EVOLUTIONARY HISTORY AND DIVERSIFICATION OF DUPLICATED FATTY-ACYL ELONGASE GENES OF ATLANTIC SALMON (*Salmo salar*)

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

by

Greta Elena Carmona-Antoñanzas (MSc)

March 2014

School of Natural Sciences
Institute of Aquaculture

Declaration

I hereby declare that this thesis has been composed entirely by myself and has not been submitted for any other higher degree or qualification. All sources of information have been suitably acknowledged in the text.

Name: Greta Carmona-Antoñanzas

Sign: …….

Date: …….

Primary Supervisor: Dr Michael J. Leaver

Sign: …….

Date: …….

Acknowledgements

It is my pleasure to thank all those people who made this project possible.

I would like to express my gratitude towards my supervisors, Dr Michael J. Leaver and Douglas R. Tocher, whose academic experience, kind support and trust in me has made an independent and mature researcher out of me.

I am deeply indebted to Dr Óscar Monroig, who introduced me to the fascinating world of molecular biology and shared with me his knowledge which would enable me to successfully pursue a research career in the field of molecular biology and evolution. My sincere thanks to Dr John B. Taggart and Dr Michaël Bekaert for their motivational and enlightening conversations about the most diverse topics, and technicians at the Institute of Aquaculture for the help provided in molecular biology and cell culture techniques.

My thanks also to the previous and current members of the L5 office, my flatmates and everyone who has crossed my life making this journey the most wonderful experience. I would also like to thank my friends Nacho, Xristoforos, Beatrix, Munevver, Joanna, Laura, Christos, Sean, Eleanna, Željka, Marie, Carina, Sara and Eric for their support and making my PhD a fun and enriching experience.

Finalmente, me gustaría especialmente dedicar mi tesis a mis padres quienes me han apoyado continuamente desde el día en el que comenzó ésta aventura.

Abstract

Background: The Atlantic salmon, *Salmo salar* L., is a prominent member of the Salmonidae family, and has been the focus of intense research because of its environmental and economic significance as an iconic sporting species and its global importance as an aquaculture species. Furthermore, salmonids constitute ideal organisms for the study of evolution by gene duplication as they are pseudotetraploid descendants of a common ancestor whose genome was duplicated some 25 to 100 million years ago. Whole-genome duplication is considered a major evolutionary force capable of creating vast amounts of new genetic material for evolution to act upon, promoting speciation by acquisition of new traits.

Recently, large-scale comparison of paralogous genes in Atlantic salmon suggested that asymmetrical selection was acting on a significant proportion of them. However, to elucidate the physiological consequences of gene and genome duplications, studies integrating molecular evolution and functional biology are crucial. To this end, sequence and molecular analyses were performed on duplicated Elovl5 fatty-acyl elongases of Atlantic salmon, as they are responsible for a rate-limiting reaction in the elongation process of long-chain polyunsaturated fatty acids (LC-PUFA), critical components of all vertebrates.

The aim of the research presented here was to investigate the role of gene duplication as an evolutionary process capable of creating genetic novelty, and to identify the potential ecological and physiological implications.

Results: Linkage analyses indicated that both fatty-acyl elongases segregated independently and located *elovl5* duplicates on different linkage groups. Genetic mapping using microsatellites identified in each *elovl5* locus assigned *elovl5a* and *elovl5b* to chromosomes ssa28 and ssa13, respectively. *In silico* sequence analysis and selection tests indicated that both salmon Elovl5 proteins were subject to purifying selection, in agreement with previous results showing indistinguishable substrate specificities.

Gene expression and promoter analysis indicated that Elovl5 duplicates differed in response to dietary lipids and tissue expression profile. Lipid biosynthesis and metabolic gene expression profiling performed in Atlantic salmon SHK-1 cells, suggested that the control of lipid homeostasis in fish is similar to that described in higher vertebrates, and revealed the particular importance of Lxr and Srebp transcription factors (TFs) in the regulation of LC-PUFA biosynthetic enzymes. Sequence comparison of upstream promoter regions of *elovl5* genes showed intense differences between duplicates. Promoter functional analysis by co-transfection and transcription factor transactivation showed that both *elovl5* duplicates were upregulated by Srebp overexpression. However, *elovl5b* exhibited a higher response and its promoter contained a duplication of a region containing response elements for Srebp and NF-Y cofactors. Furthermore, these studies indicated an Lxr/Rxr dependant response of *elovl5a*, which was not observed in *elovl5b*.

Analysis of the genomic sequences of *elovl5* duplicates by comparison to various sequence databases showed an asymmetrical distribution of transposable elements (TEs) in both introns and promoter regions. Further comparison to introns of the

single *elovl5* gene in pike indicated much higher TE distribution in salmon genes compared to the pike.

Conclusions: Although not conclusive, the most parsimonious origin for the salmon *elovl5* duplicates is that they are derived from a WGD event. This conclusion is also supported by the close similarity of two *elovl5* paralogs in the recently available rainbow trout genome. Regardless of their origin, Atlantic salmon *elovl5* genes have been efficiently retained in the genome under strong functional constraints indicating a physiological requirement for both enzymes to be functionally active. In contrast, upstream promoter regions have strongly diverged from one another, indicating a relaxation of purifying selection following the duplication event. This divergence of *cis*-regulatory regions has resulted in regulatory diversification of the *elovl5* duplicates and regulatory neofunctionalisation of *elovl5a*, which displayed a novel Lxr/Rxr-dependant response not described in sister or other vertebrate lineages. Promoter analysis indicated that the observed *elovl5* differential response to dietary variation could be partly attributed to varying transcriptional regulation driven by lipid-modulated TFs. The distribution of TEs in *elvol5* genes of Atlantic salmon shows a clear increase in TE mobilisation after the divergence of esocids and salmonids. This must have occurred after the elongase duplication and thus the salmonid WGD event and contributes to the observed regulatory divergence of *elovl5* paralogs.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Tables

Abbreviations and Acronyms

Below is a list of the most commonly used abbreviations in the text. Other abbreviated terms are explained in the text.

aa	Amino acid(s)
AA	Adrenic acid (22:4n-6)
ABCA1	ATP-binding cassette transporter 1
ACOX	Acyl-CoA oxidase
ALA	α-linolenic acid (18:3n-3)
ARA	Arachidonic acid (20:4n-6)
BHT	Butilhidroxitolueno
BLAST	Basic local alignment search tool
bp	Base pair(s)
BSA	Bovine serum albumin
cDNA	Complementary DNA
CDS	Coding DNA sequence
CYP7α1	Cholesterol 7alpha-hydroxylase
DHA	Docosahexaenoic acid (22:6n-3)
DMEM	Dulbecco's modified eagle medium
dN	Observed non-synonymous substitutions
DPA	Docosapentaenoic acid (22:5n-3)
dS	Observed synonymous substitutions
EDTA	Ethylenediaminetetraacetic acid
ELF-1α	Elongation factor 1 alpha
EPA	Eicosapentaenoic acid (20:5n-3)
ELOVL	Fatty-acyl elongase
ER	Endoplasmic reticulum

EtOH	Ethanol
FA	Fatty acid
FADS	Fatty-acyl desaturase
FAME	Fatty acyl methyl ester
FAS	Fatty acid synthase
FAF	Fatty acid free
FBS	Foetal bovine serum
FHM	Fathead minnow epithelial cells
FO	Fish oil
GC-MS	Gas chromatography-mass spectrometry
HBSS	Hank's balanced salt solution
HMG-CoAR	3-hydroxy-3-methyl-glutaryl-CoA reductase
LA	Linoleic acid (18:2n-6)
LBD	Ligand binding domain
LC-PUFA	Long-chain polyunsaturated fatty acids
LO	Linseed oil
LXR	Liver X receptor
LXRE	LXR response element
MCL	Maximum composite likelihood
mRNA	Messenger RNA
MYR	Million years
NJ	Neighbour Joining method
nt	Nucleotide(s)
ORF	Open reading frame
OTU	Operational taxonomic unit
PA	Palmitic acid
PBS	Phosphate buffer saline
PHD	Profile fed neural network systems from HeiDelberg
PPAR	Peroxisome proliferator-activated receptors

PPRE	PPAR response element
PUFA	Polyunsaturated fatty acids
qPCR	Quantitative RT-PCR
RE	Response element
RO	Rapeseed oil
RXR	Retinoid X receptor
SCD	Stearoyl CoA desaturase
SHK-1	Salmon head kidney cells
SNAP	Synonymous non-synonymous analysis program
SO	Soybean oil
SRE	Serum response element
SREBP	Sterol regulatory element binding proteins
TE	Transposable element
TF	Transcription factor
TSS	Transcription start site
UAS	Upstream activation sequence
UTR	Unstranslated region
VO	Vegetable oil

Chapter 1

General Introduction

The Atlantic salmon, *Salmo salar* L., is a prominent member of the Salmonidae, and has been the focus of intense research because of its environmental and economic importance as an iconic sporting species and its global importance as an aquaculture species. The understanding of Atlantic salmon biology has advanced considerably over recent years, with a great deal of effort directed to the study of genetics, nutrition and welfare. Much of the research on salmonid biology is focused toward the sustainable production of farmed fish, and the genetics of Atlantic salmon is now widely recognised as fundamental to the understanding of the internal biological processes responsible for development, reproduction and inheritance necessary for aquaculture (Webb *et al*., 2007). In addition to genetics, an understanding of the ecology and evolution of Atlantic salmon and other salmonids in the context of habitat preferences, feeding and reproductive strategies can also enable the inference of aspects of the species biology useful in aquaculture.

1.1 LIPIDS IN FISH NUTRITION

In fish, lipids and their constituent fatty acids (FAs), along with proteins are the major organic source of metabolic energy for growth, migration and reproduction, whereas carbohydrates play a minor functional role (Tocher, 2003). In

vertebrates, most fatty acid chains generally have an even number of carbons predominantly in the range of C14-C24, but there is also high variation in the structure of minor lipids, with odd-numbered FA, saturated or unsaturated, cyclic, straight or branched, and hydroxylated or non-hydroxylated chains (Gurr *et al*., 2002). This variability in FA structure explains their ability to perform diverse roles in biological processes. In fish, the prominence of lipids and fatty acids, particularly in the context of aquaculture, relates to the unique abundance of long-chain, highly unsaturated fatty acids of the omega-3 (n-3) series, for which there is overwhelming evidence of beneficial impact on many aspects of human health (Tocher, 2003; Turchini *et al*., 2011).

1.1.1 Biosynthesis of fatty acids

Teleosts, like all vertebrates, have the enzymatic capacity to biosynthesise saturated and monounsaturated fatty acids (single ethylenic bond) *de novo* using a complex enzymatic system consisting of enzymes to chain-elongate and desaturate fatty acid substrates that include the fatty acid synthase (FAS) and stearoyl Δ9 fatty acid desaturase (SCD1), respectively (Tocher, 2003). The final products of these reactions are 16:0, 16:1, 18:0 and 18:1. However, algae (phytoplankton) and plants with chlorophyll are the only organisms that possess Δ12 and Δ15 desaturases capable of converting monounsaturated fatty acids, such as 18:1, into the polyunsaturated fatty acids (PUFAs), linoleic acid (LA, 18:2n-6) and α-linolenic acid (ALA, 18:3n-3). PUFA such as ALA and LA are considered essential fatty acids (EFAs) for vertebrates as they cannot be biosynthesised endogenously and have to be to supplied in the diet (Agaba *et al*., 2005; Brett and Müller-Navarra, 1997; Tocher *et al*., 2001). In vertebrates, these essential dietary FAs can undergo metabolic

transformations to long-chain polyunsaturated FAs (LC-PUFAs) catalysed by multi-enzyme complexes located in the endoplasmic reticulum, which contain fatty-acyl desaturases (FADS) to form ethylenic or "double" bonds, and fatty-acyl elongases responsible for chain-elongation (Nakamura and Nara, 2004; Tocher *et al*., 2003). The elongation process consists of four successive reactions (condensation, β-reduction, dehydration and enoyl-reductions) performed by individual associated proteins whose activity results in the elongation of FA by the addition of two carbons from malonyl-CoA to an existing fatty acyl-CoA (Agbaga *et al*., 2010; Jakobsson *et al*., 2006). The first step in the elongation process consists of a condensation reaction, which is the rate-limiting reaction catalysed by proteins termed elongases of very long-chain fatty acids (ELOVL) (Jakobsson *et al*., 2006). Seven members of the ELOVL family, termed ELOVL1-7, that differ from each other in their substrate specificity and tissue distribution have been described in mammals (Jump, 2009; Tamura *et al*., 2009). Generally, ELOVL1, ELOVL3, ELOVL6, ELOVL7 elongate saturated and monounsaturated fatty acids, ELOVL4 efficiently elongates both saturated and unsaturated fatty acids, and ELOVL2 and ELOVL5 are preferentially involved in the elongation of PUFAs and LC-PUFAs.

1.1.2 Importance of long-chain polyunsaturated fatty acids (LC-PUFAs)

Naturally, the most important and abundant LC-PUFA in marine fish are those of the n-3 series, also known as omega-3, including eicosapentaenoic acid (EPA, 20:5n-3) and docosahexaenoic acid (DHA, 22:6n-3), although in freshwater fish, n-6 LC-PUFA, such as arachidonic acid (ARA, 20:4n-6), can be found at similar concentrations (Karapanagiotidis *et al*., 2006, 2007). Here, the term "LC-PUFA" is used for PUFA that have a carbon chain length comprised between C20 to

C22 that contain three or more double bonds between the individual carbon atoms of the fatty acid chain (Tocher, 2003; Turchini *et al*., 2011). Where double bonds are formed, hydrogen atoms are eliminated; hence, unsaturated chains have a lower melting point conferring increased fluidity and functionality to cell membranes, when integrated in phospholipid molecules (Kidd, 2007). For instance, membrane LC-PUFA composition is especially important in cold environments as it helps maintain membrane functionality at low temperatures (Hulbert, 2003). Abundant in the retina and cerebral grey matter, DHA (Breckenridge *et al*., 1972) has been proven essential to pre- and postnatal neural development, as well as for visual functioning and reproduction (Crawford, 1993; Kidd, 2007; Neuringer *et al*., 1986; Sargent *et al*., 1999; Tocher, 2003). Besides their generalised role in maintaining the functional integrity of cell membranes, LC-PUFAs play a more specific role as modulators of the immune response and the inflammatory process. In fish, as in mammals, ARA is the major precursor of eicosanoids, biologically active paracrine hormones majorly involved in the development of the inflammatory process (Calder *et al*., 2009). EPA, in contrast, competitively inhibits the formation of proinflammatory eicosanoids from ARA, and decreases the production of cytokines, reactive oxygen species and the expression of adhesion molecules characteristic of inflammatory processes. Long-chain n-3 PUFAs also give rise to a family of potent antiinflammatory mediators termed resolvins (Calder, 2006, 2007; Calder *et al*., 2009). High tissue ratios of ARA:EPA enhance eicosanoid actions believed to result in excessive or inappropriate inflammation contributing to a range of acute or chronic diseases such as cardiovascular diseases, blood clotting disorders, type II diabetes and others (Calder, 2006; Montero and Izquierdo, 2011; Sargent *et al*., 1999). This indicates

that the optimal requirement for any individual long-chain PUFA cannot be considered in isolation, and neither, therefore, can the optimal dietary requirements. Rather, the ratio of n-6 / n-3 PUFA should be considered and defined in a specific tissue or whole body level. Overall, n-3 LC-PUFAs are recognised as physiologically important nutrients established as beneficial to health, and factors lowering the risk of premature death (Castell *et al*., 1994; Kidd, 2007; Wang *et al*., 2006).

In addition to the critical role of LC-PUFA in human health, these compounds are also important for the normal physiology in other vertebrates including fish. The growth in fin-fish aquaculture has been made possible by the development of artificial diets or feeds formulated to satisfy essential requirements, including requirements for LC-PUFA. Traditionally, the key biological resources needed for aquaculture development, fish meal (FM) and fish oil (FO), have been extracted to the limits of sustainability from world capture fisheries. Fish oil, the key supplier of health-beneficial n-3 LC-PUFA, is a finite and limiting resource that fluctuates drastically in supply and cost, and thus constitutes a major impediment to the growth and sustainability of aquaculture. Therefore, to maintain the steady growth of the aquaculture industry under an ecologically sustainable strategy, an alternative source of protein and, especially, lipid is required, among which terrestrial plant-derived products constitute the more readily available and economic alternative (Turchini *et al*., 2011). In comparison with oils from terrestrial origin, fish oils, typically extracted from “fatty fish” species, are characterised by their high proportions of EPA (20:5n-3) and DHA (22:6n-3). Almost all vegetable oils (VO), on the other hand, predominantly extracted from rapeseed, soybean, palm, corn and sunflower are characterised by having high proportion of n-6 C18 PUFAs, generally

low n-3 PUFA and absolutely no n-3 LC-PUFA, which results in a disproportionally higher n-6 / n-3 ratios (Gunstone, 2011; Leaver *et al*., 2008b). In addition, plant-derived oils are devoid of cholesterol, vital for the normal fluidity of biological membranes. As a result of the higher inclusion of VO in fish feeds, the lipid content and composition of today's fish diets has been changing.

The replacement of FO with VO in the diets of freshwater herbivorous fish species, which have been shown to have lower requirements of n-3 LC-PUFA, has great potential (Turchini *et al*., 2009, 2011). Research on marine fish species mostly of piscivorous habit, such as turbot (*Psetta maxima*), European sea bass (*Dicentrarchus labrax*), sharpsnout sea bream (*Diplodus puntazzo*), red sea bream (Pagrus auratus), gilthead sea bream (*Sparus aurata*), humpback grouper (*Cromileptes altivelis*) and many others has shown similar encouraging results, although with more restrictive inclusions of VO by comparison (Montero and Izquierdo, 2011; Shapawi *et al*., 2008). Inclusion of high levels of plant-products in Atlantic salmon diets resulted in no major detrimental effects on fish health and welfare. However, the fatty acid composition of the fish flesh mirrored the fatty acid composition of the diets characterised by high levels of C18 and n-6 PUFA and decreased proportions of saturated fatty acids, EPA, DHA and sterols (Bell *et al*., 2003; Leaver *et al*., 2008b; Rosenlund *et al*., 2001; Torstensen *et al*., 2000), despite attempts to control the extent of fatty acid modification towards the maintenance of a general fish fatty acid composition (Montero *et al*., 2003; Turchini *et al*., 2009). In a number of cases, the inclusion of VO resulted in neutral lipid and triglyceride (TAG) deposition, which may impair intestine, liver or heart functionality, affecting not only dietary nutrient utilisation, but also immune defence and smoltification (Leaver *et*

al., 2008b; Montero and Izquierdo, 2011). The accumulation of individual fatty acids in fish tissues is dependent on various metabolic processes, such as incorporation, β-oxidation, lipogenesis and fatty acid elongation and desaturation processes (Robin *et al*., 2003; Sargent *et al*., 2002). Hence, it is not surprising that the inclusion of plant-derived products alters the activity and expression of genes involved in the uptake, transport and biosynthesis of lipids and LC-PUFA (Leaver *et al*., 2008b; Tocher *et al*., 2001, 2002). Fish, especially oily fish, are the main source of n-3 LC-PUFA in the human food basket. Thus, there has been intense interest in understanding the molecular functioning of LC-PUFA biosynthetic pathways in an attempt to produce healthy fish without compromising the health-promoting characteristics associated with their consumption.

1.1.3 Synthesis of LC-PUFA in fish

Among organisms in the biosphere, algae, and, in particular marine diatoms, cryptophytes, and dinoflagellates, rather than vascular plants, can synthesise high amounts of EPA, DHA and ARA *de novo* (Alonso and Maroto, 2000; Cohen *et al*., 1995; Heinz, 1993; Tocher *et al*., 1998). Therefore, aquatic ecosystems play a unique role as the principal source of the LC-PUFA for all animals, including inhabitants of terrestrial ecosystems as they are accumulated and transferred throughout the food chain (Gladyshev *et al*., 2009). The degree to which an animal can produce LC-PUFA from dietary C18 PUFA substrates depends on possession of the biosynthetic enzyme machinery, as well as the availability of EPA and DHA in their natural diets (Monroig *et al*., 2010b; Morais *et al*., 2009; Tocher, 2003). Thus, the ability to accumulate sufficient n-3 LC-PUFA in fish might be determined by their feeding behaviour and the genetics of the species (FIGURE 1.1).

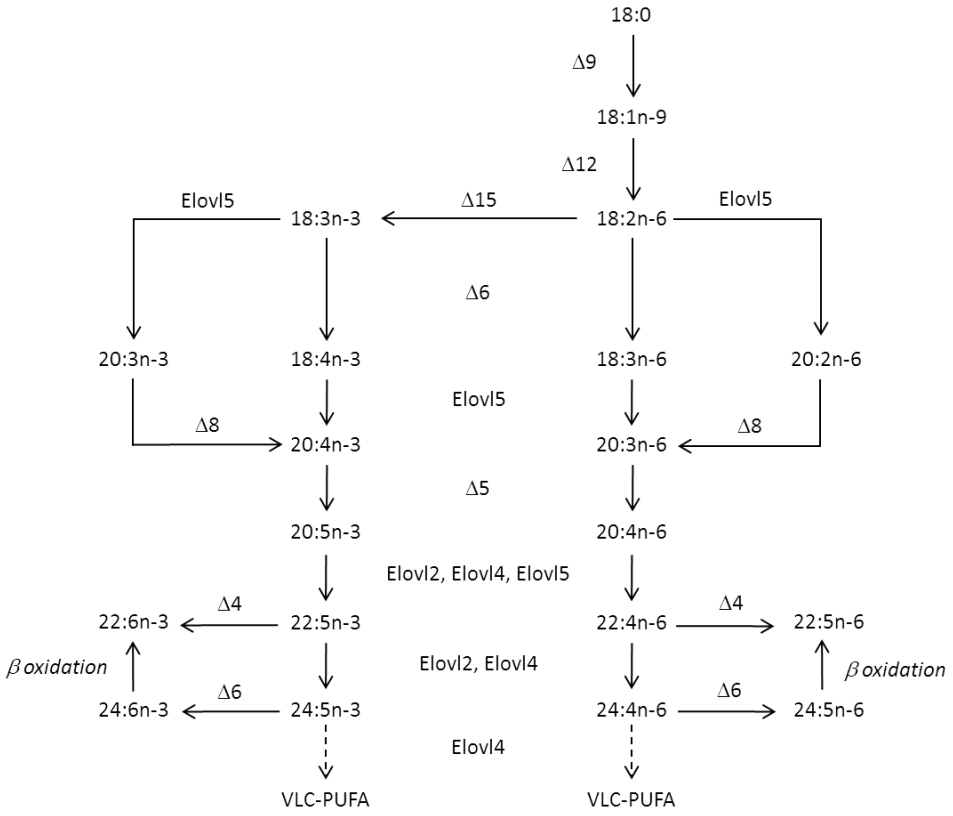

FIGURE 1.1 Pathways of biosynthesis of long-chain and very long-chain (VLC polyunsaturated fatty acids (PUFA) from n-3, n-6 and n-9 C18 precursors α-linolenic (18:3n-3), linoleic (18:2n-6) and stearic (18:0) acids. Δ5, Δ6, Δ8, Δ9, Δ12, Δ15, Fatty-acyl desaturases; Elovl, Fatty-acyl elongases. Δ9 desaturase is found in all animals and plants whereas Δ12 and Δ15 desaturases are generally only found in plants and so 18:2n-6 and 18:3n-3 are "essential" fatty acids (EFA) for many animals. Carnivores and many marine fish generally have only limited ability to carry out the conversions above due to specific deficiencies in desaturases and/or elongases and so 20:5n-3, 22:6n-3 and 20:4n-6 are their EFA.

There is a large body of evidence suggesting that the conversion of 18:3n-3 to EPA and DHA occurs poorly in most species of marine fish studied so far as they lack some essential enzymes involved in the biosynthesis of LC-PUFA, possibly lost during evolution (Castro *et al*., 2011; Tocher *et al*., 2006). Biosynthesis of LC-PUFA is a multi-stage process that requires successive actions of elongase and desaturase enzymes. For instance, synthesis of EPA is achieved by Δ6 desaturation of 18:3n-3 to produce 18:4n-3 that is elongated to 20:4n-3 followed by Δ5 desaturation (Cook and McMaster, 2004), whereas DHA synthesis requires two further elongation steps,

a second Δ6 desaturation and a peroxisomal chain shortening step (Sprecher, 2000). Genomic screenings and cellular assays have suggested that marine Acanthopterygians lack *elovl2* and there appears to be a block at the Δ5-desaturase and/or the C18-C20 elongase step, as has been observed in the European sea bass (*Dicentrarchus labrax*) and Atlantic cod, (*Gadus morhua*) (Bell *et al*., 2006; Mourente *et al*., 2005; Tocher, 2003; Tocher *et al*., 2006), consistent with their negligible ability to synthesise EPA and DHA (Mourente and Tocher, 1994). Evidence suggests that even with those enzymes that are present, such as Δ6 desaturase, gene expression may be very low and thus flux through the pathway is negligible (Tocher *et al*., 2006; Zheng *et al*., 2004).

Recent phylogenetic studies have concluded that two desaturases, FADS1 and FADS2, evolved prior the origin of jawed vertebrates *ca* 400 million years (myr) ago and prior to the evolutionary emergence of the dominant LC-PUFA producing phytoplankton (i.e. 250 myr ago) in marine environments (Castro *et al*., 2012; Knoll *et al*., 2007). Early oceans might have been relatively LC-PUFA-poor ecosystems, primarily dominated by green algae and cyanobacteria, where the acquisition of FADS desaturases might have been necessary to enable survival (Castro *et al*., 2011). Later, neopterygians ("new fins"), which include the vast majority of the modern fish (teleosts), evolved skull modifications that enabled different feeding mechanisms and consequently allowed colonisation of new ecological niches (López-Arbarello, 2012). Previously, the ability to synthesise EPA, DHA and ARA had been attributed to the marine or freshwater nature of the species (Sargent *et al*., 1999). However, further research on numerous freshwater and marine species has suggested that perhaps, rather than ecological niche, it is the feeding strategy and

dietary LC-PUFA that have shaped the genome and ability of certain teleosts to endogenously synthesise LC-PUFA. Supporting this view, the highly piscivorous freshwater esocid Northern pike (*Esox lucius* L.) has displayed only limited activity to convert C18 PUFA to EPA and ARA (Buzzi *et al*., 1997; Henderson *et al*., 1995) whereas, freshwater salmonids with a low-LC-PUFA diet display a relatively efficient ability to synthesise LC-PUFA (Schwalme, 1992; Sushchik *et al*., 2006; Tocher *et al*., 2001). Despite their close phylogenetic relationship, having diverged some 95 myr ago, esocids and salmonids including salmon, trout, charrs, whitefish and graylings exhibit different feeding strategies (Alp *et al*., 2008; Bell *et al*., 1994; Heissenberger *et al*., 2010; Winfield *et al*., 2012). In freshwater, primary production and efficiency of energy transfer in riverine aquatic food webs, where the early life stages of Atlantic salmon and other salmonid take place, is poorer and lower than in mesotrophic or eutrophic lakes which are more suitable for phytoplankton and zooplankton (i.e. copepods and cladocerans) production (Gladyshev *et al*., 2011; Heissenberger *et al*., 2010; Jonsson and Jonsson, 2011).

Generally, lipids and their constituent PUFA originate in phytoplankton at the level of primary producers and are transferred to higher trophic levels of the food web and can be accumulated in aquatic organisms at higher trophic levels (Sushchik *et al*., 2006). Thus, carnivorous/piscivorous species at the upper trophic level of food webs fuelled by LC-PUFA-rich phytoplankton, like esocids or most marine species can obtain sufficient LC-PUFA from the diet. Thus, they do not require to biosynthesise LC-PUFA *de novo* to any significant extent. Rather, the lipid deposits these fish accumulate will be derived largely if not exclusively from dietary lipid (Schwalme, 1992; Tocher, 2003). On the other hand, freshwater salmonids in

oligotrophic streams and lakes exhibit a diet primarily composed of aquatic zoobenthos and terrestrial invertebrates during the early life stages (Heissenberger *et al*., 2010; Sushchik *et al*., 2006). These ecosystems are fuelled allochthonously by terrestrial plant inputs and are therefore relatively poor in n-3 PUFA (Hanson *et al*., 1985; Makhutova *et al*., 2011; Schwalme, 1992; Sushchik *et al*., 2006), which are critical during larval developmental stages (Benitez-Santana *et al*., 2006; Navarro *et al*., 1995). Despite the dramatic difference in LC-PUFA levels between the principal prey species of salmonids and esocids, the body lipid composition of wild specimens does not vary significantly (Blanchet *et al*., 2005; Jankowska *et al*., 2008; Schwalme, 1992). In addition, the levels of LC-PUFA in salmonids inhabiting lacustrine (i.e. brown trout, *Salmo trutta*, and Arctic char, *Salvelinus alpinus*), or riverine (i.e. rainbow trout, *Oncorhynchus mykiss*, grayling, *Thymallus arcticus*) systems show no variation, in spite of high seasonal n-3 LC-PUFA fluctuations in the invertebrates composing their diets (Heissenberger *et al*., 2010; Sushchik *et al*., 2006). These results suggest that salmonids might efficiently retain DHA and/or possess higher biosynthetic LC-PUFA capability than carnivorous/piscivorous species like esocids. Yet, no clear connection between the feeding behaviour and the LC-PUFA biosynthetic ability of salmonids has been established, although habitat is indeed described as a "template" that influences life histories and physiology of organisms (Makhutova *et al*., 2011; Poff and Ward, 1990). Therefore, the n-3 LC-PUFA level in a fish depends not only on the dietary supply but also on the genetically determined ability to biosynthesise these fatty acids. For this reason, qualitative and quantitative EFA requirements of fish can vary significantly with species, season, temperature and stage of development or reproductive system (Turchini *et al*., 2011).

There is some evidence that salmonids do have higher n-3 LC-PUFA biosynthetic capability. Isolated hepatocytes from several salmonid species, including trout, salmon and charr incubated with [1-^{14}C] 18:3n-3 show efficient conversion to LC-PUFA (Tocher *et al*., 2001). Similar studies performed with freshwater non-piscivorous species Nile tilapia (*Oreochromis niloticus*) and zebrafish (*Danio rerio*) indicated that despite being able to convert labelled 18:3n-3 to LC-PUFA, the absolute enzymatic activity was significantly lower than that registered in salmonids for the given physicochemical conditions (Tocher *et al*., 2002). Atlantic salmon, in contrast with most actinopterygians that lack Δ5 desaturases and *elovl2* elongases, possess a full set of LC-PUFA biosynthetic genes that include duplicated genes. Duplicated *elovl5* and *fasd2* genes in salmon, perhaps responsible for the relatively higher ability towards LC-PUFA biosynthesis, might have arisen as a result of the whole-genome duplication (WGD) that occurred *ca* 25 to 100 myr ago in the common ancestor of all salmonids (Koop and Davidson, 2008).

1.2 GENE AND GENOME DUPLICATION (POLYPLOIDISATION)

Gene, and particularly genome duplication, creates new raw material for species to evolve by generating sufficient material to enable evolution by natural selection (Levasseur and Pontarotti, 2011; Long *et al*., 2003; Zhang, 2003). The term "gen(om)e" will be used from now on to refer to both gene and genome duplications, as genome duplication could be considered as a large-scale gene duplication event. Advanced genome sequencing and map information have confirmed that genome duplication has shaped the genome of eukaryotic organisms throughout evolution, and offered evidence to hypothesise its potential significance as a major evolutionary

mechanism for speciation and diversification (Aury *et al*., 2006; Crow and Wagner, 2006; David *et al*., 2003; Gaut and Doebley, 1997; Hoegg *et al*., 2004; Seoighe and Wolfe, 1999; Soltis *et al*., 2009). WGD, regarded as a specific type of polyploidisation by which the entire gene repertoire of an organism is doubled, can occur spontaneously in nature at low frequency by hybridisation between closely-related species (allopolyploidisation), or within a species (autopolyploidisation) *via* failure of cell division by lack of disjunction among daughter chromosomes after DNA replication during meiosis (gametic non-reduction) and mitosis (genome doubling) (Levasseur and Pontarotti, 2011; Li, 1997). Allopolyploids present enough structural dissimilarity between their basic genomes to generate bivalent pairing during meiosis, and thus disomic inheritance (David *et al*., 2003). However, in autopolyploids, originally identical chromosome sets pair up as multivalents at meiosis, which frequently result in unviable aneuploid gametes (Li, 1997). Thus, such a change in ploidy is typically expected to be deleterious and an evolutionary dead-end, which could exert physiological or developmental constraints (Carothers and Beatty, 1975; Mable, 2004).

1.2.1 Polyploidisation events in vertebrates

Polyploidy has been estimated to occur in about half the natural species of flowering plants (angiosperms) as a result of ancestral polyploidisation events, and, although wholesale duplication is much rarer in animals, several rounds of genome duplication have shaped the invertebrate and vertebrate genome, including mammals (Gallardo *et al*., 2004; Soltis *et al*., 2009). It is widely accepted that two rounds of WGD (2R hypothesis) occurred in early vertebrate evolution (Kuraku *et al*., 2009; Soltis *et al*., 2009; Wolfe, 2001) followed by one ancient genome duplication (3R)

specific to ray-finned fish (Actinopterygii) after they diverged from lobe-finned fish (Sarcopterygii) in the early Devonian (FIGURE 1.2) (Jaillon *et al*., 2004; Naruse *et al*., 2004). In terms of numbers, actinopterygians constitute the dominant radiation of vertebrates in terms of species richness with more than 30,000 species, over half the total number of living vertebrate species on the planet (Alfaro *et al*., 2013; Hoegg *et al*., 2004). Among them, most lineages are believed to have arisen during a relatively short period of time *ca*. 100 and 200 myr ago promoted by the genome duplication that preceded the teleost radiation (Comber and Smith, 2004; Crow *et al*., 2006). Additional rounds of genome duplication (4R) have occurred in several orders of teleosts (Comber and Smith, 2004; Ludwig *et al*., 2001; Taylor *et al*., 2001), and genome duplication has given rise to the complete Salmonidae family (Allendorf and Thorgaard, 1984; Koop and Davidson, 2008; Leong *et al*., 2010; Nelson, 2006).

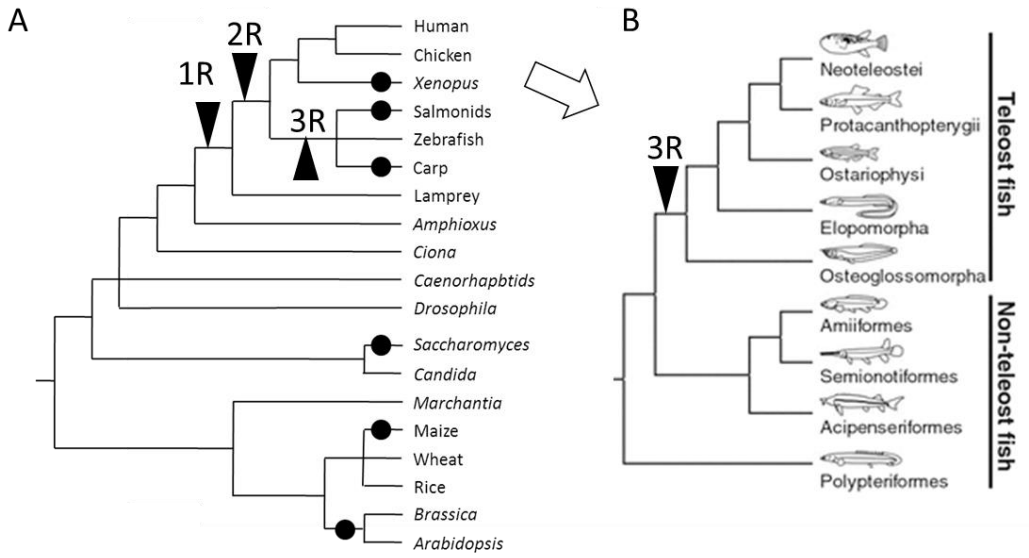

FIGURE 1.2 Whole-genome duplication (WGD) events during eukaryotic evolution. 1R, 2R, and 3R indicate first, second, and third-rounds of WGD in vertebrate evolution, respectively. A) Vertebrate phylogeny and proposed timing of the 1R-, 2R- and 3R-WGD and other WGD events occurred in vertebrate, plant and yeast lineages. B) Actinopterygian phylogeny and the estimated timing of the 3R-WGD. Black circles indicate lineage specific WGD events (Adapted from Sato and Nishida, 2010).

1.2.1.1 Whole-genome duplication in Salmonids

Extant salmonids, like maize, viscacha rats (*Tympanoctomys barrerae*) and the African clawed frog (*Xenopus laevis*), are considered pseudopolyploids, degenerate polyploids in a state of reversion towards diploidy (David *et al*., 2003; Gallardo *et al*., 2004; Helentjaris *et al*., 1988; Hughes and Hughes, 1993; Skrabanek and Wolfe, 1998). Diploidisation is the evolutionary process whereby a polyploid species "decays" to become a diploid with up to twice as many distinct chromosomes (Moen *et al*., 2008). The key event is the switch from having multiple chromosomes that form multivalents at meiosis, to having two pairs of chromosomes each of which forms a bivalent, resulting in two alleles at each of two distinct loci (disomic inheritance). Atlantic salmon, with 29 (European) or 27 (North American) pairs of chromosomes (Phillips *et al*., 2009) is considered functionally diploid although hints of multivalent chromosome formation during meiosis and tetrasome segregation can still be observed at some loci in males (Allendorf and Danzamann, 1997), consistent with the evolutionary recent duplication event in salmonids, one of the latest tetraploidisation events in vertebrates. In addition, salmon exhibit one of the largest euteleost genome sizes (3–4.5 pg) with double of that of sister orders Esociformes (0.8–1.8 pg, Northern pike) and Osmeriformes (0.7 pg, smelt) (Gregory, 2013; Koop and Davidson, 2008). Salmonids have clear homeologous chromosomal segments based on genetic maps and comparative studies using microsatellite markers (Leder *et al*., 2004; Lien *et al*., 2011), and more duplicates within gene families (Hoegg and Meyer, 2005; McKay *et al*., 2004; Moghadam *et al*., 2011; Shiina *et al*., 2005; Thorgaard *et al*., 2002). An increase in species diversity has often been observed in clades that include polyploid forms, and whole-genome duplications have been

proposed to be a lineage-splitting force because of the subsequent occurrence of gene losses, and chromosome rearrangements reducing the fecundity of hybrids (Donoghue and Purnell, 2005; Lynch and Force, 2000; Taylor *et al*., 2001).

Fish exhibit a vast range of diversity in their ecology, morphology, physiology and life history, and it has been suggested that their success and diversity is a direct consequence of gene multiplicity (Comber and Smith, 2004). For this reason, teleosts and, in particular, Atlantic salmon are considered particularly valuable for the study of evolutionary changes in vertebrates following a duplication event, with potential to elucidate the evolutionary processes of speciation and adaptation, and the mechanisms by which duplicated genomes are re-diploidised. The mechanism by which this stabilisation occurs is not well understood, but it may be through large-scale movement of repetitive sequences (Ayala and Coluzzi, 2005; Wolfe, 2001). Increased mobilisation of transposable elements (TEs) has been observed under environmental stress (de Boer *et al*., 2007; Krasnov *et al*., 2005) and inferred during speciation in salmonids. The fact that recombination at sites of mobile element insertion results in deletions, translocations and formation of dicentric chromosomes (Gray, 2000) has led to hypotheses that mobilisation of transposons could play a role in genome evolution and also diploidisation by impairing tetravalent formation during meiosis (de Boer *et al*., 2007; Hurst and Werren, 2001).

1.2.2 Fate of duplicated genes

Ohno's classical view of gene and genome duplication considered that, without duplication, the evolution of the genome tends to be conservative according to the null hypothesis of purifying selection (negative selection) (Ohno, 1970). Under

the assumption that biological evolution is a continuous and gradual process that results in optimal features for the survival of a specific population at a certain time and space, we could ideally consider that most species today are adapted to their current environments. Consequently, unless there is change in the environment, the major mechanism of selection is purifying selection (negative selection), which tends to homogenise and maintain the adapted genetic features. After gen(om)e duplication, duplicated genes are theoretically free to diverge at an accelerated rate due to relaxed purifying selection (Castillo-Davis *et al*., 2004). This, combined with the functional redundancy arising from the presence of two originally identical copies, means that the most common fate of duplicated genes is to become inactive (nonfunctionalisation) through the accumulation of deleterious mutations, or disappear as a result of chromosomal rearrangements. Nevertheless, duplicated genes within a species, also termed paralogous genes or paralogs, can be retained under strong negative selection if the extra amount of protein or RNA product is beneficial, or if they asymmetrically diverge to acquire a new beneficial function (neofunctionalisation) or if the ancestral function(s) is partitioned or shared by both duplicates (subfunctionalisation) (Otto and Whitton, 2000; Sémon and Wolfe, 2007). These different fates are governed by the nature of functionally significant alterations in the coding region of the protein, or in *cis*-regulatory elements altering expression patterns (Castillo-Davis *et al*., 2004; Force *et al*., 1999; Zhang, 2003). These types of evolutionary changes in polyploid genomes might be detected by comparing a polyploid species to a sister species which has not undergone WGD by characterising differences in substitution rates, gene expression, regulation and protein function. Thus, a recent study utilised the esocid Northern pike as the closest non-duplicated

(Ishiguro *et al*., 2003; Nelson, 2006; Ramsden *et al*., 2003) genome to identify duplicated pairs of genes in Atlantic salmon and determine patterns of change. While the salmon genome may still be in the process of reverting to a stable diploid state, studies show that many duplicated genes have been retained and purifying selection was determined as the predominant force acting on paralogous genes, consistent with retention of functionality in both copies, although some asymmetrical relaxation of purifying selection was observed in some genes, thus allowing one of the duplicates to diverge at a faster rate (Lai *et al*., 2009; Leong *et al*., 2010).

The molecular basis and control of LC-PUFA biosynthesis has been highly studied in fish driven by their crucial role as the primary source of n-3 LC-PUFA in the human diet. Hence, genomic screening of DNA libraries elucidated that there were at least four genes encoding functional Fads2 (Fads2Δ6a, Fads2Δ6b, Fads2Δ6c and Fads2Δ5) desaturases and two Elovl5 (Elovl5a and Elovl5b) elongases in Atlantic salmon (Monroig *et al*., 2010b; Morais *et al*., 2009). While no segregation data is available for *elovl5* duplicates, preliminary mapping in informative males suggested that the four *fads2* genes were tightly linked, an indication of segmental duplication origin, rather than a WGD. However, genes in male salmonids exhibit pseudolinkage and further studies on females would have to be performed to establish origin beyond doubt. Genomic and phylogenetic studies have suggested that *Fads1*-Δ5 *and Fads2*-Δ6 appeared before the gnathostome radiation, since the small-spotted catshark (*Scyliorhinus canicula*) has functional orthologues of both gene families. Consequently, the loss of Fads1 in teleosts has been hypothesised as a secondary episode in evolution (Castro *et al*., 2012). These data indicate that the one salmon *fads2* with Δ5 activity, termed *fasd2d5*, might have diverged functionally

from an ancestral Fads2Δ6 enzyme through independent mutations (Monroig *et al*., 2010b). As a result, Atlantic salmon possess the complete biosynthetic pathway capable of desaturating and elongating LC-PUFA from dietary PUFA (Hastings *et al*., 2004; Zheng *et al*., 2009a, 2009b).

Tissue expression studies have indicated that liver, intestine, and brain are the major tissues in which elongases and desaturases of LC-PUFA are expressed, thus suggesting their involvement as major biosynthetic tissues of LC-PUFA (Tocher, 2003; Zheng *et al*., 2009a). Interestingly, the tissue expression patterns displayed by duplicated LC-PUFA genes differed significantly, and both *in vivo* and *in vitro* assays have indicated that desaturase and elongase duplicates respond differently to dietary lipids and fatty acids (Minghetti *et al*., 2011; Monroig *et al*., 2010b; Morais *et al*., 2009). For instance, salmon *elovl5b* expression was downregulated in the liver of fish fed diets rich in n-3 LC-PUFA and cholesterol in comparison with salmon fed low LC-PUFA, whereas *elovl5a* expression did not respond to any of the dietary treatments (Morais *et al*., 2009). Thus, we considered these results a sign of potential divergence in the lipid- and fatty acid-mediated transcriptional regulation of *elovl5* paralogous genes. Although purifying selection was stronger in *cis*-regulatory regions in comparison with coding sequences (Castillo-Davis *et al*., 2004), differential patterns of spatial expression in paralogous genes have been observed in a large range of duplicated genes (Force *et al*., 1999; Monroig *et al*., 2010b; Sha *et al*., 2008). In addition, mechanisms of transcriptional and post-transcriptional regulation have been suggested to play an important role in maintaining the network stoichiometry (gene dosage principle) after duplication events to avoid potential unfavourable and disruptive effects resulting from unbalanced duplicated gene

products (Kondrashov and Kondrashov, 2006; Levasseur and Pontarotti, 2011; Qian and Zhang, 2008).

1.3 TRANSCRIPTION FACTOR-DEPENDENT REGULATION BY DIETARY LIPIDS

Intense research into the consequences of nutritional imbalances in physiological disorders has highlighted the crucial role that LC-PUFA and cholesterol play as signalling molecules at the transcriptional level (Albert *et al*., 2002; Marquart *et al*., 2010; Montero and Izquierdo, 2011). Dietary lipids and LC-PUFA particularly exert major effects in modulating gene regulation through interaction with key transcription factors (TFs) involved in lipid homeostasis, such as liver X receptor (LXR), peroxisome proliferator-activated receptor (PPAR) and sterol regulatory element binding protein (SREBP) (Calder and Burdge, 2004; Colliar *et al*., 2011; Espenshade, 2006; Reschly *et al*., 2008). Transcription factors are proteins capable of regulating gene expression in response to specific stimuli, such as variations in the levels of key nutrients, metabolites, hormones, cytokines etc., through binding to specific DNA sequences in target genes and promoting (as an activator) or blocking (as an inhibitor) the recruitment of RNA polymerase responsible for the transcription of genetic information from DNA to messenger RNA (FIGURE 1.3) (Latchman, 1997). A defining feature of TFs is that they contain DNA-binding domains (DBDs), which recognise specific sequences of DNA, or response elements (RE) adjacent to the genes that they regulate. Indeed, TF are frequently classified on the basis of their DNA binding domain, such as the well

characterised zinc-finger that is found in the nuclear receptor (NR) family (Berg, 1989; Klug and Schwabe, 1995), or the basic helix-loop-helix motif present in SREBPs (Minghetti *et al*., 2011) among others. Interestingly, transcription factors can be activated at two levels, namely the regulation of transcription factor activity and the regulation of transcription factor synthesis. Thus, post-translational activation of transcription factors includes a number of different mechanisms, which can involve proteolytic activation, protein-protein interactions, and/or ligand binding (Bunce and Campbell, 2010; Espenshade, 2006; Lalli and Sassone-Corsi, 1994). Following post-translational activation, transcription factors move to the nucleus where they associate with specific REs located in the promoters of the target genes. DNA dimerisation can also induce conformational changes in the TF interface that lead to different receptor conformations enabling cooperative assembly of protein dimerisation (Holmbeck *et al*., 1998; Zhao *et al*., 2000). Cofactor recruitment leads to activities such as chromatin reorganisation to convert tightly compact chromatin into lightly packed euchromatin, and/or recruitment of the basal transcription complex primarily composed by RNA polymerase II (Amemiya-Kudo *et al*., 2002; Burd and Archer, 2010; Glass *et al*., 1997; Latchman, 1997; Makishima, 2005; Son and Lee, 2005). The activity is an important control point, although it is not the only regulatory mechanisms in all transcription factors. In some cases there is a further regulatory mechanism that allows *de novo* gene transcription by the activation of pre-existing transcription factors, and some transcription factors are capable of self-regulating their own transcription level (Laffitte *et al*., 2001; Shimano, 2001). The transcriptional regulation of lipid homeostasis is a highly intricate and complex mechanism that functions in synchrony with dietary lipids. The major lipid-sensitive

modulators belong to the nuclear receptor group, and the sterol response element-binding protein group. In mammals, upon activation, LXR or PPAR are translocated into the nucleus where they recognise specific LXRE and PPRE binding sites, respectively, in the promoter regions of SREBP-1c (Fernández-Alvarez *et al*., 2011; Ou *et al*., 2001; Yoshikawa *et al*., 2002).

