Molecular Characterisation of Sex Differentiation Genes in Yellowtail Kingfish (*Seriola lalandi*)

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

A thesis
submitted in partial fulfilment
of the requirements for the degree
of

Master of Science (Research) in Biological Sciences

at

The University of Waikato

by

NICHOLAS B. ELLIOTT

2015

Abstract

Species diversification is crucial for the successful establishment and growth of a finfish aquaculture industry in New Zealand. The yellowtail kingfish *Seriola lalandi* is a promising candidate species with a fast growth rate, high market value, and is successfully farmed overseas, along with several other *Seriola* species. To truly domesticate any farmed species, understanding reproduction is vital. Sexual differentiation is a key part of early reproductive development and is characterised by the migration of primordial germ cells through the body to the site of the presumptive gonad, where they proliferate and differentiate into spermatogonia and oogonia. A number of key genes are expressed during sex differentiation, but currently very little molecular research has been conducted on *S. lalandi*.

Using bioinformatics approaches, the available gene databases and the implementation of a transcriptomic library prepared from the gonad, ovary and pituitary tissue of *S. lalandi*, a number of key sex differentiation genes were identified, which included, *Vasa, Amh* and *Cyp19a1a*. Primers were designed to each gene, to enable confirmation of the gene sequences, by PCR and RACE-PCR. Using this approach, no confirmed sequence was obtained for *Cyp19a1a*, however sequences were obtained for part of the *Vasa* gene and a large part of the *Amh* gene, including the 3'end.

The availability of these sex differentiation gene sequences allowed expression of *Vasa*, *Amh* and *Cyp19a1a* to be investigated. Primers were designed using the confirmed sequences for *Vasa* and *Amh*, as well as the unconfirmed sequence for *Cyp19a1a*. These were used in real-time PCR to examine the expression of each gene during development, within larvae at hatching, 3 days post hatch (dph), 12 dph and 18 dph. Expression of each gene was found to increase by 12 dph, but then decrease at 18dph, however, no meaningful significant difference in expression was seen statistically between any of the time points.

Lastly, a protocol was developed for embedding and sectioning of larvae from a range of different ages. A number of attempts were made to optimise an approach, with fish from 3 dph to 60 dph eventually being paraffin embedded, sectioned and stained with either toluidine blue or hematoxylin and eosin (H&E). A number of regions of interest were described, with one showing good similarity to early gonad formation, detailed in other studies. However, due to technical issues and time constraints, no PGCs or presumptive gonad tissues could be positively identified. Future work will use prepared sections to stain for the sex differentiation genes, to gain a better understanding of their role during development.

This study has established a platform for further work on *S. lalandi* by creating a framework of molecular tools for studying sex differentiation genes and their expression in this species. It has also worked towards refining the histological techniques required to successfully analyse structural development in larvae. This will be valuable for future studies with techniques such as *in-situ* hybridisation, assisting in the development of tools for the improvement of *S. lalandi* aquaculture, for applications such as monosex culture and surrogate broodstock technology.

Acknowledgements

I cannot express enough gratitude to my supervisors Simon Muncaster and Steve Bird. They have strongly supported not only this thesis, but also my scientific career and personal development. It has been fantastic to work with researchers who are inspired and motivated to improve the science surrounding aquaculture.

I must thank Jane Symonds and Alicia King of the NIWA Bream Bay Aquaculture Park for providing the larvae used in this study, and for keeping a keen interest in my progress.

The academic and technical staff throughout the University of Waikato Science Department were essential for this project. I would specifically like to thank Linda Peters and Barry O'Brien for sharing their expertise. I would also like to acknowledge Ray Cursons, Olivia Patty, Sari Karppinen, Judith Burrows and John Longmore.

Thank you to my colleagues in the lab for filling my days with humour, and for helping me to find my way in molecular biology.

Thank you to my family and close friends for your unwavering support. It has always been appreciated.

Finally I would like to acknowledge the inspiring works of Cusack, which have strongly influenced my time as a postgraduate student.

Table of Contents

List of Figures

List of Tables

List of Terms and Abbreviations

°C	degrees Celsius
11KT	11-ketotestosterone
AMH	anti-Müllerian hormone
BLAST	basic local alignment search tool
bp	base pairs
cDNA	complementary DNA
CDS	coding DNA sequence
C_T	threshold cycle
DEAD	Asp-Glu-Ala-Asp
DEPC	diethylpyrocarbonate
DNA	deoxyribose nucleic acid
dNTP	deoxynucleotide triphosphate
dpf	days post fertilisation
dph	days post hatch
DPX	di-n-butyl phthalate in xylene
E2	estradiol-17β
ENA	European Nucleotide Archive
ERE	estrogen responsive element
EtBr	3,8-diamino-5-ethyl-6-phenylphenanthridinium bromide
GM	genetically modified / genetic modification
H&E	hematoxylin and eosin
HPG axis	hypothalamic-pituitary-gonadal axis
IPTG	isopropyl beta-D-thiogalactopyranoside
LB	lysogeny broth
mRNA	messenger RNA
mQH_2O	Milli-Q ultrapure water

MT	17α-methyltestosterone
NGS	next generation sequencing
NIWA	National Institute of Water and Atmospheric Research
OCT	optimum cutting temperature solution
p-xylene	1,4-dimethylbenzene (*para*-xylene)
PBS	phosphate buffered saline
PC1	physical containment level 1
PCR	polymerase chain reaction
PFA	paraformaldehyde
PGC	primordial germ cell
qPCR	quantitative (real-time) polymerase chain reaction
RACE	rapid amplification of cDNA ends
RNA	ribose nucleic acid
RNase	ribonuclease
SOLiD	sequencing by oligonucleotide ligation and detection
TAE	tris-acetate-EDTA buffer
Taq	*Thermus aquaticus* DNA polymerase
TGFβ	transforming growth factor beta
T_m	melting temperature
Tris	(hydroxymethyl)aminomethane
TDF	testis-determining factor
TSD	temperature-dependent sex determination
TWEEN® 20	polyoxyethylene 20 sorbitan monolaurate
UniProt	Universal Protein Resource
Xgal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
YTK	yellowtail kingfish

Preface

Genetic Nomenclature

The recommended formatting of gene and protein nomenclature varies widely across species, particularly within vertebrates. This is further confounded by inconsistencies between authors in the literature. In the interests of clarity, this thesis follows the recommended formatting detailed by Mouse Genome Informatics (2014). With the exception of the housekeeping gene β-actin (*Actb*), all genes in this study are referred to in italics with first letter capitalisation (*e.g. Amh*), while proteins are non-italicised with full capitalisation (*e.g.* AMH).

Permits and Training

The University of Waikato received HSNO approval (GMD101146) from the New Zealand Environmental Protection Authority (EPA) in 2011 to develop a range of genetically modified (GM) non-pathogenic microorganisms, cell lines and zebrafish carrying genes coding for proteins involved in causation of disease, in the evolution of protein stability and cellular functions. The research in this thesis meets these requirements by developing GM organisms containing sex differentiation genes for the purposes of cloning and sequencing. The location and nature of the development and disposal of the approved GM *E. coli* were in accordance with the APP201152 application submitted to the EPA, and controls listed in the GMD101146 approval. Details can be found online at *http://www.epa.govt.nz/search-databases/Pages/applications-details.aspx?appID=APP201152*.

Training was required for numerous aspects of this project, including Physical Containment Level 1 (PC1) laboratory training at the University of Waikato, and a site and laboratory induction at the National Institute of Water and Atmospheric Research (NIWA) Bream Bay Aquaculture Park.

Chapter 1
Introduction

1.1 Background

Many of the oceans' wild fish stocks face overfishing and exploitation. As the human population exponentially increases year on year, demand for seafood increases accordingly and often vastly exceeds fisheries sustainable catch limits. This crisis of supply is being partly resolved by aquaculture, which has expanded rapidly over recent decades. While global capture fisheries have slowed to zero growth since 1980, fish aquaculture has grown by an average of 8.8% per year over the same period (FAO, 2012). Aquaculture contributed over 40% of global fish production in 2010 and has become the primary driver of total growth in the sector. Asia is currently responsible for the vast majority of this growth (89% of total aquaculture production in 2010), predominantly farming freshwater finfish species including carps and tilapia (FAO, 2012).

Freshwater fish aquaculture has grown at nearly twice the rate of fish mariculture over the last decade, reflecting the expansion of simple freshwater fish farming in developing countries. Higher value species such as salmonids and eels are typically farmed in brackish-water environments. Diadromous finfish constituted 6% of total aquaculture production in 2010, of which salmonids represent a large proportion (FAO, 2012). Atlantic salmon (*Salmo salar*) is a species of particular interest for research and development, due to its economic importance and potential use as a model species for other salmonids. Much of the growth of the aquaculture sector has been driven by market demand, environmental resource availability and favourable business circumstances (Bostock *et al.*, 2010).

Salmonids are one of a small number of high value fish species which represent a vast amount of aquaculture production globally. As a large, fast growing fish, salmonids are ideal candidates for aquaculture; the success of the industry to date is a testament to this. However, total aquaculture production for salmon has levelled off or decreased in countries such as Japan and some areas of Europe. This was recently investigated in an analysis of the Norwegian salmon aquaculture industry (Asche *et al.*, 2013), which indicated that growth had been pushed by increased demand and production, rather than an increase in efficiency and productivity. There is a need to diversify aquaculture by developing more high value species with similar growth traits to salmon, particularly in the marine environment where fish aquaculture has been under-represented.

1.2 Species Diversification

The development of new species for aquaculture offers many important benefits, such as reduced reliance on imported fish products, reduced pressure on strained wild fisheries, maximised resource utilisation, increased product diversity and economic development. This is being recognised on a global scale; for example, a five year project named DIVERSIFY is a €11.8 million European Commission funded collaborative effort between twelve European countries. The project seeks to remove the production bottlenecks restricting emerging fish species and primarily targets six species, each with different environmental requirements and culture methods. The greater amberjack (*Seriola dumerili*), Atlantic wreckfish (*Polyprion americanus*) and meagre (*Argyrosomus regius*) are of interest for warm and cool-water marine cage culture. The Atlantic halibut (*Hippoglossus hippoglossus*) is suited for cold water marine culture, while the grey mullet (*Mugil cephalus*) is a candidate for euryhaline pond culture. Pikeperch (*Sander lucioperca*) was selected for freshwater recirculation system culture (Mylonas, 2013). By using such a broad approach, the DIVERSIFY project is likely to yield significant gains for European aquaculture. However,

working with new species often presents unique challenges. Every species has different requirements and typically requires substantial research before it can be successfully and efficiently farmed. Genetic information is seldom available for new species, except in cases where similar species have already been studied. This will need to be a core focus of research to support diversification of aquaculture, particularly for countries such as New Zealand which farm only a small number of species.

1.2.1 Aquaculture in New Zealand

Three species currently dominate aquaculture in New Zealand: green-lipped mussels (*Perna canalicula*), Chinook salmon (*Oncorhynchus tshawytscha*) and Pacific oysters (*Crassostrea gigas*). Mussels are New Zealand's biggest aquaculture export product by both volume and export revenue. Chinook salmon (currently marketed as New Zealand generating total revenue of approximately NZ$13,500 per tonne in 2011. However, production volume in the same year was only 10% that of mussels (Aquaculture New Zealand, 2012). The growth of aquaculture of these three species is shown in Figure 1. Australia and Japan have particularly strict biosecurity surrounding imported King salmon) represents a high value species, but production volume was less than 15% that of mussels in 2011. Oysters have an even higher market value, seafood, which benefits New Zealand exporters; in Australia, New Zealand is the sole supplier of imported whole salmon (MBIE, 2013). Chinook salmon were successfully introduced to Canterbury rivers in New Zealand in the early 1900's, but aquaculture was stifled by public resistance until a change to legislation in the 1970's allowed the first commercial salmon farm to establish in Golden Bay. Early ocean ranching hatcheries proved ineffective when few fish returned, but sea cage ranching, where juveniles reared in freshwater hatcheries are transferred directly to saltwater cages until harvest, has proven more successful (Wassilieff, 2012).

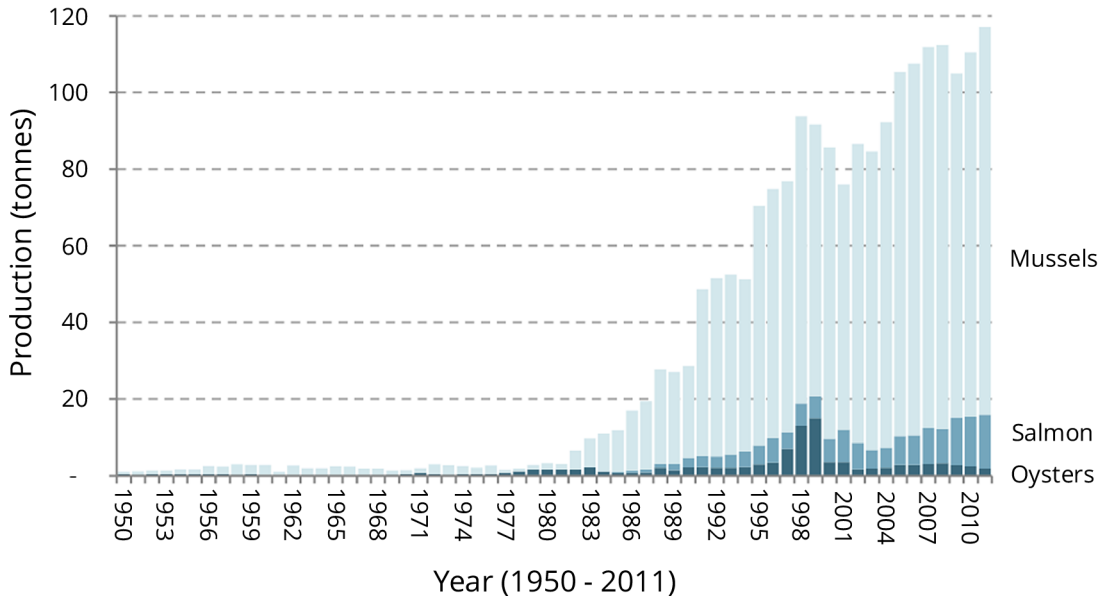


Figure 1. Total aquaculture production of green-lipped mussels, Chinook salmon and Pacific oysters in New Zealand, from 1950 to 2011. Adapted from MBIE (2013).

The need to develop more species for aquaculture has long been recognised in New Zealand. The New Zealand Aquaculture Strategy was published in 2006 which set an industry annual revenue target of $1 billion by 2025; a key focus being on innovation and commercialisation of new species (Burrell *et al.*, 2006). Yellowtail kingfish (*Seriola lalandi*) and hāpuku (*Polyprion oxygeneios*) have been identified by NIWA as suitable candidate species of finfish for aquaculture, with desirable traits such as fast growth rates, high flesh quality, and strong market opportunities both domestically and abroad (Symonds *et al.*, 2014). There are a number of other new species which have demonstrated a level of suitability for aquaculture; between NIWA, the Cawthron Institute and Plant and Food Research, notable finfish identified as candidates include butterfish (*Odax pullus*), snapper (*Pagrus auratus*), southern bluefin tuna (*Thunnus maccoyii*), New Zealand turbot (*Colistium nudipinnis*), eels (*Anguilla* spp.) and rainbow trout (*O. mykiss*) (Mussely & Goodwin, 2012). As research continues on developing new species, coupled with strong investment and changes to New Zealand's aquaculture legislation to better facilitate growth, many of these candidates may become high value products and make New Zealand much more competitive in aquaculture markets worldwide.

1.2.2 Aquaculture of *Seriola* Species

Many of the members of the *Seriola* genus are ideal candidates for aquaculture, being characterised by traits including fast growth rates, high quality nutritious white flesh and amenability to culturing, with high value product options such as sashimi and sushi (Nakada, 2008; Fernández-Palacios *et al.*, 2014; Symonds *et al.*, 2014). A number of *Seriola* species are currently farmed or being researched in countries around the world. The Japanese amberjack (*S. quinqueradiata*) is a major aquaculture product of Japan and Taiwan, while the greater amberjack (*S. dumerili*) is cultured in Japan, Saudi Arabia, Korea and in the Mediterranean (Symonds *et al.*, 2014). The yellowtail kingfish (*S. lalandi*) is farmed in Australia, Japan, Europe, and North and South America. Longfin yellowtail (*S. rivoliana*) is a species of interest for diversification in the Mediterranean (Fernández-Palacios *et al.*, 2014; Orellana *et al.*, 2014).

Global *Seriola* aquaculture is relatively stable, but faces a number of issues and limitations which will need to be overcome to ensure future stability and growth. Chiefly, these issues are disease outbreaks and reliance on wild fish stocks. In Australia for example, aquaculture of *Seriola* has had significant problems with parasitic monogean flatworms, most notably the host-specific species *Benedenia seriolae* and *Zeuxapta seriolae* (Tubbs *et al.*, 2005; Hutson *et al.*, 2007). Monogean infections are a well-documented concern in marine finfish aquaculture, with other susceptible species farmed in Australia including barramundi (*Lates calcarifer*), bream (*Pagrus major*), and mackerel (*Scomber japonicus*) (Sharp *et al.*, 2004; Militz *et al.*, 2013). *B. seriolae* has demonstrated susceptibility to sequential hydrogen peroxide baths, although their eggs remain impervious to treatment (Sharp *et al.*, 2004; Symonds *et al.*, 2014). For the Japanese *Seriola* aquaculture industry, which is estimated to produce more than US$1 billion worth of fish, management of parasitic flatworms may increase production costs by 20% or more (Rohde, 2005). This emphasises the significance of developing better parasite control treatments to reduce management costs and improve farming efficiency to make aquaculture more feasible.

The Japanese *Seriola* industry has relied heavily on wild caught juveniles (termed *Mojako* when < 50 g) and fishmeal derived from wild sardine stocks, due to both their abundance and the much higher costs associated with hatchery produced *Seriola* larvae (Nakada, 2008). Furthermore, *Seriola* hatcheries in Japan have been plagued by high mortality and deformity rates. Catch rates of *Mojako* have declined significantly since 1970 (Figure 2), prompting Japan to supplement the catch by importing *Mojako* from other countries. As demand for juvenile fish increases and supply remains limited, efficient hatcheries will need to be developed to support the industry.

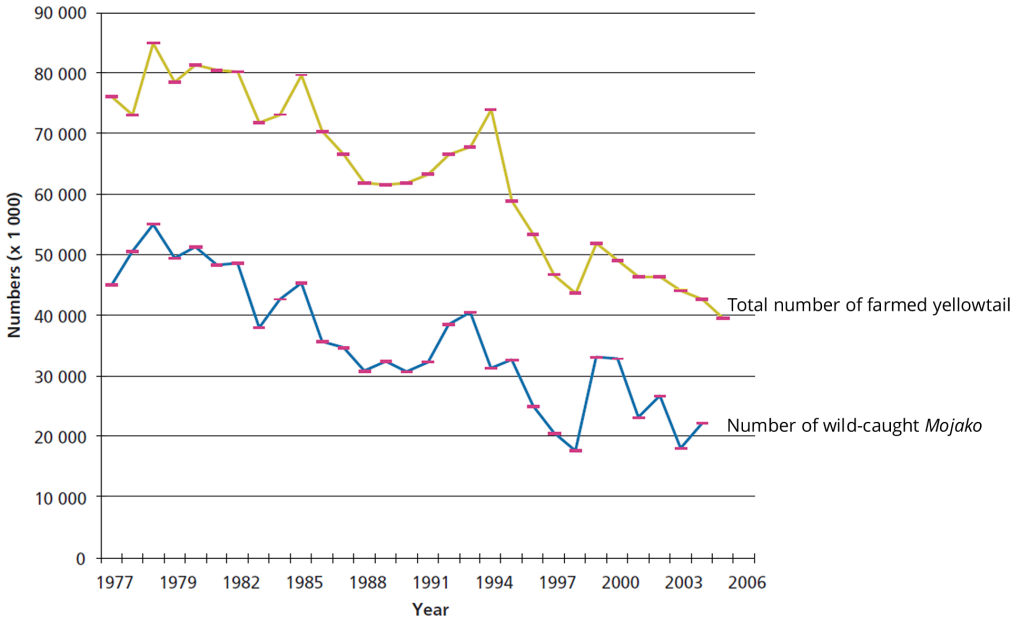


Figure 2. Temporal changes in numbers of wild-caught *S. quinqueradiata* juveniles and farm-raised individuals from 1977 to 2006. Modified from Nakada (2008), originally sourced from the Marine Aquaculture Association of Japan.

In contrast, *S. lalandi* aquaculture in Australia and New Zealand relies entirely on hatchery produced juveniles (Moran *et al.*, 2007). *S. lalandi* aquaculture research in New Zealand began in the late 1990's on Kawau Island at Moana Pacific Fisheries' Pah Farm Aquaculture facility. This facility successfully spawned wild broodstock in captivity and generated ~147,000 fertilised eggs, of which a small number of juveniles survived to be reared in sea cages. The Pah Farm *S. lalandi* operation moved to the NIWA Bream Bay Aquaculture Park in the early 2000s. NIWA has made advances in numerous areas including reproduction, broodstock conditioning, nutrition and survival (Poortenaar *et al.*, 2001; Moran *et al.*, 2009; Symonds *et al.*, 2014). As of 2014, NIWA claims to have successfully demonstrated the biological, technical and economic feasibility of *S. lalandi* farming, and has the capability to reliably produce fingerlings at commercial volumes to supply industry (Symonds *et al.*, 2014).

At present, the main factors which restrain *S. lalandi* aquaculture from establishing commercially in New Zealand are public opposition to fish farming, scarcity of investors and an absence of consolidated legislation for aquaculture. A land-based *S. lalandi* aquaculture venture was launched by the Parengarenga Incorporation in 2004, but promptly closed two years later with losses reported to exceed $15 million, citing system design flaws and insufficient capital (Barrington, 2012). Another land-based *S. lalandi* operation was planned for the Bay of Plenty by New Zealand Premium Aquaculture Limited, but resistance from local residents in 2012 led to an Environment Court hearing and was subsequently denied resource consent (Kendon, 2012). In order for an *S. lalandi* aquaculture industry to be established in New Zealand, suitable investors will need to be sought and supported by the public, the government and research institutes such as NIWA. For the industry to experience continued and accelerated growth, limitations on productivity will need to be addressed. Overcoming bottlenecks in production caused by farming practice and fish health will contribute to this aim.

1.3 Productivity Bottlenecks within Aquaculture

There are a number of factors that currently affect productivity within the aquaculture industry. Poor fish health and infectious diseases are major threats for many fish farms and hatcheries (Bondad-Reantaso *et al.*, 2005; Bricknell & Dalmo, 2005). Diets are also common limitations, with scope for species-specific optimisation through tailored nutrient profiles, improved digestibility and assimilation efficiency, as well as supplementation to address health problems (Tacchi *et al.*, 2011a; Covello *et al.*, 2012). Genetic diversity has been an issue for some species, with selective breeding approaches confounding the issue (Frost *et al.*, 2006). These bottlenecks are described in greater detail in Appendix IV.

In the context of sex-related traits, some current farming practices can restrict productivity. Often fish are grown in mixed sex cultures, but this can be inefficient for species with sexually dimorphic growth rates and differential ages of sexual maturation, such as Nile tilapia (*Oreochromis niloticus*) and European sea bass (*Dicentrarchus labrax*) (Beardmore *et al.*, 2001; Saillant *et al.*, 2001; Singh, 2013). The ability to isolate and grow the faster-growing sex in monoculture can eliminate fish which would take longer to reach marketable size, reducing production costs and improving efficiency (Singh, 2013). By improving consistency of size, this approach can also reduce cannibalism between size classes and reduce grading requirements (Devlin & Nagahama, 2002). For species in which one sex often reaches maturity precociously before reaching harvestable size, such as Atlantic cod, monosex farming can reduce the problems associated with maturation such as reduced growth rate, weight loss and reductions in flesh quality (Haugen *et al.*, 2011). Monosex farming has already started to be utilised in the case of tilapia (Beardmore *et al.*, 2001; Hafeez-Ur-Rehman *et al.*, 2008; Chakraborty & Banerjee, 2009), but the technology behind the practice is generally highly species-specific and requires research for each species of interest.

To ensure the success of the aquaculture industry in New Zealand, particularly for *S. lalandi*, research efforts need to employ modern scientific approaches to target the issues faced by aquaculture operations overseas. Disciplines of molecular biology such as genomics, transcriptomics and proteomics can offer much to aquaculture, since the acquisition of information on the molecular scale permits a deeper understanding of processes and allows for fine-scale changes to be made to farming techniques. For instance, molecular analysis of *B. seriolae* has revealed genetically distinct populations between Australia, Chile and Japan, with potentially different life cycles and characteristics (Sepúlveda & González, 2014). This has implications for aquaculture, as using an established treatment regime for a misidentified species could cause undesirable effects. The use of molecular techniques can therefore refine knowledge and lead to improved industrial solutions.

1.4 Use of Molecular Approaches in Aquaculture

Molecular studies have greatly benefitted aquaculture and the wider field of marine biology with notable advances being made in conservation, fisheries management and development of natural products (Thakur *et al.*, 2008). In aquaculture, the full genome sequences of key species are beginning to be determined. The complete 0.9 giga-base pair (Gbp) genome of the Atlantic cod was recently sequenced and immediately revealed novel information for the species, particularly regarding unique features of the immune system (Star *et al.*, 2011). Similarly, sequencing the genome of European sea bass has improved understanding of adaptations to euryhalinity in this species (Tine *et al.*, 2014). Comparative genomics can be used to relate information from known genomes to phylogenetically similar species (Sarropoulou & Fernandes, 2011). For example, the elucidation of the full Atlantic salmon genome will inevitably allow great advances in research and technology for the production of this species. This information may also provide model sequences for all salmonids (W. S. Davidson *et al.*, 2010), including Chinook salmon in the New

Zealand aquaculture industry. Such sequences will be a vital source of information for genetic analysis in a range of species.

Transcriptomic approaches in a number of studies on several fish species, including *S. lalandi*, have provided information valuable for improving productivity. For example, metamorphosis during larval development is often a bottleneck for aquaculture in many species, including the gilthead sea bream (*Sparus aurata*). The transcriptome of this species has been studied using the 454 pyrosequencing platform, revealing a significant number of transcripts with key roles in larval development (Yufera *et al.*, 2012). This will improve the understanding of the larval phase of *S. aurata*, enabling further molecular studies and possibly leading to the development of new molecular markers. The Illumina platform was used by Palstra *et al.* (2013) to resolve the transcriptome of rainbow trout skeletal muscle, with relevance to swimming-induced exercise and migration-induced sexual maturity. Immune genes were investigated in bacterially infected Japanese sea bass (*Lateolabrax japonicus*) using sequencing-based transcriptome profiling (Xiang *et al.*, 2010). This study provided an insight into marine fish immunogenetics and is highly relevant to aquaculture; the importance of understanding the immune system with respect to disease prevention, animal welfare and productivity cannot be understated. *S. lalandi* has been used as a model for a transcriptomic study of glutathione peroxidases (GPxs) in marine fish, which are important antioxidant enzymes for protecting cells against oxidative damage (Bain & Schuller, 2012). Transcriptomic studies such as these provide very precise information and may lead to productivity improvements in numerous areas including hatchery production and fish health.

Molecular studies of aquaculture conducted overseas may be relevant to New Zealand, particularly to the salmon industry. The Atlantic salmon genome sequence is likely to be an appropriate model sequence for Chinook salmon and, by using homology approaches with gene or transcript sequences, may allow many findings and advances relevant to Atlantic salmon to be applied to New Zealand farms. For

any new candidate species such as yellowtail kingfish or hāpuku, full or partial genome sequencing would be a valuable foundation to enable characterisation of important homologous genes, transcriptomic expression profiling and eventually the development of new technologies to improve farming efficiency. This should be a starting point for molecular research in these species.

It is also important to also understand pathogens of aquaculture species on the molecular scale. For instance, during 2010 a salmon farm in Queen Charlotte Sound experienced mass mortality due to a bloom of *Pseudochattonella verruculosa*. This marine flagellate algae was previously unobserved as problematic in New Zealand, but has caused problems in Atlantic salmon across Europe and in a range of species in Japan, including *Seriola* (MacKenzie *et al.*, 2011; Skjelbred *et al.*, 2011). Despite being such a widespread issue, there remains an absence of effective mitigation measures and this pathogen is still not well understood (MacKenzie *et al.*, 2011). The related *P. farcimen* is also widely problematic in aquaculture and equally poorly understood. However, ongoing molecular research has revealed an unusual fatty acid profile of this species (Dittami *et al.*, 2012), which may contribute towards developing a targeted treatment strategy with potential application for *P. verruculosa*. The acquisition of detailed molecular knowledge of pathogens and immune functioning in each species of interest to aquaculture will help to promptly develop suitable therapeutic treatments.

1.5 Next Generation Molecular Sequencing

New molecular tools have become available for genomics as second and third generation sequencing (NGS) technology has developed. Sequencing-by-synthesis is an approach used by the Illumina platform (Illumina) which can rapidly sequence hundreds of millions of base pairs by emitting fluorescence each time a nucleotide is incorporated into the sequence. The 454 pyrosequencing platform (Roche) was the

first commercial NGS system and utilises emulsion polymerase chain reaction (PCR) of DNA fragments inside microscopic beads. These are deposited into picotiter-plate wells with sequencing enzymes, where localised luminescence occurs in each well every time a nucleotide is incorporated into the sequence. The SOLiD platform (Life Technologies), Sequencing by Oligonucleotide Ligation and Detection, also uses emulsion PCR with beads. The preference of DNA ligase to match oligonucleotides generates signal information which indicates the nucleotide at each position of the sequence. The approach is highly cost effective but much slower than sequencing-by-synthesis (Sanchez-Pla *et al.*, 2012). These three platforms are summarised in Figure 3. Other established platforms include Pacific Biosciences (PacBio), HeliScope (Helicos) and Ion Torrent (Life Technologies). Each has particular advantages and disadvantages, with variations in run time (2 hours to 14 days), output (10^4 to 10^9 reads per run, 35 to over 1000 bases per read) and instrument cost (US$49,500 to $690,000) (Glenn, 2011). Each platform is optimal for particular applications, due to differences in the error rates, error types and biases (Quail *et al.*, 2012). The development of these tools has revolutionised molecular biology and enabled genomes to be studied in much more detail.

Transcriptomics focuses on RNA (particularly mRNA) transcripts as indicators of gene expression. Unlike genomes, transcriptomes are dynamic and change composition according to which genes are being expressed at a given time. Analysis of transcriptomic data can provide useful information regarding the functioning of genes. In particular, the characterisation of genes expressed during key processes (*e.g.* immune responses, sex differentiation, spawning) can lead to the recognition of functional pathways. Manipulating such pathways will potentially lead to improvements in aquaculture. Transcriptomics has been characterised by the use of expression microarrays since the 1980s, but these are limited to profiling only previously identified genes. Serial analysis of gene expression (SAGE), and its later derivative SuperSAGE, are sequence-based techniques which do not share this limitation and are widely used. However, these technologies are quickly being

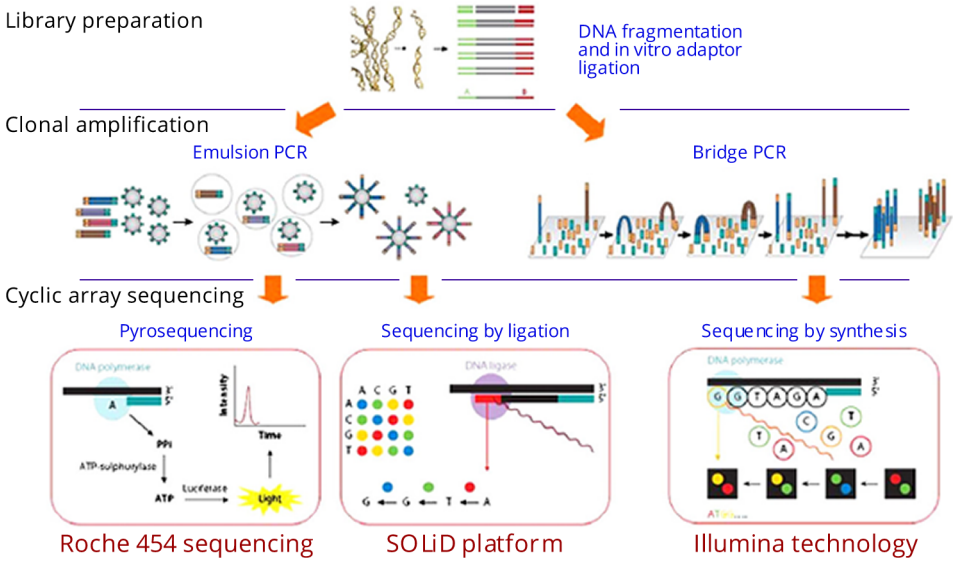


Figure 3. Modified from Sanchez-Pla *et al.* (2012), this diagram highlights the differences between three major NGS platforms. All three require initial DNA fragmentation and ligation of adaptors. PCR is used to amplify fragments prior to sequencing. Each platform uses a different sequencing process, but all three use an enzymatic reaction to generate luminescent signals from which the sequence can be determined.

displaced by next generation sequencing methods with significantly higher throughputs, such as RNA-seq (McGettigan, 2013). Transcript sequencing requires RNA to be converted to a cDNA fragment library to increase stability, allowing any of the next generation DNA sequencing platforms mentioned previously to be utilised. The benefits of superior transcript sequencing technology are wide reaching, including the ability to generate more detail of complexities indeterminable by microarrays and the sequencing of non-model organisms (McGettigan, 2013).

