP gene by 5′ RLM-RACE.

many eukaryotic RNA polymerase II-dependent promoters was located 496 bp upstream of the transcription start site in the zebrafish I-FABP gene. An element homologous to the CCAAT box is located 46 bp upstream of the transcription start site in the mouse I-FABP gene promoter (Green *et al.*, 1992), but analysis of the human and rat I-FABP promoters did not reveal conserved CCAAT box elements (Sweetser *et al.*, 1987). No CCAAT box element is located at or near this position in the zebrafish I-FABP gene.

2.3.4 Putative 5′ upstream *cis*-acting regulatory elements of the zebrafish I-FABP gene

The 891 bp sequence 5′ upstream of the transcription start site in the zebrafish I-FABP gene was screened for potential *cis*-acting regulatory elements. A 14 bp repeat element (consensus: 5′ TGAACTTTGAACTT 3′) is found in the 5′ upstream sequence of human, rat and mouse I-FABP genes (Green *et al.*, 1992, Sweetser *et al.*, 1987). In zebrafish, a 14 bp element composed of two 7 bp direct repeats (5′ TGACCTTTGACC TT 3′) was found at nt -428 to -441 in the 5′ upstream region of the I-FABP gene (Fig. 1). This sequence matches the consensus sequence for the 14 bp repeat element (Green *et al.*, 1992; Sweetser *et al.*, 1987) at 12 out of 14 positions. The 14 bp repeat element is the binding site for the two members of the steroid/thyroid hormone receptor superfamily of transcription factors, the hepatic nuclear factor-4 (HNF-4) and the apolipoprotein regulatory protein (ARP-1) (Rottman and Gordon, 1993) and have been referred to as a putative direct repeat-1 (DR-1) element (Bernlohr *et al.*, 1997). The DR-1 element ($TG^{A}/_{T}CCT$ in a direct repeat separated by one nucleotide) or the imperfect version of this element is a peroxisome-proliferator response element (PPRE) to which different peroxisome proliferator-activated receptors (PPARs) bind. PPARs are members of the

steroid/thyroid hormone receptor superfamily and are activated by fatty acids and peroxisome proliferators such as fibrate class of hypolipidemic drugs (Bernlohr *et al.*, 1997). PPARs activate gene transcription by cooperatively binding to PPREs as a heterodimer with the retinoid X receptor (RXR). PPARs and RXRs are both expressed in the intestine (Ibabe *et al.*, 2002; Ishaq *et al.*, 1998; Bernlohr *et al.*, 1997). PPREs are also found in the promoter region of other members of the iLBP multigene family including the murine adipocyte P2 gene (Bernlohr *et al.*, 1997) and rat L-FABP gene (Simon *et al.*, 1993). Members of the steroid/thyroid hormone receptor superfamily such as HNF-4, ARP-1, PPAR, retinoic acid receptor (RAR), vitamin D_3 receptor (VDR) and thyroid hormone receptors (TR) may compete for binding to PPRE *cis*-regulatory elements (Rottman and Gordon, 1993) and might be involved in regulation of the zebrafish I-FABP gene.

Further analysis of the sequence 5′ upstream of the transcription start site for zebrafish I-FABP gene using MatInspector professional (http://transfac.gbf.de/cgi-bin/matSearch/matsearch.pl; Quandt *et al.*, 1995) revealed a number of potential *cis*-regulatory elements (Table 2) including six copies of CCAAT/enhancer-binding proteins (C/EBP) binding elements. Lyons *et al.*, (2001) showed that the zebrafish C/EBP are expressed in intestine, liver and yolk syncytial nuclei. The intestine, liver and the yolk syncytial layer of teleost fishes are sites of lipid metabolism. The presence of C/EBP binding elements suggests possible involvement of C/EBP in the transcriptional regulation of zebrafish I-FABP gene and is consistent with the role of C/EBP in lipid metabolism. Furthermore, defects in lipid storage have been observed in C/EBP gene knockout mice (Tanaka *et al.*, 1997). C/EBP are important in the regulation of many

Table 2. Potential *cis*-regulatory elements in the 5′ upstream sequence of the zebrafish I-FABP gene.

Matrix Name	Further Information	Position	Strand	Core Simil.	Matrix Simil.	Sequence
GC_01	GC box elements	–115	+	1.000	0.858	gttgGGCGggtgag
CAAT_01	cellular and viral CCAAT box	–496	+	1.000	0.984	aatagCCAAtga
CEBPB_01	C/EBP binding site	–722	–	1.000	0.912	gtatatgGCAAtgc
CEBPB_01	C/EBP binding site	–587	+	1.000	0.948	tttttgtGCAAtga
CEBPB_01	C/EBP binding site	–481	+	1.000	0.911	gcatgcaGCAAgat
CEBPB_01	C/EBP binding site	–339	+	1.000	0.852	agttactGCAAcac
CEBPB_01	C/EBP binding site	–172	–	1.000	0.948	tgattgtGCAAaga
CEBPB_01	C/EBP binding site	–50	+	1.000	0.890	agataagGCAAcgc
GATA_C	GATA-binding site	–53	+	1.000	0.933	aGATAAtactc
GATA_C	GATA-binding site	–64	+	1.000	0.989	aGATAAggcaa
HFH1_01	HNF-3/Fkh Homolog 1	–409	+	1.000	0.872	aattGTTTaaaa
HFH2_01	HNF-3/Fkh Homolog 2	–284	–	1.000	0.886	aatTGTTtgctg
HFH3_01	HNF-3/Fkh Homolog 3 (= Freac-6)	–275	–	1.000	0.869	tttTGTTtgaatt
HNF3B_01	Hepatocyte Nuclear Factor 3beta	–822	+	1.000	0.901	gtaaaTATTtgagtt
HNF3B_01	Hepatocyte Nuclear Factor 3beta	–599	–	1.000	0.877	aatttTATTtatgta
HNF3B_01	Hepatocyte Nuclear Factor 3beta	–578	–	1.000	0.937	tgaaaTATTtcattg
AP1_Q2	activator protein 1	–543	–	1.000	0.869	gtTGACatgaa
OCT1_06	octamer-binding factor 1	–144	+	1.000	0.876	catgctgggATGTg
OCT1_Q6	octamer factor 1	–853	–	1.000	0.865	taaaatgaAAATttg

genes involved in intermediary lipid metabolism. They are also implicated in the regulation of a number of iLBP genes including I-FABP, Adipocyte-FABP and L-FABP (reviewed in Bernlohr *et al.*, 1997).

The transcription factors GATA-4, GATA-5 and GATA-6 interact with DNA *cis*-elements containing the consensus sequence $^{A}/_{T}GATA^{A}/_{G}$ and are involved in transcriptional activation of intestine-specific genes such as the lactase gene in human Caco-2 cells (Fang *et al.*, 2001) and the gut esterase gene (*ges-1*) in *Caenorhabditis elegans* (Egan *et al.*, 1995). The *Xenopus* I-FABP gene has been identified as an *in vitro* target for GATA-4, -5 and -6 (Gao *et al.*, 1998). Two potential GATA elements are present in the 5′ upstream sequence of the zebrafish I-FABP gene indicating that the GATA family of transcription factors may regulate the zebrafish I-FABP gene transcription in the intestine.

The winged helix transcription factor, hepatic nuclear factor 3 (HNF3), initially identified as a liver-specific transcription factor plays an important role in gut development (Ang and Rossant,1994) and is known to activate virtually all genes expressed in endoderm-derived lineages. Six putative HNF3 binding elements are present in the 5′ upstream sequence of the zebrafish I-FABP. HNF regulates another member of the iLBP multigene family, the rat L-FABP (Simon *et al.*, 1993), which is also expressed in the intestine.

Common promoter elements, including a binding site for AP-1 and two binding sites for Oct-1 factors, occur in the 5′ upstream sequence of zebrafish I-FABP gene. In addition, six copies of the binding element for SRY- related HMG box-5 (SOX-5) factor, expressed in a number of tissues including brain, liver, testis and heart (Wunderle *et al.*,

1996), were present in the zebrafish I-FABP 5′ upstream sequence (data not shown). Analysis involving footprinting and gel retardation assays would be required to access the functionality of these transcription factors in the regulation of I-FABP gene expression in zebrafish.

2.3.5 Tissue-specific expression of the I-FABP mRNA in adult zebrafish

I-FABP mRNA was detected in various tissues of adult zebrafish using RT-PCR. An I-FABP-specific product was generated by RT-PCR from total RNA extracted from zebrafish intestine. RT-PCR product was also generated from liver, brain, testis and muscle (Fig. 4A). No RT-PCR product, however, was detected in cDNA made from total RNA extracted from heart, skin and ovary. The constitutively expressed RACK1 mRNA was amplified by RT-PCR and was used as a positive control for each tissue (Fig. 4A). To estimate the relative levels of I-FABP mRNA in different adult zebrafish tissues and to confirm the tissue distribution of I-FABP mRNA revealed by conventional RT-PCR, quantitative RT-PCR of I-FABP mRNA was performed with the same tissues using another constitutively expressed gene, β-actin, as a positive control. Levels of I-FABP mRNA in each cDNA sample ranged between 1.00 copy to 2.22×10^3 copies per μL of cDNA. The β-actin RT-PCR product was amplified from each cDNA sample and ranged from 1.53×10^2 to 3.53×10^5 copies per μL of cDNA. The ratio of I-FABP mRNA/β-actin RT-PCR product was calculated for each experimental sample (Fig. 4B). The analysis demonstrates that the highest levels of I-FABP mRNA was found in intestine with diminishing steady-state levels of I-FABP mRNA in brain, muscle, testis, heart, liver, skin and the least amount in ovary. The amount of I-FABP mRNA in intestine is $\sim 3.5 \times 10^1$ times higher than in brain and $\sim 16.5 \times 10^1$ times higher in the intestine than in

Figure 4. Detection of I-FABP mRNA in developing zebrafish embryos and in adult tissues by RT-PCR. **(A)** Top panel: zebrafish I-FABP-specific primers produced a product from total RNA extracted from adult zebrafish intestine (I), liver (L), brain (B), muscle (M) and testis (T). No product was detected in adult zebrafish heart (H), skin (S) and ovary (O). A negative control (-), lacking cDNA, did not generate any RT-PCR product. Bottom panel: RT-PCR product was generated from RNA in all samples for the constitutively expressed RACK1. A negative control (-) lacking cDNA did not generate any RT-PCR product. **(B)** In quantitative RT-PCR, zebrafish I-FABP specific primers detected abundant product in RNA extracted from adult zebrafish intestine (I). The I-FABP mRNA was also detected in brain (B), muscle (M), testis (T), heart (H), liver (L), skin (S) and ovary (O). As a negative control, cDNA template was omitted from the RT-PCR reaction. **(C)** Top panel: zebrafish I-FABP-specific primers did not generate a product from total RNA of 1 hpf (1) and 3 hpf (3) embryos. An I-FABP-specific product was observed from total RNA of 12 hpf (12), 24 hpf (24) and 60 hpf (60) embryos and adult (A) zebrafish. A negative control (-) lacking cDNA did not generate any RT-PCR product. Bottom panel: RT-PCR product was generated from RNA in all samples for the constitutively expressed RACK1. A negative control (-) lacking cDNA did not generate any RT-PCR product.

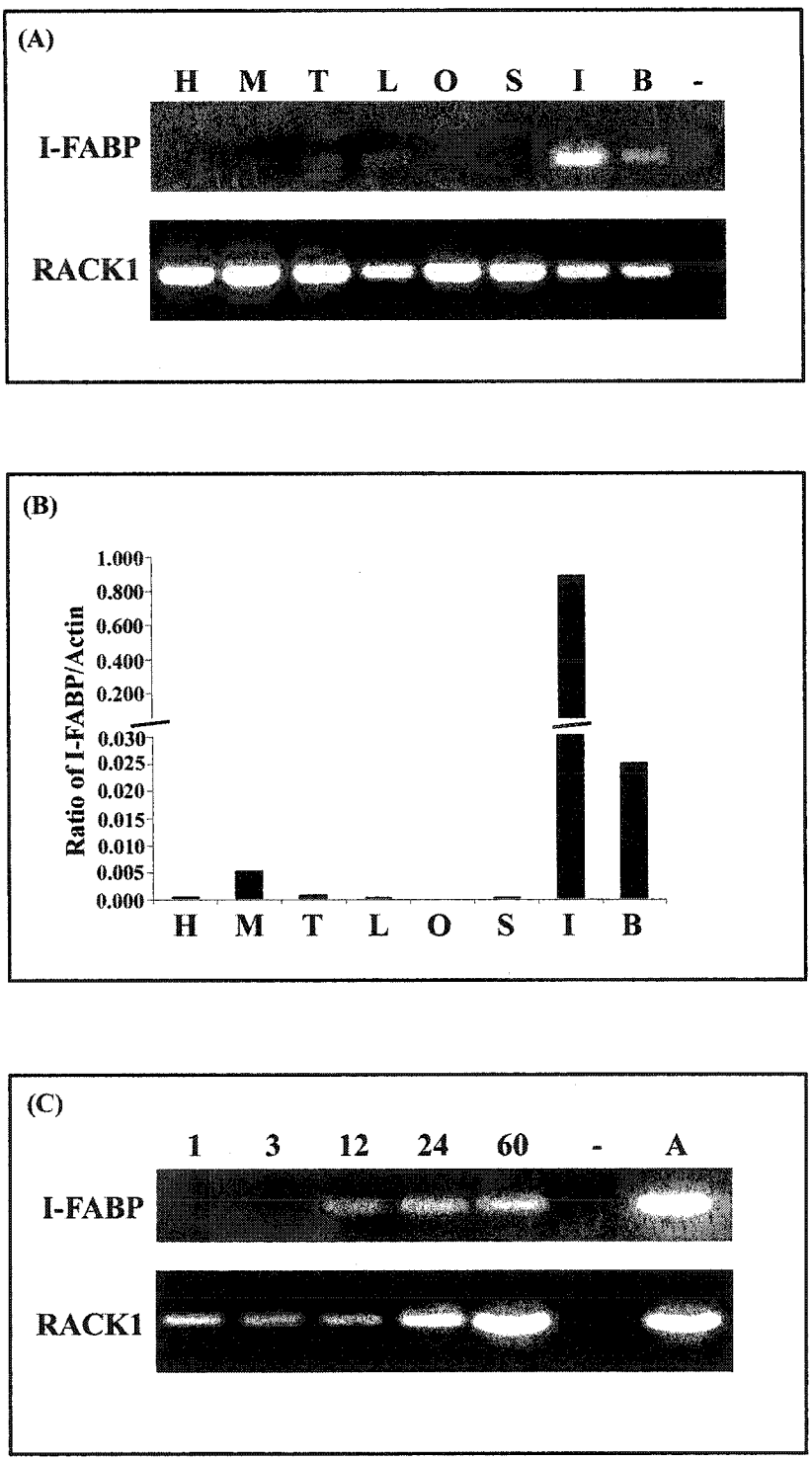


Figure 4. Detection of I-FABP mRNA in developing zebrafish embryos and in adult tissues by RT-PCR.

muscle and between 1×10^3 to 4.2×10^4 times higher in intestine than in testis, heart, liver, skin and ovary. Both conventional RT-PCR and quantitative RT-PCR using different controls, RACK1 and β-actin, showed similar tissue distribution where the zebrafish I-FABP mRNA was abundant. Conventional RT-PCR did not, however, detect I-FABP mRNA in tissues where the levels were low. Previously, I-FABP mRNA has been reported to be present in intestine using *in situ* hybridization analysis to adult zebrafish tissue sections (Pierce *et al.*, 2000; Cameron *et al.*, 2002). As suggested by the results derived from conventional and quantitative RT-PCR, the levels of I-FABP mRNA may be too low to be detected by *in situ* hybridization in tissues other than intestine but was revealed by the more sensitive RT-PCR assay. The highest level of I-FABP mRNA in the intestine could reflect a major role of I-FABP in the intestine but its expression in brain, liver and other tissues suggests additional functions of I-FABP in these tissues.

2.3.6 Temporal and spatial expression of the I-FABP gene during zebrafish embryogenesis

Expression of the I-FABP gene during embryogenesis was examined by both RT-PCR and whole-mount *in situ* hybridization. RT-PCR did not amplify an I-FABP-specific product at 1 hpf and 3 hpf, but did detect it at 12 hpf and stages thereafter (Fig. 4C). Previously, André *et al.* (2000) reported the expression of zebrafish I-FABP mRNA during gut morphogenesis (3 days post-fertilization and stages thereafter) but stages during early embryogenesis were not studied.

Whole-mount *in situ* hybridization to zebrafish embryos detected I-FABP mRNA at the beginning of somitogenesis (11 hpf) in the yolk syncytial layer (YSL) (Fig. 5A).

Figure 5. Detection of I-FABP mRNA by whole-mount *in situ* hybridization during zebrafish embryogenesis. (A) I-FABP mRNA in the yolk syncytial layer (YSL) at the 3 somite stage (11 hpf). (B) 16 somite stage (17 hpf) embryo with I-FABP mRNA expression in YSL. (C) Expression in YSL at 24 hpf embryo. (D) 36 hpf embryo showing decreased level of I-FABP mRNA in YSL as compared to panel (C). (E) Dorsal view of embryo shown in (D) with I-FABP mRNA expression in YSL and the intestinal bulb (IB). (F) 48 hpf embryo showing I-FABP mRNA expression in liver (L), intestine (I) and IB. (G) Dorsal view of embryo shown in (F) showing expression in pancreas (P), L, and IB. (H) Larvae at 5 days of development showing I-FABP mRNA expression restricted to I and IB. Scale bar 50 μM for A, B, E, G and 100 μM for C, D, F and H.

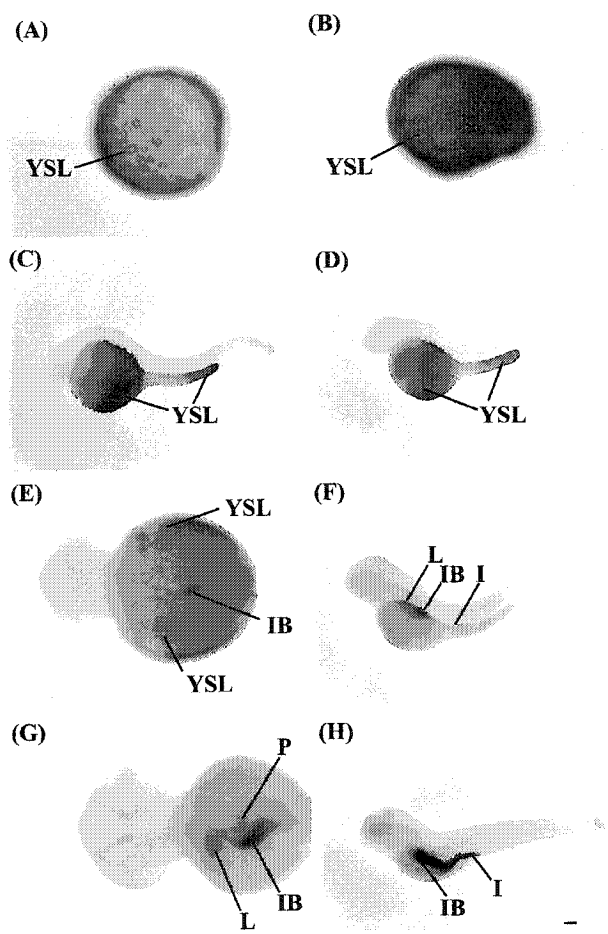


Figure 5. Detection of I-FABP mRNA by whole mount *in situ* hybridization during zebrafish embryogenesis.

much before the initiation of gut morphogenesis (André *et al.*, 2000). The localization of I-FABP mRNA in YSL implies that I-FABP could be involved in the transfer of nutrients from the yolk to the blastoderm. I-FABP transcripts became more abundant in YSL at 17 hpf (Fig. 5B). Expression in the YSL decreased after 24 hpf (Fig. 5C) but persisted at 36 hpf (Fig. 5D). At 36 hpf, I-FABP mRNA was also detected in the intestinal bulb (Fig. 5E). The intestinal bulb continued to express I-FABP mRNA at 48 hpf. At this stage I-FABP mRNA was detected in the liver and the pancreas primordium, but not in the YSL (Fig. 5F, 5G). At 5 days post-fertilization, I-FABP mRNA expression was detectable only in the intestinal bulb and in the anterior part of the intestine (Fig. 5H).

The whole-mount *in situ* hybridization demonstrated that, in addition to the intestine, the I-FABP gene is expressed in the liver and the pancreas during embryogenesis. In the adult zebrafish, I-FABP mRNA was detected by RT-PCR in a number of tissues in addition to intestine. Based on their work with gene knockout mice, Vassileva *et al.* (2000) suggest that the I-FABP is not absolutely required for the fatty acid absorption in the intestine. The expression of I-FABP much before the initiation of first feeding in zebrafish (~72 hpf) (Fig. 4C) and expression of I-FABP mRNA in tissues other than the intestine (Fig. 4 A, 4B, 5F, 5G) indicate a role for I-FABP in functions other than just the assimilation of dietary fatty acids and their intracellular transport.

Chapter 3: A basic liver fatty acid-binding protein from zebrafish: gene structure, radiation hybrid mapping and mRNA distribution during embryonic development and adulthood.

3.1 Introduction

Basic liver fatty acid-binding protein (Lb-FABP/*fabp10*) is a member of the iLBP multigene family that has been found only in non-mammalian species. The first report identifying a Lb-FABP was from the chicken liver (Scapin *et al.*, 1988). Ceciliani *et al.* (1994) later determined the primary structure of this FABP and found that, when compared to other known FABPs, it showed highest sequence similarity to ileal FABPs and not to the mammalian L-FABPs. The chicken L-FABP has an isoelectric point of 9.0, whereas other L-FABPs have an isoelectric point of about 6.0. Lb-FABP has since been found in the liver of catfish (*Rhambia sapo*; Di Pietro *et al.*, 1997), toad (*Bufo arenarum*; Di Pietro *et al.*, 2003), sea bass (*Lateolabrax japonicus*; Odani *et al.*, 2001), iguana (*Anolis pulchellus*; GenBank accession number U28756), axolotl (*Ambistoma mexicanum*; Di Pietro *et al.*, 1999), shark (*Halaetunus bivius*, Córdoba *et al.*, 1999) and lungfish (*Lepidosiren paradoxa*; Di Pietro and Santomé, 2001). Although the presently accepted name is the basic liver fatty acid binding protein, it is worth mentioning that the toad and lungfish Lb-FABP has an isoelectric point of approximately 7.0 (Di Pietro and Santomé, 2001, Di Pietro *et al.*, 2003).

Like L-FABP, the Lb-FABP shows a broad ligand-binding specificity and binds not only long chain fatty acids but a number of structurally different hydrophobic ligands (Di Pietro *et al.*, 1999; Córdoba *et al.*, 1999; Di Pietro and Santomé, 2001). Differences with regard to binding capacity of the Lb-FABPs to the *cis*- and/or *trans*-parinaric acid

have been reported. The catfish Lb-FABP can bind only one *cis*- or *trans*-parinaric acid molecule (Di Pietro *et al.*, 1997), the chicken Lb-FABP binds only one ligand (Schievano *et al.*, 1994; Beringhelli *et al.*, 2001) and the axolotl, shark and lungfish Lb-FABPs bind two molecules of *cis*-parinaric acid, but only one molecule of *trans*-parinaric acid (Di Pietro *et al.*, 1999; Córdoba *et al.*, 1999; Di Pietro and Santomé, 2001).

The isolation of both Lb-FABP and L-FABP from axolotl (Di Pietro *et al.*, 1999) provided evidence that these two liver FABPs are paralogs. Phylogenetic analysis of available iLBPs places Lb-FABP as a separate branch of the iLBP subfamily that includes L-FABP and ileal lipid-binding protein (IlLBP) (Schleicher *et al.*, 1995). The Lb-FABP, L-FABP and IlLBP are estimated to have diverged from their last common ancestral gene approximately 680 mya (Schaap *et al.*, 2002).

The identification and characterization of the zebrafish Lb-FABP gene is presented in this chapter.

3.2 Materials and methods

3.2.1 Zebrafish Lb-FABP cDNA sequence analysis

The zebrafish Lb-FABP cDNA sequence available in our laboratory (Denovan-Wright *et al.*, 2000) was used to deduce the encoded amino acid using Gene Runner V. 3.05 (Hastings Software Inc.). The deduced amino acid sequence was analyzed for isoelectric point and molecular mass. The Lb-FABP and L-FABP sequences obtained from GenBank, National Centre for Biotechnology Information (NCBI) were aligned using CLUSTALW (Thompson *et al.*, 1994).

3.2.2 Organization of the Lb-FABP gene

Zebrafish Lb-FABP cDNA sequence (Denovan-Wright *et al.*, 2000) was used to search the zebrafish genome database as described in section 2.2.2. The sequence of intron 3, not present in the zebrafish genome sequence database, was amplified by PCR with a sense primer specific to exon 3 (5′ GAGATCAAGGCTGGAGAAATG 3′; nt 316 to 336, GenBank accession number AF254642) and an antisense primer specific to exon 4 (5′ AAATAGAGTGATGGTGAAACGC 3′; nt 416 to 395, GenBank accession number AF254642) of the zebrafish Lb-FABP gene. The PCR product was size fractionated by 1% agarose gel-electrophoresis and the single band was excised from the gel. The excised band was purified using QIAquick Gel Extraction kit (QIAGEN Inc.), cloned into the plasmid vector using pGEM-T (Promega) and sequenced.

3.2.3 Lb-FABP gene promoter cloning

The 5′ upstream sequence of the zebrafish Lb-FABP gene was cloned by performing LM-PCR as described in section 2.2.3 with the following exceptions. The first round of PCR was performed with adapter primer C_1 and an external Lb-FABP gene-specific antisense primer (5′ GGCTCTGAGAAACTCCTCGT 3′; nucleotides 75 to 56, GenBank accession number AF254642). The second round of PCR (nested PCR) was performed with adapter primers C_2 and an internal Lb-FABP gene-specific antisense primer (5′ CTCCTGAGCGTAAACCTGCC 3′; nucleotides 51 to 32, GenBank accession number AF254642). The annealing temperature used in primary PCR and the secondary PCR was 61°C and 62°C, respectively. A single product of ~550 bp obtained from genomic DNA digested with *Eco* RI was cloned and sequenced.