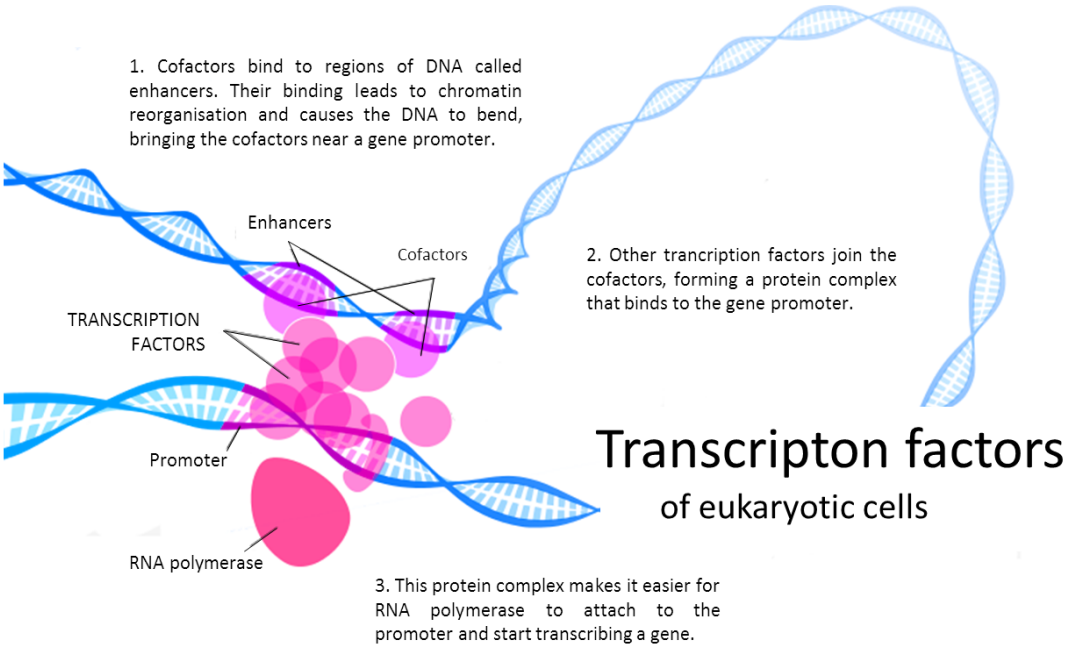

FIGURE 1.3 Transcriptional regulation by transcription factors (Adapted from Creative Commons).

1.3.1 Nuclear receptors (NR)

The nuclear receptor superfamily is a group of originally ligand-dependent transcription factors that regulate diverse physiological functions (Bunce and Campbell, 2010; Evans, 1988). More than 270 nuclear receptors clustered in six subfamilies (NR1 to NR6) have been described in different organisms (Maglich *et al*., 2001), thought to be evolutionary derived from a common ancestor. NR1 family

members are particularly involved in the regulation of lipid metabolic pathways induced in response to exposure to fatty acids, cholesterol and some secondary metabolites (e.g. steroids, corticosteroids, sex hormones, vitamin D and bile acids) (Bunce and Campbell, 2010; Reschly *et al*., 2008). The classic nuclear receptor contains several modulating domains: a conserved DNA binding domain (DBD) that interacts with response element (RE), a conserved ligand-binding domain (LBD) and a transactivation domain or cofactor-binding site, which undergoes conformation changes in response to ligand binding, allowing dimerisation and facilitating recruitment of coactivators (Barrera *et al*., 2008; Glass and Rosenfeld, 2000; Makishima, 2005). NR1 members preferentially form "permissive" heterodimers with retinoid X receptor (RXR) (Färnegårdh *et al*., 2003; Nolte *et al*., 1998; Yue *et al*., 2005; Zelcer and Tontonoz, 2006), which bind to double hexads of the canonical direct repeat (DR) AGGTCA separated by 1 or 4 nucleotides in PPAR (DR1) or LXR (DR4), respectively (Ijpenberg *et al*., 1997; Janowski *et al*., 1996).

1.3.1.1 Liver X receptor (LXR)

LXR (NR1H3) is the major sensor of oxysterols, oxidised cholesterol metabolites, in vertebrates and its activation can regulate the metabolism of cholesterol and bile acids (Reschly *et al*., 2008; Schultz *et al*., 2000; Willy *et al*., 1995). In mammals, two duplicated genes encode isoforms LXRα and LXRβ that differ in their regulation, and expression pattern (Laffitte *et al*., 2001), whereas one single copy in present in the rest of vertebrates. In sterol-replete cells, cellular oxysterols, the natural LXR agonists (Reschly *et al*., 2008), accumulate as a result of increased concentration of cholesterol, and LXR induces the transcription of genes that protect cells from cholesterol overload. Ultimately, LXR activation promotes the

inverse-transport of cholesterol out of the cells by transferring it to apolipoproteins responsible for delivering cholesterol to the liver, where it is transformed into bile acids that will be excreted with the faeces (Zhao and Dahlman-Wright, 2010).

1.3.1.2 Peroxisome proliferator-activated receptor (PPAR)

This ligand-activated family consists of three members named PPARα (NR1C1), PPARβ (NR1C2) and PPARγ (NR1C3). Each PPAR subtype is characterised by a specific tissue expression pattern, partially accounting for distinct biological functions (Sanderson and Kersten, 2010). Activating agonists for PPARs are semi-selective for the subtype, and the selectivity depends on the ligand concentration and cell type. A broad variety of endogenous and exogenous ligands include lipid-like compounds such as fibrates, as well as numerous unsaturated fatty acids and eicosanoids (Forman *et al*., 1997; Willson *et al*., 2000). The most studied PPARγ drives adipocyte differentiation, and has important regulatory roles during fat storage and glucose metabolism. PPARα is known to play an important role in mitochondrial and peroxisomal β-oxidation of fatty acids, whereas PPARβ is ubiquitously involved in basic biological processes including lipid metabolism (Aoyama *et al*., 1998; Barrera *et al*., 2008; Tontonoz *et al*., 1993).

1.3.2 Sterol response element-binding protein (SREBP)

SREBPs belong to the basic helix-loop-helix leucine zipper (bHLH-ZIP) family of transcription factors, but they differ from other bHLH-ZIP proteins in that they are synthesised as inactive TFs sequestered in the membrane of the endoplasmic reticulum (ER) (Brown and Goldstein, 1997; Goldstein *et al*., 2002). SREBP precursors are organised intro three domains: (a) the NH_2-terminal domain that

harbours the bHLH-ZIP, which possesses the ability to recognise sterol response element (SRE) and bind the DNA sequence (Hua *et al*., 1993; Yokoyama *et al*., 1993), (b) two hydrophobic transmembrane-spanning segments, and (c) a COOH-terminus that forms a tight complex with SREBP-cleavage-activating protein (Scap) that functions as the sterol sensor in this system (Horton *et al*., 2002). In sterol-replete cells Scap binds to cholesterol in the ER membranes and assumes a conformation that promotes binding to the ER-resident protein Insig (Peng *et al*., 1997). This retains the SREBP-Scap complex in the ER by preventing interaction of Scap with vesicle-formation proteins (Rawson, 2003). In sterol-depleted cells, binding to Insig is impaired and SREBP-Scap is engulfed by transport vesicles. Scap then escorts SREBP to the Golgi where two sequential proteolytic cleavage events, mediated by proteases S1P and S2P, release the soluble N-terminal domain (Espenshade, 2006; Horton *et al*., 2002). Released SREBP is transported to the nucleus as dimers by importin β through interactions with the helix-loop-helix domain (Lee *et al*., 2003). Vertebrate SREBPs are encoded by the genes *SREBF1* and *SREBF2*, which in mammals encode three isoforms. Mammalian *SREBF1* can be transcribed through the use of alternative transcription start sites that produce alternate forms of exon 1, designated SREBP1c and SREBP1a (Brown and Goldstein, 1997). The predominant forms in the liver are SREBP2 and SREBP1c, which preferentially regulate genes involved in cholesterol biosynthesis and fatty acid synthesis, respectively, whereas SREBP1a is a potent activation of all SREBP-responsive genes. However, when expressed at higher than physiological levels, all SREBP isoforms can activate all target genes (Horton *et al*., 2002). In the nucleus, SREBPs cooperate with other transcription factors and coactivators (Amemiya-Kudo

et al., 2002) to stimulate target gene expression, including that of both SREBP1c and SREBP2 (Horton *et al*., 2002). To date, SREBPs are known to enhance directly transcription of numerous genes required for uptake and synthesis of cholesterol, fatty acids, triacylglycerols, and phospholipids (Horton *et al*., 2003).

1.3.3 Atlantic salmon transcription factors

Homologous genes to the mammalian *LXR*, *SREBPs* and *PPARs* have been identified in Atlantic salmon, including the full-length sequences of *lxr*[1], *srebp1*, *srebp2*, *pparα*, *pparγ* and four isoforms of *pparβ* available in the NCBI database [www.ncbi.nlm.nih.gov/], and several unpublished duplicate sequences (Bente Ruyter, personal communication, 19 June, 2013). Intense research has focused on lipid-sensitive mammalian TFs considered key therapeutic targets for many cardiopathologies and metabolic disorders in humans. However, very little research has been directed to elucidate their role in LC-PUFA biosynthesis in Atlantic salmon, with only few studies presenting gene expression data obtained from *in vivo* nutrition analysis and *in vitro* assays (Cruz-Garcia *et al*., 2009; Kortner *et al*., 2013; Leaver *et al*., 2005, 2007, 2008b; Minghetti *et al*., 2011).

1.4 THESIS AIM AND OBJECTIVES

Recently, extensive genomic resources have facilitated the study of evolution and confirmed that gene and genome duplications have occurred at the base of speciation and divergence of many eukaryote species (Danzmann *et al*., 2005, 2008;

[1] The nomenclature applied to fish genes and proteins throughout the manuscript is in line with the guidelines established by the zebrafish ZFIN Database [http://zfin.org/cgi-bin/webdriver?MIval=aa-ZDB_home.apg]

Lien *et al*., 2011). Large-scale comparison of paralogous genes in Atlantic salmon suggested that asymmetrical selection was acting on a significant proportion of them (Leong *et al*., 2010). To elucidate the physiological consequences of this divergence of duplicated genes, studies integrating molecular evolution and functional biology are crucial; however, they are lacking in Atlantic salmon. Within the context of the above, the aim of the research presented here was to investigate the role of genome duplication as an evolutionary process capable of creating gene novelty. To this end, sequence and molecular analyses were performed on duplicated elongases of LC-PUFA, *elovl5*, of Atlantic salmon to identify the selection pressure acting at the protein and the *cis*-regulatory levels, and the ecological and physiological implications of the duplication event. Firstly, genetic studies and selection tests were applied to identify the origin and evolutionary fate of duplicated LC-PUFA protein-coding genes. Secondly, transactivation assays were developed to assess the evolution of the respective *cis*-regulatory regions. For this, the role of lipid-sensitive transcription factors was analysed both *in vitro* and *in vivo*, to provide further insights into the dietary control of the expression of LC-PUFA genes.

The specific objectives of the current thesis were,

i. To identify the genetic origin of the two *elovl5* loci in Atlantic salmon *(Chapter 2)*
ii. To identify the major selective pressure acting on the salmon Elovl5 proteins (*Chapter 2*)
iii. To characterise a critical gene and enzyme of LC-PUFA biosynthesis, Elovl5, in Northern pike and to compare with previously identified duplicate

elovl5 paralogs in Atlantic salmon *(Chapter 2)*

iv. To identify a series of natural oxysterols and synthetic agonists capable of activating the ligand-binding domain of Atlantic salmon Lxr *(Chapter 3)*

v. To identify Lxr, Ppar and Srebp target genes involved in lipid metabolism and the biosynthetic pathway of LC-PUFA (*Chapter 3*)

vi. To assess *in vivo* dietary regulation of transcription factors involved in the modulation of the expression of duplicated *elovl5* LC-PUFA genes (*Chapter 3*)

vii. To characterise the *cis*-regulatory regions of the duplicated LC-PUFA genes and detect specific Lxr and Srebp response elements responsible for the transcriptional regulation of duplicated *elovl5* LC-PUFA genes (*Chapter 4*)

viii. To assess the potential of PUFA and LC-PUFA as cellular modulators capable of controlling lipid metabolism in an Lxr-dependent manner (*Chapter 4*)

ix. To identify the presence and relative position of transposable elements in Northern pike and Atlantic salmon duplicates to assess their potential role in the divergence of the expression profiles in paralogous genes *(Chapter 4)*

Chapter 2

An evolutionary perspective on Elovl5 fatty acid elongase: comparison of Northern pike and duplicated paralogs from Atlantic salmon (*Salmo salar*)

Greta Carmona-Antoñanzas [§], Douglas R. Tocher, John B. Taggart and Michael J. Leaver

Institute of Aquaculture, School of Natural Sciences, University of Stirling, Stirling, FK9 4LA, Scotland, UK

Authors' contributions

GCA, DRT and MJL planned and coordinated the research; GCA performed laboratory analyses and data analysis; JBT conducted the genetic linkage analysis; GCA wrote the first draft of the manuscript, followed by contributions from remaining authors.

Carmona-Antoñanzas *et al. **BMC Evolutionary Biology*** 2013, **13**:85

Abstract

The ability to produce physiologically critical LC-PUFA from dietary fatty acids differs greatly among teleost species, and is dependent on the possession and expression of fatty acyl desaturase and elongase genes. Atlantic salmon, as a result of a recently duplicated genome, have more of these enzymes than other fish. Recent phylogenetic studies show that Northern pike represents the closest extant relative of the preduplicated ancestral salmonid. Here we characterise a pike fatty acyl elongase, *elovl5,* and compare it to Atlantic salmon *elovl5a* and *elovl5b* duplicates. Genetic linkage analyses showed that the two salmon *elovl5* genes reside on distinct chromosomes and are flanked by homologous genes indicating conserved synteny. Phylogenetic analyses show that the Atlantic salmon Elovl5 paralogs are evolving symmetrically, and they have been retained in the genome by purifying selection. Heterologous expression in yeast showed that Northern pike Elovl5 activity is indistinguishable from that of the salmon paralogs, efficiently elongating C18 and C20 substrates. However, in contrast to salmon, pike *elovl5* was predominantly expressed in brain with negligible expression in liver and intestine. We suggest that the predominant expression of Elovl5b in salmon liver and Elovl5a in salmon intestine is an adaptation, enabled by genome duplication, to a diet rich in terrestrial invertebrates which are relatively poor in LC-PUFA. Pike have retained an ancestral expression profile which supports the maintenance of PUFA in the brain but, due to a highly piscivorous LC-PUFA-rich diet, is not required in liver and intestine. Thus, the characterisation of Elovl5 in Northern pike provides insights into the evolutionary divergence of duplicated genes, and the ecological adaptations of salmonids, which have enabled colonisation of nutrient poor freshwaters.

2.1 INTRODUCTION

Atlantic salmon (*Salmo salar*) have been the focus of considerable research effort as a result of their widespread environmental and economic importance as a sporting and cultured species. In addition, in common with all other salmonids, they possess a comparatively recently duplicated genome, believed to have arisen as a result of a relatively recent autotetraploidisation event between 25 and 100 myr ago (Allendorf and Thorgaard, 1984; Koop and Davidson, 2008). Whole-genome duplication (WGD) has been argued as a powerful evolutionary force creating new raw material for evolution to act upon (Ohno, 1970), thus enabling the divergence and neo- or subfunctionalisation of duplicated loci promoting adaptation and speciation. The imprints of three or four ancient duplications can be detected in vertebrate genomes, including a specific event early in teleost evolution and the recent one in salmonids (Roth *et al*., 2007; Sémon and Wolfe, 2007). Esocids (members of the pike family) are regarded as having the closest extant preduplicated (diploid) genomes to salmonids, based on molecular phylogenetic studies (Li *et al*., 2008; López *et al*., 2004), karyotype data (Gregory, 2013) and comparative analyses of expressed gene sequences (Ishiguro *et al*., 2003; Leong *et al*., 2010). Therefore, Northern pike (*Esox lucius*) is representative of a sister-group to salmonids, and can be viewed as an appropriate species to study the consequences of genome duplication in salmonids. Despite their shared ancestry and overlapping habitats, Atlantic salmon and Northern pike have differing life histories and feeding behaviours where, in freshwaters, pike have a largely piscivorous diet, and salmon a diet rich in terrestrial insects (Tocher *et al*., 2001). These differences may be reflected in differing nutritional physiology and, in particular, lipid biochemistry.

Teleosts, like all vertebrates, are unable to synthesise polyunsaturated fatty acids (PUFA) *de novo*, and so they are essential and required in the diet (Holman, 1986; Sargent *et al*., 2002). However, which PUFA can satisfy the dietary requirement for essential fatty acids (EFA) varies with species. The long-chain PUFA (LC-PUFA), arachidonic acid (20:4n-6, ARA), eicosapentaenoic acid (20:5n-3, EPA) and docosahexaenoic acid (22:6n-3, DHA), which have essential functions in vertebrate immune defense systems and neuronal membranes (Calder *et al*., 2009; Dangour and Uauy, 2008; Wainwright, 2002) can be produced endogenously, in some but not all vertebrates, from the base EFA, α-linolenic acid (18:3n-3, ALA) and linoleic acid (18:2n-6, LA), by desaturation and elongation (Tocher, 2003). The capability to produce LC-PUFA from EFA varies between fish species (Sargent *et al*., 1995, 2002), and salmonids, including Atlantic salmon, brown trout (*Salmo trutta*), and Arctic charr (*Salvelinus alpinus*) have substantially higher LC-PUFA biosynthetic efficiency in comparison with other freshwater species, including zebrafish (*Danio rerio*), Nile tilapia (*Oreochromis niloticus*) (Bell *et al*., 2001; Tocher *et al*., 2001, 2002), and Northern Pike (Buzzi *et al*., 1997; Henderson *et al*., 1995; Kluytmans and Zandee, 1974; Schwalme, 1992). The ability to produce LC-PUFA is dependent on the possession and expression of fatty acyl desaturase (*fads*) and fatty acid elongase (*elovl*) genes, and salmonids, in contrast to many other fish species examined, have a complete set of genes and expressed enzymes required for the production of ARA from LA, and EPA and DHA from ALA (Monroig *et al*., 2010b; Morais *et al*., 2009). In Atlantic salmon some of these LC-PUFA biosynthetic enzymes appear to have arisen from duplicated genes, and the subsequent neo- or

subfunctionalisation has been hypothesised as an enabling adaptation for salmonids to thrive in relatively nutrient-poor freshwater environments (Leaver *et al*., 2008a).

The aim of the present study was to characterise a critical gene and enzyme of LC-PUFA biosynthesis, Elovl5, in Northern pike and to compare with previously identified, duplicated *elovl5* paralogs in Atlantic salmon. Elovl5 catalyses the first and second elongations of LA and ALA and is therefore essential for the production of LC-PUFA. Thus, functional comparison of the protein-coding sequences, activities and tissue expression of pike and salmon *elovl5s* combined with phylogenetic analyses may provide insights into mechanisms that have driven the evolution and ecological adaptations of salmonids.

2.2 MATERIALS AND METHODS

2.2.1 Biological samples

Wild Northern pike were caught in Airthrey Loch, Stirling (Scotland). The individuals were killed by a blow to the head and rapidly dissected and the tissue samples stored in RNAlater® (Invitrogen, Paisley, UK) at room temperature until tissue homogenisation was performed 24 h later. Fin clips were obtained and stored in 70% ethanol for genomic DNA extraction. Pike were obtained conforming to local ethical regulations regarding field studies on wild animals, and with the permission of the local fisheries managers.

2.2.2 Molecular cloning of pike *elovl5* cDNA

Pike *elovl5* cDNA was cloned via RNA extraction, cDNA synthesis and a variety of PCR techniques using the oligonucleotide primers listed in TABLE 2.1. Pike intestinal tissue was used as a source of total RNA, isolated by phenol/1-bromo-3-chloropropane (BCP1) extraction procedure (TriReagent, Sigma, Poole, UK). Total RNA was transcribed to cDNA using MMLV reverse transcriptase (ImProm-II, Promega) primed with oligo(dT) and random hexamers. The nucleotide sequences of salmon *elovl5a* [GenBank: AY170327] and *elovl5b* [GenBank: FJ237531], zebrafish (*Danio rerio*) [GenBank: NM_200453], Atlantic bluefin tuna (*Thunnus thynnus*) [GenBank: HQ214237] and European seabass (*Dicentrarchus labrax*) [GenBank: FR717358.1] *elovl5* cDNA were aligned, and highly conserved regions identified for the design of oligonucleotide primers. A pike *elovl5* partial cDNA was then obtained using a nested PCR design. To obtain the full-length cDNA, rapid amplification of cDNA ends (RACE) with nested Elovl5-specific primers and the SMART 3' and 5' oligonucleotide primers (SMART RACE cDNA Amplification Kit, Clontech) was performed, using TaKaRa polymerase (*TaKaRa LA Taq* Hot Start Version, Takara Bio Inc., Clonetech). PCR products were ligated into plasmid pCR2.1 (TA Cloning Kit, Invitrogen), sequenced and data assembled using SeqMan (SeqMan II, Lasergene DNASTAR) to determine the full-length cDNA sequence of pike *elovl5*. Finally, the entire coding sequence of pike *elovl5* was amplified from brain cDNA using the high fidelity Pfu Turbo DNA polymerase (Stratagene, Agilent Technologies, Cheshire, UK) by performing a nested PCR. The first PCR was performed with primers designed in the untranslated regions (UTR), PIKE5OUTF and PIKE5OUTR (TABLE 2.1). PCR conditions consisted of an initial denaturing

step at 95 °C for 1 min, followed by 30 cycles of denaturation at 95 °C for 45 s, annealing at 60 °C for 45 s, and extension at 72 °C for 2 min and a final extension at 72 °C for 5 min.

2.2.3 Sequence and phylogenetic analyses

2.2.3.1 Multiple sequence alignment

Generally, it is more informative to compare protein sequences as they can identify homologous sequences from organisms that last shared a common ancestor over 1 billion years ago compared to DNA sequence comparison (600 myr ago). This is because many changes in the DNA sequences either do not impose a change in the encoded amino acid (aa), or they result in a mismatched aa with related biophysical properties, which is accounted by scoring alignment methods. The aa sequence deduced from pike *elovl5* cDNA was aligned with salmonid orthologs, including Atlantic salmon, rainbow trout (*Oncorhynchus mykiss*) [GenBank: NM_001124636.1], and masu salmon (*Oncorhynchus masou)* [GenBank: DQ067616.1] using ClustalW (BioEdit 7.1.3, Tom Hall, Ibis Biosciences, Abbott Laboratories) (Thompson *et al*., 1994). The ClustalW algorithm (FIGURE 2.1) is the most popular progressive multiple sequence alignment, consisting of the calculation of pairwise sequence alignments scores between all the sequences in the alignment, then beginning the alignment with the two closest sequences and progressively adding more. In stage one, the global alignment approach of Needleman and Wunsch was used to create pairwise alignment scores of all sequences applying the BLOSUM62 protein similarity matrix. BLOSUM (blocks substitution matrix) is a very common set of scoring matrices that assign a score according to the probability

of mutation and the biophysical properties of amino acids. Among such matrices, BLOSUM62 performs better than the majority at detecting distant relationships between proteins and is thus utilised as the default-scoring matrix in most databases like BLAST. In the second stage, a guide tree (not considered to be a phylogenetic tree) was calculated from the similarity (or distance) matrix, which reflects the relatedness of all the proteins defined in the next section indicating the order in which the sequences have to be added to the multiple alignment. And finally, the multiple sequence alignment was created by aligning the protein sequences according to their distance in the phylogenetic tree.

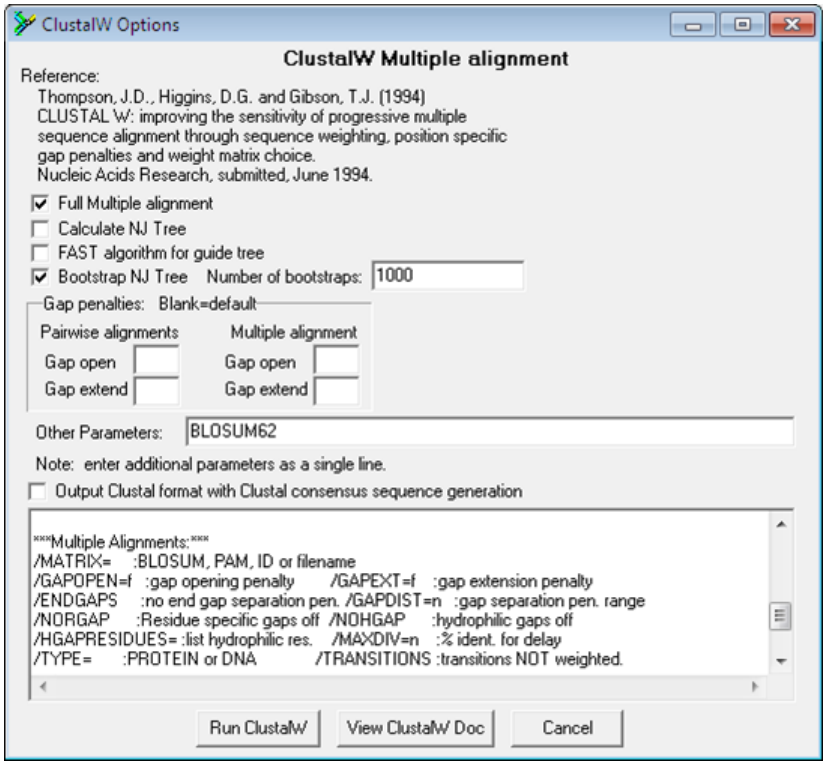

FIGURE 2.1 Screen capture of a multiple sequence alignment applying the ClustalW algorithm in BioEdit (Thompson *et al*., 1994).

TABLE 2.1 Details of primer pairs (restriction sites for *Xba*I and *Bam*HI, and fluorescent CAG and M13R binding sites underlined) used for the cloning of pike Elovl5 ORF in pYES2, qPCR analysis of tissue expression and for the genotyping of Atlantic salmon paralogous genes.

Aim	Transcript	Primer	Primer sequence	Amplicon size	Ta	Accession no.
elovl5 cDNA cloning	Elovl5	UNIE5OUTF	5'- ATGGATGGGTCCCAGAGA -3'	437 bp	55°C	JX272634[a]
		UNIE5OUTR	5'- AGTTCATAACGAACCACCAGAT -3'			
		UNIE5INNF	5'- TGGGGCCCAAGTACATGA -3'	388 bp	55°C	
		UNIE5INNR	5'- TGGACGAAGCTGTTAAGGG -3'			
5' RACE	Elovl5	PIKE5_5'F	5'- GATGGCAGAGAGCCCATAGT -3'	639 bp	55°C	JX272634[a]
		PIKE5_5'R	5'- CCACACAGCAGACACCATCT -3'	336 bp	55°C	
3' RACE	Elovl5	PIKE5_3'F	5'- AGATGGTGTCTGCTGTGTGG -3'	1,118 bp	60°C	JX272634[a]
		PIKE5_3'R	5'- ATGCTCAACATCTGGTGGTTT -3'	915 bp	60°C	
ORF cloning	Elovl5	PIKE5OUTF	5'- GCCCAGGTTCGCATCACCCAG -3'	1,337 bp	60°C	JX272634[a]
		PIKE5OUTR	5'- ATTCCGGGGGTCATTTGAGATAGACG -3'			
		PIKE5CDS_F	5'- CCCGAGCTCGGATCCAAATGGAGACTTTTAATCAGAGACTTAACACC -3'	922 bp	60°C	
		PIKE5CDS_R	5'- GGGTCTAGACTCGAGCTTCAGTCCGCCCTCACTTTCCT -3'			
qPCR	18S	qPCRp18SF	5'- TTCGAATGTCTGCCCTATCAAC -3'	128 bp	55°C	Contig2237[b]
		qPCRp18SR	5'- CCTTCCTTGGATGTGGTAGC -3'			
	Elovl5	qPCRpELO5F	5'- CCTTTGCACTGACCGTGATA -3'	195 bp	56°C	JX272634[a]
		qPCRpELO5R	5'- GCGTGTCCTGGCAGTAGAA -3'			
	Efl-a1	qPCRpEFLF	5'- AAGATCGACCATCGTTCTGG -3'	209 bp	55°C	GH265867[a]
		qPCRpEFLR	5'- CTGGCAGCCTTCTTATCGAC -3'			

[a] GenBank (http://www.ncbi.nlm.nih.gov/)

[b] University of Virginia, Centre for Medical Research (http://lucy.ceh.uvic.ca/contigs/cbr_contig_viewer.py)

2.2.3.2 Northern pike Elovl5 protein topology

Northern pike Elovl5 membrane topology was predicted using PredictProtein server (Rost *et al*., 2004), which applies the hydrophobicity PHD method (HeiDelberg). The first step in a PHD prediction is generating a multiple sequence alignment using evolutionary data from a broad spectrum of homologous sequences applying the MaxHom algorithm that generates a pairwise profile-based multiple alignment, which incorporates structural information about the homologous members of the group of proteins aligned. The transmembrane helices are predicted with an expected average accuracy of ~ 98%.

2.2.3.3 Phylogenetic tree

A phylogenetic tree was constructed between salmon *elovl5* coding sequences and other vertebrate *elovl5*-like ORF on the basis of the nucleotide sequence alignment in order to take account of single nucleotide substitutions, that is including silent aa mutations that would be otherwise missed in protein trees. The vertebrate *elovl*5 sequences used to construct the phylogenetic tree corresponded to: *Homo sapiens* [NM_021814], *Mus musculus* [GenBank: NM_134255], *Xenopus laevis* [GenBank: NM_001096414], and teleosts *T. thynnus* [GenBank: HQ214237], *Thunnus maccoyii* [GenBank: GQ204105], *Pagrus major* [GenBank: HQ415605], *D. labrax* [GenBank: FR717358], *Rachycentron canadum* [GenBank: FJ440239], *Lates calcarifer* [GenBank: GQ214180], *Oncorhynchus masou* [GenBank: DQ67616], *Oncorhynchus mykiss* [GenBank: NM_001124636], *S. salar elovl5a* [GenBank: AY170327, *S. salar elovl5b* [GenBank: FJ237531]; *Esox lucius* [GenBank: JX272634], *Gadus morhua* [GenBank: AY660881] and *D. rerio* [GenBank:

NM_200453]. The tree was rooted to the soil fungi *Mortierella alpina* PUFA elongase [GenBank: AF206662] as the outgroup species applying the neighbour-joining (NJ) method (Saitou and Nei, 1987), and using the maximum composite likelihood (MCL) substitution model in MEGA4 (Tamura *et al*., 2007) (FIGURE 2.2). The NJ method is the preferred distance-based algorithm as its fast performance allows bootstrapping, and because it does not assume the molecular clock theory. The NJ method starts by generating a starlike structure with no hierarchical structure, and then creating a matrix of pairwise distances for all the aligned OTUs (operational taxonomic units) based upon the likelihood for each DNA residue to occur. This approach, termed MCL, accounts for rate substitution differences between nucleotides (heterogeneous model, i.e. transitional and transversional substitutions) and inequality of nucleotide frequencies (gamma model), and allows estimation of the evolutionary distances between all pairs of OTUs simultaneously and independently, thus reducing errors and inferring phylogenies more accurately (Tamura and Nei, 1993; Tamura *et al*., 2007). This constitutes the main difference between a phylogenetic tree and a guide tree, used to create the multiple sequence alignment, as the guide tree does not include a model accounting for multiple substitutions. The NJ method links with a node the two most closely related OTUs that exhibit the smallest sum of branch distance. Then, these OTUs are treated as a single OTU, and the model identifies the next pair of OTUs that give the smallest sum of branch lengths according to the MCL algorithm and so on until the topology of the tree is completed. The robustness of the tree was evaluated by bootstrapping 10,000 randomised trees, which calculates the frequency with which each clade is observed in the original tree. Although the NJ method produces both a topology and

branch lengths, we omitted the branch length because nodes and OTUs represent speciation events and species, respectively, which might not correlate with the evolution of genes.

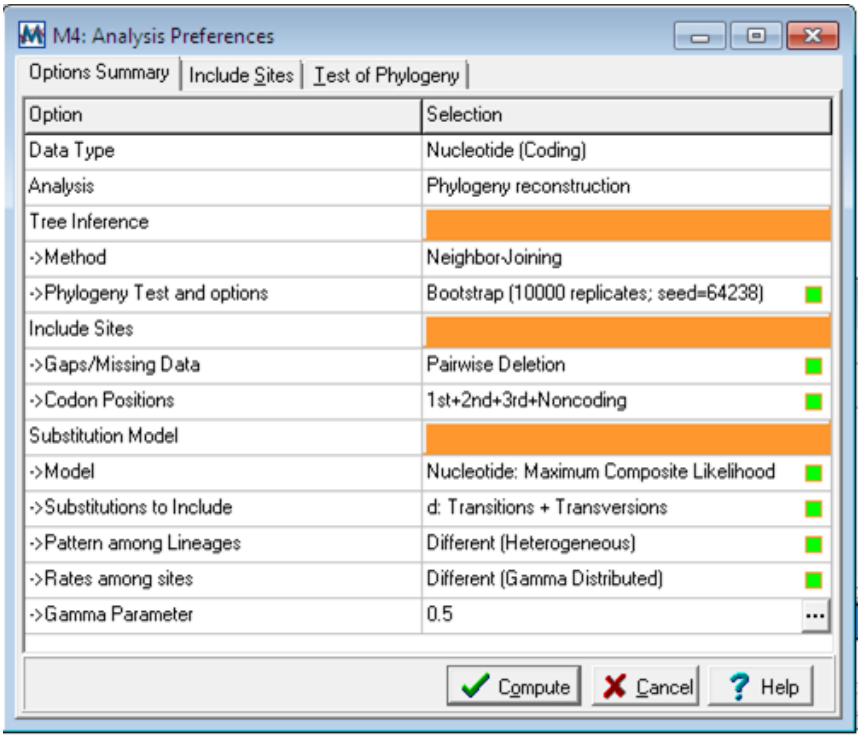

FIGURE 2.2 Screen capture of the construction of a phylogenetic tree using the Neighbour Joining (NJ) method (Saitou and Nei, 1987), and the maximum composite likelihood (MCL) substitution model using MEGA4 (Tamura *et al*., 2007).

2.2.4 Evolution rate and tests of selection

At the molecular level, evolution operates both at the DNA level and at the protein level. Duplicated genes encoding enzymes might suffer subsequent nucleotide changes allowing one to diverge at a faster rate, and to develop a novel function that is advantageous and thus selected (Pevsner, 2009).

2.2.4.1 Tajima's relative rate test

The relative evolution rate of nucleotide and amino acid substitution of the Atlantic salmon *elovl5a* (A) and *elovl5b* (B) was tested using Northern pike *elovl5* (C) as an outgroup in Tajima's relative rate test in MEGA4 (Tajima, 1993; Tamura *et al*., 2007). Tajima's test performs a chi-square (χ^2) to determine whether the rates of evolution of the two compared sequences are comparable (H_o), or not (H_A), thus disrupting the rate of the molecular clock. The molecular clock theory suggests that sister lineages diverge at an equal rate after a speciation process, although we applied the test to the divergence level of duplicated genes (Pevsner, 2009). The test measure the number of nucleotide sites m_1 in which residues in sequence A differ from those in B and C; similarly m_2 corresponds to sites in B that are different than A and C (Pevsner, 2009). Given that C is an outgroup, the expectation of m_1 must equal the expectation m_2, that is:

$$E(m_1) = E(m_2)$$

This equality is tested with a chi-square analysis:

$$\chi^2 = (m_1 - m_2)^2 / m_1 + m_2$$

2.2.4.2 Positive and negative selection

The influence of selection on *elovl5* genes in Atlantic salmon was assessed by determining the corrected rates of synonymous substitutions (silent, dS) versus nonsynonymous substitutions (amino acid-altering, dN) per site in the aa coding sequence of salmon Elovl5 and vertebrate orthologs (Nei and Gojobori, 1986). For a nucleotide change in a single codon, a synonymous substitution does not result in a change of the amino acid that is specified, whereas a nonsynonymous substitution does. The "acceptance rate" (ω = dN/dS) termed by Miyata and Yasunaga (1980) is

related to evolutionary constraints at the protein level for which a value of $\omega > 1$ indicates evidence of Darwinian positive selection, whereas $\omega < 1$ suggests negative or purifying selection, and $\omega = 1$ neutral selection (Yang *et al*., 2000). Darwin's theory of evolution suggests that at the phenotypic level, traits in a population that enhance survival are selected for (positive selection), while traits that reduce fitness are selected against (negative or purifying selection).

For greater confidence, a variety of stochastic algorithms were used to assess whether selection had occurred based on the dN/dS ratio of codon-aligned nucleotide sequences. Selection tests calculate the proportion of observed synonymous (pS) and non-synonymous (pN) substitutions between a pair of homologous sequences by comparing them codon by codon as follows:

$$pS = Sd / S$$

$$pN = Sn / N$$

Sd and Sn, the number of observed synonymous and non-synonymous substitutions, respectively; S and N, the number of potential synonymous and non-synonymous substitutions (the average of substitutions per codon for the two compared sequences), respectively; pS and pN are the proportion of observed synonymous and non-synonymous substitutions, respectively.

Then, the proportion of substitutions is corrected applying the Jukes-Cantor correction resulting in dS and dN corrected values for synonymous and non-synonymous mutations, respectively. The synonymous non-synonymous analysis programme (SNAP) (Korber, 2000) employs the Nei and Gojobori method (Nei and Gojobori, 1986) to estimate the overall dS and dN values of two sequences with a fairly accurate efficiency, unless the rates of transition and transversion are quite

different (Ota and Nei, 1994). In detecting Darwinian selection, is it often necessary to compare the means and variation of dS and dN through the construction of a phylogenetic tree; however SNAP incorporates a statistic that allows a faster and accurate estimation of the variances and covariances of dS and dN in the absence of a phylogenetic tree (Ota and Nei, 1994). This method allows the use of a best estimate of an ancestral sequence to perform pairwise comparisons between the predicted ancestor and the rest of the sequences, providing pairwise ω values and an overall ω for all the sequences analysed. Similarly, Datamonkey (Kosakovsky Pond and Frost, 2005) is a publically available set of tools implemented by HyPhy package that include robust maximum likelihood approaches to determine the total species dN/dS ratio. Instead of addressing "where" in the gene has selection acted as most methods implemented by Datamonkey, the GA-branch genetic algorithm assigns different ω values to lineages answering "when" in the past has selection acted by assigning rate classes of ω to lineages in a phylogenetic context (Delport *et al*., 2010). The HKY85 substitution model (Kosakovsky Pond and Frost, 2004) implemented by the GA-branch analysis allows for transversion and transition substitution bias and varying nucleotide frequencies between the position within a codon, opposite to SNAP.

2.2.5 Substrate specificity of pike Elovl5

2.2.5.1 Heterologous expression of pike elovl5 in yeast

The pike Elovl5 coding sequence was obtained by PCR from full-length cDNA with primers containing restriction sites, PIKE5ORF_F (*Bam* HI) and PIKE5ORF_R (*Xho* I) (TABLE 2.1). This fragment and the yeast galactose-induced

expression plasmid pYES2 (Invitrogen) were digested with the corresponding restriction endonucleases (New England BioLabs, Herts, UK), and ligated using T4 DNA Ligase (Bioline, London, UK). The resulting plasmid construct pYES2–Elovl5 was transformed into *Saccharomyces cerevisiae* (strain INVSc1) using the S.C. EasyComp Transformation kit (Invitrogen) and incubated for 3 days at 30ºC in yeast drop-out plates$^{-\text{ uracil}}$. A single colony of transgenic yeast was grown in *S. cerevisiae* minimal medium for 24 h prior supplementation with 2% D-galactose and one of the following FA substrates diluted in an aqueous detergent solution: ALA (18:3n-3), LA (18:2n-6), stearidonic acid (SDA, 18:4n-3), γ-linolenic acid (GLA, 18:3n-6), EPA (20:5n-3), ARA (20:4n-6) and docosapentaenoic acid (DPA, 22:5n-3) and adrenic acid (AA, 22:4n-6). Substrate FA final concentrations consisted of 0.5 mM for C18, 0.75 mM for C20 and 1 mM for C22 fatty acids. The fatty acid substrates (> 99% pure) and chemicals used to prepare the *S. cerevisiae* minimal medium$^{-\text{uracil}}$ were purchased from Sigma Chemical Co. Ltd. After 2 days, yeast were harvested and washed with HBSS with 1% fatty acid-free bovine serum albumin (FFA BSA) prior to lipid extraction and FA analyses. Yeast transformed with pYES2 containing no insert were cultured under the same conditions described above and used as control treatments.

2.2.5.2 FAME analysis by GC–MS

Yeast samples were homogenised in chloroform/methanol (2:1, v/v) containing 0.01% BHT as antioxidant and total lipid extracted following the Folch method (Folch *et al*., 1957). Fatty acid methyl esters (FAMEs) were subsequently prepared, extracted and purified (Christie, 2003). FAMEs were identified and quantified using a gas chromatograph (GC8000) coupled to a MD8000mass

spectrometer (ThermoFisher Scientific, Hemel Hempstead, UK) as described previously (Monroig *et al*., 2010a). Elongation activities on PUFA substrates were calculated as the proportion of substrate FA converted to elongated FA product(s):

$$\% \text{ Conversion} = [\text{product area} / (\text{total product area} + \text{substrate area})] \times 100$$

Conversion rates from 24:0 were not calculated as yeast endogenously contains several of the FA involved in the elongation pathway and thus individual saturated FA (≥ C24) contents from Elovl4-transformed yeast were compared to those of control yeast (Monroig *et al*., 2010a).

2.2.6 Tissue distribution of pike Elovl5

Tissue expression of Elovl5 was determined by quantitative real-time PCR (qPCR). Total RNA was extracted from brain, intestine, liver, adipose tissue, white muscle, kidney, spleen, heart and gill from juvenile Northern pike (n = 3) using Tri Reagent (Sigma). RNA quality and quantity were assessed by electrophoresis and spectrophotometry (Nanodrop 1000, Thermo Scientific, Wilmington, USA). One microgram of DNase (DNA-free kit, Ambion, Applied Biosystems, Warrington, UK) treated total RNA was reverse transcribed into cDNA using SuperScriptTM III Reverse Transcriptase (Invitrogen) and primed with 2.5 μM of oligo(dT) and 7.5 μM of random hexamers (AB Applied Biosystems). Following primer annealing at 25 °C for 5 min, and cDNA synthesis at 55 °C for 1 h, reactions were stopped by heating at 70 °C for 15 min, and cDNA was diluted 5-fold with nuclease-free water. The *elovl5* qPCR primers were designed to anneal to each of two predicted exons spanning an intron, whereas the *elf-1α* (reference gene) primer pair was designed in a region corresponding to the coding sequence. TABLE 2.1 shows the sequence of the

primers used, their specific annealing temperatures, the size of fragments produced, and the reference sequences used for primer design, using the Primer3 software (Rozen and Skaletsky, 1998). qPCR analyses were performed using a Mastercycler RealPlex^2 (Eppendorf, Stevenage, UK) in a final volume of 20 µl containing 5 µl of diluted cDNA (1/20 for *elf-1α* and *elovl5*, and 1/5,000 for 18S rRNA), 0.5 µM of each primer, and 10 µl of SensiMix SYBR No-ROX (Bioline). Amplifications were carried out with a systematic negative control (NTC, no template control) containing no cDNA. Standard curves with dilutions: 1/5, 1/10, 1/20, 1/50, 1/100, 1/200, and 1/500 were prepared for *elf-1α* and *elovl5*, whereas dilutions ranging from 1/10 to 1/1,000,000 were made up for 18S rRNA given its predominant abundance in the total RNA. Thermal cycling was initiated at 95 °C for 10 min, followed by 40 cycles: 15 s denaturing step at 95 °C, 15 s at the specific primer pair annealing Tm (TABLE 2.1), and 15 s extension at 72 °C. After the amplification phase, a melting curve from 60 °C to 95 °C was performed, enabling confirmation of amplification of a single product in each reaction. The qPCR product sizes were determined by agarose gel electrophoresis and their identity confirmed by sequencing. No primer-dimer formation occurred in the NTC.

Normalised expression values were generated by the ΔCt method (Pfaffl, 2001) using the geometric mean of the Ct values of the two reference genes (*elf-1α* and 18S rRNA). Results were expressed as the percentage of normalised expression relative to the sum of expression across all tissues tested. Normalised expression values of salmon *elovl5a* and *elovl5b* reported in Morais *et al*. (2009) were treated in the same way. Differences in the expression of pike *elovl5* among different tissues were determined by one-way analysis of variance (ANOVA) followed by Tukey's

multiple comparison test at the significance level of $P \leq 0.05$ (PASWS 18.0, SPSS Inc., Chicago, USA).

2.2.7 Linkage analysis of duplicated loci in Atlantic salmon

2.2.7.1 Genotyping of elovl5 microsatellites

BlastN searches of *elovl5a* [GenBank: AY170327] and *elovl5b* [GenBank: FJ237531] cDNA sequences identified larger contig fragments for both genes (*elovl5a* [GenBank: AGKD01037727], *elovl5b* [GenBank: AGKD01030045]) within the preliminary Atlantic salmon genome assembly [GenBank: GCA_000233375.1]. RepeatMasker (Smit *et al.*, 2010) was used to identify potentially informative microsatellite markers within both sequences. A large dinucleotide repeat $(CA)_{36}$ was located 0.5 kb downstream of *elovl5a*, while the closest marker identified for *elovl5b* was a trinucleotide repeat, $(TTA)_{10}$, situated 13 kb downstream. PrimerSelect software (DNASTAR Inc., USA) was used to design flanking PCR primers for allele detection (TABLE 2.1). Fluorescently-labelled tailed primers (Boutin-Ganache *et al*., 2001) were used (detailed in TABLE 2.1), the allelic products being detected and sized by capillary electrophoresis (CEQ-8800 Genetic Analysis System, Beckman Coulter Inc., USA). The loci were amplified separately, with 0.5 µL of both reactions being pooled later and electrophoresed in duplex. Each reaction (6 µL final volume) comprised 50 ng template DNA, 20 nM specific forward primer (tailed), 300 nM each of specific reverse primer and fluorescently labelled tail primer and 3 µL 2× MyTaq HS mix (a *Taq* polymerase mastermix from Bioline Reagents Ltd, UK). Hot-start PCR was carried out on a T-gradient thermocycler, (Biometra GmbH, Germany) using the following profiles 95ºC for 1 min (for Taq activation and initial

denaturation), followed by 34 cycles consisting of 95ºC for 15 s, annealing at 58ºC (*elovl5b*), or 56ºC (*elovl5a*) for 30 s, and extension at 72ºC for 45 s.

2.2.7.2 Segregation and linkage analyses

The parents of six available Atlantic salmon pedigree panels were screened for both loci, with two of these (Br5 and Br5/2) being identified as informative for linkage analysis (i.e. at least one parent being heterozygous for both loci, AABB × AA'BB'). A total of 46 progeny from family Br5 were screened (more than 400 loci having been previously mapped for this panel (SalMap project, Danzmann *et al*., 2005), while 31 progeny available for family Br5/2 were genotyped. In linkage studies, genetic markers are used to search for coinheritance of chromosomal regions within families. Two genes that are in proximity in a chromosome will frequently co-segregate during meiosis. Following the pattern of transmission on a set of markers in a linkage analysis can be used to localise a gene based on its linkage to a genetic marker locus. Thus, Mendelian and joint segregation analysis (Mather, 1951), as implemented in Taggart and Ferguson (1984) were carried out on both panels:

Test cross

a_1 = observed progenyAABB

a_2 = observed progenyAABB'

a_3 = observed progenyAA'BB

a_4 = observed progenyAA'BB'

N = total informative progeny = $a_1 + a_2 + a_3 + a_4$

χ^2_A= test for departure from disomic 1:1 segregation at the A locus

$= (a_1 + a_2 - a_3 - a_4)^2 / N$ (d.f. = 1)

χ^2_B= test for departure from disomic 1:1 segregation at the A locus

$= (a_1 - a_2 + a_3 - a_4)^2 / N$ (d.f. = 1)

χ^2_L = test for departure from independent assortment of the two loci

$= (a_1 - a_2 - a_3 + a_4)^2 / N$ (d.f. = 1)

where *r* is the fraction of non-parental genotypes = $(a_2 + a_3) / N$, SE is the standard error of $r = \sqrt{r (1 - r) / N}$ and d.f. are the degrees of freedom.

Linkage mapping was performed on the SalMap reference family (Br5) using JoinMap (Stam, 1993), which applies the LOD (logarithm of odds) score approach, especially useful with a small number of progeny. The LOD uses a maximum likelihood method to calculate the ratio of the probability of obtaining the test data if the two loci are linked, to the probability of observing the test data if the two loci are not linked (Morton, 1955; Ott, 2001) as follows:

$$LOD = log_{10} \frac{probability\ of\ birth\ sequence\ with\ a\ given\ linkage\ vaue}{probability\ of\ birth\ sequence\ with\ no\ linkage} = log_{10} \frac{(1-\theta)^{NR} \times \theta^{R}}{* 0.5^{(NR+R)}}$$

where θ is the probability of recombination between the two loci ($0 < \theta < 1$; nearly always, $\theta < 1/2$), NR is the non-recombinant offspring and R is the recombinant offspring. *Any alleles that are completely unlinked have a 50% chance of recombination, due to independent assortment.