1.6 Sex Differentiation

The ability to control the sex and maturation of captive fish offers benefits and solutions to numerous production issues in aquaculture. Such problems often include skewed sex ratios, reproductive precocity, reproduction within stock and

negative effects of sexual maturation on product quality (Piferrer *et al.*, 2012). Many species have sexually dimorphic growth rates; the faster growing sex may thereby offer an advantage during growout. Females are the faster growing sex in many species, such as European sea bass (Saillant *et al.*, 2001; Díaz *et al.*, 2013), turbot (*Scophtalmus maximus*) (Imsland *et al.*, 1997), Senegalese sole (*Solea senegalensis*) (Sánchez *et al.*, 2010), Eurasian perch (*Perca fluviatilis*) (Rougeot *et al.*, 2002) and Atlantic halibut (*Hippoglossus hippoglossus*) (Norberg *et al.*, 1991). However, in species such as Nile tilapia and channel catfish (*Ictalurus punctatus*), males grow faster and more uniformly than females (Goudie *et al.*, 1994; Bwanika *et al.*, 2007; Chakraborty & Banerjee, 2009). Identifying and selecting for a beneficial sex requires thorough knowledge of the biology of a species. Understanding sex differentiation and puberty at the genetic level may enable control of sexual maturation and sex ratios, which will ultimately improve farm efficiency.

Reproduction and sex determination strategies are highly diverse between fish taxa. Fish can be broadly defined as gonochoristic or hermaphroditic, between which exists a spectrum of reproductive strategies. Gonochoristic species have genetically or environmentally determined sexes; individuals develop into one sex and remain as that sex for the duration of their life. In contrast, hermaphroditic species are capable of producing both male and female gametes during their lifespan. Synchronous hermaphrodites are able to produce both male and female gametes simultaneously, whereas sequential hermaphrodites produce one gamete before undergoing sex reversal to produce the other. Sequential hermaphrodites which first mature as males are termed protandrous, such as the gilthead sea bream (Haffray *et al.*, 2005). Protogynous hermaphrodites begin as females, such as Ballan wrasse (*Labrus bergylta*) and butterfish (Trip *et al.*, 2011; Muncaster *et al.*, 2013). This diversity raises potential challenges for aquaculture, making it essential to understand the reproductive biology of each species.

Sex determination refers to the genetic mechanisms controlling the sex of each individual, whereas sex differentiation refers to the period where physiologically

sexless fish develop gonadal tissues. Within gonochorists, sex determination systems vary. The familiar male heterogametic karyotype XX/XY is common in fish, including Nile tilapia and Eurasian perch (Rougeot *et al.*, 2002; Rougeot *et al.*, 2008). Species such as turbot and Japanese amberjack are suspected to have female heterogamety (ZZ/ZW) (Haffray *et al.*, 2009; Fuji *et al.*, 2010). It is likely that this will be the case for other *Seriola* species, including *S. lalandi*. Sex differentiation also varies within gonochorists. Some species develop testes or ovaries directly from an undifferentiated gonad (*e.g.* European sea bass, common carp), while others may initially develop ovarian tissue before half the population subsequently changes to produce testes (*e.g.* tiger barb (*Barbus tetrazona*), zebrafish) (Devlin & Nagahama, 2002). The latter is termed false hermaphroditism and should not be confused with protogyny, but rather as an intersex phase (Penman & Piferrer, 2008). Secondary gonochorists first develop an entirely intersex gonad before differentiating into normal testes and ovaries, such as in the Nassau groper (*Epinephelus striatus*) (Sadovy & Colin, 1995). This reiterates the importance of having a fine-scale understanding of sex determination for each species for complete domestication.

The reproductive systems of fish are generally similar to other vertebrates. Sex differentiation occurs early during development and is defined by the initiation of male and female gonad formation. Early gonad development begins during the gastrula phase of fertilised eggs. A region of specialised cells, termed somatic gonadal precursor cells (SGPCs), forms a germinal ridge which eventually gives rise to the gonad. SGPCs differentiate into Sertoli cells in males, which form the testes tubules and play an important role in spermatogenesis. Conversely, SGPCs in females become ovarian granulosa cells, which later envelop oocytes to become ovarian follicles (Siegfried, 2010). Primordial germ cells (PGCs) form from specialised granular cytoplasm and migrate to the presumptive gonad, where they differentiate into spermatogonia or oogonia (Figure 4) (Devlin & Nagahama, 2002; Presslauer *et al.*, 2012; Viñas *et al.*, 2013). It is these cells which give rise to sperm and oocytes respectively following puberty.

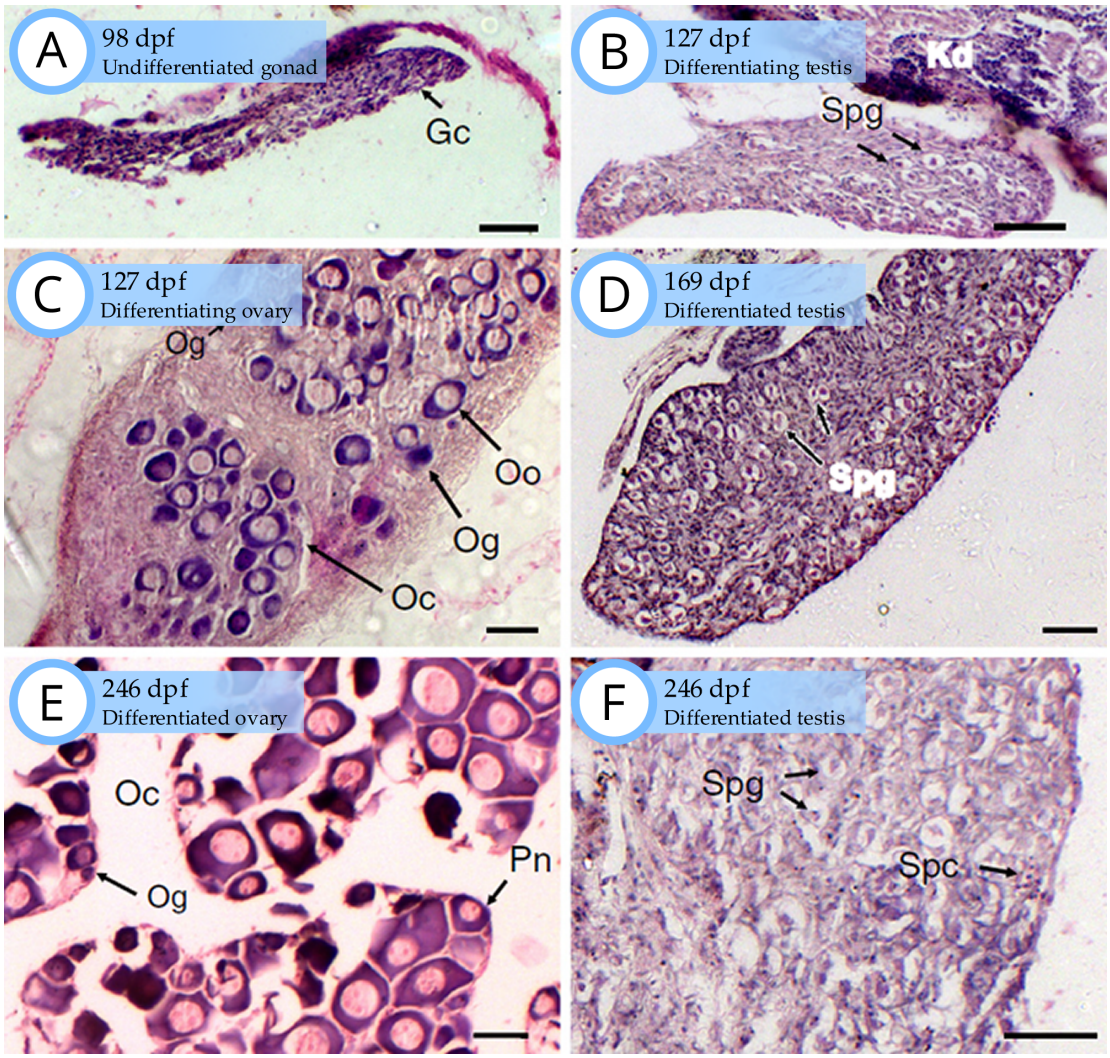


Figure 4. Photomicrographs of Senegalese sole (*Solea senegalensis*) gonad sections during sex differentiation, adapted from Viñas et al. (2013). Oogonia (Og) and spermatogonia (Spg) are visible in each sex from 127 days post fertilisation (dpf) (images B, C). Also of note are the ovarian cavity (Oc), oocytes (Oo) and perinuclear stage oocytes (Pn). Spermatocytes (Spc) can be observed in testes at 246 dpf (image F).

The reproductive system remains largely dormant until later in the life cycle when puberty occurs (Taranger *et al.*, 2010). Activation of the brain-pituitary-gonad (BPG) axis marks the onset of puberty and is characterised by distinctive hormonal changes. Neurons in the brain secrete gonadotropin-releasing hormone (GnRH), which stimulates the pituitary to produce two gonadotropins: luteinising hormone (LH) and follicle stimulating hormone (FSH). LH and FSH are released into the circulatory

system and bind to membrane receptors on the gonad (LH-R, FSH-R). This triggers the gonad to produce sex hormones which initiate and regulate gametogenesis. Estradiol-17β (E2) is the primary estrogen in most fish, prevalent in females and responsible for driving ovarian development. 11-ketotestosterone (11KT) is the primary androgen in most fish, predominantly found in males where it induces testicular development (Devlin & Nagahama, 2002; Weltzien *et al.*, 2004; Schulz *et al.*, 2010). Some aspects vary between species, such as timing of steroid release and effects of environmental cues (*e.g.* photoperiod, temperature). The genes and proteins active in these processes can be used as molecular markers to obtain a fine-scale understanding of reproductive physiology. Improving the understanding of reproductive development of a species from sexual differentiation to sexual maturity will help enable farmers to control reproduction and improve efficiency.

1.7 Applications of Sex Differentiation to Aquaculture

Gonad structure can be affected by the endocrine environment, and vice versa. This relationship can have important applications for aquaculture. For instance, manipulating E2 levels is particularly useful for generating monosex cultures. E2 is produced *via* aromatase catalysis of testosterone and stimulates oocyte development, therefore having a pivotal role in early ovarian differentiation. Estrogen treatment experiments have generated fully feminised fish in species such as Japanese black porgies (*Acanthopagrus schlegelii*) and turbot (Devlin & Nagahama, 2002; Haffray *et al.*, 2009). Conversely, aromatase inhibitors (*e.g.* fadrozole, exemestane) can be used to block E2 production and cause masculinisation of fish. This has been demonstrated on Atlantic halibut, Nile tilapia, medaka (*Oryzias latipes*) and numerous others (Devlin & Nagahama, 2002; Guiguen *et al.*, 2010; Ruksana *et al.*, 2010; Siegfried, 2010; Babiak *et al.*, 2012). The best time to administer steroid treatments is during sex differentiation; in most cases, these experiments have yielded 100% feminisation or masculinisation.

Masculinisation can also be achieved by treating with androgens. The synthetic 17α-methyltestosterone (MT) is commonly used for this purpose as it is non-aromatisable and cannot be converted to E2. It has shown a high degree of success in a wide range of species including Atlantic cod, Nile tilapia, coho salmon and turbot (Fitzpatrick *et al.*, 2005; Chakraborty & Banerjee, 2009; Haffray *et al.*, 2009; Haugen *et al.*, 2011). 11KT can also be effective at inducing masculinisation of females but may require substantially higher doses, as appeared to be the case in Siamese fighting fish (*Betta splendens*) (Kavumpurath & Pandian, 1994). This is likely to be due to differences in steroid receptor affinities or complexes (Devlin & Nagahama, 2002). The effectiveness of androgen-induced masculinisation to reliably produce neo-males has seen the practice widely adopted into aquaculture, notably in Nile tilapia.

Sex differentiation is often strongly affected by a suite of environmental factors. For instance, temperature is well known to affect the sex differentiation process and skew sex ratios in a range of species, including Nile tilapia, European sea bass and medaka (Baroiller & D'Cotta, 2001; Rougeot *et al.*, 2008; Selim *et al.*, 2009). In carp, changes in temperature of ~5°C can dramatically alter the secretion of E2 (Penman & Piferrer, 2008). Fish exhibiting temperature dependent sex determination (TSD) are typically most thermosensitive during a critical period in their development. In a number of species including Nile tilapia, Japanese flounder (*Paralichthys olivaceus*) and Atlantic salmon, high temperature treatments at this time cause down-regulation of *Cyp19a1a* and subsequent masculinsation (Watts *et al.*, 2004). However, the susceptibility to such treatments varies not only across species but even between progenies of the same stock. Behavioural changes related to stocking density can affect sex differentiation (Baroiller & D'Cotta, 2001), as can social interactions of hermaphrodite fish (Devlin & Nagahama, 2002). The critical period of sex steroid sensitivity varies between species and must be thoroughly understood if steroid treatments for monosex farming are to be successful. This requires detailed information regarding the genetic and physiological processes controlling sex differentiation.

The ability to control the sex of hatchery-produced fish offers considerable advantages to aquaculture and can eliminate many sexually dimorphic problems. Aquaculture of females is favoured for many species, including European sea bass and Atlantic cod, although other species favour males (*e.g.* Nile tilapia). Stanley (1976) presented an early basis for producing monosex populations of grass carp (*Ctenopharyngodon idella*) by combining UV-irradiated common carp milt with grass carp roe, inducing gynogenesis (the fertilization of eggs without a male genetic contribution) and subsequently producing all-female progeny. Since they only carry maternal DNA, the progeny are genetic clones. For XX/XY fish, a common method of generating monosex female populations is to masculinise females by administering MT during a critical sensitivity period. The resulting phenotypic males (pseudomales) remain genotypically female and carry only X chromosomes; crossing them with normal females results in all female progeny (Devlin & Nagahama, 2002; Haugen *et al.*, 2011). However, using a similar approach in *Seriola* or other ZZ/ZW fish with aromatase inhibitors could therefore generate a male monosex population (Desprez *et al.*, 2003). This reiterates the importance of understanding sex genetics in each species of interest. An alternative approach to produce all-female fish for ZZ/ZW species is to cross pseudomales with ordinary females, the progeny of which 25% would be superfemales (WW). Assuming this genotype is non-lethal in the species of interest, crossing the F1 superfemales with ordinary males would thus generate all-female progeny (Devlin & Nagahama, 2002). However, the WW genotype is uncommon and the survivability of the condition would therefore need to be assessed for any given species. Monosex stocks have potential to improve the efficiency of aquaculture; achieving this will rely heavily on the use of molecular-based studies employing the tools and markers outlined previously.

1.8 Genes Involved in Sex Differentiation

Numerous genes have been identified as being involved in sex differentiation. Of these, several have been identified as having key roles: *Sry, Dmy, Dmrt, Cyp19, Amh, Sox9, Dax1* and *Vasa* (Yoshizaki *et al.*, 2000; Tchoudakova *et al.*, 2001; Rodriguez-Mari *et al.*, 2005; Harikrishnan *et al.*, 2011; Li *et al.*, 2013; Smith *et al.*, 2013). The master sex determination gene in mammals is *Sry*, which is unique to the male Y chromosome and is expressed in SGPCs. Expression of *Sry* causes these cells to up-regulate *Sox9*, an autosomal gene which promotes Sertoli cell differentiation (Devlin & Nagahama, 2002; Siegfried, 2010). SGPCs will differentiate into ovarian granulosa cells in the absence of *Sry*. The master sex determination genes of fish are poorly understood due to lack of conservation among vertebrates. In medaka, *Dmy* has been identified as the master sex determination gene. *Sry* and *Dmy* are non-homologous, with the latter being similar to *Dmrt1*, a gene postulated as the master sex determination gene in birds (Siegfried, 2010). *Dmrt* genes are downstream regulators of sex determination in many metazoans, although their functions are not well understood and expression patterns appear to vary significantly between teleosts. Five forms of *Dmrt* gene were revealed during analysis of the Atlantic cod genome (Johnsen & Andersen, 2012). As the genes involved in sex differentiation are characterised in more species, the existence and role of master sex determination will become clearer.

1.8.1 *Vasa*

The formation and migration of PGCs marks the beginning of gonad formation. A host of genes have been found to be linked to this action: most notably *Vasa, Nanos* and *Sdf1*. *Vasa* is an RNA helicase, which are important for catalysing ATP hydrolysis for unwinding RNA during replication. It has been highly conserved in animals; first identified in *Drosophila, Vasa* homologues have been identified in other invertebrates (*e.g.* hydra, flatworms) and in vertebrates, including mammals and fish

(Koprunner *et al.*, 2001; Tanaka *et al.*, 2001; Li *et al.*, 2009), although the exact mode of action varies between taxa. The role of *Vasa* in PGC migration has been confirmed in fish, notably medaka and zebrafish. Knockdown experiments in medaka confirm that *Vasa* has no role in the survival or proliferation of PGCs, but defective migration occurs in *Vasa*-depleted embryos (Li *et al.*, 2009). In contrast, knockdown of the *Nanos1* gene in zebrafish causes abnormal PGC morphology and development, including expressing mRNAs typically expressed in somatic cells (Koprunner *et al.*, 2001). Thus, *Vasa* and *Nanos* are indispensable to correct germ line formation.

It is paramount to understand the precise details of sex differentiation in each species of interest to aquaculture. Molecular techniques such as real-time PCR (qPCR) and *in-situ* hybridisation are ideal for this task. The former can be used to quantify the expression of the marker genes discussed earlier, such as *Cyp19a1a* and *Amh*. *In-situ* hybridisation is particularly useful for physically locating gene expression and making associations with cell and tissue types. Presslauer *et al.* (2012) used *in-situ* hybridisation to investigate the suitability of *Vasa* and *Nanos3* as germ cell markers of PGC migration in Atlantic cod. *Vasa* was identified as an ideal marker since it was expressed strongly throughout development. In Nile tilapia, two forms of *Vasa* are present: *Vas* and *Vas-s*, the expression of which depends on sex differentiation. Both are expressed during PGC migration to the presumptive gonad, but following gonad differentiation *Vas* becomes predominant during spermatogenesis in the testes. Conversely, *Vas-s* is predominant in the ovary during oogenesis (Tohru Kobayashi *et al.*, 2002). Manipulating *Vasa* has the potential to generate sterile fish, which has lucrative benefits for aquaculture (Presslauer *et al.*, 2012; Su *et al.*, 2013). This is particularly true for Atlantic cod, European sea bass and other species which have trouble with reproductive precocity.

1.8.2 *Amh*

Anti-Müllerian hormone (AMH) is a protein which inhibits the development of Müllerian ducts in mammals during embryogenesis, and has inhibitory effects on both male and female steroidogenesis (Halm et al., 2007). If AMH is not produced, the Müllerian ducts will form and eventually become components of the female reproductive tract. The *Amh* gene is up-regulated by SOX9 in male Sertoli cells during early differentiation, but generally significantly after the onset of sexual dimorphism. Despite Müllerian ducts not occurring in teleosts, AMH remains important to sex differentiation of fish, although it has been examined far less than in mammals (Halm et al., 2007; Li et al., 2013). The mode of action of *Amh* appears to have diversified significantly during vertebrate evolution. In Nile tilapia, qPCR analysis of developing gonads revealed *Amh* is expressed at low levels in both sexes prior to differentiation, but an increase occurs in testes from 15 days post hatching (dph). It is worth noting that sexually dimorphic expression of *Amh* in this species still occurs much earlier than histologically detectable morphology differentiation (Siegfried, 2010). In juvenile zebrafish, *Amh* is expressed in the undifferentiated gonad and testes but not the ovary or any other tissue, creating a distinct expression pattern correlating with the onset of sex differentiation (Rodriguez-Mari *et al.*, 2005; Piferrer & Guiguen, 2008). Similarly, the Japanese flounder (*P. olivaceus*) exhibits low *Amh* expression in the ovary compared to testes and undifferentiated gonad (Yoshinaga et al., 2004). European sea bass express high levels of *Amh* in both juvenile testes and ovary, although lower in the latter. *Amh* expression in this species has also been detected in somatic tissues (Halm et al., 2007). Some species do not show sexual dimorphism of *Amh*, such as medaka (Piferrer & Guiguen, 2008). The extent of diversification of AMH/*Amh* has caused its exact role in sex differentiation of fish to remain unclear. Future studies should continue to explore *Amh* expression in a wider range of species to better define its role. This is particularly true of previously undescribed species, as it will help to clarify the evolutionary pathways of AMH/*Amh* and will assuredly assist in improving aquaculture efficiency.

1.8.3 *Cyp19*

Aromatase P450 is a key enzyme which catalyses the conversion of androgens to estrogens. Unlike most mammals which have a single aromatase gene, most fish have two forms of aromatase encoded by separate genes. Ovarian aromatase (*Cyp19a1a*) is expressed primarily in ovarian tissue, initially in the granulosa cells which surround developing oocytes. Brain aromatase (*Cyp19a1b*) is expressed mainly in the brain and, to a lesser extent, ovarian tissue (Devlin & Nagahama, 2002; Kwon & Kim, 2013; Li *et al.*, 2013). The terminology of these genes is consistent in the literature; in some studies they are *Cyp19a1a* and *Cyp19a1b* (Kwon & Kim, 2013; Li *et al.*, 2013), while other studies use *Cyp19a1* and *Cyp19a2* (Strobl-Mazzulla *et al.*, 2008). Some studies refer to the genes as simply *Cyp19a* and *Cyp19b* (Callard *et al.*, 2001; Tchoudakova *et al.*, 2001). However, all of these studies refer to the two genes as the gonadal and brain isoforms of aromatase. Both isoforms are important in sex differentiation, but have notable differences in function. In medaka, the expression of *Cyp19a1b* does not become sexually dimorphic until near the onset of puberty, whereas *Cyp19a1a* is up-regulated in females early in development. Sex reversal studies have revealed that the expression of *Cyp19a1b* also correlates with phenotypic sex rather than genotypic sex (Menuet *et al.*, 2005; Okubo *et al.*, 2011). This is due to the presence of a functional estrogen responsive element (ERE) in *Cyp19a1b* which causes up-regulation of the gene in response to E2. *Cyp19a1b* may have roles in the brain including neurogenesis, social modulation and inhibition of male reproductive behaviours (Blazquez & Piferrer, 2004; Garcia-Segura, 2008; Okubo *et al.*, 2011). Since *Cyp19a1b* has a much more complicated expression pattern than *Cyp19a1a*, it can be considered less suitable as a marker of sex differentiation. Regardless, the gene is instrumental to correct development and should not be overlooked.

There does not appear to be functional ERE in *Cyp19a1a* and E2 subsequently has no regulatory effects over the gene, causing *Cyp19a1a* expression to be strictly correlated with genotypic sex. Up-regulation of *Cyp19a1a* is required both for triggering and maintaining ovarian development, while down-regulation induces testicular development. For this reason, *Cyp19a1a* is considered a reliable molecular marker of early sex differentiation (Guiguen *et al.*, 2010).

1.8.4 *Sox9*

The *Sox9* gene is essential to male sex differentiation in vertebrates and is well understood in mammals. The SOX9 protein promotes *Amh* expression, inhibiting formation of the female reproductive system. *Sox9* is up-regulated directly by testis-determining factor (TDF), encoded by *Sry*. The timing of sexually dimorphic *Sox9* expression occurs around physiological sex differentiation, although it may not coincide exactly. In Nile tilapia, it occurs slightly later during SGPC differentiation. Some species have two paralogues of *Sox9*, termed *Sox9a* and *Sox9b*. In adult medaka, *Sox9a* is expressed primarily in developing oocytes while *Sox9b* is only detectable in testes. Embryonic medaka express *Sox9b* in both sexes from the formation of SGPCs onwards, gradually decreasing expression in the ovary while remaining strongly expressed in testes (Siegfried, 2010). Zebrafish, being a false hermaphrodite species, differ in their *Sox9* characteristics. *Sox9a* is first expressed in SGPCs and the initial ovary before becoming restricted to Sertoli cells as the testes form. *Sox9b* is expressed mainly in the ovary (Rodriguez-Mari *et al.*, 2005). Since *Sox9a* expression coincides with the physiological onset of sex differentiation, it is useful as molecular indicator in this species. The same cannot be said for either Nile tilapia or medaka, since *Sox9* expression occurs later than initial physiological sex dimorphism. Consequently, it is unlikely that *Sox9* has any controlling role in sex determination. However, the vital role it serves by regulating *Amh* across vertebrates necessitates an understanding of its function for species of interest to aquaculture.

1.8.5 *Dax1*

There are numerous other genes involved in sex differentiation which have only recently begun to be studied in fish. Of note is *Dax1*, an antagonist of *Sry* and a regulator of aromatase expression. It is first expressed well before to sex differentiation and in a variety of tissues, including the developing gonad, brain and liver (Zhao *et al.*, 2006; Yajima *et al.*, 2012; von Schalburg *et al.*, 2014). The DAX1 protein down-regulates some genes involved in gonadal and adrenal steroidogenic activity, notably the SF1 gene which up-regulates *Amh* (Ludbrook & Harley, 2004). It also appears to maintain the pluripotence of embryonic stem cells. *Dax1* expression varies among vertebrates but exhibits a high degree of gene content conservation (Martins et al., 2013). For instance, it has been found to be strongly expressed in the liver of fish and other non-mammalian vertebrates, but not in mammalian livers. Liver expression of *Dax1* in Pengze crucian carp (*Carassius auratus*) is suspected to regulate vitellogenesis (the formation of yolk in oocytes) and closely mirrors the expression of *Cyp19a1a* (Li et al., 2013). Although the functioning of *Dax1* has not yet been completely characterised, it certainly is important with respect to sex differentiation. Studies in mice have demonstrated this; over-expression of the gene induces phenotypic sex reversal of males, while DAX1 deficient mice fail to properly form the testis cord (Ludbrook & Harley, 2004; Li *et al.*, 2013; von Schalburg *et al.*, 2014). The expression patterns of *Dax1* homologues in several fish appear to be similar, including European sea bass, Nile Tilapia, Atlantic salmon, zebrafish and medaka (Martins et al., 2013; von Schalburg et al., 2014). In response to MT treatment, Li et al. (2013) detected an increase in *Dax1* and *Cyp19a1a* expression. This suggests a possible regulatory or feedback effect of sex steroids at some point in the expression pathway. It is likely that *Dax1* expression will have applications as an important marker of sex differentiation, although that is yet to be validated. Further studies and improvements to molecular techniques will continue to improve the understanding of this gene.

1.9 Sex Differentiation in *S. lalandi*

A selection of the key genes described above can be targeted to develop genetic tools to study and identify sex differentiation in new aquaculture species such as *S. lalandi*. The ability to identify PGCs during early development is important and can be achieved using *in-situ* hybridisation with probes designed for *Vasa*. Expression profiling by qPCR of sex-linked genes such as *Cyp19a1a* and *Amh* can help identify when ovarian and testicular differentiation are initiated during larval development. To achieve this, species-specific gene sequences must first be identified. This relatively novel approach was used in this study with the aim to develop the aforementioned genetic tools and better understand sex differentiation in *S. lalandi*. Although most studies of aquaculture species arise following industry related problems, this study was commissioned to create an early knowledge platform in an attempt to pre-emptively manage common sex-linked production bottlenecks in farmed fish.

1.10 Aims and Objectives

1.10.1 Aims

Identifying the period of sex differentiation during early development is important for fully understanding the life cycle of *S. lalandi*, and may have potential for improving production efficiency. To achieve this, the relevant genetic tools must first be developed. With this in mind, the following aims were identified:

1. Identify transcripts of key genes involved in sex differentiation in the kingfish transcriptome.

2. Confirm the nucleotide sequences of these genes in adult kingfish.

3. Profile the expression of these genes during larval development.

4. Contrast the gene expression profile to the structural development of gonad tissue during larval development.

1.10.2 Objectives

To identify key sex differentiation genes in kingfish, bioinformatics approaches were used to:

1. Identify previously described amino acid sequences of sex differentiation proteins in a range of Perciformes.

2. Identify nucleotide sequences in the kingfish transcriptome which show homology to the Perciformes sex differentiation sequences.

To confirm the full nucleotide sequences of the sex differentiation genes in kingfish and profile their expression in developing larvae, molecular biology approaches were used to:

1. Amplify nucleotide sequences of selected sex differentiation genes from gonad tissues of adult kingfish, using PCR and sequencing to confirm the sequence.

2. Use qPCR to quantify expression levels in larvae of a range of ages.

In order to contrast gene expression with physiological development, histological methods were used to:

1. Embed and section larvae of a range of ages using a variety of matrices and stains.

2. Identify developing gonad tissues and primordial germ cells.

Chapter 2
Materials and Methods

2.1 Laboratories

The majority of laboratory work described herein was conducted in the University of Waikato C2.03 Molecular Genetics Laboratory in Hamilton, New Zealand. The University of Waikato R2.43 Histology Laboratory was utilised for paraffin embedding and sectioning, while the E2.14 Mammalian Physiology Laboratory was used for cryosectioning. Photomicroscopy was conducted in the R.2.29 Confocal Microscopy Laboratory. Sample collection and processing was conducted in the egg and larvae laboratory at the National Institute of Water and Atmospheric Research (NIWA) Bream Bay Aquaculture Park in Ruakaka, New Zealand.

2.2 Sample Collection

Kingfish larvae were provided by the NIWA Bream Bay Aquaculture Park. The larvae were hatched from the fertilized eggs of captive broodstock. Eggs were held in incubator tanks until approximately ten days post hatch (dph), after which they were transferred to larger rearing tanks.

Sample collections were conducted in 2013 from February to April, occurring every three days from hatching until 21 dph, then every five days from 25 dph to 50 dph, followed by a final collection occurred at 60 dph. Larvae were captured at each collection point using a fine aquarium hand net. Non-deformed larvae were identified and anaesthetised using aquatic anaesthetic (AQUI-S®) prior to the following sampling procedures.

The mean standard length was determined by measuring at least thirty larvae. This was conducted using Leica microscopy imaging software for larvae up to 35 dph, after which electronic vernier callipers were used. Photographs of larvae were also taken at each collection point.

Four replicates of up to fifty larvae were stored in RNA*later*® solution (Life Technologies), an aqueous reagent which rapidly permeates tissues to stabilise RNA. Fish larger than 10 mm standard length were cut into pieces ~4 mm thick to ensure solution penetration. Care was taken to minimise the transfer of fluids with larvae to prevent diluting the RNA*later*® solution. RNase-free plastic 2 ml microtubes or 20 ml falcon tubes were used according to body size. Samples were stored at -18°C.