3.2.4 Linkage analysis by radiation hybrid mapping

To localize the Lb-FABP gene to a zebrafish linkage group, the LN54 panel of radiation hybrids (zebrafish DNA in a mouse background) (Hukriede *et al.*, 1999) was used. Sense (5′ GCCAGAAGAGGTCATTAAAC 3′; nucleotides 84 to 103, GenBank accession number AF254642) and antisense (5′ TGGTGGTGATTTCAGCCTC 3′; nucleotides 232 to 214, GenBank accession number AF254642) primers were used to PCR-amplify exon 2 of the zebrafish Lb-FABP gene. The PCR conditions were similar to those described in section 2.2.4 with the exception that the annealing temperature used in the PCR was 53 °C.

3.2.5 Mapping the transcription start site of the zebrafish Lb-FABP gene

To map the transcription start site of the zebrafish Lb-FABP gene, 5′ RLM-RACE was employed. The template for 5′ RLM-RACE was prepared as described in section 2.2.5. Nested PCR using two sense adapter primers (section 2.2.5) and two antisense primers specific to Lb-FABP mRNA (outer: 5′ TGGTGGTGATTTCAGCCTC 3′; nucleotides 232 to 214, GenBank accession number AF254642 and inner: 5′ GGCTCTGAGAAACTCCTCGT 3′; nucleotides 75 to 56, GenBank accession number AF254642) was performed. Ten µl of the first round PCR reaction contained 1×PCR buffer, 0.75 U of Taq DNA polymerase (MBI Fermentas), 1.5 mM $MgCl_2$, 0.25 mM each dNTP, 0.5 µM outer adapter primer, 0.5 µM outer L-FABP mRNA-specific primer and 0.5 µl of cDNA. PCR conditions involved initial denaturation at 94°C for 1 min, followed by 30 cycles of 94°C for 30s, 57°C for 30s, 72°C for 40 sec and final extension at 72°C for 10 min. Half a µl of the primary PCR product was used as template for the second round of PCR. The 20 µl reaction mixture of second round PCR contained

1×PCR buffer, 1 U of Taq DNA polymerase (MBI Fermentas), 1.5 mM $MgCl_2$, 0.25 mM each dNTP, 0.25 μM inner adapter primer and 0.25 μM inner L-FABP mRNA-specific primer. The conditions for the secondary PCR were similar to those of the primary PCR with the exception that the annealing temperature was increased to 60 °C and the reaction was subjected to 35 cycles of amplification. The product from the secondary PCR was size fractionated by 1.5 % agarose gel electrophoresis. A single band of ~125 bp in the +TAP reaction was excised from the agarose gel and purified using QIAquick Gel Extraction Kit (QIAGEN Inc.), cloned into plasmid, pGEM-T vector system (Promega) and sequenced.

3.2.6 RNA extraction and northern blotting

Total RNA was extracted from adult zebrafish using Trizol® (Gibco BRL) according to the manufacturer's instructions. Fifteen μg of total RNA 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 0.2 M formaldehyde. The resolved RNA was transferred to Hybond-N^+ membrane according to the supplier's recommendations (Amersham Pharmacia Biotech. Inc.). An antisense oligonucleotide specific to Lb-FABP mRNA (5′ ATAGAGTGATGGTGAAA CGCTTCAGATCTTCTTGC 3′; nucleotides 414 to 379, GenBank accession number AF254642) was 5′ end labelled with γ-[^{32}P]-ATP. Ten μL of labelling reaction contained 1×One-Phor-All buffer (Amersham Pharmacia Biotech. Inc.), 6 U of T4 Polynucleotide Kinase (MBI Fermentas), 30 μCi of γ-[^{32}P]-ATP (3000 ci/mmol; Amersham Pharmacia Biotech. Inc.) and 1 μM antisense oligonucleotide. The 5′-end labelling was performed at 37 °C for 45 min followed by a 10 min incubation at 68 °C to inactivate T4

Polynucleotide Kinase. After labelling, the unincorporated radionucleotides were removed using ProbeQuant™ G-50 micro columns (Amersham Pharmacia Biotech Inc.) according to the manufacturer's instructions. The membrane was prehybridized in 10 ml of hybridization solution (5×SSPE, 5×Denhardt's solution, 0.5% (w/v) SDS) and 21 μg/ml sonicated salmon sperm DNA. Prehybridization was carried out at 55°C for 3 h followed by hybridization with the purified probe in 7 ml of fresh hybridization solution at 57°C overnight. The membrane was washed once at room temperature for 10 min in 2×SSPE, 0.1% SDS and exposed to X-ray film at –70°C for three weeks.

3.2.7 Developmental and tissue-specific expression of Lb-FABP mRNA in zebrafish

Tissue-specific expression of zebrafish Lb-FABP mRNA was defined by RT-PCR as described in section 2.2.6. Developmental expression was determined by both RT-PCR and whole mount *in situ* hybridization as described in section 2.2.6. The antisense RNA probe synthesised using Lb-FABP cDNA (GenBank accession number AF254642) as template was used in whole mount *in situ* hybridization. The Lb-FABP mRNA-specific primers used in the RT-PCR were: sense: 5′ TTACGCTCAGGAGAACTACG 3′ (nt 39 to 58, GenBank accession number AF254642) and antisense: 5′ CTTCCTGATCA TGGTGGTTC 3′ (nt 378 to 358, GenBank accession number AF254642).

3.3 Results and discussion

3.3.1 cDNA sequence analysis

The zebrafish Lb-FABP cDNA sequence (Denovan-Wright *et al.*, 2000) encodes a peptide of 126 amino acids with a calculated isoelectric point of 8.8 and a molecular

weight of 14.0 kDa. The deduced amino acid sequence of Lb-FABP was aligned with the Lb-FABP and L-FABP sequences from several other species (Fig. 6). Zebrafish Lb-FABP showed 85% identity to catfish Lb-FABP and between 60% to 69% identity to Lb-FABP from salamander, chicken, lungfish, iguana, shark and toad. The identity of the zebrafish Lb-FABP to L-FABPs was in the range of 38% to 42% (Fig. 6). The alignment revealed that the zebrafish Lb-FABP is most similar to Lb-FABPs and not the L-FABPs reported from other species.

3.3.2 Organization of the zebrafish Lb-FABP gene

DNA traces with identity to a zebrafish Lb-FABP cDNA (Denovan-Wright *et al.*, 2000) were retrieved from the Wellcome Trust Sanger Institute (http://trace.ensembl.org/perl/ssahaview and http://www.ensembl.org/Danio_rerio/ssahaview). Trace zfishG-a2428c01.q1ca contained the sequence for exon 1 and part of intron 1, trace zfishC-a924e02.p1c contained part of intron 1, exon 2 and part of intron 2, sequence z06s041263 contained part of exon 2 and part of intron 2, trace zfishC-a924e02.q1c contained part of intron 2 and the trace zfishC-a2463a03.p1c contained the sequence for part of intron 2, exon 3 and part of intron 3. The sequence of intron 3, not present in the zebrafish genome sequence database was obtained from genomic DNA by PCR. Comparison of the genomic and cDNA sequence revealed that zebrafish Lb-FABP gene is organized into 4 exons separated by 3 introns (Fig. 6). The intron/exon splice junctions follow the "GT-AG" dinucleotide rule of Breathnach and Chambon (1981). The four exon/three intron organization of the Lb-FABP gene is consistent with the structure of other mammalian iLBP genes (Bernlohr *et al.*, 1997) and provides additional evidence that Lb-FABP gene

Figure 6. Sequence alignment of the zebrafish Lb-FABP with Lb-FABP and L-FABP from other species. The amino acid sequences of zebrafish Lb-FABP (zf-Lb-FABP; GenBank accession number AF254642), catfish Lb-FABP (Cf-Lb-FABP; P80856), salamander Lb-FABP (Sa-Lb-FABP; P81400), chicken Lb-FABP (Ch-Lb-FABP; P80226), iguana Lb-FABP (Ap-Lb-FABP; U28756), shark Lb-FABP (Sh-Lb-FABP;P81653), toad Lb-FABP (Td-Lb-FABP; P83409), lungfish Lb-FABP (Lf-Lb-FABP; P82289), pufferfish L-FABP (Pf-L-FABP; AAC60290), salamander L-FABP (Sa-L-FABP; P81399) pig L-FABP (Pi-L-FABP; P49924), human L-FABP (Hu-L-FABP; P07148), mouse L-FABP (Mo-L-FABP; Y14660), cow L-FABP (Bo-L-FABP; P80425) and rat L-FABP (Ra-L-FABP; P02692) were aligned using CLUSTALW (Thompson *et al.*, 1994). Dots indicate identity and dashes were introduced to maximize alignment. The percent identity between the zebrafish Lb-FABP and other Lb-FABPs or L-FABPs are shown at the end of each sequence. The intron positions relative to the amino acid sequence are indicated by "▼".

```
Zf-Lb-FABP 1 MAFSGTWQVYAQENYEEFLRAI▼SLPEEVIKLAKDVKPVTEIQQNGSDF
Cf-Lb-FABP 1 -.....  ....D.............T.N..
Sa-Lb-FABP 1 -P.N.....S.....A....V  G...DI.NV...IN.II......DN.
Ch-Lb-FABP 1 -.................K.L  A...DL..M.R.I..IV....K.D..
Ap-Lb-FABP 1 ...N.....S.....D..K..  A..DDI..A.........R.T.NT.
Sh-Lb-FABP 1 -.........S...I.D....L  ........IG..I...ID.K.T.EH.
Td-Lb-FABP 1 -..N...N........N...TV  G...DI..V....N..I..E...NE.
Lf-Lb-FABP 1 -..............A..KV.  GVA.DI.PH..EI..TI....S.NS.
Pf-L-FABP  1 .S...KY.QVS...F.P.MK..  G..D...QQV.EL.STS..E...N..
Sa-L-FABP  1 -S.A.KYELQS...F.A.MK..  G..D.L.QKG..I.S.S......KS.
Pi-L-FABP  1 .N...KY..QS...F.A.MK.V  G..D.L.QKG..I.GTS..V...KH.
Hu-L-FABP  1 .S...KY.LQS...F.A.MK..  G....L.QKG..I.G.S..V...KH.
Mo-L-FABP  1 .N...KY.LQS...F.P.MK..  G...DL.QKG..I.G.S..VHE.KKI
Bo-L-FABP  1 .N...KY..QT.....A.MK.V  GM.DDI.QKG..I.G.S..V...KH.
Ra-L-FABP  1 .N...KY..QS...F.P.MK.M  G...DL.QKG..I.G.S..VHE.KKV

Zf-Lb-FABP 50 TITSKTPGKTVTNSFTIGKEAEITTMDGKKLK▼CIVKLDG-GKLVCRT
Cf-Lb-FABP 49 V........S.................R...  .....E.-...ISE.
Sa-Lb-FABP 49 VV.....NQS.............S.G...I.  .T.V.E.-....SK.
Ch-Lb-FABP 49 VV.....RQ.......L....D.........  .T.H.AN-....TKS
Ap-Lb-FABP 50 VV.....N.S......L....DM.......V.  .T.N.VD-....AKS
Sh-Lb-FABP 49 V.VV..SQQ....E..V.......S.......  .T.Q.ED-....AKK
Td-Lb-FABP 49 VV.....KQ.HS....V...S...S.....I.  VT.Q.E.-...I.KS
Lf-Lb-FABP 49 .V..TAQK.ST..T.............N.N..R  .TINMED-.....K.
Pf-L-FABP  50 K..IT.GP.VTV.K......T.MD.IT.E.I.  TVFH...-N..KVSL
Sa-L-FABP  49 KV.VT.GS.VLE.E..L.E...LE.LT.E.V.  S...QE.DN...VNL
Pi-L-FABP  50 KL.IT.GS.V.Q.E..L.E.C.ME.LT.E.V.  TV.Q.E.DN...TTF
Hu-L-FABP  50 KF.ITAGS.VIQ.E..V.E.C.LE..T.E.V.  TV.Q.E.DN...TTF
Mo-L-FABP  50 KL.ITYGP.V.R.E..L.E.C.LE..T.E.V.  AV...E.DN.M.TTF
Bo-L-FABP  50 KFIITAGS.VIQ.E..L.E.C.MEF.T.E.I.  AV.QQE.DN...TTF
Ra-L-FABP  50 KL.ITYGS.VIH.E..L.E.C.LE..T.E.V.  AV..ME.DN.M.TTF

Zf-Lb-FABP 96 DRFSHIQEIKAG▼EMVETLTVGGTTMIRKSKKI-     100%
Cf-Lb-FABP 95 EK...K....G.  ..I.....A....V.....V-      85%
Sa-Lb-FABP 95 .Q......V.GN  ..........A.L..R..RV-      69%
Ch-Lb-FABP 95 EK...E..V.GN  .....I.F..V.L..R..RV-      69%
Ap-Lb-FABP 96 .K.I.E...VGN  .....I.S.SA.FT.R....-      67%
Sh-Lb-FABP 95 LK.T....VQGN  ..I.K..A.NA......RRM-      64%
Td-Lb-FABP 95 .K......VNGD  ....KI.I.SS.LT....RV-      64%
Lf-Lb-FABP 95 EK......VQGE  ..I....S.SA.L..R.R.V-      60%
Pf-L-FABP  95 KGIESVT.LA-D  PNTI.M.L.DVVYKTT..RM-      42%
Sa-L-FABP  96 KGITSVT.LSGD  TLIN..QK.DD.YK.I..R.-      42%
Pi-L-FABP  97 KGIKSVT.LNGD  IITS.M.L.DIVFK.I..R.-      42%
Hu-L-FABP  97 KNIKSVT.LNGD  IITN.M.L.DIVFK.I..R.-      42%
Mo-L-FABP  97 KGIKSVT.LNGD  TITN.M.L.DIVYK.V..R.-      40%
Bo-L-FABP  97 KGIKSVT.FNGD  TVTS.M.K.DVVFK.V..R.-      38%
Ra-L-FABP  97 KGIKSVT.FNGD  TITN.M.L.DIVYK.V..R.-      38%
```

Figure 6. Sequence alignment of the zebrafish Lb-FABP with Lb-FABP and L- FABP from other species.

and other members of the iLBP multigene family originated from a common ancestral gene as a result of duplication event (s) during the course of evolution.

3.3.3 Radiation hybrid mapping of the zebrafish Lb-FABP gene

The LN54 panel of radiation hybrids (Hukriede *et al.*, 1999) was used to assign zebrafish Lb-FABP gene to LG 16 at a distance of 8.23 cR from the marker Z10256 with a LOD of 17.1.

Although no mammalian Lb-FABP orthologs are reported, the comparison of a number of genes mapped to zebrafish LG 16 and their mammalian orthologs suggest that the zebrafish LG 16 has syntenic relationship with human chromosomes 1, 3, 6, 17, 19 and 20 (Table 3). The mouse chromosome 3 and rat chromosomes 1 and 2 are syntenic to zebrafish LG 16 (Table 3).

3.3.4 5′ upstream sequence and the transcription start site of the zebrafish Lb-FABP gene

A 484 bp genomic DNA fragment 5′ upstream of the initiation codon for the Lb-FABP gene was cloned by LM-PCR following digestion of genomic DNA with the restriction enzyme *Eco* RI and sequenced. The transcription start site of the zebrafish Lb-FABP gene was mapped by amplifying the capped and mature 5′ end of the Lb-FABP mRNA transcript from adult zebrafish RNA using 5′ RLM-RACE. Nested PCR produced a single band ~ 125 bp in size (Fig. 7A) that was isolated and purified from the agarose gel. Sequencing of a randomly selected clone revealed that the cDNA end was 90 bp in length excluding the 38 bp 5′ RLM-RACE adapter sequence. Alignment of the 5′ RLM-RACE sequence and the 5′ upstream sequence cloned using LM- PCR revealed

Table 3. Conserved syntenies between zebrafish LG 16 and mammalian chromosomes.

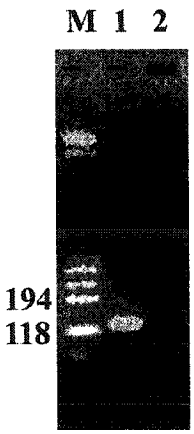
Zebrafish[1]			Human[2]		Mouse[2]		Rat[2]	
LG	**Locus**	**Accession Number**	**Gene**	**Chromosomal Position**	**Gene**	**Chromosomal Position**	**Gene**	**Chromosomal Position**
16	Lb-FABP/fabp10	AF254642	-	-	-	-	-	-
16	pklr	AW171121	PKLR	1q21	Pklr	3 41.5 cM	Pklr	2q34
16	ctss	AA549805	CTSS	1q21	Ctss	3 42.7 cM	Ctss	2q34
16	cct3	AI476862	CCT3	1q23	Cct3	3 50.0 cM	Cct3	-
16	kiaa0907	AI415764	KIAA0907	1q22	Kiaa0907	-	Kiaa0907	-
16	hypb	AI588663	HYPB	3p21.31	Hypb	-	Hybp	-
16	ctnnb	U41081	CTNNB1	3p22-p21.3	Catnb	9 72.0 cM	Catnb	-
16	vegf	AF016244	VEGF	6p12	Vegf	-	Vegf	9q12
16	pou47	Y07905	POU3F2	6q16	Pou3f2	4 6.3 cM	Pou3f2	5q21
16	rpl38	AI384538	RPL38	17q23-q25	Rpl38	-	Rpl38	-
16	ngfr	AI629342	NGFR	17q21-q22	Ngfr	11 55.6 cM	Ngfr	10q31
16	gsk3a	AJ223501	GSK3A	19q13.31	Gsk3a	-	Gsk3a	1q21
16	dnase2	AI330733	DNASE2	19p13.2	Dnase2a	8 38.6 cM	Dnase2	19q11
16	ceacam1	AI353126	CEACAM1	19q13.2	Ceacam1	7 5.5 cM	Ceacam1	1q21-q22
16	rps9	AI353657	RPS9	19q13.4	Rps9	-	Rps9	1q12
16	ck2a1	S76875	CSNK2A1	20p13	Csnk2a1	1 G3	Csnk2a1	3q41
16	rpn2	AI496816	RPN2	20q12-q13.1	Rpn2	2 91.0 cM	Rpn2	3q42

[1]Woods *et al.*, 2000.

[2] LocusLink (http://www.ncbi.nlm.nih.gov/LocusLink/list.cgi), NCBI.

Figure 7. The 5′ end of the mature zebrafish Lb-FABP mRNA and 5′ upstream sequence of the zebrafish Lb-FABP gene. **(A)** Total RNA from whole adult zebrafish was treated with calf intestinal alkaline phosphatase (CIP), tobacco acid pyrophosphatase (TAP) and ligated to a designated RNA adapter. A single product of ~125 bp (lane 1) was amplified by two rounds of nested PCR. The RNA treated with CIP, but not TAP and ligated to the designated RNA adapter did not generate a product in nested PCR (lane 2). The molecular weight marker (lane M) used was ΦX174 RF DNA/*HaeIII* fragment (Gibco BRL). The sizes of the molecular weight markers in bp are indicated to the left. **(B)** A portion of the Lb-FABP 5′ upstream sequence was cloned using LM-PCR. The internal Lb-FABP-specific primer used in the LM-PCR is underlined and the initiation codon is in bold font. The transcription start site identified using 5′ RLM-RACE is indicated (*). The sequence 5′ upstream of the transcription start site is in lower case. A putative TATA box is italicized and underlined.

(A)



(B)

5′gaattctacactatctcttttgatacatttaattgatttaaaa
aaagcagttcacccaagaaacatttcctcacactcgagtggttgt
cttgaaaacttttatgaattactttcacaaaacaacgtatttgga
agaatgttggaaaaaagcagccattgacttccatagtaacaacaa
aaaaaatactatggaagtcaatggctgtttttttcaccattcggta
tcttcattctggagcagaatttttgggtgatctgtccctttaag
tcgtcaaatcctggtgcaatattccacatgcactgatgagtcttt
atttttggtctgctactgctgtgcatgtgggggcatttacctcat
cttctgctggagttgatgaacggtgggttgttcaaacagcagcag
gtcattgactgaactcctctcga*tataaa*agctgcagatctgaag
TATA box
ctgaccttcACTTTGTGTTGAGCTTCTCCAGAAAGC**ATG**GCCTTC
AGCGGGACGTGGCAGGTTTACGCTCAGGAGAACTACGAGGAGTTT
CTCAGAGCC 3′

Figure 7. The 5′ end of the mature zebrafish Lb-FABP mRNA and the 5′ upstream sequence of the zebrafish Lb-FABP gene.

that the transcription start site of the zebrafish Lb-FABP was located 27 bp upstream of the initiation codon (Fig. 7B).

A 435 bp sequence in the Lb-FABP promoter is sufficient for the liver-specific gene expression (Her *et al.*, 2003b). Based on the transcription start site identified here, the 435 bp sequence is located between -1482 to -1916 in the 5′ upstream sequence of the zebrafish Lb-FABP gene.

3.3.5 Northern blot hybridization

The antisense oligonucleotide specific to Lb-FABP mRNA hybridized to a transcript of approximately 700 nucleotides in the northern blot (Fig. 8). Considering the average size of a poly(A) tail (150-200 nucleotides) in eukaryotic mRNAs, the difference in length between the cDNA sequence (442 bp plus a poly(A) tail of at least 91 nucleotides, GenBank accession number AF254642, Denovan-Wright *et al.*, 2000) and the detected mRNA of 700 nucleotides is likely due to the poly(A) tail. Comparison of the sequence obtained by 5′ RLM-RACE (section 3.3.4) and the Lb-FABP cDNA sequence (GenBank accession number AF254642) revealed an additional 15 bp sequence that was not present in the original cDNA clone. Probing of the northern blot demonstrated that a single mRNA transcript is encoded by the zebrafish Lb-FABP gene.

3.3.6 Temporal and spatial expression of the Lb-FABP gene during zebrafish embryonic and larval development

RT-PCR of total RNA extracted from zebrafish embryos at different postfertilization stages revealed the temporal expression of the Lb-FABP gene during

Figure 8. Northern blot and hybridization to detect Lb-FABP mRNA. Total RNA isolated from adult zebrafish was analysed by northern blot analysis. The antisense Lb-FABP specific oligonucleotide probe (1) hybridized to a mRNA of approximately 700 nucleotides. Following removal of the radiolabelled probe, the blot in 1 was hybridized to the sense oligonucleotide probe (2) that did not detect any zebrafish RNA under the hybridization conditions employed.

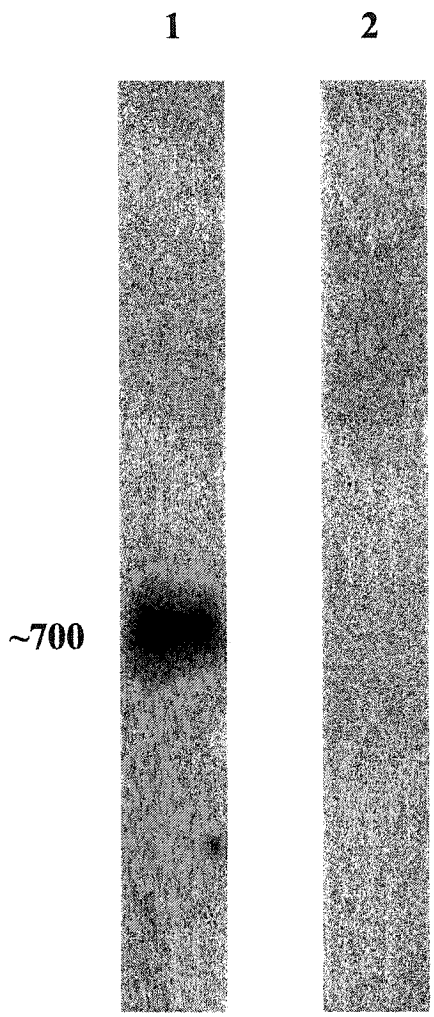


Figure 8. Northern blot and hybridization to detect Lb-FABP mRNA.

embryogenesis. Lb-FABP-specific product was amplified from total RNA extracted from zebrafish embryos at 60 and 72 hpf (Fig. 9A). No product was detected from zebrafish embryos between 1, 3, 12 and 24 hpf or in the negative control reactions (Fig. 9A). As a positive control, RT-PCR was performed to amplify the constitutively expressed zebrafish RACK1 (Hamilton and Wright, 1999).

An antisense RNA probe prepared from the zebrafish Lb-FABP cDNA did not detect a Lb-FABP-specific product in the embryos from the gastrula (approximately 6 hpf) until 36 hpf (data not shown). This suggests that Lb-FABP does not play an important role in the early stages of liver morphogenesis, at least not for a few hours after the developing hepatocytes have aggregated (~28 hpf, Ober *et al.*, 2003). The weak expression of the Lb-FABP mRNA in the ventral endoderm (near the heart chamber) of the 36 hpf zebrafish embryos (Her *et al.*, 2003a) was not observed in the present investigation and could represent a strain-specific difference. At 48 hpf, the approximate time when the liver budding process is complete (Ober *et al.*, 2003), Lb-FABP transcripts were detected in the liver of the zebrafish embryos (Fig. 10A, B, C). Lb-FABP mRNA was also detected in 5 day old larvae (Fig. 10D, E).

3.3.7 Tissue-specific expression of the Lb-FABP mRNA in adult zebrafish

RT-PCR was used to define the Lb-FABP mRNA expression in various adult zebrafish tissues. The constitutively expressed RACK1 employed as positive control, generated a product for each tissue sample analyzed by RT-PCR (Fig. 9B). A Lb-FABP-specific product was generated from total RNA extracted from liver, intestine and testes of adult zebrafish (Fig. 9B). The total RNA extracted from brain, muscle, heart, ovary and skin did not generate a Lb-FABP-specific product in the RT-PCR (Fig. 9B). We

Figure 9. RT-PCR analysis to determine the temporal and tissue-specific expression of zebrafish Lb-FABP gene. **(A)** Zebrafish Lb-FABP mRNA-specific primers did not generate a product from total RNA extracted from zebrafish embryos at 1 hpf (1), 3hpf (3), 12 hpf (12) and 24 hpf (24) stages. Lb-FABP specific product was generated from the total RNA extracted from zebrafish embryos at 60 hpf (60) stage, 72 hpf (72) stage and the total RNA extracted from whole adult zebrafish (A). No product was detected in the negative control (-) lacking RNA template in the RT-PCR (upper panel). A product for constitutively expressed RACK1 mRNA was detected in total RNA extracted from all embryonic stages (lower panel). **(B)** RT-PCR generated a Lb-FABP mRNA-specific product from total RNA extracted from adult zebrafish liver (L), intestine (I) and testes (T). No Lb-FABP mRNA product was generated from adult zebrafish heart (H), muscle (M), ovary (O), skin (S), brain (B) or the negative control (-) lacking the total RNA in the RT-PCR (upper panel). A RACK1 mRNA-specific product was generated from all the adult zebrafish tissues analyzed (lower panel).