2.3 RESULTS

2.3.1 Northern pike Elovl5 sequence and phylogenetics

A 1,434bp full-length cDNA sequence (5'UTR 72 bp, ORF 888 bp, 3'UTR 474 bp) was obtained by 5' and 3' RACE PCR and submitted to the GenBank

database under the accession number JX272634. The pike Elovl5 open reading frame (ORF) encodes a putative protein of 295 aa that shares 69.7% to 71% aa identity to mammalian and reptilian orthologues including human [GenBank: NM_021814], mouse [GenBank:NM_134255] and the frog *Xenopus laevis* [GenBank: NM_00109614]. Phylogenetic analysis shows that teleost *elovl5* genes cluster according to accepted taxonomy as displayed in the phylogenetic tree (FIGURE 2.3) with Protacanthopterygii including Salmoniformes and Esociformes forming a clade and thus in agreement with phylogenetic analysis performed upon whole mitochondrial genomes (Ishiguro *et al*., 2003). Among all teleosts, pike exhibit the highest amino acid identity scores with the salmonid Elovl5 members, with Atlantic salmon Elovl5a and Elovl5b being the most similar (86.4%) and dissimilar (83.4%), respectively. Lower identity values were observed in comparison with Elovl5 sequences of species belonging to orders other than Salmoniformes ranging from 73% (*Gadus morhua*) to 80% (*Lates calcarifer*). All fish *elovl5* grouped together with reptilian and mammalian homologs, and more distantly from other members of the *elovl* family (not included in the phylogenetic tree, Carmona-Antoñanzas *et al*., 2011).

The pike Elovl5 deduced amino acid sequence contains the three typical features present in all Elovl members: a single HXXHH histidine box motif, a carboxyl-terminal targeting signal responsible for the retention of transmembrane protein to the endoplasmic reticulum (ER), and multiple putative transmembrane-spanning domains containing hydrophobic aa stretches. The best hydrophobicity model predicted 5 transmembrane helices (transmembrane domain ≥ 20 aa) in accordance with previous analysis using the GES algorithm (Engelman *et al*., 1986;

Tvrdik *et al*., 2000). However, these two methods compute protein polarity scores based upon different chemical arguments resulting in slightly different transmembrane boundaries (± 2 aa). Thus, for greater reliability the transmembrane domains depicted in FIGURE 2.4 represent the overlapping regions described by both methods. Additionally, 16 out of the 17 aa residues that have been established to be highly conserved across 22 members of the Elovl family (Leonard *et al*., 2000) were identified in all protacanthopterygian Elovl5 proteins.

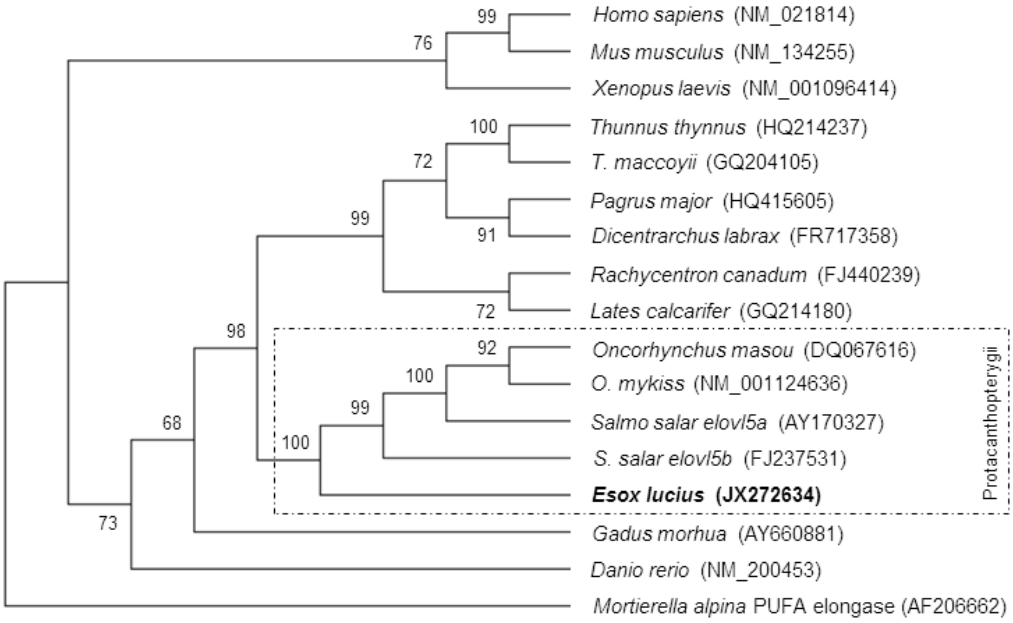

FIGURE 2.3 Phylogenetic tree revealing the relative position of Northern pike Elovl5 according to proteins from other vertebrate orthologs, and rooted to *Mortierella alpina* PUFA elongase as outgroup. The tree was constructed on the DNA sequences extracted from GenBank using MEGA4 and applying the Neighbour Joining method. The numbers on the branches represent the frequencies (%) with which the presented tree topology was obtained after bootstrapping (10,000 iterations).

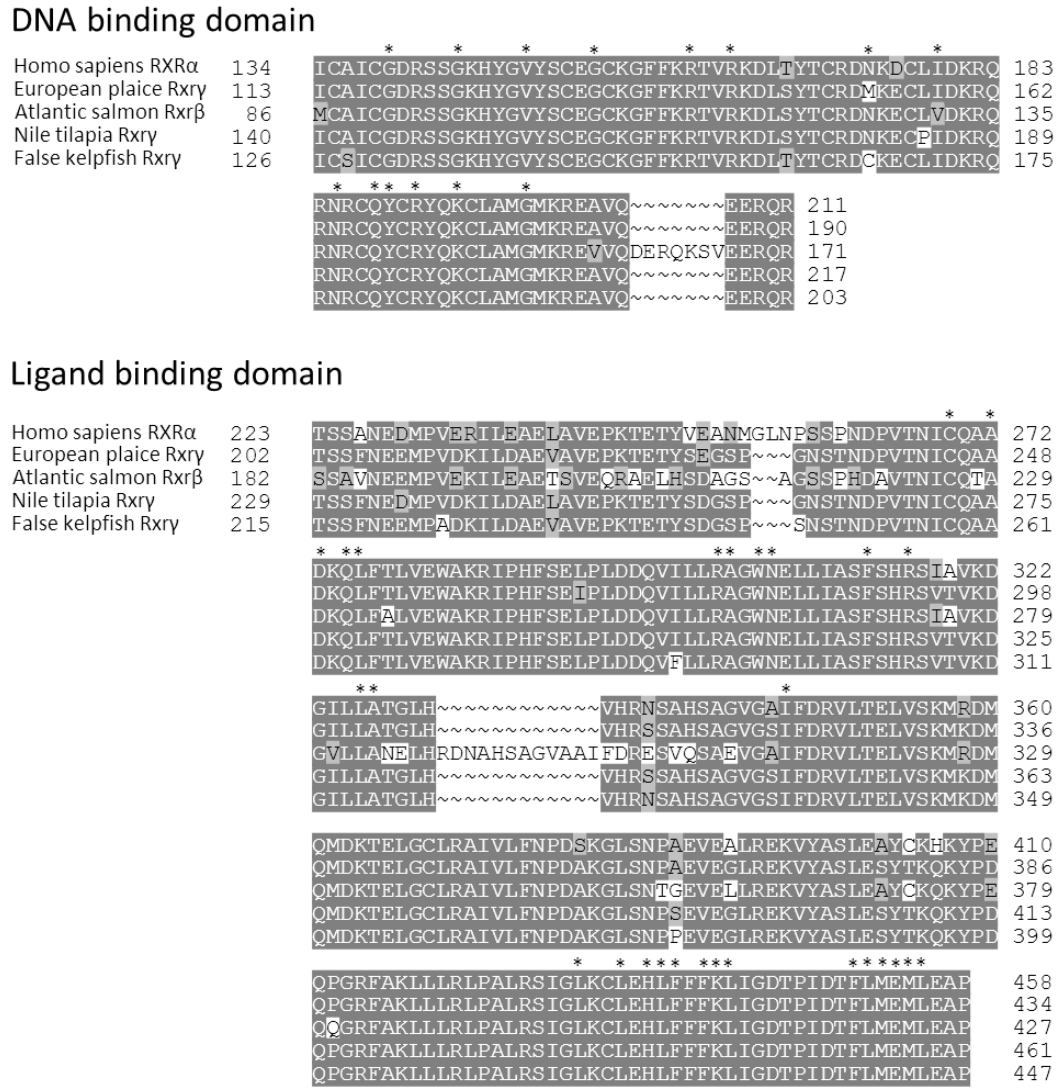

FIGURE 2.4 Alignment of the deduced amino acid sequence of the Northern pike Elovl5 with orthologs of other members of the Protacanthopterygii lineage. The amino acid alignment was performed using the BLOSUM62 matrix from BioEdit, and identity/similarity was calculated based on a 60 % identity threshold. Identical residues are shaded in grey, and altered residues are shaded in white if they exhibit the same chemical qualities, or black if they do not. Outlined are the HXXHH histidine box, and the endoplasmic reticulum retention signal (ER); the five putative transmembrane domains (I-V) are dash-underlined; the predicted "catalytic site" is indicated by the solid arrow above the sequences; and an asterisk indicates each of the 17 amino acid residues conserved across Elovl proteins (Leonard *et al*., 2000).

2.3.2 Purifying selection on salmonid Elovl5 paralogs

The number of synonymous (dS) and nonsynonymous substitutions (dN) per site was determined by comparing the Northern pike ORF sequence to each of the

duplicate Atlantic salmon ORF, and the salmon duplicate ORF to each other. The selection tests indicated that negative (purifying) selection was the major evolutionary force acting on the salmon duplicates since their divergence from Northern pike, with ω equal to 0.24, or 0.20 when using the GA-branch (dN>dS, $P < 0.01$), or SNAP (dN = 0.064, dS = 0.330) approaches, respectively. Accordingly, the average ω between all vertebrate members included in the phylogenetic tree also confirmed overall purifying selection (ω < 1) (GA-branch, $P < 0.01$; SNAP, dN = 0.178, dS = 1.434). When salmon paralogs were compared to one another using pike Elovl5 amino acid sequence as the outgroup the results indicated that both duplicates exhibit comparable evolutionary rates with molecular clock-like behaviour ($\chi^2 = 3.56$, $P > 0.05$). In contrast, the results obtained when Tajima's test was performed on the nucleotide data of the aforementioned sequences showed that the salmon *elvol5* sequences are evolving asymmetrically ($\chi^2 = 6.75$, $P < 0.05$) (TABLE 2.2).

This suggested that, despite the fact that the salmon elongases appear to be subjected to functional constraints in order to maintain the functionality of the protein, the nucleotide sequence of one of the duplicates seems to be diverging faster than the other. The ORF sequences of the vertebrate *elov5* members included in this study were codon-aligned, and the accumulated dN and dS substitutions along the coding sequence assessed using SNAP. Results revealed substantial differences in dN throughout the protein sequence. A region corresponding with exon 5 (109-165 aa), which includes the catalytic histidine box, displayed 7 to 8-fold reduction in nonsynonymous substitutions with respect to the flanking regions (FIGURE 2.5).These results indicate that selective pressure is not constant along the coding sequence.

TABLE 2.2 Evolutionary rate between Atlantic salmon *elovl5a* and *elovl5b* tested using Northern pike as an outgroup in Tajima's relative rate test in MEGA4.

Configuration		Count (nt)	Count (aa)
Identical sites in all three sequences (m_{iii})		748	241
Divergent sites in all three sequences (m_{ijk})		5	6
Unique differences in Atlantic salmon *elovl5a* (m_{ijj})		15	5
Unique differences in Atlantic salmon *elovl5b* (m_{iji})		33	13
Unique differences in Northern pike *elovl5* (m_{iij})		81	29
Statistics	χ2 test statistic	6.75	3.56
	d.f.	1	1
	P-value	0.009	0.059

P-value < 0.05 is used to reject the null hypothesis of equal rates between lineages.
nt, nucleotide sequence; aa, amino acid sequence; d.f., degrees of freedom.

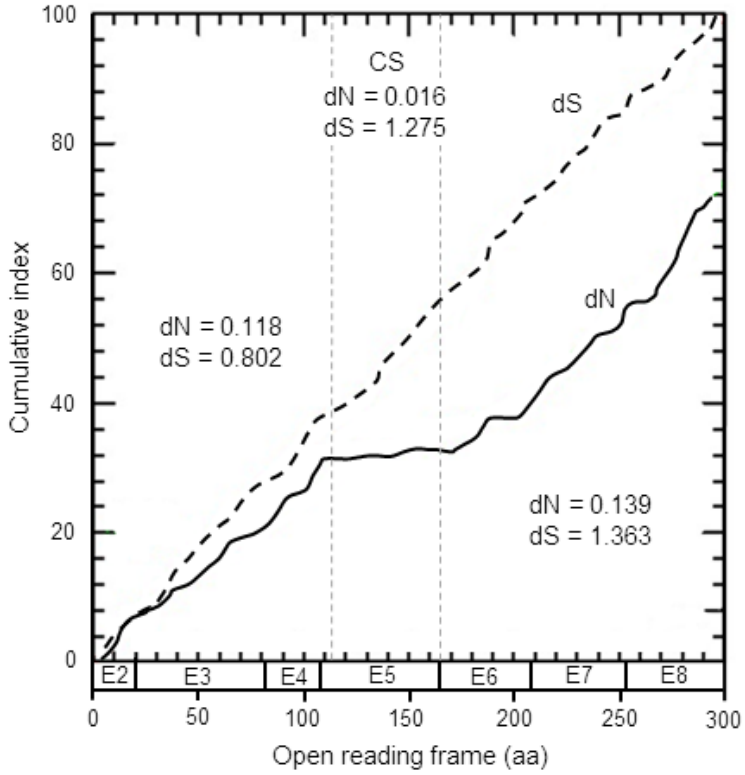

FIGURE 2.5 Average synonymous and nonsynonymous substitutions in vertebrate Elovl5 orthologs. Cumulative index of the synonymous (dS, dashed line) and nonsynonymous (dN, solid line) substitutions for all the pairwise comparisons are plotted against the open reading frame of vertebrate Elovl5 orthologs. SNAP was used under the default settings to compute the overall dS and dN for all the pairwise comparisons in the predicted "catalytic site" and flanking areas. CS, indicates the region of the predicted catalytic site. Exon boundaries are depicted along the X-axis, E2-E8.

2.3.3 Functional characterisation of pike Elovl5

The ability of Northern pike Elovl5 to elongate LC-PUFA of the n-3 and n-6 series was determined by the relative quantification of the fatty acid conversions obtained when transformed *Saccharomyces cerevisiae* containing either the empty pYES2 vector (control), or a vector with the pike Elovl5 ORF insert was grown in presence of potential PUFA substrates. Yeast cultures transformed with pYES2 containing the pike Elovl5 ORF and grown in the absence of PUFA substrates showed that pike Elovl5 is capable of efficiently converting the yeast endogenous monounsaturated fatty acids 16:1n-7 and 18:1n-9 to their elongated products determined by the presence of 18:1n-7 and 20:1n-9 constituting around 8% and 1% of the total fatty acids, respectively. The role of pike Elovl5 in the biosynthesis of LC-PUFA was investigated by culturing yeast transformed with pYES2-Elovl5 in the presence of C18 (18:3n-3, 18:4n-3, 18:2n-6, 18:3n-6), C20 (20:5n-3, 20:4n-6), or C22 (22:5n-3, 22:4n-6) PUFA substrates. Gas chromatography analysis demonstrated that the yeast transformed with empty pYES2 (control) did not have the ability to elongate LC-PUFA due to a lack of PUFA elongase activity (Agaba *et al*., 2005). However, in the presence of pike Elovl5 the C18 and C20 PUFA substrates were efficiently elongated to longer products (FIGURE 2.6), whereas C22 PUFA were elongated to a much lower extent not exceeding 4% conversion (TABLE 2.3). These results confirmed that pike Elovl5 is involved in the synthesis of LC-PUFA, and presents similar specificities to that described for both salmon Elovl5 paralogs. It was noteworthy that pike Elovl5 was able to convert 18:3n-3 and 18:2n-6 to 20:3n-3 and 20:2n-6, respectively, intermediates in the alternative (Δ8) pathway for the biosynthesis of EPA and ARA (Monroig *et al*., 2011), through subsequent

consecutive desaturations by Δ6Fad (Δ8 activity) to 20:4n-3 and 20:3n-6, and then Δ5Fad to EPA and ARA. However, Elovl5 showed higher activity towards the elongation of 18:4n-3 and 18:3n-6 with over 70% of each PUFA converted to C20 products, whereas 18:3n-3 and 18:2n-6 were elongated less efficiently with around 43% and 28% converted, respectively (TABLE 2.3).

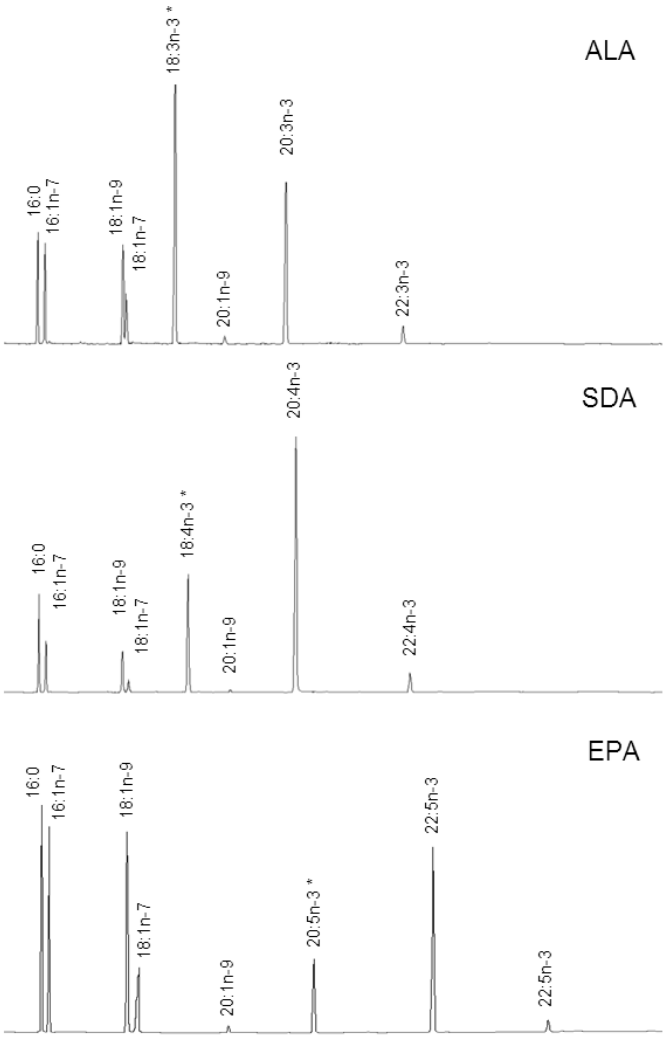

FIGURE 2.6 Identification of fatty acid elongation products in transgenic yeast (*S. cerevisiae*) transformed with pike Elovl5 ORF and grown in the presence of LC-PUFA substrates. Yeast were incubated in presence of PUFA substrates α-linolenic acid (ALA, 18:3n-3), stearidonic acid (SDA, 18:4n-3), or eicosapentaenoic acid (EPA, 20:5n-3), and fatty acid elongation products determined. Each panel shows the main endogenous fatty acids of *S. cerevisiae*, namely 16:0, 18:0, 16:1n-7 and 18:1n-9, and their elongated products 18:1n-7 and 20:1n-9. * Substrates and their corresponding elongated products are indicated accordingly in the panels. Vertical axis, FID response; horizontal axis, retention time.

TABLE 2.3 Functional characterisation of Northern pike Elovl5 elongase and its role in the biosynthesis of long-chain polyunsaturated fatty acids (LC-PUFA). Data are expressed as percentage of PUFA converted to elongated products at each step of the pathway and overall when yeast were transformed with pYES2-Elovl5 construct and incubated in presence of C18, C20 or C22 substrates. Atlantic salmon data extracted from Morais *et al*. (2009).

PUFA substrate	Product	Conversion (%)			Activity
		E. lucius Elovl5	*S. salar* Elovl5a	*S. salar* Elovl5b	
18:3n-3*	20:3n-3	39	-	-	C18 → 20
	22:3n-3	4	-	-	C20 → 22
	Total	43	-	-	
18:4n-3*	20:4n-3	67	56	58	C18 → 20
	22:4n-3	5	7	4	C20 → 22
	Total	72	63	62	
20:5n-3*	22:5n-3	67	36	68	C20 → 22
	24:5n-3	5	1	1	C22 → 24
	Total	72	37	69	
22:5n-3*	24:5n-3	4	1	1	C22 → 24
18:2n-6*	20:3n-6	26	-	-	C18 → 20
	22:3n-6	2	-	-	C20 → 22
	Total	28	-	-	
18:3n-6*	20:4n-6	67	43	65	C18 → 20
	22:4n-6	7	5	6	C20 → 22
	Total	74	48	71	
20:4n-6*	22:5n-6	35	23	48	C20 → 22
	24:5n-6	1	1	1	C22 → 24
	Total	36	24	49	
22:4n-6*	24:5n-6	1	1	1	C22 → 24

Data are expressed as percentage of PUFA converted to elongated products at each step of the pathway and overall when yeast were transformed with pYES2-Elovl5 construct and incubated in presence of C18, C20 or C22 substrates. Atlantic salmon data extracted from Morais *et al*. (2009).

2.3.4 Tissue distribution of pike Elovl5

The tissue distribution of pike Elovl5 mRNA transcripts was determined by real-time qPCR. For comparison, the normalised expression values of salmon *elovl5a* and *elovl5b* reported in Morais *et al*. (2009) were treated in the same way and plotted on the same graph. Results indicate the pike *elovl5* was expressed significantly

higher in brain ($P < 0.05$) (FIGURE 2.7). Thus, compared to expression in liver, expression of *elovl5* in brain was 1000-fold greater, and up to 30-fold higher than intestine. The expression levels in spleen, gill, kidney, white muscle, heart and adipose tissue were negligible, with liver exhibiting the lowest expression. From FIGURE 2.7 it is clear that the salmon *elovl5* genes are expressed in a very different pattern, with expression predominating in liver and intestine.

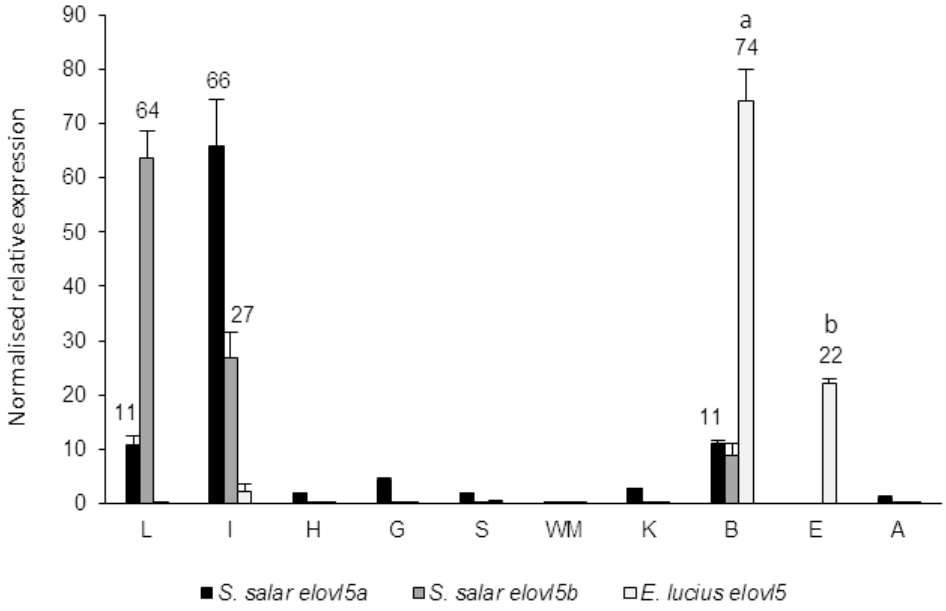

FIGURE 2.7 The tissue distribution of *elovl5* transcripts (mRNA) in Northern pike and Atlantic salmon. The relative number of the putative pike *elovl5* transcripts obtained by qPCR were normalised by the geometric mean of the relative copy numbers of the reference genes (*elf-1α* and 18S rRNA). Results are means (± SE) of analysis of three individual pike, and expressed as the percentage relative to the sum of expression across all tissues tested. Normalised expression values of salmon *elovl5a* and *elovl5b* reported in Morais *et al*. (2009) were treated in the same way. * $P \leq 0.05$ as determined by one-way ANOVA and Tukey's multiple comparison tests. L, liver; I, intestine; H, heart; G, gill; S, spleen; WM, white muscle; K, kidney; B, brain; A, adipose; E, eye (retina).

2.3.5 Segregation analysis and mapping of salmon *elovl5* duplicated loci

The amplicons derived from the *elovl5*-linked microsatellite primer sets resolved clearly and consistently. In both cases allelic size variants, consistent with

amplification of a single discrete locus, were detected. Among the pedigree panel parental samples, 12 different alleles were observed for *MS_ elovl5a* and four alleles for *MS_elovl5b*. A summary of the informative crosses is shown in TABLE 2.4 indicating some divergence from Mendelian segregation within both SalMap families, perhaps due to allele drop-out. Joint segregation analysis of the two panels (Br5 and Br5/2), informative for sire based linkage only, did not detect linkage between the two loci (H_o = independent assortment; P = 0.77 and 0.11 for Br5 and Br5/2, respectively) (TABLE 2.5). Genetic mapping of the two *elovl5* loci in the SalMap family (Br5) confirmed the Atlantic salmon paralogs to be located in distinct linkage groups (LG): *elovl5b* on LG 5 and *elovl5a* on LG 33 (LOD scores > 3.5).

2.4 DISCUSSION

The primary aim of this study was to characterise Northern pike Elovl5, a critical enzyme of LC-PUFA biosynthesis in vertebrates. The precise reasons for undertaking the work were ultimately to gain understanding of the evolutionary and ecological adaptations of salmonids. Phylogenetic evidence indicates that esocids are the nearest living relatives of salmonids, having diverged at some point prior to a WGD event in the common ancestor of all salmonids between 25 and 100 myr ago (Allendorf and Thorgaard, 1984; Koop and Davidson, 2008). As WGD has been widely suggested as a major enabling event in evolutionary innovation (Zhang, 2003), comparison of single preduplicated genes in pike with their duplicated paralogs in Atlantic salmon has the potential to shed light on evolutionary mechanisms and adaptation in salmonids. The genes of the LC-PUFA biosynthetic pathway are interesting candidates for studies of this type because both genetic and

biochemical evidence suggests that salmonids have a higher LC-PUFA biosynthetic capacity than many other fish species (Tocher *et al*., 2001, 2002). Atlantic salmon have duplicated genes for both desaturases and elongases of fatty acids and, in the case of desaturases, duplicates appear to have diverged and neo- or subfunctionalised to provide enzymatic activities for the entire LC-PUFA pathway (Leaver *et al*., 2008a; Monroig *et al*., 2010b).

In contrast, other fish species, particularly marine carnivorous species, are incapable of endogenous production of significant amounts of LC-PUFA because they lack critical genes of the biosynthetic pathway (Ghioni *et al*., 1999; Morais *et al*., 2009; Tocher, 2003). It has been suggested that this enhancement of LC-PUFA capacity in salmonids has been a factor in their success in colonising relatively nutrient poor freshwater environments (Leaver *et al*., 2008a). Salmonids can thrive, and are often the only fish group in flowing bodies of water, which are reliant on allochthonous terrestrial inputs as a major source of energy. Terrestrial inputs, whether directly from leaf litter or from insect species, are poor sources of LC-PUFA, especially DHA, in contrast with food sources in marine or eutrophic freshwater environments which are underpinned by blooms of phytoplankton rich in LC-PUFA (Ghioni *et al*., 1999; Ramsden *et al*., 2003; Tocher, 2003; Tocher *et al*., 2002). Despite the fact that the tissues of freshwater and marine fish are generally rich in C20 and C22 LC-PUFA, the strategies they utilise to fulfil such requirements vary depending on the species and the dietary input. Northern pike is a strictly freshwater species (with the exception of brackish water in Baltic Sea), whose distribution overlaps that of salmonids, and which shares a relatively recent common

TABLE 2.4 Summary of informative crosses for *elovl5a* and *elovl5b* including χ^2 test analyses for Mendelian segregation of *elovl5a* and *elovl5b* alleles.

Nº of crosses	Parental genotype		Progeny genotypes observed (expected)							*N*	X^2 value	d.f.	*P*
	Female	Male											
	elovl5b		*170/170*	*170/172*	*172/172*	*172/175*							
1[a]	170/170	170/172	13 (15.5)	18 (15.5)						31	0.806	1	0.369
1[b]	172/172	172/175			31 (23)	15 (23)				46	5.565	1	0.018
	elovl5a		*196/184*	*196/218*	*184/218*	*218/218*	*196/196*	*196/209*	*218/209*				
1[a]	196/218	184/218	6 (7.75)	5 (7.75)	9 (7.75)	11 (7.75)				31	2.935	3	0.402
1[b]	196/218	196/209		9 (11.5)			14 (11.5)	4 (11.5)	19 (11.5)	46	10.870	3	0.012

[a] SalMap family Br5/2 :

A = "196" & "218" alleles

A' = "184" allele

B = "172" allele

B' = "175" allele

[b] SalMap family Br5 :

A = "196" & "218" alleles

A' = "209" allele

B = "172" allele

B' = "175" allele

N, total informative progeny; d.f., degrees of freedom

TABLE 2.5 Joint segregation analysis between *elovl5a* and *elovl5b*. Only informative progeny is given.

Cross	Parents			Informative progeny				Joint segregation					
	Sex[a]	A locus	B locus	AA BB	AA BB'	AA' BB	AA' BB'	X^2_L	*N*	d.f.	*r*	± SE	*P*
		*elovl5*a	*elovl5*b										
Br5/2	<u>M</u>	AA'	BB'	9	4	7	11	2.613	31	1	0.35	0.09	0.106
	F	AA	BB										
Br5	<u>M</u>	AA'	BB'	16	15	7	8	0.087	46	1	0.48	0.07	0.768
	F	AA	BB										

[a] Informative parent underlined.

M, male; F, female; N, total informative progeny; d.f., degrees of freedom; r, fraction of non-parental genotypes (θ); SE, standard error.

ancestor with salmonids. However, pike differ from salmonids in exhibiting a far more piscivorous feeding behaviour. Hence, our particular interest in LC-PUFA biosynthetic genes in this species lies in comparison to those in Atlantic salmon.

Genetic linkage analyses established that the Atlantic salmon *elovl5* duplicates are located on different linkage groups: *elovl5a* on LG 33, and *elovl5b* on LG 5. Cytogenetic mapping using fluorescence *in situ* hybridisation has assigned salmon LG 33 and LG 5 to chromosomes 28 and 13, respectively (ASalBase, 2013; Phillips *et al*., 2009), both of which are acrocentric. While our data clearly show that the two *elovl5* loci are not physically linked, in a recent analysis (Lien *et al*., 2011) the Atlantic salmon chromosomes ssa28 and ssa13 were not homeologous and it is therefore not possible to conclude that the salmon *elovl5* paralogous genes are the result of a WGD. It is also possible that this duplication is unique to Atlantic salmon, since no clear evidence of duplicated *elovl5* genes in other salmonids exist in the current sequence databases. However, compared to all other salmonids which have c. 100 chromosome arms, Atlantic salmon is unique in possessing only 72-74 chromosome arms, believed to be the result of species-specific tandem fusions and other rearrangements (ASalBase, 2013). Furthermore, linkage maps show (Lien *et al*., 2011) that salmon chromosome ssa13 has homeologous regions on at least three other salmon chromosomes, and shares syntenic blocks with at least four separate chromosomes from the diploid stickleback (*Gasterosteus aculeatus*). WGD constitutes the most likely origin; however, given the complexity of the Atlantic salmon genome, rejecting a segmental duplication origin would be premature.

Phylogenetic analyses confirmed the basal nature of Northern pike within the protacanthopterygians. Analysis of the rates of nucleotide substitution in the Atlantic

salmon Elovl5 paralogous genes indicated that they are under an evolutionary regime of purifying selective pressure ($\omega < 1$), and are currently evolving at comparable evolutionary rates at the protein level, thus supporting the idea that both duplicates are physiologically required, and have the same biochemical activity as the pike Elovl5. In a larger phylogenetic study, Leong *et al*. (2010) performed pairwise dN/dS analyses on 408 sets of duplicated salmon genes using a preduplicated set of Northern pike orthologs as outgroup, and similarly concluded that salmon paralogs were predominantly exposed to purifying selection, although some loci may be showing a relaxation of selective pressure suggesting that evolution was acting asymmetrically on some paralogs due to reduced constraints. A closer inspection of the dN/dS along the coding sequences of vertebrate *elovl5* suggested the accommodation of a major catalytic site where stronger functional constraints seemed to have acted against the retention of nonsynonymous mutations that would compromise the performance of the enzyme, or impair its activity (Fernandes *et al*., 2010; Newcomer *et al*., 1984).

Here we also tested for functional similarity of pike Elovl5 to salmon paralogs by heterologous expression of the pike enzyme in yeast. Supporting the phylogenetic results, the activity of the pike Elovl5 was indistinguishable from previous assays of the salmon paralogs (Morais *et al*., 2009). Pike Elovl5 was able to lengthen PUFA substrates with chain lengths from C18 to C22. The 18:4n-3, 18:3n-6 and C20 specificity of pike Elovl5 and both salmon Elovl5 paralogs was very similar, and similar to that in other vertebrates (Agaba *et al*., 2005). Only very low, residual activity for the production of 24:5n-3 was detected in yeast transformed with the pike elongase when incubated with 22:5n-3, as previously observed in other

species, including salmonids (Agaba *et al*., 2005; Inagaki *et al*., 2002; Morais *et al*., 2009). As 24:5n-3 is an important intermediate in most vertebrates for the biosynthesis of DHA, and in salmonids C22 to C24 activity has been demonstrated in Elovl2 and Elovl4 (Carmona-Antoñanzas *et al*., 2011; Morais *et al*., 2009), it would be interesting to look for and study these elongase genes in pike.

Until recently the production of LC-PUFA from C18 PUFA was thought to proceed via an alternating desaturation/elongation cycle with the initial step being Δ6 desaturation. However, recent evidence suggests that an alternative pathway is also possible, with the initial step being C18 to C20 elongation, based upon the ability of Δ6 desaturases to also catalyse Δ8 desaturation. For such an alternative loop to exist, elongase activity on LA and ALA is required, and this has been demonstrated in other species (Gregory *et al*., 2010; Inagaki *et al*., 2002; Tu *et al*., 2012). According to the results presented here, is also shown by pike Elovl5. Although not yet demonstrated in salmon, we would expect that, given the Δ8 activity of salmon Δ6 desaturases (Monroig *et al*., 2011), salmon Elovl5 enzymes would also possess this activity.

Although the phylogenetic and functional analyses point towards the maintenance of ancestral activities, the expression profiles indicate functional partitioning in the salmon *elovl5* paralogs. Salmon *elovl5a* and *elovl5b* have different tissue expression profiles, with *elovl5a* being expressed at higher level in intestine and *elovl5b* in liver (Morais *et al*., 2009). In addition, the nutritional regulation of mRNA transcription of these genes differs in tissues from fish fed diets containing low levels of LC-PUFA. The tissue distribution of pike *elovl5* transcripts showed that the highest expression across the tissues tested was in brain, whereas most other

tissues, including liver and intestine, showed very low expression in comparison. In contrast, salmon *elovl5* transcripts were predominantly expressed in liver and intestine, with much lower expression in brain, and even less in other tissues (Morais *et al*., 2009). The pattern of pike *elovl5* tissue expression closely resembles the pattern of LC-PUFA biosynthetic gene expression in carnivorous marine fish, where expression and activity is low in liver and intestine, but high in the brain (Tocher *et al*., 2006). Brain, as with all neural tissues, has a high LC-PUFA level (Lauritzen *et al*., 2001; Tocher, 2003) but, as yet, mechanisms for the biosynthesis, or the transport of fatty acids to the brain are not fully understood (Hamilton and Brunaldi, 2007). Mammalian brain is only capable of biosynthesising a restricted set of fatty acids (Moore, 2001) and it is clear that fatty acid uptake in the brain is different to uptake in most other tissues, probably due the requirement for passage across the blood-brain barrier. Radiolabeled PUFA injected intraperitoneally in Northern pike could be detected in high concentrations throughout the body, with the exception of brain, which consistently contained the lowest amounts of injected PUFA (Henderson *et al*., 1995). Currently, most studies support an energy-dependant mechanism that facilitates and regulates fatty acid uptake influenced by chain length and degree of unsaturation (Mitchell *et al*., 2011; Schwenk *et al*., 2010). The gene expression results from pike and other carnivorous fish suggest that brain may have endogenous biosynthetic machinery for LC-PUFA to supply the high requirements of this tissue, whereas low expression in liver and intestine indicate a reduced requirement for LC-PUFA, or at least DHA, in these tissues. Although low levels of LC-PUFA biosynthesis have been detected in isolated pike hepatocytes (Buzzi *et al*., 1997), experiments *in vivo* failed to find evidence of significant conversion in liver

(Henderson *et al*., 1995). In salmon, although *elovl5* and other LC-PUFA biosynthetic genes are expressed in brain, the highest expression levels are in liver and intestine, the main tissues responsible for dietary fatty acid uptake, biosynthesis and distribution (Morais *et al*., 2009; Tocher, 2003). Salmonid liver has a comparatively high capacity for biosynthesising LC-PUFA such as EPA and DHA (Zheng *et al*., 2004), whereas marine carnivorous fish, such as sea bass (*Dicentrarchus labrax*) have negligible capacity depending on dietary LC-PUFA (Mourente and Bell, 2006).

This above discussion should be qualified by noting that studies on various salmonids and other freshwater species fed with artificial diets with varying LC-PUFA contents and compositions have shown that hepatic desaturase and elongase enzymes exhibited higher expression when the amount of dietary LC-PUFA decreased and LA and/or ALA increased (Morais *et al*., 2009; Tocher *et al*., 2001, 2002; Zheng *et al*., 2004). Although it is possible that LC-PUFA biosynthetic enzymes in liver are also under nutritional regulation in pike, the fish in the present study were obtained from wild stock and thus essentially piscivorous and so would be equivalent to salmon fed diets containing fish oil (i.e. high in LC-PUFA) (Morais *et al*., 2009).

2.4.1 Conclusions

Northern pike possess an active *elovl5* gene, which is homologous to the duplicated salmon *elovl5a* and *elovl5b* genes. There is no evidence of positive (diversifying) selection acting on the salmon genes, or on the pike gene. On the contrary there is evidence of purifying selection maintaining the activity of all of

these genes under functional constraints at the protein level despite some evidence of asymmetrical evolution acting at the DNA level of the salmon *elovl5* paralogs. Supporting this, the enzymatic activities of the pike and the two salmon enzymes are indistinguishable. However, a sharply contrasting expression profile was noted between pike *elovl5* and salmon *elovl5* paralogs. Pike *elovl5* expression was restricted to brain, in common with other carnivorous fish, whereas salmon have highest expression in liver and intestine. This may be the result of adaptations; salmon to a diet in freshwater relatively poor in LC-PUFA, while pike have a highly piscivorous diet containing higher levels of LC-PUFA. Given the complexity of the Atlantic salmon genome, genetic linkage analyses were inconclusive in elucidating the origin of the two salmon *elovl5* loci. Whereas a WGD event constitutes the most likely origin, rejecting a segmental duplication origin would be premature.

Subsequent chapters deal with mechanisms of gene expression and transcriptional regulation to investigate possible asymmetrical diversification events on the *cis*-regulatory regions of *elovl5* duplicated genes.

Chapter 3

Conservation of transcriptional regulatory networks of lipid homeostasis in fish, Atlantic salmon, and mammals

Greta Carmona-Antoñanzas [§], Douglas R. Tocher, Laura Martinez-Rubio and Michael J. Leaver

School of Natural Sciences, University of Stirling, Institute of Aquaculture, Stirling FK9 4LA, Scotland, UK

Authors' contributions

GCA, DRT and MJL planned and coordinated the research; GCA performed laboratory analyses and data analysis; LMR conducted the relative RT-PCR analysis; GCA wrote the first draft of the manuscript, followed by contributions from remaining authors.

Carmona-Antoñanzas *et al*. ***Gene*** 2013, **534**: 1–9.

Abstract

Lipid content and composition in aquafeeds have changed rapidly as a result of the recent drive to replace ecologically limited marine ingredients, fishmeal and fish oil (FO). Terrestrial plant products are the most economic and sustainable alternative; however, plant meals and oils are devoid of physiologically important cholesterol and long-chain polyunsaturated fatty acids (LC-PUFA), eicosapentaenoic (EPA), docosahexaenoic (DHA) and arachidonic (ARA) acids. Although replacement of dietary FO with vegetable oil (VO) has little effect on growth in Atlantic salmon (*Salmo salar*), several studies have shown major effects on the activity and expression of genes involved in lipid metabolism. In vertebrates, sterols and LC-PUFA play crucial roles in lipid metabolism by direct interaction with lipid-sensing transcription factors (TF) and consequent regulation of target genes. The primary aim of the present study was to elucidate the role of key TFs in the transcriptional regulation of lipid metabolism in fish, particularly in the regulation of LC-PUFA biosynthetic genes, by transfection and overexpression of TFs. The results show that the expression of genes of LC-PUFA biosynthesis (*elovl* and *fads2*) and cholesterol metabolism (*abca1*) are regulated by Lxrα and Srebp TFs in salmon, indicating highly conserved regulatory mechanisms across vertebrates. Interestingly, salmon *elovl5a* and *elovl5b* gene promoters responded differently to sterol regulatory element binding protein (Srebp) and liver X receptor (Lxr/Rxr) TFs, possibly due to divergence in their *cis*-regulatory regions. In addition, *srebp1* and *srebp2* mRNA respond to replacement of dietary FO with VO. Thus, Atlantic salmon adjust lipid metabolism in response to dietary lipid composition through the transcriptional regulation of gene expression. It may be possible to further increase efficient and

effective use of sustainable alternatives to marine products in aquaculture by considering these important molecular interactions when formulating diets.

3.1 INTRODUCTION

Lipid content and composition in feeds for farmed fish have experienced a recent and rapid change, because, in order to sustain growth of the aquaculture industry, ecologically limited marine fish meal and fish oil (FO) ingredients have been replaced by terrestrial plant-derived meals and oils. Although replacement of up to 100% dietary FO with vegetable oil (VO) has little effect on growth in Atlantic salmon (*Salmo salar*) (Torstensen *et al*., 2005), some studies have shown major effects on the expression and regulation of genes involved in fatty acid and cholesterol metabolism (Leaver *et al*., 2008b; Morais *et al*., 2009). This impact was consistent with the major compositional changes caused by feeding VO to fish, including decreased levels of dietary cholesterol and long-chain polyunsaturated fatty acids (LC-PUFA), eicosapentaenoate (EPA; 20:5n-3), docosahexaenoate (DHA; 22:6n-3) and arachidonate (ARA, 20:4n-6), which are absent in terrestrial plants (Leaver *et al*., 2008b). Cholesterol and LC-PUFA are critical functional components of cellular membranes and are important precursors of bioactive lipids required for homeostasis, cell signaling, immune and inflammatory responses (Simopoulos, 2008), and the long-term health and welfare effects of reductions in these essential dietary nutrients in fish, including salmon, are unknown.

Much effort has been directed toward the understanding of effects of dietary imbalances in LC-PUFA and cholesterol in humans and mammalian models because

of the links between dyslipidemia and a range of highly prevalent cardiovascular, metabolic and inflammatory diseases (Wilson *et al*., 2005). These studies have shown that cholesterol and fatty acids and their metabolic derivatives can exert major effects on physiology by interactions with a range of transcription factors (Desvergne *et al*., 2006). Particular attention has focused on liver X receptor (LXR), peroxisome proliferator-activated receptors (PPAR) and sterol regulatory element binding proteins (SREBP) and their activities in liver and monocyte cells. LXR and PPARs are members of the nuclear receptor (NR) family of TFs that form obligate heterodimers with the retinoid X receptor (RXR), subsequently recognising specific DNA response elements (REs) within the regulatory regions of their target genes (Willy and Mangelsdorf, 1997; Yoshikawa *et al*., 2002). LXR has a pivotal role in the control of intermediary metabolism, mediating cross-regulation between fatty acid and sterol metabolism (Jakobsson *et al*., 2012). LXR activity is activated by binding oxysterol ligands, catabolic products of cholesterol (Reschly *et al*., 2008). In response to cholesterol overloading, and consequent oxysterol production, LXR modulates intracellular cholesterol levels by transactivating the expression of cholesterol ester transfer protein, apolipoproteins, cholesterol 7alpha-hydroxylase (CYP7α1) and the ATP-binding cassette transporter 1 (ABCA1), which regulate cholesterol efflux from cells (Jakobsson *et al*., 2012).

In response to cholesterol depletion, SREBPs, a family of membrane-bound transcription factors, are activated. SREBP1 plays a crucial role in the regulation of many lipogenic genes and SREBP2 primarily regulates the transcription of cholesterogenic enzymes (Jeon and Osborne, 2012). Interactions between these pathways are to some extent mediated through LXR activating SREBP1

transcription, inducing the expression of enzymes involved in the synthesis of fatty acids, triacylglycerols and phospholipids. In addition, some important lipid metabolising genes, such as fatty acid synthase (FAS), are both LXR and SREBP1 targets (Joseph *et al*., 2002). PPARs, encoded by three genes in mammals, are activated by binding fatty acids or their oxidised derivatives and act to regulate expression of genes of lipid degradation and biosynthesis. PPARα and PPARβ, regulate the expression of genes encoding mitochondrial and peroxisomal fatty acid-catabolising enzymes, whilst PPARγ has a central role in fat storage by promoting and maintaining the adipocyte phenotype (Desvergne *et al*., 2006). Thus, LXR, SREBP and PPAR transcription factors act as lipid sensors that translate changes in cellular sterol and fatty acid content and composition into metabolic activity.

Compared to mammals, few studies have addressed the existence or roles of these transcriptional regulators of lipid metabolism in fish. Our contention is that a greater understanding of lipid-mediated gene regulatory networks in Atlantic salmon will facilitate the efficient, effective and safe use of sustainable alternatives to marine products in aquaculture feeds. Recently the genes for Atlantic salmon Lxrα, Srebp1 and Srebp2 and Ppars have been characterised (Cruz-Garcia *et al*., 2009; Leaver *et al*., 2007; Minghetti *et al*., 2011). In addition, studies on an Atlantic salmon cell line (SHK-1) have shown that several lipid metabolic genes are transcriptionally regulated in response to changes in lipid composition of the medium (Minghetti *et al*., 2011). Thus, the primary aim of the present study was to elucidate Lxrα, Srebp and Ppar gene regulatory mechanisms and key lipid metabolic target genes in Atlantic salmon with particular interest in identifying distinct patterns of transcriptional regulation between duplicated *elovl5* encoding LC-PUFA biosynthetic

enzymes. Also, we intended to determine the extent to which dietary modulation of lipid and fatty acid metabolism in salmon reflects or varies from the patterns of gene regulation described for mammals.

3.2 MATERIALS AND METHODS

3.2.1 Cell lines and cell culture

The established Atlantic salmon cell line derived from head kidney (SHK-1) was grown at 22°C in an atmosphere of 4% carbon dioxide in Dulbecco's modified eagle medium (DMEM) containing 3 g L^{-1} D-glucose and 55 mg L^{-1} sodium pyruvate, and supplemented with 10% foetal bovine serum (FBS), 50 U mL^{-1} penicillin, 50 μg mL^{-1} streptomycin, 40 μM 2-mercaptoethanol and 4 mM L-glutamine. For gene promoter transactivation assays, fathead minnow (*Pimephales promelas*; FHM) epithelial cells were maintained at 22°C in Leibovitz's L-15 with GlutaMAX^{TM}-1 medium containing 900 mg L^{-1} D+galactose and 550 mg L^{-1} sodium pyruvate and 10% FBS. All media and supplements were obtained from Life Technologies (Glasgow, UK).

For subculturing, the cell monolayer was washed twice with phosphate buffer saline (PBS) without $CaCl_2$ or $MgCl_2$ (Invitrogen, UK), cells detached by incubation with 0.05% trypsin/0.02% EDTA and re-suspended in medium. Viable cells were counted after harvesting using a Neubauer haemocytometer, 0.4% Trypan blue (Sigma, Dorset, UK) and an inverted microscope (IMT-2, Olympus).

For transcription factor (TF) ligand treatments, SHK-1 cells were seeded in 6-well clear plates (Nunc, Denmark) at a density of 4×10^5 cells per well in a volume of 3 ml Leibovitz's L-15 medium. Cells were approximately 70% confluent after 48 h growth, when medium was aspirated, cells washed twice with PBS and fresh medium containing the various treatments as ethanol solutions were added. Final concentrations were cholesterol (20 μM), WY14643 (25 μM), 2-bromopalmitate (25 μM), or LXR agonists (GW3965 and T0901317, 10 μM) or with ethanol carrier alone (100% ethanol, EtOH). After 24 h, the medium was aspirated, the cell monolayer washed twice (PBS) and cells scraped from each well in 0.5 ml of PBS. Cells were centrifuged for 5 min at 3000 x g, PBS discarded and replaced by 0.5 ml of TriReagent (Ambion, UK), followed by vigorous mixing to lyse and digest cells. Cells from two wells were pooled to produce three replicates per treatment.