Twenty larvae were collected and preserved for histology. The heads and tails were removed from fish larger than 12 mm standard length. Ten larvae were fixed in RNase-free buffered 4% formaldehyde fixative solution, while the second ten larvae were fixed in buffered 5% glutaraldehyde fixative solution. Both fixatives stabilise tissues by generating cross-links between proteins. The greater length of glutaraldehyde molecules enables longer cross-links and, subsequently, more rigid products. All samples were stored at 4°C for 12-18 hours, after which the fixatives were drained. Samples were washed by gently rocking in 4°C RNase-free phosphate buffered saline (PBS) solution for seven minutes, three times. Samples were subsequently transferred to 100% methanol and stored at room temperature. A sealed polystyrene box was used to transport all samples from the Bream Bay Aquaculture Park to the C2.03 Molecular Genetics Laboratory at the University of Waikato, Hamilton.

2.3 Database Searching and Primer Design

Geneious 7.1.7 (Biomatters Ltd.) software was used to identify potential sex differentiation genes from a previously constructed Ion Torrent RNA-Seq transcriptome library, prepared by Dr. Steve Bird from gonad and pituitary tissues of wild-caught adult *S. lalandi*. A reference protein sequence for the gene of interest from a closely related species (in the order Perciformes) was retrieved from the Universal Protein Resource (UniProt) database. The RNA-Seq library was searched using the basic local alignment search tool (BLAST) programme tBLASTn; the resulting hits, which covered as much of the gene of interest as possible, were downloaded and aligned to the reference protein sequence. The nucleotide sequence for each of these was obtained and used to construct a consensus sequence. Appropriate regions of the consensus sequence with high quality data were then used to design primers using Primer3 0.4.0 (*http://primer3.sourceforge.net*) (Rozen & Skaletsky, 2000) for subsequent amplification of the initial gene sequences and 5′- and 3′-RACE PCR. Using the recommendations for typical primer design by X. Yang *et al.* (2006) as guidelines, primers were designed with the following ideal specifications in mind: a melting temperature (T_m) of 60-62°C, length of ~20 nucleotides, GC content of 40-60%, and a 3′ end GC clamp to promote superior binding. However, due to sequence quality obtained from the library, specifications were not always able to be met for all primers designed. The following tables detail the primers used for initial sequence amplification (Table 1), RACE PCR (Table 2), and qPCR (Table 3). With the exception of M13, all primers were specific to *S. lalandi*. The *Actb* (β-actin) primers slBACTIN-rtF+R (Table 1) were designed by Dr. Steve Bird prior to this study. The *S. lalandi* nucleotide sequence for *Actb* is detailed in Appendix II.

Table 1. Initial and miscellaneous primer sequences used throughout this thesis. Excluding M13, all primers were designed to specifically target *Seriola lalandi*. Abbreviations: (F) = Forward. (R) = Reverse complement. T_m = Melting temperature.

Primer name	Nucleotide sequence	T_m (°C)	Gene target
slBACTIN-rtF	5′-CACAGACTACCTCATGAACATC-3′	52.3	*Actb*
slBACTIN-rtR	5′-CCGATCCAGACAGAGTATTTAC-3′	52.3	*Actb*
slVASA-F1	5′-GGGTTTAGAGGAGGAAGCCGAGG-3′	61.0	*Vasa*
slVASA-F2	5′-GGAGAAGATAAAGATCCAGAGAAGA-3′	50.9	*Vasa*
slVASA-F3	5′-GATCAGGGAAATATGCAGAG-3′	50.3	*Vasa*
slVASA-F4	5′-TCCACAAGCTGCGCTACCTG-3′	60.1	*Vasa*
slVASA-R1	5′-GTCATACTGCTGTCCTGAAAG-3′	57.0	*Vasa*
slVASA-R2	5′-GAACACACACTCCTCTAACCA-3′	54.1	*Vasa*
slVASA-R3	5′-CACAGCTAAGAACAGGTAGTCGGTCTTGAG-3′	61.2	*Vasa*
slVASA-R4	5′-GATGTCCTCGGGGTAGGTTG-3′	57.2	*Vasa*
slAMH-F1	5′-CTCATTGGTGAAAAAACAGGAAGT-3′	53.7	*Amh*
slAMH-F2	5′-CAATCAGCATAACTCTACTTCTG-3′	51.0	*Amh*
slAMH-F3	5′-CTGTCACTGGGCTTATCC-3′	52.4	*Amh*
slAMH-F4	5′-CCATTTCTCAGGTGCATCC-3′	53.5	*Amh*
slAMH-R1	5′-ATATCTGGTATCATGGAGAGGTAAG-3′	53.2	*Amh*
slAMH-R2	5′-GTAAGTCCCATGTTCATTCAAACC-3′	54.3	*Amh*
slAMH-R3	5′-GGTTATTGGTGTCCGCTCTGGTG-3′	60.0	*Amh*
slAMH-R4	5′-TACTCACGGGCCACTGTCTG-3′	59.0	*Amh*
M13 Forward(-20)	5′-GTAAAACGACGGCCAGT-3′	50.7	*lacZ* (*E. coli*)
M13 Forward(-40)	5′-GTTTTCCCAGTCACGAC-3′	50.9	*lacZ* (*E. coli*)
M13 Reverse	5′-AGTGTGTCCTTTGTCGATACTG-3′	47.0	*lacZ* (*E. coli*)
T7 Universal	5′-TAATACGACTCACTATAGGG-3′	47.5	*lacZ* (*E. coli*)

Table 2. RACE primer sequences used throughout this thesis. Excluding RACE adapters, inners and outers, all primers were designed to specifically target *Seriola lalandi*. Abbreviations: (F) = Forward. (R) = Reverse complement. T_m = Melting temperature.

Primer name	Nucleotide sequence	T_m (°C)	Gene target
3′ RACE adapter	5′-GCGAGCACAGAATTAATACGACTCACTATAGGT12VN-3′	61.2 - 62.6	-
3′ RACE outer	5′-GCGAGCACAGAATTAATACGACT-3′	55.1	-
3′ RACE inner	5′-CGCGGATCCGAATTAATACGACTCACTATAGG-3′	60.7	-
slVASA-3′F1	5′-TGATGTTCAGCGCAACCTACC-3′	57.7	*Vasa*
slVASA-3′F2	5′-AGAGGATGGCAGCTGACTTC-3′	56.9	*Vasa*
slAMH-3′F1	5′-CCAGTACTGTGCGTTTCTTCTGC-3′	58.3	*Amh*
slAMH-3′F2	5′-GAGTACGAGCTGCAGAGAGGAC-3′	58.9	*Amh*
slCYP-3′F1	5′-GTTTGCAGAAGACAGTGGAGGTTTG-3′	58.4	*Cyp19a1a*
slCYP-3′F2	5′-AGAGTCTGAGTCACGTGGATGTCC-3′	59.9	*Cyp19a1a*

Table 3. Real-time PCR primer sequences used throughout this thesis. All primers were designed to specifically target *Seriola lalandi*. Abbreviations: (F) = Forward. (R) = Reverse complement. Tm = Melting temperature.

Primer name	Nucleotide sequence	T_m (°C)	Gene target
Primer name	**Nucleotide sequence**	**T_m (°C)**	**Gene target**
slBACTIN-RTF1	5′-TACAACGAGCTGAGAGTTGC-3′	54.8	*Actb*
slBACTIN-RTR1	5′-GTTGAAGGTCTCGAACATGATC-3′	53.5	*Actb*
slGAPDH-RTF1	5′-GCTCATCTCTTGGTATGACAATG-3′	53.5	*Gapdh*
slGAPDH-RTR1	5′-TGCATGTACATCAGCAGGTC-3′	55.0	*Gapdh*
sl-VASA-RTF1	5′-GGATGTGATTGGACGAGGA-3′	54.2	*Vasa*
sl-VASA-RTR1	5′-AGCCCATATCCAGCATACG-3′	54.4	*Vasa*
slVASA-RTF2	5′-CGTATGCTGGATATGGGCT-3′	52.6	*Vasa*
slVASA-RTR2	5′-TGAACATCAGGGTCTGACG-3′	54.3	*Vasa*
sl-AMH-RTF1	5′-TCAACTCCTCTTTCCTTACTGTC-3′	53.9	*Amh*
sl-AMH-RTR1	5′-TACTTCCATTATCTGCATCCCA-3′	53.4	*Amh*
sl-AMH-RTF2	5′-TTCCTTACTGTCTTCTCCTTCG-3′	53.8	*Amh*
sl-AMH-RTR2	5′-CCTCCTTCTCCCTTATTACTTC-3′	54.2	*Amh*
slCYP19-RTF1	5′-GTTTGCAGAAGACAGTGGAG-3′	53.6	*Cyp19a1a*
slCYP19-RTR1	5′-AAGAGTCTGTTGGAGATGTCG-3′	54.2	*Cyp19a1a*
slCYP19-RTF2	5′-TACCTGTCGATGAGAAAGAGC-3′	54.0	*Cyp19a1a*
sl-CYP19-RTR2	5′-CAGTCCAACTTGAAGTAGATGTC-3′	53.1	*Cyp19a1a*

2.4 RNA Isolation

Two different RNA extraction lysis reagents were used in this study: TRIzol® reagent (Ambion), and R&A-Blue™ reagent (iNtRON Biotechnology). Both reagents are composed primarily of guanidine isothiocyanate and phenol. The former is a protein denaturant which inhibits RNase activity and lyses cells, while phenol serves to partition RNA into an aqueous supernatant when combined with chloroform. Due to the similarity of the protocols for these reagents, both will be detailed together in this section.

2.4.1 Tissue Homogenisation and Phase Separation

Kingfish tissues were added to 2 ml polypropylene RNase/DNase free microcentrifuge tubes containing 1 ml of either TRIzol® reagent or R&A-Blue™ reagent, and approximately 50 µl of 0.1 mm and 0.5 mm diameter glass beads. Larvae of 21 dph or younger were used whole, while older larvae were dissected into fragments to isolate trunk or gonadal tissues. These fragments were no larger than approximately 0.1 g. A Mini-Beadbeater™ (Alphatech) was used for 10 second bursts at 4800 oscillations per minute to homogenise the tissues. Samples were left at room temperature for five minutes to permit the nucleoprotein complex to completely dissociate.

For tissues homogenised in TRIzol®, 200 µl of chloroform was added to each tube. This was followed by vigorous shaking by hand for 15 seconds, incubation at room temperature for 2-3 minutes, and centrifugation at 12,000 x g at 4°C for 15 minutes. For tissues homogenised in R&A-Blue™, 400 µl of chloroform was added to each tube. This was followed by vigorous vortexing for 15 seconds before centrifuging at 13,000 rpm at 4°C for 10 minutes.

Centrifugation caused the mixture to separate into a lower phenol-chloroform phase, a cloudy white interphase containing DNA, and a colourless upper aqueous phase. RNA remained exclusively in the aqueous phase; this was carefully removed using an autopipette and transferred into a new 1.7 ml RNase/DNase free microcentrifuge tube. Care was taken not to disturb the interphase or lower phase as to avoid sample contamination.

2.4.2 RNA Precipitation and Resuspension

To each tube, 100% isopropanol was added (TRIzol® samples: 500 µl, R&A-Blue™ samples: 400 µl) before incubating at room temperature for at least 10 minutes. This incubation period was extended for small tissues (*e.g.* larvae younger than 12 dph) where RNA yield was expected to be low, up to a maximum of 18 hours at -18°C. The tubes were centrifuged for 10 minutes at 4°C (TRIzol® samples: 12,000 g, R&A-Blue™ samples: 13,000 rpm), generating a pellet of RNA which adhered to the bottom of the tube. The supernatant was decanted into a waste container, taking care not to disturb the pellet.

The pellet was washed by adding 1 ml of 75% ethanol (25% water) to the tube, inverting 5-6 times by hand followed by centrifuging (TRIzol® samples: 7,600 x g for 5 minutes at 4°C, R&A-Blue™ samples: 13,000 rpm for 1 minute at room temperature). The ethanol was removed by pipette and left to air dry for no more than 10 minutes, as allowing the RNA pellet to dry completely may have caused insolubility. The pellet was resuspended in 30 µl of RNase-free water and incubated at 60°C for 15 minutes to promote complete dissolution. RNA concentration and purity were determined by spectrophotometric analysis.

2.5 Spectrophotometric Analysis

A Thermo Scientific NanoDrop™ 2000 spectrophotometer was used to determine the concentration and purity of RNA, cDNA and plasmid DNA. In contrast to traditional spectrophotometers which use cuvettes and require large sample volumes, the NanoDrop utilises surface tension to hold samples in place between two optical fibres and only requires 1 μl of sample. The machine was blanked using 1 μl of DEPC-treated water, after which 1 μl of each sample was loaded and read. Kimtech Science™ KimWipes™ were used to clean the optical fibre ends between each sample. The software calculated sample concentrations by comparing absorbance readings at 260 nm between the blank and each sample. Purity was assessed by the 260 nm/280 nm absorbance ratio. Nucleic acids absorb at 260 nm whereas phenol and contaminant proteins absorb at 280 nm, so the ratio of these two wavelengths provides a useful indication of purity. A ratio of ~2.0 was considered pure for RNA, while ~1.8 was considered pure for DNA; the difference being attributed to the presence of uracil in RNA, which has a higher individual 260/280 ratio of 4.00 compared to thymine at 1.47 (Leninger *et al.*, 1975).

2.6 Complementary DNA (cDNA) Synthesis

cDNA was produced from total RNA extracted from kingfish tissues. Two kits were utilised: qScript™ Flex cDNA Synthesis Kit (Quanta Biosciences), and Tetro cDNA Synthesis Kit (Bioline). In both cases, RNA samples and kit reagents were thawed, mixed well and held on ice. The recommended maximum quantity of RNA to use for the Flex kit was 1 μg, whereas the Tetro kit permitted up to 5 μg. For simplicity, 1 μg was used for all cDNA synthesis reactions. The volume of RNA solution required to deliver 1 μg was determined using the formula:

$$RNA\ Volume\ (\mu l) = \frac{0.001}{RNA\ Concentration\ (ng/\mu l)}$$

For the Flex kit, RNA was primed with 2 µl of Oligo dT primers and 2 µl of Random Hexamer primers in thin walled 200 µl RNase/DNase free PCR tubes. Each tube was made up to 15 µl with DEPC-treated water, mixed and centrifuged. Tubes were incubated at 65°C for 5 minutes to denature the RNA, and then snap chilled on ice. The remaining reaction components were added: 4 µl of qScript Flex Reaction Mix (5X) and 1 µl of qScript Reverse Transcriptase, before mixing and centrifuging. Using a Biorad T100 Thermal Cycler, tubes were incubated at 25°C for ten minutes to permit annealing of the shorter random primers, followed by 42°C for 90 minutes for the reverse transcriptase to synthesis cDNA. A final incubation at 85°C for 5 minutes was performed to cease enzyme activity. All cDNA was initially held at 4°C before storage at -18°C.

For the Tetro kit, all reaction components were added to 200 µl RNase/DNase free PCR tubes. These components and their volumes are listed in Table 4. A master mix excluding the RNA template and reverse transcriptase was used when multiple cDNA syntheses were conducted simultaneously. As in the Flex kit, the Tetro kit utilised both Oligo dT primers and Random Hexamer primers. An RNase inhibitor was also included to prevent the activity of any RNase which may have been introduced. Using a Biorad T100 thermal cycler, tubes were incubated at 45°C for 30 minutes to permit reverse transcriptase to synthesise cDNA, then at 85°C for 5 minutes to cease enzyme activity.

2.7 Polymerase Chain Reaction (PCR)

PCR reactions were conducted on kingfish cDNA to amplify partial nucleotide sequences from sex differentiation genes of interest. A Solis BioDyne HOT FIREPol® DNA Polymerase kit was used for all reactions (Table 5). This kit contains a

Table 4. cDNA synthesis reaction components, using the Bioline Tetro cDNA Synthesis Kit.

Reaction component	Volume used per reaction (µl)
RNA template (1 µg)	Variable
Tetro Reverse Transcriptase (200µ/µl)	1.0
Oligo dT primer	1.0
Random hexamer	1.0
dNTP mix (10 mM)	1.0
5x RT Buffer	4.0
RiboSafe RNase inhibitor	1.0
DEPC-treated water	Make up to 20 µl

Table 5. PCR reaction components, using the Solis Biodyne HOT FIREPol® DNA Polymerase Kit.

Reaction component	Volume used per reaction (µl)
HOT FIREPol® (5 U/µl)	0.25
10x Buffer B1 (Tris-HCl, $(NH_4)_2SO_4$)	2.5
25 mM $MgCl_2$	2.5
20 mM dNTP mix	1.0
DNA template	1.0
Forward primer	1.0
Reverse primer	1.0
DEPC-treated water	Make up to 25 µl

chemically modified DNA polymerase (*Taq*) which is activated by a 15 minute incubation at 95°C. The kit also contains a buffer which primarily serves to keep pH optimal, and $MgCl_2$, a source of Mg^{2+} ions which are co-factors required by NTP-binding proteins. The forward and reverse primers specifically designed to each gene target (Table 1) were used for each PCR.

All solutions were defrosted and held on ice while in use. A UV-laminar flow hood dedicated to PCR was utilised when preparing master mixes. Each reaction was conducted in a 200 µl RNase/DNase free PCR tube using a Biorad T100 Thermal Cycler. The incubation protocol varied between individual PCR reactions depending on the primers being used, but a standard protocol is described in Table 6.

Table 6. Standard PCR incubation protocol.

Reaction step		Temperature (°C)	Time (min:sec)
Initial denaturation and *Taq* activation		94	15:00
Cycle steps (x35)	- Denaturation	94	00:30
	- Annealing	*	00:30
	- Extension	72	*
Final extension		72	10:00

* Varied depending on primer pairs being used.

2.7.1 Nested PCR

Nested PCR was used in many cases to reduce the amplification of non-specific products caused by non-specific binding of primers to cDNA. This approach involved two rounds of PCR as described above, but with the second round utilising 1 µl of first round PCR product in place of cDNA template. While the first round yielded both target and non-specific products, the second round primers were designed to amplify a smaller target within the first round target product. The likelihood of non-specific products binding to both sets of primers is reduced, making the technique very effective at amplifying a single DNA band.

2.8 Rapid Amplification of cDNA Ends (RACE) PCR

To obtain more nucleotide sequence for each of the sex differentiation genes, RACE PCR was used. The sequences of all RACE primers used are detailed in Table 2. For 3′ RACE, cDNA was prepared as described above, except that 2 μl of 3′ RACE adapter primer was used in place of Oligo dT and random hexamers. The adapter primer consists of a poly(T) tail, which complements the 3′ poly(A) tail of eukaryotic mRNA and also includes an additional adapter sequence, to which different 3′ RACE reverse primers can bind to and are used with forward primers specific for each sex differentiation gene. Nested PCR was also used for 3′ RACE, where the second round utilised 1 μl of first round PCR product in place of cDNA template, along with a forward and reverse primer that were designed inside of the initial primers.

2.9 Gel Electrophoresis

All PCR products were visualised by gel electrophoresis. A 1.5% agarose gel was prepared by dissolving 525 mg of HyAgarose™ LE agarose powder (Hydragene) in 35 ml of 1x TAE buffer in a 200 ml conical flask. The solution was heated in a domestic microwave on full power until it began to boil. As the solution started to cool, 1 μl of 10 mg/ml ethidium bromide (EtBr) was added. EtBr is a fluorescent nucleic acid stain which binds to DNA in the gel and fluoresces under ultraviolet light. The solution was carefully poured into a gel mould to avoid bubbles being introduced and a 12 well comb was placed into the mould. The gel was left to set for approximately 20 minutes and the gel combs were carefully removed, as to not collapse the wells. The gel was placed in an electrophoresis mini gel system tank (VWR) with 1x TAE buffer covering the gel. Matching the number of PCR samples, 3 μL drops of 6x gel loading buffer (30% (v/v) glycerol, 0.25% (w/v) bromophenol blue) were placed on Parafilm® (Sigma-Aldrich). Loading the gel involved adding 10 μl of PCR product to one of the 3 μl loading buffer drops and mixing by pipetting up and

down. Product mixtures were loaded onto the gel by pipetting 10 μl into each lane, excluding the first; this lane was always loaded with 5 μl of DNA ladder solution as a band size reference. A Genscript or Solis Biodyne 100 bp ladder (0.1 μg/μl) was usually used for this purpose (Appendix III). An Owl Scientific Lightning Volt™ OSP-250L power supply was connected to the electrophoresis tank and activated at 90 V for 30 minutes, causing the DNA to travel towards the negative electrode and separate according to size. Following electrophoresis, the gel was placed on a Life Technologies™ TFX-35M UV Transilluminator at 100% power, causing the EtBr bound to DNA bands to fluoresce. Photos were captured using a Cohu High Performance CCD Camera and edited using Scion Image v0.4 (Scion Corporation).

2.9.1 Gel Purification

Where it proved difficult to amplify a single PCR product and contaminating bands were present, bands of interest were isolated from non-specific products using a Zymoclean™ Gel DNA Recovery Kit (Zymo Research). Each band of interest was cut from the gel using a scalpel, placed into an RNase/DNase free 1.7 ml tube and weighed on a HR-200 analytical balance (AND). Agarose dissolving buffer was added at three times the volume of gel used (*e.g.* 300 μl per 100 mg gel slice). Tubes were incubated at 50°C for up to 10 minutes until the gel had completely dissolved. The solution was transferred to Zymo-Spin columns and centrifuged at 13,000 rpm for 30 seconds. The flow through was discarded and the columns subsequently washed by adding 200 μl of DNA wash buffer to each, centrifuging at 13,000 rpm for 30 seconds and discarding the flow through. This wash step was repeated and then the samples were centrifuged at 13,000 rpm for 30 seconds without adding DNA wash buffer, to ensure complete removal of the buffer. Each column was placed into a new RNase/DNase free 1.7 ml tube and 10 μl of diethylpyrocarbonate (DEPC) treated water added directly to the column matrix. DNA was eluted from the column by centrifuging at 13,000 rpm for 30 seconds.

2.10 Cloning

In order to facilitate nucleotide sequencing and other downstream applications, all PCR products of interest were cloned using chemically competent *E. coli* cell cultures. This involved the ligation of PCR products into plasmid vectors, assimilation of these plasmids into competent bacterial cells, before growing the cells and extracting the cloned plasmids for sequencing.

2.10.1 Bacterial Culture Media

Lysogeny broth (LB) was selected as a culture medium for bacterial cloning. First described by Bertani (1951), LB is a nutritionally rich medium widely used for growing bacteria in laboratories. Liquid broth was prepared in 500 ml batches by dissolving powdered LB in distilled water at 25 g/l, then autoclaved (media cycle). Liquid LB was stored at room temperature in sealed 500 ml glass reagent bottles (Duran®) to reduce the risk of contamination.

Similarly, LB agar was prepared as above but with the inclusion of agar at 15 g/l. Immediately following autoclaving, solutions were incubated in a 60°C water bath for 45 minutes, allowing the solution to cool to a level which accommodated addition of ampicillin, but keeping the LB agar from solidifying. Ampicillin (Calbiochem) was added to the solution using an autopipette at 100 mg/ml inside a positive flow PCR cabinet. The solution was subsequently poured into plastic Petri dishes, taking care to minimise the introduction of air bubbles. Once cooled and solidified, the dishes were wrapped in aluminium foil to protect them from light and stored inside a media-dedicated fridge at 4°C.

2.10.2 Ligation

A ligation reaction was set up using a pLUG Prime® TA-Cloning Vector Kit (iNtRON Biotechnology). This kit includes a vector with a 3′-end T-overhang which binds to the 3′-end A-overhang generated by *Taq* polymerase during the PCR, thereby allowing PCR products to be inserted directly (see Appendix III). The kit contents were combined into a master mix and added to PCR tubes with selected PCR products (Table 7). The mixtures were kept at 4°C overnight to provide sufficient time for ligation to occur.

Table 7. pLUG Prime® TA-Cloning Vector Kit components and aliquot volumes used for vector ligation.

Reaction component	Aliquot volume (µl)
Ligation buffer A	1.0
Ligation buffer B	1.0
pLUG Prime® TA-Cloning Vector	2.0
T4 DNA Ligase	1.0
PCR product	5.0

2.10.3 Transformation

E. cloni® 10G chemically competent cells (Lucigen®) were used for cloning plasmid vectors contaning PCR products. Tubes containing 40 uL of E. cloni® subcloning grade cells were removed from storage in the -80°C freezer and placed into ice to thaw for 10 minutes. Ligated plasmids were initially heated at 70°C for 15 minutes to terminate DNA ligase activity, before adding 4 µl of plasmid to 40 µl of competent cells in 2 ml plastic tubes and incubating on ice for 30 minutes. Each mixture was subsequently heat shocked for 45 seconds in a 42°C water bath to increase cell membrane permeability and facilitate the assimilation of plasmids into the cells, before immediately being snap-chilled on ice for 2 minutes. To aid cell recovery, 960

μl of recovery medium was added to each tube. Tubes were placed in a 37°C shaking incubator for at least two hours to allow the cells to establish ampicillin resistance.

Previously prepared LB plates were spread with Xgal (40 μl at 20 mg/ml) and IPTG (40 μl at 0.1 M) and warmed to 37°C for at least one hour. Each plate was inoculated with 100 μl of transformed cells and incubated at 37 °C overnight. Transformation efficiency was visualised using blue-white colony screening. In the presence of Xgal and IPTG, cells missing the insert form blue colonies, due to the expression of β-galactosidase, whereas cells with plasmids containing the gene insert cannot express this enzyme and form white colonies.

2.10.4 Colony PCR

Successful transformation of white colonies with the correct PCR product ligated into the plasmid was confirmed using colony PCR. Cells were collected by touching a 10 μl pipette tip to each white colony of interest, taking care not to remove any agar. The cells were added to RNase/DNase free PCR tubes containing 2 μl of DEPC-treated water and mixed by drawing the water up and down the pipette tip. Each tube was made up to 25 μl according to the HOT FIREPol® protocol (Table 5), using M13 Forward(-40) and M13 Reverse primers (Table 1). All tubes were incubated according to the standard PCR protocol (Table 6), using an annealing temperature of 54°C and an extension time of one minute. PCR products were visualised by gel electrophoresis (section 2.9). Bands were expected to be ~170 bp larger than the insert, due to the addition of plasmid DNA amplified between the priming sites of M13 primers. Colonies containing the insert were grown up in LB media overnight. For each selected colony, a 14 ml sterile culture tube (Greiner Bio One) with filled with 5 ml of LB broth and 5 μl of 100 mg/ml ampicillin (Calbiochem). A sterile 10 μl pipette tip was used to transfer the colony containing the correct PCR insert from the agar plate to the culture tube. Once a colony had been picked, the tip was dropped

into the culture tube. The tubes were then left in a shaking incubator (Bioline) at 37°C at 200 rpm overnight for at least 12 hours.

2.10.5 Plasmid Extraction

Plasmids were extracted from cell cultures using a Geneaid Presto™ Mini Plasmid Kit. Up to 1.5 ml of cell culture was transferred into an RNase/DNase free 1.7 ml microcentrifuge tube and spun at 13,000 rpm for 1 minute to form a cell pellet. The supernatant was decanted and discarded. Pellets were resuspended by adding 200 µl of PD1 Buffer and 2 µl of TrueBlue Lysis Buffer, followed by vigorous vortexing. Cells were lysed by adding 200 µl of PD2 Buffer and inverting gently ten times. Tubes were left to stand for at least two minutes at room temperature to ensure lysates were homologous, indicated by a homogeneous blue colouration. Lysis was neutralised by adding 300 µl of PD3 Buffer and gently inverting ten times, resulting in the solution becoming colourless.

Tubes were centrifuged at 13,000 rpm for 30 seconds, after which each supernatant was transferred by pipette to a PD Column inside a 2 ml collection tube. Each column was spun at 13,000 rpm for 30 seconds and the flow through discarded. To optimise the product for downstream DNA sequencing, 400 µl of W1 Buffer was added to each column before centrifuging at 13,000 rpm for 30 seconds and discarding the flow through.

This step was repeated using 600 µl of Wash Buffer. Each column was centrifuged for an additional 3 minutes to dry the column matrix. To each column, 600 µl of Wash Buffer was added to each PD column and centrifuged at 14-16,000 x g for 30 seconds, the flow through discarded and the PD column centrifuged at 13,000 rpm for 3 minutes to dry the column matrix. To elute the plasmid DNA, 30 µl of DEPC-treated H_2O was added directly the each column matrix. Columns were removed from the collection tubes and placed inside a fresh RNase/DNase free 1.7 ml

centrifuge tube, before centrifuging at 13,000 rpm for two minutes. The flow through was transferred back to the column matrix before centrifuging for a further two minutes. The eluted DNA flow through was retained and stored at 4°C.

The NanoDrop spectrophotometer was used to determine plasmid concentration and purity (section 2.5). Plasmids were required to be concentrated at ~150 ng/µl for sequencing. Samples with higher concentrations were diluted in DEPC-treated water. Samples with concentrations lower than 100 ng/µl were concentrated using a Speed Vac® and Gel Pump® vacuum centrifuge system (Savant). Plasmid samples were loaded into the centrifuge and spun under vacuum at 43°C for up to 30 minutes, or until a visible reduction in sample volume was observed. Concentrated samples were re-tested on the NanoDrop to confirm concentration.

2.10.6 Restriction Digest

A final check to ensure that each plasmid contained the correct PCR target sequence was carried out by digesting the plasmids using restriction enzymes, PstI and EcoRI (Roche). The digest reaction components were added to RNase/DNase free 1.7 ml tubes according to Table 8 and centrifuged to collect the contents. Each tube was incubated overnight at 37°C. Digested samples were visualised by gel electrophoresis (section 2.9). Due to the inclusion of vector sequence between the PstI and EcoRI cut sites, bands were expected to be ~45 bp larger than the insert.

Table 8. Roche restriction digest kit reaction components.

Reaction component	Aliquot volume (µl)
PstI enzyme	2
EcoRI enzyme	2
SureCut Buffer	2
H_2O	17.5
Plasmid DNA	1

2.10.7 Sequencing

Upon confirmation of successful ligation by restriction digest, the plasmids were prepared for sequencing. The required volume and concentration of the plasmid was 12 µl at a concentration of 150 ng/µl. The volume of plasmid required to make up a 12 µl aliquot was calculated using the formula:

$$V_1\ (Plasmid\ volume) = \frac{C_2V_2}{C_1} = \frac{(150\ \text{ng/µl} \times 12\ \text{µl})}{Plasmid\ Concentration}$$

where C_1 is the concentration of plasmid, V_1 is the required volume of plasmid, C_2 is the final concentration and V_2 is the final volume. The 12 µl aliquot of plasmid was then added to a labelled RNase/DNase free 0.7 ml PCR tube and DEPC water was added to make the final volume 12 µl. The tubes were sent to the Waikato DNA Sequencing Facility where the plasmid DNA was sequenced using BigDye® Terminator v3.1 Cycle Sequencing kit (Life Technologies) and was run on a 3130XL Genetic Analyser System (Life Technologies). Plasmids were sequenced in one direction using the T7 universal primer (Table 1).

2.11 Sequence Analysis

The ExPASy translate tool (*http://web.expasy.org/translate*) was used to translate sequences. Clustal Omega 1.2.0 (*http://www.ebi.ac.uk/Tools/msa/clustalo*) was used to construct alignments. Amino acid identities were calculated using MatGAT v2.02 (Campanella *et al.*, 2003). Phylogenetic trees were constructed by first aligning sequences using ClustalX 2.1 (*http://www.clustal.org/clustal2)* using the BLOSUM cost matrix, then using Treeview 1.6.6 (*http://taxonomy.zoology.gla.ac.uk/rod/treeview.html*) and FigTree 1.4.2 (*http://tree.bio.ed.ac.uk/software/figtree*) software to generate the trees. Locations of protein signatures were predicted using InterProScan (*http://www.ebi.ac.uk/interpro/search/sequence-search*) and Clustal Omega.