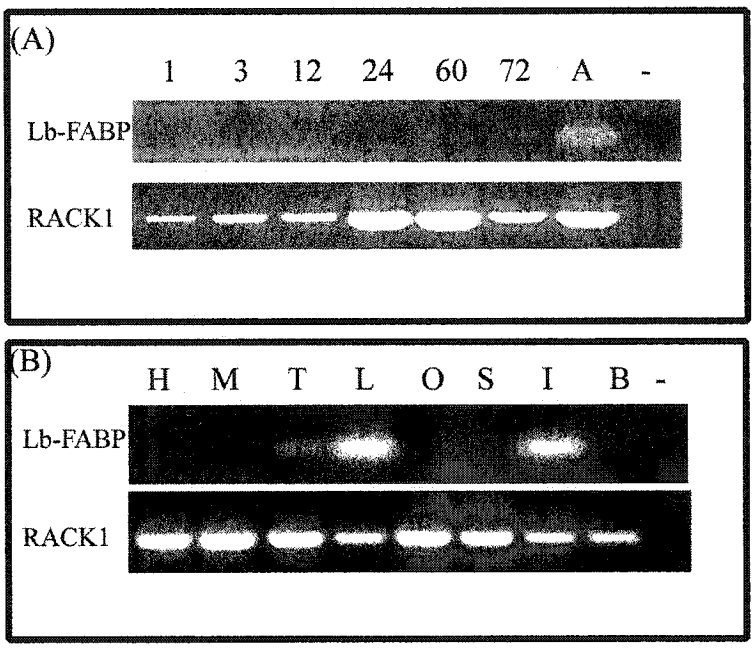


Figure 9. RT-PCR analysis to determine the temporal and tissue-specific expression of zebrafish Lb-FABP gene.

Figure 10. Zebrafish Lb-FABP mRNA detection by whole mount *in situ* hybridization during embryonic and larval development. (A) Lb-FABP mRNA in the liver (L) of the 48 hpf zebrafish embryo. (B) Dorsal view of 48 hpf embryo showing expression in the liver. (C) Side view of embryo shown in (B). (D) Expression of Lb-FABP in the liver of 5 day old larvae. (E) Dorsal view of larvae at 5 days of development showing Lb-FABP mRNA expression in the liver.

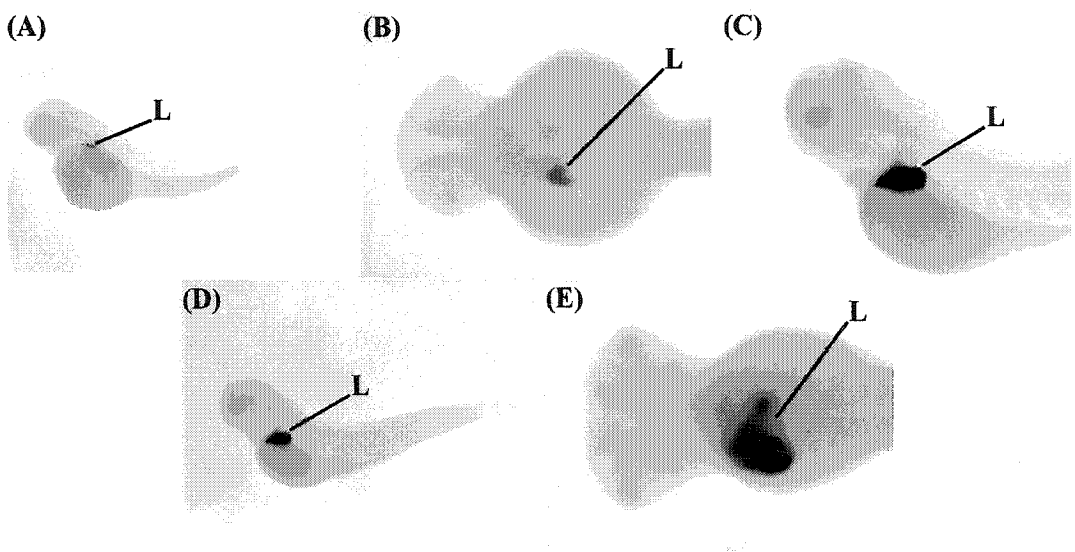


Figure 10. Zebrafish Lb-FABP mRNA detection by whole mount *in situ* hybridization during embryonic and larval development.

have previously detected Lb-FABP mRNA in the liver of adult zebrafish by tissue-section *in situ* hybridization (Denovan-Wright *et al.*, 2000). The highly sensitive RT-PCR assay, however, also detected Lb-FABP mRNA in the intestine and testes. L-FABP, a paralogous gene that is phylogenetically closely related to Lb-FABP, is believed to have duplicated from a common ancestral gene approximately 680 mya (Schaap *et al.*, 2002) and is expressed in the intestine and liver of adult rats (Sweetser *et al.*, 1986). The Lb-FABP gene has not been reported for any mammalian species, however, both L-FABP and Lb-FABP are expressed in a few non-mammalian species including chicken, salamander and catfish. The expression pattern of the paralogous L-FABP gene in zebrafish will further help elucidate the division of expression pattern and possible mechanism of its retention in zebrafish lineage after the duplication event. The zebrafish L-FABP gene structure, mRNA expression pattern, linkage analysis and potential *cis*-acting regulatory elements are described in chapter 4.

Chapter 4: An ortholog of mammalian liver fatty acid-binding protein from zebrafish: cDNA cloning, gene structure, linkage analysis and mRNA distribution.

4.1 Introduction

Liver fatty acid-binding protein (L-FABP/*fabp1*) was first isolated from the cytosol of rat liver (Ockner *et al.*, 1972) and has since been identified in the intestine and kidney of mammals (Hertzel and Bernlohr, 2000). It has been proposed that L-FABP is involved in uptake of fatty acids (Hertzel and Bernlohr, 2000), modulation of enzyme activity by altering lipid levels (Coe and Bernlohr, 1998), protection of cells against the harmful effects of excess free fatty acids by acting as a buffer (Besnard *et al.*, 2002), regulation of the expression of specific genes and regulation of cell growth and differentiation (Veerkamp and Maatman, 1995). In addition to long chain fatty acids, L-FABP binds lysophospholipids, prostaglandins, phytanic acid, eicosanoids, heme and acyl-CoAs (reviewed in: Coe and Bernlohr, 1998; Zimmerman and Veerkamp, 2002, Wolfrum *et al.*, 1999). Several studies have shown that L-FABP binds two fatty acids per molecule (Thompson *et al.*, 1997b; Di Pietro *et al.*, 1999; Nemecz *et al.*, 1991) whereas other iLBPs have the capacity to bind a single ligand (Richieri *et al.*, 1994; Nemecz *et al.*, 1991). The exception to this is the basic liver fatty-acid binding proteins from axolotl (Di Pietro *et al.*, 1999), shark (Córdoba *et al.*, 1999), lungfish (Di Pietro *et al.*, 2001) and toad (Di Pietro *et al.*, 2003) that bind two molecules of *cis*-parinaric acid but not *trans*-parinaric acid. The L-FABP binds the broadest range of ligands among iLBPs studied.

Di Pietro and co-workers (1997) were the first to report the presence of L-FABP in a fish species. In western blot analysis, antibodies to rat L-FABP were used to demonstrate the restricted expression of the catfish L-FABP in the intestine (Di Pietro *et al.*, 1997). Recently, a conceptual translation of the pufferfish L-FABP from the genomic sequence has been submitted to the GenBank (accession number AAC60290). To date, these are the only two reports available on a piscine L-FABP. In this chapter, I discuss the cloning of the cDNA for the zebrafish L-FABP and examine L-FABP mRNA expression during embryogenesis and in a number of adult zebrafish tissues. Sequence analysis, gene organization and linkage analysis of zebrafish L-FABP are presented. The putative *cis*-acting regulatory elements of the zebrafish L-FABP gene are discussed in light of the observation that L-FABP/*fabp1* is expressed in adult zebrafish intestine but not in the liver.

4.2 Materials and methods

4.2.1 5' RLM-RACE and 3' RACE to obtain L-FABP cDNA sequence

Database searches at the NCBI (http://www.ncbi.nlm.nih.gov/) identified a zebrafish EST sequence (GenBank accession number BI846703) described as similar to human L-FABP. The 3' Rapid Amplification of cDNA Ends (3' RACE) and 5' RLM-RACE were performed to obtain the zebrafish L-FABP cDNA sequence. One μg of total RNA from adult zebrafish was used to synthesize first strand cDNA using reverse transcriptase Superscript II (Gibco BRL) and an oligo dT primer with an adapter (5' GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTT 3'). Sense primer (5' GGGATCTCCTGAAGCTGAAC 3'; nt 9 to 28, Fig. 12) specific to L-FABP mRNA was used in the first round of PCR along with the adapter primer (5' GGCCACGCGTCGACT

AGTAC 3') to amplify the 3' end of the cDNA. The 25 µl reaction contained 1×PCR buffer, 1.25 U of Taq DNA polymerase (MBI Fermentas), 1.5 mM $MgCl_2$, 0.2 mM each dNTP, 0.4 µM each primer and 1 µl of cDNA. PCR conditions involved initial denaturation at 94°C for 2 min, followed by 32 cycles of 94°C for 30s, 58°C for 30s, 72°C for 40s and final extension at 72°C for 5 min. Half a µl of the primary PCR product was used as template for the second round of PCR with a L-FABP mRNA-specific internal sense primer (5' TGGGAAATATCAGCTGGAGTC 3'; nt 69 to 89, Fig. 12) and the adapter primer. The conditions for the secondary PCR were similar to those of the primary PCR with the exception that the reaction was subjected to 30 cycles of amplification. The product from the secondary PCR was size fractionated by 1% agarose gel electrophoresis. A single band of ~800 bp was excised from the agarose gel and purified using QIAquick Gel Extraction Kit (QIAGEN), cloned into plasmid pGEM-T vector system (Promega) and sequenced.

5′ RLM-RACE was performed to amplify the 5′ end of the L-FABP mRNA and to map the transcription start site of the zebrafish L-FABP gene. The template for 5′ RLM-RACE was prepared as described in section 2.2.5. Nested PCR using two sense adapter primers (section 2.2.5) and two antisense primers specific to L-FABP mRNA (outer: 5′ CGTCTGCTGATCCTCTTGTAG 3′; nt 431 to 411, Fig. 12 and inner: 5′ CGACCTCATCATCCGGCAC 3′; nt 145 to 127, Fig. 12) was performed. The PCR conditions were similar to those described in section 3.2.5 with the exception that 58°C was used as the annealing temperature in both primary and secondary PCR and the extension time was increased to 1 min. The product from the secondary PCR was size fractionated by 1.5 % agarose gel electrophoresis. A single major band of ~200 bp in the

+TAP reaction was excised from the agarose gel and purified using QIAquick Gel Extraction Kit (QIAGEN), cloned into plasmid pGEM-T vector system (Promega) and sequenced.

The complete cDNA sequence was determined by aligning the overlapping sequences obtained from 3′ RACE and 5′ RLM-RACE using CLUSTALW (Thompson *et al.*, 1994). A BLASTX search of the cDNA sequence was performed at GenBank (NCBI). The cDNA sequence was analyzed for the open-reading frames, protein molecular mass and isoelectric point using Gene Runner V 3.05 (Hastings Software, Inc.). The deduced amino acid sequence of the open-reading frame was aligned with other reported sequences using CLUSTALW (Thompson *et al.*, 1994) and an output of percentage sequence identity was generated. CLUSTALX (Thompson *et al.* 1997a) was used to generate a bootstrap neighbour-joining phylogenetic tree to test the position of zebrafish L-FABP among different iLBP sequences obtained from GenBank (NCBI). Human Von Ebner's gland protein (LCN1, GenBank accession number NP_002288) that belongs to the lipocalin family of the calycins was used as outgroup.

4.2.2 Sequence of the L-FABP gene and the 5' upstream sequence

Partial zebrafish L-FABP gene sequence including 1287 bp of 5′ upstream sequence was obtained from the zebrafish genome database at the Wellcome Trust Sanger Institute (http://www.sanger.ac.uk/Projects/ D_rerio/). The L-FABP gene sequence was aligned with the cDNA sequence to determine the intron/exon junctions. The 1287 bp 5' upstream sequence of the L-FABP gene was analysed for potential *cis*-acting regulatory elements by computer-assisted analysis using MatInspector professional at http://transfac.gbf.de/cgi-bin/matSearch/matsearch.pl (Quandt *et al.*, 1995).

4.2.3 Linkage analysis by radiation hybrid mapping

LN54 radiation hybrid panel (Hukriede *et al.*, 1999) was used to localize the L-FABP gene to a zebrafish linkage group. Radiation hybrid mapping was performed as described in section 2.2.4 with the following exceptions: Sense (5′ AAGCAGGAAGTT CTCATCGG 3′) and antisense (5′ AATCCCCTTGACAA ACGCTG 3′) primers were used to amplify part of intron 2, exon 3, intron 3, and part of exon 4 of the zebrafish L-FABP gene and 55°C was used as annealing temperature in the PCR.

4.2.4 Developmental and tissue-specific expression of L-FABP mRNA in zebrafish

Developmental and tissue-specific expression of the zebrafish L-FABP gene was performed by RT-PCR and whole-mount *in situ* hybridization as described in section 2.2.5. RT-PCR was performed as described in section 2.2.6 with the following exceptions: the L-FABP mRNA-specific sense primer (5′ TGGGAAATATCAGCTGGA GTCT 3′; nucleotides 69 to 90, Fig. 12) and antisense primer (5′ CGTCTGCTGATCCT CTTGTAG 3′; nucleotides 431 to 410, Fig. 12) were used in the PCR and 57 °C was used as the annealing temperature.

4.3 Results and discussion

4.3.1 L-FABP cDNA sequence and analysis

The sequence of the zebrafish L-FABP cDNA was determined by performing 3′ RACE and 5′ RLM-RACE using L-FABP mRNA-specific primers, based on a zebrafish EST (GenBank accession number BI846703) described as being similar to the human L-FABP. A single product of ~ 800 bp was obtained in 3′ RACE (Fig. 11A) and a single

Figure 11. 3′ RACE and 5′ RLM-RACE to clone the zebrafish L-FABP cDNA. (A) Agarose gel electrophoresis of 3′ RACE product for zebrafish L-FABP. The single band ~800 bp (lane 1) was excised, cloned and sequenced. The molecular weight markers (lane L) used were the 100 bp DNA ladder from MBI Fermentas. **(B)** 5′ RLM-RACE amplified a single major product of ~200 bp from the mature and capped L-FABP mRNA treated with calf intestinal phosphatase (CIP) and tobacco acid pyrophosphatase (TAP) (lane 1). The negative control, total RNA treated with CIP but not TAP, is shown in lane 2. The molecular weight marker (Lane L) used was 100 bp DNA ladder from MBI Fermentas.

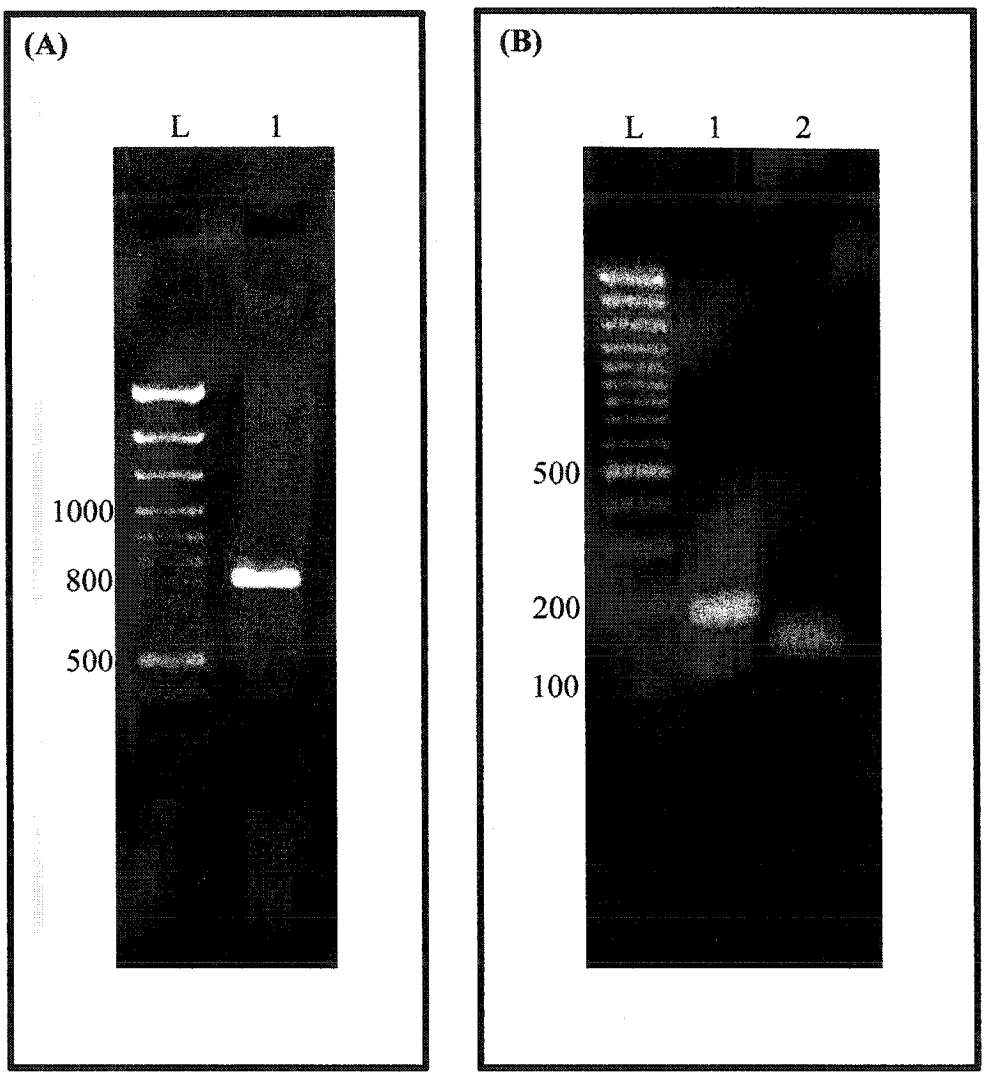


Figure 11. 3′ RACE and 5′ RLM-RACE to clone zebrafish L-FABP cDNA.

major product of ~200 bp was obtained in 5′ RLM-RACE (Fig. 11B). A minor product possibly from unprocessed or degraded mRNA was generated in the mRNA treated with CIP but not TAP (Fig. 11B) and was not cloned. Three clones for L-FABP 3′ RACE product and three clones for L-FABP 5′ RACE product were sequenced. Analysis of the sequences revealed that the L-FABP cDNA was 827 bp (excluding the poly A tail) in length (Fig. 12). An open reading frame of 384 bp from nucleotide 58 to 441, including the stop codon, was identified that codes for a polypeptide of 127 amino acids with a molecular mass of approximately 14.1 kDa and a calculated isoelectric point of 4.97. The 5′ UTR and 3′ UTR were 57 bp and 386 bp, respectively. A poly adenylation signal (AATAAA) was located at nucleotides 809-814. In one of the 5′ RLM-RACE clones, the nucleotide at position 70 was an adenine in place of guanine resulting in a change in the encoded amino acid from Gly to Arg (Fig. 12). Analysis of the zebrafish L-FABP cDNA using BLASTX revealed significant sequence similarity to L-FABP sequences from other species available in the NCBI database. Alignment of the deduced zebrafish L-FABP amino acid sequence to L-FABP and Lb-FABP sequences from different species (Fig. 13) using CLUSTALW (Thompson *et al.*, 1994) revealed 64% identity to human L-FABP suggesting that the cDNA clone codes for L-FABP in zebrafish. A phylogenetic analysis using CLUSTALX (Thompson *et al.*, 1997a) placed the zebrafish L-FABP in the same clade as L-FABP from other species (Fig. 14).

4.3.2 Organization of L-FABP gene

Searches of the zebrafish genome database (assembly Zv2) at the Wellcome Trust Sanger Institute (http://www.sanger.ac.uk/Projects/ D_rerio/) using zebrafish L-FABP

Figure 12. Zebrafish L-FABP cDNA sequence and the location of the introns in the zebrafish L-FABP gene. The 827 bp L-FABP cDNA sequence (excluding the poly A tail) was determined by cloning and sequencing the 3′ RACE and 5′ RLM-RACE products. The cDNA sequence contained an open reading frame of 384 nucleotides coding for a protein of 127 amino acids. A variation between the 5′ RLM-RACE products is shown in bold font with the variation indicated above. The intron positions obtained by aligning the genomic sequence obtained from contig ctg30243.1 (assembly Zv2, http://www.sanger.ac.uk/Projects/ D_rerio/) and the cDNA sequence are indicated by "▼". A polyadenylation signal sequence, AATAAA, is italicized and in bold font.

```
5'atttctgcgggatctcctgaagctgaacacacacacacacacacacaccacag▼gctg    57

                  A
ATG GCG TTC ACT GGG AAA TAT CAG  CTG GAG TCT CAC GAG AAC TTT   102
Met Ala Phe Thr Gly Lys Tyr Gln  Leu Glu Ser His Glu Asn Phe

GAG GCA TTC ATG AAG GCA GTC G▼GT GTG CCG GAT GAT GAG GTC GAG   147
Glu Ala Phe Met Lys Ala Val Gly  Val Pro Asp Asp Glu Val Glu

AAA GGC AAA GAC ATT AAG AGC ATC  TCT GAG ATC CAT CAG GAC GGA   192
Lys Gly Lys Asp Ile Lys Ser Ile  Ser Glu Ile His Gln Asp Gly

AAA GAC TTC AAG GTG ACG GTG ACA  GCC GGA ACT AAA GTC ATC CTG   237
Lys Asp Phe Lys Val Thr Val Thr  Ala Gly Thr Lys Val Ile Leu

TAC TCC TTC ACT GTG GGC GAG GAG  TGT GAG CTG GAG ACG TTC ACT   282
Tyr Ser Phe Thr Val Gly Glu Glu  Cys Glu Leu Glu Thr Phe Thr

GGA GAC AGA GCT AAA▼ACT GTG GTT CAA ATG GAT GGT AAT AAG CTG   327
Gly Asp Arg Ala Lys  Thr Val Val Gln Met Asp Gly Asn Lys Leu

ACA GCG TTT GTG AAG GGG ATT GAG  TCT GTG ACG GAG CTG GAT GGA   372
Thr Ala Phe Val Lys Gly Ile Glu  Ser Val Thr Glu Leu Asp Gly

GAC ACT ATC▼AGT AAC ACT CTC AGC TTT AAT GGT ATC GTC TAC AAG   417
Asp Thr Ile  Ser Asn Thr Leu Ser Phe Asn Gly Ile Val Tyr Lys

AGG ATC AGC AGA CGC ATC TCA TGA tcatcatcatcattttcatctgtgtgt    468
Arg Ile Ser Arg Arg Ile Ser  *

ttgtaaagcgacgctcatgggtttgctggtgagggcggcttcacgcatgctgtgataaa    527
catgttcgttctgcttgaacatgcgcagctccatctctgtctgcctccgcatcgcctga    586
aaacacacacacacaccagcatcacacacactcttttatattcttctgctattgatggg    645
atcataaacatgattgatagactgtttagttgcttcatggtgagacagcaaaggttaaa    704
tcagtttttcatgcactctatactgtatttcagactagcggagaatgctgaatagacct    763
ggatgtgtttgttgttaaatgtttgcatctgtagagttcaatgac***aataaa***tatcttca    822
gtttgaaaaaaaaaaaaaaaaaa 3'                                    846
```

Figure 12. Zebrafish L-FABP cDNA sequence and the location of the introns in the L-FABP gene.