3.2.2 Fish, diets and sampling protocols

Four diets (4 mm pellets) with the same basal protein composition, but coated with four different oils were formulated at Skretting Technology Centre (Stavanger, Norway) to satisfy the nutritional requirements of salmonid fish (NRC, 2011). The oils used were FO (anchovy oil), or 100% replacement with rapeseed oil (RO), linseed oil (LO) or soybean oil (SO). Atlantic salmon post-smolts (130 g) were randomly distributed into 16 tanks at the Skretting Aquaculture Research Centre (Stavanger, Norway). After a conditioning period of 3 weeks during which the fish received a commercial diet containing FO, the fish were fed the experimental diets to satiation for a period of 16 weeks. Full descriptions of the diet compositions and experimental conditions were reported previously (Leaver *et al*., 2008b) (TABLE 3.1). At the end of the trial fish were anaesthetised with metacain (50 mg/L) and

pyloric caeca (intestine), a major organ involved in uptake and transport of lipids, were dissected from five randomly selected fish from each dietary treatment. Samples of 0.5 g of caeca were immediately and rapidly disrupted in 5 ml of TriReagent (Ambion, UK) using an Ultra-Turrax homogeniser (Fisher Scientific, UK), and stored at – 80ºC prior to RNA extraction. The dietary trial and all procedures on Atlantic salmon conformed to European ethical regulations regarding the care and use of farmed animals in research.

TABLE 3.1 Proximate composition (percentage of total weight) and fatty acid compositions (percentage of total FA weight) of experimental diets (Adapted from Leaver *et al*., 2008b).

	FO	LO	RO	SO
Total Protein	46.9	46	48	47
Total Fat	33.2	33.2	34.3	34.1
Ash	6.7	7.6	7	7.8
Moisture	6.2	5.8	5.7	4.5
14:0	7.4	0.9	1.2	0.9
16:0	18.7	7.4	7	12.6
18:0	3.9	3.5	1.9	3.1
Total saturated [a]	30.9	11.6	10.1	16.6
16:1n-7	7.5	0.9	1.3	0.9
18:1n-9	8.8	17.4	47.2	20.8
18:1n-7	2.7	1	2.7	1.7
24:1n-9	0.5	0.2	0.3	0.2
Total monoenes	26.3	23	56.8	27.3
18:2n-6	3.5	14.2	17.8	44.4
20:4n-6	0.9	0.1	0.1	0.1
Total n-6 PUFA [b]	4.7	14.4	18	44.6
18:3n-3	1.2	44.9	8.1	5.7
18:4n-3	3.4	0.7	0.8	0.6
20:4n-3	0.8	0.1	0.1	0.1
20:5n-3	14.7	2	2.5	2
22:5n-3	1.7	0.2	0.2	0.2
22:6n-3	16.3	2.9	3.4	2.9
Total n-3 PUFA	38.1	50.8	15.1	11.5
n-6 / n-3	0.12	0.28	1.19	3.89
Total PUFA	42.8	65.2	33.1	56.1

Data are means of duplicate analyses.

[a] contains 15:0 and 17:0, present in some samples at up to 0.5%

[b] totals include 18:3n-6, 20:2n-6, 22:4n-6 and 22:5n-6 present in some samples up to 0.5%.

PUFA, polyunsaturated fatty acids; FO, fish oil; LO, linseed oil; RO, rapeseed oil; SO, soy bean oil.

3.2.3 Plaice retinoic X receptor (Rxr)

A ready-to-use plaice (*Pleuronectes platessa*) retinoic X receptor inserted into a constitutively expressed vector was facilitated by a member of the Institute of Aquaculture (University of Stirling). Prior co-transfection with Lxrα, sequence conservation and functional isoform were determined.

3.2.3.1 Phylogenetic analysis of plaice Rxr

All database sequence details of the OTUs used in the multiple sequence alignment and the construction of the phylogenetic tree are included in FIGURE 3.1.

3.2.3.1.1 Multiple sequence alignment

European plaice *rxrγ* ORF was used to perform a local blast search in the NCBI database to detect homologous RXR sequences. The aa sequence from plaice *rxrγ* cDNA and homologous proteins from various teleost species including Atlantic salmon Rxrβ [GenBank: BT043993], Japanese medaka Rxrα (*Oryzias latipes*) [GenBank: XM_004068342], False kelpfish Rxrγ (*Sebastiscus marmoratus*) [GenBank: EU888277], Nile tilapia Rxrγ (*Oreochromis niloticus*) [GenBank: XM_005479079], and *Homo sapiens* RXRα [GenBank: HQ692843] were aligned using ClustalW (BioEdit 7.1.3) as described in *Section 2.2.3.1*. The sequence similarity between plaice Rxrγ and other Rxr orthologs, was calculated by performing pairwise sequence comparisons using the online EMBOSS pairwise alignment server (EMBOSS, 2013). The EMBOSS tool applies the Needleman-Wunsch algorithm to compute identity and similarity scores accounting for the biophysical properties of the amino acid substitutions. We applied the BLOSUM80

substitution scoring matrix suitable for less divergent alignments and used the default gap penalty.

3.2.3.1.2 Phylogenetic tree

A phylogenetic tree was constructed using the aa (open reading frame, ORF) sequences of multiple fish Rxrα, Rrxβ and Rrxγ orthologous, and human RXR members. The Neighbour-Joining method was applied using MEGA4 (Saitou and Nei, 1987; Tamura *et al*., 2007), as described in *Section 2.2.3.3*, with adjustments to create an unrooted phylogenetic cladogram using protein-coding sequences. For this, the JTT substitution matrix (Jones *et al*., 1992) that integrates observed probabilities of amino acid substitutions obtained from local alignments of large protein databases was utilised, and uniform evolutionary rates among lineages assumed.

3.2.4 Ligand-activation of fish nuclear receptors

3.2.4.1 Construction of an Atlantic salmon Gal4-Lxrα LBD chimera

The ligand binding domain (LBD) (amino acid residues 191-462) [GenBank: FJ470290] of Lxrα was amplified by PCR from salmon pyloric caeca cDNA using primers (TABLE 3.2) with 5' restriction sites to allow in-frame subcloning between the *Bam*HI and *Xba*I sites of the pBIND vector (Promega, Southampton, UK), which contains the yeast Gal4 DNA-binding domain (Reschly *et al*., 2008). Positive clones were selected by enzymatic digestion and sequenced (CEQ-8800 Beckman Coulter Inc., USA) prior to bulk purification of the plasmid construct using gravity-flow anion-exchange purification (QIAfilter plasmid midi kit, Qiagen). The Gal4-LxrαLBD plasmid construct was co-transfected with a reporter gene plasmid in which the *Firefly* luciferase gene is under the control of a promoter containing UAS

(upstream activation sequence) (pGL4.31, Promega), which is recognised by Gal4. To control differences in transfection efficiency, a constitutively expressed control reporter construct encoding for *Renilla* luciferase (pGL4.75, Promega) was included.

3.2.4.2 Construction of a European plaice Gal4-Rxrγ LBD chimera

The original European plaice (*Pleuronectes platessa*) transcription factor Rxrγ was cloned from plaice liver cDNA using primers derived from a genomic clone isolated from a plaice genomic library in lambda FIX II (TABLE 3.2). Previously, the library had been screened with a DNA probe complementary to the DNA binding domain of a mouse *Pparα* cDNA, which recognised other similar receptors including the plaice *rxrγ* gene. Using the identified plaice genomic sequence, primers designed to amplify the entire coding sequence of plaice Rxrγ, which included suitable restriction sites, *Eco*RI and *Xho*I (TABLE 3.2), were applied to plaice liver cDNA. The resulting Rxrγ cDNA was cloned into pcDNA3. The Rxrγ LBD (amino acid residues 183-438) was amplified by PCR from pcDNA3-plaiceRxrγ construct (*See Section 3.2.6.1*) using primer ppRXRFbind and pcDNA3.1RP oligo (TABLE 3.2) to allow in-frame subcloning between the *Sal*I and *Xba*I sites of the pBIND vector (Promega, Southampton, UK), which contains the yeast Gal4 DNA-binding domain (Reschly *et al*., 2008). After insert confirmation and ultrapurification, the resulting Gal4-RxrγLBD chimeras were co-transfected with pGL4.31 and pGL4.75 as described in detail in *Section 3.2.3.4*.

3.2.4.3 Preparation of albumin-bound fatty acids

Stock solutions of ~ 2mM of saturated palmitic acid (PA), and unsaturated n-3 ALA, EPA and DHA and n-6 LA and ARA fatty acids were prepared by diluting

TABLE 3.2 Details of primer pairs (restriction sites underlined) used for cloning and qRT-PCR analyses.

Aim	Transcript	Primer sequence	Amplicon size	Ta	Accession no.
Lxrα LBD cloning	*lxrα*	5'- CGGGGATCCTGGAGCAATGCGTTCTCT -3' 5'- TTTCTAGAGGGCATCAGCGGAGTCACT -3'	879 bp	62 °C	FJ470290 [a]
Rxrγ LBD cloning	*rxrγ*	5'- AGTCGACCAGCAGTTTCAACGAGGA -3' 5'- CAAACAACAGATGGCTGGC -3'	764 bp	60 °C	*unpublished*
ORF cloning	*lxrα*	5'- TTTGAATTCGTAATGACAAGCTGC -3' 5'- TTCCTCGAGGGCATCAGCGGAGTC -3'	1,445 bp	55 °C	FJ470290 [a]
	rxrγ	5'- TAGAATTCGATCCAATGAGTGCGGCCCA -3' 5'- TACTCGAGGTCGACGGTATCGATAAGCTTGA -3'	1,372 bp	62 °C	*unpublished*
	srebp1	5'- AACGAATTCGCAAAAATGAACTTGTCTTTTGACG -3' 5'- ATGCTCGAGGTCAGGGTAGCTCGCTCTTTA -3'	1,422 bp	60 °C	TC148424 [b]
	srebp2	5'- TTCTGAATTCCATTATGGACAGTAACG -3' 5'- GAGGCTCGAGGTCACATCATCAGCATC -3'	1,165 bp	57 °C	TC166313 [b]
	pparα	5'- TTTGAATTCCGACATGGCGAGCCACT -3' 5'- TTCCTCGAGTATCTCAGTACATGTCCCT -3'	1,421 bp	60 °C	AM230809 [a]
	pparβ1a	5'- TTTGAATTCGGACATGAAGTCTCATGACG -3' 5'- TTCCTCGAGCCAGTCAGTACATGTCC -3'	1,358 bp	60 °C	AJ416953 [a]
Promoter cloning	*elovl5a*	5'- AATGAGCTCAGCTCTGCAAAGCCATGTG -3' 5'- AATCTCGAGTTCTGACCTAAATAGACAGATGAAAG -3'	4,910 bp	63 °C	GU238431 [a]
	elovl5b	5'- TATGAGCTCTGCTCTGTGTAGCCATG -3' 5'- CTCCATGGCTAACCTGAATAGACAGATG -3'	3,150 bp	62 °C	GU324549 [a]
	fads2d6a	5'- GTTGGAGCTCGGAGCATAAGAATATC -3' 5'- AATCCATGGCCTCGGTTCTCTCTGCTCCACTCACAC -3'	4238 bp	60 °C	AY736067 [a]
qRT-PCR	*abca1*	5'- GGACGAACCCTGTGTCTGTT -3' 5'- ATTTGCATTGCGTTTCAGTG -3'	203 bp	60 °C	EG836783 [a]
	fads2d5	5'- GTGAATGGGGATCCATAGCA -3' 5'- AAACGAACGGACAACCAGA -3'	192 bp	56 °C	AF478472 [a]
	fads2d6a	5'- CCCCAGACGTTTGTGTCAG -3' 5'- CCTGGATTGTTGCTTTGGAT -3'	180 bp	60 °C	AY458652 [a]
	fads2d6b	5'- ATAGAGGGTTTATATAGTAGGGCC -3'	204 bp	58 °C	GU207400 [a]

	5’- GGTGGGACGCTAGAAGTTAA -3’			
elovl2	5’- CGGGTACAAAATGTGCTGGT -3’	145 bp	60 °C	TC91192 [b]
	5’- TCTGTTTGCCGATAGCCATT -3’			
elovl4	5'- TTGTCAAATTGGTCCTGTGC -3'	191 bp	61 °C	HM208347 [a]
	5'- TTAAAAGCCCTTTGGGATGA -3'			
elovl5a	5’- ACAAGACAGGAATCTCTTTCAGATTAA -3’	137 bp	60 °C	AY170327 [a]
	5’- TCTGGGGTTACTGTGCTATAGTGTAC -3’			
elovl5b	5’- ACAAAAAGCCATGTTTATCTGAAAGA -3’	141 bp	60 °C	DW546112 [a]
	5’- AAGTGGGTCTCTGGGGCTGTG -3’			
fas	5'- ACCGCCAAGCTCAGTGTGC -3'	155 bp	60 °C	DW551395 [a]
	5'- CAGGCCCCAAAGGAGTAGC -3'			
lxrα	5'- GCCGCCGCTATCTGAAATCTG -3'	210 bp	58 °C	FJ470290 [a]
	5'- CAATCCGGCAACCAATCTGTAGG -3'			
pparα	5'- TCCTGGTGGCCTACGGATC -3’	111 bp	60 °C	DQ294237 [a]
	5’- CGTTGAATTTCATGGCGAACT -3’			
pparβ	5’- GAGACGGTCAGGGAGCTCAC -3’	151 bp	60 °C	AJ416953 [a]
	5’- CCAGCAACCCGTCCTTGTT -3’			
pparγ	5’- CATTGTCAGCCTGTCCAGAC -3’	144 bp	60 °C	AJ416951 [a]
	5’- TTGCAGCCCTCACAGACATG -3’			
srebp1	5'- GCCATGCGCAGGTTGTTTCTTCA -3'	151 bp	63 °C	TC148424 [b]
	5'- TCTGGCCAGGACGCATCTCACACT -3'			
srebp2	5'- TCGCGGCCTCCTGATGATT -3'	180 bp	63 °C	TC166313 [b]
	5'- AGGGCTAGGTGACTGTTCTGG -3'			
cyp7α1	5'- TCCTCAACACCCTGGAGAAC -3'	125 bp	63 °C	BT059202 [a]
	5'- CAGCATGGTCTTAGCCAGGT -3'			
acox	5'- AAAGCCTTCACCACATGGAC -3'	230 bp	60 °C	TC145297 [b]
	5'- TAGGACACGATGCCACTCAG -3'			
hmgCoAR	5'- CCTTCAGCCATGAACTGGAT -3'	224 bp	60 °C	BT058953 [a]
	5'- TCCTGTCCACAGGCAATGTA -3'			
ß-actin	5’- ACATCAAGGAGAAGCTGTGC -3’	141 bp	56 °C	AF012125 [a]
	5’- GACAACGGAACCTCTCGTTA -3’			
elf-1α	5’- CTGCCCCTCCAGGACGTTTACAA -3’	176 bp	59 °C	AF321836 [a]
	5’- CACCGGGCATAGCCGATTCC -3’			

[a] *GenBank (http://www.ncbi.nim.nih.gov/).*

the lyophilised free fatty acid in fatty acid free foetal bovine albumin (FAF FBA). Each solution was filter-sterilised and stored in multiple aliquots at – 20ºC protected from light in tubes evacuated under nitrogen gas. To test the antagonistic effect of fatty acids in ligand-activated Lxrα, final concentration of fatty acids at 50 μM (FAF FBA for the control) were used reflecting sub-toxic concentrations determined in previous MTT cytotoxicity assays (Hansen *et al*., 1989) performed in Chinook cell lines (*unpublished data*).

3.2.4.4 Co-transfection assays

Functional characterisation of Lxrα and Rxrγ LBDs was determined by a luciferase-based functional assay using the FHM cell line as described previously (Colliar *et al*., 2011). Twenty-four hours prior to transfection, 4×10^4 cells per well were seeded in a 96-well black-sided, clear-bottom microtitre plate (Corning, NY, USA). Transfection mixtures contained, per well; 100 ng Gal4-LBD construct (pBIND), 60 ng of *luc2*P/GAL4UAS reporter plasmid (pGL4.31), 20 ng of *hRluc*/CMV internal control reporter plasmid (pGL4.75) and 1.5 μl of Polyfect transfection reagent (Qiagen) in 100 μl of L-15 medium. Within each experiment, treatments were performed in triplicate. Experimental controls included treatments in which the Gal4-LBD construct was replaced by empty pBIND (Gal4) vectors during transfection, as well as cells transfected with the appropriate Gal4-LxrαLBD construct and reporters and treated with ethanol, or FAF FBS carrier only.

3.2.4.4.1 Lxrα activation: oxysterols and synthetic agonists

After 24 h, transfection mixes were removed and replaced with media containing treatment vehicle (ethanol), or serial dilutions of cholesterol, or one of the

following LXR agonists: natural oxysterols 20(S)-hydroxycholesterol (20S-OH) and 22(R)-hydroxycholesterol (22R-OH), synthetic agonists T0901317 and GW3965, and the fungal molecule paxilline (Sigma Aldrich, UK). Compounds of interest were diluted into L-15 from ethanolic stock and 100 µl of the dilution was added per well and incubated for 24 h.

3.2.4.4.2 Lxrα inactivation: free fatty acids and synthetic agonist

To test the antagonistic effect of fatty acids on the salmon Lxrα activation, after 24 h, transfection mixes were removed and replaced with 100 µl of L-15 media containing LXR agonist 22R-OH or GW3965 at a final concentration of 20 µM, and one of the aforementioned fatty acids at the non-toxic final concentrations of 50 µM. Treatment vehicles were ethanol and FAF FBS used at equal volumes for LXR agonists and fatty acids, respectively.

3.2.4.4.3 Plaice Rxrγ activation

As previously indicated, transfection mixtures were removed and media replaced with a variety of potential RXR agonists. Among them were the long-chain fatty acid DHA (50 µM), aquatic contaminants perfluorooctanoic (PFOA, 20 µM) and tributyl tin oxide (TBTO, 50 µM) known as a potent antagonist of fish Ppar (Colliar *et al*., 2011), the LXR agonist GW3965 (10 µM), and acknowledged mammalian RXR agonists 9cis-retinoic acid (9cRA, 10 µM) and synthetic LG268 (10 µM). Chemical concentrations reflected literature values for mammalian studies giving maximal activation in the absence of toxicity. All chemicals were diluted in ethanol, used as the control carrier in luciferase assays.

3.2.4.5 Luciferase assays

Before harvesting, cells were washed with PBS and subsequently lysed by 10 min incubation in 75 µl per well of 1x Passive Lysis Buffer (Promega) to release luciferase proteins into the supernatant. To quantify the luciferase activities of *Firefly* and *Renilla,* activating reagents were freshly prepared using an assay protocol described previously (Colliar *et al*., 2011). First, the Firefly luciferase protein responsible for the characteristic yellow light emission was activated by an alkaline solution containing the Firefly substrate luciferin (stored in aliquots at – 80ºC), which in the presence of ATP and magnesium is oxygenised and emits light at 560 nm (McElroy and Green, 1956). Subsequently, the Renilla reagent is added with the intention of quenching the Firefly activity and activating Renilla luciferase, which catalyses metabolism of coelenterazine producing an emission peak at 480 nm. This was achieved by combining the Renilla buffer formulation of Dyer (Dyer *et al*., 2000) with a balanced mixture containing phosphate and ammonium ions and DTT to inactivate Firefly. Transactivation activities were obtained using VICTOR X Multilabel plate reader (PerkinElmer, USA) and data were normalised to the *Renilla* luciferase activities. Data are presented as means of raw transactivation activities of three independent assays. The significance of effects of treatments were tested by one-way analysis of variance (ANOVA), followed by multiple comparisons using Tukey's *post hoc* test at a significance level of $P \leq 0.05$ (PASWS 18.0, SPSS Inc., USA).

3.2.5 Quantitative RT-PCR (qPCR)

Salmon caecal samples from fish fed different oil sources (n = 5), or SHK-1 cells exposed to LXR- or PPAR-ligands (n = 3) were used for relative and absolute qPCR analyses, respectively.

3.2.5.1 RNA extraction and cDNA synthesis

For gene expression analysis, samples were immediately and rapidly disrupted in TriReagent, and stored at – 80ºC prior to RNA extraction. Total RNA was extracted applying phenol/1-bromo-3-chloropropane (BCP1) extraction following the manufacturer's instructions (TRI Reagent, Sigma, UK), and the quantity and quality of isolated RNA determined by electrophoresis and spectrophotometry (Nanodrop 1000, Thermo Scientific, USA). Two micrograms of total RNA was reverse transcribed into cDNA using the Verso cDNA synthesis kit (Thermo Scientific, UK) and primed with 7.5 μM random hexamers and 2.5 μM oligo(dT) in a 3:1 molar ratio. The resulting cDNA was diluted 20-fold with nuclease-free water.

3.2.5.2 Preparation of qPCR reactions

For quantitative qPCR, oligonucleotide primers for target genes and housekeeping genes (*elf-1α* and *β-actin*) (TABLE 3.2) were used at 0.3 μM with one-fortieth of the cDNA synthesis reaction (5 μl of a 1:20 dilution) and 10 μl of SYBR-green qPCR mix (ABgene, UK) in a total volume of 20 μl. Reactions were run in a Mastercycler RealPlex^2 (Eppendorf, UK). Amplifications were carried out including systematic negative controls containing no cDNA (NTC, no template control) and omitting reverse transcriptase enzyme (-RT) to check for DNA contamination.

Thermal cycle and melting curves were performed as described previously with specific annealing temperatures for each primer pair shown in TABLE 3.2 (Carmona-Antoñanzas *et al*., 2011). The qPCR product sizes were checked by agarose gel electrophoresis and the identity of random samples was confirmed by sequencing.

3.2.5.2.1 Specifics for absolute qPCR

Absolute quantification was achieved by including a parallel set of reactions containing spectrophotometrically-determined standards consisting of serial dilutions of known copy numbers of linearised pCR2.1 TOPO plasmid (Invitrogen) that contained the predicted amplification product for each measured gene. The standard dilutions were prepared in λTE buffer (50 ng/µl λDNA), and the copy numbers calculated based on the construct size, the linearised plasmid concentration and the dilution factor applied.

3.2.5.3 Statistical analyses

Results were expressed as mean normalised ratios (± SE) between the copy number of target genes and the mean copy number of the reference genes (*elf-1α* and *β-actin*). Differences in the expression of target genes among different treatments were determined by one-way analysis of variance (ANOVA), followed by multiple comparisons using Tukey's *post hoc* test (PASWS 18.0, SPSS Inc., USA). In contrast, the effects of diet on TF expression were analysed for statistical significance using the relative expression software tool REST-MCS (Pfaffl *et al*., 2002), and normalised by two housekeeping genes (*elf-1α* and *β-actin*). Gene expression was presented as the relative expression ratio of each gene in fish fed one

of the VOs or FO. A significance of $P \leq 0.05$ was applied to all statistical tests performed.

3.2.6 Expression constructs and cell transfection assays

3.2.6.1 Construction of Atlantic salmon expression vectors

Atlantic salmon transcription factors Lxrα [GenBank: FJ470290], Pparα [GenBank: AM230809], Pparβ1a [GenBank: AJ416953], Srebp1 [GenBank: HM561860] and Srebp2 [GenBank: HM561861] had been previously described (Cruz-Garcia *et al*., 2009; Minghetti *et al*., 2011). The entire open reading frame of Lxrα (Lxrα ORF, 1-462 aa), Pparα (Pparα ORF, 1-464 aa), Pparβ1a (Pparβ ORF, 1-443 aa), and the soluble N-terminal domains of Srebp1 (nSrebp1, 1-476 aa) and Srebp2 (nSrebp2, 1-460 aa) were amplified from caecal cDNA samples using the primers detailed in TABLE 3.2, which included suitable restriction sites, *Eco*RI and *Xho*I, for subsequent insertion into the constitutive expression vector pcDNA3 (Invitrogen). Positive clones were selected by enzymatic digestion and sequenced (CEQ-8800 Beckman Coulter Inc., USA).

3.2.6.3 Construction of Atlantic salmon promoter reporter constructs

The putative promoter regions, first non-coding exon, first intron and the ATG start codon of Atlantic salmon *elovl5a* (-3618, [GenBank: GU238431]), *elovl5b* (-3141, [GenBank: GU324549]), and *fads2d6*a (-1791, [GenBank: AY736067]) were amplified from genomic DNA using high-fidelity PfuTurbo DNA Polymerase (Statagene, UK) and primers containing restriction sites (TABLE 3.2). These fragments were cloned into *Sac*I/*Xho*I or *Nco*I/*Xho*I sites encompassing the luciferase

start codon of a promoterless reporter plasmid [pGL4.10, *luc2*] (Promega, USA), which encodes *Firefly* luciferase.

3.2.6.4 Luciferase assays

For luciferase assays, FHM cells were co-transfected, under the conditions described previously (*Section 3.2.4.4*) with the salmon *elovl5s* or *fads2d6*a reporter constructs, and Lxrα, Pparα, Pparβ1a, nSrebp1, nSrebp2 or plaice Rxrγ expression constructs (pcDNA3). Briefly, stock FHM cells were seeded in 96-well opaque plates at a density of 2×10^4 per well 24 h prior to transfection. Transfection mixtures consisted of 50 ng of pGL4.10 reporter construct (empty pGL4.10 vector in controls), 30 ng of pcDNA3 expression construct (empty pcDNA3 vector in controls), 20 ng of *Renilla* pGL4.75 and 1 µl of Polyfect transfection reagent to 100 µl of L-15 medium. Within each transfection experiment, each treatment was performed in triplicate. After 24 h, the medium was aspirated, monolayer washed twice with PBS, cells lysed, and *Firefly* and *Renilla* luciferase activities quantified (Colliar *et al*., 2011). Transactivation activities were obtained, data normalised and statistical analyses performed as described above (*Section 3.2.3.5*).

3.3 RESULTS

3.3.1. European plaice Rxrγ sequence and phylogenetics

The *P. platessa* Rxrγ cDNA encodes a protein of 438 aa (1,314 bp) that contains the two typical features that characterise nuclear receptors: a DNA binding domain (DBD) and a ligand binding domain (LBD). Plaice Rxrγ DBD (Ile_{113} –

Arg_{190}) and LBD (Thr_{202} – Pro_{434}) were predicted by sequence alignment with the homologous human RXRα, thoroughly studied by X-ray crystallography (Egea *et al*., 2000; Zhao *et al*., 2000). Multiple sequence alignment among plaice Rxrγ and various RXR orthologs indicated that DBD and LBD are highly conserved among distantly related species including mammals (FIGURE 3.2). Plaice DBD shares 86.4% – 97% aa identity to the vertebrate homologues depicted in FIGURE 3.1, LBD was shown to share 82.7% – 97.9% aa identity with predicted mammalian and fish Rxr.

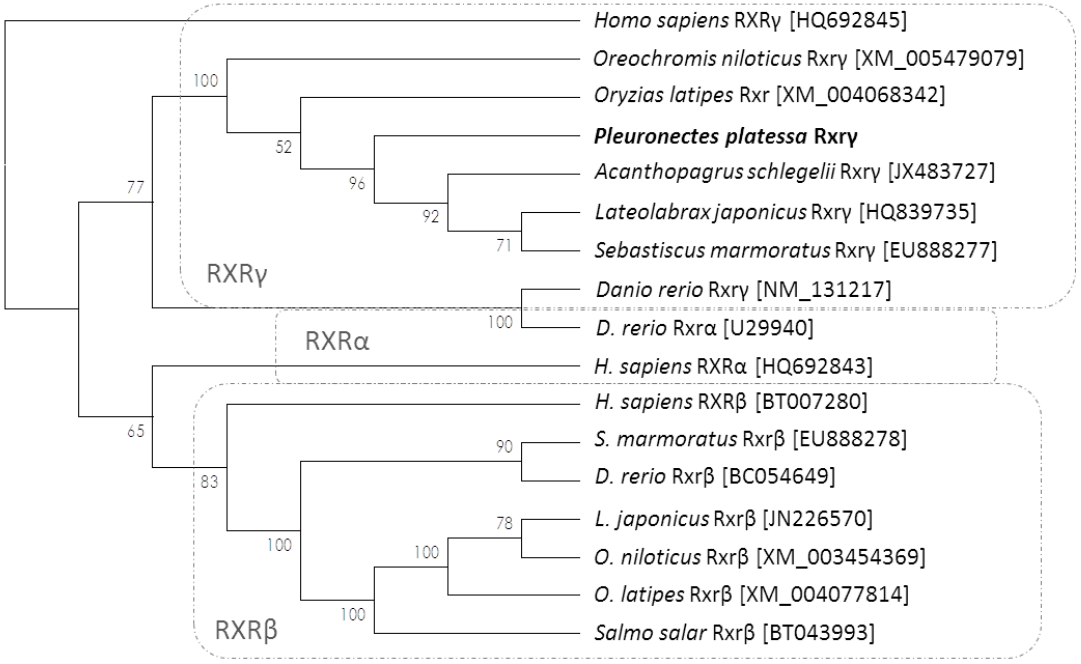

FIGURE 3.1 Evolutionary relationships of 17 vertebrate operational taxonomic units (OTUs). The evolutionary history was inferred using the Neighbour-Joining method. The percentage of replicate trees in which the associated OTUs clustered together in the bootstrap test (10,000 replicates) are shown next to the branches (Saitou and Nei, 1987). The evolutionary distances were computed using the JTT matrix-based method (Jones *et al*., 1992) and are in the units of the number of amino acid substitutions per site. Phylogenetic analyses were conducted in MEGA4 (Tamura *et al*., 2007).

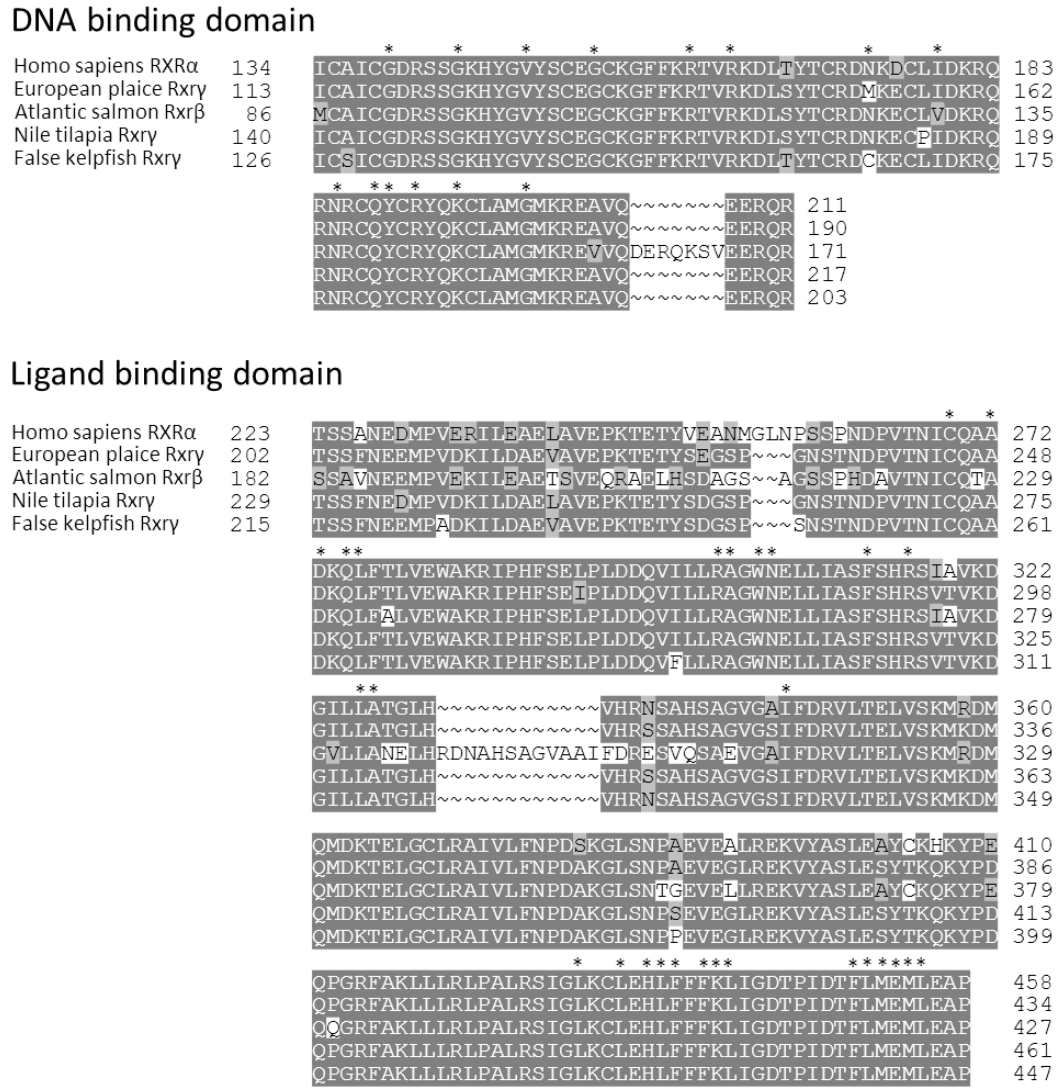

FIGURE 3.2 Multiple sequence alignment of the deduced amino acid (aa) sequence of the Rxrγ from European plaice with that of *H. sapiens* RXRα and homologous Rxr from fish Atlantic salmon (*Salmo salar*), Nile tilapia (*Oreochromis niloticuls*), Japanese medaka (*Oryzias lapites*) and False kelpfish (*Sebasticus marmoratus*). The aa regions corresponding to the DNA binding domain and the ligand binding domain were aligned using ClustalW (BLOSUM62 substitution matrix), and identity/similarity shading was based on a 60% identity threshold. Identical residues are *shaded dark grey* and similar residues are *shaded light grey*. *Indicates the main aa residues involved in the protein-protein interactions, DNA dimerisation and RXR interaction with ligand (Egea *et al*., 2000; Zhao *et al*., 2000); ~, represents a gap in the sequence.

Predictably, the greatest LBD and DBD identity scores corresponded to the pairwise comparisons performed between the plaice putative Rxrγ and another fish (Nile tilapia) Rxrγ, whereas mammalian RXR isoforms exhibited the lowest identity

scores in average. Differentiation among plaice Rxrγ protein and RXR isoforms α, β and γ present in vertebrates is reflected in the phylogenetic analysis (FIGURE 3.1) indicating that plaice Rxrγ clusters more closely with other vertebrate RXRγ members. All fish Rxrβ grouped together with the human homologue, and clustered more distantly from other members of the RXR family such as RXRα and RXRγ.

3.3.2 Ligand specificity of Atlantic salmon Lxrα

3.3.2.1 Atlantic salmon Lxrα activation

Prior to testing for Lxr-dependent gene expression the ligand binding specificity of *S. salar* Lxrα was determined. The salmon Gal4-LxrαLBD chimera was activated by synthetic LXR ligands, including T0901317, GW3965 and paxilline, and also by the physiological oxysterols, 22R-OH and 20S-OH (FIGURE 3.3). The concentration-response curve for activation of Lxrα indicated that T0901317 and GW3965 were the most potent agonists (FIGURE 3.3), followed by natural oxysterol 22R-OH. T0901317 and GW3965 activated salmon Lxrα at micromolar concentrations with maximal effect observed at 10 μM. The synthetic agonist paxilline exhibited the lowest effect, whereas cholesterol had no effect. No response was detected at 100 μM for most agonists, perhaps due to toxic effects.

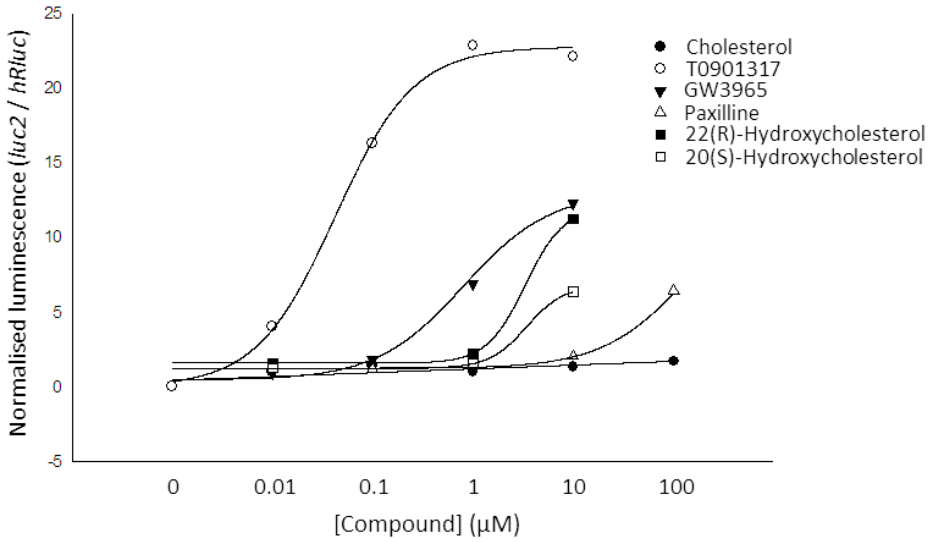

FIGURE 3.3 Activation of Atlantic salmon Lxrα. Concentration-response curve for activation of Atlantic salmon Lxrα ligand-binding domain by T091317, GW3965, paxilline, cholesterol and oxysterols (20S-OH and 22R-OH). FHM cells were transfected with Gal4-LxrαLBD constructs, firefly luciferase reported plasmid pGL4.31 and an internal *Renilla* luciferase reporter to correct for transfection efficiencies. The ordinate represents activation of Lxrα as arbitrary units of firefly luciferase normalised to *Renilla* luciferase. Data points represent the mean of three independent experiments (n = 9).

3.3.2.2 Atlantic salmon Lxrα inactivation

The identification of Lxrα activating compounds was performed prior to assays to investigate the antagonistic effect of free fatty acids in the salmon Lxrα activation, which allowed us to choose functionally active concentrations for the strongest natural oxysterol, 22R-OH and for the synthetic agonist, GW3965. Our results, though preliminary, indicated that saturated fatty acids, PUFA or LC-PUFA significantly reduced the 22R-OH-driven Lxrα activation. The highest reduction in activity was observed when FHM cells were incubated with linoleic acid (LA), whereas DHA displayed the lowest inhibitory effect (FIGURE 3.4A). On the contrary, when GW3965-activated Lxrα was incubated together with either EPA or

DHA a reduction in the activation of the salmon Lxrα LBD was not observed when compared with cells incubated with the corresponding amounts of GW3965 and FAF FBS carrier alone (FIGURE 3.4B).

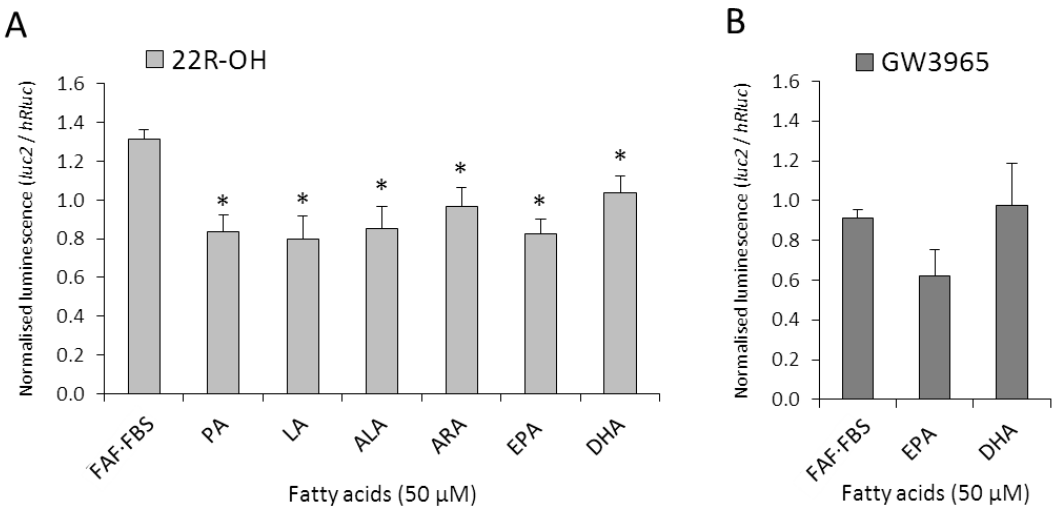

FIGURE 3.4 Inhibitory effects of various saturated and polyunsaturated fatty acids on Atlantic salmon Lxrα activation. Gal4-LxrαLBD was co-transfected into FHM cells with pGL4.31 and *hRluc* as a reference plasmid. Various fatty acids (50 µM) dissolved in fatty acid free foetal bovine serum (FAF FBS) together with either A) 22(R)-hydroxycholesterol (22R-OH, 20 µM), or B) GW3965 (20 µM) were added to the cells after transfection in medium with 10% FBS 24 h prior to the assay. After incubation, *Firefly* activity was measured and normalised to *Renilla* activity. The data shown constitute means ± SE of one independent experiment performed in triplicate. PA, palmitic acid (16:0); LA, linoleic acid (18:2n-6); ALA, α-linolenic acid (18:3n-3); ARA, arachidonic acid (20:4n-6); EPA, eicosapentaenoic acid (20:5n-3); DHA, docosahexaenoic acid (22:6n-3).

3.3.3 Ligand specificity of European plaice Rxr

Luciferase assays using the plaice Gal4-RxrγLBD chimera showed that plaice Rxrγ is significantly activated by the previously reported RXR agonists in mammals, 9cRA and LG268, the pollutant TBTO, the synthetic LXR agonist GW3965 and DHA (FIGURE 3.5). The response diagram indicated that 9cRA and LG268 were the most potent activators exhibiting increases of 50- and 30-fold, respectively. TBTO increased plaice Rxrγ activity 9-fold, whereas the synthetic LXR agonist GW3965

and n-3 DHA displayed the smallest effect with values double those observed under incubation with the carrier (EtOH) alone. No response was detected for PFOA.

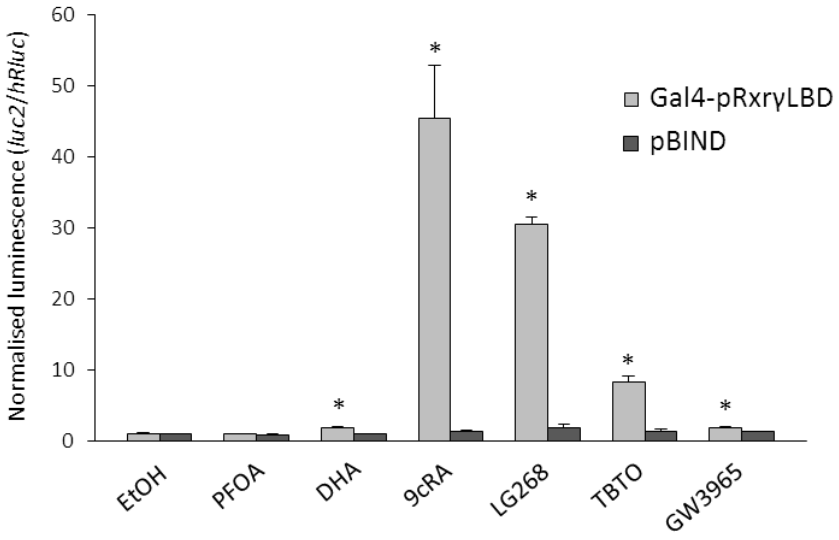

FIGURE 3.5 Activation of European plaice Rxrγ. Response to activation of European plaice Rxr ligand-binding domain by fatty acid DHA (22:6n-3, 50 μM), aquatic contaminants perfluorooctanoic (PFOA, 20 μM) and tributyl tin oxide (TBTO, 50 μM), Lxr agonist GW3965 (10 μM), known mammalian RXR agonists 9*cis*-retinoic acid (9cRA, 10 μM) and LG268 (10 μM), or ethanol carrier (EtOH). FHM cells were transfected with Gal4-RxrγLBD constructs, *Firefly* luciferase reported plasmid pGL4.31 and an internal *Renilla* luciferase reporter to correct for transfection efficiencies. The ordinate represents activation of Rxrγ as arbitrary units of firefly luciferase normalised to *Renilla* luciferase. Data points represent the mean of one independent experiments (n = 3 ± SE).

3.3.4 Gene expression in SHK-1 cells

Salmon SHK-1 cells were incubated with LXR (T091317 and GW3965) and PPAR (WY14643 and 2-bromopalmitate) agonists to identify respective target genes. Incubation with synthetic LXR ligands caused potent induction of *srebp* genes, with up to 9-fold increase for *srebp1* and 2-fold for *srebp2* (FIGURE 3.6). Lxrα agonists GW3965 and T091317 increased the expression of *fas* mRNA 3- and 6-fold, respectively (FIGURE 3.7). Fatty acyl desaturases, *fads2d6a* (~3.3-fold), *fads2d6b* (~2.4-fold) and *fads2d5* (4-fold) mRNAs were increased by GW3965. Only *fads2d5*

expression was affected by T091317 being increased < 2-fold (FIGURE 3.7). The expression of mRNA for *abca1* transporter, which regulates cholesterol efflux, was strongly upregulated by both Lxrα agonists, with ~14-fold increases (FIGURE 3.7). In the presence of Ppar agonists, only *srebp2* expression responded, relatively weakly, to 2-bromopalmitate (FIGURE 3.6), which is a dual Pparα and Pparβ agonist. However, WY14643, a specific Pparα agonist had no significant effect on *srebp2* expression. The expression of acyl-CoA oxidase (*acox*), HMG-CoA reductase (*hmgCoAR*) and fatty acid elongases, *elovl4*, *elovl5a* and *elovl5b* (FIGURE 3.7) *and ppars* (FIGURE 3.6) did not change significantly after incubation with LXR or PPAR ligands.

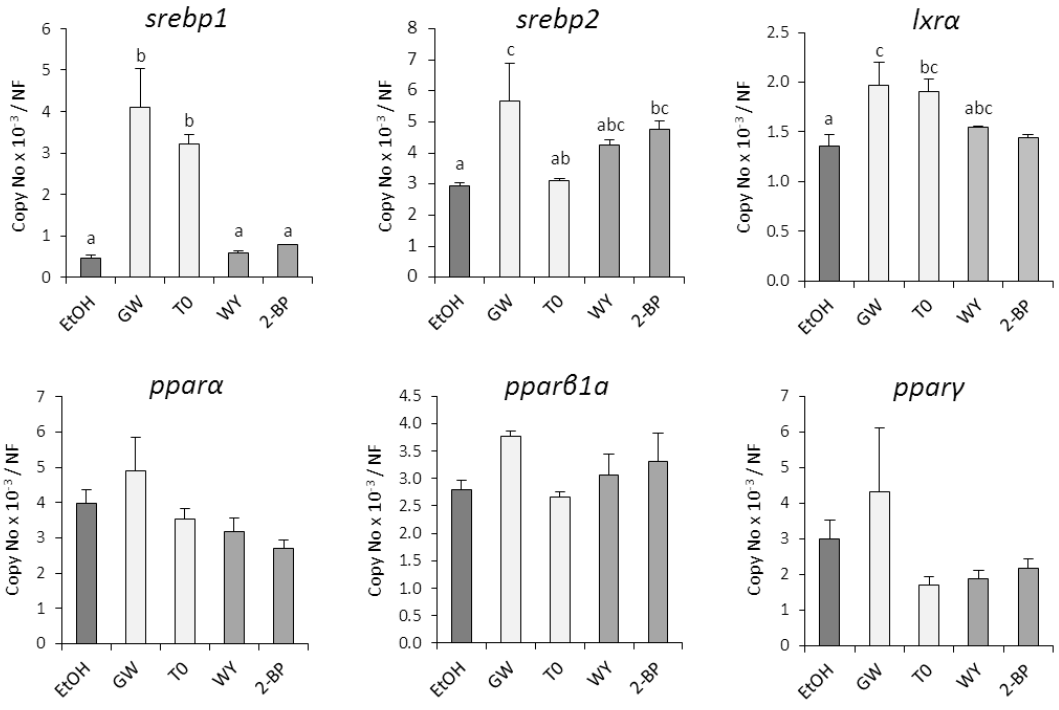

FIGURE 3.6 Effects of Lxr and Ppar agonists on transcription factor gene expression in Atlantic salmon SHK-1 cells. Expression of *lxrα*, *srebp1*, *srebp2*, *pparα*, *pparβ1a* and *pparγ* in SHK-1 cells exposed to 10 μM of Lxr agonists T0901317 (T0) or GW3965 (GW), 25 μM of Ppar agonists 2-bromopalmitate (2-BP) or WY14643 (WY), or carrier alone (EtOH). Gene expression is expressed as mean normalised ratios (n = 3, ± SE) between the copy numbers of target genes and the mean copy number of the reference genes (*elf-1α* and *β-actin*). Bars bearing different letters are significantly different (ANOVA; Tukey's test; $P < 0.05$).