2.12 Real-Time PCR

2.12.1 Primer design

Forward and reverse primers for housekeeping genes and sex differentiation genes of interest were carefully designed to ensure the best possible amplification of the targeted genes using real-time PCR (Table 3). To do this, the online software Primer3 (*http://primer3.sourceforge.net*) was used, adhering to the following guidelines where possible: GC content was between 40-60% to ensure maximum product stability. The amplicon length was 50-150 bases, as very long amplicons can lead to poor amplification efficiency. Primers ended with a C or G residue, because A or T residues can bind more easily to DNA in a non-specific way. The melting temperature (T_m) of each primer was between 54–60°C, with the Tm of the forward and reverse primers being within 1°C of each other. In addition, the number of possible interactions between nucleotides within primers and between the forward and reverse primer was low. This was to prevent the formation of hair pin loops and primer dimers. Lastly, both the primer and the amplicon sequences were BLAST searched against the public database and RNA-Seq transcriptome, to help ensure that the correct target was being amplified.

2.12.2 RNA Extraction

Total RNA was prepared from a number of different larvae and tissues. Testes and ovary tissues were targeted for use in determining the specificity of each primer set designed and for testing their efficiencies. RNA was extracted from four replicates of 50 larvae aged 0 dph, 3 dph, 12 dph and 18 dph, for a preliminary examination of the genes of interest during development.

Preparation of high purity RNA with no genomic DNA contamination is crucial for achieving the best results from real-time PCR and helps to prevent the amplification of non-specific products. To achieve this, the Quick-RNA™ MiniPrep kit (Zymo) was used, which has been designed for the easy, reliable and rapid isolation of DNA-free RNA from a wide range of cell and tissue samples. It uses a unique buffer system with CleanSpin column technology to allow high quality total RNA isolation in about 10 minutes. The RNA isolation consisted of three steps: sample lysis and homogenisation, sample clearing and gDNA removal, and finally RNA purification.

Kingfish tissues or pooled larvae from each time point were added to 2 ml polypropylene RNase/DNase free microcentrifuge tubes containing 600 µl of RNA Lysis Buffer. To this, approximately 50 µl of 0.1 mm and 0.5 mm diameter glass beads were added and the tubes were placed into a Mini-Beadbeater™ (Alphatech) and the tissues homogenised at 4800 oscillations per minute for 10 seconds. This was repeated if the tissue did not break down completely. Homogenisation caused the lysis buffer to foam, which was cleared by centrifugation at 13,000 x g for 1 minute. The supernatant was transferred into a Spin-Away™ Filter placed into a collection tube, and centrifuged at ≥10,000 x g for one minute to remove the majority of gDNA, which remained in the filter.

To the flow-through, which contained the sample in RNA Lysis Buffer, one volume of 100% ethanol was added (1:1 with flow-through volume) and mixed well. This mixture was transferred to a Zymo-Spin™ IIICG Column in a Collection Tube, centrifuged 13,000 x g for 30 seconds and the flow-through discarded, leaving the total RNA within the column. To ensure complete removal of any trace genomic DNA, the column underwent DNase I Treatment. The column was prewashed with 400 µl of RNA Wash Buffer, centrifuged at 13,000 x g for 30 seconds, and the flow-through discarded. For each sample to be treated, a DNase I Reaction Mix was prepared in a 2 ml polypropylene RNase/DNase free microcentrifuge tube, which contained 5 µl of DNase I DNA and 75 µl of Digestion Buffer. This 80 µl DNase I Reaction Mix was added directly to the column matrix and incubated at room

temperature (20-30°C) for 15 minutes. The column was centrifuged 13,000 x g for 30 seconds before discarding the flow-through, adding 400 µl RNA Prep Buffer to the column, again centrifuging at 13,000 x g for 30 seconds, before discarding the flow-through to ensure complete removal of the DNase I Reaction Mix. To wash the RNA, 700 µl RNA Wash Buffer was added to the column and centrifuged 13,000 x g for 30 seconds, discarding the flow-through. A second aliquot of 400 µl RNA Wash Buffer was added to the column, and centrifuged 13,000 x g for 2 minutes to ensure complete removal of the wash buffer. Finally, the column was transferred to a 1.5 ml polypropylene RNase/DNase free microcentrifuge with the hinge removed before adding 30 µl of DNase/RNase-free water directly to the column matrix. The column was centrifuged at 13,000 x g for 30 seconds, eluting RNA in the flow-through. The eluted RNA was measured using the Nanodrop 2000 (Thermo Scientific) and the 260/280 absorbance ratio measured to ensure purity was approximately ≥1.8. The RNA was immediately used for cDNA synthesis to minimise RNA degradation.

2.12.3 cDNA Synthesis

To synthesise cDNA, the HiSenScript™ RH(-) cDNA Synthesis Kit (iNtRON) was used. This kit has been designed for the sensitive and reproducible detection and analysis of full-length cDNA copies from a total RNA sample. The reaction components were added to 200 µl RNase/DNase free PCR tubes as described in Table 9.

Table 9. cDNA synthesis reaction components, using the HiSenScript™ RH(-) cDNA Synthesis Kit.

Reaction component	Volume used per reaction (µl)
Total RNA (1 µg)	Variable
2X RT Reaction Solution	10
Enzyme Mix Solution	1
DNase/RNase free water	Make up to 20 µl

Each tube was thoroughly vortexed before being spun down by centrifugation. The tubes were placed into a T100 thermal cycler (Biorad) and incubated at 42°C for one hour to allow reverse transcription, followed by an inactivation step of 85°C for five minutes. Each cDNA obtained from the testes and ovary was stored at 4°C until being used in real-time PCR for testing primer efficiencies. cDNA produced from pooled larvae was immediately diluted into 300 µl of DNase/RNase free water and subsequently stored at 4°C, until being used to measure the expression of the genes of interest in real-time PCR.

2.12.4 Real-Time PCR

Real-time amplification was performed using the Rotor-Gene 6000 (Corbett). The RealMOD™ GH Green Real-time PCR Master Mix Kit (iNtRON) was used following the manufacturer's instructions. The reaction components were added to a 200 µl RNase/DNase free, thin walled, clear PCR tube (Axygen) as listed in Table 10. This was subsequently used for real-time PCR using the programme described in Table 11.

Table 10. Real-time PCR reaction components, using the RealMOD™ GH Green Real-time PCR Master Mix Kit (iNtRON).

Reaction component	Volume used per reaction (µl)
cDNA template	8
Forward primer (10 µM)	1
Reverse primer (10 µM)	1
2X RealMOD™ GH Green Real-time PCR Master Mix	10

Table 11. Real-time PCR incubation programme.

Reaction step		Temperature (°C)	Time (min:sec)
Initial denaturation		94	5:00
Cycle steps (40-50)	- Denaturation	94	0:20
	- Annealing/extension	60	0:30

Fluorescence outputs were measured and recorded at 80°C, and a melt curve for each sample was performed between 72 and 94°C to ensure that only a single product had been amplified. Using the graphs generated, a threshold line was set on the amplification curves to generate the threshold cycle (C_T) value for each reaction, which is where a detectable amount of amplicon product has been generated during the early exponential phase of the reaction.

2.12.4.1 Primer Testing and Efficiencies

Before the previously designed primers were used for determining the expression of the genes of interest in the developing larvae, the efficiencies of each pair of primers were tested to determine the amplification efficiency. cDNA prepared from adult kingfish testes tissues underwent serial dilutions to give cDNA that was undiluted

(U), or diluted 1:4, 1:16, and 1:64. Each pair of primers was run in duplicate for each cDNA dilution, and each run contained a negative control. The negative control contained everything in the reaction mix except for the cDNA template, which was replaced with DNase/RNase water at the same volume to check for contamination. C_T values were obtained from each amplification curve at each dilution and plotted against the initial amounts of input material on a semi-$\log_{10}$ plot. Linear regression was performed on the data points and the slope (S) calculated, along with the R^2 value, to determine the goodness of fit of the points. Efficiencies (E) of the primers were calculated using the equation:

$$Efficiency\ (E) = 10^{\left(\frac{-1}{Slope\ (S)}\right)} - 1$$

In addition, real-time products were run on a 1.5% agarose gel using the protocol described in Section 2.9, to confirm amplification of a single product.

2.12.4.2 Measuring gene expression during development

Using the primers for which the efficiencies had been determined, the C_T values for the sex differentiation genes *Vasa*, *Amh* and *Cyp19a1a*, and the housekeeping genes *Gadph* and *Actb* (β-actin), were determined in duplicate for each sample prepared from each larval stage. Both housekeeping gene results were used for normalisation of the gene expression in each pool of larvae, and the geometric means of these two genes were used in the expression analysis (Vandesompele *et al.*, 2002). The relative expression levels were determined using the comparative ΔCq method (Silver *et al.*, 2006), which is a variation of the Livak $\Delta\Delta C_T$ method (Livak & Schmittgen, 2001) which evaluates expression of genes, where $\Delta Cq = 2^{Cq(ref) - Cq(target)}$. Assessment of statistical significance was analysed using one-way ANOVA followed by the independent Student's t-test, where values $P < 0.05$ were considered significant.

2.13 Histological Sections

Two histological approaches were explored to find a reliable means of producing high quality tissue sections: paraffin embedding and cryo-sectioning.

2.13.1 Preparation of Gelatine-Coated Slides

In order to adhere and retain tissue sections during washing and staining, microscope slides were coated in gelatine adhesive. In a 500 ml glass beaker (Kimax), 0.5 g of gelatine powder (Sigma-Aldrich) was added and left to dissolve in 100 ml of deionised water, warmed to 45°C and stirring on a Cimarec® heating plate (Barnstead Thermolyne). Once the gelatine dissolved, 0.05 g of $CrK(SO_4)_2$ (Ajax Chemicals) was added and also left to completely dissolve. The gelatine solution was cooled at room temperature and filtered through 0.45 μm hydrophilic Minisart syringe filters (Sartorius Stedim) using 10 cc/ml plastic syringes (Terumo). A glass staining dish was filled with gelatine solution, to which a rack filled with 26 mm by 76 mm glass slides (Fronine) was dipped five times for five seconds. The slides were placed in a drawer to prevent dust settling on the gelatine, and left to dry for at least 48 hours at room temperature.

2.13.2 Sample Preparation for Paraffin Embedding

Paraformaldehyde-fixed larvae were removed from methanol, transferred into new plastic tubes and put through a series of washes to prepare them for paraffin embedding. Appropriate tubes and wash volumes were selected according to the tissue size. For larvae up to 25 dph, 2.0 ml plastic tubes and 1.5 ml wash volumes were used. Samples were rehydrated by washing for two minutes with PBS with 0.2% TWEEN® 20. This was repeated three times, followed by a further three washes with PBS at for five minutes at 4°C. In a fume hood, the samples were

removed from PBS and put through a dilution series of ethanol at room temperature: 5% ethanol in PBS for five minutes, 25% ethanol in PBS for five minutes, 75% ethanol in PBS for five minutes, and 100% ethanol for one minute. The samples were washed twice in 100% ethanol for 30 minutes and thrice in mixed isomer drum-grade xylene for 30 minutes. Samples were immersed in 1:1 xylene and liquid Surgipath® Paraplast® paraffin wax (Leica), and held in a Thermotec 2000 incubator (Contherm) set to 60°C for at least 45 minutes. Three 100% paraffin washes were conducted at 60°C for at least 20 minutes each, allowing the wax to infiltrate into the tissues.

2.13.2.1 Protocol Revisions

The paraffin embedding preparation protocol was revised in an attempt to improve quality during sectioning. The PBS wash steps were omitted in favour of transferring samples directly from methanol to 100% ethanol for five minutes, before being left in a second ethanol wash overnight. In a fume hood, the ethanol was poured off and replaced with xylene. A fresh stock of *p*-xylene (BDH Chemicals) was used in favour of the mixed isomer drum-grade xylene used previously. After three xylene washes of 30 minutes each, samples were immersed in 1:1 xylene and paraffin wax, and held in an incubator set to 60°C. After allowing the paraffin wax to melt completely (at least 30 minutes), the tubes were inverted at least three times to ensure thorough mixing of the xylene and paraffin wax. The three subsequent 100% paraffin washes at 60°C were extended to ~45 minutes each in an effort to improve wax infiltration.

2.13.3 Paraffin Embedding

A Histostar embedding station (Thermo Scientific) was used to embed prepared samples into paraffin blocks. The wax reservoir was filled with paraffin wax and held at 60°C, with additional wax added as required. The unit was set to cool the cold spot to 5°C, while the hot spot, mould storage drawer and tissue storage drawer

were set to 60°C. A sufficient number of metal moulds of appropriate size for the samples were selected and placed inside the mould storage drawer prior to each embedding session. Each tissue sample was removed from liquid paraffin and placed into an appropriately sized metal embedding mould in the tissue storage drawer. Using pre-warmed tweezers, care was taken to orient each sample in the vertical or horizontal plane, according to whether transverse or longitudinal sections were desired respectively. For the former, samples were oriented posterior-up in the mould in order to cut head to tail, as the bottom of the mould represented the initial cutting surface. A small amount of liquid paraffin was dispensed into the mould and allowed to solidify slightly on the cold spot to retain the sample in position. For longitudinal sections, this step was not generally required as gravity would cause most samples to lay flat in the mould. The mould was filled to the upper lip with liquid paraffin, a labelled Shandon tissue cassette (Thermo Scientific) placed on top, and additional paraffin added to saturate the cassette. The mould was placed on the cold module for at least one hour to completely solidify. Embedded tissues were removed from the mould by gently flexing the edges of the mould.

2.13.4 Paraffin Sectioning

Immediately preceding sectioning, gelatine coated slides were placed on an HI1220 heated flattening table (Leica) set to 37°C. Using a 1 ml transfer pipette (Sarstedt), a small volume of 0.2% ethanol was pooled onto the slide surface. A razor blade was used to trim the paraffin block into a trapezoid shape to encourage cut sections to form ribbons instead of adhering to the blade. The block was loaded into a RM2055 rotary microtome (Leica) with the long side of the trapezoid facing down towards the cutting edge. After removing the protective cover, an S35 stainless steel blade (Feather) was placed into the microtome and clamped into position. The specimen holder was fully retracted using the microtome's electronic control panel, and the knife holder base positioned to closely align the blade with the edge of the block.

With the section thickness was set at 20-40 micron, the control panel was used to cycle the microtome at a fast speed (~85%), cutting the block until the embedded tissue was reached. The section thickness was reduced to 6 – 10 micron, selecting the thinnest setting which consistently produced good quality sections. The speed was reduced to a moderate setting (up to 75%). Sections were cut to produce short ribbons no longer than the length of the ethanol pool on the slide. Ribbons were carefully transferred from the microtome to the pool using camel-hair brushes and needles, ensuring that the section orientation was conserved. No more than two ribbons were floated on any single slide. Depending on the size of the tissue, sections were taken at regular intervals through the sample until the entire region of interest had been sectioned. Due to tissue compression and wrinkling during sectioning, the ribbons were left floating on ethanol at 37°C overnight, allowing the sections to expand to their original proportions. By this point the ethanol had evaporated and the slides could be easily transferred into slide racks for storage prior to staining.

2.13.5 Paraffin Staining

In order to add contrast and highlight structures of interest, slides were stained using one of two staining procedures: toluidine blue, or hematoxylin and eosin (H&E). In both cases, slides were loaded into dipping racks and put through a series of wash baths as detailed in Table 12. Xylene was used to dissolve paraffin wax, while the ethanol dilution series served to rehydrate the tissues to allow the aqueous reagents of each stain to penetrate into cells. Between each wash step, the rack was removed from solution and permitted to drip dry for up to 30 seconds. The rack was touched to a paper towel to remove any excess solution before immersion in the next wash solution. After the final wash in water, slides were left in the rack to air dry for at least ten minutes before being laid on a paper towel.

Table 12. Wash protocol for preparing slides of paraffin sections for toluidine blue or H&E staining.

Solution	Immersion time (minutes)
Xylene	10
Xylene	10
100% ethanol	5
95% ethanol	2
70% ethanol	2
35% ethanol	1
Tap water	At least 30 minutes

To stain with toluidine blue, drops of 1% toluidine blue solution were applied to individual sections on each slide using an autopipette. Care was taken to ensure each section was completely immersed. After approximately two minutes, each slide was held upside-down under a gently running tap to remove the toluidine blue solution. The slides were placed upright on a paper towel to dry the undersides, before using an Axiostar Plus compound microscope (Zeiss) to assess the quality of sectioning and staining. Good quality sections were put through a reverse dilution series of ethanol (35% for 20 seconds, 70% for 20 seconds, 95% for 20 seconds, 100% twice for 20 seconds) before two five minute washes in xylene. A drop of DPX mounting media (Sigma Aldrich) was applied to each section, taking care not to introduce air bubbles, before placing a 22x40 mm cover slip (Menzel-Glaser) on each slide. Slides were left upright overnight to allow the mounting solution to set completely.

To stain with H&E, drops of fresh hematoxylin solution were applied to individual sections on each slide using an autopipette. Care was taken to ensure each section was completely immersed. After approximately five minutes, each slide was held upside-down under a gently running tap to remove the hematoxylin solution, before being placed in a slide rack and immersed in a bath of tap water for a further ten

minutes. The rack was removed from solution and permitted to drip dry for up to 30 seconds, then touched to a paper towel to remove any remaining excess solution. The rack was dipped in 0.1% HCl three times, dipped in tap water three times, dipped in 0.1% NH_4OH three times, and finally dipped in tap water a further three times. Slides were removed from the rack and placed upright on a paper towel. Drops of eosin solution were applied to individual sections in the same manner as hematoxylin solution. After three minutes, the slides were placed back in the slide rack and dipped in a series of ethanol, acetone and xylene baths as detailed in Table 13. The slides were again removed from the rack and placed upright on a paper towel, before using an Axiostar Plus compound microscope to assess the quality of sectioning and staining. For good quality sections, a drop of DPX mounting media (Sigma Aldrich) was applied to each section, taking care not to introduce air bubbles, before placing a 22x40 mm cover slip (Menzel-Glaser) on each slide. Slides were left upright overnight to allow the mounting solution to set completely.

Table 13. Wash protocol for eosin cytoplasmic counterstaining and dehydration of slides of paraffin sections for H&E staining.

Solution	Number of dips / Immersion time
100% ethanol + 0.1% acetic acid	5 x 5 seconds
100% ethanol	5 x 5 seconds
100% ethanol	5 x 5 seconds
Acetone	5 x 5 seconds
Acetone	5 x 5 seconds
Xylene	5 x 5 seconds
Xylene	5 x 5 seconds

2.13.6 Cryostat Embedding and Sectioning

Glutaraldehyde-fixed larvae were removed from methanol, placed into individual wells of plastic six-well culture plates (Sigma Aldrich) and put through a series of washes. Samples were immersed in PBS at room temperature three times for ten minutes, followed by two ten-minute washes in a 30% sucrose solution. Samples were put through a final sucrose wash at 4°C overnight. Approximately two thirds of the sucrose bath was removed and replaced by FSC 22™ frozen section compound (Surgipath) optimal cutting temperature (OCT) solution, remaining at 4°C for at least one hour. Larvae were removed from the wells, placed in plastic Tissue-Tek® Cryomolds® (Sakura) and covered with OCT. Care was taken to orient the larvae correctly and remove any air trapped in the fluid. The moulds were sprayed from beneath with Frostbite cold spray (Leica) to solidify the OCT, forming the embedding matrix.

A CM1850 UV cryostat (Leica) was used for sectioning. Each embedded sample was removed from the mould and seated onto a cryostat specimen disc layered with OCT. This was placed on the quick freeze shelf inside the cryostat chamber at -25°C until the OCT had solidified, binding the embedded larvae to the sample holder. The specimen disc was placed on the Peltier element and the heat extractor accessory placed atop the OCT block. The Peltier element was activated for a ten minute cycle to ensure the embedded sample was completely frozen prior to sectioning. The specimen disc was inserted into the specimen head, rotated to orient vertically, before securing the disc using the locking screw. A razor blade was used to trim the block into a smooth rectangular shape to produce uniform sections. A pre-cooled sectioning blade was inserted into the blade holder, clamped into position and aligned with the sample outer edge of the sample. The handwheel was unlocked and turned at constant speed to begin cutting the sample. A thickness setting of 20 μm was used initially to trim the sample, after which the thickness was gradually reduced as low as was possible without compromising section quality. The chamber

was kept between -25°C and -30°C to produce optimal sections. A glass anti-roll plate was utilised to encourage sections to form ribbons. Suitable sections were adhered to glass slides by pressing the inverted slide against the stage.

2.13.7 Cryostat Section Staining

The slides were placed in a slide rack and soaked in a PBS bath for at least 20 minutes to dissolve the OCT. Toluidine blue stain was chosen to add contrast and highlight structures of interest. Due to failure of 1% toluidine blue solution to successfully stain OCT sections, an alternative solution was prepared using ethanol and urea (see Appendix I). Drops of this solution were applied to individual sections on each slide using an autopipette. Care was taken to ensure each section was completely immersed. After approximately two minutes, each slide was held upside-down under a gently running tap to remove the toluidine blue solution. The slides were placed upright on a paper towel to dry the undersides, before using an Axiostar Plus compound microscope (Zeiss) to assess the quality of sectioning and staining. Good quality sections were preserved by applying a drop of Shandon Immu-Mount™ (Thermo Scientific) to each section, taking care not to introduce air bubbles, before placing a 22x40 mm cover slip (Menzel-Glaser) on each slide. Slides were left upright overnight to allow the mounting solution to set completely.

2.13.8 Photomicroscopy

A DMRE microscope (Leica) with a DP70 camera attachment (Olympus) was used to photograph histology slides. The microscope was equipped with HC Plan 25mm 10x eyepieces and HC Plan Fluotar objectives, of which three were used in this study: 5x, 10x and 20x. The upper and lower filter prisms were set to bright field. Light brightness was set no higher than 6 to ensure the longevity of the bulb. To prevent

accidental collision of the slide with the objectives, the motorised focus upper threshold was calibrated at the beginning of each session. This was done by placing a slide on the stage and, under the 10x objective, using the manual focusing wheel to bring the slide into focus. The upper threshold calibration button was held once to activate set mode, then again to set the threshold to the current focus position. Slides were positioned using the X-Y control knobs. When sections were positioned and focused correctly, microphotographs were captured using DP imaging software (Olympus) on a computer attached to the DP70 camera. The exposure was adjusted according to the objective used and section thickness. The white balance tool and levels adjustment tool were used to maximise image quality.

Chapter 3
Results

3.1 Larval Collection

The mean standard lengths of collected *S. lalandi* larvae are shown in Table 14. The external morphology of a selection of developing larvae is shown in Figure 5.

Table 14. Mean standard lengths of *S. lalandi* larvae collected at NIWA. No measurement data were provided by NIWA for 30 dph larvae.

Age (dph)	Mean standard length (mm)	Standard deviation (mm)
3	4.97	0.070
6	5.17	0.170
9	5.70	0.191
12	7.37	0.263
15	6.92	0.527
18	8.21	0.602
21	9.68	1.286
25	9.81	1.041
30	-	-
35	18.95	3.328
40	22.75	3.722
45	39.10	4.483
50	54.99	6.928
60	66.96	11.498

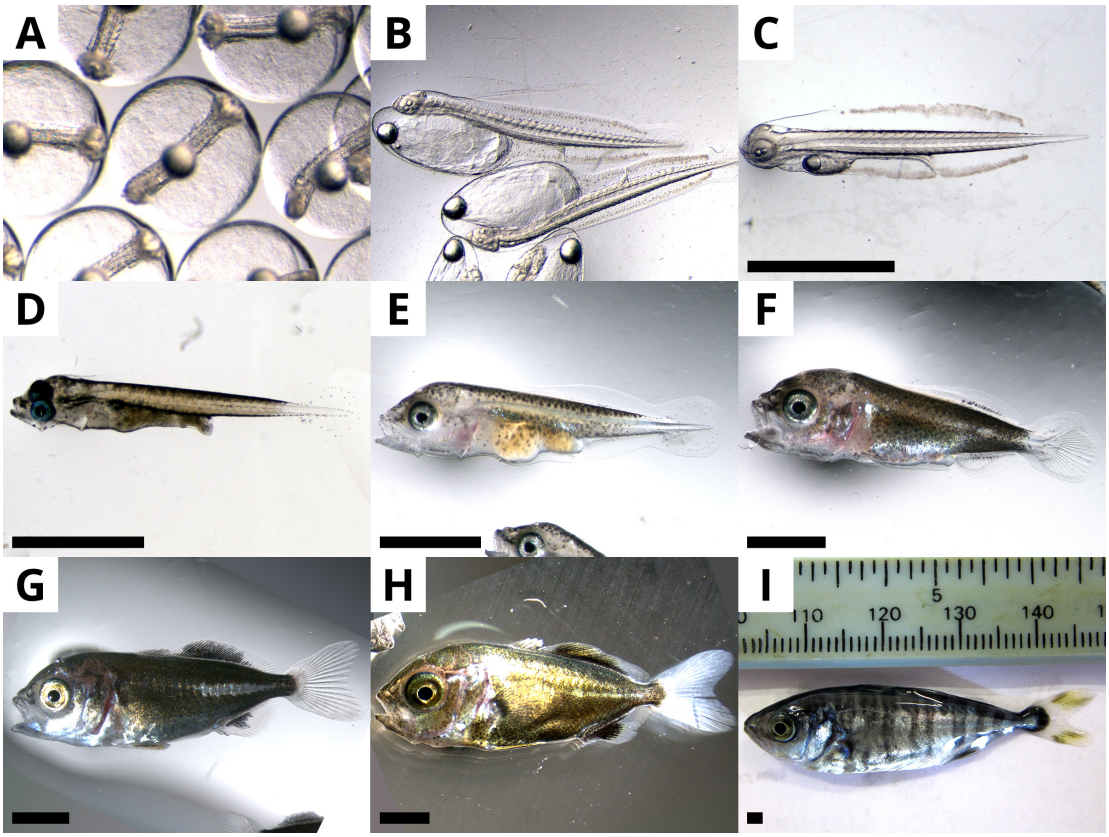


Figure 5. External development of larval *S. lalandi*. Photographs provided by A. King, NIWA, 2013. Scale bar = 2 mm. **A:** Developing embryos, 24 hours after fertilization. **B:** Newly hatched yolk-sac larvae. **C:** 3 dph yolk-sac larva. **D:** 6 dph preflexion larva with open mouth and increasing pigmentation. **E:** 12 dph preflexion larva with full gut. **F:** 18 dph flexion larva. **G:** 26 dph postflexion larva. **H:** 32 dph postflexion larva. **I:** 46 dph transformation stage larva undergoing metamorphosis to juvenile.

3.2 Bioinformatics

Known amino acid sequences for the sex differentiation genes *Vasa, Amh, Cyp19a1a, Dmrt1, Sox9b* and *Dax1* in Perciformes were used to BLAST search the gonad and pituitary Ion Torrent RNA-Seq library. Sufficient reads (up to 500) were obtained for *Vasa, Amh, Dax1* and *Dmrt1* to assemble a nucleotide sequence covering most of the gene. The BLAST search results for *Cyp19a1a* only returned 25 reads while *Sox9b* returned 173 reads; the assembled nucleotide sequences were short and only covered a small region of the gene. However, a partial amino acid sequence for *Cyp19a1a* in *S. lalandi* was subsequently found on the UniProt database.

3.2.1 *Vasa*

VASA is very well described in fish; amino acid sequences are available on the UniProt database for numerous Perciformes, notably including the Carangids *Seriola quinqueradiata* and *Trachurus japonicus*. The *S. quinqueradiata* VASA amino acid sequence was used to BLAST search the *S. lalandi* gonad and pituitary Ion Torrent RNA-Seq library. The resulting nucleotide sequences were subsequently translated and aligned to the *S. quinqueradiata* VASA amino acid sequence, showing 84.2% amino acid identity (Figure 6). The consensus sequence was used to design primers for RACE PCR and qPCR, and is detailed in Appendix II.

3.2.2 *Amh*

AMH is well described in fish, with amino acid sequences available on the UniProt database for numerous Perciformes. The Nile tilapia (*O. niloticus*) and freshwater angelfish (*Pterophyllum scalare*) AMH amino acid sequences were used to BLAST search the *S. lalandi* gonad and pituitary Ion Torrent RNA-Seq library. The resulting nucleotide sequences were translated and aligned to the tilapia AMH amino acid sequence, showing 38.9% amino acid identity (Figure 7). The consensus sequence was used to design primers for RACE PCR and qPCR, and is detailed in Appendix II.

3.2.3 *Dmrt1*

Only a small number of gene sequences for *Dmrt1* in Perciformes were available on the UniProt database at the time of this study. The sablefish (*Anoplopoma fimbria*) DAX1 amino acid sequence was used to BLAST search the *S. lalandi* gonad and pituitary Ion Torrent RNA-Seq library. The resulting nucleotide sequence was subsequently translated and aligned to the sea bass DMRT1 amino acid sequence, showing 51.5% amino acid identity (Figure 8). The consensus sequence is detailed in Appendix II.

```
S_quinqueradiata      MDEWEEEETSTSSVTLTSHASSQGSQGDFMNSDGEFGRGRGGRGRGRRGGFSSSFSSGGD
S_lalandi             ------------------------------------------------------------


S_quinqueradiata      EHGNGGDSWNNTGGERDGFRGRGGRGRGRGFGGMDCIEFGGDDGVCENGFRGGSRGGRGS
S_lalandi             ----------------------------------------GGDDGVCENGFRGGSRGGRGS
                                                               ********************

S_quinqueradiata      RGGRGFRQGGDQGGRGGFGGGYRGKDEQIFARGEDKDSEKKDGNDGDRPKVTYVPPTLPE
S_lalandi             RGGRGFRQGGDQGGRGGFGGGYRGKDEQIFARGEDKDPEKKDGNDGDRPKVTYVPPTLPE
                      ************************************ **********************

S_quinqueradiata      DEDSIFAHYESGINFNKYDDILVDVSGTNPPQAIMTFAEAALCESLSKNVSKSGYVKPTP
S_lalandi             DEDSIFAHYESGINFNKYDDILVDVSGTNPPQAIMTFAEAALCESLSKNVSKSGYVKPTP
                      ************************************************************

S_quinqueradiata      VQKHGIPIISAGRDLMACAQTGSGKTAAFLLPILQQLMADGVAASQFSELQEPEAIIVAP
S_lalandi             VQKHGIPIISAGRDLMACAQTGSGKTAAFLLPILQQLMADGVAASQFSELQEPEAIIVAP
                      ************************************************************

S_quinqueradiata      TRELINQIYLEARKFAFGTCVRPVVVYGGVSTGHQIREICRGCNVLCGTLGRLLDVIGRG
S_lalandi             TRELINQIYLEARKFAFGTCVRPVVVYGGVSTGHQIREICRGCNVLCGTPGRLLDVIGRG
                      ********************************************** **********

S_quinqueradiata      KVGLHKLRYLVLDEADRMLDMGFEPDMRRLVGSPGMPSKEKRQTLMFSATYPEDIQRMAA
S_lalandi             KVGLHKLRYLVLDEADRMLDMGFEPDMRRLVGSPGMPSKEQRQTLMFSATYPEDIQRMAA
                      ************************************ ******************

S_quinqueradiata      DFLKTDYLFLAVGVVGGACSDVEQKFIEVTKFSKREQLLDILKTTGTERTMVFVETKRQA
S_lalandi             DFLKTDYLFLAVGVVGGACSDVEQKFIEVTKFSKREQLLDILKTTGTERTMVFVETKRQA
                      ************************************************************

S_quinqueradiata      DFIATFLCQEKVPTTSIHGDREQREREQALADFRSGRCPVLVATSVAARGLDIPDVQHVV
S_lalandi             DFIATFLCQEKVPTTSIHGDREQREREQALADFRSGRCPVLVATSVAARGLDIPDVQHVV
                      ************************************************************

S_quinqueradiata      NFDLPNNIDEYVHRIGRTGRCGNTGRAVSFYDPDNDGQLAGSLVSILSKAQQEVPSWLEE
S_lalandi             NFDLPNNIDEYVHRIGRTGRCGNTGRAVSFYDPDNDGQLAGSLVSILSKAQQEVPSWLEE
                      ************************************************************

S_quinqueradiata      CVFSGSGVNPSRRTFASTDSRKGPQGSSFQDSSMTSQPAVPAAADNEDWE
S_lalandi             CVFSGSGVNPSRRTFASTDSRKGPQGSSFQDSSMTSQPAVPAAADDEDWE
                      ******************************************** ****

  [*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 6. Clustal Omega 1.2.0 pairwise alignment of translated nucleotide sequences constructed from the *S. lalandi* gonad transcriptome, to the *S. quinqueradiata* VASA amino acid sequence (UniProt accession number D4P9H5).