Figure 13. Sequence alignment of zebrafish L-FABP with L-FABP and Lb-FABP sequences from other species. The amino acid sequences of zebrafish L-FABP (Zf-L-FABP), human L-FABP (Hu-L-FABP; GenBank accession number P07148), chick L-FABP(Ch-L-FABP; AAK58095), pig L-FABP (Pi-L-FABP; P49924), salamander L-FABP (Sa-L-FABP; P81399), rat L-FABP (Ra-L-FABP; P02692), mouse L-FABP (Mo-L-FABP; Y14660), cow L-FABP (Bo-L-FABP; P80425), pufferfish L-FABP (Pf-L-FABP; AAC60290), salamander Lb-FABP (Sa-Lb-FABP; P81400), iguana Lb-FABP (Ap-Lb-FABP; U28756), chick Lb-FABP (Ch-Lb-FABP; P80226), Lb-FABP (Zf-Lb-FABP; AF254642), lungfish Lb-FABP (Lf-Lb-FABP; P82289), catfish Lb-FABP (Cf-Lb-FABP; P80856), shark Lb-FABP (Sh-Lb-FABP; P81653) and toad Lb-FABP (Td-Lb-FABP; P83409) were aligned using CLUSTALW (Thompson *et al.*, 1994). Dots indicate identity and dashes were introduced to maximize alignment. The percent identity between the zebrafish L-FABP and other L-FABPs or Lb-FABPs are shown at the end of each sequence.

```
Zf-L-FABP      1 MAFTGKYQLESHENFEAFMKAVGVPDDEVEKGKDIKSISEIHQDGKDFK
Hu-L-FABP      1 .S.S.....Q.Q.........I.L.EELIQ......GV...V.N..H..
Ch-L-FABP      1 .S.....E.Q......P....L.L...QIQ...........V.N.NK..
Pi-L-FABP      1 .N.S....VQ.Q...........L..ELIQ......GT...V.N..H..
Sa-L-FABP      1 -S.A...E.Q.Q.........I.L..ELIQ.......V...Q.N..S..
Ra-L-FABP      1 .N.S....VQ.Q....P....M.L.E.LIQ......GV...VHE..KV.
Mo-L-FABP      1 .N.S.....Q.Q....P....I.L.E.LIQ......GV...VHE..KI.
Bo-L-FABP      1 .N.S....VQTQ..Y........M...IIQ......GV...V.N..H..
Pf-L-FABP      1 .S.S....QV.Q....P....I.L..EVIQQV.EL..T...E.N.N...
Sa-Lb-FABP     1 -P.N.TW.VY.Q..Y...LR...L.E.IINVA...NP.I..Q.N.DN.V
Ap-Lb-FABP     1 ...N.TW.VY.Q..Y.D.L..IAL...IIKAA..V.PVT..R.T.NT.V
Ch-Lb-FABP     1 -..S.TW.VYAQ..Y.E.L..LAL.E.LIKMAR...P.V..Q.K.D..V
Zf-Lb-FABP     1 ...S.TW.VYAQ..Y.E.LR.ISL.EEVIKLA..V.PVT..Q.N.S..T
Lf-Lb-FABP     1 -..S.TW.VYAQ..Y...L.VI..AE.IIPHA.E..PTI..Q.S.NS.T
Cf-Lb-FABP     1 -..S.TW.VYAQ..Y.E.LR.ISL.E.VIKLA..V.PVT..Q.T.N..V
Sh-Lb-FABP     1 -..S.TW.VY.Q..I.D.LR.LSL.EEVIKI.....PVID.K.T.EH.V
Td-Lb-FABP     1 -..N.TWNVYAQ..Y.N.LRT..L.E.IIKVA..VNPVI..E.N.NE.V

Zf-L-FABP     50 VTVTAGTKVILYSFTVGEECELETFTGDRAKTVVQMDG-NKLTAFVKGI
Hu-L-FABP     50 F.I...S...QNE...........M..EKV.....LE.D...VTTF.N.
Ch-L-FABP     50 I...T.S..MTNE..I.....M.LL..EK..CI.N.E.N...V.NL..L
Pi-L-FABP     50 L.I.T.S..VQNE..L.....M..L..EKV.....LE.D...VTTF...
Sa-L-FABP     49 ....T.S..LENE..L...A....L..EKV.SI.KQE.D...VVNL...
Ra-L-FABP     50 L.I.Y.S...HNE..L........M..EKV.A..K.E.D..MVTTF...
Mo-L-FABP     50 L.I.Y.P..VRNE..L........M..EKV.A..KLE.D..MVTTF...
Bo-L-FABP     50 FII...S...QNE..L.....M.FM..EKI.A...QE.D...VTTF...
Pf-L-FABP     50 I.I.T.P..TVNK..I.K.T.MD.I..EKI...FHL..-...KVSL...
Sa-Lb-FABP    49 ..SKTPNQSVTN...I.K.A.ITSMG.KKI.CT.VLE.-G..VSKTDQF
Ap-Lb-FABP    50 ..SKTPN.SVTN...L.K.ADMT.MD.KKV.CT.NLVD-G..V.KSDKF
Ch-Lb-FABP    49 ..SKTPRQTVTN...L.K.ADIT.MD.KKL.CT.HLAN-G..VTKSEKF
Zf-Lb-FABP    50 I.SKTPG.TVTN...I.K.A.IT.MD.KKL.CI.KL..-G..VCRTDRF
Lf-Lb-FABP    49 ..S..QK.STTNT..I.K.A.IT.MN.NKLRCTIN.ED-G..VCKTEKF
Cf-Lb-FABP    49 I.SKTPG.SVTN...I.K.A.IT.MD.RKL.CI.KLE.-G..ISETEKF
Sh-Lb-FABP    49 IV.KTSQQTVTNE....K.A.ITSMD.KKL.CT..LED-G..V.KKLKF
Td-Lb-FABP    49 ..SKTPKQTHSN.....K.S.ITSMD.KKI.VT..LE.-G..ICKSDKF

Zf-L-FABP     98 ESVTEL-DGDTISNTLSFNGIVYKRISRRIS              100%
Hu-L-FABP     99 K.....-N..I.T..MTLGD..F....K..-              64%
Ch-L-FABP     99 K.....-N....TH.MTKGDLT.....K..-              62%
Pi-L-FABP     99 K.....-N..I.TS.MTLGD..F....K..-              62%
Sa-L-FABP     98 T.....-S...LI...QKGDDT.....K..-              61%
Ra-L-FABP     99 K....F-N....T..MTLGD.....V.K..-              59%
Mo-L-FABP     99 K.....-N....T..MTLGD.....V.K..-              59%
Bo-L-FABP     99 K....F-N...VTS.MTKGDV.F..V.K..-              58%
Pf-L-FABP     97 ......A.PN..T--MTLGDV...TT.K.M-              52%
Sa-Lb-FABP    97 SHIQ.V-K.NEMVE..TVG.ATLI.R.K.V-              36%
Ap-Lb-FABP    98 IHEQ.I-V.NEMVE.ITSGSATFT.R.KK.-              34%
Ch-Lb-FABP    97 SHEQ.V-K.NEMVE.IT.G.VTLI.R.K.V-              34%
Zf-Lb-FABP    98 SHIQ.I-KAGEMVE..TVG.TTMI.K.KK.-              34%
Lf-Lb-FABP    97 SHIQ.V-Q.EEMIE..TSGSATLI.R..KV-              34%
Cf-Lb-FABP    97 SHKQ.I-K.GEMIE..TVA.TTMV.K.KKV-              33%
Sh-Lb-FABP    97 THIQ.V-Q.NEMIEK.TAGNATMI.K...M-              32%
Td-Lb-FABP    97 SHIQ.V-N..EMVEKITIGSSTLT.K.K.V-              32%
```

Figure 13. Sequence alignment of zebrafish L-FABP with L-FABP and Lb-FABP sequences from other species.

Figure 14. Phylogenetic analysis to test the position of zebrafish L-FABP in the iLBP multigene family. The bootstrap neighbour-joining phylogenetic tree was constructed with CLUSTALX (Thompson *et al.*, 1997a) using human Von Ebner's gland protein (LCN1, GenBank accession number NP_002288) that belongs to the lipocalin family of the calycins as outgroup. The bootstrap values supporting the branch points are shown as number per 1000 duplicates. Branch points supported by boot strap value of at least 700 are indicated. The inclusion of zebrafish L-FABP (Zf-L-FABP) in the L-FABP clade is highly supported. The sequences used in the analysis in addition to those mentioned in figure 13 include cow epidermal-FABP (Bo-E-FABP; P55052), human E-FABP(Hu-E-FABP; Q01469), cow adipocyte-FABP (Bo-A-FABP; P48035), human A-FABP(Hu-A-FABP; P15090), human brain-FABP (Hu-B-FABP; O15540), zebrafish B-FABP (Zf-B-FABP; AF237712), mouse testis-LBP (Mo-TLBP; O08716), rat TLBP (Ra-TLBP; P55054), human CRABPI (Hu-CRABPI; NM_004378), human CRABPII (Hu-CRABPII; M68867), human CRBPI (Hu-CRBPI; NP_002890), human CRBPII (Hu-CRBPII; P50120), human intestinal-FABP (Hu-I-FABP; P12104), zebrafish I-FABP (Zf-I-FABP; AF180921), orange-spotted grouper FABP (Ec-L-FABP; AAM22208), pig Ileal-LBP (Pi-IlLBP; P10289), human IlLBP (Hu-IlLBP; NP51161), Mouse IlLBP (Mo-IlLBP; NP51162)and rat IlLBP (Ra-IlLBP; P80020). Scale bar = 0.1 substitutions per site.

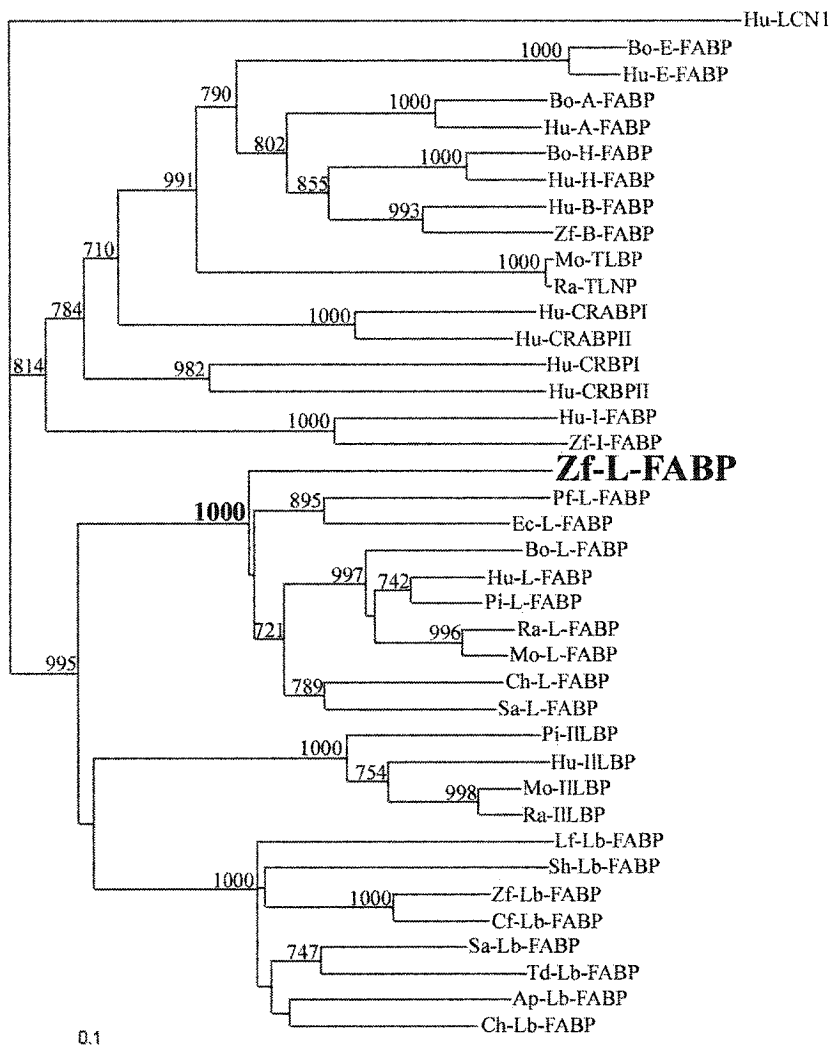


Figure 14. Phylogenetic analysis to test the position of zebrafish L-FABP in the iLBP multigene family.

cDNA sequence identified a contig, ctg30243.1 containing sequence of the zebrafish L-FABP gene. Comparison of the genomic sequence from the contig and the cDNA sequence reveal that the complete sequences of exons 1-4, introns 1-3 and partial sequence of intron 4 of the zebrafish L-FABP gene were present in the contig. The sequence of exon 5 not present in the contig sequence was obtained from the zebrafish L-FABP cDNA sequence. Therefore, it appears that the zebrafish L-FABP gene contains five exons separated by four introns (Fig. 12). Four of the five exons of the zebrafish L-FABP gene constitute the coding region while one exon is present in the 5′ UTR (Fig. 12). Rat L-FABP is the only other L-FABP for which full cDNA and gene sequence is available (Sweetser *et al.*, 1986). The exon present in zebrafish L-FABP 5′ UTR (exon 1) is not present in the rat L-FABP gene. The introns in the coding region of the zebrafish L-FABP gene are located at a position comparable to the rat L-FABP gene and the exons in the coding region have the coding capacity similar to the rat L-FABP exons. The vertebrate iLBP genes generally contain four exons interrupted by three introns (Bernlohr *et al.*, 1997), and the presence of more than four exons in the zebrafish L-FABP gene is an exception to common iLBP gene organizations. The exception to the four exon-three intron organization of the iLBPs gene has previously been reported for the muscle-type FABP gene from desert locust and a putative muscle-type FABP gene from *Drosophila*, both of which lack intron 2 (Wu *et al.*, 2001) but the five exon-four intron organization for an iLBP gene has not so far been reported for any other species. The additional intron in the zebrafish L-FABP gene could be a result of insertional mutation in the finned-fish lineage after its divergence from the lobe-finned fish lineage. It is also possible that the insertion occurred in the zebrafish lineage independently of other fish lineages.

The presence of two separate genes for L-FABP and Lb-FABP in zebrafish provide evidence that the two genes are paralogs. Also, the presence of both L-FABP and Lb-FABP genes in zebrafish suggests that the two liver FABP types diverged much before the fish-tetrapod split (approximately 400 mya). Therefore, providing additional support to the estimated divergence time of 680 mya from the last common ancestral gene (Schaap *et al.*, 2002).

4.3.3 Radiation hybrid mapping of the zebrafish L-FABP gene

The L-FABP gene was assigned to linkage group 5 at a distance of 10.2 centiRay from marker Z22208 with a LOD of 14.5 using the LN 54 panel of radiation hybrids (Hukriede *et al.*, 1999) and the primers specific to L-FABP gene (Fig. 15). Comparative analysis of the genes mapped to zebrafish linkage group 5 and orthologous genes in the human genome revealed that the zebrafish linkage group 5 has syntenic relationship with human chromosome 2. The genes for L-FABP, paired box gene 8 (*PAX8*) and minichromosome maintenance deficient 6 (*MCM6*) are located on zebrafish LG 5 and on human chromosome 2 (Table 4). In the rat genome, the L-FABP gene is located on chromosome 4, the *Pax8* on chromosome 3 and the mini chromosome maintenance deficient 6 [*S. cerevisiae*] (*Mcmd6*) on chromosome 13 while in the mouse genome the L-FABP gene is located on chromosome 6, the *Pax8* on chromosome 2 and the *Mcm6* on chromosome 1 (Table 4). The observations suggest that rearrangements and linkage breaks have occurred in these regions of the mouse and the rat genome since they diverged from the common ancestor of mammals approximately 70-100 mya (http://www.cbs.dtu.dk/~gorm/modelorganisms.html#mouse). The synteny, is however, maintained between zebrafish and humans.

Figure 15. Location of the primers used in radiation hybrid mapping of zebrafish L-FABP gene. The 517 bp sequence including the part of intron 2, exon 3, intron 3 and part of exon 4 of the L-FABP gene amplified by using sense (1) and antisense (2) primers to map the zebrafish L-FABP gene to linkage group 5. The sequences of the introns are in lower case and italicized. The sequence is obtained from contig, ctg30243.1 (assembly Zv2, http://www.sanger.ac.uk/ Projects/ D_rerio/).

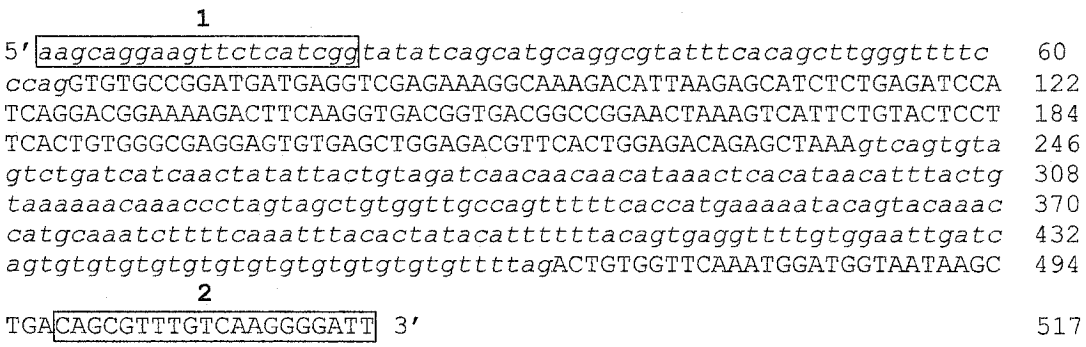


Figure 15. Location of primers used in radiation hybrid mapping of zebrafish L-FABP gene.

Table 4. Zebrafish-mammals conserved syntenies defined by the L-FABP gene.

Zebrafish[1]			Human[2]		Mouse[2]		Rat[2]	
LG	**Locus zebrafish**	**Accession Number**	**Gene**	**Chromosomal Position**	**Gene**	**Chromosomal Position**	**Gene**	**Chromosomal Position**
5	fabp1/ L-FABP		*FABP1*/ L-FABP	2p11	*Fabp1*/ L-FABP	6 (30.0 cM)	*Fabp1*/ L-FABP	4q33
5	pax8	AF072549	PAX8	2q12-q14	Pax8	2 (13.5 cM)	Pax8	3p13
5	mcm6	AI415815	MCM6	2q21	Mcm6	1 (E4)	Mcmd6	13q12

[1]Woods *et al.*, 2000.

[2] LocusLink (http://www.ncbi.nlm.nih.gov/LocusLink/list.cgi), NCBI.

4.3.4 Developmental expression of the L-FABP gene during zebrafish embryogenesis

RT-PCR using total RNA extracted from zebrafish embryos (Fig. 16A) and whole mount *in situ* hybridization to zebrafish embryos (data not shown) did not detect L-FABP mRNA in any of the embryonic stages investigated. The constitutively expressed RACK1 mRNA was amplified by RT-PCR and was used as a positive control for each embryonic stage (Fig. 16A). In the rat intestine and liver, L-FABP mRNA is first detectable at day 17 to 19 of gestation (late fetal life) (Gordon *et al.*, 1985; Börchers and Spener, 1994). A proximal to distal gradient in the levels of rat L-FABP mRNA is established in the intestine during the late fetal stage. The mRNA concentration in the gut sharply increases 3-4 fold within 24 h after birth and increases another 2 fold during the suckling period. In hepatocytes, the L-FABP mRNA is induced during the first postnatal day but remains relatively constant during the suckling and weaning period, up to 35 days postpartum (Gordon *et al.*, 1985). RT-PCR and whole mount *in situ* hybridization analyses demonstrated that the zebrafish L-FABP mRNA reported here is not detected during embryogenesis and therefore may not perform a function equivalent to the mammalian L-FABP during zebrafish embryogenesis.

4.3.5 Tissue-specific expression of the L-FABP mRNA in adult zebrafish

An L-FABP-specific product was generated by RT-PCR from total RNA extracted from zebrafish intestine. No RT-PCR product was generated from total RNA extracted from liver, brain, testis, muscle, heart, skin or ovary (Fig. 16B). The constitutively expressed RACK1 mRNA was amplified by RT-PCR and was used as a positive control for each tissue (Fig. 16B). Di Pietro and co-workers (1997) studied the

Figure 16. Developmental and tissue-specific expression of the zebrafish L-FABP mRNA analyzed by RT-PCR. (A) Zebrafish L-FABP mRNA-specific primers did not generate a product from total RNA extracted from zebrafish embryos at 1 hpf (1), 3hpf (3), 12 hpf (12), 24 hpf (24), 60 hpf (60) or 72 hpf (72) stages. L-FABP specific product was generated from the total RNA extracted from whole adult zebrafish (A). No product was detected in the negative control (-) lacking RNA template in the RT-PCR (upper panel). A product for constitutively expressed RACK1 mRNA was detected in total RNA extracted from all embryonic stages (lower panel). **(B)** RT-PCR generated a L-FABP mRNA-specific product from total RNA extracted from adult zebrafish intestine (I). No L-FABP mRNA-specific product was generated from adult zebrafish heart (H), muscle (M), testes (T), liver (L), ovary (O), skin (S), brain (B) or the negative control (-) lacking the total RNA in the RT-PCR (upper panel). A RACK1 mRNA-specific product was generated from all the adult zebrafish tissues analyzed (lower panel).

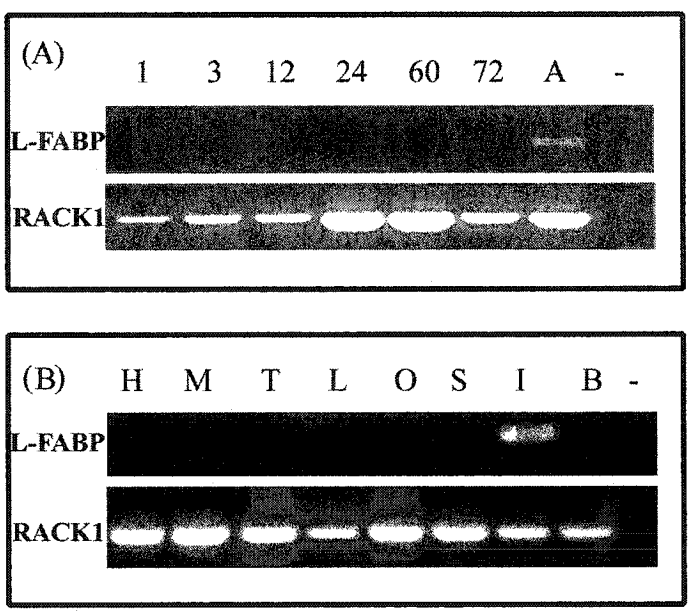


Figure 16. Developmental and tissue-specific expression of the zebrafish L-FABP mRNA analysed by RT-PCR.

expression of L-FABP in catfish tissues by western blot analysis using antibodies to rat L-FABP. Catfish L-FABP expression is restricted to intestine as observed by RT-PCR analysis in the adult zebrafish. Liver-FABP mRNA is found in adult rat intestine and liver (Sweetser *et al.*, 1986). In adult zebrafish, however, L-FABP mRNA was not detected in liver.

4.3.6 Transcription start site of the zebrafish L-FABP gene

The transcription start site of the zebrafish L-FABP gene was determined using 5′ RLM-RACE. One abundant product of approximately 200 bp was detected in the reaction with CIP/TAP treated RNA amplified in nested PCR with the adapter primer and the L-FABP mRNA-specific primer (Fig. 11B). Alignment of the 5′ RLM-RACE sequence with the zebrafish L-FABP gene sequence localized the transcription start site 407 bp upstream of the initiation codon (Fig. 17). The analysis also revealed that the 5′ RLM-RACE product had ten "CA" di-nucleotide repeats while the genomic sequence had only four "CA" di-nucleotide repeats (Fig. 17). The differences in the number of repeats could represent zebrafish strain-specific or individual differences. Di-nucleotide repeats are highly polymorphic and are used to detect genetic variation in a number of fish species (O'Rielly *et al.*, 1996; Carleton *et al.*, 2002). The sequence CAT at location -1 to +2 was observed in L-FABP gene (Fig. 17). This sequence fits the consensus for RNA polymerase II promoter start sites where the first nucleotide of the mRNA is an "A" and the corresponding A in the genomic sequence is flanked by pyrimidines (Lewin, 2000).

Figure 17. The sequence 5′ upstream of the initiation codon of the zebrafish L-FABP gene. The 1694 bp sequence 5′ upstream of the initiation codon for the zebrafish L-FABP gene was obtained from zebrafish genome database at the Wellcome Trust Sanger Institute (contig: ctg30243.1, assembly Zv2, http://www.sanger. ac.uk /Projects/ D_rerio/). The transcription start site was identified from the capped and mature L-FABP mRNA by 5′ RLM-RACE and is marked by an arrow sign. The initiation codon is boxed, intron 1 sequence is in lower case and italicized and the 1287 bp sequence 5′ upstream of the transcription start site is in lower case. The six "CA" di-nucleotide repeats not present in the genomic sequence are inserted based on the 5′ RLM-RACE sequence and are underlined. A single nucleotide variation between the 5′ RLM-RACE sequence and the genomic sequence is in bold font with the variation mentioned over it.

Figure 17. The sequence 5′ upstream of the initiation codon of the zebrafish L-FABP gene.

4.3.7 Putative 5′ upstream *cis*-acting regulatory elements of the zebrafish L-FABP gene

Analysis of the sequence 5′ upstream of the transcription start site for zebrafish L-FABP gene using MatInspector professional (http://transfac.gbf.de/cgi-bin/matSearch/matsearch.pl; Quandt *et al.*, 1995) revealed a number of potential *cis*-regulatory elements (Table 5). The zebrafish L-FABP 5′ upstream sequence does not contain TATA or CAAT box elements present in the rat L-FABP promoter (Simon *et al.*, 1993). Four potential GATA1 binding sites were present in the 5′ upstream sequence of the zebrafish L-FABP gene. A putative binding site for transcription factor GATA is present in the rat L-FABP promoter.

Five putative C/EBP binding elements, but no PPREs known to be involved in the regulation of the rat L-FABP (Bernlohr et al., 1997) were identified in the zebrafish L-FABP promoter. Binding elements for hepatocyte nuclear factor (HNF)-1α shown to be involved in expression of L-FABP gene in the mouse liver (Akiyama *et al.*, 2000) were not present in the zebrafish L-FABP 5′ upstream sequence. Akiyama and co-workers (2000) demonstrated that targeted disruption of HNF-1α eliminates L-FABP expression in the liver. Simon and co-workers (1993) analyzed the PPRE in the rat L-FABP promoter and demonstrated that it regulates the L-FABP gene expression in liver but does not have any significant effect on the regulation of L-FABP in the intestine. The absence of PPRE and the HNF-1α binding elements in the zebrafish L-FABP promoter may be responsible for the loss of expression of the L-FABP in the zebrafish liver.

Table 5. Potential *cis*-regulatory elements in the 5′ upstream sequence of the zebrafish L-FABP gene.

Matrix Name	Further Information	Position	Strand	Core Simil.	Matrix Simil.	Sequence
GATA1_02	GATA-binding factor 1	-1089	(-)	1.000	0.970	aaattGATAgggca
GATA1_02	GATA-binding factor 1	-849	(-)	1.000	0.923	ctgagGATActgac
GATA1_02	GATA-binding factor 1	-788	(-)	1.000	0.907	tttaaGATAaataa
GATA1_02	GATA-binding factor 1	-61	(+)	1.000	0.947	gcattGATAggttt
CEBPB_01	CCAAT/enhancer binding protein	-1093	(-)	1.000	0.876	tgataggGCAAata
CEBPB_01	CCAAT/enhancer binding protein	-1025	(-)	1.000	0.898	aaataatGCAAaaa
CEBPB_01	CCAAT/enhancer binding protein	-971	(+)	1.000	0.907	aaatcaaGCAAgct
CEBPB_01	CCAAT/enhancer binding protein	-768	(-)	1.000	0.911	aattaatGCAAtat
CEBPB_01	CCAAT/enhancer binding protein	-358	(-)	1.000	0.972	atgttttGCAAgat

Chapter 5: The zebrafish cellular retinoic acid-binding protein type II: cDNA sequence, phylogenetic analysis, mRNA expression and gene linkage analysis.

5.1 Introduction

Retinoic acid, one of the derivatives of vitamin A, is important for normal vertebrate development, cell growth, and differentiation. Retinoic acid regulates cell proliferation and differentiation by interacting with nuclear receptors, the retinoic acid receptors (RARs), which exhibit high affinity for all-*trans* and 9-*cis* retinoic acid, and the retinoid X-receptors (RXRs), which have moderate affinity for 9-*cis* retinoic acid (reviewed in Lampron *et al.*, 1995). Knockout RAR or RXR mouse mutants show a wide range of defects, similar to those reported for vitamin A deficient fetuses, thereby revealing a pivotal role of RARs and RXRs in the retinoic acid signalling pathway (Kastner *et al.*, 1995). RARs and RXRs act as inducible transcriptional *trans*-regulators by binding to RAR and RXR response elements (called RAREs and RXREs, respectively) in the control regions of specific genes. Some RAR targets are the mouse and human RAR (α, β, γ), human class1 alcohol dehydrogenase 3, mouse and rat CRBPI, mouse CRABPII, rat and bovine growth hormone, *Xenopus* vitellogenin A2, and human oxytocin dehydrogenase genes (Mangelsdorf *et al.*, 1994). RXR targets include the rat CRBPII, chick ovalbumin, rat acyl-CoA oxidase, human apolipoprotein AI, mouse CRABPII, human hepatitis B virus, and human medium chain acyl-coenzyme A dehydrogenase genes (Mangelsdorf *et al.*, 1994).