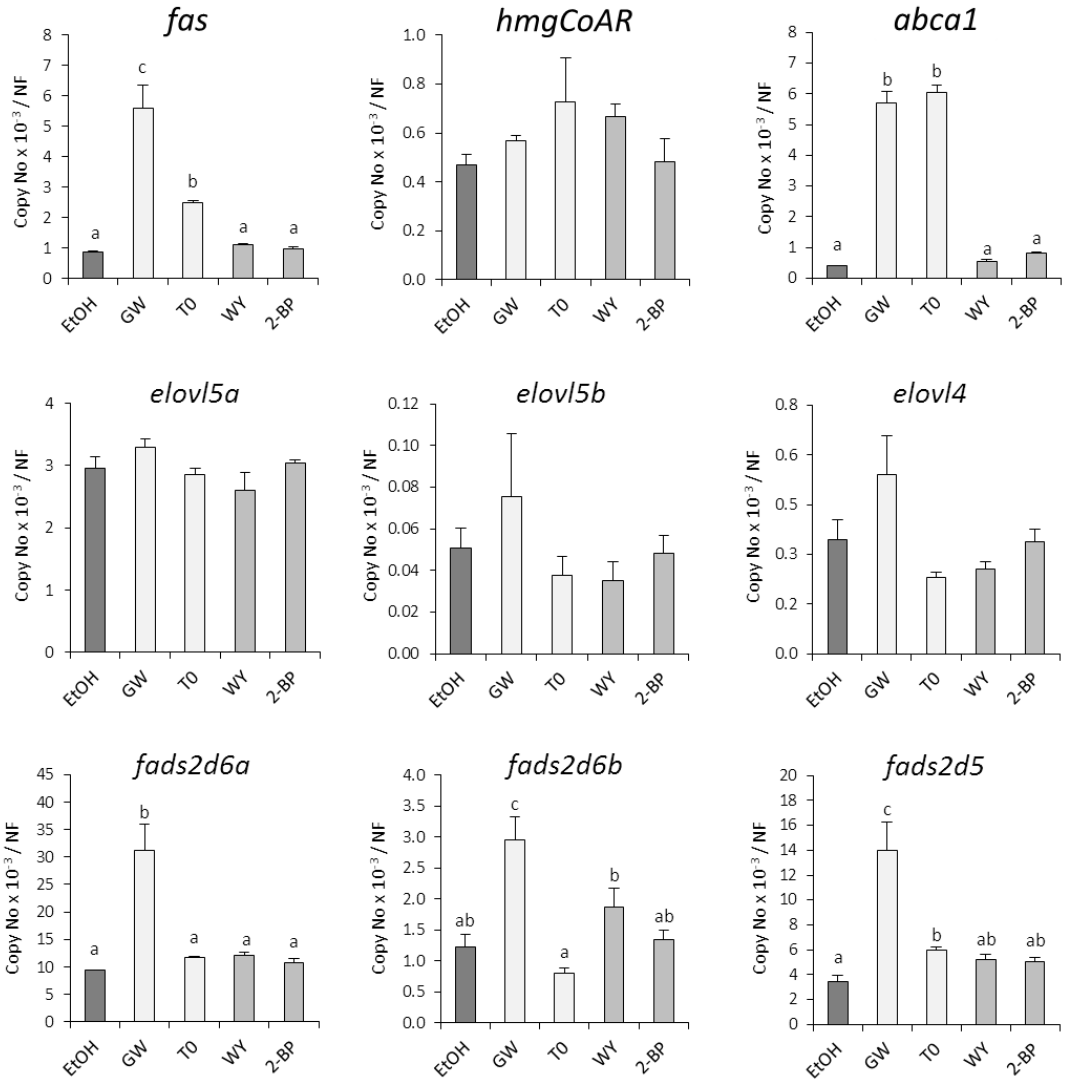

FIGURE 3.7 Effects of Lxr and Ppar agonists on lipid metabolic gene expression in Atlantic salmon SHK-1 cells. Expression of Atlantic salmon lipid metabolic genes in SHK-1 cells exposed to 10 μM of Lxr agonists T0901317 (T0) or GW3965 (GW), 25 μM of Ppar agonists 2-bromopalmitate (2-BP) or WY14643 (WY), or carrier alone (EtOH). Gene expression was expressed as mean normalised ratios (n = 3, ± SE) between the copy numbers of target genes and the mean copy number of the reference genes (*elf-1α* and *β-actin*). Bars bearing different letters are significantly different (ANOVA; Tukey's test; $P < 0.05$).

3.3.5 Trans-regulation of genes related to LC-PUFA metabolism by Srebp, Lxr and Ppar

To investigate the regulatory role of Srebp, Lxr and Ppar on the expression of key genes of LC-PUFA biosynthesis, we established Atlantic salmon promoter

luciferase assays in FHM cells. The Srebp transcriptional regulation of duplicated *elovl5a* and *elovl5b* elongases, and *fads2d6a* desaturase promoters was assessed by co-transfecting with the active nuclear DNA-binding region of either Srebp1 or Srebp2. Consistent with previous *in vitro* observations that f*ads2d6a* contains an SRE response element in the promoter region (Zheng *et al*., 2009b) the data indicate that both Srebp1 and Srebp2 promote the expression of *fads2d6a*. The promoters of both *elovl5a* and *elovl5b* duplicated genes were also highly activated (FIGURE 3.8) and Srebp2 showed higher activity than Srebp1 for *elovl5b* and *fads2d6a*, whereas *elovl5a* was activated equally by both Srebp mature proteins. Co-transfection of FHM cells with ligand-activated salmon Ppars did not stimulate expression from *elovl5a*, *elovl5b*, or *fads2d6a*; promoters, however the *fads2d6a* promoter was significantly activated by Lxr (FIGURE 3.9).

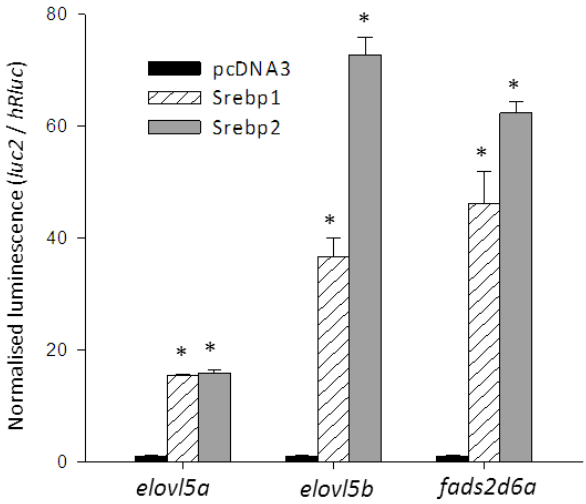

FIGURE 3.8 Effects of Srebps on LC-PUFA gene promoters. Co-transfection in FHM cells with Srebp1, or Srebp2 expression constructs (empty pcDNA3 expression vector as control), and reporter pGL4.10 [*luc2*] directed by the promoters of salmon *elovl5a* (- 3618 nt), *elovl5b* (- 3141 nt), or *fads2d6a* (- 1791 nt). Error bars indicate SE between data points (n = 3) of independent luciferase assays. *Significant differences between treatments and control are indicated (ANOVA; Tukey's test; $P < 0.05$).

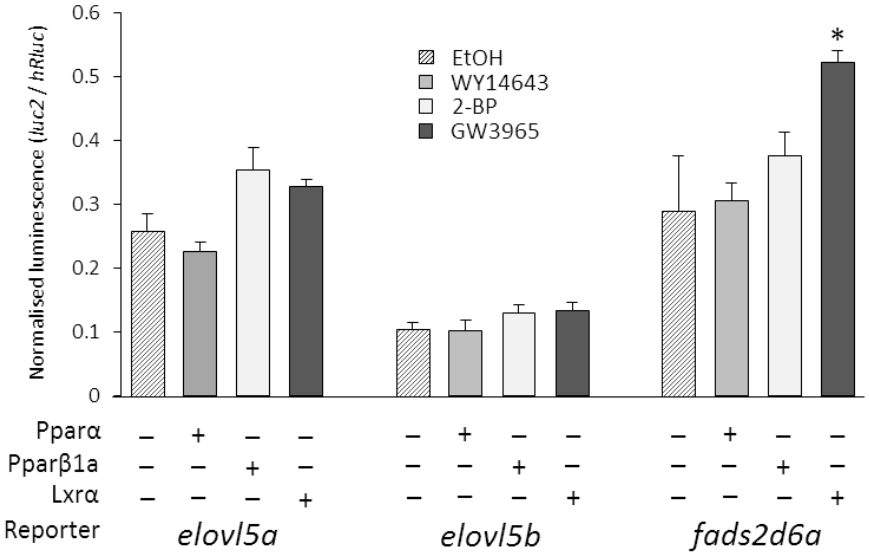

FIGURE 3.9 Effects of Lxr and Ppars on LC-PUFA gene promoters. Co-transfection in FHM cells with Pparα or Pparβ expression constructs (empty pcDNA3 expression vector as control), and reporter pGL4.10 [*luc2*] directed by the promoters of salmon *elovl5a* (- 3618 nt), *elovl5b* (- 3141 nt), or *fads2d6a* (- 1791 nt). After transfection, FHM cells were incubated with Ppar agonists WY14643 (25 µM) or 2-bromopalmitate (2-BP, 25 µM), Lxr synthetic agonist GW3965 (10 µM), or ethanol carrier (EtOH). Error bars indicate SE between data points (n = 3) of independent luciferase assays. *Significant differences between treatments and control are indicated (ANOVA; Tukey's test; $P < 0.05$).

3.3.5.1 Trans-regulation of *elovl5* duplicates by co-expression with salmon Lxr and plaice Rxr

In mammals, LXR functions by forming obligate heterodimers with the RXR. Thus, to investigate more thoroughly the regulatory role of Lxr on the expression of salmon duplicated *elovl5* genes, luciferase assays were carried out in FHM under simultaneous transfection of Rxrγ and Lxrα. In contrast to the above results, *elolv5a* showed a significantly increased response to ligand-activated Lxrα when it was in the presence of Rxrγ, whereas *elovl5b* did not seem to be regulated under any of the conditions included in the assay (FIGURE 3.10). In a similar experiment we exclusively tested the response of the salmon *elovl5a* promoter in the presence of

either Rxr or Lxr agonists. These results indicated that neither 9cRA nor LG268 Rxr agonists were capable of regulating the expression of *elovl5a* in Lxrα and Rxrγ replete cells. Only in cells transfected with Rxrγ (alone or in the presence of Lxrα) and incubated with GW3965 did we observed a significant response elicited by the *elovl5a* promoter, although activation was considerably higher when Lxrα and Rxrγ were simultaneously overexpressed (FIGURE 3.11).

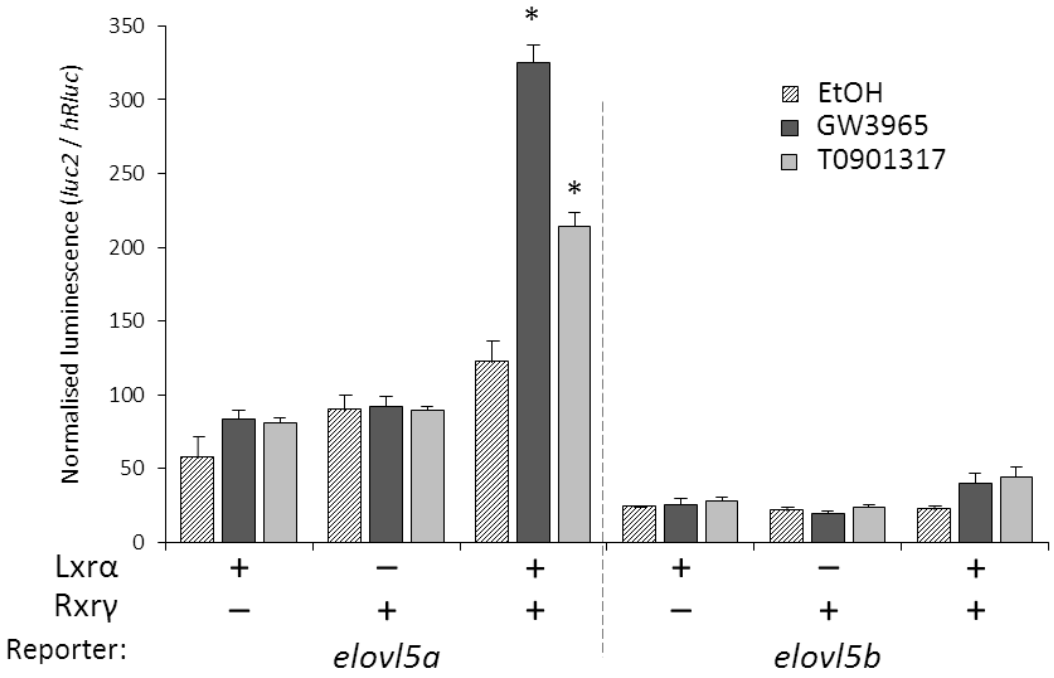

FIGURE 3.10 Effects of Lxr and Rxr on *elovl5a* and *elovl5b* LC-PUFA duplicated gene promoters. Co-transfection in FHM cells with salmon Lxrα and/or plaice Rxrγ expression constructs (empty pcDNA3 expression vector as control), and reporter pGL4.10 [*luc2*] directed by the promoters of salmon *elovl5a* (- 3618 nt) or *elovl5b* (- 3141 nt). After transfection, FHM cells were incubated with Lxr synthetic agonists GW3965 (10 µM) or T0901317 (10 µM), or ethanol carrier (EtOH). Error bars indicate SE between data points (n = 3) of independent luciferase assays. *Significant differences between treatments and control are indicated (ANOVA; Tukey's test; $P < 0.05$).

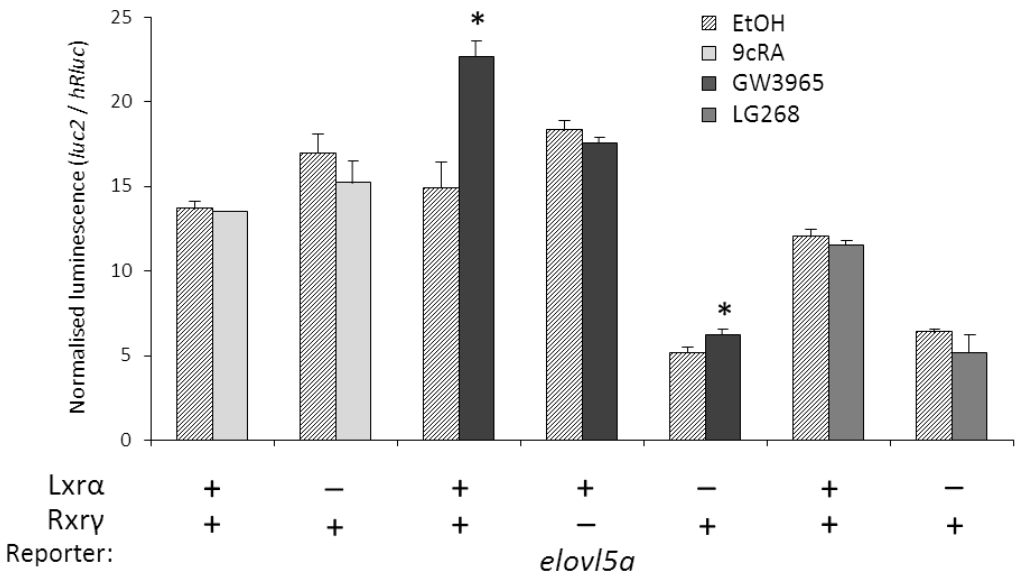

FIGURE 3.11 Effects of Lxr and Rxr on *elovl5a* LC-PUFA gene promoters. Co-transfection in FHM cells with salmon Lxrα and/or plaice Rxrγ expression constructs (empty pcDNA3 expression vector as control), and reporter pGL4.10 [*luc2*] directed by the salmon *elovl5a* (- 3618 nt) promoter. After transfection, FHM cells were incubated with Rxr agonists 9cRA (10 µM) or LG268 (10 µM), Lxr synthetic agonist GW3965 (10 µM), or ethanol carrier (EtOH). Error bars indicate SE between data points (n = 3) of independent luciferase assays. *Significant differences between treatments and control are indicated (ANOVA; Tukey's test; $P < 0.05$).

3.3.6 Nutritional regulation of the expression of transcription factors

Regulation of TF genes in response to dietary lipid composition was examined in pyloric caeca from salmon post-smolts fed diets containing either FO rich in EPA, DHA and cholesterol, or VOs rich in C18 FA, 18:1n-9 (RO), 18:2n-6 (SO), or 18:3n-3 (LO) but lacking LC-PUFA and cholesterol (Leaver *et al*., 2008b) (TABLE 3.1). Compared to the FO-fed group, there was a significant increase of *srebp1* and *srebp2* transcripts in the pyloric caeca of fish fed the RO and LO diets (FIGURE 3.12). The same trend was observed in fish fed the SO diet suggesting biological significance although not statistically. No differences were found between the dietary groups with regard to the expression of *lxrα* or *ppar* α, β or γ.

3.4 DISCUSSION

Elucidating the regulation of lipid and fatty acid metabolism in fish at a fundamental level is critical to understanding the relationships between lipid biosynthesis and metabolism and dietary lipid supply in fish, and will be crucially important to sustain the growth of aquaculture against a background of static or diminishing supplies of fish oil derived from wild fisheries. These relationships might be especially complex in fish such as Atlantic salmon, where extensive gene duplication and subsequent neo- or subfunctionalisations could give rise to many more gene regulatory interactions, than in vertebrates with unduplicated genomes. To address these issues we have undertaken this study to investigate the role of key lipid-regulated TFs on several salmon genes in an Atlantic salmon cell line, SHK-1, with a focus on duplicated *elovl5* genes of the LC-PUFA biosynthetic pathway. The SHK-1 cell line was initially developed from salmon head kidney tissue and is possibly leucocyte-derived, showing both macrophage- or dendritic-like phenotypes (Dannevig *et al*., 1997). However, as a suitable salmon liver- or intestine-like cell line is lacking, the SHK-1 cell line was recently tested, and successfully utilised, as a model for studying salmon lipid metabolism (Minghetti *et al*., 2011) on the basis that the critical transcriptional regulators of macrophage lipid homeostasis were the same as those in liver and they regulated similar target genes (Caimari *et al*., 2010; Castrillo and Tontonoz, 2004). Thus, the response of this cell line to supplementation or depletion of cholesterol or fatty acids suggested the involvement of Srebp and Lxr in regulating several critical lipid biosynthetic genes (Minghetti *et al*., 2011). To extend these studies and investigate the roles of individual TFs and their target genes more specifically we first characterised the ligand-activation dependency of fish

Lxrα and Rxrγ and developed promoter/reporter gene constructs to investigate the role of Srebp, Lxr and Ppar in Atlantic salmon.

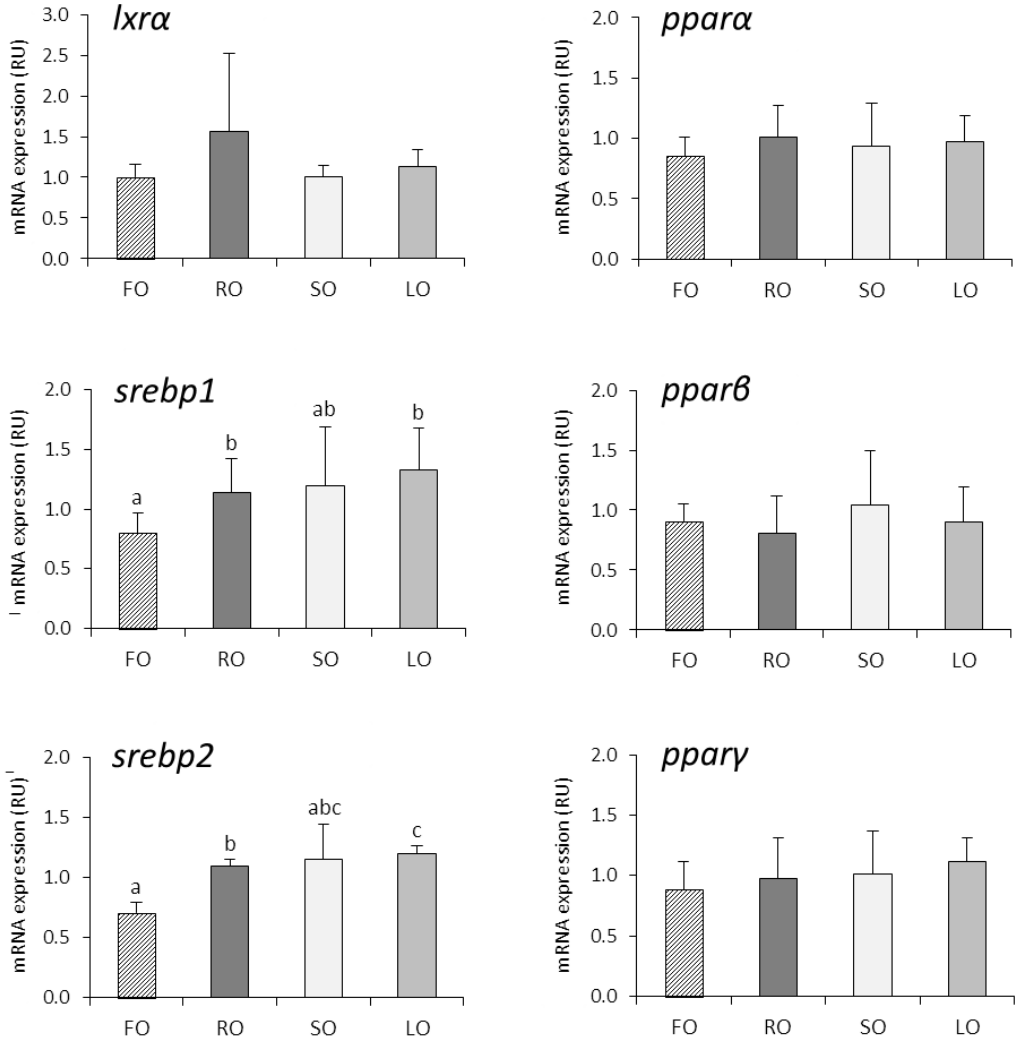

FIGURE 3.12 Nutritional regulation of transcription factor mRNA in Atlantic salmon intestine. *lxrα*, *srebp1*, *srebp2*, *pparα*, *pparβ* and *pparγ* mRNA expression in Atlantic salmon fed diets containing fish oil (FO), rapeseed oil (RO), soybean oil (SO), or linseed oil (LO) in the pyloric caeca determined by RT-qPCR. The results shown are means (n = 5) ± SE of normalised expression (β-actin and elongation factor-1alpha reference genes) in relative units (RU). Bars bearing different letters represent significant differences between dietary treatments for the respective transcripts (REST-MCS 2009 V2; $P < 0.05$).

A single Lxr subtype has been identified in various fish species (compared to two in mammals), and all show a highly conserved structure across teleost fish, amphibians, birds and mammals (Reschly *et al*., 2008). Although a salmon *lxrα* cDNA had been characterised the response of the corresponding receptor to activating ligands had not been tested (Cruz-Garcia *et al*., 2009). Accordingly, concentration-response curves for Lxr ligands indicated that natural oxysterols including 20S-OH and 22R-OH (but not cholesterol) and the fungal metabolite paxilline were activators of Atlantic salmon Lxrα. As with other vertebrate LXRs (Reschly *et al*., 2008), the synthetic agonists T0901317 and GW3965 were strong activators of Atlantic salmon Lxrα. Thus the synthetic agonists T0901317 and GW3965 were considered suitable tools for cellular assays of Lxr function. Ou and colleagues (2001) established that fatty acids acted as competitive antagonists of natural oxysterols, whereas synthetic agonists rapidly reversed the suppressive effect of unsaturated fatty acids, possibly due to their stronger binding affinities (Ou *et al*., 2001). Incubation of mammalian cells with fatty acids of different nature in the presence of LXR ligand binding domain-containing construct indicated that LC-PUFA exhibited the strongest inhibitory effect (Yoshikawa *et al*., 2002). On the contrary, our results indicated that all fatty acids tested including saturated and unsaturated had a statistically undistinguishable antagonistic effect in the ligand-dependent activation of salmon Lxrα. The ability of synthetic LXR ligand to overcome FA antagonism was observed with salmon Lxrα, as it has been reported with mammalian LXRs.

European plaice has been intensively studied as an ecological sentinel, and intense research has been directed into understanding the disruption/interaction of

aquatic pollutants and metabolic networks. Hence, available nuclear receptors such as Ppars and Rxr had been previously cloned from this species at the Institute of Aquaculture (Colliar *et al*., 2011). Following a thorough sequence analysis we observed that the DBD and LBD sequences are highly conserved among vertebrates. Accordingly, our results suggested that plaice Rxrγ heterodimerise with salmon Lxrα and bind to DR4 elements (LXRE) indicating once more that the DNA binding properties of the C-domain of these receptors are conserved in lower vertebrates (Leaver *et al*., 2005; Zhao *et al*., 2000). The characterisation of the ligand-binding domain indicated that similarly to what has been observed in higher vertebrates, plaice Rxrγ is strongly activated by mammalian RXR agonists (Tóth *et al*., 2011), and slightly by docosahexaenoic acid (DHA) (Lengqvist *et al*., 2004; Urquiza, 2000), but also by LXR agonist GW3965 and the pollutant TBTO. These results indicate that TBTO, used as preservative in marine antifouling paints, is potentially a potent Rxr mediated endocrine-disrupting compound in fish, as previously reported in mammals (Grün *et al*., 2006; Hiromori *et al*., 2009). The genetic control of lipid metabolism involves a complex interplay of transcription factors and regulatory loops acting on many genes of lipid metabolism and transport. For example, in mammals, LXR can mediate the regulation of lipogenesis through the direct activation of genes involved in lipid biosynthesis, or in a SREBP1c-dependent manner (Joseph *et al*., 2002; Repa *et al*., 2000a, 2000b). In mammals, two SREBP1 isoforms, SREBP1a and SREBP1c, are encoded by a single gene through the use of alternative start transcription sites, while a separate gene encodes Srebp2 (Hua *et al*., 1995). Given this complexity, elucidated from studies in human and mammalian models, the present study aimed to determine if similar regulatory systems exist in

fish due to the increasing importance of understanding basic lipid metabolic processes in modern aquaculture. In agreement with recent studies performed in rainbow trout (*Oncorhynchus mykiss*) (Cruz-Garcia *et al*., 2011, 2012), treatment of Atlantic salmon SHK-1 cells with LXR agonists resulted in the upregulation of mRNAs for several important lipid metabolic genes, including *fas* involved in *de novo* biosynthesis of fatty acids and *abca1*, which controls the reverse cholesterol efflux (See diagram in FIGURE 3.13). Ligand-activated Lxrα also induced the expression of both *srebp1* and *srebp2* in SHK-1 cells. Similarly, *SREBP1c* and also *FAS* and *ABCA1* are established direct LXR targets in various human and rodent systems (Joseph *et al*., 2002; Laffitte *et al*., 2001; Repa *et al*., 2000b). However, the LXR-mediated induction of *Srebp2* is not observed in mammals (Repa *et al*., 2000a), and may be specific to Atlantic salmon. According to our results, several groups have reported LXR autoregulatory behaviour in mammals through binding to LXRE (LXR response element) present in LXRα promoter (Laffitte *et al*., 2001).

Although vertebrates have the capability for endogenous synthesis of monounsaturated fatty acids from the saturated fatty acid products of FAS through the action of stearoyl CoA desaturase (SCD or Δ9 desaturase), they are incapable of creating LC-PUFA *de novo*. Vertebrates thus require either dietary LC-PUFA directly or, depending upon species, they can produce the physiologically important ARA, EPA and DHA by desaturation and elongation of dietary shorter chain PUFA, linoleic (18:2n-6) and α-linolenic (18:3n-3) acids, through the action of Fads1 and 2 and Elovl5 and Elovl2 (Tocher, 2003). Intense interest has been paid to vertebrate LC-PUFA enzymes given their hormonal, ontogenetic and/or nutritional regulation in vertebrates (Wang *et al*., 2005, 2006). Salmon Elovl5 enzymes are encoded by two

very similar duplicated genes (Morais *et al*., 2009). Neither mRNA was affected by Lxr agonist in SHK-1 cells, despite the *elovl5* gene promoters being activated by co-transfection with nSrebps, which indicates that, although Srebps were induced in the cell line, under the experimental condition used they were not processed to transcriptionally active forms. In mammals, *Elovl5* is not an LXR target, but is secondarily activated through LXR effects on Srebp1c (Qin *et al*., 2009). In salmon, on the other hand, *in vitro* co-expression of Lxrα and Rxrγ suggested that *elovl5a*, but not *elovl5b*, is likely to be a target of Lxr-Rxr heterodimers. Normally, mammalian LXRs act as permissive heterodimers with RXR, which means that upon formation the heterodimer can be activated by either LXR or RXR agonists raising the possibility of pleiotropic signalling (Son and Lee, 2009; Széles *et al*., 2010; Zelcer and Tontonoz, 2006). This does not seem to be the case in SHK1 cells because the Lxr-Rxr heterodimer was not capable of driving the same response observed with Lxr agonists when it was incubated with RXR agonists, suggesting that Lxrα is likely to be the main element responsible for the response observed on *elovl5a*. However, given that Rxrγ LBD was significantly although mildly upregulated by Lxr agonist GW3965 the exact regulatory mechanism of Lxr/Rxr on *elovl5a* remains unclear. Salmon *fads2* were significantly increased by Lxr agonist *in vitro*, and *fasd2d6a* was strongly activated by nSrebps indicating that this gene is a direct target of both transcription factors. In mammals, LXR agonists increased *Fads1* and *Fads2* expression although the results were attributed to a secondary LXR-dependent induction of SREBP1c (Nara *et al*., 2002). The promoter element of the *Fads2* gene conferring Srebp1 response is highly conserved in several vertebrates

including Atlantic salmon, with a potential SRE site and a NF-Y cofactor site (Zheng *et al*., 2009b).

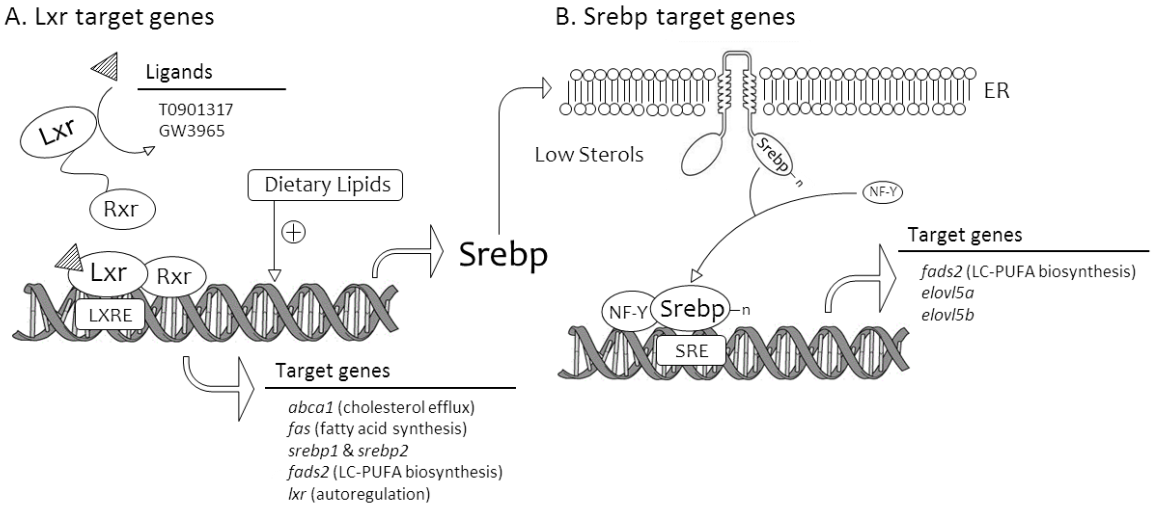

FIGURE 3.13 Diagram indicating Lxr and Srebp target genes in Atlantic salmon. A) SHK-1 cells incubated with synthetic Lxr agonists (T0901317 and GW3965) showed that Lxrα regulates multiple genes of the lipid biosynthetic pathway, mediating cross-regulation between cholesterol and fatty acid biosynthesis. Transcriptional expression of *srebp* was induced by Lxr agonist *in vitro* and dietary lipids (vegetable oils) in salmon intestine. B) Co-transfection assays in FHM cells indicated that Srebp1 and Srebp2 promote the transcription of *elovl5* and *fads2*, genes of the LC-PUFA biosynthesis; ER, endoplasmic reticulum; n, amino-terminus.

Notably, Atlantic salmon *elovl5* duplicates exhibit differential tissue expression patterns and respond differently to dietary nutrients (Carmona-Antoñanzas *et al*., 2013b; Morais *et al*., 2009). Accordingly, our results show that *elovl5a* is equally regulated by both Srebp1 and Srebp2, whereas *elovl5b* regulation by Srebp2 is 2-fold greater than that exhibited in response to Srebp1. In addition, whereas *elovl5b* showed no response to ligand-activated Lxr/Rxr heterodimer, the *Firefly* gene driven by the salmon *elovl5a* promoter was significantly upregulated by Lxr/Rxr in the presence of Lxr agonists, possibly as a result of sequence divergence in the promoter regions of these genes.

PPARs are also central to the transcriptional control of lipid metabolism and several interactions with LXRs and SREBPs have been proposed (Martens *et al*., 2008; Michalik *et al*., 2006). Mice knockout assays and subsequent *in silico* analysis confirmed that PPARα induces *Srebp1* and *Srebp2* expression through the interaction with PPRE (PPAR response element) in the mammalian *Srebp2* promoter (Martens *et al*., 2008). In SHK-1 cells, few effects were observed upon treatment with WY14643, a Pparα specific ligand, or 2-bromopalmitate, a non-metabolisable fatty acid activator of teleost Pparα and Pparβ (Colliar *et al*., 2011). However, bromopalmitate induced the expression of *srebp2*, whereas WY14643 had no effect. This suggests a Pparβ-mediated effect, although it is also possible that the lack of response to the Pparα-specific WY14643 is due to a limited *ppar* subtype expression profile in SHK-1 cells. Despite this, no evidence was found to suggest a regulatory role of Ppars in the transcriptional control of *elovl5* or *fads2* expression, although, given the results with Lxrα, further analyses including co-transfection with Rxr should be carried out.

The *in vivo* significance of the results from the SHK-1 cell line and promoter assays was previously shown by the observation that hepatic *srebp2* mRNA levels and LC-PUFA biosynthetic gene expression and biosynthesis increased in salmon fed diets containing high levels of VO (Leaver *et al*., 2008b). In this study, similar results for the expression of LC-PUFA biosynthetic genes were observed in the intestine (pyloric caeca) of salmon fed VO (Morais *et al*., 2009). The pyloric caeca of salmon constitutes the major tissue in terms of nutrient uptake, and lipid digestion and transport (Krogdahl *et al*., 1999), and is also the tissue that exhibited the highest level of expression of *srebp*s and *lxr* mRNAs in Atlantic salmon (Minghetti *et al*.,

2011). Therefore, if the LC-PUFA biosynthetic genes, *elovl5* and *fads2*, are driven by Srebps *in vivo*, then the same increase in Srebps should be observed in intestine from VO-fed fish. As the present study has shown, this was indeed the case. Our results indicated that, among the transcription factors studied, only *srebp1* and *srebp2* expression was increased significantly in the pyloric caeca of Atlantic salmon fed rapeseed oil (RO) and linseed oil (LO) compared to the expression in fish fed FO. Plant-derived products constitute the perfect activation ground for SREBPs. In the absence of sterols, the TF amino-terminus is proteolytically released from the endoplasmic reticulum membrane and transported into the nucleus where it binds SRE (serum response element) of specific sets of target genes (Espenshade, 2006). In mammals, SREBP1 recognises SRE sites in the promoters of *Srepb1* and *Srebp2* (Amemiya-Kudo *et al*., 2002), thus it is likely that salmon Srebp1 activated in low sterol conditions, induced by the inclusion of VO, promoted its own expression and that of *srebp2* in a positive-feedback mechanism. Although other TF responses to dietary lipid variations in fish have been reported at specific developmental stages (Cruz-Garcia *et al*., 2009), no effects of VO inclusion were observed in the expression of *lxr* or *ppar* in this study.

3.4.1 Conclusions

The results of the present study showed that the fatty acyl elongases and desaturases responsible for endogenous production of LC-PUFA from PUFA in Atlantic salmon were primarily regulated by Srebps, with a possible contribution from Lxr, whereas there was no evidence for a direct role of Ppars, at least in the salmon cell line tested. ABCA1, a gene central to the process of reverse cholesterol transport, and which is a direct LXR target in mammals, was also a target of Lxr in

salmon. Furthermore *fas* and both *srebp* genes, responsible for the major steps of lipogenesis, are direct Lxr targets in salmon as they are in mammals. Thus, overall, the transcriptional regulatory systems that drive cholesterol transport from the cellular space to lipoprotein, initiate lipogenesis, and regulate LC-PUFA biosynthesis in mammals are largely conserved in Atlantic salmon. This knowledge will be key to deriving a conceptual framework for future experiments designed to answer more applied questions related to lipid metabolism and nutrition with regard to the development and optimisation of more sustainable aquaculture feeds. Importantly the present results also showed that these lipid regulatory factors and the genes that they target are ancient, likely to be have arisen early in vertebrate evolution. Thus, further basic studies of evolutionarily conserved pathways of lipid metabolic control across vertebrates might also elucidate, through definition of fundamental regulatory mechanisms, medically relevant aspects of human lipid nutrition and metabolism.

Chapter 4

Functional promoter analyses of duplicated *elovl5* LC-PUFA biosynthetic genes in Atlantic salmon

Greta Carmona-Antoñanzas [§], Xiaozhong Zheng, Douglas R. Tocher, and Michael J. Leaver

School of Natural Sciences, University of Stirling, Institute of Aquaculture, Stirling FK9 4LA, Scotland, UK

Authors' contributions

GCA, DRT and MJL planned and coordinated the research; GCA performed laboratory analyses and data analysis; XZ conducted the gene cloning; GCA wrote the first draft of the manuscript, followed by contributions from remaining authors.

Publication in preparation.

Abstract

Control and coordination of eukaryotic gene expression relies on transcriptional regulation of *cis*-regulatory elements by transcription factors. It has been hypothesised that asymmetrical evolution of regulatory regions enable duplicated genes to develop new functions under relaxed purifying selection. The transcriptional functionality of upstream regulatory regions of Atlantic salmon (*Salmo salar*) duplicated elongases of fatty acids, *elovl5a* and *elovl5b*, were compared in cellular transfection assays. Both Elovl5 duplicates possessed conserved regulatory elements that promoted Srebp1- and Srebp2-dependent transcription. Differences in the magnitude of Srebp response between promoters could be attributed to a tandem duplication of SRE and NF-Y cofactor sites in *elovl5b*. In addition, the *elovl5a,* promoter, was shown to have a site conferring an Lxr/Rxr-dependent response. This regulatory region is in a non-conserved sequence likely inserted into the upstream region of *elovl5a* shortly after the gene duplication event. This region is highly repeated in the Atlantic salmon genome and has sequence similarity to mobile DNA elements. Comparison of the detailed sequence structures of *elovl5a*, *elovl5b* and the single *elovl5* gene in the preduplicated relative, Northern pike, show intense and asymmetrical insertion of transposons and other repeated elements throughout the salmon *elovl5* duplicate genes.

4.1 INTRODUCTION

Extant salmonids are pseudotetraploid descendants of a common ancestor whose genome underwent duplication between 25 to 100 myr ago (Koop and Davidson, 2008). As a proposed outcome of the ancestral genome duplication, Atlantic salmon possess several copies of Fads2 and Elovl5, hypothesised to have functionally diverged and thus physiologically enabled Atlantic salmon to thrive in LC-PUFA-poor environments (Carmona-Antoñanzas *et al*., 2013b; Leaver *et al*., 2008b). Recent functional and phylogenetic analyses indicated that duplicated Elovl5 LC-PUFA proteins are subject to strong functional constraints as suggested by comparative studies with the closest extant preduplicated genome, Northern pike (*Esox lucius*) (Carmona-Antoñanzas *et al*., 2013b; Morais *et al*., 2009). However, the nutritional regulation of paralogous LC-PUFA genes differ in tissues from fish fed diets containing varying levels of LC-PUFA and sterols (Leaver *et al*., 2008b), and their expression also exhibit different responses to overexpression of transcription factors *in vitro* (Carmona-Antoñanzas *et al*., 2013a; Morais *et al*., 2009).

Transcriptional regulation constitutes a sophisticated mechanism to control the differential spatial and temporal expression patterns in specific cell types during development, and in response to a wide variety of external signals. Gene promoters, responsible for initiation of transcription, consist of a core motif required for binding of the transcription preinitiation complex recognised by RNA polymerase (Javahery *et al*., 1994; Smale and Kadonaga, 2003; Xi *et al*., 2007), and stimulating upstream *cis*-regulatory elements. Evolutionary divergence of promoter regions constitutes one of the major mechanisms of transcriptional variation resulting in adaptation and speciation. Romero *et al*. (2012), and Castillo-Davis *et al*. (2004) identified a

dramatic acceleration in the evolution of regulatory regions of duplicated genes, and consequential divergence or acquisition of new *cis*-regulatory regions has recently been suggested to be responsible for the differential expression patterns of duplicated genes (Force *et al*., 1999; Gu *et al*., 2002).

DNA-bound transcription factors (TF) constitute major transcriptional regulators of metabolic homeostasis (Desvergne *et al*., 2006). In vertebrates, particular attention has focused on sterol regulatory element binding proteins (SREBP) and liver X receptor (LXR), given their involvement in nutritional regulation, and cross-regulatory role between cholesterol and fatty acid biosynthesis (Carmona-Antoñanzas *et al*., 2013a; Matsuzaka *et al*., 2002; Qin *et al*., 2009; Zhao and Dahlman-Wright, 2010). SREBPs are a type of membrane-bound basic helix-loop-helix leucine zipper (bHLHZip) TF that bind sterol regulatory element (SRE) DNA sequence '*TCACNCCACA'*, adjacent to particular enhancer E-box, and cofactors NF-Y and Sp1, thus promoting transcriptional activation of target genes in a sterol or carbohydrate-regulated manner (Amemiya-Kudo *et al*., 2002; Espenshade, 2006; Kim *et al*., 1995; Magaña *et al*., 2000). SREBP1 and SREBP2 duplicated prior to vertebrate divergence (Rawson, 2003) vary in nuclear location, specificity for SRE recognition and towards sterol-activation (Amemiya-Kudo *et al*., 2000, 2002,; Hasty *et al*., 2000; Shimano, 2001), partially explaining their different *in vivo* functions. Promoter assays in fish cell lines indicated that salmon *elovl5a* responded similarly to Srebp1 and Srebp2, whereas *elovl5b* displayed a significantly increased response only to Srebp2 (Carmona-Antoñanzas *et al*., 2013a). In mammals, SREBP2 preferentially binds classic SRE, present in cholesterogenic genes, while lipogenic

genes exhibit a more diverse consensus sequence designated SRE-like sequence and can be regulated by either SREBP1 or SREBP2 (Amemiya-Kudo *et al*., 2002).

In contrast, LXR is a member of the nuclear hormone receptor family and, as such, its activation is dependent on ligand binding. In the case of LXR the physiological agonists are thought to be oxysterols, and LXR is also antagonised by the competitive interaction of fatty acids with the ligand binding domain (LBD) (Carmona-Antoñanzas *et al*., 2013a; Ou *et al*., 2001; Reschly *et al*., 2008; Yoshikawa *et al*., 2002). The activity of nuclear receptors is controlled by conformation changes induced upon ligand binding, and DNA dimerisation (Markov *et al*., 2010), which enable the formation of coactivator complexes that induce recruitment of basal transcriptional machinery and initiate transcription (Holmbeck *et al*., 1998; Zhao *et al*., 2000). In the nucleus, LXR recognises specific response elements, termed LXRE (LXR response elements) (Willy and Mangelsdorf, 1997; Yoshikawa *et al*., 2002). LXRE are direct repeats (DR4) defined as a variant of two '*AGGTCA'* sequences spaced by 4 nucleotides, whereas RXR homodimers preferentially bind DR1 and DR2 sequences (Egea *et al*., 2000; Zhao *et al*., 2000). Previous studies suggested that mammalian LXR was directly involved in the stimulation of LC-PUFA biosynthesis through targeting of ELOVL5 and FADS genes (Wang *et al*., 2006). Nevertheless, further analyses attributed such ability to the SREBP1-dependent regulation of LC-PUFA (Horton *et al*., 2002, 2003) when a LXRE was identified in the promoter region of SREBP1 (Ou *et al*., 2001; Qin *et al*., 2009; Wang *et al*., 2006). In *Chapter 3* evidence was presented that suggested LC-PUFA desaturase *fads2d6a* and elongase *elovl5a* promoters of Atlantic salmon were

transcriptionally regulated by Lxr and/or Rxr, whereas the paralog *elovl5b* promoter did not show a significant response.

This present study aimed to determine the structure of *elovl5* LC-PUFA duplicated genes in Atlantic salmon to identify marks of asymmetrical gene divergence, especially in *cis*-regulatory elements.

4.2 MATERIALS AND METHODS

4.2.1 FHM cell line and culture conditions

Fathead minnow (*Pimephales promelas*; FHM) epithelial cells were maintained at 25°C in Leibovitz's L-15 with GlutaMAX^{TM}-1 medium containing 900 mg L^{-1} D+galactose and 550 mg L^{-1} sodium pyruvate and supplemented with 10% FBS and antibiotics (Carmona-Antoñanzas *et al*., 2013a). All media and supplements were obtained from Life Technologies (Glasgow, UK). For subculturing, the cell monolayer was washed twice with PBS without $CaCl_2$ or $MgCl_2$ (Invitrogen, UK), cells detached by incubation with 0.05% Trypsin and 0.02% EDTA and re-suspended in supplemented medium. Viable cells were counted after harvesting using 0.4% Trypan blue (Sigma, Dorset, UK), a Neubauer haemocytometer and inverted microscope (IMT-2, Olympus). For gene promoter transactivation assays, FHM cells were harvested and seeded in 96-well black-sided clear-bottom microtitre plates (Corning, NY, USA) at a density of 3×10^4 cells per well in a volume of 100 µl of Leibovitz's L-15 medium. Cells were approximately 80% confluent after 24 h growth, when medium was aspirated, cells washed twice with PBS and medium replaced with transfection mixtures and agonists.

4.2.2 Cloning and fine-scale *elovl5* gene structure analysis

4.2.2.1 Atlantic salmon gene structure

An Atlantic salmon genomic DNA library was constructed in lambda FIX II (Stratagene, USA, Zheng *et al*., 2009). The salmon DNA library was screened with full-length cDNA probes of the salmon *elovl5* paralogs [*elovl5a*, GenBank: AY170327] and [*elovl5b*, GenBank: FJ237531]. Inserts of positive recombinant phages were isolated and subcloned into the pBluescript II KS vector (Life Technologies, USA) for sequencing using intrinsic *Sac I* restriction sites at the 5' end of the *elovl5* promoters. The full putative elongase genomic nucleotide sequences were assembled using the SeqMan II 6.1 module of the Lasergene (DNASTAR Inc., USA). Sequence homology between *elovl5* duplicates was identified using the MegaBLAST programme to perform pairwise alignments and search against the Atlantic salmon whole-genome shotgun contigs (wgs) database available in NCBI. The discontinuous MegaBLAST algorithm employs the "discontinuous word" strategy of Ma *et al*. (2002) that ignores some bases (allowing mismatches) to rapidly detect homologous genomic sequences ranging from about 75% to 90% nucleotide identity.

4.2.2.2 Northern pike gene structure

Biological samples were obtained as described previously (*Salmo salar*, Leaver *et al*., 2008b; *Esox lucius*, Carmona-Antoñanzas *et al*., 2013b). DNA was extracted from fin clips stored in 70% EtOH using RealPure Genomic DNA Extraction kit according to the manufacturer's instructions (REAL Durviz, Spain)

and quality tested by electrophoresis and spectrophotometry (Nanodrop 1000, Thermo Scientific, USA).

The exons flanking intronic regions in the Northern pike *elovl5* [GenBank: JX272634] were identified by aligning the pike cDNA sequence with the genomic sequences of Atlantic salmon *elovl5* paralogous genes (*elovl5a*, [GenBank: GU238431.1] and *elovl5b*, [GenBank: GU324549.1]). Amplification of pike introns was performed using MyTaq Mix (Bioline) with primer pairs designed in exons flanking the corresponding introns (TABLE 4.1). The purified PCR products were cloned into pCR2.1 vector (TA Cloning Kit, Invitrogen) using T4 DNA Ligase (Promega), and subsequently used to transform TOP10 Chemically Competent *Escherichia coli* cells (Invitrogen). Purified plasmids including pike introns were fully sequenced using GenomeLab DTCS Quick Start Kit and the CEQ-8800 Genetic Analysis System (Beckman Coulter Inc, USA).