```
O_niloticus      MLGLLVLYSEALTLCWTLQPAQDPTVTEYSLPSAKTPSSPSSSSAAAPHAAPCFVEDIFA
S_lalandi        ------------------------------------------------------------


O_niloticus      ALRDGVGDSGELTNSSLVLFGFCSQSARSSASVSLDLANKKSSLEVLHPAAVHVSEEEEQ
S_lalandi        -----------------------------------------------------MKKSLKDLL-


O_niloticus      GTITLTFDLPRPPSLMTNPVLLLVFENPLARGDLEVAFTSQFLQPNTQAVCISGDTQYVL
S_lalandi        ------------------------------------------------------------


O_niloticus      LTGKSSEGSVNDRWQITAQTKLPHMKQNLKSILIGEKSGSNISMSPLLLFSGGTGTDTRC
S_lalandi        --------------------------------IGEKTGSTISITLLLLFSGGTGTDTRQ
                                                 **** ** **   ************


O_niloticus      ASG--SPPASLQTSFLCE-MKRFLGAVLPQEHFTSPPLPLDSLQSLPPLSLGLSSSETLL
S_lalandi        THVSGSSLASSQTFSFLCELKRFLGEVLPQXHPESPPFQLESLPSMPPLSLGLSSSETLL
                      *  ** **       ***** **** *  ***  * ** * **************


O_niloticus      AVMINSTAPTVFGFTSWGSVLPVCHGELALSAALLEELRQRLDQTLVQMTEIIREEEVSL
S_lalandi        AGLINSSFLTVFSFARWGSISQVHPGQLAMSPALVEEVRQRLADXXMQIMEVIREKEVGH
                 *  ***   *** *  ***   *  * ** * ** ** ****     *  * *** **


O_niloticus      GAKESLGRLKELSALQEKEHATGGSQFRVFLLLKALQTVAQTYDAQRKLRATRADPSSSV
S_lalandi        RATERLERLQEFSALPTMEQAAGRSQYCAFLLLKALQTVAREYELQRGLRATRADTNNPV
                  * * * ** * ***   * * * **   ***********  *  ** *******    *


O_niloticus      RGGVCGLKALTVSLTKLLVGPSSANINNCHGSCTFPLTNGNNHAILLNSHIETGNADERS
S_lalandi        RGSVCSVRSLTVNLERRLVGPSAANIKNCHGSCAFPLTNANDHVVLLHFHIESGNMDERA
                 ** **    *** *   ***** *** ****** ***** * *  **  *** ** ***


O_niloticus      PCCVPVAYEALEVVDWNADGTFISIKPDAVARECGCR
S_lalandi        PCCVPVAYEPLEVVGLNEHGTYLSMIPDMVAKECGCR
                 ********* ****  *  **  *  ** ** *****

[*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 7. Clustal Omega 1.2.0 pairwise alignment of translated nucleotide sequences constructed from the *S. lalandi* gonad transcriptome, to the *O. niloticus* AMH amino acid sequence (UniProt accession number A1XPM9).

3.2.4 *Dax1*

Only a small number of gene sequences for *Dax1* in Perciformes were available on the UniProt database at the time of this study. The European sea bass (*D. labrax*) DAX1 amino acid sequence was used to BLAST search the *S. lalandi* gonad and pituitary Ion Torrent RNA-Seq library. The resulting nucleotide sequences were subsequently translated and aligned to the sea bass DAX1 amino acid sequence, showing 41.4% amino acid identity (Figure 9). The consensus sequence is detailed in Appendix II.

```
A_fimbria      MSKDKPSKQVPECPGPLSPAKGQKPPRMPKCSRCRNHGYVSPLKGHKRFCNWRDCQCPKC
S_lalandi      ------------------------PPRMPKCSRCRNHGYVSPLKGHKRFCNWKDCQCPKC
                                       ****************************:*******

A_fimbria      KLIAERQRVMAAQVALRRQQAQEEELGICSPVALSGPEVMVKNEAAADCLFSVDGRSPSP
S_lalandi      KLIAERQRVMAAQVALRRQQAQEEELGICSPMTLSVPEVLVKNEAGADCLFSVEGRSLTP
               ********************************::** ***:*****.*******:*** :*

A_fimbria      TP----SLAVTGGRSALSPSLSVGSRSHTDGSSDLLLETSYYNFYQPSRYPAYYGNLYNY
S_lalandi      TSISTSAVAVTGSRSASSSSPSAAARAHTEGASDLLMETSYYNFYQPPRYXCLLQQPLQL
               *     ::****.*** * * *..:*:**:*:****:********** ** .   :  :

A_fimbria      QQYQQMPHADGR----LS-SHNMSSQYRMHSYY-PAGTYL-TQG-LGSTTCVPPL-----
S_lalandi      PA---IPAGAPW-WPTVRPQHVLSV---PHAFLLPSQCXX-XRAWAPPPVCRPSSAWRTT
                    :* .   *   :  .* :*     *::  *:    * :.     .* *

A_fimbria      --FSLDDNNNNNCSETMATSFSPSSTTTGHDSTMTCRSIMVNSD-VNACEAGGETANFTV
S_lalandi      VTTTTTTTTTTIAXETMAASFSPSSITGAHDSTMTCGSIISLVNSVVKPEASDETANFSN
                  :   .... . ****:****** * .******* **:   : *   **. *****:

A_fimbria      GSIIDGDAAK
S_lalandi      S---------

[*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 8. Clustal Omega 1.2.0 pairwise alignment of translated nucleotide sequences constructed from the *S. lalandi* gonad transcriptome, to the *A. fimbria* DMRT1 amino acid sequence (UniProt accession number M4TB31).

```
D_labrax       MATLEGCRCRGASGRNNNSILYNILKSDSLATAEEQQQQQQQQQQQQQQHPQQQQTLQQLF
S_lalandi      ------------------------------------------------------------


D_labrax       HKTSSTAAPASLQELRQQTCSCGSTRRRGILRFPQVTCKA----ASAVLVKTLRFVKNVP
S_lalandi      -------------------------RLHTAPRCPSLPAGDVQSRVGGSGEDAAVRXKTSP
                                         * :   * *.: .      ...   .:    *. *

D_labrax       CFRELPEDDQLMLIRSGWAPLLVLGLAQDRVDFETTETVEPSMLQRILTGLPDRQSEVPA
S_lalandi      CFRELPEDDQLMLIRSGWAPLLVLGLAQDRVDFETTETVEPSMLQRILTGLPDRQSEALA
               ************************************************** *

D_labrax       GQ-SRGAAGVSVVDIEAIKAFLKKCWSVDISTKEYAYLKGAVLFNPDVEGLRCLHYIQSL
S_lalandi      GQSSRGAVGVSVVDIEAIKAFLKKCWSVDISTKEYAYLKGAVLFNPDLEGLXLPPLHPVS
               ** ****.****************************************:***

D_labrax       RREAHQALNEH--VRLIHREDTTRFAKLL-I-ALS-MLRAINPLVVAQLFFRPVIGAVNI
S_lalandi      ASGGAPGSERACQADPSRRHNTVRQAAHSSVHAEGHQPAGGRPALLQTRH----------
                  .  . :.   .   :*.:*.* *    : * .    . .* ::   .

D_labrax       EEVLMEMFYGK
S_lalandi      -----------

[*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 9. Clustal Omega 1.2.0 pairwise alignment of translated nucleotide sequences constructed from the *S. lalandi* gonad transcriptome, to the *D. labrax* DAX1 amino acid sequence (UniProt accession number Q59AE8).

3.2.5 *Sox9b*

A reasonable number of amino acid sequences for *Sox9b* in Perciformes were available on the UniProt database. The gilthead sea bream (*Sparus aurata*) SOX9B amino acid sequence was used to BLAST search the *S. lalandi* gonad and pituitary Ion Torrent RNA-Seq library. The resulting nucleotide sequences were subsequently translated and aligned to the sea bream amino acid sequence, showing 14.9% amino acid identity (Figure 10). The consensus sequence is detailed in Appendix II.

```
S_aurata       MNLLDPYLKMTEEQEKCHSDAPSPSMSEDSAGSPCPSGSGSDTENTRPSDNHLLLGAEYK
S_lalandi      ------------------------------------------------------------


S_aurata       KEGEEEKFPVCIRDAVSQVLKGYDWTLVPMPVRVNGSSKNKPHVKRPMNAFMVWAQAARR
S_lalandi      ----------------------------------------MDKVKRPMNAFMVWSRGQRR
                                                       :***********::. **

S_aurata       KLADQYPHLHNAELSKTLGKLWRLLNEVEKRPFVEEAER-LRVQHKKDHPDYKYQPRRRK
S_lalandi      KMAQENPKMHNSEISKRLGAEWKLLTDAEKRXLHRRGQXGCGPST-RSXXXYKYXPRRKT
               *:*:: *::**:*:** **  *:**.:.*** : ...:     .  :.   *** ***:.

S_aurata       SVKNGQNDPEDGEQTHISPNAIFKALQQADSPASSLGEVHSPGEHSGQSQGPPTPPTTPK
S_lalandi      XAAAQEG-RSGGQVPAVSG----------EPAG--GGGAGPGW-SED--G---------
                .   :   . *:   :*           .**.  *   .**  * :  *

S_aurata       TDLPASKADLKREGRPMQEGTSRQLNIDFGAVDIGELSSEVISNMGSFDVDEFDQYLPPH
S_lalandi      ------------------------------------------------------------


S_aurata       SHAGMPGVTQAGYTGSYGISSSSVGQGPNVGAHAWMSKQQQQQQQQHSLTTLGGGGEQGQ
S_lalandi      ------------------------------------------------------------


S_aurata       QGQQRTTQIKTEQLSPSHYSEQQGSPQHVTYGSFNLQHYSASSYPSITRAQYDYSDHQGG
S_lalandi      ------------------------------------------------------------


S_aurata       ANSYYSHAAGQGSGLYSTFSYMSPSQRPMYTPIADNTGVPSVPQTHSPQHWEQQPIYTQL
S_lalandi      ------------------------------------------------------------


S_aurata       SRP
S_lalandi      ---

[*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 10. Clustal Omega 1.2.0 pairwise alignment of translated nucleotide sequences constructed from the *S. lalandi* gonad transcriptome, to the *S. aurata* SOX9B amino acid sequence (UniProt accession number R4IPZ8).

3.2.6 *Cyp19a1a*

A partial amino acid sequence for gonadal aromatase in *S. lalandi* was available on the UniProt database (accession number I1SV64) (Figure 11). The corresponding nucleotide sequence, obtained from the European Nucleotide Archive (ENA; accession number HQ449733), is listed in Appendix II and was used to design primers for RACE PCR and qPCR.

A number of amino acid sequences for aromatase in other Perciformes were also available on the UniProt datase, although for most it was not specified if they were for the gonadal aromatase (*Cyp19a1a*) or brain aromatase (*Cyp19a1b*) genes.

```
WINGEETLIISRPSAVHHVLKNGLYCSRFGSKQGLSCVGMNERGLIFNNNVTLWKKMRTYFTKALTGP
GLQKTVEVCVSSTQTHLDDLQSLSHVDVLGLLRCIVVDISNRLFLGVPVDEKELLVKIHKYFDTWQSV
LIKPDIYFKLDWIHRRHKTAAQELRDAIESLVEQKRRDVEQADKLDNINFTTELIFAQNHGELSAENV
LQCVLEM
```

Figure 11. Partial amino acid sequence of gonadal aromatase in *S. lalandi*, obtained from the UniProt database (I1SV64).

3.3 Sequence Amplification and Cloning

3.3.1 Validation of cDNA Quality

The quality of adult *S. lalandi* ovary and testes cDNA was verified by PCR using housekeeping primers specific for β-actin (slBACTIN-rtF and slBACTIN-rtR). The standard PCR programme was followed (Table 6), using an annealing temperature of 52°C and an extension time of 45 seconds. The expected band size was 476 bp; a strong band of approximately this size was observed in ovary cDNA, while a weak band of similar size was observed in testes cDNA (Figure 12). Similarly, the quality of adult *S. lalandi* ovary and testes 3′ RACE cDNA was verified by PCR using β-actin housekeeping primers (slBACTIN-rtF and slBACTIN-rtR). The standard PCR

programme was followed (Table 6), using an annealing temperature of 52°C and an extension time of 30 seconds. Bands of ~330 bp was observed in both ovary and testes cDNA (Figure 13).

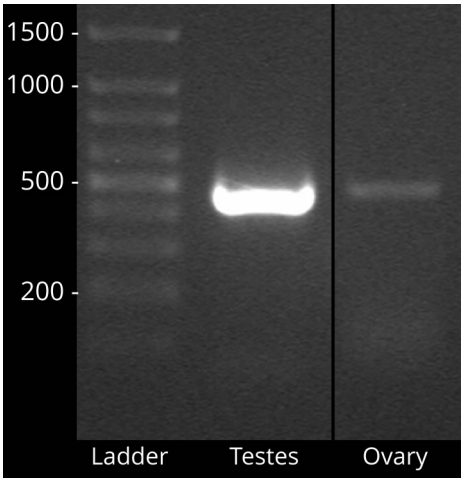


Figure 12. Gel electrophoresis analysis of β-actin in testes and ovary cDNA. Bands of expected size (~470 bp) were obtained for both tissues.

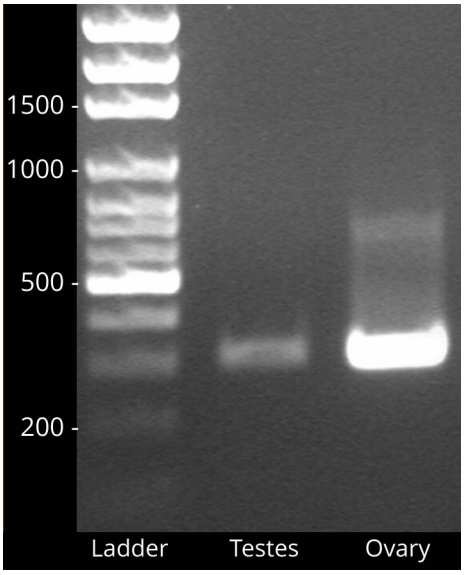


Figure 13. Gel electrophoresis analysis of β-actin in testes and ovary 3′ RACE cDNA. Bands of ~330 bp were obtained for both tissues.

3.3.2 *Vasa*

Primers specific to *Vasa* (slVASA-F1-4, R1-4) were used to amplify sections of ovary and testes cDNA using PCR. Combinations of F1-2 and R1-2 primers were expected to amplify fragments of ~1350-1580 bp, but despite extensive optimisation, only generated smaller non-specific products and smearing (Figure 14). This led to the F3-4 and R3-4 primers being designed, which targeted smaller fragments of ~150-290 bp. In testes cDNA, combinations of F3-4 and R3-4 primers produced weak bands of expected size as well as larger non-specific products (Figure 15). By optimising the annealing temperature to 52°C, F3+R3 primers produced a strong specific band of ~290 bp (Figure 16). This band was gel purified and ligated for cloning. Only weak bands were obtained for ovary cDNA.

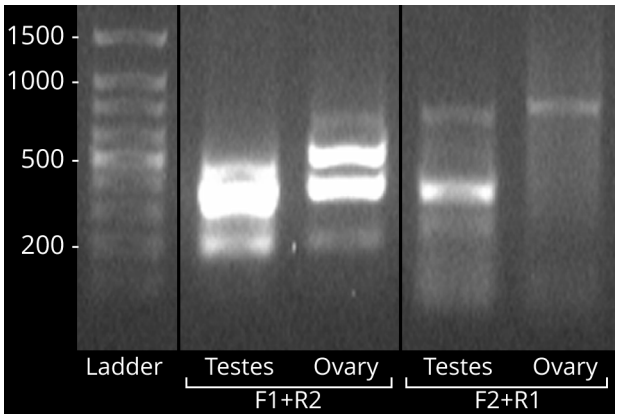


Figure 14. Gel electrophoresis analysis of second round nested PCR of *Vasa* in ovary and testes cDNA, using combinations of slVASA-F1-2 and R1-2 primers. Only small non-specific bands were produced.

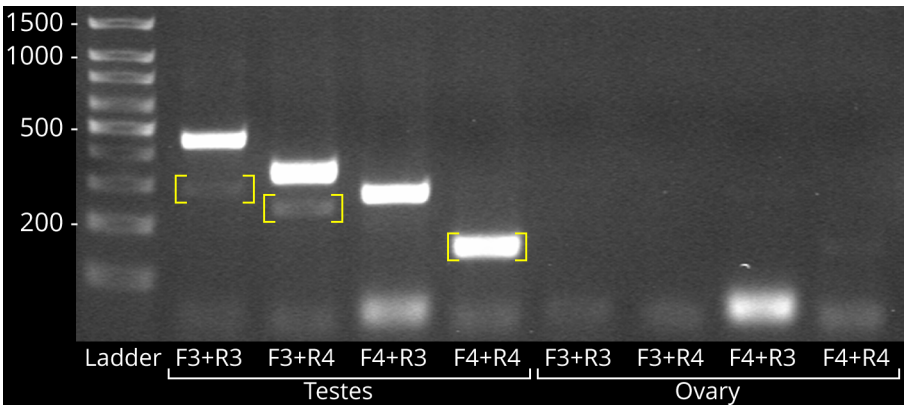


Figure 15. Gel electrophoresis analysis of *Vasa* in ovary and testes cDNA, using combinations of slVASA-F3-4 and R3-4 primers. For testes, weak bands of expected size were produced for F3+R3 and F3+R4 (highlighted), as well as strong bands of larger non-specific products. The F4+R4 primers produced a strong band of expected size for testes (highlighted), but only a very weak band for ovary.

The ligated product of *Vasa* F3+R3 from Figure 16 was used to transform chemically competent cells, which were visualised using blue and white colony screening. Colony PCR revealed which white colonies contained correctly sized inserts of ~460 bp. Of the ten colonies screened, nine showed bands of expected size (Figure 17). Colonies 1-5 were cultured overnight for plasmid extraction. Colony #4 failed to grow and was omitted. Following extraction, the concentration and quality of each

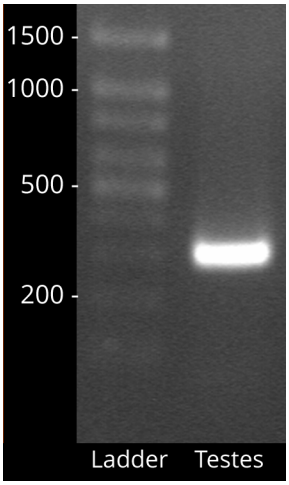


Figure 16. Gel electrophoresis analysis of *Vasa* in testes cDNA, using slVASA-F3 and slVASA-R3 primers. A strong specific band of ~290 bp matched the expected size for this primer combination.

plasmid was measured on the NanoDrop (Table 15). The presence of a correctly sized insert in the plasmids was verified by restriction digest. All four digest products showed weak but correctly sized bands of ~350 bp, in addition to the larger band of >1500 bp which represented the remaining plasmid DNA (Figure 18). All four plasmids were sent for sequencing.

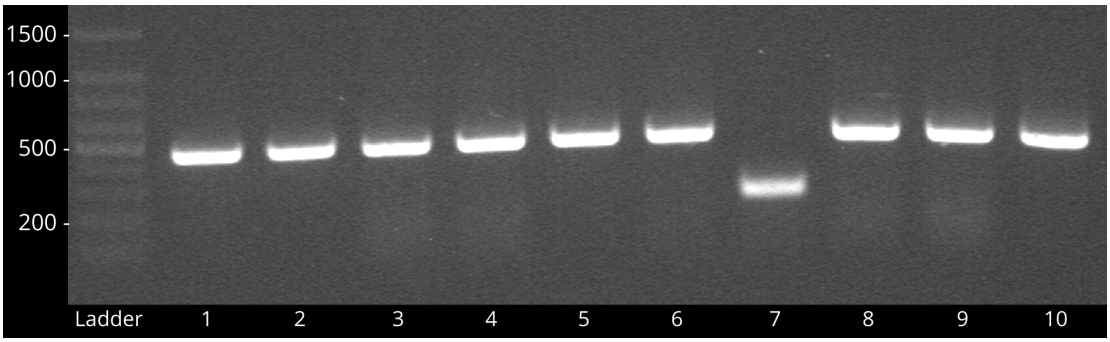


Figure 17. Gel electrophoresis analysis of *Vasa* products from colony screening, using M13 forward(-40) and reverse primers. With the exception of colony #7, all colonies produced bands of expected size and were cultured overnight for plasmid extraction.

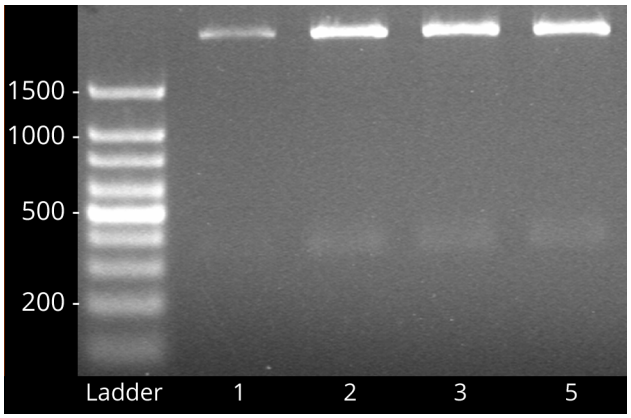


Figure 18. Gel electrophoresis analysis of *Vasa* restriction digest. All four products show a faint but correctly sized band of ~350 bp, representing the product of slVASA-F3+R3 primers. The larger band at >1500 bp represents the remaining plasmid DNA.

Table 15. Nanodrop 2000 analysis of *Vasa* plasmid DNA concentration and quality.

Plasmid DNA	DNA concentration (ng/µl)	260/280	260/230
Vasa F3+R3 Colony 1	162.5	1.91	2.08
Vasa F3+R3 Colony 2	319.6	1.85	2.01
Vasa F3+R3 Colony 3	245.8	1.87	2.17
Vasa F3+R3 Colony 5	276.2	1.85	2.16
Vasa 3'F1+R1 Colony 4b	177.0	1.88	2.21
Vasa 3'F1+R1 Colony 5b	238.6	1.80	1.44
Vasa 3'F1+R1 Colony 7	366.7	1.79	1.17
Vasa 3'F1+R1 Colony 8	246.6	1.97	2.35
Vasa 3'F1+R1 Colony 9	371.5	1.80	1.18

The similarity of these sequences was confirmed by multiple alignments, alongside the *Vasa* nucleotide sequence constructed from the Ion Torrent transcriptome data (Figure 19).

```
VASA-IT      CCAGATCAGGGAAATATGCAGAGGATGCAACGTCCTGTGTGGAACTCCAGGGAGACTGCT
VASA-1       ---GATCAGGG-AATATGCAGAGGATGCAACGTCCTGTGTGGAACTCCAGGGAGACTGCT
VASA-2       ---GATCAGGG-AATATGCAGAGGATGCAACGTCCTGTGTGGAACTCCAGGGAGACTGCT
VASA-3       ---GATCAGG-AAATATGCAGAGGATGCAACGTCCTGTGTGGAACTCCAGGGAGACTGCT
VASA-5       ---GATCAGGG-AATATGCAGAGGATGCAACGTCCTGTGTGGAACTCCAGGGAGACTGCT
                  *******  **********************************************

VASA-IT      GGATGTGATTGGACGAGGAAAGGTGGGGCTCCACAAGCTGCGCTACCTGGTGCTGGATGA
VASA-1       GGATGTGATTGGACGAGGAAAGGTGGGGCTCCACAAGCTGCGCTACCTGGTGCTGGATGA
VASA-2       GGATGTGATTGGACGAGGAAAGGTGGGGCTCCACAAGCTGCGCTACCTGGTGCTGGATGA
VASA-3       GGATGTGATTGGACGAGGAAAGGTGGGGCTCCACAAGCTGCGCTACCTGGTGCTGGATGA
VASA-5       GGATGTGATTGGACGAGGAAAGGTGGGGCTCCACAAGCTGCGCTACCTGGTGCTGGATGA
             ************************************************************

VASA-IT      GGCCGACCGTATGCTGGATATGGGCTTTGAGCCTGACATGCGCCGCCTGGTGGGCTCCCC
VASA-1       GGCCGACCGTATGCTGGATATGGGCTTTGAGCCTGACATGCGCCGCCTGGTGGGCTCCCC
VASA-2       GGCCGACCGTATGCTGGATATGGGCTTTGAGCCTGACATGCGCCGCCTGGTGGGCTCCCC
VASA-3       GGCCGACCGTATGCTGGATATGGGCTTTGAGCCTGACATGCGCCGCCTGGTGGGCTCCCC
VASA-5       GGCCGACCGTATGCTGGATATGGGCTTTGAGCCTGACATGCGCCGCCTGGTGGGCTCCCC
             ************************************************************

VASA-IT      TGGAATGCCCTCCAAAGAGCAGCGTCAGACCCTGATGTTCAGCGCAACCTACCCCGAGGA
VASA-1       TGGAATGCCCTCCAAAGAGCAGCGTCAGACCCTGATGTTCAGCGCAACCTACCCCGAGGA
VASA-2       TGGAATGCCCTCCAAAGAGCAGCGTCAGGCCCTGATGTTCAGCGCAACCTACCCCGAGGA
VASA-3       TGGAATGCCCTCCAAAGAGCAGCGTCAGGCCCTGATGTTCAGCGCAACCTACCCCGAGGA
VASA-5       TGGAATGCCCTCCAAAGAGCAGCGTCAGACCCTGATGTTCAGCGCAACCTACCCCGAGGA
             ***************************.********************************

VASA-IT      CATCCAGAGGATGGCAGCTGACTTCCTCAAGACCGACTACCTGTTCTTAGCTGTGGGC
VASA-1       CATCCAGAAGATGGCAGCTGACTTCCTCAAGACCGACTACCTGTTCTTAGCTGTG---
VASA-2       CATCCAGAGGATGGCAGCTGACTTCCTCAAGACCGACTACCTGTTCTTAGCTGTG---
VASA-3       CATCCAGAGGATGGCAGCTGACTTCCTCAAGACCGACTACCTGTTCTTAGCTGTG---
VASA-5       CATCCAGAGGATGGCAGCTGACTTCCTCAAGACCGACTACCTGTTCTTAGCTGTG---
             ********.*********************************************

[*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 19. Clustal Omega 1.2.0 pairwise alignment of *S. lalandi Vasa* nucleotide sequences obtained by cloning (colonies 1a, 2a, 3a and 5a), and the corresponding fragment of *Vasa* nucleotide sequence obtained from the RNA-Seq transcriptome library (VASA-IT).

In order to obtain more of the *Vasa* gene sequence identified in the transcriptome, further PCRs were conducted using slVASA-3'F1 and slVASA-R1. The standard PCR protocol was followed (Table 6), but with an annealing temperature of 54°C and an extension time of one minute. Strong bands of approximately expected size (~690 bp) were obtained from testes cDNA (Figure 20). These bands were gel purified and ligated to transform chemically competent cells. Cell colonies were visualised using blue and white colony screening.

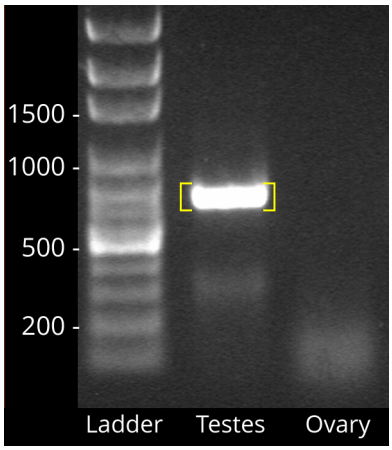


Figure 20. Gel electrophoresis analysis of *Vasa* in ovary and testes cDNA, using slVASA-3'F1 and slVASA-R1 primers. A strong specific band of ~690 bp (highlighted) in testes cDNA approximately matched the expected size for this primer combination.

Colony PCR revealed which white colonies contained correctly sized inserts of ~860 bp. Five colonies were initially screened, two colonies showing bands of approximately expected size. An image of this gel was not obtained due to temporary failure of the computer connected to the Cohu High Performance CCD Camera. A further four colonies were screened; all showed bands of approximately expected size, although the gel ran smeared due to incorrectly prepared TAE buffer (Figure 21). All six colonies were cultured overnight for plasmid extraction. The concentration and quality of each extracted plasmid was measured on the NanoDrop (Table 15). Restriction digest of the extracted plasmids revealed insert bands of ~740 bp (Figure 22).

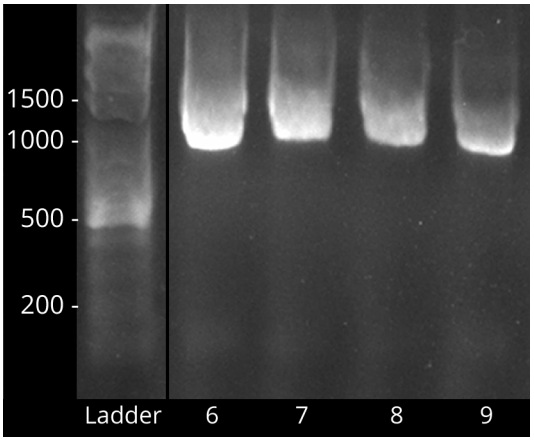


Figure 21. Gel electrophoresis analysis of *Vasa* products from colony screening, using M13 forward(-40) and reverse primers. All four colonies produced bands of approximately expected size (~860 bp) and were cultured overnight for plasmid extraction. The smearing of this gel was a result of incorrectly prepared TAE buffer; due to time constraints, the gel was not repeated.

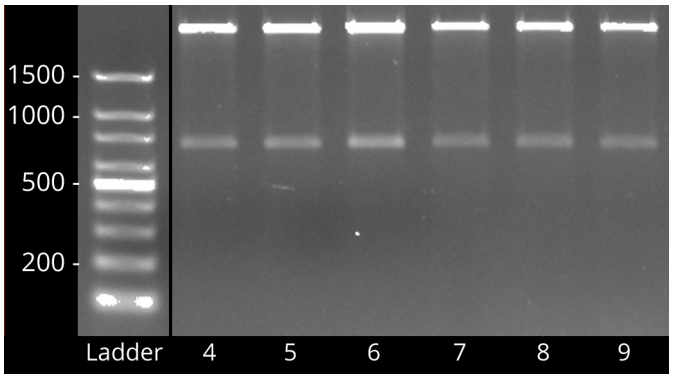


Figure 22. Gel electrophoresis analysis of *Vasa* restriction digest. All products show a correctly sized band of ~740 bp representing the insert. The larger band at >1500 bp represents the remaining plasmid DNA.