As retinoic acid is hydrophobic, it remains unclear how retinoic acid traverses the aqueous milieu of the cytoplasm to reach its receptors in the nucleus. In addition to RARs and RXRs, retinoic acid is bound by two intracellular proteins named CRABPI and

CRABPII. CRABPs are low molecular mass (~15 kDa) proteins that belong to the multigene family of intracellular lipid-binding proteins (iLBP) (reviewed in Bernlohr *et al.*, 1997).

While the function(s) of CRABPI and II is not well understood, CRABPI and II bind all *trans*-retinoic acid. The affinity of CRABPII for retinoic acid is lower than that of CRABPI (Ong *et al.*, 1994) and both proteins bind 9-*cis* retinoic acid with much lower affinity than all-*trans* retinoic acid (Fogh *et al.*, 1993). Of the two CRABPs, CRABPI is present in the seminal vesicle, testis, skin, brain and kidney of adult rat while CRABPII is present in the skin, ovary, brain and uterus (Bucco *et al.*, 1996 and references therein). In adult humans, CRABPII, but not CRABPI gene expression is induced in skin and in cultured adult human skin fibroblasts by retinoic acid (Åström *et al.*, 1991).

In an effort to elucidate the function(s) of CRABPs, knockout mice for CRABPI and CRABPII, and double knockouts of these genes, have been generated and found to be normal except for some minor limb deformities (Lampron *et al.*, 1995; Gorry *et al.*, 1994). Contrary to their previously proposed functions, these results suggest that CRABPs are not essential in the retinoic acid signalling pathway or other members of the iLBP family compensate for the lack of CRABP. The CRABPs are, therefore, considered dispensable for retinoic acid intracellular homeostasis and protection against retinoic acid excess. Based on findings to date, the function of CRABPs in cell physiology remains to be defined. Currently, there is no information on the expression of CRABPs from fishes, the largest and most evolutionary diverse group of vertebrates. Retinoic acid and other derivatives of vitamin A are clearly important for normal vertebrate development. Since zebrafish is promoted as a model system for vertebrate development, an understanding of

the function of CRABP in zebrafish is of importance to developmental biologists. Comparative studies between fishes and mammals may also provide important insight into the function and evolution of CRABPs.

In this chapter, I report the cDNA sequence, gene organization, transcription start site, spatio-temporal expression during zebrafish development, tissue-specific expression in adult zebrafish, and gene-linkage mapping of CRABPII from zebrafish.

5.2 Materials and methods

5.2.1 CRABPII cDNA sequence and analysis

Searches of the zebrafish database in GenBank identified a cDNA clone (GenBank accession number AI883746) that was similar to a cellular retinoic acid-binding protein from *Xenopus laevis*. This clone (fc69e09.y1) was purchased from Incyte Genomics Inc. and the complete nucleotide sequence was determined. The cDNA sequence was analyzed as described in section 3.2.1. Sequences of the pufferfish CRABPs were retrieved from scaffold 1427, 932, 23 and 7 by searching the pufferfish (*Fugu rubripes*) genome database v.3.0 at http://genome.jgi-psf.org/fugu6/fugu6.info.html (Aparicio *et al.*, 2002) using human CRABPI and CRABPII as query sequences.

For phylogenetic analysis, amino acid sequences of intracellular lipid-binding proteins were obtained from GenBank, pufferfish genome database v.3.0 (Aparicio *et al.*, 2002) and the published literature (Table 6). The amino acid sequences were aligned with the zebrafish CRABP (AF497478) by CLUSTALX (Thompson *et al.*, 1997a). The sequences were edited manually and protein maximum likelihood (ML) analysis (Quartet Puzzling, QP) was done using Treepuzzle version 4.0 (Strimmer and von Haeseler,

Table 6. Amino acid sequences of intracellular lipid-binding proteins used to analyze the zebrafish CRABP.

Protein	Organism	Sequence	Protein	Organism	Sequence
I-FABP	Human	P12104	CRABPI	Human	NM_004378
	Frog	AAC38012		Mouse	NM_013496
	Zebrafish	AF180921		Zebrafish	AY242125 (R-Z. Liu and J.M. Wright, unpublished data).
Lb-FABP	Axolotl	P81400	CRABPIa	Pufferfish	O42386
	Iguana	Q90239	CRABPII	Human	M68867
	Zebrafish	AAF67743		Mouse	M35523
L-FABP	Human	P07148		Rat	U23407
	Cow	P80425	xCRABP	Frog	Dekker *et al.*, 1994
	Axolotl	P81399	XCRABP-b	Frog	S74933
	Pufferfish	AAC60290	CRABP	Cow	M36808
	Shark	P81653		Zebrafish	AF497478
IILBP	Human	P51161		Shrimp	AF458289
	Pig	P10289		Tobacco hornworm	U75307
TLBP	Mouse	O08716	CRABPIIa	Pufferfish	Scaffold 1427, genome database v.3.0 (Aparicio *et al.*, 2002).
	Rat	P55054	CRABPIIb	Pufferfish	Scaffold 932, genome database v.3.0 (Aparicio *et al.*, 2002).
E-FABP	Human	Q01469	CRABPIb	Pufferfish	Scaffold 7, genome database v.3.0 (Aparicio *et al.*, 2002).
	Cow	P55052	CRBPI	Human	NP_002890
A-FABP	Human	P15090		Cow	P02694
	Cow	P48035	CRBPII	Human	P50120
H-FABP	Human	P05413		Zebrafish	Aam95336
	Cow	CAA31212	CRBPIII	Human	P82980
	Trout	O13008	CRBPIV	Mouse	Q9EPC5
B-FABP	Human	O15540	LBP-2	Nematode	AAA81435
	Mouse	P51880			
	Zebrafish	AF237712			

The sequences were obtained from GenBank, the original references or the pufferfish genome sequence database v.3.0 (Aparicio *et al.*, 2002).

1996). ML distance analyses were performed using Treepuzzle, Puzzleboot (A. Roger and M. Holder; http://members.tripod.de/korbi/puzzle/), and programs from the PHYLIP package (Felsenstein, 1995). Branch lengths were estimated by finding the best ML distance topology using Treepuzzle and PHYLIP. All Treepuzzle and PHYLIP analyses used the following parameters: JTT model of substitutions, gamma-distributed rates (8 categories), α-parameter 3.41 (estimated from the data), global rearrangement, and ten times jumbling. Parsimony analysis was performed using PAUP* Version 4.0b10 (Swofford, 2002) with the following options: heuristic search, tree bisection-reconnection (TBR) branch swapping, 100 replicates, and random addition.

5.2.2 Sequence of CRABPII gene and the 5' upstream sequence

Partial sequence of the zebrafish CRABPII gene including 1807 bp 5′ upstream of the initiation codon, was obtained from the zebrafish genome database at the Wellcome Trust Sanger Institute (assembly Zv2, http://www.sanger.ac.uk/Projects/ D_rerio/). The intron/exon junctions of the zebrafish CRABPII gene sequence were determined by aligning the cDNA sequence with the genomic sequence. The 5' sequence upstream of the transcription start site was analysed for potential *cis*-acting regulatory elements by computer-assisted analysis performed using MatInspector professional at http://transfac.gbf.de/cgi-bin/matSearch/matsearch.pl (Quandt *et al.*, 1995).

5.2.3 Mapping the transcription start site of the zebrafish CRABPII gene

The 5′ RLM-RACE using capped and mature mRNA was performed to map transcription start site of the zebrafish CRABPII gene. The template for 5′ RLM-RACE was prepared as described in section 2.2.5. Nested PCR using two sense adapter primers

(refer to section 2.2.5) and three antisense primers specific to CRABPII mRNA (outer: 5′ GCATCATTGGTTATTTCCCTG 3′; nt 628 to 608, Fig. 18, inner 1: 5′ GCTCTTCAT TTTCCAGGTACC 3′; nt 311 to 291, Fig. 18 and inner 2: 5′ CCAGGTACCAGCAAA ATCGG 3′; nt 280 to 299, Fig. 18) was performed. The primary and secondary PCR conditions were similar to those described in section 3.2.5 with the exception that 58°C was used as annealing temperature in the primary PCR and the inner 1 CRABPII mRNA-specific primer was used in the secondary PCR. Half a μL of the secondary PCR product was used as template in the third round of PCR using inner adapter primer and the inner 2 CRABPII mRNA-specific primer. The conditions for tertiary PCR were similar to those for secondary PCR. The product from the tertiary PCR was size fractionated by 1.5 % agarose gel electrophoresis. A single major band of ~350 bp in the +TAP reaction was excised from the agarose gel, cloned and sequenced as described in section 3.2.5.

5.2.4 Linkage analysis of zebrafish CRABPII gene by radiation hybrid mapping

The LN54 radiation hybrid panel (zebrafish DNA in a mouse background) (Hukriede *et al.*, 1999) was used to assign the CRABPII gene to a zebrafish linkage group. Sense (5′ GGAAGAACTGACTGAAAGCC 3′; nt 24 to 5, Fig. 18) and antisense (5′ GCTCTTCATTTTCCAGGTACC 3′; nt 311 to 291, Fig. 18) oligonucleotides synthesized to amplify exon 1 of the zebrafish CRABP gene were used in a PCR. The PCR conditions were similar to those described in section 2.2.4 with the exception that the annealing temperature used in PCR was 62.5 °C. The RH-mapping panel was scored and analyzed as described in section 2.2.4.

5.2.5 RNA extraction and northern blotting to determine the size of zebrafish CRABPII transcript.

Total RNA was extracted from adult zebrafish and the northern blot was prepared as described in section 3.2.6. The CRABPII cDNA generated by RT-PCR from total RNA of adult zebrafish with sense (5′ GGAAGAACTGACTGAAAGCC 3′; nt 24 to 5, Fig. 18) and antisense (5′ GCATCATTGGTTATTTCCCTG 3′; nt 628 to 608, Fig. 18) oligonucleotides was labelled with α-[^{32}P]-dATP and used as a hybridization probe. The membrane was prehybridized in 10 ml of hybridization solution and 21 μg/ml sonicated salmon sperm DNA (refer to section 3.2.6). Prehybridization was carried out at 55°C for 2 h followed by hybridization with the denatured probe in 7 ml of fresh hybridization solution at 55°C overnight. The blot was washed twice at room temperature for 15 min in 2×SSPE, 0.1% SDS and at 55°C for 10 min in 1×SSPE, 0.1% SDS. Following hybridization, the membrane was exposed four days to X-ray film at –70°C.

5.2.6 Developmental and tissue-specific expression of CRABPII mRNA in zebrafish

The tissue-specific expression of the zebrafish CRABPII was determined by using total RNA extracted from adult zebrafish heart, brain, muscle, kidney, testes, liver, ovary, skin, intestine and brain. Total RNA extracted from 1, 3, 12, 24, 51, 60 and 72 hpf and adult zebrafish using Trizol® (Gibco BRL), was used to define the temporal expression of zebrafish CRABPII using RT-PCR. The cDNA synthesis and PCR were performed as described in section 2.2.6 with the exception that 61.5°C was used as annealing temperature. Whole mount *in situ* hybridization using antisense riboprobe based on CRABPII cDNA clone was performed as described in section 2.2.6.

Quantitative RT-PCR for zebrafish CRABPII and β-actin was performed as described in section 2.2.6. To determine the standard curve of the crossing points for the amplification of CRABP and β-actin from tissue-specific cDNA samples, five dilutions of the CRABP and β-actin product ranging from 9.2×10^5 to 9.2×10^1 copies per reaction were used in individual quantitative PCR reactions.

5.3 Results and discussion

5.3.1 CRABPII cDNA sequence analysis

The complete cDNA sequence coding for a CRABP from zebrafish was 2545 bp in length excluding the 51 bp poly A tail (Fig. 18). An open reading frame of 426 bp from nucleotide 264 to 689 (including the stop codon) was identified that codes for a polypeptide of 142 amino acids with a molecular mass of approximately 15.8 kDa and a calculated isoelectric point of 5.2. The 5′ UTR is 263 bp and the 3′ UTR is 1853 bp. A poly adenylation signal (AATAAA) is located at nucleotide 2527. Analysis of the zebrafish CRABP cDNA sequence using BLASTX revealed significant sequence similarity to other CRABP sequences in the NCBI (http://www.ncbi.nlm.nih.gov /BLAST/) database. Alignment of the deduced zebrafish CRABP amino acid sequence to CRABPI and CRABPII sequences from rat, mouse, human, cow, zebrafish, pufferfish, tobacco hornworm, shrimp and *Xenopus* (Fig. 19) using CLUSTALW (Thompson *et al.*, 1994) revealed 74% identity to human CRABPII suggesting that the cDNA clone codes for CRABPII in zebrafish. Zebrafish CRABPII also had 72% identity to *Xenopus* CRABP (xCRABP-b) (Ho *et al.*, 1994), but only 69% identity to *Xenopus* CRABP (xCRABP) (Dekker *et al.*, 1994), suggesting that the zebrafish CRABPII is more closely related to xCRABP-b than xCRABP.

Figure 18. Nucleotide and deduced amino acid sequences of zebrafish CRABPII. The 2545 bp cDNA sequence (excluding the 51 bp poly A tail) contained an open-reading frame of 426 nucleotides coding for a protein of 142 amino acids. A polyadenylation signal sequence, aataaa, is italicized and in bold font. The asterisk indicates the stop codon of the CRABPII coding sequence. The intron positions obtained by aligning the genomic sequence obtained from the contig, ctg15237.1 (assembly Zv2, http://www.sanger.ac.uk/Projects/D_rerio/) and the cDNA sequence are indicated by "▼". The GenBank accession number for this sequence is AF497478.

```
      5'ttttggaagaactgactgaaagcctgagaccggaggaaaagggcacagtttgacatctc     59
cactttcctttgtcaacggtaaaaggaaaaagagctgtagaaggtttgaggaaggaggaaaatcacgt   127
tggctttatcagtgggaatttatcactgctttttagaatcagagtggcaggtacttttagagagtgtc   195
aggttcaaagactgtgactcatcggagaaaagtgataggttactttcagaggaacgacagaacggaag   263

ATG GAT CGT AAA ATT CCC GAT TTT GCT GGT ACC  TGG AAA ATG AAG AGC TCT    314
Met Asp Arg Lys Ile Pro Asp Phe Ala Gly Thr  Trp Lys Met Lys Ser Ser

GAG AAC TTC GAA GAG CTT CTT AAA GCA CTA G▼GT GTG AAC GTG ATG CTC CGT   365
Glu Asn Phe Glu Glu Leu Leu Lys Ala Leu Gly  Val Asn Val Met Leu Arg

AAG ATT GCG GTT GCA GCT GCA TCA AAA CCA TCC  GTG GAG ATT ACA CAG GAG   416
Lys Ile Ala Val Ala Ala Ala Ser Lys Pro Ser  Val Glu Ile Thr Gln Glu

GGA GAG ACA CTG ACA ATC AAG ACG TCC ACC TCA  GTG AGG ACC ACC AAC GTC   467
Gly Glu Thr Leu Thr Ile Lys Thr Ser Thr Ser  Val Arg Thr Thr Asn Val

ACC TTC ACA GTT GGA CAG GAG TTT AAT GAG GCC  ACT GTG GAT GGA CGT CCC   518
Thr Phe Thr Val Gly Gln Glu Phe Asn Glu Ala  Thr Val Asp Gly Arg Pro

TGC ACG▼AGC TTT CCT CGC TGG GTA ACA GAC AGC AAG ATT AGC TGC GAA CAG   569
Cys Thr  Ser Phe Pro Arg Trp Val Thr Asp Ser Lys Ile Ser Cys Glu Gln

ACT TTG CAG AAG GGC GAG GGT CCA AAA ACC TCA  TGG ACC AGG GAA ATA ACC   620
Thr Leu Gln Lys Gly Glu Gly Pro Lys Thr Ser  Trp Thr Arg Glu Ile Thr

AAT GAT GCT GAA TTG ATT CTG▼ACC ATG ACT GCT GAC GAT GTG GTG TGT ACA   671
Asn Asp Ala Glu Leu Ile Leu  Thr Met Thr Ala Asp Asp Val Val Cys Thr

AGA GTT TAT GTC AGA GAG TGA acatataaggaaaaaggaagaatatccatccagattggag   732
Arg Val Tyr Val Arg Glu  *

tctttcctcattaatcacaatgactgtgacttgtttccttttttcctaatctcatctttcacacttt   800
ttattgacaatgtatttgaatagtacgtcagagagccatgcaaatatagttccagagattcaagcaat   868
tttgttgttttctgattggtcaatttaaaactacatcatcttgggttcaaaagcacttttaaagattg   936
atctaaaatagatgtagacctggaaatctgaagcttggaaaccagttattggtggtatagttcataat  1004
cacagctgaaaataaagtgctaaaggtacagtcacatttagcattgttttgtcattttcatgggcaaa  1072
atagttttttttaagagccatctatacagttattcatgctgaccattagaaacccagcgagtacacaac  1140
atcataaaacgttaatattagactagattcagatcatgacgtcaggtgaccaaaattcaatgtctaac  1208
cagcatctaatgacaacgtaattttgacgtctaataatgatgtcaaattacgtttatatttggttgat  1276
tttaggttgtgttgtaaagtgactaaaatccaacgtctgatggatgtcatggtgttaacgtccgcaca  1344
acgtcaaggtgcaacatgattagatgttgattttaggttggacattgatgtcggcatgacgttgggtt  1412
ctgatgtcaacccgattttatttccaaacataatccaacatccccgtgacattggggtacattggggt  1480
acaatgtcaatatgacatcctgttgacgtcttgtgcctgcagagaaggtgcttaggaataaagtgact  1548
tctgcatcattcagtgttacacaaattacattagacaatggacatttttgtcctgaaatgtctctcca  1616
taatatggccacaccattagaatcttctcaaaacatagcagtttagtttgttatttttttcctaattta  1684
atgctatgcaaagatccaccttaaaaatgttacatccactgtccccatccatttctcagacagacctt  1752
ttttattatataactgtatggagtgtctgtttttatattttgaaggaccattcacgctaaggatgata  1820
actgtaatgataatggtatacttccaaattttagaaatataaaacaattacagaaacctgatcccaat  1888
gaatgatattgctttatttttttcaaatcagattccatttagctaaaacaagctcaagcatttaaagtg  1956
tcagacaaaaaacttgcctaacttgtttttaataaacaaaatgctatcattcatttttgtaaacttta  2024
atattgatgtcatcatagttataattctaggtgtgaatgtgccttttttttatcgtgcccgaaatatat  2092
gttcatttcatttgattgtgagaatttgtttgacaaagattatattcattagaacataatgtaacagt  2160
ttatttttaagaaaaatgacagataaaatgcatttcaacttaagttcattccatctatgtaatatccaa  2228
tcaatattttaacatttttttaagtagacccagtcatgccattaatgaaaatgtccgattcattcctat  2296
gtgaacttttttgtctgaattaaatacagcacaaaaagagatggagttaagaaaaagggtcatatgtga  2364
ttagaatcccccttagaatctcagaaaacaactgtcttccatctttacttttttgtatatattttaata  2432
aaacatgaatctgaaaagtttcatccaaacagtttttttattatgtttttatataaacagttccagca  2500
gtcatcactcatttctgtgcatggac***aataaa***agtttctgccctg(a)$_{51}$  3'              2596
```

Figure 18. Nucleotide and deduced amino acid sequences of zebrafish CRABPII.

To obtain orthologous sequences from pufferfish (*Fugu rubripes*), searches of the pufferfish genome sequence database v.3.0 (Aparicio *et al.*, 2002) was performed. The data mining revealed CRABP genes in scaffold 1427, 932, 23 and 7. The deduced amino acid sequence of the pufferfish CRABP from scaffold 23 was greater than 99% identical (data not shown) to the published pufferfish CRABPI (Kleinjan *et al.*, 1998; GenBank accession number O42386) and is most likely the same sequence. In this chapter, CRABPIa refers to the published pufferfish CRABPI (Kleinjan *et al.*, 1998) and the CRABP sequence from scaffold 23 was not included in any analysis as we presumed that it is the same sequence as CRABPIa reported by Kleinjan *et al.* (1998). The deduced amino acid sequence for pufferfish CRABP from scaffold 7 was similar to the human CRABPI by BLASTP analysis at NCBI (http://www.ncbi.nlm.nih.gov/ BLAST/) and is referred to as CRABPIb in this chapter. BLASTP analyses for the deduced amino acid sequence for pufferfish CRABPs from scaffold 1427 and 932 demonstrated that these sequences are similar to the *Xenopus* CRABP and the human CRABPII. The pufferfish CRABP from scaffold 1427 and 932 are referred to as CRABPIIa and CRABPIIb, respectively. Zebrafish CRABPII had 83% and 79% identity to pufferfish CRABPIIa and CRABPIIb, respectively, and only 68% and 65% identity to pufferfish CRABPIa and CRABPIb, respectively.

We analysed the position of the zebrafish CRABPII within the iLBP family including FABPs, CRBPs and CRABPs using three different phylogenetic methods (i.e., Quartet puzzling, ML Distance and Parsimony). The phylogenetic tree for different members of the ILBP multigene family showed that the zebrafish CRABPII clusters with CRABPs from other species (bootstrap values 94/100/100) and not with CRBPs and

```
      5′ttttggaagaactgactgaaagcctgagaccggaggaaaagggcacagtttgacatctc     59
cactttcctttgtcaacggtaaaaggaaaaagagctgtagaaggtttgaggaaggaggaaaatcacgt   127
tggctttatcagtgggaatttatcactgctttttagaatcagagtggcaggtacttttagagagtgtc   195
aggttcaaagactgtgactcatcggagaaaagtgataggttactttcagaggaacgacagaacggaag   263

ATG GAT CGT AAA ATT CCC GAT TTT GCT GGT ACC  TGG AAA ATG AAG AGC TCT   314
Met Asp Arg Lys Ile Pro Asp Phe Ala Gly Thr  Trp Lys Met Lys Ser Ser

GAG AAC TTC GAA GAG CTT CTT AAA GCA CTA G▼GT GTG AAC GTG ATG CTC CGT   365
Glu Asn Phe Glu Glu Leu Leu Lys Ala Leu Gly  Val Asn Val Met Leu Arg

AAG ATT GCG GTT GCA GCT GCA TCA AAA CCA TCC  GTG GAG ATT ACA CAG GAG   416
Lys Ile Ala Val Ala Ala Ala Ser Lys Pro Ser  Val Glu Ile Thr Gln Glu

GGA GAG ACA CTG ACA ATC AAG ACG TCC ACC TCA  GTG AGG ACC ACC AAC GTC   467
Gly Glu Thr Leu Thr Ile Lys Thr Ser Thr Ser  Val Arg Thr Thr Asn Val

ACC TTC ACA GTT GGA CAG GAG TTT AAT GAG GCC  ACT GTG GAT GGA CGT CCC   518
Thr Phe Thr Val Gly Gln Glu Phe Asn Glu Ala  Thr Val Asp Gly Arg Pro

TGC ACG▼AGC TTT CCT CGC TGG GTA ACA GAC AGC AAG ATT AGC TGC GAA CAG   569
Cys Thr  Ser Phe Pro Arg Trp Val Thr Asp Ser Lys Ile Ser Cys Glu Gln

ACT TTG CAG AAG GGC GAG GGT CCA AAA ACC TCA  TGG ACC AGG GAA ATA ACC   620
Thr Leu Gln Lys Gly Glu Gly Pro Lys Thr Ser  Trp Thr Arg Glu Ile Thr

AAT GAT GCT GAA TTG ATT CTG▼ACC ATG ACT GCT GAC GAT GTG GTG TGT ACA   671
Asn Asp Ala Glu Leu Ile Leu  Thr Met Thr Ala Asp Asp Val Val Cys Thr

AGA GTT TAT GTC AGA GAG TGA acatataaggaaaaaggaagaatatccatccagattggag   732
Arg Val Tyr Val Arg Glu  *

tctttcctcattaatcacaatgactgtgacttgtttccttttttcctaatctcatctttcacacttttt   800
ttattgacaatgtatttgaatagtacgtcagagagccatgcaaatatagttccagagattcaagcaat   868
tttgttgttttctgattggtcaatttaaaactacatcatcttgggttcaaaagcacttttaaagattg   936
atctaaaatagatgtagacctggaaatctgaagcttggaaaccagttattggtggtatagttcataat  1004
cacagctgaaaataaagtgctaaaggtacagtcacatttagcattgttttgtcattttcatgggcaaa  1072
atagtttttttaagagccatctatacagttattcatgctgaccattagaaacccagcgagtacacaac  1140
atcataaaacgttaatattagactagattcagatcatgacgtcaggtgaccaaaattcaatgtctaac  1208
cagcatctaatgacaacgtaattttgacgtctaataatgatgtcaaattacgtttatatttggttgat  1276
tttaggttgtgttgtaaagtgactaaaatccaacgtctgatggatgtcatggtgttaacgtccgcaca  1344
acgtcaaggtgcaacatgattagatgttgattttaggttggacattgatgtcggcatgacgttgggtt  1412
ctgatgtcaacccgattttatttccaaacataatccaacatccccgtgacattggggtacattggggt  1480
acaatgtcaatatgacatcctgttgacgtcttgtgcctgcagagaaggtgcttaggaataaagtgact  1548
tctgcatcattcagtgttacacaaattacattagacaatggacatttttgtcctgaaatgtctctcca  1616
taatatggccacaccattagaatcttctcaaaacatagcagtttagtttgttatttttttcctaattta  1684
atgctatgcaaagatccaccttaaaaatgttacatccactgtccccatccatttctcagacagacctt  1752
ttttattatataactgtatggagtgtctgtttttatatttgaaggaccattcacgctaaggatgata  1820
actgtaatgataatggtatacttccaaattttagaaatataaaacaattacagaaacctgatcccaat  1888
gaatgatattgctttatttttttcaaatcagattccatttagctaaaacaagctcaagcatttaaagtg  1956
tcagacaaaaaacttgcctaacttgtttttaataaacaaaatgctatcattcatttttgtaaacttta  2024
atattgatgtcatcatagttataattctaggtgtgaatgtgccttttttatcgtgcccgaaatatat  2092
gttcatttcatttgattgtgagaatttgtttgacaaagattatattcattagaacataatgtaacagt  2160
ttattttaagaaaaatgacagataaaatgcatttcaacttaagttcattccatctatgtaatatccaa  2228
tcaatattttaacatttttttaagtagacccagtcatgccattaatgaaaatgtccgattcattcctat  2296
gtgaacttttttgtctgaattaaatacagcacaaaaagagatggagttaagaaaaagggtcatatgtga  2364
ttagaatccccttagaatctcagaaaacaactgtcttccatctttacttttttgtatatattttaata  2432
aaacatgaatctgaaaagtttcatccaaacagtttttttattatgtttttatataaacagttccagca  2500
gtcatcactcatttctgtgcatggac***aataaa***agtttctgccctg(a)$_{51}$  3′                2596
```

Figure 18. Nucleotide and deduced amino acid sequences of zebrafish CRABPII.