4.2.2.3 Identification of interspersed (transposon-derived) sequences

The described Atlantic salmon *elovl5* genomic sequences and pike *elovl5* intronic regions were screened against a reference collection of interspersed (transposon-derived) repeats and low complexity DNA sequences including closely related species, zebrafish (*Danio rerio*) and Fugu pufferfish (*Takifugu rubripes*). RepeatMasker (Smit *et al*., 2010) and CENSOR (Kohany *et al*., 2006) web-based servers were used under default options to search a DNA query of interests against Repbase, a database of repetitive element consensus sequences in eukaryotic DNA (Jurka *et al*., 2005). RepeatMasker and CENSOR implement position-specific scoring matrices using alignment logarithms, Smith-Waterman and BLAST,

respectively, to create local pairwise alignments between canonical (reference) repeat sequences and the query (Pevsner, 2009). The local alignment algorithms resemble global algorithms (Needleman-Wunsch algorithm, *Chapter 2 & 3*) in that two sequences are arranged in a matrix; however, a local alignment does not contain the entire sequence of each OTU (operational taxonomic unit), but the optimal result is reduced to the regions of greatest similarity between two OTUs. The results were stringently filtered applying similarity (>70%) and alignment (>800) cutoff scores.

4.2.3 Promoter constructs, deletions and mutations

The regulatory regions of *elovl5a* (-4898 bp relative to ATG initiation codon; GenBank: GU238431.1) and *elovl5b* (-3143 bp relative to initiation codon; GenBank: GU324549.1) were amplified from genomic DNA using a proof-reading enzyme (*Pfu* DNA Polymerase, Promega, UK) and primers containing suitable restriction sites (TABLE 4.1) such that the ATG initiation codon of the luciferase gene in pGL4.10, *luc2* (Promega) was replaced by the initiation codon for each *elovl5* gene. The upstream limit for the putative promoter sequence was selected on the basis of the presence of a conserved *Sac*I site immediately beyond which no clear homology between the two *elovl5* sequences could be detected. Thus, the tested promoter regions, in addition to upstream untranscribed sequence, contained transcriptional start sites (TSS), an upstream non-coding exon and an ATG initiation codon residing within the boundary of the second exon. Each promoter construct was sequenced (Sanger ABI 8730xl, GATC Biotech) to confirm sequence identity and purified using anion-exchange purification columns (QIAfilter plasmid midi kit, Qiagen) for high transfection efficiency. The vectors containing the wild type

TABLE 4.1 Details of primer pairs (restriction sites for *Sac*I, *Xho*I and *Nco*I and mutated sites underlined) used for the cloning of Northern pike *elovl5* introns and the construction of promoter deletions and mutations.

Aim	Construct	Forward primer (5' → 3')	Reverse primer (5' → 3')
Pike *elovl5* cloning	Intron4	CCTTTGCACTGACCGTGATA	GCGTGTCCTGGCAGTAGAA
	Intron6	CTGCACATCTACCACCATGC	GATGGCAGAGAGCCCATAGT
Promoter deletions	SEA1	AATGAGCTCAGCTCTGCAAAGCCATGTG	AATCTCGAGTTCTGACCTAAATAGACAGATG
	SEA2	AATGAGCTCTCGGCACATGATCGGCC	
	SEA3	AATGAGCTCGAACACGGTGTGGTTGTCTG	
	SEA4	AATGAGCTCACGATCATACCACGCACAAA	
	SEA5	AATGAGCTCTCCCACAATGCAACACACTT	
	SEA6	AATGAGCTCGTATGCGTTTGGGACTGCTT	
	SEA7	AATGAGCTCCCATACATTTCAGGTCTTTC	
	SEB1	TATGAGCTCTGCTCTGTGTAGCCATG	CTCCATGGCTAACCTGAATAGACAGATG
	SEB2	AATGAGCTCCGGCTTGCTAACTGATTTATTA	
	SEB3	TATGAGCTCACAACTAAGATTTTCACTACT	
	SEB4	TATGAGCTCTGGAAACCTAGTTTAAGTGGG	
	SEB5	TATGAGCTCTGTGTTTCTCCCGCTGTTTCC	
	SEB6	TATGAGCTCTGCGTTTGGGACTGCTTGAGC	
	SEB7	TATGAGCTCAAATCAGGGTTCCCCAACTGG	
Promoter mutations	SEA1_M1	GATGCAGTGATTTGACCAATTTCCACTGACGAG	CTCGTCAGTGGAAATTGGTCAAATCACTGCATC
	SEA1_M2	CGAGATGTGAATAAGAAGACACTCGGCACATG	CATGTGCCGAGTGTCTTTTTATTCACATCTCG
	SEA1_M3	GTCATAACTTCATGAAATTTGATCACATGGT	ACCATGTGATCAAATTTCATGAAGTTATGAC
	SEA1_M4	TTCATGACCTTTGATTACATAGTTTTTGTCAAG	CTTGACAAAAACTATGTAATCAAAGGTCATGAA
	SEA1_M5	CATGGTTTTTGTCAAATTAATGCAGTCGTCATC	GATGACGACTGCATTAATTTGACAAAAACCATG
	SEA1_M6	GTTCATGCAGTCGTCATCTGTAGTCGGAAATCG	CGATTTCCGACTACAGATGACGACTGCATGAAC
	SEB1_M1 / SEB1_M3	GATGCAGTGATTTGACCAATCTCCACTGACGA	TCGTCAGTGGAGATTGGTCAAATCACTGCATC
	SEB1_M3 / SEB1_M2	GATGCAGTGATTTGACCAATCTCCACTGATGC	GCATCAGTGGAGATTGGTCAAATCACTGCATC
	SEB1_M4 / SEB1_M6	CGAGAAATGAATAAGAAGACATCGGCTTGCTAACT	AGTTAGCAAGCCGATGTCTTTTTATTCATTTCTCG
	SEB1_M5 / SEB1_M6	TGCGAAATGAATAAGAAGACATCGGCTTGCTTACC	GGTAAGCAAGCCGATGTCTTTTTATTCATTTCGCA

full-length promoters, pGL4.10-*elovl5a* and pGL4.10-*elovl5b*, were termed SEA1 and SEB1, respectively.

4.2.3.1 Deletion constructs

Progressive deletions of *elovl5* gene upstream sequences were constructed using constructs SEA1 and SEB1 as template for PCR amplification and primers containing restriction sites specified in TABLE 4.1. Nine or seven deletion constructs were produced from each, respectively.

4.2.3.2 Selection of sites for mutation

The selection of regions and sites for mutations was based on the results obtained from promoter deletion analysis using the *in silico* online MATCH™, PATCH public 1.0 (Matys *et al*., 2006) and TFSEARCH tools (Akiyama, 1995), which use similar weight matrix algorithms that work with the TRANSFAC® binding site database (Matys *et al*., 2006) to predict potential transcription factor binding sites. The TRANSFAC® database allows the computation prediction of potential transcription factor binding sites. Based on empirical data collected from different eukaryote organisms, TRANSFAC®, similar to CENSOR and RepeatMasker described above, utilises a position-specific scoring matrix created by aligning all sites that refer to one TF to further perform a multiple local sequence alignment using the sample OTU (Fu and Weng, 2004). Before transfection, all clones were purified using the Qiagen Plasmid Midi Kit (Qiagen) for high transfection efficiency, and constructs verified for accuracy by restriction and sequencing (Sanger ABI 8730xl, GATC Biotech).

4.2.3.2.1 Site-directed mutagenesis

Site-directed mutations were performed using the QuickChange II site-directed mutagenesis kit (Stratagene) according to the manufacturer's protocol. This kit utilises oligonucleotide primers containing the desired mutation. The primers (TABLE 4.1), each complementary to opposite strands of the vector, were extended by *PfuUltra* HF DNA polymerase at high annealing temperature (72ºC). The generated amplification product consisted of a mutated circular vector containing staggered nicks at the 5' end of the amplified strand. Following temperature cycling, the product was digested with 10U of *Dnp* I endonuclease, specific for methylated DNA, for 3 h at 37ºC to digest the parental DNA template, thus selecting for the mutated vector. The nicked vector was then transformed into *Escherichia coli* competent TOP10 cells according to the manufacturer's instructions (Invitrogen), which repaired the nick as if it were a DNA polymerase error.

4.2.4 Transfection assays

For luciferase assays, FHM cells were harvested and seeded as described in *Section 4.1*. To assess the transcriptional control of Lxr, Rxr, or Srebp TFs in salmon *elovl5* gene, transcription was studied by co-transfecting FHM cells with the resulting pGL4.10-*elovl5* constructs (wild promoters, deletion or site-directed mutants) and salmon nSrebp1 (1-470 aa), nSrebp2 (1-459 aa), Lxrα (1-462 aa) and/or plaice Rxrγ (1-438 aa) expression constructs described in *Chapter 3* (Carmona-Antoñanzas *et al*., 2013a) to overexpress the protein product. A constitutively expressed control reporter construct (pGL4.75, *hRluc*/CMV; Promega) encoding *Renilla* luciferase was also used as an internal control vector to normalise variations

in transfection efficiency. Transfection mixtures consisted of 60 ng of pGL4.10-*elovl5* reporter construct (empty pGL4.10 vector in controls), 40 ng of expression vector pcDNA3 (empty pcDNA3 vector in controls), 20 ng of pGL4.75 reporter plasmid (Promega) and 1 µl of Polyfect (Qiagen) transfection reagent. Cells co-transfected with salmon Lxrα and plaice Rxrγ were further treated with the Lxr agonist GW3965 (10 µM) or ethanol carrier added 24 h after transfection. Forty eight hours after all tranfections, the medium was aspirated, monolayer washed twice with PBS, cells lysed by 10 min incubation in 75 µl per well of 1x Passive Lysis Buffer (Promega), and Firefly and *Renilla* luciferase activities were quantified as described in *Section 3.2.4.5* (FIGURE 4.1). Within each transfection experiment, each treatment was performed in triplicate. Transactivation activities were obtained using VICTOR X Multilabel Plate reader (PerkinElmer, USA) and data was normalised to the *Renilla* luciferase activities. Data are presented as means of transactivation activities of the triplicate assays.

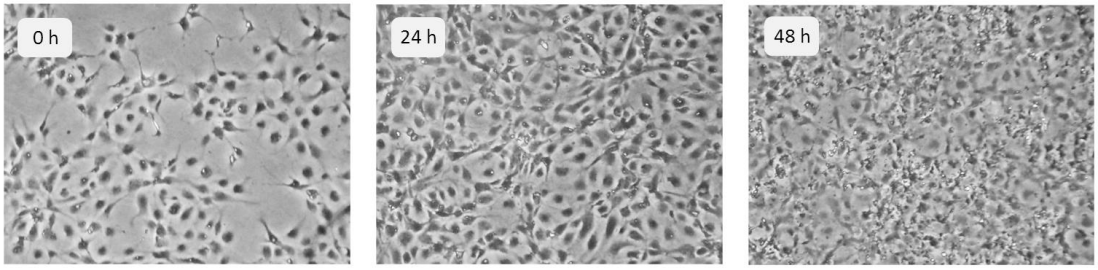

FIGURE 4.1 Fathead minnow (FHM) cell morphology during the course of luciferase assays. Cell lines were seeded at 3×10^4 cells per well in 96-well plates, to achieve 80% confluency 24 hours later when transfection was performed. After 48 hours, cells had reached ~ 90–100% confluency and were ready to be harvested and lysed to proceed with the luciferase assay.

4.2.6 Statistical analyses

All data are presented as means ± SE with three replicates per condition. The effects of deletions or mutations on promoter activities were determined by one-way analysis of variance (ANOVA), followed by multiple comparisons using Tukey's *post hoc* test. A significance of $P \leq 0.05$ (PASWS 18.0, SPSS Inc., USA) was applied to all statistical tests performed.

4.3 RESULTS

4.3.1 Organisation of *elovl5* elongase genes

4.3.1.1 Gene structure of Atlantic salmon elovl5 duplicates

It was previously reported that the full-length cDNAs for the salmon elongase duplicates contained open reading frames of 888 bp and 885 bp, corresponding to coding proteins of 295 and 294 aa, respectively (Hastings *et al*., 2004; Morais *et al*., 2009). Four lambda phage salmon genomic inserts with homologous regions to the previously described salmon elongase cDNAs were obtained after screening the library. These four phage inserts resolved into two distinct *elovl5* genes, which corresponded to the previously cloned and functionally characterised elongase cDNAs, and so the genome sequences were termed *elovl5a* [GenBank: GU238431.1] and *elovl5b* [GenBank: GU324549.1]. The complete gene sequence for *elovl5a* was obtained, but only a partial sequence corresponding to the first five introns and exons deduced from the *elovl5b* cDNA sequence [GenBank: FJ237531] was obtained. Both gene sequences were subsequently compared to and extended with data obtained

from the recently upgraded Atlantic salmon whole genome shotgun database with 3× coverage. Both salmon elongase genes had highly similar coding exon structures, being comprised of one 5'UTR non-coding exon and 7 coding exons that shared a high degree of sequence identity. Identification of two salmon elongases as separate genes was confirmed as each had distinct sizes and sequences of introns and so do not represent alleles (TABLE 4.2). The putative transcription start sites for these genes were identified according to the published cDNA sequences obtained by 5′ SMART RACE PCR (Hastings *et al*., 2004; Morais *et al*., 2009) and were used to delineate the start of untranslated exon 1. Subsequent exon/intron boundaries were identified by comparison to the cDNA sequences. The translation start sites (ATG) were located in exons 2 and termination codons in exons 8. However, exon 8 of *elovl5b* was three nucleotides shorter than that for *elovl5a* resulting in the previously reported proteins of 294 and 295 aa, respectively. The completed salmon *elovl5a*, and *elovl5b* genes spanned 19,150 bp and 12,678 bp of genomic DNA, respectively.

4.3.1.2 Northern pike *elovl5* gene

The full-length cDNA of Northern pike *elovl5* [GenBank: JX272634] was aligned with salmon gene sequences to predict exon boundaries and amplify introns 4 and 6 that were 2.3 and 0.6 kb, respectively (TABLE 4.2).

4.3.2 Fine-scale analysis of salmon *elovl5* duplicates

Fine-scale analysis of salmon *elovl5* genomic sequences was performed using computational studies that included the use of pairwise local alignments and software tools to identify the presence of conserved regions between the promoter regions of salmon *elovl5* duplicates and to identify potential mobile elements.

TABLE 4.2 Comparison of the sizes of exons and introns of Atlantic salmon and Northern pike *elovl5* orthologous genes.

	Exon (bp)			Intron (bp)		
No	salmon *elovl5a*	salmon *elovl5b*	pike *elovl5*	salmon *elovl5a*	salmon *elovl5b*	pike *elovl5*
				3618*	1307*	
1	86[a]	88[a]	72[d]	1194	1746	
2	65[b]	65[b]	65[d]	381	686	
3	188	188	188[d]	8096	2351	2376
4	78	78	78[d]	1212	719	
5	172	172	172[d]	2144	3565	672
6	125	125	125[d]	350	486	
7	135	135	135[d]	211	172	
8	1095[c]	795[c]	577[c]			

* 5′ upstream sequences.
[a] The exon is 5′UTR.
[b] The exon includes 5′UTR of 7 bp.
[c] The exons include 129, 126 and 129 coding nt.
[d] Exon size predicted from the cDNA sequence.

4.3.2.1 Analysis of *cis*-regulatory elements in salmon duplicates

Genes for *elovl5a* and *elovl5b* [GenBank: GU238431, GU324549], including 4.9 kb and 3.1 kb upstream of the initiation codons, respectively, were investigated to identify homologous regions among the two sequences, and were also compared (BlastN) to the whole-genome shotgun Atlantic salmon and Repbase databases to identify intergenic repeated sequences and transposon-like elements present elsewhere in the salmon genome. Highly conserved regions in both *elovl5* duplicates made the largest contribution in size to the *elovl5b* upstream sequence (43%) and accounted for 26% of the size of the *elovl5a* promoter with nucleotide identities ranging between 75% and 93%, including the first exon, containing the majority of the 5'UTR region, which shared 83% identity. Highly repeated elements with more than 30 hits throughout the salmon genome and > 80% identity accounted for 23% of

elovl5b promoter and constituted the largest contribution to the *elovl5a* length with 39%. The greater size of the *elovl5a* promoter appears to have originated from a 3 kb insertion scored as a highly repeated (> 85%) sequence throughout the salmon genome and, within this region, signs of ancestral transposable elements could be identified (FIGURE 4.2). RepeatMasker and CENSOR hits and NCBI megablast analysis, indicated the presence of a non-autonomous Tc1-like element, sharing 82% identity with a previously reported Tc1-like transposase pseudogene, Tcb2 [GenBank: BT059074] from Atlantic salmon. Regions that were not similar between the two *elovl5* genes and were not highly repeated in the salmon genome accounted for 35% and 34% of the total promoter size of *elovl5a* and *elovl5b*, respectively.

4.3.2.2 Distribution of transposable elements in *elovl5* homologs

The intronic regions of the salmon *elovl5* genes exhibited a divergent set of transposable elements (TE). Three non-autonomous Tc1-like DNA transposons and one non-LTR (long terminal repeat) retrotransposable element (short interspersed nuclear element, or SINE) were predicted in the salmon *elovl5a* gene, whereas in the salmon *elovl5b* only one autonomous Tc1-like DNA transposon was identified in intron 6 (FIGURE 4.2). All mobile elements identified in salmon were located in introns 1, 4 and 6 regions, and so the sequences corresponding to pike intron 4 and 6 were cloned and sequenced using specific primers designed in flanking exons. No obvious repeated or transposon-like elements were detected in the pike introns (FIGURE 4.2). However, the pike introns showed considerable homology to the regions of the corresponding salmon *elovl5* introns not matched to repeated or transposon-like sequences. All mobile elements identified in the Atlantic salmon

elongases were represented as highly repetitive sequences found across the salmon genome.

4.3.3 Promoter deletion analysis

To determine the regions of each salmon *elovl5* gene responsible for TF-dependent transcription, the DNA fragment encompassing upstream, untranscribed exon and initiation codon sequence, and deletions thereof, was fused to a promoterless luciferase reporter gene [pGL4.10] and co-transfected with expression vectors containing the factors of interest. Results were normalised using the *Renilla* luciferase [pGL4.75] and negative controls transfected with promoterless pGL4.10 and mock expression vector [pcDNA3] lacking the insert.

4.3.3.1 Regulation of salmon *elovl5* duplicates by Lxr and Rxr

Deletion constructs incubated with ligand-activated Lxr and Rxr showed significant differences between *elovl5a* and *elovl5b* responses, supporting the results obtained in *Chapter 3*. Lxr/Rxr and GW3965-dependent stimulation of *elovl5a* promoter activity was evident when up to 153 nt upstream of the TSS were included in the construct (SEA5). Upstream deletion constructs SEA1, SEA3 and SEA4 also showed significant expression upregulation differences, with maximal activity observed when up to 1121 nt upstream of the TSS were included in the construct. In contrast, all deletion constructs excluding the promoter region upstream + 9 of the TSS (SEA7, SEA8 and SEA9) exhibited much reduced activity and no significant differences between the expression levels of the GW3965-activated TF and the mock activated TF with ethanol (FIGURE 4.3). These results suggested that *elovl5a* likely possessed an LXRE within a fragment of 162 nt, located between 153 nt upstream of

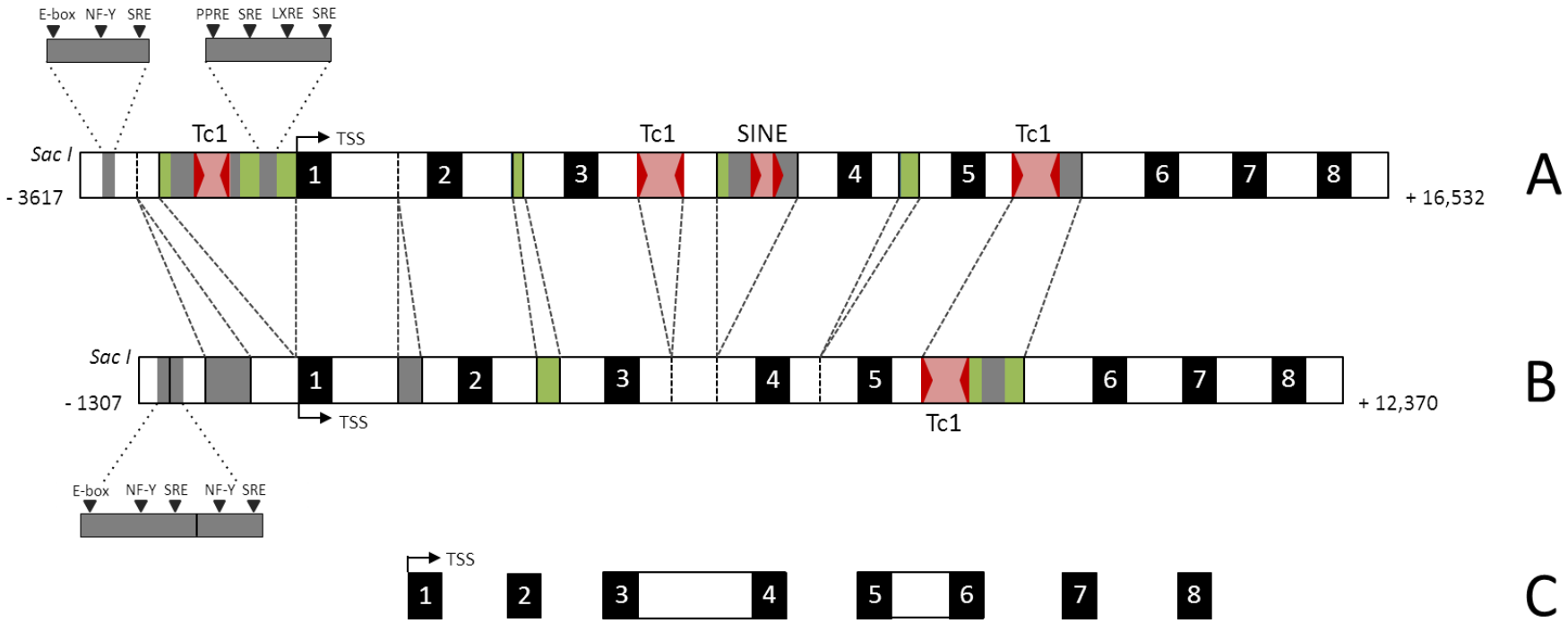

FIGURE 4.2 Gene structures of Atlantic salmon A) *elovl5a* and B) *elovl5b* and C) Northern pike *elovl5* homologs. Gene structure was determined from genomic sequences obtained from a genomic DNA lambda FIX II library and complemented with sequences from the NCBI Atlantic salmon whole genome shotgun database and PCR cloning. Coding exons (TABLE 4.2) predicted from alignments with cDNAs are indicated by black boxes connected by intergenic sequences (not to scale) classified as conserved (unfilled boxes), non-conserved (green boxes) or highly repeated (grey boxes) regions among the *elovl5* orthologs. Sequence regions showing high levels of identify across the Atlantic salmon promoters were identified using MegaBlast and the corresponding regions highlighted (open boxes). The position of predicted non-autonomous Tc1-like transposons and non-LRT (long terminal repeat) retrotransposons from the short-interspersed nuclear element (SINE) family that share extensive similarity to RepBase entries from a reference collection of interspersed elements are indicated by light red boxes flanked by inverted or direct short repeats depicted by maroon arrowheads (Wessler, 2006). The results were screened for hits with > 70% identity values and alignment scores > 800. Srebp (SRE), Lxr (LXRE) and other TF binding sites predicted from mutagenesis assays are indicated. *Sac*I indicate the endogenous restriction sites used to clone the DNA libraries into sequencing vectors. TSS, transcriptional start site.

the TSS and 9 nt downstream of the TSS. In contrast, the *elovl5b* promoter was not stimulated by Lxr/Rxr and GW3965 and a significant downregulation, not detected in *elovl5a*, was observed at deletions beyond +743 (FIGURE 4.3).

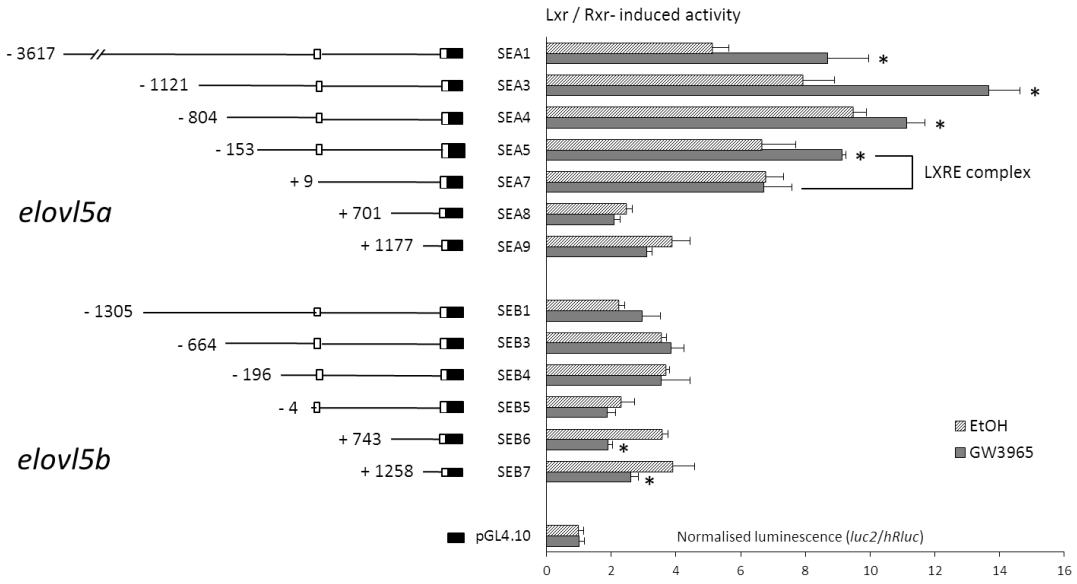

FIGURE 4.3 Deletion analyses of salmon *elovl5a* and *elovl5b* gene promoters in the presence of overexpressed salmon Lxrα and plaice Rxrγ transcription factors. Deletion constructs are represented on the left. Non-coding exon is indicated with open boxes and *Firefly* luciferase coding-sequence by closed boxes. Sequences are numbered relative to the first base of the transcription start site, assumed to be the first base of the 5' non-coding exon. Promoter activity of constructs is represented on the right with the values representing normalised Firefly activity (*luc2*) to Renilla activity (*hRluc*). *Indicate the effect of GW3965-activated Lxr/Rxr is significant compared to ethanol (EtOH) carrier on the same construct (One-way ANOVA; $P < 0.05$). The results are representative of three independent experiments.

4.3.3.2 Regulation of salmon *elovl5* duplicates by Srebps

Both *elovl5* deletion constructs were also tested for response to Srebp1 and Srebp2 overexpression. Maximal Srebp1 and Srebp2 dependent stimulation of luciferase was observed on the largest *elovl5a* (SEA1) and *elovl5b* (SEB1) promoter constructs. The *elovl5a* promoter showed 3-fold and 2-fold increases stimulated by Srebp2 and Srebp1, respectively, when up to 3617 nucleotides upstream of the transcription start site (TSS) were included in the wild-type reporter construct

(SEA1) (FIGURE 4.4). A deletion with 3538 nucleotides upstream of the TSS (SEA2) abolished the Srebp1 and Srebp2-dependent stimulation. The *elovl5b* promoter exhibited the highest activity in the presence of Srebp2 (5-fold induction), 3-fold higher than that of Srebp1 and twice as high as the maximum Srebp2-induced activity observed for *elovl5a* (SEA1) (FIGURE 4.4) confirming the results obtained in *Chapter 3*. A deletion, less than 80 nucleotides shorter, including 1226 nucleotides upstream of the TSS (SEB2) displayed around 50% reduction in Srebp2 dependent activity, and the Srebp1 dependent effect disappeared with respect to activity in the absence of Srebp. No significant differences in activity were observed in further *elovl5b* promoter deletions. The promoterless pGL4.10 (negative control) exhibited a much-reduced activity in all analyses. These results indicated that Srebp1 and Srebp2 are capable of binding *elovl5a* promoter within a 79 nt region located between 3617 nt (SEA1) and 3538 nt (SEA2) upstream of the TSS. The salmon *elovl5b* promoter seemed to respond to Srebp1 within a fragment of 79 nt between 1305 nt (SEB1) and 1226 nt (SEB2) upstream of the TSS, whereas an SRE responsible solely for Srebp2 regulation appeared to be located within a larger region of 641 nt located between 1305 nt (SEB1) and 664 nt (SEB2) upstream of the TSS.

4.3.4 Promoter mutagenesis analysis

Candidate transcriptional regulatory regions in the salmon *elovl5* duplicates identified by promoter deletions as containing potential Lxr response elements (LXRE) (FIGURE 4.5) or Srebp response elements (SRE) (FIGURE 4.6) were subjected to *in silico* analysis (TABLE 4.3) to identify potential TF binding elements. Then, regulatory activity was investigated in FHM cells transfected with mutants targeting potential response elements.

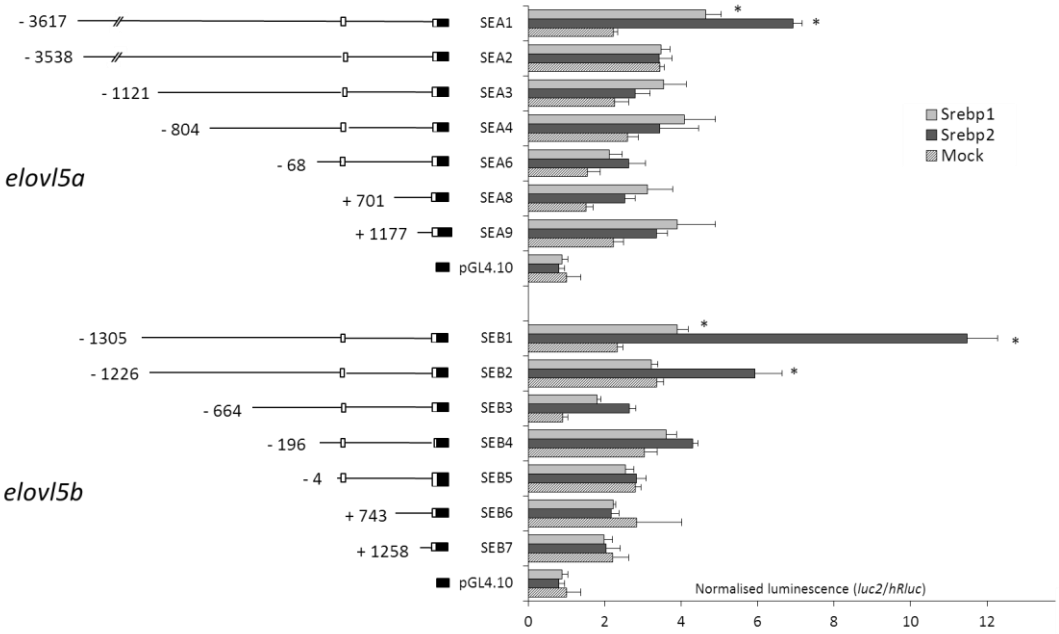

FIGURE 4.4 Deletion analyses of salmon *elovl5a* and *elovl5b* gene promoters in the presence of overexpressed nuclear Srebp factors. Deletion constructs are represented on the left. Non-coding exon is indicated with open boxes and *Firefly* luciferase coding-sequence by closed boxes. Sequences are numbered relative to the first base of the transcription start site, assumed to be the first base of the 5' non-coding exon. Promoter activity of constructs is represented on the right with the values representing normalised Firefly activity (*luc2*) to Renilla activity (*hRluc*). *Indicate the effect of Srebp1 or Srebp2 is significant compared to a construct containing no insert (mock) co-transfected with on the same reporter construct (ANOVA; $P < 0.05$). The results are representative of three independent experiments.

Salmon *elovl5a* promoter

```
                                      aa       t     a
-160  GACTGAAGTCATAACTTCATGACCTTTGATCACATGGT  -123
                                  SITE 3       SITE 4

                 a  a              a    a
      TTTTGTCAAGTTCATGCAGTCGTCATCTGTAGTCGGAA  -85
                SITE 5            SITE 6

      ATCGGAACATTTTTTATTCCCACAATGCAACACACTTC  -47
      GCAAGGCGTTGACCGGACAAAAGTGATCTGCTCGACGC  -9
      GCCCTGTGTTTCTCCC                        +8
```

FIGURE 4.5 Intergenic nucleotide sequence of the salmon *elovl5a* promoter recognised by Lxr/Rxr. Numbers are given relative to the first base of the transcription start site (TSS). Exonic region corresponding to the 5' non-coding region is shown in bold letters (+1 to +8). Potential transcriptional binding motifs investigated by mutational analyses are indicated by SITE 3-6, with nucleotide mutations indicated by lower-case letters above the wild-type sequences.

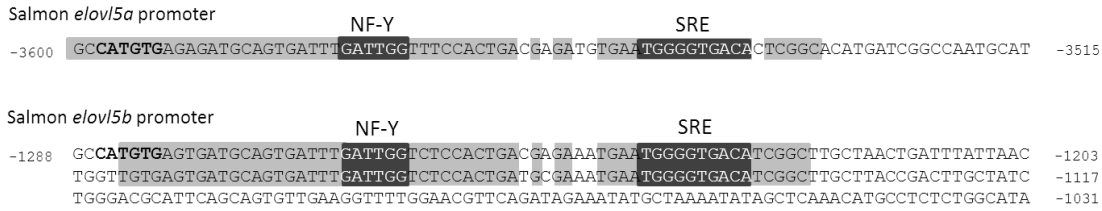

FIGURE 4.6 Alignment of the salmon *elovl5* duplicated *cis*-regulatory regions. The numbers indicate the sequences position relative to the transcriptional start sites based on gene sequence information for *elovl5a* and *elovl5b*. Shading indicates that nucleotides at that position are identical. E-box is indicated in bold and NF-Y and SRE binding sites contained in closed boxes. NF-Y, nuclear factor Y binding sites; SRE, sterol regulatory element; E-box, enhancer box.

4.3.4.1 Identification of Lxr response elements (LXRE) in *elovl5a*

For salmon *elovl5a*, mutations at sites 3, 5 and 6 caused a reduction of promoter activity of 16%, 22% and 25% relative to the wild type activity (SEA1), respectively, although those were not statistically significant (FIGURE 4.7). None of the site mutations had a significant effect on promoter activity compared to the wild type activity.

TABLE 4.3 Mutated promoter sequences of Atlantic salmon *elovl5* duplicated genes corresponding to transcription factor binding sites.

Gene	Construct name	Mutation sites	Binding factor	Sequence	Mutated sequence	Position from TSS
elovl5a	SEA1 M1	1	NF-Y	GATTGGT	GAccaaT	-3576
	SEA1 M2	2	SREBP	TGGGGTGACA	ATaaGaaGAC	-3549
	SEA4 M1	3	PPARα	TGACCT	TGAaaT	-141
	SEA4 M2	4	SREBP1	CACATG	tACATa	-130
	SEA4 M3	5	LXRα	AGTTCA	AaTTaA	-114
	SEA4 M4	6	SREBP1	CATCTG	aATCTa	-99
elovl5b	SEB1 M1	1	NF-Y	GATTGGT	GAccaaT	-1264
	SEB1 M2	3	NF-Y	GATTGGT	GAccaaT	-1178
	SEB1 M3	1, 3				
	SEB1 M4	2	SREBP	TGGGGTGACA	ATaaGaaGAC	-1237
	SEB1 M5	4	SREBP	TGGGGTGACA	ATaaGaaGAC	-1151
	SEB1 M6	2, 4				

Transcriptional binding sites were identified in *elovl5a* [GU238431.1] and *elovl5b* [GU324549.1] promoter regions using the TRANSFAC® binding site database (Matys *et al*., 2006). TSS, transcription start site.

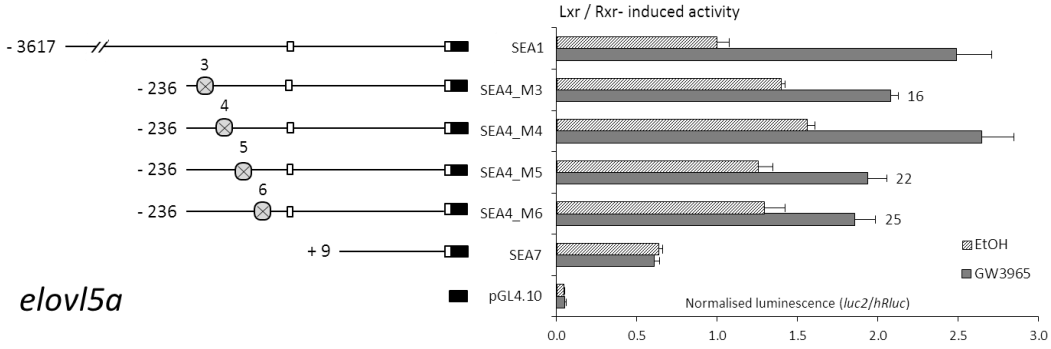

FIGURE 4.7 Mutation analyses of salmon *elovl5a* gene promoter in the presence of overexpressed salmon Lxrα and plaice Rxrγ transcription factors. Various mutations of the salmon promoter deletions SEA4 (-236, FIGURE 4.3) were generated according to TABLE 4.3 and the resulting reporter gene activity expressed as the normalised Firefly luminescence (*luc2* / *hRluc*). Relative positions of binding sites, as described in the text and in TABLE 4.3, are shown and mutation sites within each construct indicated by a closed crossed box. Values represent the percentage expression reduction with respect of the corresponding non-mutated (wild type) promoter of one independent experiment.

4.3.4.2 Identification of Srebp response elements (SRE) in *elovl5* duplicates

For salmon elovl5a, mutations at sites identified by *in silico* analysis as potential NF-Y transcription factor binding and SREs caused significant reductions of promoter activity of 36% and 55%, respectively, compared to activity of Srebp1 on the wild type sequence (FIGURE 4.8). No effect was observed for Srebp2. In *elovl5b* two NF-Y sites (sites 1 and 3) and two SRE sites (sites 2 and 4) were identified and mutated. For Srebp1-dependent activity on *elovl5b*, independent mutation of NF-Y at site 1 or 3 caused non-significant reductions of 43% or 53% of the wild type activity (SEB1), respectively, whereas the simultaneous mutation of both NF-Y sites (SEB1 M3) caused a significant reduction of 76% compared to wild type activity. SRE mutation at site 2 or 4 caused significant reductions of 86% or 83%, and the simultaneous mutations in sites 2 and 4 reduced the wild type activity by 90%. Similar results were observed for *elovl5b* promoter mutations when they were co-expressed with Srebp2. Mutations in sites 1, 3, 2 and 4 caused reductions of

36%, 80%, 92% and 92% respectively whereas the co-mutation of NF-Y sites 1 and 3, or SRE sites 2 and 4 simultaneously caused the activities to drop by 90% and 95% compared to the wild type activity.

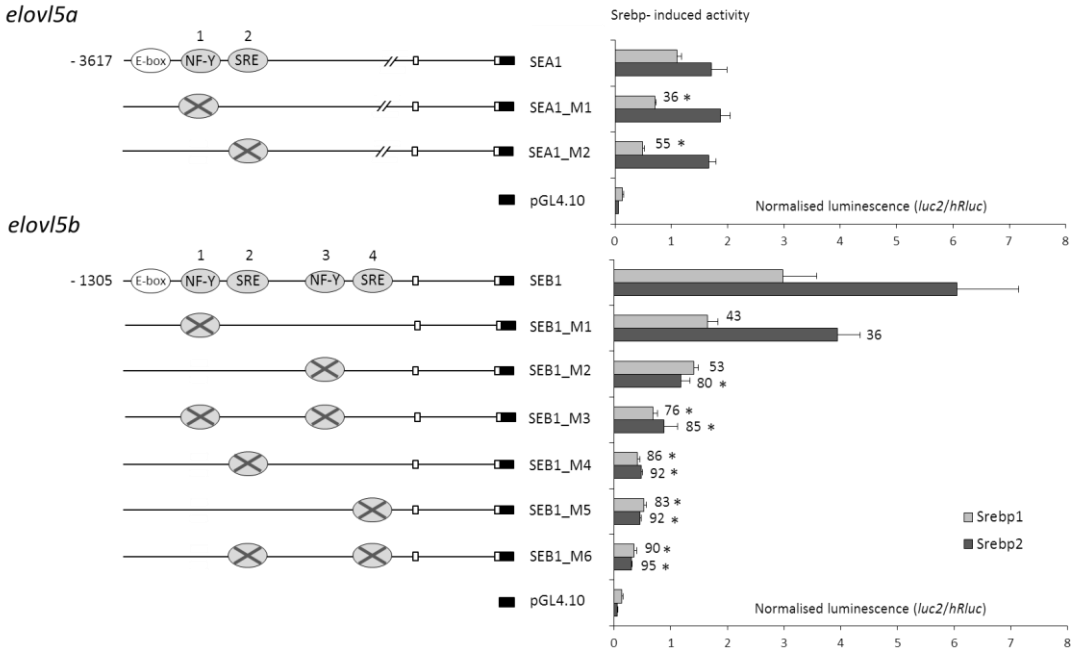

FIGURE 4.8 Mutation analyses of salmon *elovl5a* and *elovl5b* gene promoters in the presence of overexpressed nuclear Srebp factors. Various mutations of the salmon promoter deletions SEA1 (-3617, FIGURE 4.3) and SEB1 (-1305, FIGURE 4.3) were generated according to TABLE 4.3 and the resulting reporter gene activity expressed as the normalised Firefly luminescence (*luc2* / *hRluc*) relative to an expression construct containing no insert. Relative positions of E-box, NF-Y and SRE sites, as described in text and in TABLE 4.3, are shown and mutation sites within each construct indicated by a closed crossed box. Values represent the percentage expression reduction with respect of the corresponding non-mutated (wild type) promoter of one independent experiment. *Indicate the effect of Srebp1 or Srebp2 is significant compared to the non-mutated wild-type construct (One-way ANOVA; $P < 0.05$).

4.4 DISCUSSION

This study described the sequence, structure and regulation of duplicated Atlantic salmon *elovl5* genes and aimed to infer the evolutionary history and explain

the sub- or neofunctionalisation of these genes by comparative analysis of their regulation, and also comparison to the corresponding pre-duplicated *elovl5* gene in pike. Salmon and pike are useful species to study these aspects of evolution since Atlantic salmon and all other salmonids are descended from an ancestor that underwent whole genome duplication (WGD) event shortly after salmonids and esocids shared a common ancestor (Ishiguro *et al*., 2003; Koop and Davidson, 2008; Leong *et al*., 2010). Since WGD is hypothesised to play a major role in evolution and speciation (yeast, Maclean and Greig, 2011), providing a larger complement of ready-made genes on which selection can act, the study of the consequences of WGD might shed light on the processes of diploidisation (Lönnig and Saedler, 2002) and of sub- or neofunctionalisation. Sequence analysis clearly showed that the salmon *elovl5* genes had each been colonised by a distinct range of transposon-like elements, at distinct sites. These transposon-like elements are highly repeated in the salmon genome. The limited comparison with pike further indicated that *elovl5* introns in this species have not been significantly colonised by transposons, and that these pike introns can be aligned with the non-repeated, non-transposon-like regions of the corresponding exons in both salmon *elovl5* genes. The asymmetrical distribution pattern of Tc1-like transposons and non-LTR retrotransposons in paralogous *elovl5* genes clearly suggested that Atlantic salmon *elovl5* genes likely gained most TEs after the species diverged from a common ancestor with the closest pre-duplicated extant relative, Northern pike. Furthermore the salmon genes must have acquired transposable elements after the *elovl5* duplication event.

Recent studies have suggested that the bulk of the genome in higher vertebrates is ultimately derived from transposable elements (TEs) and their

remnants (Brosius, 1999, 2003; Smit, 1999) as they are highly dynamic elements capable of making duplicated copies of themselves, colonising the genome, mostly in intergenic and intronic regions due to relaxed functional constraints. Although mechanisms exist for silencing the activity of such elements, they are still capable of promoting illegitimate recombinations and chromosome rearrangements (Brosius, 2003; Fedoroff, 2012), although they have also been suggested to promote re-diploidisation after WGD through impaired multivalent formation during meiosis (de Boer *et al*., 2007; Parisod *et al*., 2009). Furthermore, it has become clear that TEs can contribute promoters and regulatory elements to existing genes (Bejerano *et al*., 2006; Makałowski, 2000; Muotri *et al*., 2007; Nishihara *et al*., 2006) and influence the regulatory divergence of duplicated genes (Herpin *et al*., 2010). Genome-scale bioinformatics analyses have shown that many promoters are derived from specific TEs postulating that insertion of TE harbouring "ready-to-use" *cis*-regulatory sequences probably contributed to the establishment of specific patterns of gene expression (Ferrigno *et al*., 2001; Mariño-Ramírez and Jordan, 2006; Mariño-Ramírez *et al*., 2005).

The salmon *elovl5* genes are present in highly syntenic regions on separate chromosomes (Carmona-Antoñanzas *et al*., 2013b) indicating, although not conclusively, that they have arisen as a result of the salmonid whole genome duplication event timed at about 25-100 myr ago. In plants, physiological and genetic stresses, such as polyploidisation events, can "boost" transpositional expansion of silent transposons within genomes, possibly caused by disruption of DNA methylation (Capy *et al*., 2000; Fedoroff, 2012; Yaakov and Kashkush, 2011). Accelerated mobilisation of TE was detected following the salmonid whole-genome

duplication, hypothesised to play an important role in the genome re-diploidisation process (Bento *et al*., 2013; Huminiecki and Heldin, 2010; Parisod *et al*., 2009; Wolfe, 2001) and salmonid speciation (de Boer *et al*., 2007; Macqueen and Johnston, 2014). Our results indicated that mobilisation of transposable elements occurred after the *elovl5* paralogs were duplicated in Atlantic salmon based on comparative studies with the closest preduplicated genome. Increased amounts of mobile element insertions are believed to drive non-homologous recombination during meiosis causing insertions, deletions and inversions that, in the long term, might impair multivalent formation (Gray, 2000; Lönnig and Saedler, 2002). Thus, elevated transposon activity following polyploidisation might play a role in driving diploidisation, which is required to enable genes duplicated via a WGD to diverge under relaxed selective pressures on at least one of the duplicates.

To define the mechanisms involved in the differential regulation of salmon *elovl5* duplicated genes, luciferase-reporter gene assays were performed in FHM cells to compare the transactivities of nuclear Srebps and Lxr/Rxr heterodimers on a battery of promoter deletions and mutations. The results showed that Srebp1 and Srebp2 were directly involved in the stimulation of both *elovl5* duplicated promoters, which contained inverted SRE-like sequences in close proximity with NF-Y cofactor sites and an E-box site, recognised as an insulin or carbohydrate response element (Wang and Sul, 1997). In general, Srebp2 stimulated higher response than Srebp1, whereas in mammals the response of SREBP2 for lipogenic genes is significantly lower than that of SREBP1 (Amemiya-Kudo *et al*., 2002). Yet, the most striking observation was the differential response of the wild type *elovl5* promoters. As demonstrated by sequence analysis, *elovl5b* possesses adjacent regions with

duplicated dyads of SRE and NF-Y binding sites that confer greater Srebp-dependant stimulation than the single dyad in *elovl5a*. It was also apparent that both intact NF-Y and Srebp sites were required for maximal Srebp-dependent activity. Furthermore on these particular sites, Srebp2 appeared to promote more transcriptional activity than Srebp1. It is, however, difficult to explain why mutation of either the NF-Y-binding site or the SRE in *elovl5a* had no effect on Srebp-dependent activity. The same mutations reduced Srebp1-dependent activity and also both Srebp1- and Srebp2-dependent activity on the *elovl5b* promoter. The results are complicated to some extent in that a further downstream site must exist in *elovl5a*, which is only responsive to Srebp2. Unfortunately this site was not characterised in detail. Nevertheless, overall the transactivation results showed that these *elovl5* promoters were regulated by Srebps through SREs, likely in cooperation with NF-Y. In mammals, Srebps also require recruitment of NF-Y cofactors to sites adjacent to some SREs (Jackson *et al*., 1995, 1998; Näär *et al*., 1998) for maximal activity. The conserved sequence in *elovl5a* and *elovl5b* also contained an E-box. However, although SREBPs have been shown to efficiently bind E-boxes (Atchley and Fitch, 1997; Kumar *et al*., 2003), their subsequent ability to activate the downstream gene is negligible for mammalian SREBPs (Amemiya-Kudo *et al*., 2000, 2002). Thus, it is unlikely that E-boxes are responsible for the differential Srebp2-dependent stimulation of the salmon duplicates, which could be possibly attributed to the presence of an undefined factor in *elovl5a* promoter, or different Srebp binding efficiencies masked by the presence of duplicated SRE and NF-Y recognition sites in the promoter of *elovl5b*. In all, the results indicated that Srebps overexpression drove

different responses between *elovl5* duplicate promoters most likely caused by asymmetrical divergence in *cis*-regulatory regions.