The similarity of these sequences was confirmed by multiple alignments, alongside the *Vasa* nucleotide sequence constructed from the Ion Torrent transcriptome data (Figure 23).

```
VASA-IT      CCCTGATGTTCAGCGCAACCTACCCCGAGGACATCCAGAGGATGGCAGCTGACTTCCTCA
VASA-4b      ---TGATGTTCAGCGCAACCTACCCCGAGGACATCCAGAGGATGGCAGCTGACTTCCTCA
VASA-5b      ---TGATGTTCAGCGCAACCTACCCCGAGGACATCCAGAGGATGGCAGCTGACTTCCTCA
VASA-7       ---TGATGTTCAGCGC-ACCTACCCCGAGGACATCCAGAGGATGGCAGCTGACTTCCTCA
VASA-8       ---TGATGTTCAGCGC-ACCTACCCCGAGGACATCCAGAGGATGGCAGCTGACTTCCT-A
VASA-9       ---TGATGTTCAGCGC-ACCTACCCCGAGGACATCCAGAGGATGGCAGCTGACTTCCTCA
                  ************ ****************************************** *

VASA-IT      AGACCGACTACCTGTTCTTAGCTGTGGGCGTGGTGGGCGGAGCTTGCAGTGATGTGGAGC
VASA-4b      AGACCGACTACCTGTTCTTAGCTGTGGGCGTGGTGGGCGGAGCTTGCAGTGATGTGGAGC
VASA-5b      AGACCGACTACCTGTTCTTAGCTGTGGGCGTGGTGGGCGGAGCTTGCAGTGATGTGGAGC
VASA-7       AGACCGACTACCTGTTCTTAGCTGTGGGCGTGGTGGGCGGAGCTTGCAGTGATGTGGAGC
VASA-8       AGA-AGGCTACCTGTTCTTAGCTGTGGGCGTGGTGGGCGGAGCTTG-AGTGGTGTGGAGC
VASA-9       AGACCGACTACCTGTTCTTAGCTGTGGGCGTGGTGGGCGGAGCTTGCAGTGATGTGGAGC
             *** .*.************************************ ****.********

VASA-IT      AGAAGTTTATTGAAGTAACTAAGTTCTCCAAGAGGGAGCAGCTTCTTGACATCCTGAAGA
VASA-4b      AGAAGTTTATTGAAGTAACTAAGTTCTCCAAGAGGGAGCAGCTTCTTGACATCCTGAAGA
VASA-5b      AGAAGTTTATTGAAGTAACTAAGTTCTCCAAGAGGGAGCAGCTTCTTGACATCCTGAAGA
VASA-7       AGAAGTTTATTGAAGTAACTAAGTTCTCCAAGAGGGAGCAGCTTCTTGACATCCTGAAGA
VASA-8       AGAAGTTTATTGAAGTAACTAAGTTCTCCAAGAGGGAGCAGCTTCTTGACATCCTGAAGA
VASA-9       AGAAGTTTATTGAAGTAACTAAGTTCTCCAAGAGGGAGCAGCTTCTTGACATCCTGAAGA
             ************************************************************

VASA-IT      CAAC-AGGAACAGAGCGCACCATGGTGTTTGTGGAGACCAAGAGACAGGCTGATTTTATT
VASA-4b      CAACA-GGAACAGAGCGCACCATGGTGTTTGTGGAGACCAAGAGACAGGCTGATTTTATT
VASA-5b      CAACCCCGCACAGAGCGCACCATGGTGTTTGTGGAGACCAAGAGACAGGCTGATTTTATT
VASA-7       CAACA-GGAACAGAGCGCACCATGGTGTTTGTGGAGACCAAGAGACAGGCTGATTTTATT
VASA-8       CAACA-GGAACAGAGCGCACCATGGTGTTTGTGGAGACCAAGAGACAGGCTGATTTTATT
VASA-9       CAACA-GGAACAGAGCGCACCATGGTGTTTGTGGAGACCAAGAGACAGGCTGATTTTATT
             ****   *.************************************************

VASA-IT      GCCACATTCTTGTGCCAGGAGAAGGTTCCAACTACCAGCATCCATGGGGACCGGGAGCAG
VASA-4b      GCCACATTCTTGTGCCAGGAGAAGGTTCCAACTACCAGCATCCATGGGGACCGGGAGCAG
VASA-5b      GCCACATTCTTGTGCCAGGAGAAGGTTCCAACTACCAGCATCCATGGGGACCGGGAGCAG
VASA-7       GCCACATTCTTGTGCCAGGAGAAGGTTCCAACTACCAGCATCCATGGGGACCGGGAGCAG
VASA-8       GCCACATTCTTGTGCCAGGAGAAGGTTCCAACTACCAGCATCCATGGGGACCGGGAGCAG
VASA-9       GCCACATTCTTGTGCCAGGAGAAGGTTCCAACTACCAGCATCCATGGGGACCGGGAGCAG
             ************************************************************

VASA-IT      CGGGAACGGGAGCAGGCCCTGGCAGACTTCCGCTCTGGGAGATGTCCAGTCCTGGTGGCA
VASA-4b      CGGGAACGGGAGCAGGCCCTGGCAGACTTCCGCTCTGGGAGATGTCCAGTCCTGGTGGCA
VASA-5b      CGGGAACGGGAGCAGGCCCTGGCAGACTTCCGCTCTGGGAGATGTCCAGTCCTGGTGGCA
VASA-7       CGGGAACGGGAGCAGGCCCTGGCAGACTTCCGCTCTGGGAGATGTCCAGTCCTGGTGGCA
VASA-8       CGGGAACGGGAGCAGGCCCTGGCAGACTTCCGCTCTGGGAGATGTCCAGTCCTGGTGGCA
VASA-9       CGGGAACGGGAGCAGGCCCTGGCAGACTTCCGCTCTGGGAGATGTCCAGTCCTGGTGGCA
             ************************************************************

VASA-IT      ACCTCTGTTGCTGCCCGTGGTCTGGATATTCCAGATGTTCAGCATGTGGTGAACTTTGAC
VASA-4b      ACCTCTGTTGCTGCCCGTGGTCTGGATATTCCAGATGTTCAGCATGTGGTGAACTTTGAC
VASA-5b      ACCTCTG-------------------------------------------------------
VASA-7       ACCTCTGTTGCTGCCCGTGGTCTGGATATTCCAGATGTTCAGCATGTGGTGAACTTTGAC
VASA-8       ACCTCTGTTGCTGCCCGTGGTCTGGATATTCCAGATGTTCAGCATGTGGTGAACTTTGAC
VASA-9       ACCTCTGTTGCTGCCCGTGGTCTGGATATTCCAGATGTTCAGCATGTGGTGAACTTTGAC
             *******

VASA-IT      CTCCCCAACAATATAGACGAATACGTCCACCGTATTGGGAGAACTGGCCGCTGTGGAAAC
VASA-4b      CTCCCCAACAATATAGACGAATACGTCCACCGTATTGGGAGAACTGGCCGCTGTGGAGAC
VASA-5b      ------------------------------------------------------------
VASA-7       CTCCCCAACAATATAGACGAATACGTCCACCGTATTGGGAGAACTGGCCGCTGTGGAAAC
VASA-8       CTCCCCAACAATATAGACGAATACGTCCACCGTATTGGGAGAACTGGCCGCTGTGGAAAC
VASA-9       CTCCCCAACAATATAGACGAATACGTCCACCGTATTGGGAGAACTGGCCGCTGTGGAAAC


VASA-IT      ACTGGAAGGGCGGTGTCGTTCTATGACCCCGACAATGATGGA-CAGCTGGCTGGGTCCCT
VASA-4b      ACTGGAAGGGCGGTGTCGTTCTATGACCCCGACAATGATGGA-CAGCTGGCTGGGTCCCT
VASA-5b      ------------------------------------------------------------
VASA-7       ACTGGAAGGGCGGTGTCGTTCTATGACCCCGACAATGATGGA-CAGCTGGCTGGGTCCCT
VASA-8       ACTGGAAGGGCGGTGTCGTTCTATGACCCCGACAATGATGCCCCCGCTGCCTGGGTCCTT
VASA-9       ACTGGAAGGGCGGTGTCGTTCTATGACCCCGACAATGATGGA-CAGCTGGCTGGGTCCCT
```

```
VASA-IT      TGTCAGCATCCTGTCCAAGGCCCAGCAGGAAGTGCCCTCCTGGTTAGAGGAGTGTGTGTT
VASA-4b      TGTCAGCATCCTGTCCAAGGCCCAGCAGGAAGTGCCCTCCTGGTTAGAGGAGTGTGTGTT
VASA-5b      ------------------------------------------------------------
VASA-7       TGTCAGCATCCTGTCCAAGGCCCAGCAGGAAGTGCCCTCCTGGTTAGAGGAGTGTGTGTT
VASA-8       TGTCCGCATCCTGTCCAAGGCCCAGCAG-AAGTGCCCT----------------------
VASA-9       TGTCAGCATCCTGTCCAAGGCCCAGCAGGAAGTGCCCTCCTGGTTAGAGGAGTGTGTGTT


VASA-IT      CAGCGGCTCAGGTGTGAACCCCTCCAGGAGGACCTTTGCGTCCACAGACTCCAGGAAGGG
VASA-4b      CAGCGGCTCAGGTGTGAACCCCTCCAGGAGGACCTTTGCGTCCACAGACTCCAGGAAGGG
VASA-5b      ------------------------------------------------------------
VASA-7       CAGCGGCTCAGGTGTGAACCCCTCCAGGAGGACCTTTGCGTCCACAGACTCCAGGAAGGG
VASA-8       ------------------------------------------------------------
VASA-9       CAGCGGCTCAGGTGTGAACCCCTCCAGGAGGACCTTTGCGTCCACAGACTCCAGGAAGGG


VASA-IT      TCCACAGGGCAGCTCCTTTCAGGACAGCAGTATGACGAG
VASA-4b      TCCACAGGGCAGCTCCTTTCAGGACAGCAGTATGAC---
VASA-5b      ---------------------------------------
VASA-7       TCCACAGGGCAGCTCCTTTCAGGACAGCAGTATGAC---
VASA-8       ---------------------------------------
VASA-9       TCCACAGGGCAGCCCCTTTCAG-ACAGCAGTATGAC---

[*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 23. Clustal Omega 1.2.0 pairwise alignment of *S. lalandi Vasa* nucleotide sequences obtained by cloning (colonies 4b, 5b, 7, 8 and 9), and the corresponding fragment of *Vasa* nucleotide sequence obtained from the RNA-Seq transcriptome library (VASA-IT).

Despite extensive PCR optimisation, 3′ RACE PCR was unsuccessful at amplifying the 3′ end of the *Vasa* nucleotide sequence. After limited success with the standard PCR protocol (Table 6), an alternative PCR program was tested which produced a band of ~550 bp in testes 3′ RACE cDNA (Figure 24). This band was gel purified and transformed, despite being smaller than expected. Colony PCR produced large bands of ~900 bp in several colonies, while others showed small bands of ~350 bp (Figure 25). One small and three large colonies were cultured overnight for plasmid extraction, but only one culture (#7) successfully grew. Restriction digest of the extracted plasmid produced multiple bands. The plasmid was sent for sequencing, but only returned vector sequence.

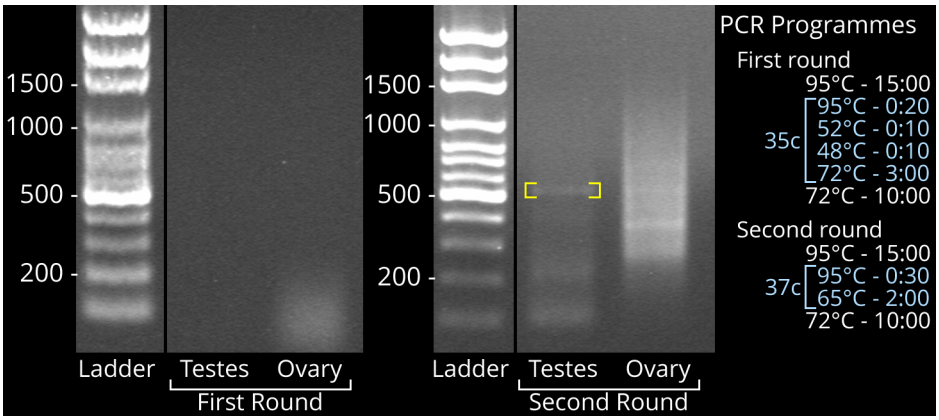


Figure 24. Gel electrophoresis analysis of first and second rounds of 3′ RACE PCR for *Vasa*, using an alternative PCR program (detailed in figure). The second round band in testes (highlighted) was ligated for cloning and sequencing.

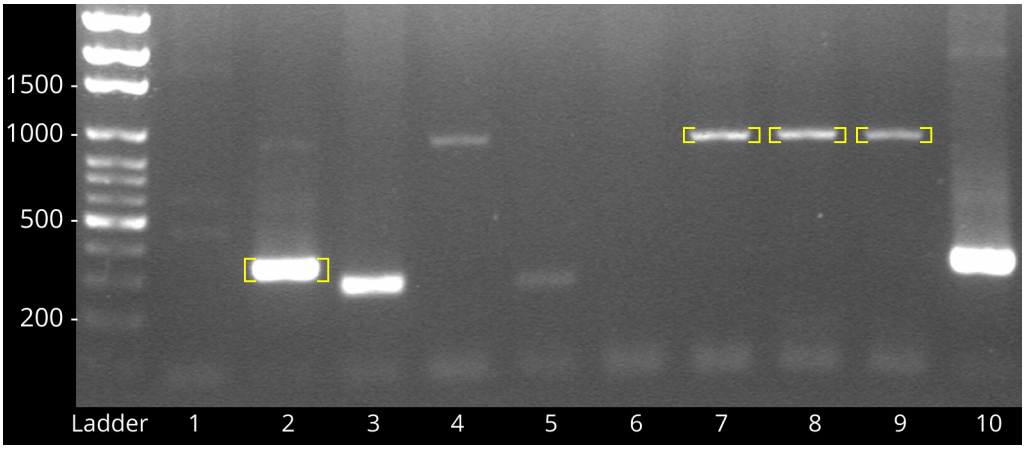


Figure 25. Gel electrophoresis analysis of *Vasa* 3′ RACE products from colony PCR, using M13 forward(-40) and reverse primers. The highlighted colonies were cultured overnight for plasmid extraction, but cultures 2, 8 and 9 failed to grow.

The standard PCR protocol (Table 6) was revisited, using an annealing temperature of 54°C and an extension time of two minutes. Several bands were generated in ovary 3′ RACE cDNA; the largest of which (~700 bp) was gel purified and ligated for transformation. Colony PCR revealed which white colonies contained inserts of expected size. Ten colonies were screened, with nine showing larger than expected inserts of ~950 bp. Six colonies were cultured overnight for plasmid extraction, but only one successfully grew. This plasmid was sent for sequencing but no nucleotide sequence was detected.

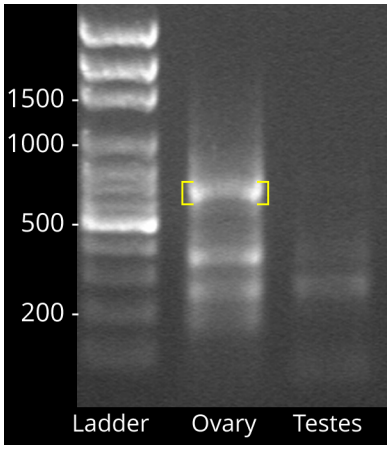


Figure 26. Gel electrophoresis analysis of 3′ RACE PCR for *Vasa*. The ~700 bp band in ovary cDNA was gel purified.

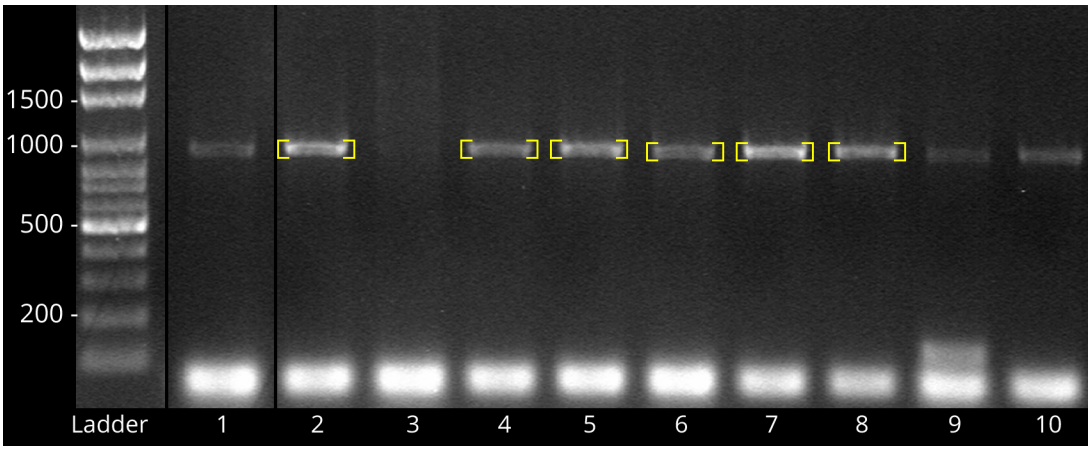


Figure 27. Gel electrophoresis analysis of *Vasa* 3′ RACE products from colony PCR, using M13 forward(-40) and reverse primers. The highlighted colonies were cultured overnight for plasmid extraction.

3.3.3 *Amh*

Combinations of primers specific to *Amh* in *S. lalandi* (slAMH-F1-4, R1-4) were used to amplify sections of ovary and testes cDNA using PCR. Combinations of F1-2 and R1-2 primers were expected to amplify fragments of ~850-890 bp, but despite extensive optimisation, only generated smaller non-specific products and smearing (Figure 28). This led to the F3-4 and R3-4 primers being designed, which targeted

smaller fragments of ~250-390 bp. In testes cDNA, combinations of F3-4 and R3-4 primers produced a cascade of products, all ~200 bp larger than expected (Figure 29). No bands were obtained from ovary cDNA. Optimising the annealing temperature to 52°C and extension time to 45 seconds with the F3+R3 primers produced a strong specific band of ~390 bp (Figure 30). This band was gel purified and ligated for cloning.

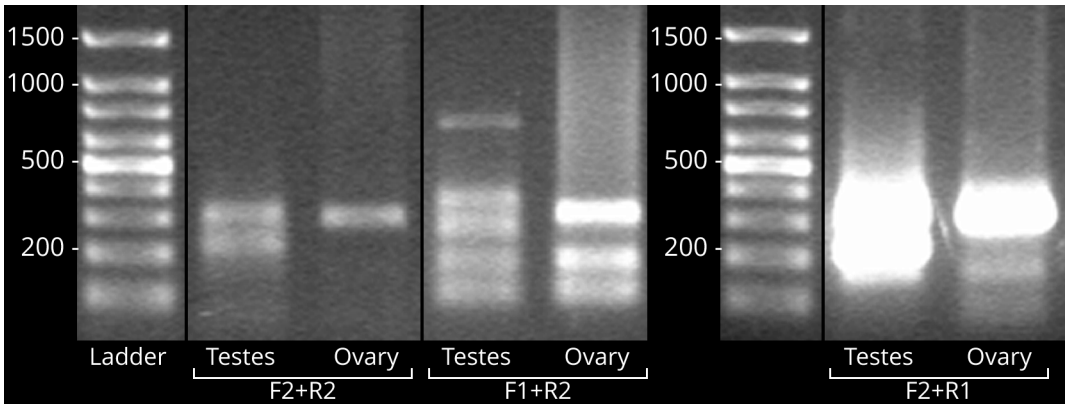


Figure 28. Gel electrophoresis analysis of *Amh* in ovary and testes cDNA, using combinations of slAMH-F1-2 and R1-2 primers. Only non-specific bands were produced.

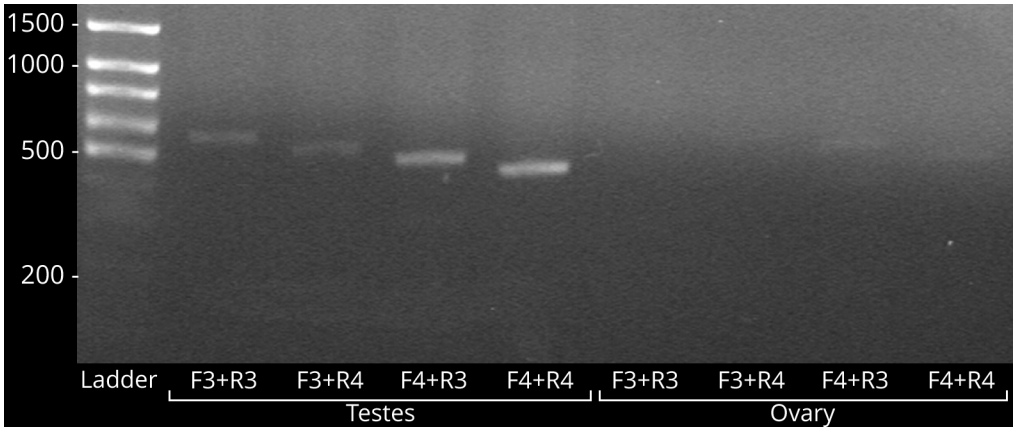


Figure 29. Gel electrophoresis analysis of ovary and testes cDNA, using slAMH-F3-4 and R3-4 primers specific to *S. lalandi*. For testes, each primer combination generated a single band of ~150 bp greater size than expected. No strong bands were produced for ovary cDNA.

Colony PCR revealed which white colonies contained correctly sized inserts. Eight colonies were initially screened; two showed bands of expected size (~560 bp) (Figure 31). A further eight colonies were screened, with another two showing bands of expected size (Figure 32). All four colonies were cultured overnight for plasmid extraction. The concentration and quality of each extracted plasmid was measured on the NanoDrop (Table 16). Restriction digest was conducted on each plasmid, revealing insert bands of expected size (~440 bp) for colonies 6 and 8, but smaller bands of ~280 bp for colonies 15 and 16 (Figure 33). All four plasmids were sent for sequencing.

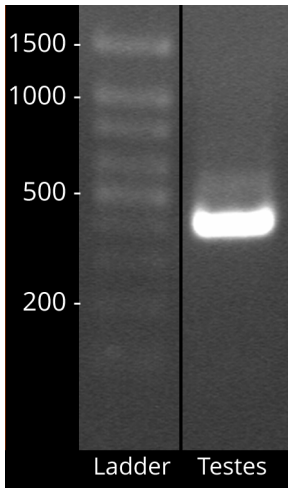


Figure 30. Gel electrophoresis analysis of *Amh* in testes cDNA, using slAMH-F3 and slAMH-R3 primers specific to *S. lalandi*.

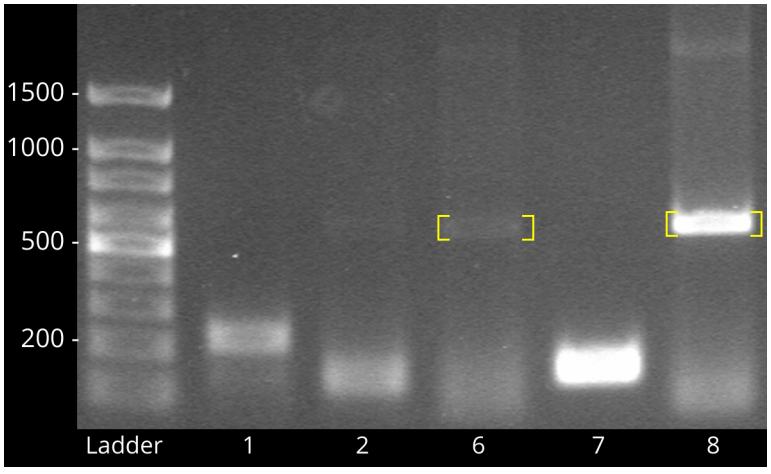


Figure 31. Gel electrophoresis analysis of *Amh* products from colony screening, using M13 forward(-40) and reverse primers. Colonies 6 and 8 produced bands of expected size (~560 bp, highlighted) and were cultured overnight for plasmid extraction.

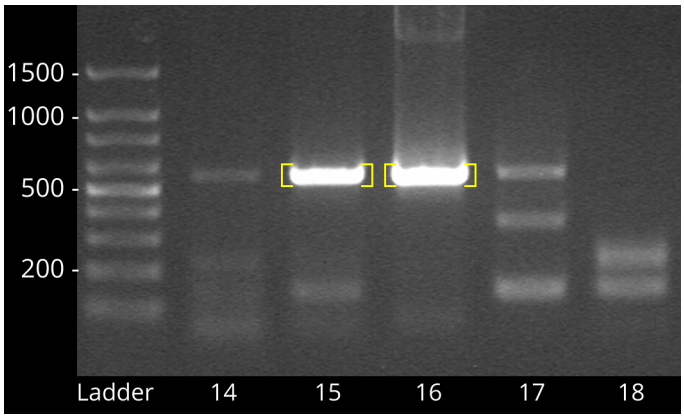


Figure 32. Gel electrophoresis analysis of *Amh* products from colony screening, using M13 forward(-40) and reverse primers. Colonies 15 and 16 produced strong bands of expected size (~560 bp, highlighted) and were cultured overnight for plasmid extraction.

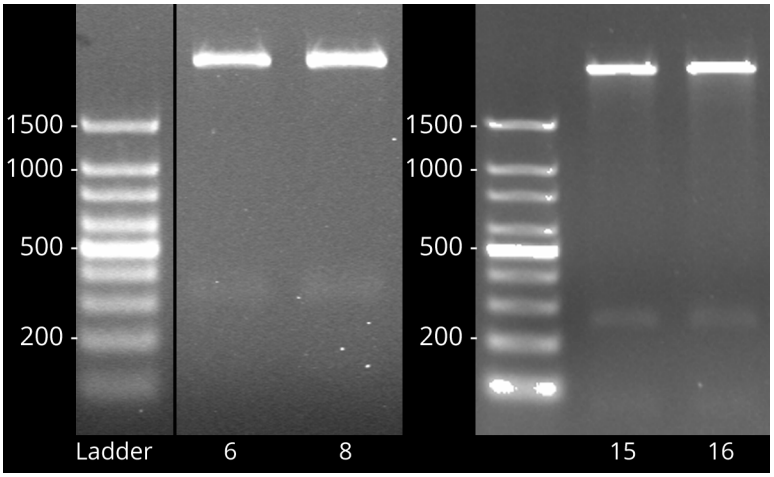


Figure 33. Gel electrophoresis analysis of *Amh* restriction digest. Products 6 and 8 show a correctly sized band of ~440 bp representing the insert, while products 15 and 16 show a smaller band of ~280 bp. The bands at >1500 bp represent the remaining plasmid DNA.

Table 16. Nanodrop 2000 analysis of *Amh* plasmid DNA concentration and quality.

Plasmid DNA	DNA concentration (ng/µl)	260/280	260/230
Amh F3+R3 Colony 1	233.9	1.84	2.07
Amh F3+R3 Colony 8	153.7	1.85	2.13
Amh F3+R3 Colony 15	314.7	1.90	2.01
Amh F3+R3 Colony 16	150.2	1.95	2.04
Amh 3′F2+3′ inner Colony 2	112.4	1.86	2.01
Amh 3′F2+3′ inner Colony 3	77.6	1.84	1.62
Amh 3′F2+3′ inner Colony 4	128.5	1.88	2.06
Amh 3′F2+3′ inner Colony 5	178.4	1.77	1.23
Amh 3′F2+3′ inner Colony 6	147.4	1.89	2.14
Amh 3′F2+3′ inner Colony 7	108.4	1.88	2.19

The similarity of these sequences was confirmed by multiple alignments, alongside the *Amh* nucleotide sequence constructed from the Ion Torrent transcriptome data (Figure 34).

```
AMH-IT      CCCCTGTCACTGGGCTTATCCTCCAGTGAGACCCTGCTGGCAGGACTGATCAACTCCTCT
AMH-1       ---CTGTCACTGGGCTTATCCTCCAGTGAGACCCTGCTGGCAGGACTGATCAACTCCTCT
AMH-8       ---CTGTC-CTGGGCTTATCCTCCAGTGAGACCCTGCTGGCAGGACTGATCAACTCCTCT
AMH-15      ---CTGTC-CTGGGCTTATCCTCCAGTGAGACCCTGCTGGCAGGACTGATCAACTCCTCT
AMH-16      ---CTGTCACTGGGCTTATCCTCCAGTGAGACCCTGCTGGCAGGACTGATCAACTCCTCT
               ***** **************************************************

AMH-IT      TTCCTTACTGTCTTCTCCTTCGCTAGATGGGGCTCCATTTCTCAGGTGCATCCTGGACAG
AMH-1       TTCCTTACTGTCTTCTCCTTCGCTAGATGGGGCTCCATTTCTCAGGTGCATCCTGGACAG
AMH-8       TTCCTTACTGTCTTCTCCTTCGCTAGATGGGGCTCCATTTCTCAGGTGCATCCTGGACAG
AMH-15      TTCCTTACTGTCTTCTCCTTCGCTAGATGGGGCTCCATTTCTCAGGTGCATCCTGGACAG
AMH-16      TTCCTTACTGTCTTCTCCTTCGCTAGATGGGGCTCCATTTCTCAGGTGCATCCTGGACAG
            ************************************************************

AMH-IT      CTGGCCATGTCTCCTGCACTGGTGGAGGAGGTCAGGCAGAGATTGGCAGATNTGNGGATG
AMH-1       CTGGCCATGTCTCCTGCACTGGTGGAGGAGGTCAGGCAGAGATTGGGGCAGATTGGGATG
AMH-8       CTGGCCATGTCTCCTGCACTGGTGGAGGAGGTCAGGCAGAGATTGGGGCAGATTGGGATG
AMH-15      CTGGCCATGTCTCCTGCACTGGTGGAGGAGGTCAGGCAGAGATTGGGGCAGATTGGGATG
AMH-16      CTGGCCATGTCTCCTGCACTGGTGGAGGAGGTCAGGCAGAGATTGGGGCAGATTGGGATG
            ********************************************* . *  *  *****

AMH-IT      CAGATAATGGAAGTAATAAGGGAGAAGGAGGTGGGTCACAGGGCCACGGAGAGGCTGGAG
AMH-1       CAGATAATGGAAGTAATAAGGGAGAAGGAGGTGGGTCACAGGGCCACGGAGAGGCTGGAG
AMH-8       CAGATAATGGAAGTAATAAGGGAGAAGGAGGTGGGTCACAGGGCCACGGAGAGGCTGGAG
AMH-15      CAGATAATGGAAGTAATAAGGGAGAAGGAGGTGGGTCACAGGGCCACGGAGAGGCTGGAG
AMH-16      CAGATAATGGAAGTAATAAGGGAGAAGGAGGTGGGTCACAGGGCCACGGAGAGGCTGGAG
            ************************************************************
```

```
AMH-IT      AGGCTTCAAGAATTCAGTGCATTACCGACGATGGAACAAGCAGCAGGACGGAGCCAGTAC
AMH-1       AGGCTTCAAGAATTCAGTGCATTACCGACGATGGAACAAGCAGCAGGACGGAGCCAGTAC
AMH-8       AGGCTTCAAGAATTCAGTGCATTACCGACGATGGAACAAGCAGCAGGACGGAGCCAGTAC
AMH-15      AGGCTTCAAGAATTCAGTGCATTACCGACGATGGAACAAGCAGCAGGACGGAGCCAGTAC
AMH-16      AGGCTTCAAGAATTCAGTGCATTACCGACGATGGAACAAGCAGCAGGACGGAGCCAGTAC
            ************************************************************

AMH-IT      TGTGCGTTTCTTCTGCTGAAGGCCCTGCAGACAGTGGCCCGTGAGTACGAGCTGCAGAGA
AMH-1       TGTGCGTTTCTTCTGCTGAAGGCCCTGCAGACAGTGGCCCGTGAGTACGAGCTGCAGAGA
AMH-8       TGTGCGTTTCTTCTGCTGAAGGCCCTGCAGACAGTGGCCCGTGAGTACGAGCTGCAGAGA
AMH-15      TGTGCGTTTCTTCTGCTGAAGGCCCTGCAGACAGTGGCCCGTGAGTACGAGCTGCAGAGA
AMH-16      TGTGCGTTTCTTCTGCTGAAGGCCCTGCAGACAGTGGCCCGTGAGTACGAGCTGCAGAGA
            ************************************************************

AMH-IT      GGACTGCGAGCCACCAGAGCGGACACCAATAACCCGG
AMH-1       GGACTGCGAGCCACCAGAGCGGACACCAATAACC---
AMH-8       GGACTGCGAGCCACCAGAGCGGACACCAATAACC---
AMH-15      GGACTGCGAGCCACCAGAGCGGACACCAATAACC---
AMH-16      GGACTGCGAGCCACCAGAGCGGACACCAATAACC---
            **********************************

[*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 34. Clustal Omega 1.2.0 pairwise alignment of *S. lalandi Amh* nucleotide sequences obtained by cloning (colonies 1, 8, 15 and 16), and the corresponding fragment of *Amh* nucleotide sequence obtained from the RNA-Seq transcriptome library (AMH-IT).