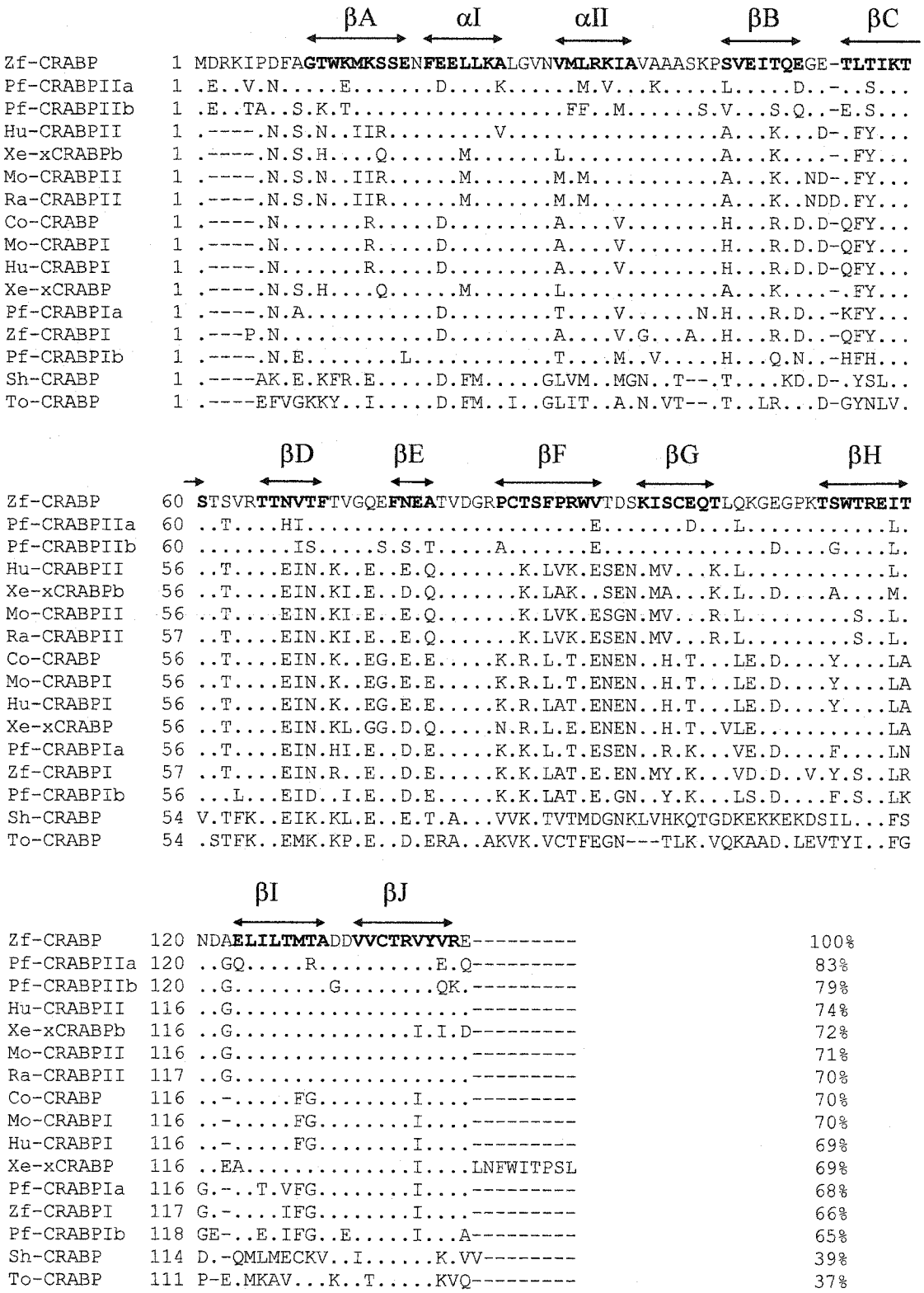


Figure 19. Sequence alignment of zebrafish CRABPII with CRABPs from other species.

FABPs (Fig. 20). To determine the orthology of the zebrafish CRABPII with the pufferfish CRABPs, the pufferfish CRABPIa, CRABPIb, CRABPIIa and CRABPIIb were included in the phylogenetic analysis along with other members of ILBP multigene family. The amino acid sequence for the zebrafish CRABPI (GenBank accession number AY242125; R-Z. Liu and J.M. Wright, unpublished data) was also included in the phylogenetic analysis. The phylogenetic tree demonstrated that the pufferfish CRABPIIa and CRABPIIb clusters with the zebrafish CRABPII, whereas, the pufferfish CRABPIa and CRABPIb clusters with the zebrafish CRABPI. This tree supports the hypothesis that the zebrafish CRABPII is a member of the CRABPII subfamily and not CRABPI.

Alignment of the zebrafish CRABPII amino acid sequence with FABPs and cellular retinoid (retinol and retinoic acid) binding protein available at FUGUE (http://www-cryst.bioc.cam.ac.uk/fugue/, Shi *et al.*, 2001) revealed that the zebrafish CRABPII contains ten β- strands (βA - βJ) and two α- helices (αI and αII) in the same organization as for all polypeptides encoded by this multigene family. Chen *et al.* (1998) proposed that three arginine residues in human CRABPII (Arg29, Arg59, and Arg132) are involved in directing retinoic acid entry in the binding pocket. Based on the amino acid sequence alignment, these three arginine residues correspond to arginine residues at positions 34, 64 and 137 in the zebrafish CRABPII sequence reported here (Fig. 19).

5.3.2 Organization of the zebrafish CRABPII gene

The zebrafish CRABPII cDNA was used to search the assembly Zv2 of the zebrafish genome database at the Wellcome Trust Sanger Institute (http://www.sanger.ac.uk/Projects/D_rerio/) to determine the organization of the

Figure 20. Phylogenetic tree of intracellular lipid-binding proteins to determine orthology of the zebrafish CRABPII. Unrooted topology is from the best ML distance tree (Treepuzzle), arbitrarily rooted with nematode LBP-2. The inclusion of the zebrafish CRABP (indicated in bold type) within the CRABP family is highly supported; support values are given in Quartet Puzzling/ML Distance/Parsimony bootstraps. Other branch points supported by a bootstrap of at least 75 in two or more methods are indicated by an asterisk (*). Subfamilies are indicated on the right; organism abbreviations: Ig – iguana; Ax – Axolotl; Ne – nematode; Co – cow; Sk – shark; Hu – human; mo – mouse; Pf – pufferfish; Pi – pig; Ra – rat; Rt – rainbow trout; Xe – Xenopus; Zf – zebrafish. Scale bar = 0.1 substitutions per site. Refer to Table 6 for source of amino acid sequences.

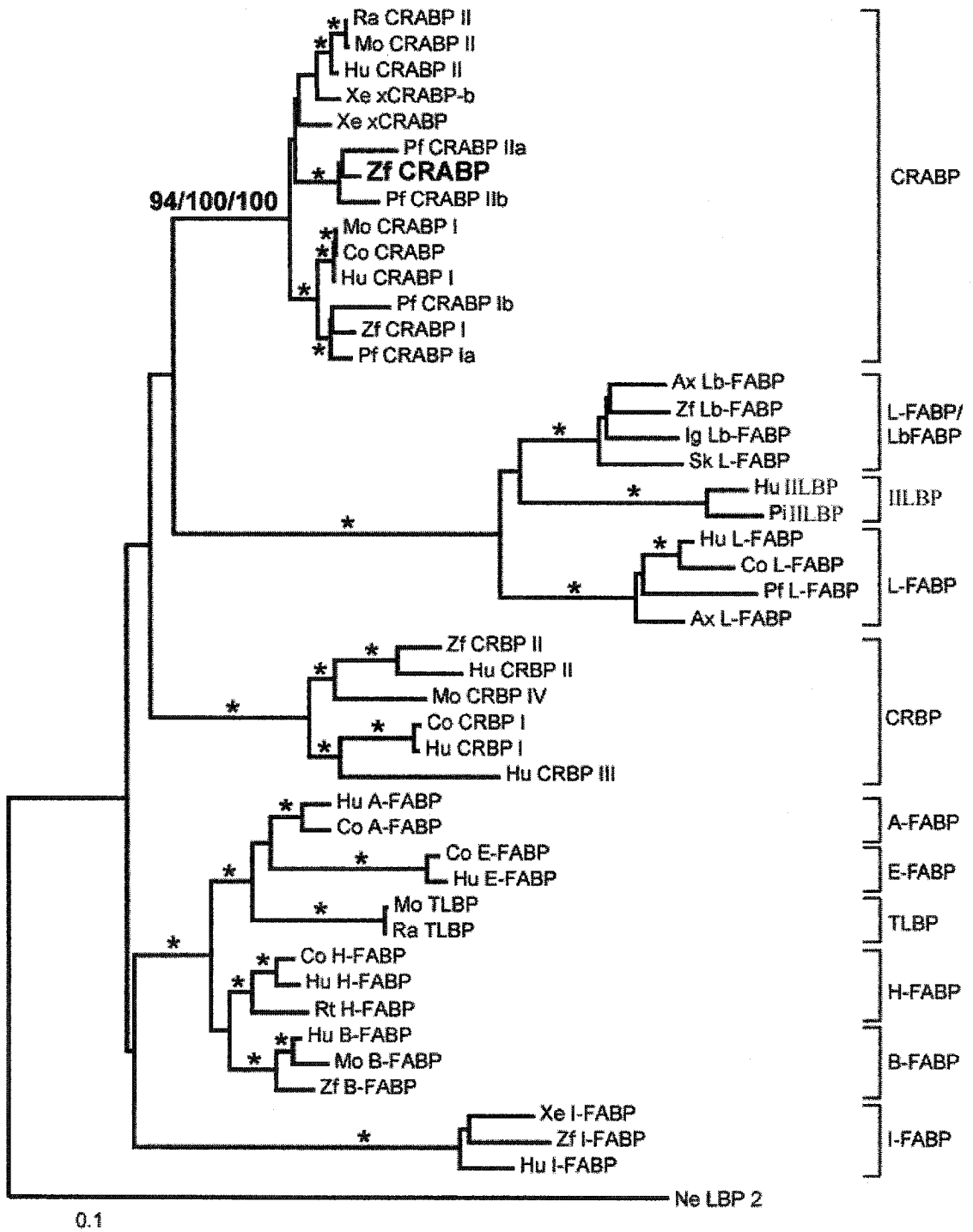


Figure 20. Phylogenetic tree of intracellular lipid-binding proteins to determine orthology of the zebrafish CRABPII

CRABPII gene. The contig, ctg15237.1 was identified as containing the sequence of CRABPII gene. The exon/intron junctions of the CRABPII gene were identified (Fig. 18) by aligning the cDNA sequence with the genomic sequence obtained from the contig, ctg15237.1. The zebrafish CRABPII gene consists of 4 exons separated by three introns. The organization of the zebrafish CRABPII gene is similar to other vertebrate iLBP genes described in literature (Bernlohr *et al.*, 1997; Zimmerman and Veerkamp, 2002) including the human CRABPII (Åström *et al.*, 1992).

5.3.3 Transcription start site and the putative 5′ upstream *cis*-acting regulatory elements of the zebrafish CRABPII gene

The transcription start site for the zebrafish CRABPII gene was determined by 5′ RLM-RACE using 5′ capped mRNA (Fig. 21). A single product of ~350 bp was obtained from mRNA treated with CIP/TAP (Fig. 21). This product representing the 5′ end of the mature and capped CRABPII mRNA was cloned and sequenced. Sequence alignment of the 317 bp 5′ RLM-RACE product, excluding the 38 bp of adapter sequence, and the 5′ upstream sequence of the zebrafish CRABPII gene obtained from contig, ctg15237.1 (assembly Zv2, http://www.sanger.ac.uk/Projects/D_rerio/) identified the transcription start site at 281 bases upstream from the ATG translational initiation codon (Fig. 22).

Analysis of the 1526 bp sequence 5′ upstream of the transcription start site revealed a potential TATA box (Fig. 22). The TATA box element was reported in the human CRABPII gene (Åström *et al.*, 1992) but not in the mouse CRABPI gene (Wei *et al.*, 1990). An imperfect version of the retinoic acid-responsive element, DR-1

Figure 21. Determination of the transcription start site of zebrafish CRABPII gene by 5′ RLM-RACE. Total RNA from adult zebrafish was sequentially treated with calf intestinal phosphatase (CIP). One aliquot was treated with tobacco acid pyrophosphatase (TAP) and ligated to an RNA adapter (+TAP, lane 1). A second aliquot was not treated with TAP prior to RNA adapter ligation (-TAP, lane 2). Nested PCR with TAP-treated and TAP-untreated total RNA was performed with CRABPII specific primers and the adapter primers. A transcript of ~350 bp was amplified in the +TAP but not the –TAP reaction. The size marker used (L) was 100 bp from MBI Fermentas.

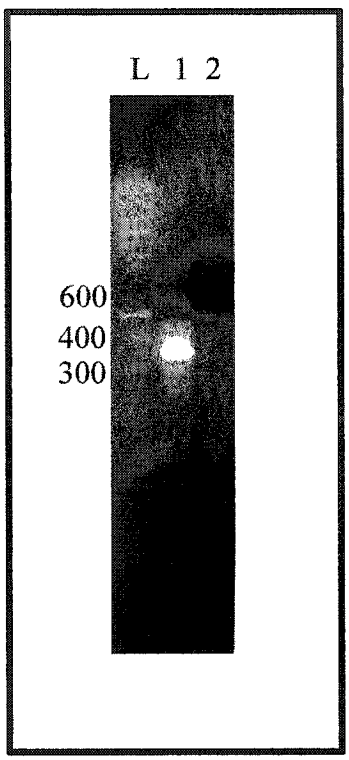


Figure 21. Determination of the transcription start site of zebrafish CRABPII gene by 5′ RLM-RACE.

Figure 22. The 5′ upstream sequence and putative *cis*-acting elements of the zebrafish CRABPII gene. The sequence 5′ of the translation initiation codon was obtained from contig, 15237.1. The transcription start site (arrowed) was mapped to 281 bp 5′ of the initiation codon by 5′ RLM-RACE. The sequence upstream of the transcription start site was analyzed for the putative *cis*-acting regulatory elements using MatInspector V2.2 (http://transfac.gbf.de/cgi-bin/matSearch/matsearch.pl). The putative *cis*-regulatory elements identified are underlined and are in bold font. The translation initiation codon is boxed and the sequence upstream of the transcription start site is in lower case.

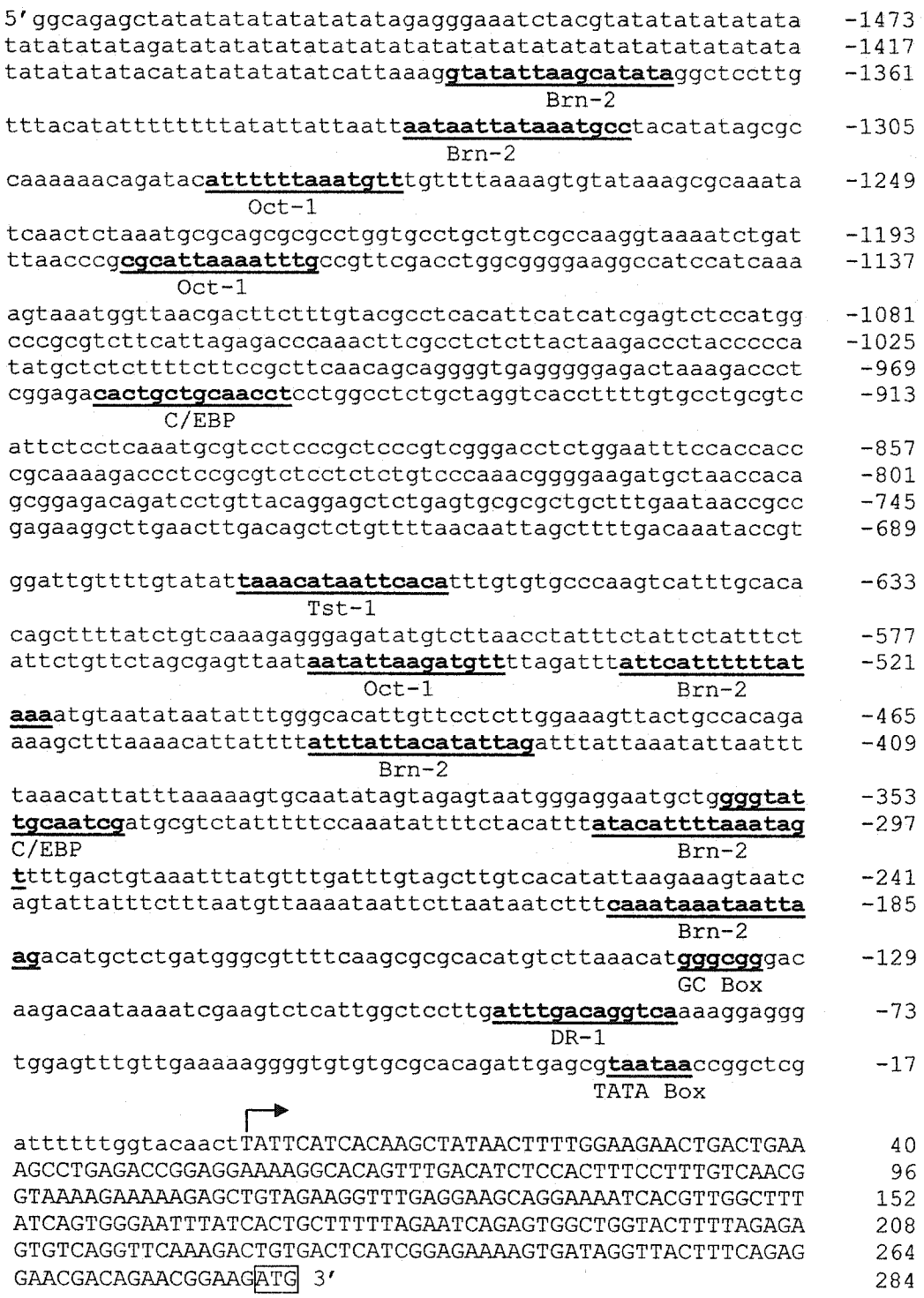


Figure 22. The 5′ upstream sequence and putative *cis*-acting elements of the zebrafish CRABPII gene.

(ATTTGAcAGGTCA) was identified in the upstream sequence of the zebrafish CRABPII gene. Sequence analysis using MatInspector V2.2 (http://transfac.gbf.de/cgi-bin/matSearch/matsearch.pl) revealed potential regulatory elements including one potential SP1 binding site (GC box) and two copies of C/EBP binding element (Fig. 22). Several POU-domain binding elements were present in the 5′ upstream sequence of the zebrafish CRABPII gene including five Brain-2 (Brn-2), three octamer-binding factor (Oct-1) and one Testis-1 (Tst-1) elements (Fig. 22). POU-domain transcription factors are widely expressed in the developing mammalian brain (He *et al.*, 1989) and are implicated in the regulation of central nervous system (CNS)-specific gene expression (Josephson *et al.*, 1998). The AP-2 transcription factor binding site present in the human CRABPII promoter (Åström *et al.*, 1992) was however, not present in the zebrafish CRABPII 5′ upstream sequence.

5.3.4 Northern blot hybridization

To determine the size of the zebrafish CRABPII mRNA transcript, northern blot and hybridization was employed. Hybridization of the zebrafish CRABPII cDNA to total RNA from adult zebrafish detected a transcript of ~2600 nucleotides (Fig. 23A). The size of the zebrafish transcript is larger than that reported for CRABP transcripts from other organisms with the exception of the *Xenopus* CRABP-b (the xCRABP-b transcript is approximately 3 kb) (Ho *et al*, 1994). While the coding regions for iLBPs are highly conserved, the transcripts for these CRABPs are large owing to the length of the 5′ and 3′ untranslated regions.

Figure 23. Northern blot-hybridization and RT-PCR analysis of CRABPII mRNA expression. **(A)** Northern blot analysis of total RNA isolated from whole adult zebrafish using the CRABP cDNA as a hybridization probe detected a single transcript of approximately 2600 nucleotides. **(B)** Top panel: CRABP-specific primers did not generate a product from total RNA of 1 hpf (1) and 3 hpf (3) embryos. CRABP-specific product was observed from total RNA of 12 hpf (12), 24 hpf (24), 51 hpf (51), 60 hpf (60) and 72 hpf (72) embryos and adult (A) zebrafish. A negative control (-) lacking cDNA did not generate any RT-PCR product. Bottom panel: RT-PCR product was generated from RNA in all samples for the constitutively expressed RACK1. A negative control (-) lacking cDNA did not generate any RT-PCR product. **(C)** Top panel: CRABP-specific primers produced a product from total RNA extracted from muscle (M), testes (T) and skin (S) and to a lesser extent heart (H), ovary (O) and brain (B). No product was detected in kidney (K), liver (L) and intestine (I). A negative control (-), lacking cDNA, did not generate any RT-PCR product. Bottom panel: RT-PCR product was generated from RNA in all tissue samples for the constitutively expressed RACK1. A negative control (-) lacking cDNA did not generate any RT-PCR product. **(D)** Quantitative RT-PCR was performed to determine the relative steady-state levels of zebrafish CRABPII mRNA in adult tissues. The histogram shows the ratio of CRABP mRNA to β-actin mRNA in various tissues with highest steady-state levels of CRABP mRNA seen in RNA extracted from adult muscle (M), followed by testes (T), skin (S), ovary (O) and almost undetectable levels in heart. CRABP mRNA was undetectable in RNA extracted from adult kidney (K), liver (L), intestine (I) and brain (B).

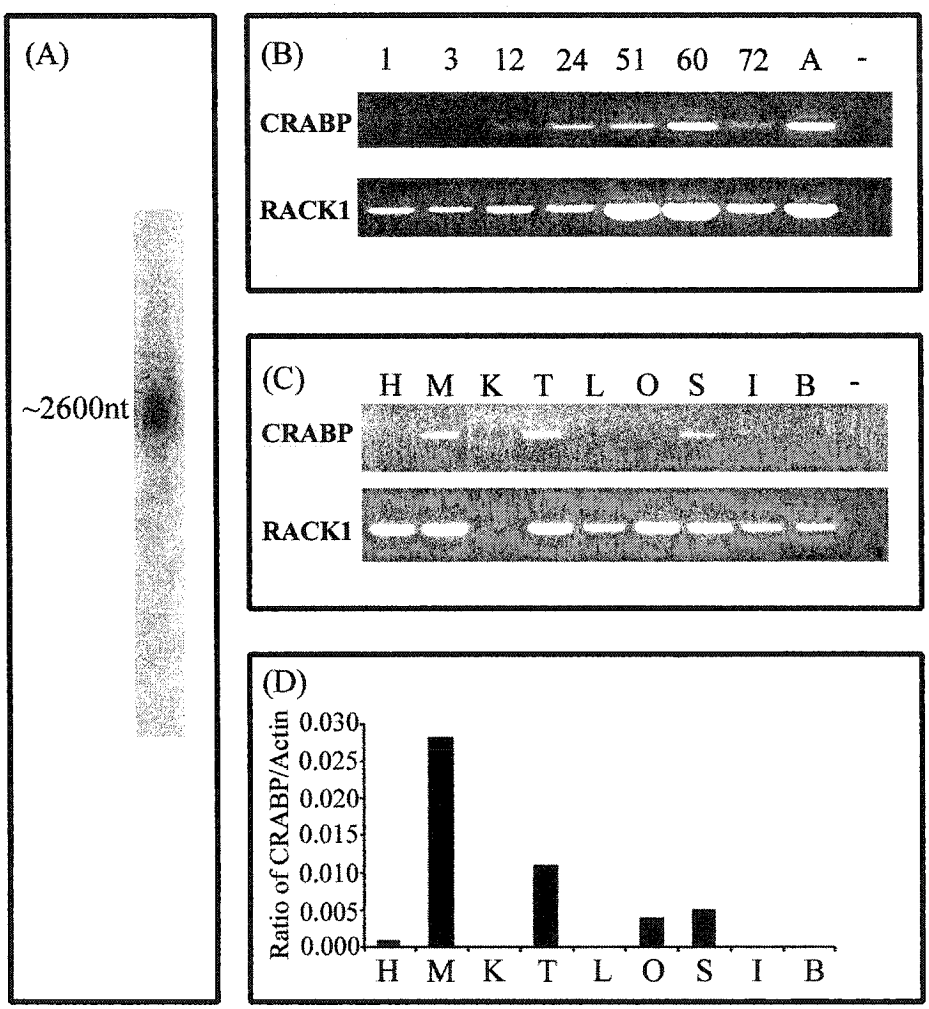


Figure 23. Northern blot-hybridization and RT-PCR analysis of CRABPII mRNA expression.