Supporting this argument, the deletion and mutation constructs allowed us to identify a region located few hundred base pairs upstream of the transcriptional start site that was stimulated by Lxr agonist GW3965 in the presence of Lxr and obligated heterodimer Rxr (Carmona-Antoñanzas *et al*., 2013a). More important was the fact that such a response was only detected on one of the duplicates, *elovl5a*, when a fragment identified as a non-conserved homologous region in close proximity to a nonautonomous Tc1-like element was included in the reporter construct. In all, these results presented evidence of functional divergence of regulatory *cis*-elements after gene duplication resulting in a novel function not previously described in vertebrates (Wang *et al*., 2005, 2006; Yoshikawa *et al*., 2002). From a nutritional perspective, the identification of RE in lipid targets responsible for the metabolic response of fish to low-sterol and LC-PUFA diets, could be exploited with the intention to elucidate the regulatory mechanisms of LC-PUFA biosynthesis from an evolutionary point of view, and as a potential asset to optimise and efficiently use sustainable alternative sources to marine products in aquaculture feeds. Sequence analysis allowed this regulatory region to be traced back to a non-conserved sequence likely inserted into the upstream promoter shortly after the duplication event. Normally, *cis*-regulatory elements retain functional motifs across divergent homologous genes with or without identifiable sequence similarity (Barrière *et al*., 2011; Hare *et al*., 2008; Romano and Wray, 2003). However, following gene duplication, relaxed purifying selection has been suggested to allow rapid changes in the DNA sequences (Castillo-Davis *et al*., 2004) that might result in gene neo- or subfunctionalisation events (Herpin *et al*.,

2010; Zhang, 2003). Functional constrains seem to contain the extent of change in protein-coding genes to prevent deleterious functional transformation, although intergenic regions achieve major sequence changes through insertion of transposable-derived elements (Brosius, 2003; Fedoroff, 2012; Nekrutenko and Li, 2001; Pray, 2008; Wang *et al*., 2013).

The fine-scale study of duplicated gene structure and organisation, especially in the context of whole genome duplication, enables the reconstruction of the molecular events allowing initially redundant duplicated genes to diverge from one another under relaxed purifying selection. Such studies are fundamental to understanding the processes that provide the necessary variation for speciation by natural selection.

4.4.1 Conclusions

The study of the gene structure of duplicated *elovl5* genes in Atlantic salmon identified signs of increased transposition following the recent evolutionary divergence with the esocid sister-lineage. Deduced interspersed sequences were scattered among the intronic regions of *elovl5* paralogous, although distribution patterns suggested fastest divergence of *elovl5a*, possibly attributed to weaker purifying selection in agreement with data for amino acid-coding sequences (Carmona-Antoñanzas *et al*., 2013b). Detailed sequence analysis of the target gene promoters presented evidence of asymmetrical divergence of *cis*-regulatory regions in *elovl5a* and *elovl5b*, which resulted in different transactivation responses to transcription factors, Lxr and Srebp, involved in the regulation of lipid homeostasis. Although we detected shared motifs with high local similarity and signs of ancestral

homology, fine-scale sequence analyses enabled the delineation of regions of non-conservation associated with transposon insertion as responsible for gene regulatory differences observed between duplicates.

Chapter 5

General Discussion

The present studies have investigated the molecular and evolutionary divergence of recently duplicated Atlantic salmon genes involved in the elongation of n-3 LC-PUFA, essential dietary nutrients and components of all vertebrates. Functional analyses combined with computational studies of gene families were used to reveal signs of asymmetrical divergence between the paralogous *elovl5* genes. An overall discussion on the main outcomes, limitations and future prospects is reviewed as follows:

5.1 CONCLUDING REMARKS

Evolutionists have directed intense efforts towards discovering the mechanisms required for optimal survival of a population at a certain time and in a certain space. Such mechanisms, according to neo-Darwinian theory, must involve heritable changes in genomes. Extant ray-finned fish have been hypothesised to have arisen through three rounds (3R) of ancient whole-genome duplications (WGDs), the first at the base of the vertebrate lineage, and the third at the base of the teleosts (Hashimoto *et al*., 2008). This has probably been the mechanism by which many different sub-families of genes, including *elovls*, have arisen and diversified from preduplicated ancestral genes. Although the evidence for ancient genome duplications is widely accepted, subsequent gene loss, chromosomal rearrangements

and further intra-chromosomal duplications, as well genetic drift and the ubiquitous effect of negative (or purifying) selection in coding genes and other functionally important regions (Hughes, 2007; Kimura, 1983) have made it very difficult to track the diversification of individual genes over long evolutionary timescales. Without detailed understanding of how gene duplicates, particularly those derived from WGD, diversify and sub- or neofunctionalise in the context of processes of diploidisation and selection, the idea that WGD leads to innovation and speciation remains no more than a hypothesis. However, a few vertebrates have undergone more recent WGD, including an ancestor of all extant salmonids. Thus, Salmonids, with one of the most recent tetraploid (now pseudotetraploid) genomes within vertebrates (Allendorf and Thorgaard, 1984; Koop and Davidson, 2008), are an ideal model lineage to answer questions such as how do newly duplicated genes survive and acquire novel functions, and what role does gene duplication play in the evolution of genomes and organisms? Accordingly, Atlantic salmon have been shown to possess many duplicated genes compared to the nearest living preduplicated relative, the pike (Ishiguro *et al*., 2003; Leong *et al*., 2010), and *elovl5a* and *elovl5b* are most likely examples of genes duplicated by WGD and retained in salmonids (Leonard *et al*., 2000, 2004; Morais *et al*., 2009). Generally, fully functional duplicates can be effective buffers against deleterious mutations that could otherwise have detrimental effects (Jakobsson *et al*., 2006; Moon *et al*., 2009; Stryke *et al*., 2003; Zadravec *et al*., 2011) but, most importantly, duplicated genes might constitute an ‘excess’ source of genetic material for evolution to act upon.

Following a gene duplication event, purifying selection might be expected to greatly relax on at least one of the duplicates, allowing one copy to accumulate

formerly "forbidden" mutations and evolve asymmetrically. This process is hypothesised to occasionally result in a gene locus with a novel and beneficial function (neofunctionisation), but is more likely to result in nonfunctionalisation, pseudogenisation and eventually loss (Ohno, 1970; Zhang, 2003). Initial analysis showed signs of purifying selection evident in both *elovl5* duplicates and, in stable environments, purifying selection is regarded the major evolutionary mechanism homogenising and maintaining the genetic features adapted for survival. These results indicated that purifying selection was the major force driving the evolution of *elovl5* orthologs at the protein level, not only in Atlantic salmon paralogs, but in all vertebrates studied. Thus, the Atlantic salmon fatty-acyl elongase duplicates have been retained in the salmon genome under strong functional constraints, supported by heterologous functional characterisation assays that presented no evidence of functional diversification (Morais *et al*., 2009).

However other pressures have also been suggested to maintain paralogous genes, for example they might be retained by concerted evolution when higher (extra) amounts of product are beneficial (e.g. rRNA and histones) (Zhang, 2003). Furthermore, significant changes at the phenotypic level are not necessarily caused by changes in protein sequences, but by changes in the patterns of gene expression contolled by *cis*-regulatory elements located upstream the coding sequences, such as the loss of eyes in cavefish (Jeffery, 2005) and the changes in bill morphology in Darwin's finches (Abzhanov *et al*., 2004). In fact, previous studies have shown that *elovl5a* and *elovl5b* exhibit different tissue expression patterns and dietary regulation (Leaver *et al*., 2008b; Morais *et al*., 2009). Hence, analysis of the *cis*-regulatory regions of these duplicates seemed appropriate. Comparative sequence analysis

showed regions upstream of the putative transcriptional start sites of these *elovl5* duplicates had indeed diverged from one another with sequence insertions/deletions and apparent colonisation by transposable elements.

As explained above, after a duplication event, according to theory, relaxed purifying selection should allow accelerated evolution of duplicated genes, permitting previously maladaptive changes to occur at both the protein and regulatory levels (Castillo-Davis *et al*., 2004; Hughes *et al*., 2000). However, whereas changes that affect proteins might have a detrimental effect on the activity of the elongase, most changes in promoters might not inflict an impact on the organism physiology, thus allowing some divergence in expression pattern. Changes in functional motifs, for example transcription factor response elements, are normally subjected to stronger functional constrains, as the high conservation of these sites across distant vertebrate species reveals (Cheffers *et al*., 1997). Similarly, the results presented in this thesis indicated that the transcriptional regulatory networks of lipid metabolic genes are, in general, highly conserved among vertebrate species.

Asymmetrical divergence of duplicated genes not only affected length and structure of promoter regions, but also modified and created new binding sites of transcription factors involved in lipid homeostasis. We identified that a region containing response elements for Srebp and cofactors had been duplicated in tandem, which induced a greater response of *elovl5b* to Srebps. Most importantly, an Lxr/Rxr response element was identified in the salmon *elovl5a* elongase, located within a non-conserved region. Identifying the ancestral function is key to recognise neofunctionalisation, as the origin of a new function not present in the ancestor, or a subfunctionalisation outcome, regarded as the partition of ancestral functions,

normally used to exemplify division of gene expression after duplication (Zhang, 2003). A clear example of a subfunctionalisation is that of *engrailed-1* and *engrailed-1b* paralogs in zebrafish, expressed in the pectoral appendage bud and the hindbrain/spinal cord, respectively, whereas the murine ortholog is expressed in both sites (Force *et al*., 1999). In vertebrates, *Elovl5* orthologs are regulated by SREBPs although no evidence of LXR-dependent response has been identified in any species other than Atlantic salmon. Hence, our results indicate that the presence of an Lxr/Rxr binding site in one of the Atlantic salmon duplicates constitutes a novel regulatory mechanism not described previously in any other vertebrate or teleost ortholog. The new response element might have resulted from the novel acquisition in the salmon *elovl5a* promoter (neofunctionalisation) after the gene duplication process or through loss of function in the other copy. The study of promoters in closely related species has the potential to clarify the origin of the Lxr/Rxr binding site in the Atlantic salmon *elovl5a*.

Neofunctionalisation might incur a high detrimental physiological cost if it results in the expression of a gene in an inappropriate cell, or developmental stage (Castillo-Davis *et al*., 2004; Wray, 2007). Thus, only beneficial functions are likely to be maintained in the genome. At the molecular level, salmon *elovl5* paralogs are directly controlled by Lxr and Srebp TFs and indirectly by Ppars in an Lxr-dependent manner, thus creating a complex regulatory network. Fatty-acyl elongase tissue expression (Morais *et al*., 2009) mirrors the expression patterns of the *lxr* and *srebp*, which are mostly expressed in Atlantic salmon liver, pyloric caeca (intestine) and brain under control conditions (Cruz-Garcia *et al*., 2009; Leaver *et al*., 2007; Minghetti *et al*., 2011). Identification of loss of function, neo- or subfunctionalisation

processes is especially complicated when studying continuous traits, such as tissue expression levels. The expression profile of Northern pike *elovl5*, predominantly transcribed in brain, might be caused by lack of promoter elements that enable expression in liver and intestine, or through a novel element that upregulates the expression in brain.

SREBPs and LXRs or PPARs can also be regulated post-transcriptionally by dietary sterols and fatty acids. Unlike FO diets, VO diets lack cholesterol, repressor of SREBPs activation and thus LC-PUFA biosynthesis, and LC-PUFA the major agonists and antagonists of PPAR and LXR, respectively. Thus, when Atlantic salmon smolts were fed vegetable-derived diets the expression of *elovl5b* was significantly increased in comparison to *elovl5a* (Leaver *et al*., 2008b), as would be expected as a result of duplicated SRE binding sites. Therefore, it is likely that the expression differences previously observed between the fatty-acyl elongase duplicates in Atlantic salmon are partly caused by the sequence differences in the *cis*-regulatory regions identified as response elements recognised by key transcription factors involved in lipid homeostasis. In addition, two elongase paralogs have been recently identified in the rainbow trout genome assembly draft and transcriptome (Palti *et al*., 2012). The analysis of the trout *elovl5* promoters might elucidate the evolution of structures and regulation of duplicated Atlantic salmon *elovl5*s.

Previously, in studying LC-PUFA biosynthesis, greater importance has generally been given to fatty-acyl desaturases, which have well defined hormonal and nutritional responses in mammals (Wang *et al*., 2005, 2006). However, the results of the studies described in this thesis suggest that salmon fatty-acyl elongases

are similarly important in the biosynthesis of LC-PUFA, and both sterols and fatty acids might play an important role in the endogenous production of LC-PUFA in Atlantic salmon, acting through conserved vertebrate transcriptional machinery. This suggested that it could be possible to manipulate the dietary supply of these products to stimulate the biosynthesis of n-3 LC-PUFA. Until recently, interest has focused on the amount of specific fatty acids in fish diets, which were considered the main regulator of LC-PUFA biosynthesis. If, however, both PUFAs and LC-PUFAs have an antagonistic effect on the activation of salmon Lxr, similar to mammals and shown in this thesis, then the downregulation of fatty-acyl enzymes in fish fed diets rich in EPA and DHA cannot be totally attributed to the presence of LC-PUFA (Ou *et al*., 2001; Yoshikawa *et al*., 2002). Hence, our results raise a question regarding what the role and ideal amount of cholesterol is in aquafeeds. If Srebp, inactivated in sterol-replete cells, are direct transcriptional regulators of LC-PUFA biosynthetic gene expression, thus excess dietary cholesterol could actually decrease fatty-acyl elongase transcription resulting in reduced EPA and DHA biosynthesis. The overarching idea is that there is a potential for greater biosynthesis of n-3 fatty acids through optimisation of feed composition for fish.

On another note, duplicated genes can also be maintained under functional constraints when extra amounts of product are beneficial. One of the drawbacks of duplication is the potential imbalance between products within the same metabolic network (stoichiometry) (Huminiecki and Heldin, 2010; Levasseur and Pontarotti, 2011). Atlantic salmon possess four fatty-acyl desaturases, responsible for the limiting step in LC-PUFA desaturation, whereas most vertebrates present only one member. Accordingly, the extra Elovl5 product might be beneficial to compensate

for gene dosage effects in fatty-acid biosynthetic tissues such as liver and intestine allowing a faster and possibly more efficient production of EPA and DHA.

A WGD preceded Salmonid diversification and speciation possibly enabling them to evolve new physiological features, which have previously been related to acquisition of an anadromous life history (Allendorf and Utter, 1976; Ishiguro *et al*., 2003), enabling exploitation of the more productive marine environment (Gross *et al*., 1988; Ramsden *et al*., 2003). Conversely, the results presented here suggest that the acquisition of a complete and more diverse set of fatty-acyl biosynthetic enzymes might have enabled Atlantic salmon, and other salmonids like rainbow trout, to colonise less productive riverine environments where the majority of biomass is of terrestrial origin poor in LC-PUFAs (Hanson *et al*., 1985; Schwalme, 1992; Sushchik *et al*., 2006), which are critical during freshwater developmental stages (Benitez-Santana *et al*., 2006; Navarro *et al*., 1995). Despite the dramatic difference in LC-PUFA levels between the principal prey species of insectivorous salmonids (Heissenberger *et al*., 2010; Sushchik *et al*., 2006) and carnivorous species, the body composition of wild specimens does not vary greatly (Blanchet *et al*., 2005; Jankowska *et al*., 2008; Schwalme, 1992) and neither does the body composition given the high seasonal n-3 LC-PUFA fluctuations experienced by the zoomacrobenthos composing their diet. These results suggest that salmonids have a higher biosynthetic capacity for LC-PUFA than carnivorous/piscivorous species, which might have enabled them to thrive in riverine biomes. Supporting evidence was obtained from tissue expression studies, indicating that salmonid fatty-acyl elongases and desaturases are greatly expressed in lipid biosynthetic tissues, opposite

to members of species with piscivorous/LC-PUFA rich such as the sister preduplicated lineage Northern pike.

Nevertheless, linkage analysis was inconclusive in resolving the origin of the Atlantic salmon duplicated targets given intense chromosome rearrangements experienced by this species after the whole-genome duplication. Moreover, with less chromosome arms, as a consequence of intense fusions, than most salmonid species it is difficult without a full genome sequence to discern between a WGD and a segmental duplication followed by intense rearrangements that placed both paralogs in non-homologous chromosomes (Phillips *et al*., 2009). Interestingly, the discovery of duplicated *elovl5* in rainbow trout suggested a more parsimonious origin by WGD, indicating that a single elongase present in the preduplicated ancestor might have been duplicated in all salmonids rather than originating from independent duplication events following the genome duplication. Increasing genomic resources from teleost species and the discovery of genetic markers, SNPs, microsatellites and AFLPs, allows integration of linkage maps and physical maps, which are highly valuable in comparative genomics. In addition, the expansion of the genome resources from a close preduplicated group would be an invaluable asset in phylogenetic and evolutionary studies aimed at elucidating the fate and evolutionary consequences of a recent WGD process in vertebrates, and providing a more accurate reconstruction of the ancestral salmonid genome.

In addition to WGD, acceleration of transposable element activation was observed after divergence from the closest preduplicated lineage (de Boer *et al*., 2007; Phillips *et al*., 2009) suggesting that mobile elements might have played an important role in the genome rediploidisation and speciation. Accordingly, our data

showed that the impact of mobile elements was clearly evident on the structure of salmon fatty-acyl elongase genes compared to the gene in pike, the closest preduplicated ortholog. The absence of transposable elements in pike *elovl5* introns suggests that these elements have accumulated in the salmonid genome after WGD, and unequally in duplicated segments. Such asymmetric accumulations might faciliate re-diploidisation, essential for divergence and ultimately speciation, by preventing the formation of tetravalents during meiosis. However, in the absence of specific information about the pike *elovl5* promoter, none of the functional differences observed in the *cis*-regulatory regions of the salmon *elovl5* genes could be firmly attributed to the presence of mobile elements, although the pike tissue expression data indicated the absence of transcriptional regulatory sequences driving significant expression of *elovl5* in liver or intestine.

5.2 FUTURE RESEARCH DIRECTIONS

These studies have shown that duplicated *elovl5* genes in Atlantic salmon have diverged, and that some of this divergence can be attributed to mobilisation of transposable elements in the salmonid line. To investigate further some of the hypotheses discussed above it would be desirable to extend similar studies to other salmonid genes. From an evolutionary perspective, defining the origin of genes in Atlantic salmon is challenging given the intense chromosome fusions that have occurred and the absence of full genomes. Yet, analysis of synteny or gene order could be carried out to disprove an origin from tandem duplication and look for large conserved synteny blocks within non-homologous chromosomes that would point

towards a WGD as the most parsimonious event. Such studies would be greatly improved by the final assemblies of the Atlantic salmon genome and the genome of the sister species rainbow trout. This would in turn enable targeted and more comprehensive study of homologous preduplicated representative genes from the closest preduplicated lineage, Northern pike. In relation to the transcriptional control of fatty-acyl elongases, sequence studies and functional assays of homologous genes would elucidate the origin and evolutionary significance of the functional divergence observed between the Atlantic salmon paralogs.

In polyploids, new phenotypes and speciation processes result not only from genomic rearrangements and gene expression changes, but are also associated with epigenetic changes, considered key in the spatiotemporal expression of targeted genes. Computational studies have indicated that duplicated genes are enriched for microRNA control (Li *et al*., 2008) suggesting that they are likely to be involved in the dosage regulation of duplicated genes, and DNA methylation alterations have been associated with gene expression changes after genome duplication (Chang and Liao, 2012; Ha *et al*., 2009), as well as with transposon mobilisation (Herpin *et al*., 2010). Hence, it would be interesting to compare the microRNA repertoire and DNA methylation patterns in Atlantic salmon and its closest preduplicated sister group in order to identify and understand their potential regulatory roles. A comparative study would provide insights into the role of these epigenetic elements in gene dosage buffering after WGD and into the origin of novel miRNA binding sites.

BIBLIOGRAPHY

Abzhanov, A., Protas, M., Grant, B.R., Grant, P.R., Tabin, C.J., 2004. Bmp4 and morphological variation of beaks in Darwin's finches. Science 305, 1462–1465.

Agaba, K.M., Tocher, D.R., Zheng, X., Dickson, C.A., Dick, J.R., Teale, A.J., 2005. Cloning and functional characterisation of polyunsaturated fatty acid elongases of marine and freshwater teleost fish. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 142, 342–352.

Agbaga, M.P., Brush, R.S., Mandal, M.N., Elliott, M.H., Al-Ubaidi, M.R., Anderson, R.E., 2010. Role of elovl4 protein in the biosynthesis of docosahexaenoic acid. Adv. Exp. Med. Biol. 664, 233–242.

Akiyama, Y., 1995. TFSEARCH: Searching transcription factor binding Sites [www.cbrc.jp/research/db/TFSEARCH.html].

Albert, C.M., Campos, H., Stampfer, M.J., Ridker, P.M., Manson, J.E., Willett, W.C., Ma, J., 2002. Blood levels of long-chain n-3 fatty acids and the risk of sudden death. N. Engl. J. Med. 346, 1113–1118.

Alfaro, M.E., Faircloth, B.C., Sorenson, L., Santini, F., 2013. A Phylogenomic perspective on the radiation of ray-finned fishes based upon targeted sequencing of ultraconserved elements (UCEs). PLoS One 8, e65923.

Allendorf, F.W., Danzamann, R.G., 1997. Secondary tetrasomic segregation of MDH-B and preferential pariring of homeologues in rainbow trout. Genetics 145, 1083–1092.

Allendorf, F.W., Thorgaard, G.H., 1984. Tetraploidy and the evolution of salmonid fishes, in: Turner, B.J. (Ed.), *Evolutionary Genetics of Fishes*. Plenum Press, New York, pp. 1–53.

Allendorf, F.W., Utter, F.M., 1976. Gene duplication in the family Salmonidae. Hereditas 82, 19–24.

Alonso, D.L., Maroto, F.G., 2000. Plants as "chemical factories" for the production of polyunsaturated fatty acids. Biotechnol. Adv. 18, 481–497.

Alp, A., Yeğen, V., Apaydin Yağci, M., Uysal, R., Biçen, E., Yağci, A., 2008. Diet composition and prey selection of the pike, *Esox lucius*, in Çivril Lake, Turkey. J. Appl. Ichthyol. 24, 670–677.

Amemiya-Kudo, M., Shimano, H., Hasty, A.H., Yahagi, N., Yoshikawa, T., Matsuzaka, T., Okazaki, H., Tamura, Y., Iizuka, Y., Ohashi, K., *et al*., 2002. Transcriptional activities of nuclear SREBP-1a, -1c, and -2 to different target promoters of lipogenic and cholesterogenic genes. J. Lipid Res. 43, 1220–1235.

Amemiya-Kudo, M., Shimano, H., Yoshikawa, T., Yahagi, N., Hasty, A.H., Okazaki, H., Tamura, Y., Shionoiri, F., Iizuka, Y., Ohashi, K., *et al*., 2000. Promoter analysis of the mouse sterol regulatory element-binding protein-1c gene. J. Biol. Chem. 275, 31078–31085.

Aoyama, T., Peters, J.M., Iritani, N., Nakajima, T., Furihata, K., Hashimoto, T., Gonzalez, F.J., 1998. Altered constitutive expression of fatty acid-metabolizing enzymes in mice lacking the peroxisome proliferator-activated receptor α (PPARα) 273, 5678–5684.

ASalBase, 2013. ASalBase [www.asalbase.org/sal-bin/index].

Atchley, W.R., Fitch, W.M., 1997. A natural classification of the basic helix-loop-helix class of transcription factors. Proc. Natl. Acad. Sci. U. S. A. 94, 5172–5176.

Aury, J.M., Jaillon, O., Duret, L., Noel, B., Jubin, C., Porcel, B.M., Segurens, B., Daubin, V., Anthouard, V., Aiach, N., *et al*., 2006. Global trends of whole-genome duplications revealed by the ciliate Paramecium tetraurelia. Nature 444, 171–178.

Ayala, F.J., Coluzzi, M., 2005. Chromosome speciation: humans, Drosophila, and mosquitoes. Proc. Natl. Acad. Sci. USA 102, 6535–6542.

Barrera, G., Toaldo, C., Pizzimenti, S., Cerbone, A., Pettazzoni, P., Dianzani, M.U., Ferretti, C., 2008. The role of PPAR ligands in controlling growth-related gene expression and their interaction with lipoperoxidation products. PPAR Res. 2008, 524671.

Barrière, A., Gordon, K.L., Ruvinsky, I., 2011. Distinct functional constraints partition sequence conservation in a cis-regulatory element. PLoS Genet. 7, e1002095.

Bejerano, G., Lowe, C.B., Ahituv, N., King, B., Siepel, A., Salama, S.R., Rubin, E.M., Kent, W.J., Haussler, D., 2006. A distal enhancer and an ultraconserved exon are derived from a novel retroposon. Nature 441, 87–90.

Bell, J.G., Ghioni, C., Sargent, J.R., 1994. Fatty acid compositions of 10 freshwater invertebrates which are natural food organisms of Atlantic salmon parr (*Salmo salar*): a comparison with commercial diets. Aquaculture 128, 301–313.

Bell, J.G., McEvoy, J., Tocher, D.R., McGhee, F., Campbell, P.J., Sargent, J.R., 2001. Replacement of fish oil with rapeseed oil in diets of Atlantic salmon (*Salmo salar*) affects tissue lipid compositions and hepatocyte fatty acid metabolism. J. Nutr. 131, 1535–1543.

Bell, J.G., Strachan, F., Good, E.J., Tocher, D.R., 2006. Effect of dietary echium oil on growth, fatty acid composition and metabolism, gill prostaglandin production and macrophage activity in Atlantic cod (*Gadus morhua* L.). Aquac. Res. 37, 606–617.

Bell, J.G., Tocher, D.R., Henderson, R.J., Dick, J.R., Crampton, V.O., 2003. Altered fatty acid compositions in Atlantic salmon (*Salmo salar*) fed diets containing linseed and rapeseed oils can be partially restored by a subsequent fish oil finishing diet. J. Nutr. 133, 2793–2801.

Benitez-Santana, T., Masuda, R., Juárez Carrillo, E., Ganuza, E., Valencia, A., Hernández-Cruz, C.M., Izquierdo, M.S., 2006. Dietary n-3 HUFA deficiency induces a reduced visual response in gilthead seabream *Sparus aurata* larvae. Aquaculture 264, 408–417.

Bento, M., Tomás, D., Viegas, W., Silva, M., 2013. Retrotransposons represent the most labile fraction for genomic rearrangements in polyploid plant species. Cytogenet. Genome Res. 140, 286–294.

Berg, J.M., 1989. DNA binding specificity of steroid receptors. Cell 57, 1065–1068.

Blanchet, C., Lucas, M., Julien, P., Morin, R., Gingras, S., Dewailly, E., 2005. Fatty acid composition of wild and farmed Atlantic salmon (*Salmo salar*) and rainbow trout (*Oncorhynchus mykiss*). Lipids 40, 529–531.

Boutin-Ganache, I., Raposo, M., Raymond, M., Deschepper, C.F., 2001. M13-tailed primers improved the readability and usability of microsatellites analyses performed with two different allele-sizing methods. Biotechniques 31, 24–28.

Breckenridge, W.C., Gombos, G., Morgan, I.G., 1972. The lipid composition of adult rat brain synaptosomal plasma membranes. Biochim. Biophys. Acta 266, 695–707.

Brett, M., Müller-Navarra, D., 1997. The role of highly unsaturated fatty acids in aquatic foodweb processes. Freshw. Biol. 38, 483–499.

Brosius, J., 1999. RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. Gene 238, 115–134.

Brosius, J., 2003. The contribution of RNAs and retroposition to evolutionary novelties. Genetica 118, 99–116.

Brown, M.S., Goldstein, J.L., 1997. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. Cell 89, 331–340.

Bunce, C.M., Campbell, M.J., 2010. *Nuclear Receptors: Current Concepts and Future Challenges*. Springer, Dordrecht.

Burd, C.J., Archer, T.K., 2010. Nuclear receptors and ATP dependent chromatin remodelin: a compex story, in: Bunce, C.M., Campbell, M.J. (Eds.), *Nuclear Receptors: Current Concepts and Future Challenges*. Springer, Dordrecht, pp. 345–363.

Buzzi, M., Henderson, R.J., Sargent, J.R., 1997. The biosynthesis of docosahexaenoic acid [22:6(n-3)] from linolenic acid in primary hepatocytes isolated from wild northern pike. J. Fish Biol. 51, 1197–1208.

Caimari, A., Oliver, P., Keijer, J., Palou, A., 2010. Peripheral blood mononuclear cells as a model to study the response of energy homeostasis-related genes to acute changes in feeding conditions. OMICS 14, 129–141.

Calder, P.C., 2006. n-3 polyunsaturated fatty acids, inflammation, and inflammatory diseases. Am. J. Clin. Nutr. 83, 1505–1519.

Calder, P.C., 2007. Immunomodulation by omega-3 fatty acids. Prostaglandins, Leukot. Essent. Fat. Acids 77, 327–335.

Calder, P.C., Albers, R., Antoine, J.M., Blum, S., Bourdet-Sicard, R., Ferns, G.A., Folkerts, G., Friedmann, P.S., Frost, G.S., Guarner, F., *et al*., 2009. Inflammatory disease processes and interactions with nutrition. Br. J. Nutr. 1, 1–45.

Calder, P.C., Burdge, G.C., 2004. Fatty acids, in: Nicolaou, A., Kokotos, G. (Eds.), *Bioactive Lipids*. Oily Press, Bridgewater, pp. 1–36.

Capy, P., Gasperi, G., Biémont, C., Bazin, C., 2000. Stress and transposable elements : co-evolution or useful parasites ? Heredity 85, 101–106.

Carmona-Antoñanzas, G., Monroig, O., Dick, J.R., Davie, A., Tocher, D.R., 2011. Biosynthesis of very long-chain fatty acids (C>24) in Atlantic salmon: cloning, functional characterisation, and tissue distribution of an Elovl4 elongase. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 159, 122–129.

Carmona-Antoñanzas, G., Tocher, D.R., Martinez-Rubio, L., Leaver, M.J., 2013a. Conservation of lipid metabolic gene transcriptional regulatory networks in fish and mammals. Gene 534, 1–9.

Carmona-Antoñanzas, G., Tocher, D.R., Taggart, J.B., Leaver, M.J., 2013b. An evolutionary perspective on Elovl5 fatty acid elongase: comparison of Northern pike and duplicated paralogs from Atlantic salmon. BMC Evol. Biol. 13, 85.

Carothers, A.D., Beatty, R.A., 1975. The recognition and incidence of haploid and polyploid spermatozoa in man, rabbit and mouse. J. Reprod. Fertil. 44, 487–500.

Castell, J.D., Bell, J.G., Tocher, D.R., Sargent, J.R., 1994. Effects of purified diets containing different combinations of arachidonic and docosahexaenoic acid on survival, growth and fatty acid composition of juvenile turbot (*Scophthalmus maximus*). Aquaculture 128, 315–333.

Castillo-Davis, C.I., Hartl, D.L., Achaz, G., 2004. cis-Regulatory and protein evolution in orthologous and duplicate genes. Genome Res. 14, 1530–1536.

Castrillo, A., Tontonoz, P., 2004. Nuclear receptors in macrophage biology: at the crossroads of lipid metabolism and inflammation. Annu. Rev. Cell Dev. Biol. 20, 455–480.

Castro, L.F.C., Monroig, Ó., Leaver, M.J., Wilson, J., Cunha, I., Tocher, D.R., 2012. Functional desaturase Fads1 (Δ5) and Fads2 (Δ6) orthologues evolved before the origin of jawed vertebrates. PLoS One 7, e31950.

Castro, L.F.C., Wilson, J.M., Goncalves, O., Galante-Oliveira, S., Rocha, E., Cunha, I., 2011. The evolutionary history of the stearoyl-CoA desaturase gene family in vertebrates. BMC Evol. Biol. 11, 132.

Chang, A.Y.F., Liao, B.Y., 2012. DNA methylation rebalances gene dosage after mammalian gene duplications. Mol. Biol. Evol. 29, 133–144.

Cheffers, U., Kraskok, A., Pancer, Z., Müller, W.E.G., 1997. High conservation of the serum response factor within Metazoa: cDNA from the sponge *Geodia cydonium*. Biol. J. Linn. Soc. 61, 127–137.

Christie, W.W., 2003. *Lipid Analysis*. Oily Press, Bridgwater.

Cohen, Z., Norman, H.A., Heimer, Y.M., 1995. Microalgae as a source of omega 3 fatty acids. World Rev. Nutr. Diet. 77, 1–31.

Colliar, L., Sturm, A., Leaver, M.J., 2011. Tributyltin is a potent inhibitor of piscine peroxisome proliferator-activated receptor α and β. Comp. Biochem. Physiol. C Toxicol. Pharmacol. 153, 168–173.

Comber, S.C.L., Smith, C., 2004. Polyploidy in fishes: patterns and processes. Biol. J. Linn. Soc. 82, 431–442.

Cook, H.W., McMaster, R.C.R., 2004. Fatty acid desaturation and chain elongation in eukaryotes, in: Vance, D.E., Vance, J.E. (Eds.), *Biochemistry of Lipids, Lipoproteins and Membranes*. Elsevier, Amsterdam, pp. 181–204.

Crawford, M.A., 1993. The role of essential fatty acids in neural development: implications for perinatal nutrition. Am. J. Clin. Nutr. 57, 703–710.

Crow, K.D., Stadler, P.F., Lynch, V.J., Amemiya, C., Wagner, G.P., 2006. The "fish-specific" Hox cluster duplication is coincident with the origin of teleosts. Mol. Biol. Evol. 23, 121–136.

Crow, K.D., Wagner, G.P., 2006. What is the role of genome duplication in the evolution of complexity and diversity? Mol. Biol. Evol. 23, 887–892.

Cruz-Garcia, L., Minghetti, M., Navarro, I., Tocher, D.R., 2009. Molecular cloning, tissue expression and regulation of liver X receptor (LXR) transcription factors of Atlantic salmon (*Salmo salar*) and rainbow trout (*Oncorhynchus mykiss*). Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 153, 81–88.

Cruz-Garcia, L., Sánchez-Gurmaches, J., Gutiérrez, J., Navarro, I., 2011. Regulation of LXR by fatty acids, insulin, growth hormone and tumor necrosis factor-a in rainbow trout myocytes. Comp. Biochem. Physiol. Part A Mol. Integr. Physiol. 160, 125–136.

Cruz-Garcia, L., Sánchez-Gurmaches, J., Gutiérrez, J., Navarro, I., 2012. Role of LXR in trout adipocytes: Target genes, hormonal regulation, adipocyte differentiation and relation to lipolysis. Comp. Biochem. Physiol. Part A Mol. Integr. Physiol. 163, 120–126.

Dangour, A.D., Uauy, R., 2008. N-3 long-chain polyunsaturated fatty acids for optimal function during brain development and ageing. Asia Pac. J. Clin. Nutr. 1, 185–188.

Dannevig, B.H., Brudeseth, B.E., Gjøen, T., Rode, M., Wergeland, H.I., Evensen, Ø., Press, C.M., 1997. Characterisation of a long-term cell line (SHK-1) developed from the head kidney of Atlantic salmon (*Salmo salar* L.). Fish Shellfish Immunol. 7, 213–226.

Danzmann, R.G., Cairney, M., Davidson, W.S., Ferguson, M.M., Gharbi, K., Guyomard, R., Holm, L.E., Leder, E., Okamoto, N., Ozaki, A., *et al.*, 2005. A comparative analysis of the rainbow trout genome with 2 other species of fish (Arctic charr and Atlantic salmon) within the tetraploid derivative Salmonidae family (subfamily: Salmoninae). Genome 48, 1037–1051.

Danzmann, R.G., Davidson, E.A., Ferguson, M.M., Gharbi, K., Koop, B.F., Hoyheim, B., Lien, S., Lubieniecki, K.P., Moghadam, H.K., Park, J., *et al.*, 2008. Distribution of ancestral proto-Actinopterygian chromosome arms within the genomes of 4R-derivative salmonid fishes (Rainbow trout and Atlantic salmon). BMC Genomics 9, 557.

David, L., Blum, S., Feldman, M.W., Lavi, U., Hillel, J., 2003. Recent duplication of the common carp (*Cyprinus carpio* L.) genome as revealed by analyses of microsatellite loci. Mol. Biol. Evol. 20, 1425–1434.

De Boer, J.G., Yazawa, R., Davidson, W.S., Koop, B.F., 2007. Bursts and horizontal evolution of DNA transposons in the speciation of pseudotetraploid salmonids. BMC Genomics 8, 422.

Delport, W., Poon, A.F.Y., Frost, S.D.W., Kosakovsky Pond, S.L., 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26, 2455–2457.

Desvergne, B., Michalik, L., Wahli, W., 2006. Transcriptional regulation of metabolism. Physiol. Rev. 86, 465–514.

Donoghue, P.C.J., Purnell, M.A., 2005. Genome duplication, extinction and vertebrate evolution. Trends Ecol. Evol. 20, 312–319.

Dyer, B.W., Ferrer, F.A., Klinedinst, D.K., Rodriguez, R., 2000. A noncommercial dual luciferase enzyme assay system for reporter gene analysis. Anal. Biochem. 282, 158–161.

Egea, P.F., Mitschler, A., Rochel, N., Ruff, M., Chambon, P., Moras, D., 2000. Crystal structure of the human RXR a ligand-binding domain bound to its natural ligand : 9- cis retinoic acid. EMBO J. 19, 2592–2601.

EMBOSS, 2013. EMBOSS Needle Pairwise Sequence Alignment [www.ebi.ac.uk/Tools/psa/emboss_needle/].

Engelman, D.M., Steitz, T.A., Goldman, A., 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu. Rev. Biophys. Biophys. Chem. 15, 321–353.

Espenshade, P.J., 2006. SREBPs: sterol-regulated transcription factors. J. Cell Sci. 119, 973–976.

Evans, R.M., 1988. The steroid and thyroid hormone receptor superfamily. Science 240, 889–895.

Färnegårdh, M., Bonn, T., Sun, S., Ljunggren, J., Ahola, H., Wilhelmsson, A., Gustafsson, J.Å., Carlquist, M., 2003. The three-dimensional structure of the liver X receptor beta reveals a flexible ligand-binding pocket that can accommodate fundamentally different ligands. J. Biol. Chem. 278, 38821–38828.

Fedoroff, N. V., 2012. Transposable elements, epigenetics, and genome evolution. Science 338, 758–767.

Fernandes, J.M.O., Ruangsri, J., Kiron, V., 2010. Atlantic Cod piscidin and its diversification through positive selection. PLoS One 5, e9501.

Fernández-Alvarez, A., Alvarez, M.S., Gonzalez, R., Cucarella, C., Muntané, J., Casado, M., 2011. Human SREBP1c expression in liver is directly regulated by peroxisome proliferator-activated receptor alpha (PPARalpha). J. Biol. Chem. 286, 21466–21477.

Ferrigno, O., Virolle, T., Djabari, Z., Ortonne, J.P., White, R.J., Aberdam, D., 2001. Transposable B2 SINE elements can provide mobile RNA polymerase II promoters. Nat. Genet. 28, 77–81.

Folch, J., Lees, M., Sloane Stanley, G.H., 1957. A simple method for the isolation and purification of total lipides from animal tissues. J. Biol. Chem. 226, 497–509.

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

Forman, B.M., Chen, J., Evans, R.M., 1997. Hypolipidemic drugs, polyunsaturated fatty acids, and eicosanoids are ligands for peroxisome proliferator-activated receptors alpha and delta. Proc. Natl. Acad. Sci. U. S. A. 94, 4312–4317.

Fu, Y., Weng, Z., 2004. Improvement of TRANSFAC matrices using multiple local alignment of transcription factor binding site sequences. Conf. Proc. IEEE Eng. Med. Biol. Soc. 4, 2856–2859.

Gallardo, M.H., Kausel, G., Jiménez, A., Bacquet, C., González, C., Figueroa, J., Köhler, N., Ojeda, R., 2004. Whole-genome duplications in South American desert rodents (Octodontidae). Biol. J. Linn. Soc. 82, 443–451.

Gaut, B.S., Doebley, J.F., 1997. DNA sequence evidence for the segmental allotetraploid origin of maize. Proc. Natl. Acad. Sci. USA 94, 6809–6814.

Ghioni, C., Tocher, D.R., Bell, M. V, Dick, J.R., Sargent, J.R., 1999. Low C18 to C20 fatty acid elongase activity and limited conversion of stearidonic acid, 18:4(n-3), to

eicosapentaenoic acid, 20:5(n-3), in a cell line from the turbot, *Scophthalmus maximus*. Biochim. Biophys. Acta 1437, 170–181.

Gladyshev, M.I., Arts, M.T., Sushchik, N.N., 2009. Preliminary estimates of the export of omega-3 highly unsaturated fatty acids (EPA + DHA) from aquatic to terrestrial ecosystems, in: Arts, M.T., Brett, M.T., Kainz, M. (Eds.), *Lipids in Aquatic Ecosystems*. Springer, New York, pp. 179–209.

Gladyshev, M.I., Sushchik, N.N., Anishchenko, O. V, Makhutova, O.N., Kolmakov, V.I., Kalachova, G.S., Kolmakova, A.A, Dubovskaya, O.P., 2011. Efficiency of transfer of essential polyunsaturated fatty acids versus organic carbon from producers to consumers in a eutrophic reservoir. Oecologia 165, 521–531.

Glass, C.K., Rose, D.W., Rosenfeld, M.G., 1997. Nuclear receptor coactivators. Curr. Opin. Cell Biol. 9, 222–232.

Glass, C.K., Rosenfeld, M.G., 2000. The coregulator exchange in transcriptional functions of nuclear receptors. Genes Dev. 14, 121–141.

Goldstein, J.L., Rawson, R.B., Brown, M.S., 2002. Mutant mammalian cells as tools to delineate the sterol regulatory element-binding protein pathway for feedback regulation of lipid synthesis. Arch. Biochem. Biophys. 397, 139–148.

Gray, Y.H., 2000. It takes two transposons to tango: transposable-element-mediated chromosomal rearrangements. Trends Genet. 16, 461–468.

Gregory, M.K., See, V.H.L., Gibson, R.A., Schuller, K.A., 2010. Cloning and functional characterisation of a fatty acyl elongase from southern bluefin tuna (*Thunnus maccoyii*). Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 155, 178–185.

Gregory, T.R., 2013. Animal genome size database [www.genomesize.com].

Gross, M.R., Coleman, R.M., McDowall, R.M., 1988. Aquatic productivity and the evolution of diadromous fish migration. Science 239, 1291–1293.

Grün, F., Watanabe, H., Zamanian, Z., Maeda, L., Arima, K., Cubacha, R., Gardiner, D.M., Kanno, J., Iguchi, T., Blumberg, B., 2006. Endocrine-disrupting organotin compounds are potent inducers of adipogenesis in vertebrates. Mol. Endocrinol. 20, 2141–2155.

Gu, Z., Nicolae, D., Lu, H.H.S., Li, W.H., 2002. Rapid divergence in expression between duplicate genes inferred from microarray data. Trends Genet. 18, 609–613.

Gunstone, F.D., 2011. The world's oils and fats, in: Turchini, G.M., Ng, W.K., Tocher, D.R. (Eds.), *Fish Oil Replacements and Alternative Lipid Sources in Aquaculture Feeds*. CRC, Boca Raton, pp. 61–98.

Gurr, M.I., Harwood, J.L., Frayn, K.N., 2002. *Lipid Biochemistry : An Introduction*. Blackwell Science Ltd, Oxford.

Ha, M., Kim, E.D., Chen, Z.J., 2009. Duplicate genes increase expression diversity in closely related species and allopolyploids. Proc. Natl. Acad. Sci. U. S. A. 106, 2295–2300.

Hamilton, J., Brunaldi, K., 2007. A model for fatty acid transport into the brain. J. Mol. Neurosci. 33, 12–17.

Hansen, M.B., Nielsen, S.E., Berg, K., 1989. Re-examination and further development of a precise and rapid dye method for measuring cell growth/cell kill. J. Immunol. Methods 119, 203–210.

Hanson, B.J., Cummins, K.W., Cargill, A.S., Lowry, R.R., 1985. Lipid content fatty-acid composition and the effect of diet on fats of aquatic insects. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 80, 257–276.

Hare, E.E., Peterson, B.K., Iyer, V.N., Meier, R., Eisen, M.B., 2008. Sepsid even-skipped enhancers are functionally conserved in Drosophila despite lack of sequence conservation. PLoS Genet. 4, e1000106.

Hashimoto, K., Yoshizawa, A.C., Okuda, S., Kuma, K., Goto, S., Kanehisa, M., 2008. The repertoire of desaturases and elongases reveals fatty acid variations in 56 eukaryotic genomes. J. Lipid Res. 49, 183–191.

Hastings, N., Agaba, M.K., Tocher, D.R., Zheng, X., Dickson, C.A., Dick, J.R., Teale, A.J., 2004. Molecular cloning and functional characterization of fatty acyl desaturase and elongase cDNAs involved in the production of eicosapentaenoic and docosahexaenoic acids from α-linolenic acid in Atlantic salmon (*Salmo salar*). Mar. Biotechnol. 6, 463–474.

Hasty, A.H., Shimano, H., Yahagi, N., Amemiya-Kudo, M., Perrey, S., Yoshikawa, T., Osuga, J., Okazaki, H., Tamura, Y., Iizuka, Y., *et al*., 2000. Sterol regulatory element-binding protein-1 is regulated by glucose at the transcriptional level. J. Biol. Chem. 275, 31069–31077.

Heinz, E., 1993. Biosynthesis of polyunsaturated fatty acids, in: Moore, T.S.J. (Ed.), *Lipid Metabolism in Plants*. CRC Press, Boca Raton, pp. 33–89.

Heissenberger, M., Watzke, J., Kainz, M.J., 2010. Effect of nutrition on fatty acid profiles of riverine, lacustrine, and aquaculture-raised salmonids of pre-alpine habitats. Hydrobiologia 650, 243–254.

Helentjaris, T., Weber, D., Wright, S., 1988. Identification of the genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. Genetics 118, 353–363.

Henderson, R.J., Park, M.T., Sargent, J.R., 1995. The desaturation and elongation of 14C-labelled polyunsaturated fatty acids by pike (*Esox lucius* L.) in vivo. Fish Physiol. Biochem. 14, 223–235.

Herpin, A., Braasch, I., Kraeussling, M., Schmidt, C., Thoma, E.C., Nakamura, S., Tanaka, M., Schartl, M., 2010. Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements. PLos Genet. 6, e1000844.

Hiromori, Y., Nishikawa, J., Yoshida, I., Nagase, H., Nakanishi, T., 2009. Structure-dependent activation of peroxisome proliferator-activated receptor (PPAR) gamma by organotin compounds. Chem. Biol. Interact. 180, 238–244.

Hoegg, S., Brinkmann, H., Taylor, J.S., Meyer, A., 2004. Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. J. Mol. Evol. 59, 190–203.

Hoegg, S., Meyer, A., 2005. Hox clusters as models for vertebrate genome evolution. Trends Genet. 21, 421–424.

Holman, R.T., 1986. Control of polyunsaturated acids in tissue lipids. J. Am. Coll. Nutr. 5, 183–211.

Holmbeck, S.M., Dyson, H.J., Wright, P.E., 1998. DNA-induced conformational changes are the basis for cooperative dimerization by the DNA binding domain of the retinoid X receptor. J. Mol. Biol. 284, 533–539.

Horton, J.D., Goldstein, J.L., Brown, M.S., 2002. SREBPs : activators of the complete program of cholesterol and fatty acid synthesis in the liver. J. Clin. Invest. 109, 1125–1131.

Horton, J.D., Shah, N.A., Warrington, J.A., Anderson, N.N., Park, S.W., Brown, M.S., Goldstein, J.L., 2003. Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes. Proc. Natl. Acad. Sci. U. S. A. 100, 12027–12032.

Hua, X., Wu, J., Goldstein, J.L., Brown, M.S., Hobbs, H.H., 1995. Structure of the human gene encoding sterol regulatory element binding protein-1 (SREBF1) and localization of SREBF1 and SREBF2 to chromosomes 17p11.2 and 22q13. Genomics 25, 667–673.

Hua, X., Yokoyama, C., Wu, J., Briggs, M.R., Brown, M.S., Goldstein, J.L., Wang, X., 1993. SREBP-2, a second basic-helix-loop-helix-leucine zipper protein that stimulates transcription by binding to a sterol regulatory element. Proc. Natl. Acad. Sci. U. S. A. 90, 11603–11607.

Hughes, A.L., 2007. Looking for Darwin in all the wrong places: the misguided quest for positive selection at the nucleotide sequence level. Heredity 99, 364–373.

Hughes, A.L., Green, J.A., Garbayo, J.M., Roberts, R.M., 2000. Adaptive diversification within a large family of recently duplicated, placentally expressed genes. Proc. Natl. Acad. Sci. USA 97, 3319–3323.

Hughes, M.K., Hughes, A.L., 1993. Evolution of duplicate genes in a tetraploid animal, *Xenopus laevis*. Mol. Biol. Evol. 10, 1360–1369.

Hulbert, A.J., 2003. Life, death and membrane bilayers. J. Exp. Biol. 206, 2303–2311.