3′ RACE PCR was successful at amplifying the 3′ end of the *Amh* nucleotide sequence, although substantial optimisation was required. Several initial attempts at nested PCR using the standard protocol (Table 6) generated a number of non-specific bands. An alternative PCR program was tested which appeared to increase product specificity (Figure 35). However, cloning and sequencing revealed these bands did not match the *Amh* sequence from the Ion Torrent transcriptome data. Using the standard protocol with 38 cycles, an extension time of 2 minutes, and annealing temperatures of 54°C (first round) and 60°C (second round), a band of ~680 bp was obtained from 3′ RACE testes cDNA (Figure 36). Reducing the second round annealing temperature to 58°C amplified only a smaller band of ~330 bp in 3′ RACE ovary cDNA (Figure 37). Both bands were ligated for cloning.

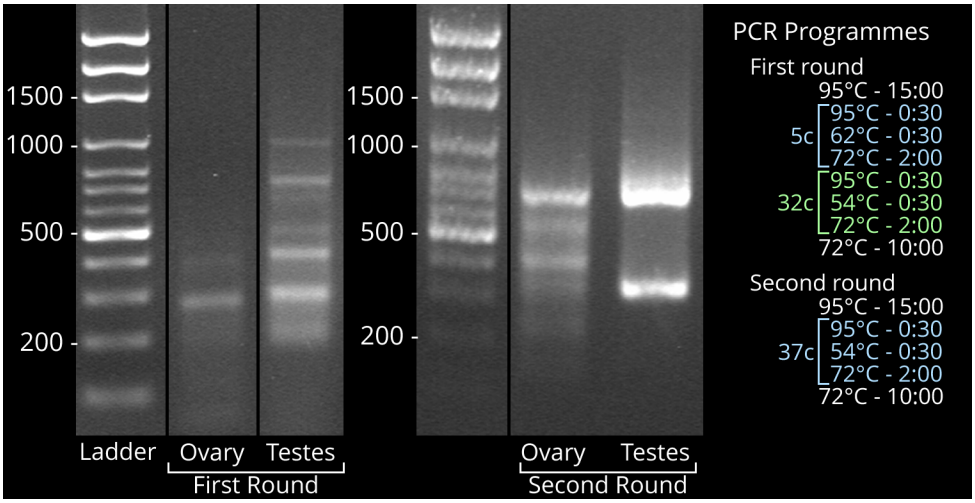


Figure 35. Gel electrophoresis analysis of first and second rounds of 3′ RACE PCR for *Amh*, using an alternative PCR program (detailed in figure). The second round bands highlighted were ligated, cloned and sequenced, but did not match the *Amh* sequence from the Ion Torrent transcriptome data.

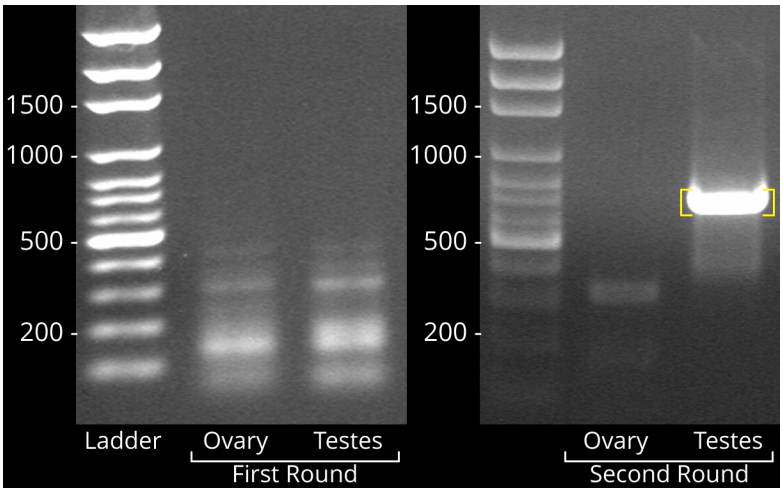


Figure 36. Gel electrophoresis analysis of first and second rounds of 3′ RACE PCR for *Amh*. A ~680 bp band (highlighted) was generated in testes 3′ RACE cDNA using a second round annealing temperature of 60°C.

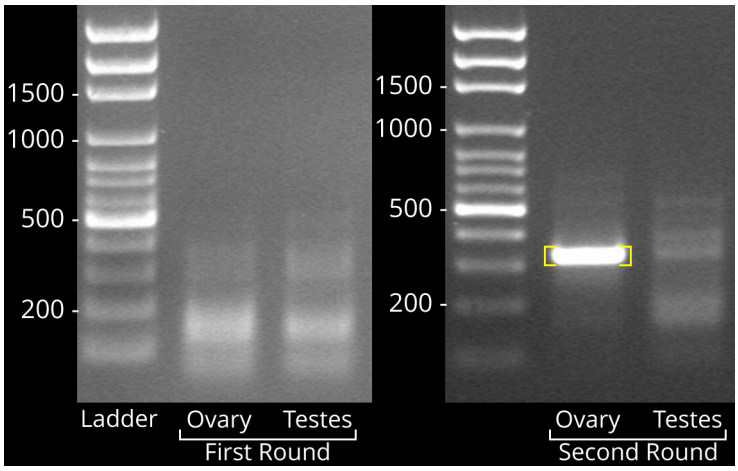


Figure 37. Gel electrophoresis analysis of first (A) and second (B) rounds of 3′ RACE PCR for *Amh*. A ~330 bp band (highlighted) was generated in ovary 3′ RACE cDNA using a second round annealing temperature of 58°C.

Colony PCR revealed which white colonies contained inserts of expected size. Ten colonies were screened for each transformed product. For testes colonies, nine colonies showed inserts of expected length (~850 bp) (Figure 38). Colonies 2-7 were cultured overnight for plasmid extraction. The concentration and quality of each plasmid was measured on the NanoDrop (Table 15). Restriction digest revealed insert bands of ~730 bp (Figure 39). All six plasmids were sent for sequencing. For ovary colonies, only two colonies showed inserts of expected length (~500 bp) (Figure 40). These were not pursued for plasmid extraction due to their small size.

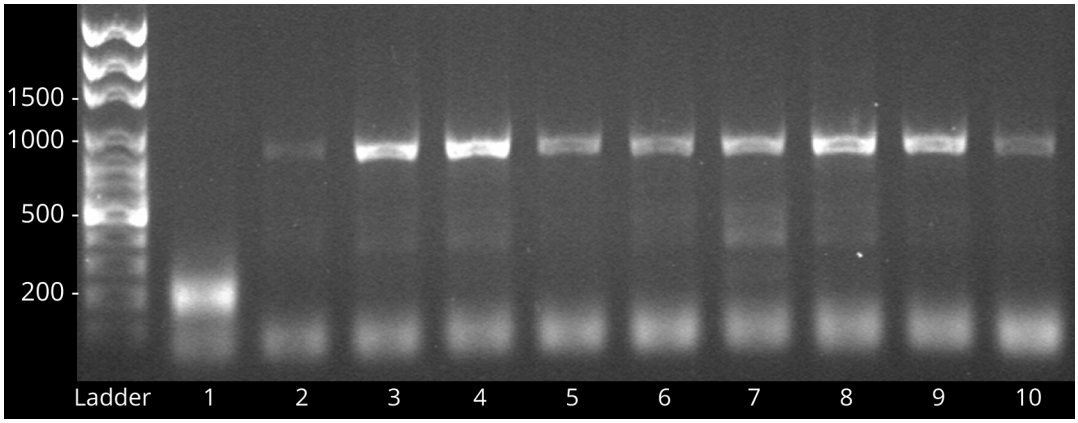


Figure 38. Gel electrophoresis analysis of *Amh* testes 3′ RACE products from colony screening, using M13 forward(-40) and reverse primers. Colonies 2-10 produced bands of expected size. Colonies 2-7 were cultured overnight for plasmid extraction.

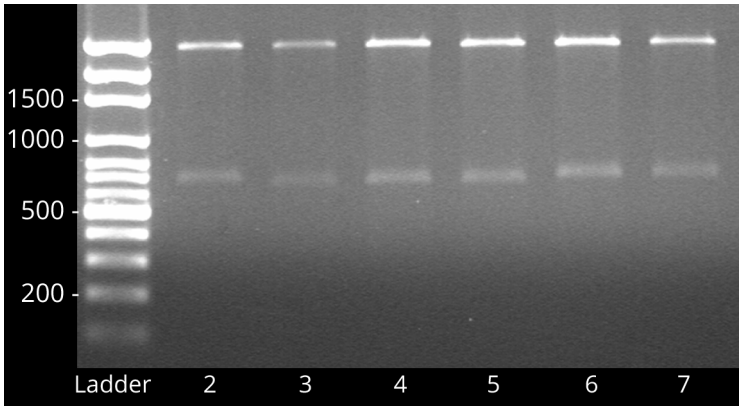


Figure 39. Gel electrophoresis analysis of *Amh* testes 3′ RACE restriction digest. Bands of ~730 bp represent the insert, while the larger bands at >1500 bp represent the remaining plasmid.

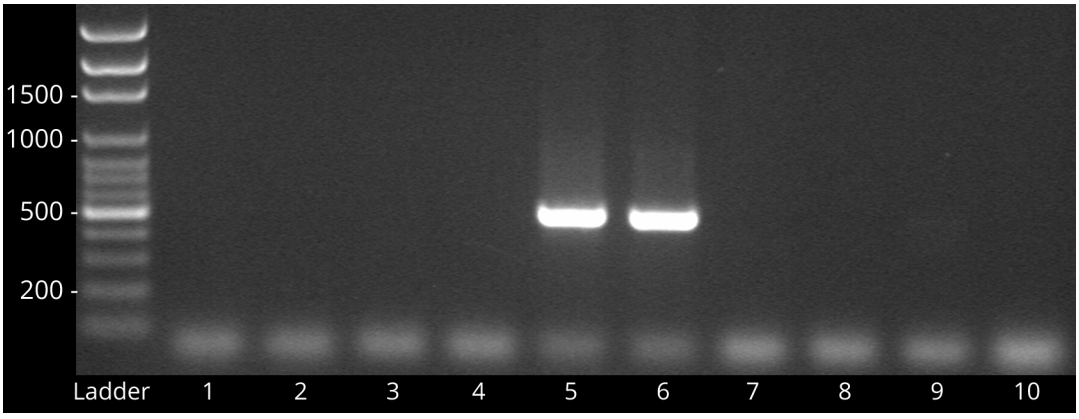


Figure 40. Gel electrophoresis analysis of *Amh* ovary 3′ RACE products from colony screening, using M13 forward(-40) and reverse primers. Colonies 5 and 6 produced bands of expected, but smaller than desired, size, and were not pursued for plasmid extraction.

The similarity of these sequences was confirmed by multiple alignments, alongside the *Amh* nucleotide sequence constructed from the Ion Torrent transcriptome data (Figure 41).

```
AMH-IT      CGTGAGTACGAGCTGCAGAGAGGACTGCGAGCCACCAGAGCGGACACCAATAACCCGGTG
AMH-3'2     ---GAGTACGAGCTGCAGAGAGGACTGCGAGCCACCAGAGCGGACACCAATAACCCGGTG
AMH-3'3     ---GAGTACGAGCTGCAGAGAGGACTGCGAGCCACCAGAGCGGACACCAATAACCCGGTG
AMH-3'4     ---GAGTACGAGCTGCAGAGAGGACTGCGAGCCACCAGAGCGGACACCAATAACCCGGTG
AMH-3'5     ---GAGTACGAGCTGCAGAGAGGACTGCGAGCCACCAGAGCGGACACCAATAACCCGGTG
AMH-3'6     ---GAGTACGAGCTGCAGAGAGGACTGCGAGCCACCAGAGCGGACACCAATAACCCGGTG
AMH-3'7     ---GAGTACGAGCTGCAGAGAGGACTGCGAGCCACCAGAGCGGACACCAATAACCCGGTG
               ********************************************************

AMH-IT      AGGGGCAGTGTCTGCAGTGTGAGAAGCCTCACCGTCAACCTTGAAAGGCGTCTTGTGGGT
AMH-3'2     AGGGGCAGTGTCTGCAGTGTGAGAAGCCTCACCGTCAACCTTGAAAGGCGTCTTGTGGGT
AMH-3'3     AGGGGCAGTGTCTGCAGTGTGAGAAGCCTCACCGTCAACCTTGAAAGGCGTCTTGTGGGT
AMH-3'4     AGGGGCAGTGTCTGCAGTGTGAGAAGCCTCGCCGTCAACCTTGAAAGGCGTCTTGTGGGT
AMH-3'5     AGGGGCAGTGTCTGCAGTGTGAGAAGCCTCACCGTCAACCTTGAAAGGCGTCTTGTGGGT
AMH-3'6     AGGGGCAGTGTCTGCAGTGTGAGAAGCCTCACCGTCAACCTTGAAAGGCGTCTTGTGGGT
AMH-3'7     AGGGGCAGTGTCTGCAGTGTGAGAAGCCTCACCGTCAACCTTGAAAGGCGTCTTGTGGGT
            ******************************.*****************************

AMH-IT      CCAAGCGCTGCCAACATCAAAAACTGCCATGGCTCTTGTGCTTTCCCCCTGACCAATGCA
AMH-3'2     CCAAGCGCTGCCAACATCAAAAACTGCCATGGCTCTTGTGCTTTCCCCCTGACCAATGCA
AMH-3'3     CCAAGCGCTGCCAACATCAAAAACTGCCATGGCCCTTGTGCTTTCCCCCTGACCAATGCA
AMH-3'4     CCAAGCGCTGCCAACATCAAAAACTGCCATGGCTCTTGTGCTTTCCCCCTGACCAATGCA
AMH-3'5     CCAAGCGCTGCCAACATCAAAAACTGCCATGGCTCTTGTGCTTTCCCCCTGACCAATGCA
AMH-3'6     CCAAGCGCTGCCAACATCAAAAACTGCCATGGCTCTTGTGCTTTCCCCCTGACCAATGCA
AMH-3'7     CCAAGCGCTGCCAACATCAAAAACTGCCATGGCTCTTGTGCTTTCCCCCTGACCAATGCA
            ******************************** ************************

AMH-IT      AACGACCATGTTGTCCTGCTCCACTTCCACATTGAGAGCGGGAATATGGATGAGCGGGCT
AMH-3'2     AACGACCATGTTGTCCTGCTCCACTTCCACATTGAGAGCGGGAATATGGATGAGCGGGCT
AMH-3'3     AACGACCATGTTGTCCTGCTCCACTTCCACATTGAGAGCGGGAATATGGATGAGCGGGCT
AMH-3'4     AACGACCATGTTGTCCTGCTCCACTTCCACATTGAGAGCGGGAATATGGATGAGCGGGCT
AMH-3'5     AACGACCATGTTGTCCTGCTCCACTTCCACATTGAGAGCGGGAATATGGATGAGCGGGCT
AMH-3'6     AACGACCATGTTGTCCTGCTCCACTTCCACATTGAGAGCGGGAATATGGATGAGCGGGCT
AMH-3'7     AACGACCATGTTGTCCTGCTCCACTTCCACATTGAGAGCGGGAATATGGATGAGCGGGCT
            ************************************************************

AMH-IT      CCGTGCTGTGTGCCTGTTGCCTATGAACCCCTTGAGGTGGTGGGTTTGAATGAACATGGG
AMH-3'2     CCGTGCTGTGTGCCTGTTGCCTATGAACCCCTTGAGGTGGTGGGTTTGAATGAACATGGG
AMH-3'3     CCGTGCTGTGTGCCTGTTGCCTATGAACCCCTTGAGGTGGTGGGTTTGAATGAACATGGG
AMH-3'4     CCGTGCTGTGTGCCTGTTGCCTATGAACCCCTTGAGGTGGTGGGTTTGAATGAACATGGG
AMH-3'5     CCGTGCTGTGTGCCTGTTGCCTATGAACCCCTTGAGGTGGTGGGTTTGAATGAACATGGG
AMH-3'6     CCGTGCTGTGTGCCTGTTGCCTATGAACCCCTTGAGGTGGTGGGTTTGAATGAACATGGG
AMH-3'7     CCGTGCTGTGTGCCTGTTGCCTATGAACCCCTTGAGGTGGTGGGTTTGAATGAACATGGG
            ************************************************************

AMH-IT      ACTTACCTCTCCATGATACCAGATATGGTAGCAAAGGAGTGTGGATGCCGC---------
AMH-3'2     ACTTACCTCTCCATGATACCAGATATGGTAGCAAAGGAGTGTGGATGCCGCTGAAGCTGT
AMH-3'3     ACTTACCTCTCCATGATACCAGATATGGTAGCAAAGGAGTGTGGATGCCGCTGAAGCTGT
AMH-3'4     ACTTACCTCTCCATGATACCAGATATGGTAGCAAAGGAGTGTGGATGCCGCTGAAGCTGT
AMH-3'5     ACTTACCTCTCCATGATACCAGATATGGTAGCAAAGGAGTGTGGATGCCGCTGAAGCTGT
AMH-3'6     ACTTACCTCTCCATGATACCAGATATGGTAGCAAAGGAGTGTGGATGCCGCTGAAGCTGT
AMH-3'7     ACTTACCTCTCCATGATACCAGATATGGTAGCAAAGGAGTGTGGATGCCGCTGAAGCTGT
            ************************************************

AMH-IT      ------------------------------------------------------------
AMH-3'2     TTTCTTTTTGGAGTGTGTTTGTGGCATGAGAGATATTTCACTTGTTTGGTTTGATAGTCA
AMH-3'3     TTTCTTTTTGGAGTGTGTTTGTGACATGAGAGATATTTCACTTGTTTGGTTTGATAGTCA
AMH-3'4     TTTCTTTTTGGAGTGTGTTTGTGACATGAGAGATATTTCACTTGTTTGGTTTGATAGTCA
AMH-3'5     TTTCTTTTTGGAGTGTGTTTGTGACATGAGAGATATTTCACTTGTTTGGTTTGATAGTCA
AMH-3'6     TTTCTTTTTGGAGCGTGTTTGTGACATGAGAGATATTTCACTTGTTTGGTTTGATAGTCA
AMH-3'7     TTTCTTTTTGGAGTGTGTTTGTGACATGAGAGATATTTCACTTGTTTGGTTTGATAGTCA
```

```
AMH-IT      ------------------------------------------------------------
AMH-3'2     CTGACTGTAGCTTGTTAGGTGGAGGAAAAGTGTTATTTAATGTTTTTTTTTATTCACCTAT
AMH-3'3     CTGACTGTAGCTTGTTAGGTGGAGGAAAAGTGTTATTTAATGTTTTTTTTTATTCACCTAT
AMH-3'4     CTGACTGTAGCTTGTTAGGTGGAGGAAAAGTGTTATTTAATGTTTTTTTTTATTCACCTAT
AMH-3'5     CTGACTGTAGCTTGTTAGGTGGAGGAAAAGTGTTATTTAATGTTTTTTTTTATTCACCTAT
AMH-3'6     CTGACTGTAGCTTGTTAGGTGGAGGAAAAGTGTTATTTAATGTTTTTTTTTATTCACCTAT
AMH-3'7     CTGACTGTAGCTTGTTAGGTGGAGGAAAAGTGTTATTTAATGTTTTTTTTTATTCACCTAT

AMH-IT      ------------------------------------------------------------
AMH-3'2     TCACATATGTTAATTCAGAGAGACTTTTT-TTATAAACCTACCTAGATGCACACAGCAAA
AMH-3'3     TCACATATGTTAATTCAGAGAGACTTTTT-TTATAAACCTACCTAGATGCACACAGCAAA
AMH-3'4     TCACATATGTTAATTCAGAGAGACTTTTT-TTATAAACCTACCTAGATGCACACAGCAAA
AMH-3'5     TCACATATGTTAATTCAGAGAGACTTTTT-TTATAAACCTACCTAGATGCACACAGCAAA
AMH-3'6     TCACATATGTTAATTCAGAGAGACTTTTT-TTATAAACCTACCTAGATGCACACAGCAAA
AMH-3'7     TCACATATGTTAATTCAGAGAGACTTTTTTTTTATAAACCTACCTAGATGCACACAGCAAA

AMH-IT      ------------------------------------------------------------
AMH-3'2     ATGTTTCACACTTAGAAATAATCAAGTTTATTGAAGACAAAAATATATGTGTTAACAGCT
AMH-3'3     ATGTTTCACACTTAGAAATAATCAAGTTTATTGAGGACAAAAATATATGTGTTAACAGCT
AMH-3'4     ATGTTTCACACTTAGAAATAATCAAGTTTATTGAAGACAAAAATATATGTGTTAACAGCT
AMH-3'5     ATGTTTCACACTTAGAAATAATCAAGTTTATTGAAGACAAAAATATATGTGTTAACAGCT
AMH-3'6     ATGTTTCACACTTAGAAATAATCAAGTTTATTGAAGACAAAAATATATGTGTTAACAGCT
AMH-3'7     ATGTTTCACACTTAGAAATAATCAAGTTTATTGAAGACAAAAATATATGTGTTAACAGCT

AMH-IT      ------------------------------------------------------------
AMH-3'2     ATTTATTATTTTAAACTTCTTCTATAACACCATATTTTGTTTTATAATGAAGATTCCCCG
AMH-3'3     ATTTATTATTTTAAACTTCTTCTATAACACCATATTTTGCTTTATAATGAAGATTCCCCG
AMH-3'4     ATTTATTATTTTAAACTTCTTCTATAACACCATATTTTGTTTTATAATGAAGATTCCCCG
AMH-3'5     ATTTATTATTTTAAACTTCTTCTATAACACCATATTTTGTTTTATAATGAAGATTCCCCG
AMH-3'6     ATTTATTATTTTAAACTTCTTCTATAACACCATATTTTGTTTTATAATGAAGATTCCCCG
AMH-3'7     ATTTATTATTTTAAACTTCTTCTATAACACCATATTTTGTTTTATAATGAAGATTCCCCG

AMH-IT      --------------------------------
AMH-3'2     TATTAAAAATATTGTTCTAAAAAAAAAA---
AMH-3'3     TATTAAAAATATTGTTCTAAAAAAAAA----
AMH-3'4     TATTAAAAATATTGTTCTAAAAAAAAA----
AMH-3'5     TATTAAAAATATTGTTCTAAAAAAAAAA---
AMH-3'6     TATTAAAAATATTGTTCTAAAAAAAAAAAAA
AMH-3'7     TATTAAAAATATTGTTCTAAAAAAAAAAA-

[*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 41. Clustal Omega 1.2.0 pairwise alignment of *S. lalandi Amh* nucleotide sequences obtained by cloning (3′ RACE colonies 2-7), and the corresponding fragment of *Amh* nucleotide sequence obtained from the RNA-Seq transcriptome library (AMH-IT).

3.3.4 *Cyp19a1a*

Despite extensive PCR optimisation, 3′ RACE PCR was unsuccessful at amplifying the 3′ end of the *Cyp19a1a* nucleotide sequence. After limited success with the standard PCR protocol (Table 6), an alternative PCR program was tested. This produced numerous non-specific bands, including strong two bands of ~620 bp and ~420 bp (labelled A and B respectively) in testes 3′ RACE cDNA (Figure 42). These bands were gel purified and transformed. Colony PCR was used to screen ten colonies; seven for the larger product and three for the smaller product. Two colonies generated bands of relative size for each product, but the absolute size could not be determined due to a ladder being omitted from the gel (Figure 43). The absence of a ladder could not be rectified due to time constraints. All four colonies were cultured overnight for plasmid extraction. Restriction digest of the extracted plasmids produced only small, faint bands of ~250 bp (Figure 44). All four were sent for sequencing.

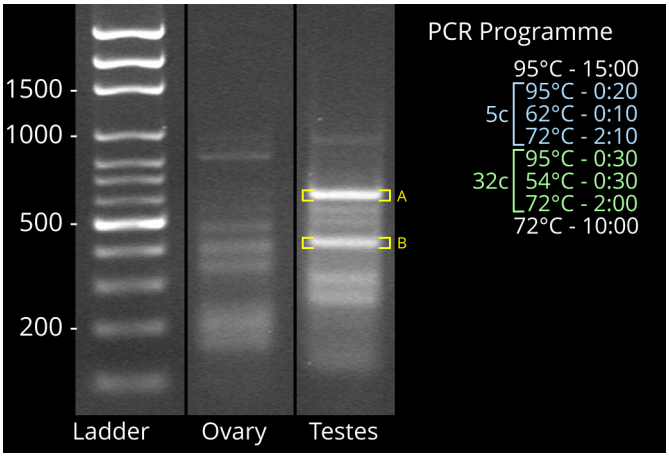


Figure 42. Gel electrophoresis analysis of first round 3′ RACE PCR for *Cyp19a1a*, using an alternative PCR program (detailed in figure). The two bands highlighted were gel purified and ligated for transformation.

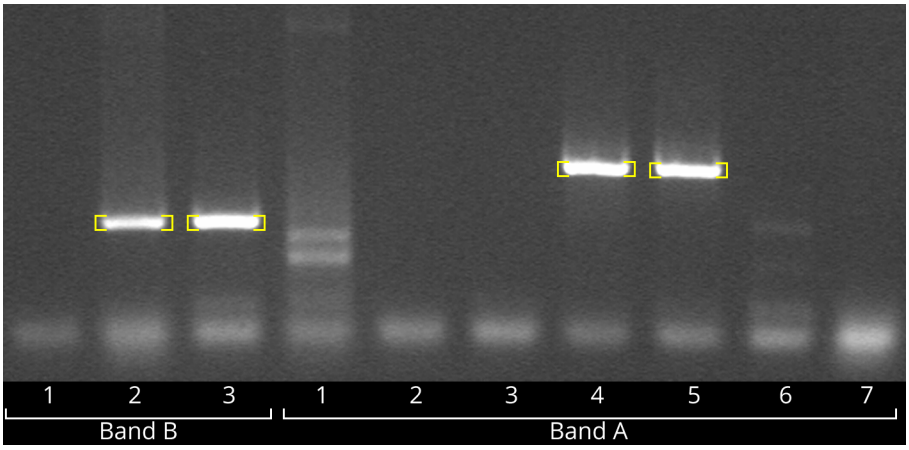


Figure 43. Gel electrophoresis analysis of *Cyp19a1a* 3′ RACE products from colony PCR screening, using M13 forward(-40) and reverse primers. Colonies B2, B3, A4 and A5 (highlighted) were cultured overnight for plasmid extraction. The absence of a ladder could not be rectified due to time constraints.

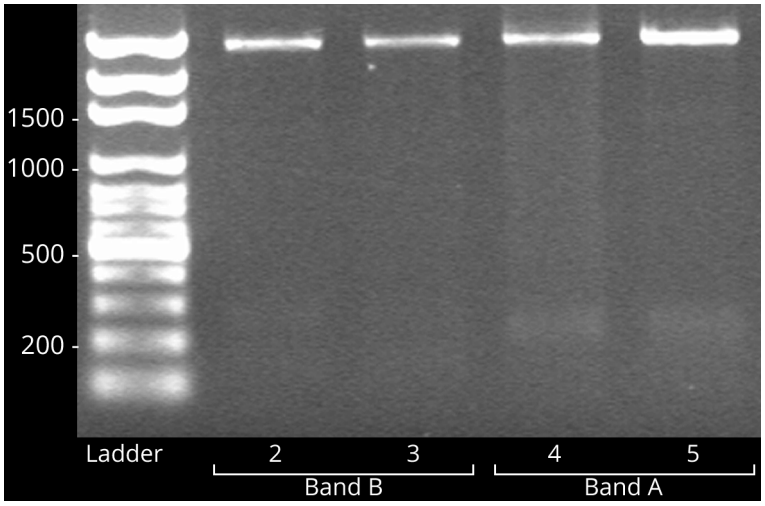


Figure 44. Gel electrophoresis analysis of *Cyp19a1a* 3′ RACE restriction digest. Small, faint bands of ~250 bp were detected.

Table 17. Nanodrop 2000 analysis of *Cyp19* plasmid DNA concentration and quality.

Plasmid DNA	DNA concentration (ng/µl)	260/280	260/230
Cyp 3'F1 (A) Colony 4	110.5	1.88	1.53
Cyp 3'F1 (A) Colony 5	237.2	1.86	2.09
Cyp 3'F1 (B) Colony 2	355.2	1.87	2.36
Cyp 3'F1 (B) Colony 3	213.2	1.88	2.56

The sequencing results from 3' RACE PCR did not generate any homologous sequences for *Cyp19a1a*, but the sequences for colonies 4 and 5 showed homology with zona pellucida (ZP) sperm binding protein. The similarity of these sequences was confirmed by multiple alignments, alongside the nucleotide sequence for *ZP* in *D. labrax* obtained from the UniProt database (K4RH98) (Figure 45). Because *ZP* was not a gene of interest to this study, further clones were not obtained.

```
D.labrax_ZP      CAATCCTAACAGTCTGCCACAGTGGGATATTCTGGTTGATGGGTGTTCCTACCATGATGA
CYP-A4           ---GTTCGGGGGCGGATCAAGCTTAGAGTCTGAGTCACGTGGATGTCCCTACCACGATGA
CYP-A5           ------GCCAGGGCGCATCAGCTTAGAGTCT-GAGTCCGTGGATGTCCCTACCACGATGA
                           *          *  **   *        *** *** ******* *****

D.labrax_ZP      CCGTTACTTGACCACA--GTGGTGCCTGTGGATGGATCGTCTGGACTCCATTACCCAACT
CYP-A4           CCGTTACACAACCACAG--TAGTGCCTGTGGATGGCTCCTCTGGGCTTCAGTATCCATCA
CYP-A5           CCGTTACACAACCACTTCTAGTGTTTTGTGGATGGCTCCTCTGGGCTTCAGTATCCATCA
                 *******   *****           ********* ** ***** ** ** ** *** *

D.labrax_ZP      CACTACAAGCGTTTTGTCATTAAAATGTTTACATTTGTGGATCAAAACACCTTCACTCCT
CYP-A4           CACCACAAACGTTTCATCATCAAAATGTTCACATTTGTGGATCAAAACACCTTTACTACT
CYP-A5           CACCACAAACGTTTCATCATCAAAATGTTCACATTTGTGGATCAAAACACCTTTACTACT
                 *** **** *****  **** ******** ********************** *** **

D.labrax_ZP      CAGAAAGAGACGGTGTTCATCCACTGTGCTACAGCAGTGTGTTATCCCAGCAGCACAAAC
CYP-A4           CAACAGGACACGGTGTTCATCCACTGTGCTACTGCAGTGTGTTACCCCAGCAGCACAAAC
CYP-A5           CAACAGGACACGGTGTTCATCCACTGTGCTACTGCAGTGTGTTACCCCACCAGCACAAAC
                 **  * ** ********************** ********** **** **********

D.labrax_ZP      TCTTGTCAACAGCAATGCCACAGGCAACGAAGAGCTGTAGCTGCAGTGAGGAAAGTTTCC
CYP-A4           TCTTGTGCACAGCCATGCCACCGGCAACGGAGAGCAGTGGCTGCAGTGAGGAAAGCTTCC
CYP-A5           TCTTGTGCACAGCCATGCCACCGGCAACGGAGAGCAGTGGCTGCAGTGAGGAAAGCTTCC
                 ******  ***** ******* ******* ***** ** ********************

D.labrax_ZP      TCAAGTCAGAGGGCTTTAGTCTCAAGCGGTGAGGTGATCCTGACCGAGCCGAGGGCTCCA
CYP-A4           TCAAACCAGAGGGCTGTGGTCTCAAGTGGAGAGGTGATCCTGACAGAGGAGAGGACTACG
CYP-A5           TCAAACCAGAGGGCTGTGGTCTCAAGTGGAGAGGTGATCCTGACAGAGGAGAGGACTACG
                 ****  ********* * ******** ** ************** ***  **** ** *

D.labrax_ZP      TCTGCTTCCAACACCAAATTGAAGAGATGACACTCGACTGGCAGCCATATCTGGCCGTGA
CYP-A4           TCTGCTCCCGACACCAAATCGAACAGATGACTGTCCATATCTGGCCATGC----------
CYP-A5           TCTGCTCCCGACACCAAATCGAACAGATGACTGTCCATATCTGGGCATGC----------
                 ****** ** ********* *** *******  ** *      * ***
```

```
D.labrax_ZP      CTTTTAATAAAAGCTC-ATGCTCACAAAAAAAAAAAAAAAAAAAAA--------------
CYP-A4           ATTTTAATAAAAAGCTTGTACTAAAAAAAAAAAACCTATAGTGAGTCGTATTAATTCGGAT
CYP-A5           ATTTTAATAAAAAGCTTGTACTAAAAAAAAAAAACCTATAGTGAGTCGTATTAATTCGGAT
                  ***********      * ** * ********   * *   *

D.labrax_ZP      -----
CYP-A4           CCGCG
CYP-A5           CCGCG

[*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 45. Clustal Omega 1.2.0 pairwise alignment of *S. lalandi* nucleotide sequences obtained by cloning the products of 3′ RACE PCR for *Cyp19a1a,* and a corresponding fragment of *ZP* nucleotide sequence in *D. Labrax* obtained from the UniProt database.