5.3.5 Tissue-specific expression of the CRABPII mRNA in adult zebrafish

RT-PCR was performed to determine the tissue distribution of the CRABPII mRNA in adult zebrafish. A CRABPII-specific product was generated by RT-PCR from total RNA extracted from muscle, testes, and skin, and a very faint band was detected following RT-PCR using total RNA from heart, ovary and brain (Fig. 23C). No RT-PCR product, however, was generated from RNA extracted from kidney, liver or intestine suggesting that the gene coding for this CRABPII mRNA was not expressed in these tissues in zebrafish. To estimate the relative levels of CRABPII mRNA in different zebrafish tissues and to confirm the tissue distribution of CRABPII mRNA revealed by the conventional RT-PCR, we performed quantitative RT-PCR of CRABPII mRNA from the same tissues using another constitutively expressed gene, β-actin, as a positive control. Levels of CRABPII mRNA in each cDNA sample ranged between undetectable to 1.3×10^3 copies per μL of cDNA. The β-actin RT-PCR product was amplified from each cDNA sample and ranged from 1.5×10^2 to 3.5×10^5 copies per μL of cDNA. The ratio of CRABP mRNA/β-actin RT-PCR was calculated for each experimental sample (Fig. 23D). This analysis showed that CRABPII mRNA was more abundant in muscle, testes and skin than in heart and ovary. CRABPII mRNA was barely detectable or undetectable in brain, kidney, liver and intestine. The CRABPII mRNA levels were approximately three, six, seven and thirty-five times higher in muscle than in testes, skin, ovary and heart, respectively. Both conventional RT-PCR and quantitative RT-PCR using two constitutively expressed mRNAs as positive controls, i.e. RACK1 and β-actin mRNA, showed similar patterns of tissue-distribution where the zebrafish CRABPII mRNA was abundant, but not in brain where the levels of CRABPII mRNA were very

low or undetectable. Amino acid sequence identity with known CRABPs from several species indicated that the zebrafish CRABPII is more closely related to CRABPII than to CRABPI (Fig. 19). In adult human, CRABPII mRNA is present in skin (Tavakkol *et al.*, 1992). In adult rat, CRABPII mRNA is detected in skin, ovary, brain and uterus (Bucco *et al.*, 1996). In adult mouse, CRABPII mRNA has also been localized in skin and at lower levels in brain (Giguere *et al.*, 1990). However, using the highly sensitive method of RT-PCR to examine a range of tissues, we found the zebrafish CRABPII mRNA distribution more extensive than that reported for mammalian CRABPII.

5.3.6 Spatio-temporal expression of the CRABPII gene during zebrafish embryogenesis.

RT-PCR of total RNA extracted from zebrafish at various embryonic stages revealed the temporal expression of the CRABPII gene during embryogenesis. CRABPII-specific product was not detected in total RNA from embryos at 1 hpf and 3hpf, but was detected at 12 h and stages thereafter (Fig. 23B).

The antisense DIG-labelled cRNA probe revealed the spatial expression of the zebrafish CRABPII gene during embryogenesis. At early gastrula stages, the CRABPII mRNA was detected in the posterior epiblast, adaxial epiblast (two longitudinal stripes on both sides of axial epiblast), presumptive epidermis, adaxial neuroectoderm and neuroectoderm (Fig. 24A, B). During late gastrula, bud stage (10 hpf), the CRABPII transcripts became more abundant in adaxial epiblast, adaxial neuroectoderm and posterior epidermis (Fig. 24C, D). The CRABPII mRNA expression was detected in the posterior neural plate (Fig. 24C).

Figure 24. Spatio-temporal distribution of zebrafish CRABPII mRNA during early embryonic development. Whole mount *in situ* hybridization showing zebrafish CRABPII mRNA in the embryos during early gastrula at approximately 6 hpf (A, B), then during late gastrula at 10 hpf (C, D), during early and middle somitogenesis at 11 hpf (E, F) and 17 hpf (G-O). A, C: dorsal view, anterior is to the top; B, D: side view, anterior is to the top; E, G, H, L: dorsal view, anterior is to the left; F, K, I, O: side view, anterior is to the left; J: frontal view showing forebrain (Fb) and retina (Re), dorsal is to top; M, N: transverse, dorsal is to top. The CRABPII mRNA expression in adaxial cells (Ac); adaxial epiblast (Ae); adaxial neuroectoderm (An); chordo-neural hinge (Cnh); diencephalon (Di); epidermis (Ep); floor plate (Fp); hindbrain (Hb); notochord (No); neural plate border (Npb); posterior epiblast (Pe); posterior epidermis (PEp); posterior neural plate (Pnp); rhombomere (r) 2, 4, 6, 7; roof plate (Rp); spinal chord (Sc); somites (So); telencephalon (Te) and ventral spinal chord (VSo) is indicated.

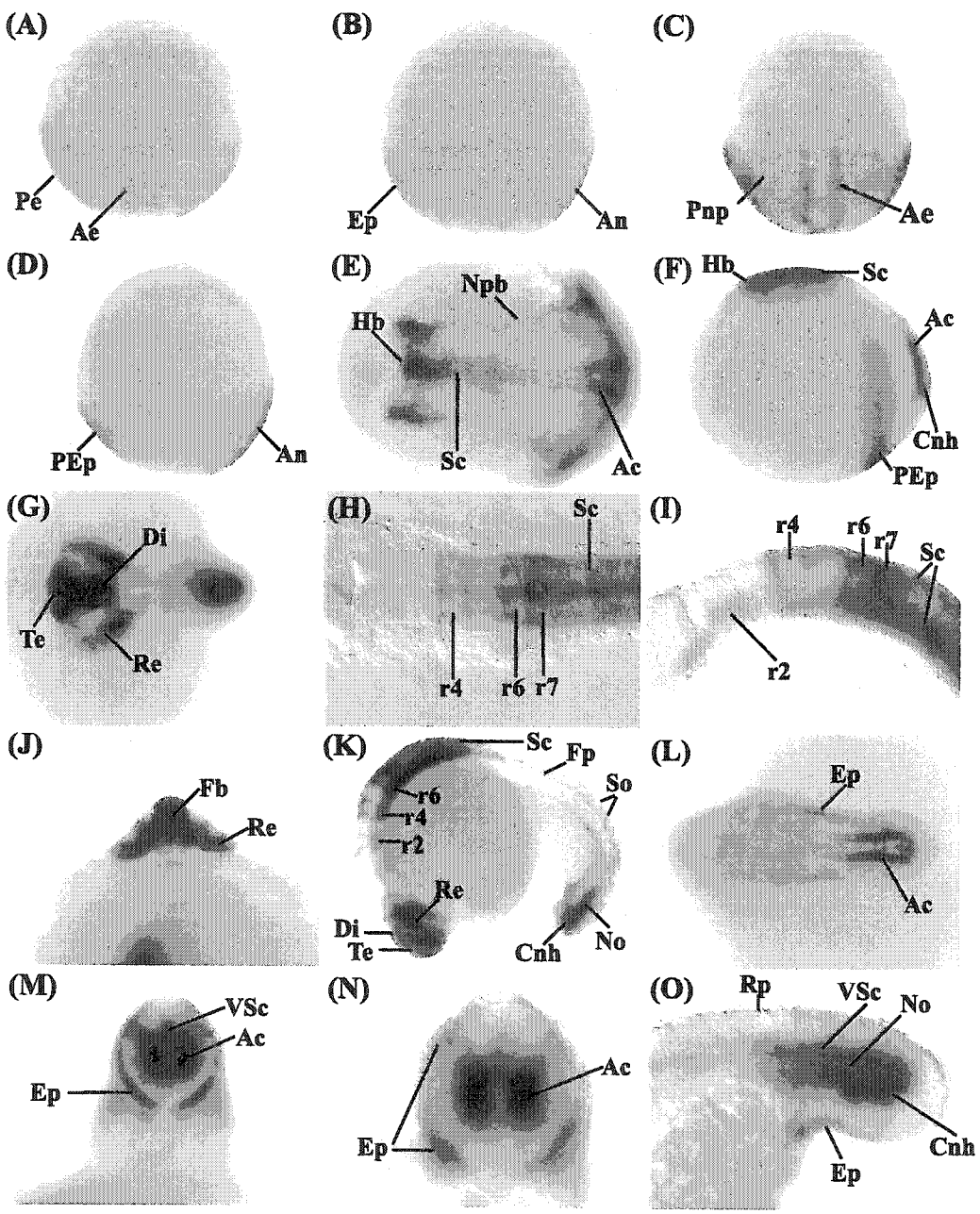


Figure 24. Spatio-temporal distribution of zebrafish CRABPII mRNA during early embryonic development.

During early stages of somitogenesis (11 hpf), CRABPII continued to be expressed in the adaxial cells (Fig. 24E, F) and the expression was also observed in the tail bud (chordo-neural hinge) (Fig. 24F). In the ectoderm, CRABPII mRNA expression was detected at the border of neural plate, posterior hindbrain, spinal chord and the epidermis (Fig. 24E, F). During mid-somitogenesis (17 hpf), the CRABPII gene was expressed in many parts of the developing central nervous system (CNS). The CRABPII mRNA-specific hybridization signal was detected in the presumptive forebrain (diencephalon and telencephalon) (Fig. 24G, J, K), hindbrain (weak in rhombomere 4, strong in rhombomere 6 and 7) (Fig. 24H, I, K), presumptive spinal chord (Fig. 24H, I, K, M, O), floor plate (Fig. 24K) and the retina (Fig. 24G, J, K). The CRABPII gene continued to express in the epidermis during mid-somitogenesis (Fig. 24L-O). In the mesoderm, CRABPII mRNA expression persisted in the adaxial cells and the tail bud (Fig. 24K-O). In the tail bud, CRABPII transcripts were detected in the notochord and the chordo-neural hinge (Fig. 24K, O).

At 24 hpf, CRABPII transcript continued to be detected in the zebrafish CNS. Zebrafish forebrain (Fig. 25A-D), proximal part of the retina (Fig. 25A, B, D), optic stalk (Fig. 25B, C), hindbrain and dorsal spinal chord (Fig. 25B, C, E) expressed CRABPII mRNA at 24 hpf. In the hindbrain, rhombomeres 2 and 4 had weak hybridization signal while the rhombomere 6 had a strong CRABPII-specific hybridization signal. CRABPII mRNA was detected in the chordo-neural hinge (Fig. 25B, H) and at much higher levels in the lateral somites (slow muscles fibres) (Fig. 25B, G, H) when compared to levels observed during mid somitogenesis (Fig. 24K). Weak CRABPII mRNA specific signal was detected in the branchial arches and the pectoral fin

Figure 25. Distribution of zebrafish CRABPII mRNA during pharyngula stage of embryonic development. Whole mount and LM section *in situ* hybridization showing zebrafish CRABPII mRNA in the embryos at 24 hpf (A-H), 36 hpf (I-M) and 48 hpf (N-U) stages. All results are from whole mount *in situ* except L, M and N. A, E, G, I, N, U: dorsal view, anterior is to left; B, C, F, H, J, K, P-R, T: side view, anterior is to left; D: frontal, dorsal is to top; O: frontal, anterior is to the left; L, M, S: transverse section, dorsal is to the top. The CRABPII mRNA expression in branchial arches (Ba); cerebellum (Ce); chordo-neural hinge (Cnh); dorsal diencephalon (DDi); diencephalon (Di); dorsal hindbrain (Dhb); dorsal spinal chord (Dsc); forebrain (Fb); foor plate (Fp); hindbrain (Hb); midbrain (Mb); notochord (No); oesophagus (Oe); optic nerve (On); optic stalk (Os); pectoral fin (Pf); proliferative zone of retina (PZRe); rhombomere (r) 4, 6; retina (Re); roof plate (Rp); spinal chord (Sc); somite(s) (So); telencephalon (Te); ventral mesoderm (Vm); ventral mesenchyme (VMe) and ventral somite (VSo) is indicated.

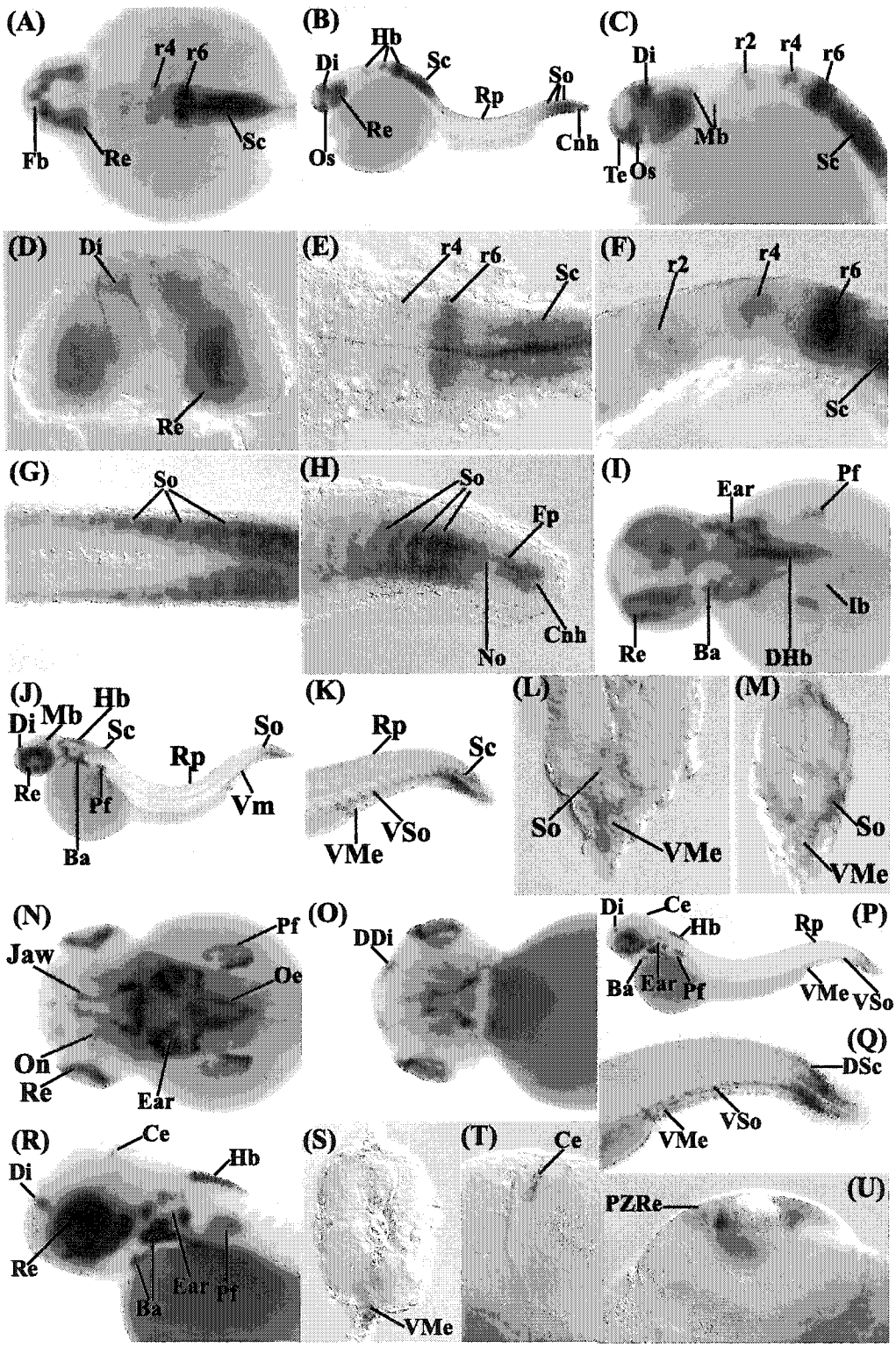


Figure 25. Distribution of zebrafish CRABPII mRNA during pharyngula stage of embryonic development.

bud of the 24 hpf zebrafish embryos (data not shown).

At 36 hpf, whole mount *in situ* hybridization detected CRABPII mRNA in the CNS and the musculature system. The CRABPII mRNA persisted in the dorsal spinal chord, dorsal hindbrain, retina, and the dorsal forebrain (Fig. 25I, J, K). The branchial arches (Fig. 25I, J), mesenchyme around the otic vesicle (inner ear) (Fig. 25I), ventral tail mesoderm (Fig. 25J), ventromedial region of somites (Fig. 25L, M), ventral mesenchyme (Fig. 25L, M) and the epidermis of the pectoral fin (Fig. 25I, J) expressed CRABPII mRNA at 36 hpf. CRABPII transcripts were also detected in the intestinal bulb (Fig. 25I).

Zebrafish embryos at 48 hpf continued to express CRABPII mRNA in the CNS specifically in the diencephalon, dorsal hindbrain (Fig. 25O, P, R), roof plate (Fig. 25P), dorsal spinal chord (Fig. 25Q) and the proliferative zone of retina (Fig. 25U). The CRABPII transcripts persisted in the ventral mesenchyme (Fig. 25S), epidermis of pectoral fin (Fig. 25N, P, R), branchial arches (Fig. 25P, R), ventral somites (Fig. 25P, Q) and the inner ear (Fig. 25N, P, R). At this stage CRABPII mRNA was also detected in the cerebellum (hindbrain) (Fig. 25P, R, T) oesophagus and the jaw (Fig. 25N) of the developing zebrafish embryo.

5.3.7 Radiation hybrid mapping of the zebrafish CRABPII gene

To localize the CRABPII gene to a zebrafish linkage group, radiation hybrid mapping using the LN 54 panel of radiation hybrids was performed (Hukriede *et al.*, 1999). The zebrafish CRABPII gene was assigned to linkage group16 at a distance of 7.69 cR from the marker Z3104 with a LOD of 14.2.

Schleicher *et al.* (1995) speculate that the family of iLBPs arose by at least fourteen gene duplications. Existence of CRABP in the invertebrate species, *Manduca sexta,* (Mansfield *et al.*, 1998) suggests that the CRABPs appeared during the evolution of the iLBP multigene family before the vertebrate/invertebrate divergence some 500 million years ago (Campbell *et al.*, 1999). Two types of CRABPs have been reported for most vertebrates including human (see for e.g. GenBank accession number NM_004378 and M68867), mouse (NM_013496 and M35523), rat (U23407 and Pravenec *et al.*, 1997) and *Xenopus* (S74933 and Dekker *et al.*, 1994). The presence of only one CRABP gene in an invertebrate species, but two in the vertebrate species indicates that the CRABPI and CRABPII genes arose by gene duplication from a common ancestral CRABP gene after the divergence of vertebrates and invertebrates, or one of the two CRABP genes was lost from the invertebrate lineage. Comparison of the mapped zebrafish CRABPII and human CRABPII gene revealed that zebrafish linkage group 16 has a syntenic relationship with human chromosome 1. The genes for pyruvate kinase liver and RBC (PKLR), cathepsin S (CTSS), CRABPII, KIAA0907 protein (KIAA0907) and chaperonin containing TCP1 subunit 3 (CCT3) are located on chromosome 1 in human and linkage group 16 in zebrafish (Table 7). The syntenic relationship between human chromosome 1 and zebrafish linkage group 16, and phylogenetic analysis, suggests that fish and mammals share a common ancestral CRABPII gene. Human CRABPI has been mapped to chromosome 15. Based on data available at ZFIN (http://zdb.wehi.edu.au:8282//) and published by Woods *et al.* (2000), no other gene located on zebrafish LG 16 is present on human chromosome 15. The higher percentage identity of the zebrafish CRABP to human CRABPII (74%) as compared to human CRABPI (69%) and no apparent syntenic

Table 7. Zebrafish-human conserved syntenies defined by the CRABPII gene

Zebrafish[1]			Human[2]	
LG	**Locus**	**Accession Number**	**Gene**	**Chromosomal position**
16	pklr	AW171121	PKLR	1q21
16	ctss	AA549805	CTSS	1q21
16	**crabpII**	AF497478	**CRABPII**	1q21.3
16	kiaa0907	AI415764	KIAA0907	1q21.3
16	cct3	AI476862	CCT3	1q23

[1] Woods *et al.*, 2000.

[2] LocusLink (http://www.ncbi.nlm.nih.gov/LocusLink/list.cgi), NCBI

relationship between zebrafish LG 16 and human chromosome 15 suggests that the zebrafish CRABP is more closely related to human CRABPII than the CRABPI. In summary, based on phylogenetic analysis and syntenic relationship to the CRABPII gene in human, the zebrafish cDNA clone appears to code for a type II CRABP. However, the CRABPII mRNA distribution in the CNS of the developing zebrafish embryos and the mouse embryos are comparable but differences in the expression pattern in the epidermis and the liver was observed (Ruberte *et al.*, 1992; Lyn and Giguère, 1994). In addition, using the highly sensitive method of RT-PCR and by examining a range of adult tissues, we found the tissue-distribution of the zebrafish CRABPII mRNA was more extensive than that reported for mammalian CRABPII.

Chapter 6: Concluding remarks

Multigene families may hold the key to the evolution of organismal complexity. It is believed that gene duplications, subsequent divergence and gene loss may have resulted in the many species of organisms that inhabit earth today (Lynch and Conery, 2000; Lynch and Force, 2000; Taylor *et al.*, 2001a). I studied four members of iLBP multigene family in the zebrafish with the hope of gaining insight into the evolution of the iLBP multigene family.

The four members of the iLBP multigene family studied here were mapped to LG 1 (I-FABP), LG 16 (Lb-FABP and CRABPII) and LG 5 (L-FABP). In the human genome, I-FABP has been mapped on chromosome 4, CRABPII on chromosome 1 and L-FABP on chromosome 2 (Bernlohr *et al.*, 1997; Flagiello *et al.*, 1997). The ortholog of Lb-FABP has not been found in the human genome. The dispersal of iLBP genes on different linkage groups in the zebrafish genome and different chromosomes in the human genome suggests that unlike the different members of α- and β- globin multigene family (Burt, 2002) or *hox* genes and others, tandem or local duplication did not play a major role in creating different members of iLBP multigene family. Although Lb-FABP and CRABPII are localized on the same linkage group in zebrafish, it is highly unlikely that they arose directly as a result of tandem duplication. The amino acid composition and the number of amino acids in the zebrafish L-FABP and Lb-FABP suggest that they are more closely related to each other than to CRABPII. The zebrafish CRABPII is 142 amino acids in length while L-FABP and Lb-FABP are 127 and 126 amino acids, respectively. Moreover, in a phylogenetic analysis, zebrafish Lb-FABP and other known Lb-FABPs clusters in a major clade which includes L-FABP and IlLBPs, but not the

CRABPs (Fig. 14). This analysis suggests that Lb-FABP is phylogenetically more closely related to L-FABP and IlLBP than to CRABPs and may have arisen as a result of duplication of the L-FABP or IlLBP genes and not a CRABPII gene. Taken together, these data suggest that tandem gene duplication is not the primary mechanism in the evolution of iLBP multigene family, but more likely, this multigene family arose by genome duplication(s) and/or chromosomal duplication events.

The exon/intron structure for the mammalian iLBPs is defined as four exons separated by three introns (Bernlohr et al., 1997). In invertebrates, the muscle-type FABP gene of desert locust and *Drosophila*, have a three exon-two intron organization as a result of loss of intron 2 that is present in the mammalian iLBP genes (Wu *et al.*, 2001). I analyzed the organization of the iLBP genes in zebrafish to understand whether the second intron was acquired by the mammalian iLBPs during the course of evolution or the intron 2 was lost independently in the insect lineage. All four of the zebrafish iLBP genes analyzed (Fig. 1, 6, 12, 18) showed an intron in a position comparable to the location of intron 2 in the mammalian iLBPs. This suggests that either intron 2 was lost from the insect lineage or was acquired in the vertebrate lineage before the divergence of ray-finned fish lineage from the lineage leading to mammals approximately 450 mya (Kumar and Hedges, 1998). The L-FABP in zebrafish, however, has an additional intron in the 5′ UTR that has not been reported for the mammalian L-FABP gene (Sweetser *et al.*, 1986). This additional intron, therefore, seems to be the result of an insertional mutation either in the finned-fish lineage or just in zebrafish. Determination of the L-FABP gene sequences from other actinopterygians will be helpful in determining the timing of this change as it relates to iLBP gene organization during vertebrate evolution.

Different theories are offered to explain the fates of duplicated genes (Prince and Pickett, 2002; Force *et al.*, 1999; Lynch and Conery, 2000). Immediately after the duplication event, duplicate genes have redundant function, and three different fates are suggested for duplicated gene pairs:

(1) *Nonfunctionalization.* One gene copy becomes silenced by degenerative mutations or is lost from the genome owing to vagaries of chromosomal remodeling, locus deletion or point mutations. Gene loss through these processes is possible because only one of the duplicates is required to maintain the function provided by the single, ancestral gene, leaving one gene copy free to accumulate loss-of-function mutations (Prince and Pickett, 2002).

(2) *Neofunctionalization.* One copy acquires a novel, beneficial function, because of alterations in coding or regulatory sequences while the other copy retains the original function. Mutations that lead to neofunctionalization are assumed to be extremely rare, so this model predicts that few duplicates should be retained in the genome over the long term (Prince and Pickett, 2002).

(3) *Subfunctionalization.* Both gene copies accumulate mutation to the point at which their total activity represents the combined function of the single-copy ancestral gene i.e. both the genes share the function of the ancestral gene (Force *et al.*, 1999; Lynch and Conery, 2000).

Both duplicates can be retained in a genome as a result of subfunctionalization or neofunctionalization. Force and co-workers (1999) proposed that subfunctionalization is the usual mechanism of duplicate gene retention. Since, many duplicated members of the iLBP multigene family are retained in vertebrate genomes, analysis of their expression

pattern may help define the possible mechanism of their retention. The expression patterns of specific iLBP genes differ between mammalian and non-mammalian species. The L-FABP was not expressed during embryonic development in zebrafish (Fig. 16A) and in adult zebrafish L-FABP expression was restricted to intestine (Fig. 16B). In contrast, mammalian L-FABP is expressed in the liver and intestine of the developing embryos and adults (Veerkamp and Maatman, 1995). A paralogous liver-type FABP, Lb-FABP found only in non-mammalian species and believed to share a common ancestral gene with the L-FABP (Schaap *et al.*, 2002) is expressed in the zebrafish liver, intestine and testes (Fig. 9). The expression of Lb-FABP in liver and intestine of adult zebrafish is more similar to the expression pattern of the mammalian L-FABP gene. In salamander, L-FABP is expressed in the liver and the intestine and Lb-FABP is expressed in the liver (Di Pietro *et al.*, 1999). Western blot analysis revealed that the catfish Lb-FABP is expressed in the liver, but not in the heart, brain, intestinal mucosa, muscle and skin. The catfish L-FABP is expressed only in the intestinal mucosa (Di Pietro *et al.*, 1997). With the assumption that the mammalian L-FABP represents the ancestral state of the L-FABP gene from which Lb-FABP arose as a result of gene duplication, it seems that the pattern of expression of the mammalian L-FABP is retained in the liver and intestine of the amphibian, salamander (Di Pietro *et al.*, 1999). The duplicated Lb-FABP gene expression is restricted to liver, possibly due to division of function (subfunctionalization) between the L-FABP and Lb-FABP in the salamander liver. In catfish, however, the L-FABP and Lb-FABP are expressed in one of the two tissues (Di Pietro *et al.*, 1997). This division of expression pattern could represent the subfunctionalization model of duplicate gene retention in catfish. In zebrafish, however,

the expression pattern of the Lb-FABP in the liver and intestine is more similar to the L-FABP expression in amphibians and mammals. The expression of Lb-FABP in zebrafish testis could represent neofunctionalization as a possible mode of duplicate Lb-FABP gene retention in zebrafish.