Huminiecki, L., Heldin, C.H., 2010. 2R and remodeling of vertebrate signal transduction engine. BMC Biol. 8, 146.

Hurst, G.D., Werren, J.H., 2001. The role of selfish genetic elements in eukaryotic evolution. Nat. Rev. Genet. 2, 597–606.

Ijpenberg, A., Jeannin, E., Wahli, W., Desvergne, B., 1997. Polarity and specific sequence requirements of peroxisome proliferator-activated receptor (PPAR)/retinoid X receptor heterodimer binding to DNA. A functional analysis of the malic enzyme gene PPAR response element. J. Biol. Chem. 272, 20108–20117.

Inagaki, K., Aki, T., Fukuda, Y., Kawamoto, S., Shigeta, S., Ono, K., Suzuki, O., 2002. Identification and expression of a rat fatty acid elongase involved in the biosynthesis of C18 fatty acids. Biosci. Biotechnol. Biochem. 66, 613–621.

Ishiguro, N.B., Miya, M., Nishida, M., 2003. Basal euteleostean relationships: a mitogenomic perspective on the phylogenetic reality of the "Protacanthopterygii". Mol. Phylogenet. Evol. 27, 476–488.

Jackson, S.M., Ericsson, J., Mantovani, R., Edwards, P.A., 1998. Synergistic activation of transcription by nuclear factor Y and sterol regulatory element binding protein. J. Lipid Res. 39, 767–776.

Jackson, S.M., Ericsson, J., Osborne, T.F., Edwards, P.A., 1995. NF-Y has a novel role in sterol-dependent transcription of two cholesterogenic genes. J. Biol. Chem. 270, 21445–21448.

Jaillon, O., Aury, J.M., Brunet, F., Petit, J.L., Stange-Thomann, N., Mauceli, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C., Bernot, A., *et al*., 2004. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. Nature 431, 946–957.

Jakobsson, A., Westerberg, R., Jacobsson, A., 2006. Fatty acid elongases in mammals: their regulation and roles in metabolism. Prog. Lipid Res. 45, 237–249.

Jakobsson, T., Treuter, E., Gustafsson, J.A., Steffensen, K.R., 2012. Liver X receptor biology and pharmacology: new pathways, challenges and opportunities. Trends Pharmacol. Sci. 33, 394–404.

Jankowska, B., Zakęś, Z., Żmijewski, T., Szczepkowski, M., 2008. Fatty acid composition of wild and cultured northern pike (*Esox lucius*). J. Appl. Ichthyol. 24, 196–201.

Janowski, B.A., Willy, P.J., Devi, T.R., Falck, J.R., Mangelsdorf, D.J., 1996. An oxysterol signalling pathway mediated by the nuclear receptor LXR alpha. Nature 383, 728–731.

Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., Smale, S.T., 1994. DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol. Cell. Biol. 14, 116–127.

Jeffery, W.R., 2005. Adaptive evolution of eye degeneration in the Mexican blind cavefish. J. Hered. 96, 185–196.

Jeon, T.I., Osborne, T.F., 2012. SREBPs: metabolic integrators in physiology and metabolism. Trends Endocrinol. Metab. 23, 65–72.

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275–282.

Jonsson, B., Jonsson, N., 2011. *Ecology of Atlantic salmon and Brown trout: Habitat as a template for life histories*. Springer, Dordrecht.

Joseph, S.B., Laffitte, B.A., Patel, P.H., Watson, M.A., Matsukuma, K.E., Walczak, R., Collins, J.L., Osborne, T.F., Tontonoz, P., 2002. Direct and indirect mechanisms for regulation of fatty acid synthase gene expression by liver X receptors. J. Biol. Chem. 277, 11019–11025.

Jump, D.B., 2009. Mammalian fatty acid elongases. Methods Mol. Biol. 579, 375–389.

Jurka, J., Kapitonov, V.V, Pavlicek, A., Klonowski, P., Kohany, O., Walichiewicz, J., 2005. Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. 110, 462–427.

Karapanagiotidis, I.T., Bell, M.V, Little, D.C., Yakupitiyage, A., 2007. Replacement of dietary fish oils by alpha-linolenic acid-rich oils lowers omega 3 content in tilapia flesh. Lipids 42, 547–560.

Karapanagiotidis, I.T., Bell, M.V, Little, D.C., Yakupitiyage, A., Rakshit, S.K., 2006. Polyunsaturated fatty acid content of wild and farmed tilapias in Thailand: effect of aquaculture practices and implications for human nutrition. J. Agric. Food Chemestry 54, 4304–4310.

Kidd, P.M., 2007. Omega-3 DHA and EPA for cognition, behavior, and mood: clinical findings and structural-functional synergies with cell membrane phospholipids. Altern. Med. Rev. 12, 207–227.

Kim, J.B., Spotts, G.D., Halvorsen, Y.D., Shih, H.M., Ellenberger, T., Towle, H.C., Spiegelman, B.M., 1995. Dual DNA binding specificity of ADD1/SREBP1 controlled by a single amino acid in the basic helix-loop-helix domain. Mol. Cell. Biol. 15, 2582–2588.

Kimura, M., 1983. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.

Klug, A., Schwabe, J.W., 1995. Protein motifs 5. Zinc fingers. FASEB J. 9, 597–604.

Kluytmans, J.H.F.M., Zandee, D.I., 1974. Lipid metabolism in the northern pike (*Esox lucius* L.) in vivo incorporation of 1-14C-acetate in the lipids. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 48, 641–649.

Knoll, H.A., Summons, R.E., Waldbauer, J.R., Zumberge, J.E., 2007. The geological succession of primary producers in the oceans, in: Falkowski, P.G., Knoll, A.H. (Eds.), *Evolution of Primary Producers in the Sea*. Academic Press, New York, pp. 133–163.

Kohany, O., Gentles, A.J., Hankus, L., Jurka, J., 2006. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinformatics 7, 474.

Kondrashov, F.A., Kondrashov, A.S., 2006. Role of selection in fixation of gene duplications. J. Theor. Biol. 239, 141–151.

Koop, B.F., Davidson, W.S., 2008. Genomics and the genome Duplication in salmonids, in: Tsukamoto, K., Kawamura, T., Takeuchi, T., Beard, T.D.J., Kaiser, M.J. (Eds.), *Fisheries for Global Welfare and Environment*, 5th World Fisheries Congress. Terrapub, Tokyo, pp. 77–86.

Korber, B., 2000. HIV signature and sequence variation analysis, in: Rodrigo, A.G., Learn, G.H. (Eds.), *Computational and Evolutionary Analysis of HIV Molecular Sequences*. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 55–72.

Kortner, T.M., Gu, J., Krogdahl, A., Bakke, A.M., 2013. Transcriptional regulation of cholesterol and bile acid metabolism after dietary soyabean meal treatment in Atlantic salmon (*Salmo salar* L.). Br. J. Nutr. 30, 1–12.

Kosakovsky Pond, S.L., Frost, S.D.W., 2004. A genetic algorithm approach to detecting lineage-specific variation in selection pressure. Mol. Biol. Evol. 22, 478–485.

Kosakovsky Pond, S.L., Frost, S.D.W., 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21, 2531–2533.

Krasnov, A., Koskinen, H., Afanasyev, S., Mölsä, H., 2005. Transcribed Tc1-like transposons in salmonid fish. BMC Genomics 6, 107.

Krogdahl, Å., Nordrum, S., Sørensen, M., Brudeseth, L., Røsjø, C., 1999. Effects of diet composition on apparent nutrient absorption along the intestinal tract and of subsequent fasting on mucosal disaccharidase activities and plasma nutrient concentration in Atlantic salmon Salmo salar L. Aquac. Nutr. 5, 121–133.

Kumar, M.S., Hendrix, J.A., Johnson, A.D., Owens, G.K., 2003. Smooth muscle alpha-actin gene requires two E-boxes for proper expression in vivo and is a target of class I basic helix-loop-helix proteins. Circ. Res. 92, 840–847.

Kuraku, S., Meyer, A., Kuratani, S., 2009. Timing of genome duplications relative to the origin of the vertebrates: did cyclostomes diverge before or after? Mol. Biol. Evol. 26, 47–59.

Laffitte, B.A., Joseph, S.B., Walczak, R., Pei, L., Wilpitz, D.C., Collins, J.L., Tontonoz, P., 2001. Autoregulation of the human liver X receptor alpha promoter. Mol. Cell. Biol. 21, 7558–7568.

Lai, Y.Y.Y., Lubieniecki, K.P., Phillips, R.B., Chow, W., Koop, B.F., Davidson, W.S., 2009. Genomic organization of Atlantic salmon (*Salmo salar*) fatty acid binding protein (fabp2) genes reveals independent loss of duplicate loci in teleosts. Mar. Genomics 2, 193–200.

Lalli, E., Sassone-Corsi, P., 1994. Signal transduction and gene regulation: the nuclear response to cAMP. J. Biol. Chem. 269, 17359–17362.

Latchman, D.S., 1997. Transcription factors: an overview. Int. J. Biochem. Cell Biol. 29, 1305–1312.

Lauritzen, L., Hansen, H.S., Jorgensen, M.H., Michaelsen, K.F., 2001. The essentiality of long chain n-3 fatty acids in relation to development and function of the brain and retina. Prog. Lipid Res. 40, 1–94.

Leaver, M., Bautista, J.M., Björnsson, B.T., Jönsson, E., Krey, G., Tocher, D.R., Torstensen, B.E., 2008a. Towards fish lipid nutrigenomics: current state and prospects for fin-fish aquaculture. Rev. Fish. Sci. 16, 73–94.

Leaver, M., Boukouvala, E., Antonopoulou, E., Diez, A., Favre-Krey, L., Ezaz, M.T., Bautista, J.M., Tocher, D.R., Krey, G., 2005. Three peroxisome proliferator-activated receptor isotypes from each of two species of marine fish. Endocrinology 146, 3150–3162.

Leaver, M., Ezaz, M.T., Fontagne, S., Tocher, D.R., Boukouvala, E., Krey, G., 2007. Multiple peroxisome proliferator-activated receptor β subtypes from Atlantic salmon (*Salmo salar*). J. Mol. Endocrinol. 38, 391–400.

Leaver, M., Villeneuve, L.A.N., Obach, A., Jensen, L., Bron, J.E., Tocher, D.R., Taggart, J.B., 2008b. Functional genomics reveals increased cholesterol and highly unsaturated fatty acid biosynthesis after dietary substitution of fish oil with vegetable oils in Atlantic salmon (*Salmo salar*). BMC Genomics 9, 299.

Leder, E.H., Danzmann, R.G., Ferguson, M.M., 2004. Comparison of GNRH3 genes across salmonid genera. Anim. Genet. 35, 126–129.

Lee, S.J., Sekimoto, T., Yamashita, E., Nagoshi, E., Nakagawa, A., Imamoto, N., Yoshimura, M., Sakai, H., Chong, K.T., Tsukihara, T., *et al*., 2003. The structure of importin-beta bound to SREBP-2: nuclear import of a transcription factor. Science 302, 1571–1575.

Lengqvist, J., Mata De Urquiza, A., Bergman, A.C., Willson, T.M., Sjövall, J., Perlmann, T., Griffiths, W.J., 2004. Polyunsaturated fatty acids including docosahexaenoic and arachidonic acid bind to the retinoid X receptor alpha ligand-binding domain. Mol. Cell. Proteomics 3, 692–703.

Leonard, A.E., Bobik, E.G., Dorado, J., Kroeger, P.E., Chuang, L.T., Thurmond, J.M., Parker-Barnes, J.M., Das, T., Huang, Y.S., Mukerji, P., 2000. Cloning of a human cDNA encoding a novel enzyme involved in the elongation of long-chain polyunsaturated fatty acids. Biochem. J. 350, 765–770.

Leonard, A.E., Pereira, S.L., Sprecher, H., Huang, Y.S., 2004. Elongation of long-chain fatty acids. Prog. Lipid Res. 43, 36–54.

Leong, J.S., Jantzen, S.G., von Schalburg, K.R., Cooper, G.A., Messmer, A.M., Liao, N.Y., Munro, S., Moore, R., Holt, R.A., Jones, S.J., *et al*., 2010. *Salmo salar* and *Esox lucius* full-length cDNA sequences reveal changes in evolutionary pressures on a post-tetraploidization genome. BMC Genomics 11, 279.

Levasseur, A., Pontarotti, P., 2011. The role of duplications in the evolution of genomes highlights the need for evolutionary-based approaches in comparative genomics. Biol. Direct 6, 11.

Li, C., Lu, G., Orti, G., 2008. Optimal data partitioning and a test case for ray-finned fishes (Actinopterygii) based on ten nuclear loci. Syst. Biol. 57, 519–539.

Li, J., Musso, G., Zhang, Z., 2008. Preferential regulation of duplicated genes by microRNAs in mammals. Genome Biol. 9, R132.

Li, W.H., 1997. *Molecular Evolution*. Sinauer Associates, Sunderland.

Lien, S., Gidskehaug, L., Moen, T., Hayes, B.J., Berg, P.R., Davidson, W.S., Omholt, S.W., Kent, M.P., 2011. A dense SNP-based linkage map for Atlantic salmon (*Salmo salar*) reveals extended chromosome homeologies and striking differences in sex-specific recombination patterns. BMC Genomics 12, 615.

Long, M., Betran, E., Thornton, K., Wang, W., 2003. The origin of new genes: glimpses from the young and old. Nat. Rev. Genet. 4, 865–875.

Lönnig, W.E., Saedler, H., 2002. Chromosome rearrangements and transposable elements. Annu. Rev. Genet. 36, 389–410.

López, J.A., Chen, W.J., Ortí, G., Wood, R.M., 2004. Esociform phylogeny. Copeia 3, 449–464.

López-Arbarello, A., 2012. Phylogenetic interrelationships of ginglymodian fishes (Actinopterygii: Neopterygii). PLoS One 7, e39370.

Ludwig, A., Belfiore, N.M., Pitra, C., Svirsky, V., Jenneckens, I., 2001. Genome duplication events and functional reduction of ploidy levels in sturgeon (Acipenser, Huso and Scaphirhynchus). Genetics 158, 1203–1215.

Lynch, M., Force, A., 2000. The origin of interspecies genomic incompatibility via gene duplication. Am. Nat. 156, 590–605.

Ma, B., Tromp, J., Li, M., 2002. PatternHunter: faster and more sensitive homology search. Bioinformatics 18, 440–445.

Mable, B.K., 2004. "Why polyploidy is rarer in animals than in plants": myths and mechanisms. Biol. J. Linn. Soc. 82, 453–466.

Maclean, C.J., Greig, D., 2011. Reciprocal gene loss following experimental whole-genome duplication causes reproductive isolation in yeast. Evolution 65, 932–945.

Macqueen, D.J., Johnston, I.A., 2014. A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. Proc. R. Soc. 281, 20132881.

Magaña, M.M., Koo, S.H., Towle, H.C., Osborne, T.F., 2000. Different sterol regulatory element-binding protein-1 isoforms utilize distinct co-regulatory factors to activate the promoter for fatty acid synthase. J. Biol. Chem. 275, 4726–4733.

Maglich, J.M., Sluder, A., Guan, X., Shi, Y., McKee, D.D., Carrick, K., Kamdar, K., Willson, T.M., Moore, J.T., 2001. Comparison of complete nuclear receptor sets from the human, Caenorhabditis elegans and Drosophila genomes. Genome Biol. 2, Research0029.

Makałowski, W., 2000. Genomic scrap yard: how genomes utilize all that junk. Gene 259, 61–67.

Makhutova, O.N., Sushchik, N.N., Gladyshev, M.I., Ageev, A.V, Pryanichnikova, E.G., Kalachova, G.S., 2011. Is the fatty acid composition of freshwater zoobenthic invertebrates controlled by phylogenetic or trophic factors? Lipids 46, 709–721.

Makishima, M., 2005. Nuclear receptors as targets for drug development: regulation of cholesterol and bile acid metabolism by nuclear receptors. J. Pharmacol. Sci. 97, 177–183.

Mariño-Ramírez, L., Jordan, I.K., 2006. Transposable element derived DNaseI-hypersensitive sites in the human genome. Biol. Direct 1, 20.

Mariño-Ramírez, L., Lewis, K.C., Landsman, D., Jordan, I.K., 2005. Transposable elements donate lineage-specific regulatory sequences to host genomes. Cytogenet. Genome Res. 110, 333–341.

Markov, G., Bonneton, F., Laudet, V., 2010. What does evolution teach us about nuclear receptors?, in: Bunce, C.M., Campbell, M.J. (Eds.), *Nuclear Receptors*. Springer, Dordrecht, pp. 14–29.

Marquart, T.J., Allen, R.M., Ory, D.S., Baldán, A., 2010. miR-33 links SREBP-2 induction to repression of sterol transporters. Proc. Natl. Acad. Sci. U. S. A. 107, 12228–12232.

Martens, K., Ver Loren van Themaat, E., van Batenburg, M.F., Heinäniemi, M., Huyghe, S., Van Hummelen, P., Carlberg, C., Van Veldhoven, P.P., Van Kampen, A., Baes, M., 2008. Coordinate induction of PPARa and SREBP2 in multifunctional protein 2 deficient mice. Biochim. Biophys. Acta 1781, 694–702.

Mather, K., 1951. *The Measurement of Linkage in Heredity*. Methuen, London.

Matsuzaka, T., Shimano, H., Yahagi, N., Amemiya-Kudo, M., Yoshikawa, T., Hasty, A.H., Tamura, Y., Osuga, J., Okazaki, H., Iizuka, Y., *et al*., 2002. Dual regulation of mouse

Delta(5)- and Delta(6)-desaturase gene expression by SREBP-1 and PPARalpha. J. Lipid Res. 43, 107–114.

Matys, V., Kel-Margoulis, O.V, Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K., *et al*., 2006. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res. 34, 108–110.

McElroy, W.D., Green, A., 1956. Function of adenosine triphosphate in the activation of luciferin. Arch. Biochem. Biophys. 64, 257–271.

McKay, S.J., Trautner, J., Smith, M.J., Koop, B.F., Devlin, R.H., 2004. Evolution of duplicated growth hormone genes in autotetraploid salmonid fishes. Genome 47, 714–723.

Michalik, L., Auwerx, J., Berger, J.P., Chatterjee, V.K., Glass, C.K., González, F.J., Grimaldi, P.A., Kadowaki, T., Lazar, M.A., O'Rahilly, S., *et al*., 2006. International Union of Pharmacology. LXI. Peroxisome proliferator-activated receptors. Pharmacol. Rev. 58, 726–741.

Minghetti, M., Leaver, M.J., Tocher, D.R., 2011. Transcriptional control mechanisms of genes of lipid and fatty acid metabolism in the Atlantic salmon (*Salmo salar* L.) established cell line, SHK-1. Biochim. Biophys. Acta 1811, 194–202.

Mitchell, R.W., On, N.H., Del Bigio, M.R., Miller, D.W., Hatch, G.M., 2011. Fatty acid transport protein expression in human brain and potential role in fatty acid transport across human brain microvessel endothelial cells. J. Neurochem. 117, 735–746.

Miyata, T., Yasunaga, T., 1980. Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. J. Mol. Evol. 16, 23–36.

Moen, T., Hayes, B., Baranski, M., Berg, P.R., Kjoglum, S., Koop, B.F., Davidson, W.S., Omholt, S.W., Lien, S., 2008. A linkage map of the Atlantic salmon (*Salmo salar*) based on EST-derived SNP markers. BMC Genomics 9, 223.

Moghadam, H.K., Ferguson, M.M., Danzmann, R.G., 2011. Whole genome duplication: challenges and considerations associated with sequence orthology assignment in Salmoninae. J. Fish Biol. 79, 561–574.

Monroig, O., Li, Y., Tocher, D.R., 2011. Delta-8 desaturation activity varies among fatty acyl desaturases of teleost fish: High activity in delta-6 desaturases of marine species. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 159, 206–213.

Monroig, O., Rotllant, J., Cerdá-Reverter, J.M., Dick, J.R., Figueras, A., Tocher, D.R., 2010a. Expression and role of Elovl4 elongases in biosynthesis of very long-chain fatty acids during zebrafish Danio rerio early embryonic development. Biochim. Biophys. Acta 1801, 1145–1154.

Monroig, O., Zheng, X., Morais, S., Leaver, M.J., Taggart, J.B., Tocher, D.R., 2010b. Multiple genes for functional 6 fatty acyl desaturases (Fad) in Atlantic salmon (*Salmo*

salar L.): gene and cDNA characterization, functional expression, tissue distribution and nutritional regulation. Biochim. Biophys. Acta 1801, 1072–1081.

Montero, D., Izquierdo, M.S., 2011. Welfare and health of fish fed vegetable oils as alternative lípid sources to fish oil, in: Turchini, G.M., Ng, W.K., Tocher, D.R. (Eds.), *Fish Oil Replacements and Alternative Lipid Sources in Aquaculture Feeds*. CRC Press, Boca Raton, pp. 439–485.

Montero, D., Kalinowski, T., Obach, A., Robaina, L., Tort, L., Caballero, M.J., Izquierdo, M.S., 2003. Vegetable lipid sources for gilthead seabream (*Sparus aurata*): effects on fish health. Aquaculture 225, 353–370.

Moon, Y.A., Hammer, R.E., Horton, J.D., 2009. Deletion of ELOVL5 leads to fatty liver through activation of SREBP-1c in mice. J. Lipid Res. 50, 412–423.

Moore, S.A., 2001. Polyunsaturated fatty acid synthesis and release by brain-derived cells in vitro. J. Mol. Neurosci. 16, 195–200.

Morais, S., Monroig, O., Zheng, X.Z., Leaver, M.J., Tocher, D.R., 2009. Highly unsaturated fatty acid synthesis in Atlantic salmon: Characterization of ELOVL5-and ELOVL2-like elongases. Mar. Biotechnol. 11, 627–639.

Morton, N.E., 1955. Sequential tests for the detection of linkage. Am. J. Hum. Genet. 7, 277–318.

Mourente, G., Bell, J.G., 2006. Partial replacement of dietary fish oil with blends of vegetable oils (rapeseed, linseed and palm oils) in diets for European sea bass (*Dicentrarchus labrax* L.) over a long term growth study: effects on muscle and liver fatty acid composition and effectiv. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 145, 389–399.

Mourente, G., Dick, J.R., Bell, J.G., Tocher, D.R., 2005. Effect of partial substitution of dietary fish oil by vegetable oils on desaturation and oxidation of [1-14C]18:3n-3 (LNA) and [1-14C] 20:5n-3 (EPA) in hepatocytes and enterocytes of European sea bass (*Dicentrarchus labrax* L.). Aquaculture 248, 173–186.

Mourente, G., Tocher, D.R., 1994. In vivo metabolism of [1-14C]linolenic acid (18:3(n-3)) and [1-14C]eicosapentaenoic acid (20:5(n-3)) in a marine fish: time-course of the desaturation/elongation pathway. Biochim. Biophys. Acta 1212, 109–118.

Muotri, A.R., Marchetto, M.C.N., Coufal, N.G., Gage, F.H., 2007. The necessary junk: new functions for transposable elements. Hum. Mol. Genet. 16, 159–167.

Näär, A.M., Beaurang, P.A., Robinson, K.M., Oliner, J.D., Avizonis, D., Scheek, S., Zwicker, J., Kadonaga, J.T., Tjian, R., 1998. Chromatin, TAFs, and a novel multiprotein coactivator are required for synergistic activation by Sp1 and SREBP-1a in vitro. Genes Dev. 12, 3020–3031.

Nakamura, M.T., Nara, T.Y., 2004. Structure, function, and dietary regulation of delta6, delta5, and delta9 desaturases. Annu. Rev. Nutr. 24, 345–376.

Nara, T.Y., He, W.S., Tang, C., Clarke, S.D., Nakamura, M.T., 2002. The E-box like sterol regulatory element mediates the suppression of human Delta-6 desaturase gene by highly unsaturated fatty acids. Biochem. Biophys. Res. Commun. 296, 111–117.

Naruse, K., Tanaka, M., Mita, K., Shima, A., Postlethwait, J., Mitani, H., 2004. A medaka gene map: the trace of ancestral vertebrate proto-chromosomes revealed by comparative gene mapping. Genome Res. 14, 820–828.

Navarro, J.C., McEvoy, L.A., Amat, F., Sargent, J.R., 1995. Effects of diet on fatty acid composition of body zones in larvae of the sea bass *Dicentrarchus labrax*: a chemometric study. Mar. Biol. 124, 177–183.

Nei, M., Gojobori, T., 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418–426.

Nekrutenko, A., Li, W.H., 2001. Transposable elements are found in a large number of human protein-coding genes. Trends Genet. 17, 619–621.

Nelson, J.S., 2006. *Fishes of the World*. John Wiley & Sons, New York.

Neuringer, M., Connor, W.E., Lin, D.S., Barstad, L., Luck, S., 1986. Biochemical and functional effects of prenatal and postnatal omega 3 fatty acid deficiency on retina and brain in rhesus monkeys. Proc. Natl. Acad. Sci. U. S. A. 83, 4021–4025.

Newcomer, M.E., Liljas, A., Sundelin, J., Rask, L., Peterson, P.A., 1984. Crystallization of and preliminary X-ray data for the plasma retinol binding protein. J. Biol. Chem. 259, 5230–5232.

Nishihara, H., Smit, A.F.A., Okada, N., 2006. Functional noncoding sequences derived from SINEs in the mammalian genome. Genome Res. 16, 864–874.

Nolte, R.T., Wisely, G.B., Westin, S., Cobb, J.E., Lambert, M.H., Kurokawa, R., Rosenfeld, M.G., Willson, T.M., Glass, C.K., Milburn, M.V., 1998. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. Nature 395, 137–143.

NRC, 2011. Nutrient requirements of fish and shrimp, National Academies Press. National Research Council (NRC), Washington D.C.

Ohno, S., 1970. *Evolution by Gene Duplication*. Springer, Berlin.

Ota, T., Nei, M., 1994. Variance and covariances of the numbers of synonymous and nonsynonymous substitutions per site. Mol. Biol. Evol. 11, 613–619.

Ott, J., 2001. Major strengths and weaknesses of the LOD score method. Adv. Genet. 42, 125–132.

Otto, S.P., Whitton, J., 2000. Polyploid incidence and evolution. Annu. Rev. Genet. 34, 401–437.

Ou, J., Tu, H., Shan, B., Luk, A., DeBose-Boyd, R.A., Bashmakov, Y., Goldstein, J.L., Brown, M.S., 2001. Unsaturated fatty acids inhibit transcription of the sterol regulatory element-binding protein-1c (SREBP-1c) gene by antagonizing ligand-dependent activation of the LXR. Proc. Natl. Acad. Sci. U. S. A. 98, 6027–6032.

Palti, Y., Genet, C., Gao, G., Hu, Y., You, F.M., Boussaha, M., Rexroad, C.E., Luo, M.C., 2012. A second generation integrated map of the rainbow trout (*Oncorhynchus mykiss*) genome: analysis of conserved synteny with model fish genomes. Mar. Biotechnol. 14, 343–357.

Parisod, C., Alix, K., Just, J., Petit, M., Sarilar, V., Mhiri, C., Ainouche, M., Chalhoub, B., Grandbastien, M.A., 2009. Impact of transposable elements on the organization and function of allopolyploid genomes. New Phytol. 186, 37–45.

Peng, Y., Schwarz, E.J., Lazar, M.A, Genin, A., Spinner, N.B., Taub, R., 1997. Cloning, human chromosomal assginment, and adipose and hepatic expression of the CL-6/INSIG1 gene. Genomics 43, 278–284.

Pevsner, J., 2009. *Bioinformatics and Functional Genomics*. John Wiley & Sons, Hoboken.

Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29, e45.

Pfaffl, M.W., Horgan, G.W., Dempfle, L., 2002. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res. 30, e36.

Phillips, B.R., Keatley, A.K., Morasch, R.M., Ventura, B.A., Lubieniecki, P.K., Koop, F.B., Danzmann, G.R., Davidson, S.W., 2009. Assignment of Atlantic salmon (*Salmo salar*) linkage groups to specific chromosomes: Conservation of large syntenic blocks corresponding to whole chromosome arms in rainbow trout (*Oncorhynchus mykiss*). BMC Genet. 10, 46.

Poff, N.L., Ward, J.V, 1990. Physical habitat template of lotic systems: recovery in the context of historical pattern of spatiotemporal heterogeneity. Environ. Manage. 14, 629–645.

Pray, L., 2008. Transposons, or jumping genes: Not junk DNA? Nat. Educ. 1.

Qian, W., Zhang, J., 2008. Gene dosage and gene duplicability. Genetics 179, 2319–1324.

Qin, Y., Dalen, K.T., Gustafsson, J.A., Nebb, H.I., 2009. Regulation of hepatic fatty acid elongase 5 by LXRalpha-SREBP-1c. Biochim. Biophys. Acta 1791, 140–147.

Ramsden, S.D., Brinkmann, H., Hawryshyn, C.W., Taylor, J.S., 2003. Mitogenomics and the sister of Salmonidae. Trends Ecol. Evol. 18, 607–610.

Rawson, R.B., 2003. The SREBP pathway - insights from Insigs and insects. Nat. Rev. Mol. Cell Biol. 4, 631–640.

Repa, J.J., Liang, G., Ou, J., Bashmakov, Y., Lobaccaro, J.M., Shimomura, I., Shan, B., Brown, M.S., Goldstein, J.L., Mangelsdorf, D.J., 2000a. Regulation of mouse sterol regulatory element-binding protein-1c gene (SREBP-1c) by oxysterol receptors, LXRalpha and LXRbeta. Genes Dev. 14, 2819–2830.

Repa, J.J., Turley, S.D., Lobaccaro, J.A., Medina, J., Li, L., Lustig, K., Shan, B., Heyman, R.A., Dietschy, J.M., Mangelsdorf, D.J., 2000b. Regulation of absorption and ABC1-mediated efflux of cholesterol by RXR heterodimers. Science 289, 1524–1529.

Reschly, E.J., Ai, N., Welsh, W.J., Ekins, S., Hagey, L.R., Krasowski, M.D., 2008. Ligand specificity and evolution of liver X receptors. J. Steroid Biochem. Mol. Biol. 110, 83–94.

Robin, J.H., Regost, C., Arzel, J., Kaushik, S.J., 2003. Fatty acid profile of fish following a change in dietary fatty acid source: model of fatty acid composition with a dilution hypothesis. Aquaculture 225, 283–293.

Romano, L.A., Wray, G.A., 2003. Conservation of Endo16 expression in sea urchins despite evolutionary divergence in both cis and trans-acting components of transcriptional regulation. Development 130, 4187–4199.

Romero, I.G., Ruvinsky, I., Gilad, Y., 2012. Comparative studies of gene expression and the evolution of gene regulation. Nat. Rev. Genet. 13, 505–516.

Rosenlund, G., Obach, A., Sandberg, M.G., Standal, H., Tveit, K., 2001. Effect of alternative lipid sources on long-term growth performance and quality of Atlantic salmon (*Salmo salar* L.). Aquac. Res. 32, 323–328.

Rost, B., Yachdav, G., Liu, J., 2004. The PredictProtein server. Nucleic Acids Res. 32, 321–326.

Roth, C., Rastogi, S., Arvestad, L., Dittmar, K., Light, S., Ekman, D., Liberles, D.A., 2007. Evolution after gene duplication: models, mechanisms, sequences, systems, and organisms. J. Exp. Zool. B. Mol. Dev. Evol. 308, 58–73.

Rozen, S., Skaletsky, J.H., 1998. Primer3 [http://bioinfo.ut.ee/primer3/].

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sanderson, L.M., Kersten, S., 2010. PPARs: important regulators in metabolism and inflammation, in: Bunce, C.M., Campbell, M.J. (Eds.), *Nuclear Receptors: Current Concepts and Future Challenges*. Springer, Dordrecht, pp. 259–285.

Sargent, J., Bell, G., McEvoy, L., Tocher, D., Estevez, A., 1999. Recent developments in the essential fatty acid nutrition of fish. Aquaculture 177, 191–199.

Sargent, J.R., Bell, J.G., Bell, M. V, Henderson, R.J., Tocher, D.R., 1995. Requirement criteria for essential fatty acids. J. Appl. Ichthyol. 11, 183–198.

Sargent, J.R., Tocher, D.R., Bell, J.G., 2002b. The lipids, in: Halver, J.E. (Ed.), *Fish Nutrition*. Academic Press, San Diego, California, pp. 181–257.

Sato, Y., Nishida, M., 2010. Teleost fish with specific genome duplication as unique models of vertebrate evolution. Environ. Biol. Fishes 88, 169–188.

Schultz, J.R., Tu, H., Luk, A., Repa, J.J., Medina, J.C., Li, L., Schwendner, S., Wang, S., Thoolen, M., Mangelsdorf, D.J., Lustig, K.D., Shan, B., 2000. Role of LXRs in control of lipogenesis. Genes Dev. 14, 2831–2838.

Schwalme, K., 1992. A quantitative comparison between diet and body fatty acid composition in wild northern pike (*Esox lucius* L.). Fish Physiol. Biochem. 10, 91–98.

Schwenk, R.W., Holloway, G.P., Luiken, J.J.F.P., Bonen, A., Glatz, J.F.C., 2010. Fatty acid transport across the cell membrane: regulation by fatty acid transporters. Prostaglandins, Leukot. Essent. Fat. Acids 82, 149–154.

Sémon, M., Wolfe, K.H., 2007. Reciprocal gene loss between Tetraodon and zebrafish after whole genome duplication in their ancestor. Trends Genet. 23, 108–112.

Seoighe, C., Wolfe, K.H., 1999. Updated map of duplicated regions in the yeast genome. Gene 238, 253–261.

Sha, Z., Xu, P., Takano, T., Liu, H., Terhune, J., Liu, Z., 2008. The warm temperature acclimation protein Wap65 as an immune response gene: Its duplicates are differentially regulated by temperature and bacterial infections. Mol. Immunol. 45, 1458–1469.

Shapawi, R., Mustafa, S., Ng, W.K., 2008. Effects of dietary fish oil replacement with vegetable oils on growth and tissue fatty acid composition of humpback grouper, *Cromileptes altivelis* (Valenciennes). Aquac. Res. 39, 315–323.

Shiina, T., Dijkstra, J.M., Shimizu, S., Watanabe, A., Yanagiya, K., Kiryu, I., Fujiwara, A., Nishida-Umehara, C., Kaba, Y., Hirono, I., *et al*., 2005. Interchromosomal duplication of major histocompatibility complex class I regions in rainbow trout (*Oncorhynchus mykiss*), a species with a presumably recent tetraploid ancestry. Immunogenetics 56, 878–893.

Shimano, H., 2001. Sterol regulatory element-binding proteins (SREBPs): transcriptional regulators of lipid synthetic genes. Prog. Lipid Res. 40, 439–452.

Simopoulos, A.P., 2008. The importance of the omega-6/omega-3 fatty acid ratio in cardiovascular disease and other chronic diseases. Exp. Biol. Med. 233, 674–688.

Skrabanek, L., Wolfe, K.H., 1998. Eukaryote genome duplication - where's the evidence? Curr. Opin. Genet. Dev. 8, 694–700.

Smale, S.T., Kadonaga, J.T., 2003. The RNA polymerase II core promoter. Annu. Rev. Biochem. 72, 449–479.

Smit, A.F., 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr. Opin. Genet. Dev. 9, 657–663.

Smit, A.F.A., Hubley, R., Green, P., 2010. RepeatMasker Open 3.0 [http://www.repeatmasker.org].

Soltis, D.E., Albert, V.A., Leebens-Mack, J., Bell, C.D., Paterson, A.H., Zheng, C., Sankoff, D., Depamphilis, C.W., Wall, P.K., Soltis, P.S., 2009. Polyploidy and angiosperm diversification. Am. J. Bot. 96, 336–348.

Son, Y.L., Lee, Y.C., 2005. Ligand-binding regulation of LXR/RXR and LXR/PPAR heterodimerizations : SPR technology-based kinetic analysis correlated with molecular dynamics simulation. Protein Sci. 14, 812–822.

Son, Y.L., Lee, Y.C., 2009. Molecular determinants of the interactions between LXR/RXR heterodimers and TRAP220. Biochem. Biophys. Res. Commun. 384, 389–393.

Sprecher, H., 2000. Metabolism of highly unsaturated n-3 and n-6 fatty acids. Biochim. Biophys. Acta 1486, 219–231.

Stam, P., 1993. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J. 3, 739–744.

Stryke, D., Kawamoto, M., Huang, C.C., Johns, S.J., King, L.A., Harper, C.A., Meng, E.C., Lee, R.E., Yee, A., L'Italien, L., *et al*., 2003. BayGenomics: a resource of insertional mutations in mouse embryonic stem cells. Nucleic Acids Res. 31, 278–281.

Sushchik, N.N., Gladyshev, M.I., Kalachova, G.S., Makhutova, O.N., Ageev, A.V, 2006. Comparison of seasonal dynamics of the essential PUFA contents in benthic invertebrates and grayling *Thymallus arcticus* in the Yenisei river. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 145, 278–287.

Széles, L., Póliska, S., Nagy, G., Szatmari, I., Szanto, A., Pap, A., Lindstedt, M., Santegoets, S.J.A.M., Rühl, R., Dezsö, B., *et al*., 2010. Research resource: transcriptome profiling of genes regulated by RXR and its permissive and nonpermissive partners in differentiating monocyte-derived dendritic cells. Mol. Endocrinol. 24, 2218–2231.

Taggart, J.B., Ferguson, A., 1984. Allozyme variation in the brown trout (*Salmo trutta* L.): Single locus and joint segregation inheritance studies. Heredity 53, 339–359.

Tajima, F., 1993. Simple methods for testing molecular clock hypothesis. Genetics 135, 599–607.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.

Tamura, K., Makino, A., Hullin-Matsuda, F., Kobayashi, T., Furihata, M., Chung, S., Ashida, S., Miki, T., Fujioka, T., Shuin, T., *et al*., 2009. Novel lipogenic enzyme ELOVL7 is involved in prostate cancer growth through saturated long-chain fatty acid metabolismo. Cancer Res. 69, 8133–8140.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512–526.

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

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

Thorgaard, G.H., Bailey, G.S., Williams, D., Buhler, D.R., Kaattari, S.L., Ristow, S.S., Hansen, J.D., Winton, J.R., Bartholomew, J.L., Nagler, J.J., 2002. Status and opportunities for genomics research with rainbow trout. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 133, 609–646.

Tocher, D.R., 2003. Metabolism and functions of lipids and fatty acids in teleost fish. Rev. Fish. Sci. 11, 107–184.

Tocher, D.R., Agaba, M., Hastings, N., Bell, J.G., Dick, J.R., Teale, A.J., 2002. Nutritional regulation of hepatocyte fatty acid desaturation and polyunsaturated fatty acid composition in zebrafish (*Danio rerio*) and tilapia (*Oreochromis niloticus*). Fish Physiol. Biochem. 24, 309–320.

Tocher, D.R., Bell, J.G., MacGlaughlin, P., McGhee, F., Dick, J.R., 2001. Hepatocyte fatty acid desaturation and polyunsaturated fatty acid composition of liver in salmonids: effects of dietary vegetable oil. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 130, 257–270.

Tocher, D.R., Bell, J.G., McGhee, F., Dick, J.R., Fonseca-Madrigal, J., 2003. Effects of dietary lipid level and vegetable oil on fatty acid metabolism in Atlantic salmon (S*almo salar*) over teh entire production cycle. Fish Physiol. Biochem. 29, 193–209.

Tocher, D.R., Leaver, M.J., Hodgson, P.A., 1998. Recent advances in the biochemistry and molecular biology of fatty acyl desaturases. Prog. Lipid Res. 37, 73–117.

Tocher, D.R., Zheng, X., Schlechtriem, C., Hastings, N., Dick, J.R., Teale, A.J., 2006. Highly unsaturated fatty acid synthesis in marine fish: cloning, functional characterization, and nutritional regulation of fatty acyl delta 6 desaturase of Atlantic cod (*Gadus morhua* L.). Lipids 41, 1003–1016.

Tontonoz, P., Kim, J.B., Graves, R.A., Spiegelman, B.M., 1993. ADD1: a novel helix-loop-helix transcription factor associated with adipocyte determination and differentiation. Mol. Cell. Biol. 13, 4753–4759.

Torstensen, B.E., Bell, J.G., Rosenlund, G., Henderson, R.J., Graff, I.E., Tocher, D.R., Lie, Ø., Sargent, J.R., 2005. Tailoring of atlantic salmon (*Salmo salar* L.) flesh lipid composition and sensory quality by replacing fish oil with a vegetable oil blend. J. Agric. Food Chem. 53, 10166–10178.

Torstensen, B.E., Lie, O., Froyland, L., 2000. Lipid metabolism and tissue composition in Atlantic salmon (*Salmo salar* L.) - Effects of capelin oil, palm oil, and oleic acid-enriched sunflower oil as dietary lipid sources. Lipids 2000 35, 653–664.

Tóth, K., Sarang, Z., Scholtz, B., Brázda, P., Ghyselinck, N., Chambon, P., Fésüs, L., Szondy, Z., 2011. Retinoids enhance glucocorticoid-induced apoptosis of T cells by facilitating glucocorticoid receptor-mediated transcription. Cell Death Differ. 18, 783–792.

Tu, W.C., Muhlhausler, B.S., James, M.J., Stone, D.A.J., Gibson, R.A., 2012. An alternative n-3 fatty acid elongation pathway utilising 18:3n-3 in barramundi (*Lates calcarifer*). Biochem. Biophys. Res. Commun. 423, 176–182.

Turchini, G.M., Ng, W.K., Tocher, D.R., 2011. *Fish oil replacements and alternative lipid sources in aquaculture feeds*. CRC Press, Boca Raton.

Turchini, G.M., Torstensen, B.E., Ng, W.K., 2009. Fish oil replacement in finfish nutrition. Rev. Aquac. 1, 10–57.

Tvrdik, P., Westerberg, R., Silve, S., Asadi, A., Jakobsson, A., Cannon, B., Loison, G., Jacobsson, A., 2000. Role of a new mammalian gene family in the biosynthesis of very long chain fatty acids and sphingolipids. J. Cell Biol. 149, 707–717.

Urquiza, A.M.D., 2000. Docosahexaenoic acid, a ligand for the retinoid X receptor in mouse brain. Science 290, 2140–2144.

Wainwright, P.E., 2002. Dietary essential fatty acids and brain function: a developmental perspective on mechanisms. Proccedings Nutr. Soc. 61, 61–69.

Wang, C., Harris, W.S., Chung, M., Lichtenstein, A.H., Balk, E.M., Kupelnick, B., Jordan, H.S., Lau, J., 2006. n-3 Fatty acids from fish or fish-oil supplements, but not alphalinolenic acid, benefit cardiovascular disease outcomes in primary- and secondary-prevention studies: a systematic review. Am. J. Clin. Nutr. 84, 5–17.

Wang, D., Sul, H.S., 1997. Upstream stimulatory factor binding to the E-box at -65 is required for insulin regulation of the fatty acid synthase promoter. J. Biol. Chem. 272, 26367–26374.

Wang, W., Zheng, H., Fan, C., Li, J., Shi, J., Cai, Z., Zhang, G., Liu, D., Zhang, J., Vang, S., *et al*., 2006. High rate of chimeric gene origination by retroposition in plant genomes. Plant Cell 18, 1791–1802.

Wang, X., Weigel, D., Smith, L.M., 2013. Transposon variants and their effects on gene expression in Arabidopsis. PLos Genet. 9, e1003255.

Wang, Y., Botolin, D., Christian, B., Busik, J., Xu, J., Jump, D.B., 2005. Tissue-specific, nutritional, and developmental regulation of rat fatty acid elongases. J. Lipid Res. 46, 706–715.

Wang, Y., Botolin, D., Xu, J., Christian, B., Mitchell, E., Jayaprakasam, B., Nair, M.G., Peters, J.M., Busik, J. V, Olson, L.K., *et al*., 2006. Regulation of hepatic fatty acid elongase and desaturase expression in diabetes and obesity. J. Lipid Res. 47, 2028–2041.

Webb, J., Verspoor, E., Aubin-Horth, N., Romakkaniemi, A., Amiro, P., 2007. The Atlantic Salmon, in: Verspoor, E., Stradmeyer, L., Nielsen, J.L. (Eds.), *The Atlantic Salmon: Genetics, Conservation and Management*. Blackwell Science Ltd, Oxford, pp. 17–56.

Wessler, S.R., 2006. Transposable elements and the evolution of eukaryotic genomes. PNAS 103, 17600–17601.

Willson, T.M., Brown, P.J., Sternbach, D.D., Henke, B.R., 2000. The PPARs: from orphan receptors to drug discovery. J. Med. Chem. 43, 527–550.

Willy, P.J., Mangelsdorf, D.J., 1997. Unique requirements for retinoid-dependent transcriptional activation by the orphan receptor LXR. Genes Dev. 11, 289–298.

Willy, P.J., Umesono, K., Ong, E.S., Evans, R.M., Heyman, R. a, Mangelsdorf, D.J., 1995. LXR, a nuclear receptor that defines a distinct retinoid response pathway. Genes Dev. 9, 1033–1045.

Wilson, P.W., D'Agostino, R.B., Parise, H., Sullivan, L., Meigs, J.B., 2005. Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. Circulation 112, 3066–3072.

Winfield, I.J., Fletcher, J.M., Ben James, J., 2012. Long-term changes in the diet of pike (*Esox lucius*), the top aquatic predator in a changing Windermere. Freshw. Biol. 57, 373–383.

Wolfe, K.H., 2001. Yesterday's polyploids and the mystery of diploidization. Nat. Rev. Genet. 2, 333–341.

Wray, G.A., 2007. The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. 8, 206–216.

Xi, H., Yu, Y., Fu, Y., Foley, J., Halees, A., Weng, Z., 2007. Analysis of overrepresented motifs in human core promoters reveals dual regulatory roles of YY1. Genome Res. 17, 798–806.

Yaakov, B., Kashkush, K., 2011. Methylation, transcription, and rearrangements of transposable elements in synthetic allopolyploids. Int. J. Plant Genomics 2011, 569826.

Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.M., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155, 431–449.

Yokoyama, C., Wang, X., Briggs, M.R., Admon, A., Wu, J., Hua, X., Goldstein, J.L., Brown, M.S., 1993. SREBP-1, a basic-helix-loop-helix-leucine zipper protein that controls transcription of the low density lipoprotein receptor gene. Cell 75, 187–197.

Yoshikawa, T., Shimano, H., Yahagi, N., Ide, T., Amemiya-Kudo, M., Matsuzaka, T., Nakakuki, M., Tomita, S., Okazaki, H., Tamura, Y., *et al*., 2002. Polyunsaturated fatty acids suppress sterol regulatory element-binding protein 1c promoter activity by inhibition of liver X receptor (LXR) binding to LXR response elements. J. Biol. Chem. 277, 1705–1711.

Yue, L., Ye, F., Gui, C., Luo, H., Cai, J., Shen, J., Chen, K., Shen, X., Jiang, H., 2005. Ligand-binding regulation of LXR/RXR and LXR/PPAR heterodimerizations: SPR technology-based kinetic analysis correlated with molecular dynamics simulation. Protein Sci. 14, 812–822.

Zadravec, D., Tvrdik, P., Guillou, H., Haslam, R., Kobayashi, T., Napier, J.A., Capecchi, M.R., Jacobsson, A., 2011. ELOVL2 controls the level of n-6 28:5 and 30:5 fatty acids in testis, a prerequisite for male fertility and sperm maturation in mice. J. Lipid Res. 52, 245–255.

Zelcer, N., Tontonoz, P., 2006. Liver X receptors as integrators of metabolic and inflammatory signaling. J. Clin. Invest. 116, 607–614.

Zhang, J., 2003. Evolution by gene duplication: an update. Trends Ecol. Evol. 18, 292–298.

Zhao, C., Dahlman-Wright, K., 2010. Liver X receptor in cholesterol metabolism. J. Endocrinol. 204, 233–240.

Zhao, Q., Chasse, S.A., Devarakonda, S., Sierk, M.L., Ahvazi, B., Rastinejad, F., 2000. Structural basis of RXR-DNA interactions. J. Mol. Biol. 296, 509–520.

Zheng, X., Ding, Z., Xu, Y., Monroig, O., Morais, S., Tocher, D.R., 2009a. Physiological roles of fatty acyl desaturases and elongases in marine fish: Characterisation of cDNAs of fatty acyl Δ6 desaturase and elovl5 elongase of cobia (*Rachycentron canadum*). Aquaculture 290, 122–131.

Zheng, X., Leaver, M.J., Tocher, D.R., 2009b. Long-chain polyunsaturated fatty acid synthesis in fish: Comparative analysis of Atlantic salmon (*Salmo salar* L.) and Atlantic cod (*Gadus morhua* L.) Delta6 fatty acyl desaturase gene promoters. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 154, 255–263.

Zheng, X., Tocher, D.R., Dickson, C.A., Bell, J.G., Teale, A.J., 2004. Effects of diets containing vegetable oil on expression of genes involved in highly unsaturated fatty acid biosynthesis in liver of Atlantic salmon (*Salmo salar*). Aquaculture 236, 467–483.