The finned-fish lineage is believed to have undergone a whole genome duplication after divergence from the lobe finned-fish lineage (Amores *et al.*, 1998; Wittbrodt *et al.*, 1998; Taylor *et al.*, 2001b; Van de Peer *et al.*, 2003). It is expected, therefore, that zebrafish had two genes for each of the mammalian orthologs at the time of whole genome duplication. Database searches at the NCBI and the Wellcome Trust Sanger Institute revealed the putative second copy of L-FABP (L-FABPb) and CRABPII (CRABPIIb) genes in the zebrafish genome (data not shown). The sequence for the duplicated copy of L-FABP (L-FABPb) was obtained from an EST submitted to GenBank (accession number BQ075349, NCBI) and the sequence for the duplicated copy of CRABPII (CRABPIIb) was obtained from genomic contig, ctg9369 (assembly Zv2, http://www.sanger.ac.uk/Projects/ D_rerio/). No sequence representing the second copy of the zebrafish Lb-FABP and I-FABP was found in the database so far. Postlethwait and co-workers (2000) estimated that approximately 20% of the duplicated gene pairs were retained in the zebrafish lineage after the whole genome duplication event. It is possible that the second copies of I-FABP and Lb-FABP genes accumulated null mutation (nonfunctionalization) and were lost from the zebrafish genome.

Preliminary investigation using the primers designed on the basis of the available L-FABPb and CRABPIIb sequences mapped the duplicated members to zebrafish LG 8 (L-FABPb) and LG 19 (CRABPIIb) by radiation hybrid mapping (data not shown). The

RT-PCR analysis revealed that L-FABPb is expressed in the developing zebrafish embryos and like zebrafish L-FABP the expression of L-FABPb is restricted to the intestine of adult zebrafish (data not shown). It seems that L-FABPb was retained in the zebrafish genome as a result of subfunctionalization because L-FABPb mRNA, but not L-FABP mRNA, was detected in the developing zebrafish embryos. Similar to CRABPII, the CRABPIIb mRNA was not detected in total RNA from zebrafish embryos at 1 hpf and 3 hpf, but was detected at 12 hpf and stages thereafter. Unlike CRABPII, the CRABPIIb mRNA was not detected in any of the adult zebrafish tissues investigated including heart, muscle, testes, liver, ovary, skin, intestine or brain. The CRABPIIb mRNA-specific product was, however, detected in total RNA extracted from adult zebrafish (data not shown). The retention of both CRABPII and CRABPIIb in the zebrafish genome may be a result of subfunctionalization as CRABPII is expressed in the muscle, testis, skin, heart, ovary and brain and the CRABPIIb is expressed in the adult zebrafish, but not in any of these tissues.

Although this thesis provides insight into the evolution of the iLBP multigene family, many questions remain. Future studies may include, characterization of the duplicated copies of the iLBP genes in zebrafish, functional analysis of the iLBP gene promoters to delineate the different *cis*-acting regulatory elements, identification of various hydrophobic ligands that are involved in the regulation of iLBP genes in zebrafish and identification of the transcription factors that bind to the *cis*-acting elements in the iLBP promoters.

Summary

1. The dispersed distribution of iLBP genes in the zebrafish and human genomes indicated that tandem duplication did not play a major role in creating the iLBP multigene family. More likely, this multigene family arose by genome duplication(s) and/or chromosomal duplication events.
2. The expression pattern and potential function of iLBP genes seems to be conserved between mammals and fish in some tissues but not in others.
3. Changes in expression pattern reflect changes in the factors regulating the iLBP genes such as the *cis*-elements in their gene promoters.
4. The proposed whole genome duplication event in the finned-fish lineage is supported by the presence of a second copy of the CRABPII and L-FABP genes.
5. Subfunctionalization most likely explains the retention of the duplicated iLBP genes.

References

Akiyama, T.E., Ward, J.M., Gonzalez, F.J., 2000. Regulation of the liver fatty acid-binding protein gene by hepatocyte nuclear factor 1α (HNF1α). Alterations in fatty acid homeostasis in HNF1α-deficient mice. J. Biol. Chem. 275, 27117-27122.

André, M., Ando, S., Ballagny, C., Durliat, M., Poupard, G., Briançon, C., Babin, P.J., 2000. Intestinal fatty acid binding protein gene expression reveals the cephalocaudal patterning during zebrafish gut morphogenesis. Int. J. Dev. Biol. 44, 249-252.

Ang, S-L., Rossant, J., 1994. HNF-3 beta is essential for node and notochord formation in mouse development. Cell 78, 561-574.

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. and Postlethwait, J.H., 1998. Zebrafish *hox* clusters and vertebrate genome evolution. Science 282, 1711-1714.

Aparicio, S., Chapman, J., Stupka, E., Putnam, N., Chia, J.M., Dehal, P., Christoffels, A., Rash, S., Hoon, S., Smit, A., Gelpke, M.D., Roach, J., Oh, T., Ho, I.Y., Wong, M., Detter, C., Verhoef, F., Predki, P., Tay, A., Lucas, S., Richardson, P., Smith, S.F., Clark, M.S., Edwards, Y.J., Doggett, N., Zharkikh, A., Tavtigian, S.V., Pruss, D., Barnstead, M., Evans, C., Baden, H., Powell, J., Glusman, G., Rowen, L., Hood, L., Tan, Y.H., Elgar, G., Hawkins, T., Venkatesh, B., Rokhsar, D., Brenner, S., 2002. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. Science 297, 1301-1310.

Åström, A, Pettersson, U, Voorhees, J.J., 1992. Structure of the human cellular retinoic acid-binding protein II gene. Early transcriptional regulation by retinoic acid. J. Biol. Chem. 267, 25251-25255.

Baier, L.J., Bogardus, C., Sacchettini, J.C., 1996. A polymorphism in the human intestinal fatty acid binding protein alters fatty acid transport across Caco-2 cells. J. Biol. Chem. 271, 10892-10896.

Beringhelli, T., Goldoni, L., Capaldi, S., Bossi, A., Perduca, M., Monaco, H,L., 2001. Interaction of chicken liver basic fatty acid-binding protein with fatty acids: a 13C NMR and fluorescence study. Biochemistry 40, 12604-12611.

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.

Besnard, P., Niot, I., Poirier, H., Clément, L., Bernard, A., 2002. New insights into the fatty acid-binding protein (FABP) family in the small intestine. Mol. Cell Biochem. 239, 139-147.

Börchers, T, Spener, F, 1994. Fatty acid binding proteins, p. 261-294. In Current topics in membranes, vol. 40, (eds. Kleinzeller, A., Benos, D. J.), Academic Press, New York.

Boullu-Sanchis, S., Lepretre, F., Hedelin, G., Donnet, J.P., Schaffer, P., Froguel, P., Pinget, M., 1999. Type 2 diabetes mellitus: association study of five candidate genes in an Indian population of Guadeloupe, genetic contribution of *FABP2* polymorphism. Diabetes Metab. 25, 150-156.

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

Bucco, R.A., Zheng, W.L., Wardlaw, S.A., Davis, J.T., Sierra-Rivera, E., Osteen, K.G., Melner, M.H., Kakkad, B.P., Ong, D.E., 1996. Regulation and localization of cellular retinol-binding protein, retinol-binding protein, cellular retinoic acid-binding protein (CRABP), and CRABP II in the uterus of the pseudopregnant rat. Endocrinology 137, 3111-3122.

Burt, D.W., 2002. Encyclopedia of life sciences, p. 1-5. Macmillan Publishers Ltd., Nature Publishing Group (www.els.net).

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.

Campbell, N.A., Reece, J.B., Mitchell, L.G., 1999. Biology, 5th Edition. Addison Wesley Longman, Inc., CA.

Carleton, K.L., Streelman, J.T., Lee, B.-Y., Garnhart, N., Kidd, M., Kocher, T.D., 2002. Rapid isolation of CA microsatellites from tilapia genome. Animal Genetics 33, 140-144.

Ceciliani, F., Monaco, H.L., Ronchi, S., Faotto, L., Spadon, P., 1994. The primary structure of a basic (p*I* 9.0) fatty acid-binding protein from liver of *Gallus domesticus*. Comp. Biochem. Physiol. 109B, 261-271.

Chen, X., Tordova, M., Gilliland, G.L., Wang, L., Li, Y., Yan, H., Ji, X., 1998. Crystal structure of apo-cellular retinoic acid-binding protein type II (R111M) suggests a mechanism of ligand entry. J. Mol. Biol. 278, 641-653.

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

Córdoba, O.L., Sánchez, E.I., Santomé, J.A., 1999. The main fatty acid-binding protein in the liver of the shark (*Halaetunus bivius*) belongs to the liver basic type. Isolation, amino acid sequence determination and characterization. Eur. J. Biochem. 265, 832-838.

Dekker, E-J., Vaessen, M-J., van den Berg, C., Timmermans, A., Godsave, S., Holling, T., Nieuwkoop, P., Geurts van Kessel, A., Durston, A., 1994. Overexpression of a cellular retinoic acid binding protein (xCRABP) causes anteroposterior defects in developing *Xenopus* embryos. Development 120, 973-985.

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., Santomé, 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.

Di Pietro, S.M., Santomé, J.A., 2001. Structural and biochemical characterization of the lungfish (*Lepidosiren paradoxa*) liver basic fatty acid binding protein. Arch. Biochem. Biophys. 388, 81-90.

Di Pietro, S.M., Veerkamp, J.H., Santomé, J.A., 1999. Isolation, amino acid sequence determination and binding properties of two fatty-acid-binding proteins from axolotl (*Ambistoma mexicanum*) liver. Evolutionary relationship. Eur. J. Biochem. 259, 127-134.

Di Pietro, S.M., Corsico, B., Perduca, M., Monaco, H.L., Santomé, J.A., 2003. Structural and biochemical characterization of toad liver fatty acid-binding protein. Biochemistry. 42, 8192-8203.

Egan, C.R., Chung, M.A., Allen, F.L., Heschl, M.F., Van Buskirk, C.L., McGhee, J.D., 1995. A gut-to-pharynx/tail switch in embryonic expression of the *Caenorhabditis elegans* ges-1 gene centers on two GATA sequences. Dev. Biol. 170, 397-419.

Fang, R., Olds, L.C., Santiago, N.A., Sibley, E., 2001. GATA family transcription factors activate lactase gene promoter in intestinal Caco-2 cells. Am. J. Physiol. Gastrointest. Liver Physiol. 280, G58-G67.

Felsenstein, J., 1995. PHYLIP (phylogeny inference package). Version 3.5c. Distributed by the author, Department of Genetics, University of Washington, Seattle.

Fiorella, P.D., Napoli, J.L., 1991. Expression of cellular retinoic acid binding protein (CRABP) in *Escherichia coli*. Characterization and evidence that holo-CRABP is a substrate in retinoic acid metabolism. J. Biol. Chem., 266, 16572-16579.

Flagiello, D., Apiou, F., Gibaud, A., Poupon, M.F., Dutrillaux, B., Malfoy, B., 1997. Assignment of the genes for cellular retinoic acid binding protein 1 (CRABP1) and 2 (CRABP2) to human chromosome band 15q24 and 1q21.3, respectively, by *in situ* hybridization. Cytogenet. Cell Genet. 76, 17-18.

Fogh, K., Voorhees, J.J., Åström, A., 1993. Expression, purification, and binding properties of human cellular retinoic acid binding protein type I and type II. Arch. Biochem. Biophys. 300, 751-755.

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

Gao, X., Sedgwick, T., Shi, Y.B., Evans, T., 1998. Distinct functions are implicated for the GATA-4, -5, and -6 transcription factors in the regulation of intestine epithelial cell differentiation. Mol. Cell Biol. 18, 2901-2911.

Ghyselinck, N.B., Båvik, C., Sapin, V., Mark, M., Bonnier, D., Hindelang, C., Dierich, A., Nilsson, C.B., Hakansson, H., Sauvant, P., Azaïs-Braesco, V., Frasson, M., Picaud, S., Chambon, P., 1999. Cellular retinol-binding protein I is essential for vitamin A homeostasis. EMBO J. 18, 4903-4914.

Giguère, V., Lyn, S., Yip, P., Siu, C-H., Amin, S., 1990. Molecular cloning of cDNA encoding a second cellular retinoic acid-binding protein. Proc. Natl. Acad. Sci. USA 87, 6233-6237.

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

Gordon, J.I., Elshourbagy, N., Lowe, J.B., Liao, W.S., Alpers, D.H., Taylor, J.M., 1985. Tissue specific expression and developmental regulation of two genes coding for rat fatty acid binding proteins. J. Biol. Chem. 260, 1995-1998.

Gorry, P., Lufkin, T., Dierich, A., Rochette-Egly, C., Décimo, D., Dollé, P., Mark, M., Durand, B., Chambon, P., 1994. The cellular retinoic acid binding protein I is dispensable. Proc. Natl. Acad. Sci. USA 91, 9032-9036.

Green, R.P., Cohn, S.M., Sacchettini, J.C., Jackson, K.E., Gordon, J.I., 1992. The mouse intestinal fatty acid binding protein gene: nucleotide sequence, pattern of developmental and regional expression, and proposed structure of its protein product. DNA Cell Biol. 11, 31-41.

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

Haunerland, N., Jagschies, G., Schulenberg, H. and Spener, F., 1984. Fatty-acid-binding proteins. Occurrence of two fatty-acid-binding proteins in bovine liver cytosol and their binding of fatty acids, cholesterol, and other lipophilic ligands. Hoppe Seyler's Z. Physio.l Chem. 365, 365-376.

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(6228), 35-41.

Hegele, R.A., Harris, S.B., Hanley, A.J., Sadikian, S., Connelly, P.W., Zinman, B., 1996. Genetic variation of intestinal fatty acid-binding protein associated with variation in body mass in aboriginal Canadians. J. Clin. Endocrinol. Metab. 81, 4334-4337.

Her, G.M., Chiang, C.C., Chen, W.Y., Wu, J.L., 2003a. *In vivo* studies of liver-type fatty acid binding protein (L-FABP) gene expression in liver of transgenic zebrafish (*Danio rerio*). FEBS Letters 538, 125-133.

Her, G.M., Yeh, Y.H., Wu, J.L., 2003b. 435-bp liver regulatory sequence in the liver fatty acid binding protein (L-FABP) gene is sufficient to modulate liver regional expression in transgenic zebrafish. Dev. Dyn. 227, 347-356.

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.

Ho, L., Mercola, M., Gudas, L.J., 1994. *Xenopus laevis* cellular retinoic acid-binding protein: temporal and spatial expression pattern during early embryogenesis. Mech. Dev. 47, 53-64.

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

Ibabe, A., Grabenbauer, M., Baumgart, E., Fahimi, H.D., Cajaraville, M.P., 2002. Expression of peroxisome proliferator-activated receptors in zebrafish (*Danio rerio*). Histochem. Cell Biol. 118, 231-239.

Inagami, S. and Ong, D.E., 1992. Purification and partial characterization of cellular retinol-binding protein, type two, from human small intestine. J. Nutr. 122, 450-456.

Ishaq, M., Zhang, Y.M., Natarajan, V., 1998. Activation-induced down-regulation of retinoid receptor RXRalpha expression in human T lymphocytes. Role of cell cycle regulation. J. Biol. Chem. 273, 21210-21216.

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.

Kastner, P., Mark, M., Chambon, P., 1995. Nonsteroid nuclear receptors: what are genetic studies telling us about their role in real life? Cell 83, 859-869.

Kleinjan, D.A., Dekker, S., Guy, J.A., Grosveld,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.

Kleywegt, G.J., Bergfors, T., Senn, H., Le Motte, P., Gsell, B., Shudo, K. and Jones, T.A., 1994. Crystal structures of cellular retinoic acid-binding proteins I and II in complex with all-*trans*-retinoic acid and a synthetic retinoid. Structure 2, 1241-1258.

Kumar, S. and Hedges, S.B., 1998. A molecular timescale for vertebrate evolution. Nature 392, 917-920.

Lampron, C., Rochette-Egly, C., Gorry, P., Dollé, P., Mark, M., Lufkin, T., LeMeur, M., Chambon, P., 1995. Mice deficient in cellular retinoic acid binding protein II (CRABPII) or in both CRABPI and CRABPII are essentially normal. Development. 121, 539-548.

Levin, M.S., Locke, B., Yang, N.C., Li, E., Gordon, J.I., 1988. Comparison of the ligand binding properties of two homologous rat apocellular retinol-binding proteins expressed in *Escherichia coli*. J. Biol. Chem. 263, 17715-17723.

Levy, E., Menard, D., Delvin, E., Stan, S., Mitchell, G., Lambert, M., Ziv, E., Feoli-Fonseca, J.C., Seidman, E., 2001. The polymorphism at codon 54 of the *FABP2* gene increases fat absorption in human intestinal explants. J. Biol. Chem. 276, 39679-39684.

Lewin, B., 2000. Genes VII, p. 627-632. Oxford University Press Inc., NY, USA.

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.

Lyn, S., Giguère, V., 1994. Localization of CRABP-I and CRABP-II mRNA in the early mouse embryo by whole-mount *in situ* hybridization: implications for teratogenesis and neural development. Dev Dyn. 199, 280-291.

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

Lynch, M., and Force, A., 2000. The origin of interspecific genomic incompatibility via Gene duplication. Am. Nat. 156:590-605.

Lyons, S.E., Shue, B.C., Lei, L., Oates, A.C., Zon, L.I., Liu, P.P., 2001. Molecular cloning, genetic mapping, and expression analysis of four zebrafish *c/ebp* genes. Gene 281, 43-51.

Mangelsdorf, D.J., Umesono, K., Evans, R.M., 1994. The retinoid receptors. In: Sporn, M.B., Roberts, A.B., Goodman, D.S. (Eds.). The retinoids: biology, chemistry and medicine, 2nd Edition, Raven Press, NY, pp. 319-349.

Mansfield, S.G., Cammer, S., Alexander, S.C., Muehleisen, D.P., Gray, R.S., Tropsha, A., Bollenbacher, W.E., 1998. Molecular cloning and characterization of an invertebrate cellular retinoic acid binding protein. Proc. Natl. Acad. Sci. USA 95, 6825-6830.

Mayer, A., Mosler, G., Just, W., Pilgrim, C., Reisert, I., 2000. Developmental profile of *Sry* transcripts in mouse brain. Neurogenetics 3, 25-30.

McArthur, M.J., Atshaves, B.P., Frolov, A., Foxworth, W.D., Kier, A.B., Schroeder, F., 1999. Cellular uptake and intracellular trafficking of long chain fatty acids. J. Lipid Res. 40, 1371-1383.

Miller, K.R. and Cistola, D.P., 1993. Titration calorimetry as a binding assay for lipid-binding proteins. Mol. Cell Biochem. 123, 29-37.

Nemecz, G., Hubbell, T., Jefferson, J.R., Lowe, J.B., Schroeder, F., 1991. Interaction of fatty acids with recombinant rat intestinal and liver fatty acid-binding proteins. Arch. Biochem. Biophys. 286, 300-309.

Ober, E.A., Field, H.A., Stainier, D.Y., 2003. From endoderm formation to liver and pancreas development in zebrafish. Mech. Dev. 120, 5-18.

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.

Odani, S., Baba, K., Tsuchida, Y., Aoyagi, Y., Wakui, S., Takahashi, Y., 2001. Hepatic fatty acid-binding proteins of a teleost, *Lateolabrax japonicus*. The primary structures and location of a disulfide bond. J. Biochem. 129, 69-76.

Ong, D.E., Chytil, F., 1978. Cellular retinoic acid-binding protein from rat testis. Purification and characterization. J. Biol. Chem. 253, 4551-4554.

Ong, D.E., Newcomer, M.E., Chytil, F., 1994. Cellular retinoid-binding proteins. In: Sporn, M.B., Roberts, A.B., Goodman, D.S. (Eds.). The retinoids: biology, chemistry and medicine, 2nd Edition, Raven Press, NY, pp. 283-317.

O'Rielly, P.T., Hamilton, L.C., McConnell, S.K., Wright, J.M., 1996. Rapid analysis of genetic variation in atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. Can. J. Fish. Aquat. Sci. 53, 2292-2298.

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.

Pravenec, M., Křen, V., Wang, J.-M., Kurtz, T.W., 1997. Linkage mapping of the cellular retinoic acid-binding protein 1 (Crabp1) gene to rat chromosome 8. Mamm. Genome 8, 455-456.

Prince, V.E. and Pickett, F.B., 2002. Splitting pairs: the diverging fates of duplicated genes. Nat. Rev. Genet. 3, 827-837.

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.

Raza, H., Pongubala, J.R. and Sorof, S., 1989. Specific high affinity binding of lipoxygenase metabolites of arachidonic acid by liver fatty acid-binding protein. Biochem. Biophys. Res. Commun. 161, 448-455.

Richieri, G.V., Ogata, R.T., Kleinfeld, A.M., 1994. Equilibrium constants for the binding of fatty acids with fatty acid-binding proteins from adipocyte, intestine, heart, and liver measured with the fluorescent probe ADIFAB. J. Biol. Chem. 269, 23918-23930.

Rottman, J.N., Gordon, J.I., 1993. Comparison of the patterns of expression of rat intestinal fatty acid binding protein/human growth hormone fusion genes in cultured intestinal epithelial cell lines and in the gut epithelium of transgenic mice. J. Biol. Chem. 268, 11994-12002.

Ruberte E, Friederich V, Morriss-Kay G, Chambon P., 1992. Differential distribution patterns of CRABP I and CRABP II transcripts during mouse embryogenesis. Development. 115, 973-987.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Scapin, G., Spadon, P., Pengo, L., Mammi, M., Zanotti, G., Monaco, H.L., 1988. Chicken liver basic fatty acid-binding protein (p*I* = 9.0). Purification, crystallization and preliminary X-ray data. FEBS Letters 240, 196-200.

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.

Schievano, E., Quarzago, D., Spadon, P., Monaco, H.L., Zanotti, G., Peggion, E., 1994. Conformational and binding properties of chicken liver basic fatty acid binding protein in solution. Biopolymers 34, 879-887.

Schleicher, C.H., Córdoba, 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.

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

Shi, J., Blundell, T.L., Mizuguchi, K., 2001. FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J. Mol. Biol. 310, 243-257.

Simon, T.C., Roth, K.A., Gordon, J.I., 1993. Use of transgenic mice to map *cis*-acting elements in the liver fatty acid-binding protein gene (*Fabpl*) that regulate its cell lineage-specific, differentiation-dependent, and spatial patterns of expression in the gut epithelium and in the liver acinus. J. Biol. Chem. 268, 18345-18358.

Sorof, S., 1994. Modulation of mitogenesis by liver fatty acid-binding protein. Cancer Metastasis Rev. 13, 317-336.

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

Storch, J., Thumser, A.E.A, 2000. The fatty acid transport function of fatty acid-binding proteins. Biochim. Biophys. Acta 1486, 28-44.

Strimmer, K. and von Haeseler A., 1996. Quartet Puzzling: A Quartet Maximum-Likelihood Method for Reconstructing Tree Topologies. Mol. Biol. Evol. 13, 964-969.

Sweetser, D.A., Lowe, J.B., Gordon, J.I., 1986. The nucleotide sequence of the rat liver fatty acid-binding protein gene. Evidence that exon 1 encodes an oligopeptide domain shared by a family of proteins which bind hydrophobic ligands. J. Biol. Chem. 261, 5553-5561.

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.

Swofford D.L., 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tanaka, T., Yoshida, N., Kishimoto, T., Akira, S., 1997. Defective adipocyte differentiation in mice lacking the C/EBPβ and/or C/EBPδ gene. EMBO J. 16, 7432-7443.

Tataranni, P.A., Baier, L.J., Paolisso, G., Howard, B.V., Ravussin, E., 1996. Role of lipids in development of noninsulin-dependent diabetes mellitus: lessons learned from Pima Indians. Lipids 31 (Suppl.), S267-S270.

Tavakkol, A., Griffiths, C.E., Keane, K.M., Palmer, R.D., Voorhees, J.J., 1992. Cellular localization of mRNA for cellular retinoic acid-binding protein II and nuclear retinoic acid receptor-gamma 1 in retinoic acid-treated human skin. J. Invest. Dermatol. 99, 146-150.

Taylor, J.S., Van de Peer, Y. and Meyer, A., 2001a. Genome duplication, divergent resolution and speciation. Trends Genet. 17:299-301.

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

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-4680.

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

Thompson, J., Winter, N., Terwey, D., Bratt, J., Banaszak, L., 1997b. The crystal structure of the liver fatty acid-binding protein. A complex with two bound oleates. J. Biol. Chem. 272, 7140-7150.

Van de Peer, Y., Taylor, J.S., Meyer, A., 2003. Are all fishes ancient polyploids? J. Struct. Funct. Genomics. 3:65-73.

Vassileva, G., Huwyler, L., Poirier, K., Agellon, L.B., Toth, M.J., 2000. The intestinal fatty acid binding protein is not essential for dietary fat absorption in mice. FASEB J. 14, 2040-2046.

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

Wei, L-N., Tsao, J.L., Chu, Y.S., Jeannotte, L., Nguyen-Huu, M.C., 1990. Molecular cloning and transcriptional mapping of the mouse cellular retinoic acid-binding protein gene. DNA Cell Biol. 9, 471-478.

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

Wittbrodt, J., Meyer, A., Schartl, M., 1998. More genes in fish? BioEssays 20: 511-515.

Wolfrum, C., Ellinghaus, P., Fobker, M., Seedorf, U., Assmann, G., Borchers, T., Spener, F., 1999. Phytanic acid is ligand and transcriptional activator of murine liver fatty acid binding protein. J Lipid Res. 40, 708-714.

Wolf, G., 1991. The intracellular vitamin A-binding proteins: an overview of their function. Nutrition reviews 49, 1-12.

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 Research 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.

Yamada, K., Yuan, X., Ishiyama, S., Koyama, K., Ichikawa, F., Koyanagi, A., Koyama, W., Nonaka, K., 1997. Association between Ala54Thr substitution of the fatty acid-binding protein 2 gene with insulin resistance and intra-abdominal fat thickness in Japanese men. Diabetologia 40, 706-710.

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

Zimmerman, A.W., Veerkamp, J.H., 2002. New insights into the structure and function of fatty acid-binding proteins. Cell Mol. Life Sci. 59, 1096-1116.