3.4 Sequencing of Plasmid Inserts

3.4.1 *Vasa*

Using the nucleotide sequences in Figure 19 and Figure 23, a fragment of *Vasa* coding DNA sequence was determined. As the 3′ and 5′ ends of the *Vasa* transcript were not sequenced, sequence from the Ion Torrent transcriptome data was included (Figure 46). This sequence was translated using the ExPASy Translate Tool to construct a sequence of 551 amino acids (Figure 47).

```
GGAGGAGATGATGGAGTGTGTGAAAATGGGTTTAGAGGAGGAAGCCGAGGAGGACGAGGCAGCAGGGG
AGGACGAGGCTTCAGACAAGGTGGAGACCAGGGAGGCAGAGGAGGCTTTGGAGGAGGTTACAGAGGAA
AAGATGAACAGATCTTTGCTCGAGGAGAAGATAAAGATCCAGAGAAGAAGGATGGGAACGATGGTGAC
AGACCAAAGGTCACCTACGTCCCTCCAACCCTCCCTGAGGATGAGGACTCCATCTTTGCCCACTATGA
GTCAGGCATCAACTTCAACAAGTATGACGACATTCTGGTTGACGTCAGTGGAACCAACCCACCACAGG
CCATCATGACCTTTGCTGAGGCAGCACTGTGCGAGTCCCTGAGTAAAAACGTCAGTAAATCTGGTTAC
GTGAAGCCGACTCCTGTGCAGAAGCACGGCATCCCCATCATCTCTGCCGGCAGAGATCTCATGGCCTG
TGCCCAGACTGGATCTGGTAAAACGGCTGCGTTCCTGCTGCCCATCCTACAGCAGCTGATGGCAGACG
GCGTGGCAGCCAGTCAGTTCAGTGAACTGCAGGAGCCTGAAGCAATCATCGTCGCTCCAACCAGGGAG
CTCATCAACCAGATCTACCTGGAGGCCAGGAAGTTCGCCTTTGGGACCTGTGTGCGCCCAGTGGTGGT
TTATGGTGGAGTCAGCACTGGACACCAGATCAGGGAAATATGCAGAGGATGCAACGTCCTGTGTGGAA
CTCCAGGGAGACTGCTGGATGTGATTGGACGAGGAAAGGTGGGGCTCCACAAGCTGCGCTACCTGGTG
CTGGATGAGGCCGACCGTATGCTGGATATGGGCTTTGAGCCTGACATGCGCCGCCTGGTGGGCTCCCC
TGGAATGCCCTCCAAAGAGCAGCGTCAGACCCTGATGTTCAGCGCAACCTACCCCGAGGACATCCAGA
GGATGGCAGCTGACTTCCTCAAGACCGACTACCTGTTCTTAGCTGTGGGCGTGGTGGGCGGAGCTTGC
AGTGATGTGGAGCAGAAGTTTATTGAAGTAACTAAGTTCTCCAAGAGGGAGCAGCTTCTTGACATCCT
GAAGACAACAGGAACAGAGCGCACCATGGTGTTTGTGGAGACCAAGAGACAGGCTGATTTTATTGCCA
CATTCTTGTGCCAGGAGAAGGTTCCAACTACCAGCATCCATGGGGACCGGGAGCAGCGGGAACGGGAG
CAGGCCCTGGCAGACTTCCGCTCTGGGAGATGTCCAGTCCTGGTGGCAACCTCTGTTGCTGCCCGTGG
```

TCTGGATATTCCAGATGTTCAGCATGTGGTGAACTTTGACCTCCCCAACAATATAGACGAATACGTCC
ACCGTATTGGGAGAACTGGCCGCTGTGGAAACACTGGAAGGGCGGTGTCGTTCTATGACCCCGACAAT
GATGGACAGCTGGCTGGGTCCCTTGTCAGCATCCTGTCCAAGGCCCAGCAGGAAGTGCCCTCCTGGTT
AGAGGAGTGTGTGTTCAGCGGCTCAGGTGTGAACCCCTCCAGGAGGACCTTTGCGTCCACAGACTCCA
GGAAGGGTCCACAGGGCAGCTCCTTTCAGGACAGCAGTATGACGAGCCAGCCGGCTGTTCCTGCTGCA
GCTGATGATGAAGACTGGGAGTGAGAGGAGCATGAAGC

Figure 46. Partial nucleotide sequence of *Vasa* in *S. lalandi*, obtained from the consensus sequence of nine different colonies containing *Vasa* inserts. Non-confirmed sequence obtained from the Ion Torrent transcriptome data is highlighted in grey. The stop codon is highlighted in black, with the 3'UTR underlined.

GGDDGVCENGFRGGSRGGRGSRGGRGFRQGGDQGGRGGFGGGYRGKDEQIFARGEDKDPEKKDGNDGD
RPKVTYVPPTLPEDEDSIFAHYESGINFNKYDDILVDVSGTNPPQAIMTFAEAALCESLSKNVSKSGY
VKPTPVQKHGIPIISAGRDLMACAQTGSGKTAAFLLPILQQLMADGVAASQFSELQEPEAIIVAPTRE
LINQIYLEARKFAFGTCVRPVVVYGGVSTGHQIREICRGCNVLCGTPGRLLDVIGRGKVGLHKLRYLV
LDEADRMLDMGFEPDMRRLVGSPGMPSKEQRQTLMFSATYPEDIQRMAADFLKTDYLFLAVGVVGGAC
SDVEQKFIEVTKFSKREQLLDILKTTGTERTMVFVETKRQADFIATFLCQEKVPTTSIHGDREQRERE
QALADFRSGRCPVLVATSVAARGLDIPDVQHVVNFDLPNNIDEYVHRIGRTGRCGNTGRAVSFYDPDN
DGQLAGSLVSILSKAQQEVPSWLEECVFSGSGVNPSRRTFASTDSRKGPQGSSFQDSSMTSQPAVPAA
ADDEDWE

Figure 47. Partial amino acid sequence of *Vasa* in *S. lalandi*, predicted from the nucleotide sequence using the ExPASy Translate Tool.

3.4.2 *Amh*

Using the nucleotide sequences in Figure 34 and Figure 41, a fragment of *Amh* DNA sequence was determined, including the 3' end and the majority of the CDS. As the 5' end of the *Amh* transcript was not sequenced, sequence from the Ion Torrent transcriptome data was included (Figure 48). This sequence was translated using the ExPASy Translate Tool to construct a sequence of 551 amino acids (Figure 49).

ATGAAGAAGAGCCTAAAAGACCTCCTCATTGGTGAAAAAACAGGAAGTACAATCAGCATAACTCTACT
TCTGCTTTTCTCTGGGGGAACGGGAACTGATACAAGACAAACACACGTTTCAGGTTCATCTCTGGCCT
CTTCACAGACCTTCTCCTTCCTCTGTGAGCTGAAACGGTTCCTAGGTGAAGTCCTGCCTCAGYACCAC
CCTGAGTCCCCTCCATTCCAGCTGGAGTCCTTACCCTCCATGCCACCCCTGTCACTGGGCTTATCCTC
CAGTGAGACCCTGCTGGCAGGACTGATCAACTCCTCTTTCCTTACTGTCTTCTCCTTCGCTAGATGGG
GCTCCATTTCTCAGGTGCATCCTGGACAGCTGGCCATGTCTCCTGCACTGGTGGAGGAGGTCAGGCAG
AGATTGGCAGATNTGNGGATGCAGATAATGGAAGTAATAAGGGAGAAGGAGGTGGGTCACAGGGCCAC
GGAGAGGCTGGAGAGGCTTCAAGAATTCAGTGCATTACCGACGATGGAACAAGCAGCAGGACGGAGCC
AGTACTGTGCGTTTCTTCTGCTGAAGGCCCTGCAGACAGTGGCCCGTGAGTACGAGCTGCAGAGAGGA
CTGCGAGCCACCAGAGCGGACACCAATAACCCGGTGAGGGGCAGTGTCTGCAGTGTGAGAAGCCTCAC
CGTCAACCTTGAAAGGCGTCTTGTGGGTCCAAGCGCTGCCAACATCAAAAACTGCCATGGCTCTTGTG
CTTTCCCCCTGACCAATGCAAACGACCATGTTGTCCTGCTCCACTTCCACATTGAGAGCGGGAATATG
GATGAGCGGGCTCCGTGCTGTGTGCCTGTTGCCTATGAACCCCTTGAGGTGGTGGGTTTGAATGAACA
TGGGACTTACCTCTCCATGATACCAGATATGGTAGCAAAGGAGTGTGGATGCCGCTGAAGCTGTTTTC
TTTTTGGAGTGTGTTTGTGACATGAGAGATATTTCACTTGTTTGGTTTGATAGTCACTGACTGTAGCT
TGTTAGGTGGAGGAAAAGTGTTATTTAATGTTTTTTTTATTCACCTATTCACATATGTTAATTCAGAG
AGACTTTTTTTATAAACCTACCTAGATGCACACAGCAAAATGTTTCACACTTAGAAATAATCAAGTTT
ATTGAAGACAAAAATATATGTGTTAACAGCTATTTATTATTTTAAACTTCTTCTATAACACCATATTT
TGTTTTATAATGAAGATTCCCCGTATTAAAAATATTGTTCTAAAAAAAAAA

Figure 48. Partial nucleotide sequence of *Amh* in *S. lalandi,* obtained from the consensus sequence of ten different colonies containing *Amh* inserts. Non-confirmed sequence obtained from the Ion Torrent transcriptome data is highlighted in grey. The stop codon is highlighted in black, with the 3'UTR underlined.

MKKSLKDLLIGEKTGSTISITLLLLFSGGTGTDTRQTHVSGSSLASSQTFSFLCELKRFLGEVLPQXH
PESPPFQLESLPSMPPLSLGLSSSETLLAGLINSSFLTVFSFARWGSISQVHPGQLAMSPALVEEVRQ
RLADXXMQIMEVIREKEVGHRATERLERLQEFSALPTMEQAAGRSQYCAFLLLKALQTVAREYELQRG
LRATRADTNNPVRGSVCSVRSLTVNLERRLVGPSAANIKNCHGSCAFPLTNANDHVVLLHFHIESGNM
DERAPCCVPVAYEPLEVVGLNEHGTYLSMIPDMVAKECGCR

Figure 49. Partial amino acid sequence of *Amh* in *S. lalandi,* predicted from the nucleotide sequence using the ExPASy Translate Tool.

3.5 Gene Sequence Analysis

3.5.1 *Vasa*

The *S. lalandi* partial VASA amino acid sequence was aligned to other available VASA amino acid sequences in Perciformes using Clustal Omega. Conserved regions and important structural characteristics were identified using InterProScan and BLAST. VASA in *S. lalandi* contains DEADc (DEAD-box helicase) and HELICc (helicase superfamily C-terminal) domains characteristic of VASA proteins in vertebrates (Figure 50, 51).

```
S.lalandi           ------------------------------------------------------------
S.quinqueradiata    MDEWEEEETSTSSVTLTSHASSQGSQGDFM-NSDGEFGRGRGGRGRGRRGGFSSSFSSGG
T.japonicus         MDEWNEESTPASSASYSANDSS---GKDFRNGDSGEFGRGRGRGGRGRRGGFSSSFSAGG
S.australasicus     MDEWEEEGTVS-TIPLTSYTPNEGTQGDSWNTDAGECARGRGGRGRGR-GGFKSSYSSGG
T.orientalis        MDEWEEEGNTS-TITLTSRTSSEGTQGDFWNTNGGEFGRGRGGRGRGR-GGFKSSYSSGG


S.lalandi           ------------------------------------------GGDDGVCENGFRGGSRG
S.quinqueradiata    DEH--GNGGDSWNNTGGERDGFRGRGGRGRGRGFGGMDCIEF-GGDDGVCENGFRGGSRG
T.japonicus         DDDEGGDSSSSWNNSGGDRGGFRGRG-RGRGRGFGGGDRGDRGDGDDGVSENGFRGGRGG
S.australasicus     D----GNDGDKWNNAGGERGGFRGRGGQGRGRGFGRTDHSEVNGDDSGVCENGFRGGR--
T.orientalis        D----GNDEDKWNNAGGERGGFRGRGGQGRGRGFCRMDQSEFNGDDNGMRENGYRGGR--
                                                          *.*: ***:***


S.lalandi           GRGSRGG-RGFRQGGDQGGRGGFG---GGYRGKDEQIFARGEDKDPEKKDGNDGDRPKVT
S.quinqueradiata    GRGSRGG-RGFRQGGDQGGRGGFG---GGYRGKDEQIFARGEDKDSEKKDGNDGDRPKVT
T.japonicus         GRGSRGGGGGFRQGGDEGGRGGFGGGGGGYRGKDEQVFSKGDDKDGENKDGAEDERPRVT
S.australasicus     --GGRGRGGSFSQGRDSGGRGGFG---GGYRGKDEEIFSQGEDKDQDKKD--DSERPKVT
T.orientalis        --GGRGRGGGFSQGGDQGGRGGFG---GGYRGKDEEIFTRGEDKDPEKKD--DSDRPKIT
                      *.**   .* ** *.*******   ********::*:::*::*** ::**  : :**::*
                                                                 DEADc →
S.lalandi           YVPPTLPEDEDSIFAHYESGINFNKYDDILVDVSGTNPPQAIMTFAEAALCESLSKNVSK
S.quinqueradiata    YVPPTLPEDEDSIFAHYESGINFNKYDDILVDVSGTNPPQAIMTFAEAALCESLSKNVSK
T.japonicus         YIPPALPDDEDSIFAHYESGINFNKYDDILVDVSGSNPPQAVMTFEEAALCESLRKNVSK
S.australasicus     YVPPTLPEDEDSIFSHYETGINFDKYDDIMVDVSGTNPPQAIMTFDEAQLCESLRKNVSK
T.orientalis        YVPPTLPEDEDSIFSHYETGINFDKYDDIMVDVSGTNPPQAVMTFDEAALCESLRKNVSK
                    *:**:**:******:***:****:*****:*****:*****:*** ** ***** *****

S.lalandi           SGYVKPTPVQKHGIPIISAGRDLMACAQTGSGKTAAFLLPILQQLMADGVAASQFSELQE
S.quinqueradiata    SGYVKPTPVQKHGIPIISAGRDLMACAQTGSGKTAAFLLPILQQLMADGVAASQFSELQE
T.japonicus         SGYVKPTPVQKHGIPIISAGRDLMACAQTGSGKTAAFLLPILQKLMADGVAASQFSELQE
S.australasicus     SGYVKPTPVQKHGIPIISAGRDLMACAQTGSGKTAAFLLPILQQLMADGVAASRFSELQE
T.orientalis        SGYVKPTPVQKHGIPIISAGRDLMACAQTGSGKTAAFLLPILQQLMADGVAASRFSELQE
                    ****************************************:*********:******

S.lalandi           PEAIIVAPTRELINQIYLEARKFAFGTCVRPVVVYGGVSTGHQIREICRGCNVLCGTPGR
S.quinqueradiata    PEAIIVAPTRELINQIYLEARKFAFGTCVRPVVVYGGVSTGHQIREICRGCNVLCGTLGR
T.japonicus         PEAIIVAPTRELINQIYLEARKFAFGTCVRPVVVYGGVASGHQIRDICRGCNILCGTPGR
S.australasicus     PEAIIVAPTRELINQIYLEARKFAFGTCVRPVVVYGGVSTGHQIRDIERGCNIVCGTPGR
T.orientalis        PEAIIVAPTRELINQIYLEARKFAFGTCVRPVVVYGGVSTGHQIREIERGCNVVCGTPGR
                    ***************************************::*****:* ****::*** **

S.lalandi           LLDVIGRGKVGLHKLRYLVLDEADRMLDMGFEPDMRRLVGSPGMPSKEQRQTLMFSATYP
S.quinqueradiata    LLDVIGRGKVGLHKLRYLVLDEADRMLDMGFEPDMRRLVGSPGMPSKEKRQTLMFSATYP
T.japonicus         LLDVIGRGKVGLEKLRYLVLDEADRMLDMGFEPDMRRLVGSPGMPTKENRNTLMFSATYP
S.australasicus     LLDMIGRGKVGVSKLRYLVLDEADRMLDMGFEPDMRRLVGSPGMPSKENRQTLMFSATYP
T.orientalis        LLDMIGRGKVGLSKLRYLVLDEADRMLDMGFEPDMRRLVGSPGMPSKENRQTLMFSATYP
                    ***:*******: ********************************:**:*:*********
                                                 HELICc →
S.lalandi           EDIQRMAADFLKTDYLFLAVGVVGGACSDVEQKFIEVTKFSKREQLLDILKTTGTERTMV
S.quinqueradiata    EDIQRMAADFLKTDYLFLAVGVVGGACSDVEQKFIEVTKFSKREQLLDILKTTGTERTMV
T.japonicus         EDIQRMAADFLKTDYLFLAVGVVGGACSDVEQTFIEVTKFSKREQLVDLLKATGSERTMV
S.australasicus     EDIQRMAADFLKTDYLFLAVGVVGGACSDVEQTFIQVTKFSKREQLLDLLKTTGMERTMV
T.orientalis        EDIQRMAADFLKTDYLFLAVGVVGGACSDVEQTFIQVTKFSKREQLLDLLKTTGTERTMV
                    *********************************.**:**********:*:**:** *****

S.lalandi           FVETKRQADFIATFLCQEKVPTTSIHGDREQRERERQALADFRSGRCPVLVATSVAARGLD
S.quinqueradiata    FVETKRQADFIATFLCQEKVPTTSIHGDREQRERERQALADFRSGRCPVLVATSVAARGLD
T.japonicus         FVETKRQADFIATFLSQTKIPTTSIHGDREQRERERQALADFRSGKCPVLVATSVAARGLD
S.australasicus     FVETKRQADFIATFLCQEKVPTTSIHGDREQRERERQALADFRSGKCPVMVATSVAARGLD
T.orientalis        FVETKRQADFIATFLCQEKVPTTSIHGDREQRERERQALADFRSGKCPVLVATSVAARGLD
                    ***************.* *:*************************:***:***********

S.lalandi           IPDVQHVVNFDLPNNIDEYVHRIGRTGRCGNTGRAVSFYDPDNDGQLAGSLVSILSKAQQ
S.quinqueradiata    IPDVQHVVNFDLPNNIDEYVHRIGRTGRCGNTGRAVSFYDPDNDGQLAGSLVSILSKAQQ
T.japonicus         IPDVQHVVNFDLPNNIDEYVHRIGRTGRCGNTGRAVSFYDQDADGQLARALVTILSKAQQ
S.australasicus     IPDVQHVVNFDLPNNIDEYVHRIGRTGRCGNTGRAVSFYDPDGDSQLASSLVTVLSKAQQ
T.orientalis        IPDVQHVVNFDLPNNIDEYVHRIGRTGRCGNTGRAVSFYDPDADGQLARSLVTVLSKAQQ
                    ************************************ * *.*** :**::******
```

```
S.lalandi              EVPSWLEECVFSGSGVN---PSRRTFASTDSRKGPQGSSFQDSSMTSQ-PAVPAA-ADDE
S.quinqueradiata       EVPSWLEECVFSGSGVN---PSRRTFASTDSRKGPQGSSFQDSSMTSQ-PAVPAA-ADNE
T.japonicus            TVPSWLEESAFSGSGGAGI-NLSRTFASTDSRKGAQGDSFQDGGRNAAAAAASAA-ADDE
S.australasicus        EVPSWLEESAFSGSTSSSFKPPRKNFASTDSRKG---GSFQDNSMQSQEAVQPAANDDDE
T.orientalis           EVPSWLEESAFSGPATTGFNPPRKNFASTDSRKR---GSFQDNSVKSQPAVQTAA-DDDE
                        *******..***          :.********     **** .  :   .  **  *:*

S.lalandi              DWE
S.quinqueradiata       DWE
T.japonicus            EWE
S.australasicus        EWE
T.orientalis           EWE
                       :**

  [*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 50. Alignment of the *S. lalandi* partial VASA amino acid sequence with four other VASA sequences in Perciformes. The conserved DEADc domain is highlighted in blue, while the HELICc domain is highlighted in green.

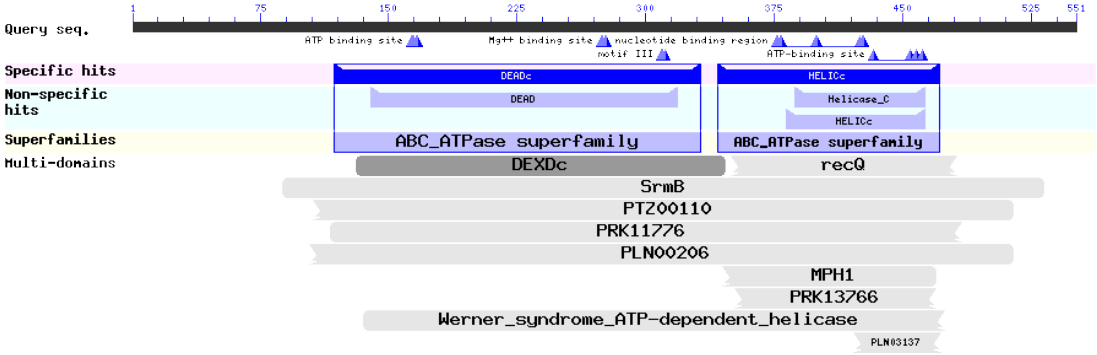


Figure 51. Conserved domains of VASA in *S. lalandi*, generated by BLAST searching the partial amino acid sequence. The DEADc and HELICc domains are characteristic of VASA proteins in vertebrates.

To confirm the identity of the sequence, a phylogenetic tree was constructed using the *S. lalandi* partial VASA amino acid sequence and other known VASA sequences in vertebrates obtained from the UniProt database (Figure 52).

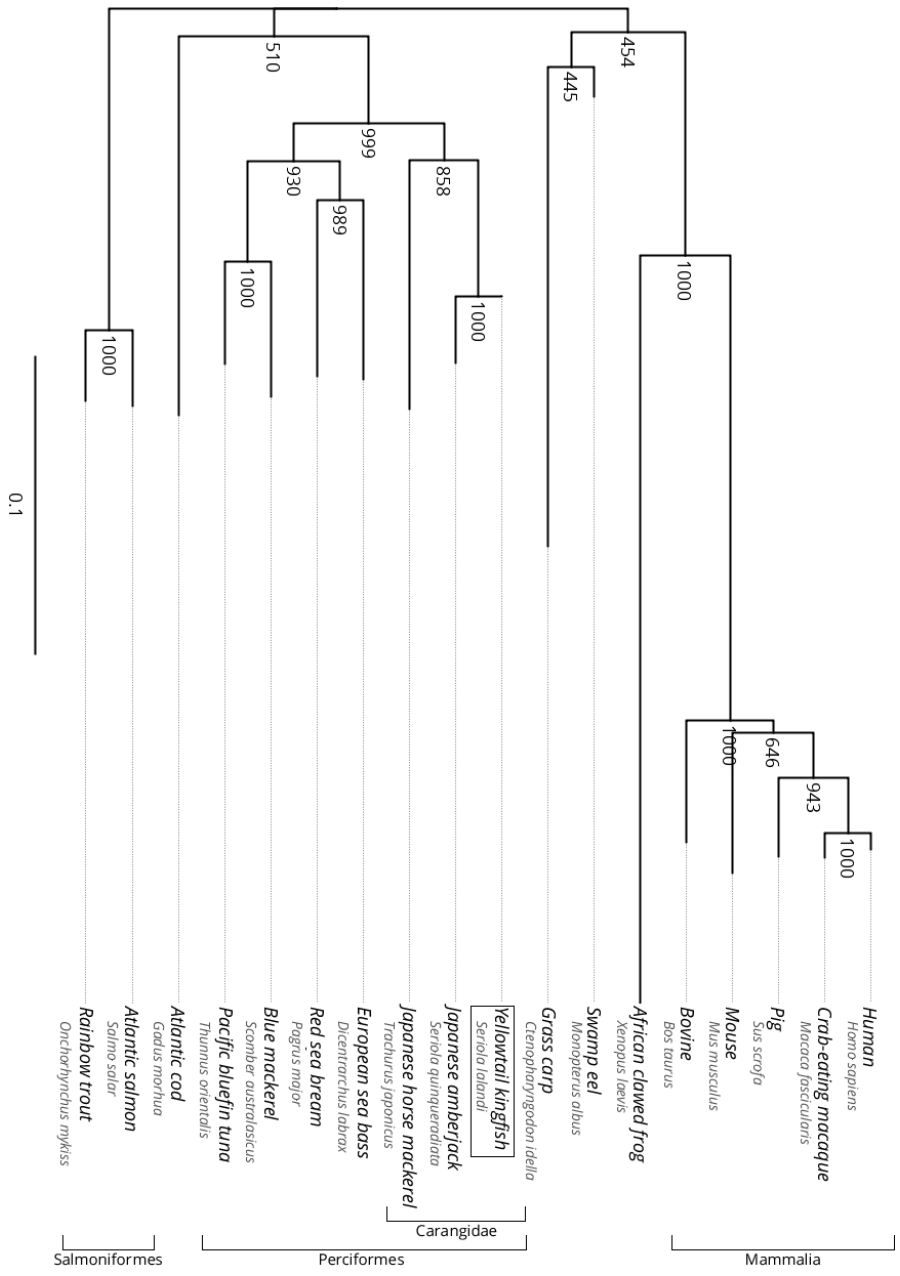


Figure 52. Phylogenetic tree constructed in FigTree 1.4.2 using the neighbour joining method from an alignment in ClustalX 2.1, showing the relationship between VASA proteins across a range of vertebrate species. Black lines represent genetic distance. Grey lines are species labels. In addition to the *S. lalandi* partial amino acid sequence (boxed), sequences used were: *H. sapiens* (Q9NQI0), *M. fascicularis* (Q4R5S7), *S. scrofa* (Q6GWX0), *M. musculus* (Q61496), *B. taurus* (F1MYC6), *X. laevis* (Q91372), *T. japonicus* (B5U9B5), *S. quinqueradiata* (D4P9H5), *S. australasicus* (D4P8B9), *T. orientalis* (B6YZ85), *D. labrax* (F1ARL6), *P. major* (E3W9C6), *O. mykiss* (Q9PT10), *G. morhua* (E7CQY3), *S. salar* (M1F1H3), *C. idella* (C5MR43) and *M. albus* (Q48HI1).

3.5.2 *Amh*

The *S. lalandi* partial AMH amino acid sequence was aligned to other available AMH amino acid sequences in Perciformes using Clustal Omega (Figure 53). Conserved regions and important structural characteristics were identified using InterProScan and BLAST. AMH in *S. lalandi* contains the distinctive AMH_N (Anti-Müllerian hormone N terminal region), as well as the TGFβ (Transforming growth factor beta) domain (Figure 54).

```
S.lalandi         ------------------------------------------------------------
O.niloticus       MLGLLVLYSEALTLCW-------TLQPAQDP------TVTEYSLPSAK----------T-
P.scalare         MLGLLVIYSELMLLCCTRLSVVLQLNPAQDS------PITEYSLPTLK----------T-
A.schlegelii      MLGVDVFYCGALMLCWTRLCVALHVSHGQPLIPAHDPEMTDLHAASSTETGDGLEIK---
L.japonicus       ------------------MCVALQVSPGRRLVPTTT---GDHHSAGSTKAGDGLEIRNAE

S.lalandi         ------------------------------------------------------------
O.niloticus       ---------PSSPSSSSAAAPHAAPCFVEDIFAALRDGVGDSGELTNSSLVLFGFCSQSA
P.scalare         ---------S-----AAAAVPRIAPCFVDDMFTVLREGVGDGDELINSTLALFGICTVSD
A.schlegelii      -----------------NHAPHHEPCFVDNILAALREAVGNGGELTNRSLTLFGICTATG
L.japonicus       NTPPSTQTSLDLISDVSHHVPRRAPCFVEDIFAALREGVGDDGELTNRSLSLFGICTVSG

S.lalandi         ------------------------------------------------------------
O.niloticus       RSSASVSLDLANKK-----SSLEVLHPAAVHVSEEEEQGTITLTFDLPRPPSLMTNPVLL
P.scalare         HSS--RLLDLANKK-----DSLGVLHPAAAHVSEEEGQGTLTLTFDLPQSPLLVTNPVLL
A.schlegelii      DPSGSVLLELTNEASR---NGLEALHPAAVHLSEEDERGALKLTFDLEQSPLLERKPVLL
L.japonicus       NSSGSILLELSKETSRNQGNGLEALHPAGVLLGEEDERGTLKLTFDLPQSSLLKLNPVLL

S.lalandi         ------------------------------------------------------------
O.niloticus       LVFENPLARGDLEVAFTSQFLQPNTQAVCISGDTQYVLLTGKSSEGSVNDRWQITAQTKL
P.scalare         LAFESPLTQGDLDVAFTSPSLQPNTQMVCISGGTQYILLTGKSPEGSANQKWQISAEAKI
A.schlegelii      LAFKSPLSGGNLDIMFTSQSLQPNTQPVCISGETQYMVLTGKASEDNAHQNWRVSADTKS
L.japonicus       LGFESPLTGGNLDVTFTGQSLHPNTQSVCISGETQYIMLTAKASEGSVHQTWRISVEANS

                      AMH_N→
S.lalandi         --MKKSLKDLLIGEKTGSTISITLLLLFSGGTGTDTRQTHVSGSSLASSQTFSFLCELKR
O.niloticus       PHMKQNLKSILIGEKSGSNISMSPLLLFSGGTGTDTRCAS--GSPPASL-QTSFLCEMKR
P.scalare         PAMKQNLKDILIGGKPGSNISMTPLLLFSRERGTDSRCVP--GSSPASL-QTSFLCELKR
A.schlegelii      PDMKQRLRDVLTGGKSGSSISLIPLLLFSGGRGTDTRHVHGSGS----SQTSSFLCELKR
L.japonicus       PDMKQSLKDILIGGKSGSNISMTPLLLFSGERGTDTRRTHVSGSSPASPQTSSFLCELKR
                    **: *:.:* * * **.**:  *****   ***:* .   **        *****:**

S.lalandi         FLGEVLPQXHPESPPFQLESLPSMPPLSLGLSSSETLLAGLINSSFLTVFSFARWGSISQ
O.niloticus       FLGAVLPQEHFTSPPLPLDSLQSLPPLSLGLSSSETLLAVMINSTAPTVFGFTSWGSVLP
P.scalare         FLGAALPQKHPISPPLPLDSLQSLPPLTLGLSSSETLLAAIINSTAPTIFAFTSWGSIFP
A.schlegelii      FLGDVMPQGQPESPPLQLTSLQSMPPLTLGVSSSETLLAGLINSSALTIFSFSTCCSVSQ
L.japonicus       FLGDVLPQDHPESPPLQLDSVQSLPPLTLGLSSSETVLAGLINSSAPTVFSFTSLGPGFQ
                  *** .:** :  ***: * *: *:***:**:*****:** :***:  *:*.*:

S.lalandi         VHPGQLAMSPALVEEVRQRLADXXMQIMEVIREKEVGHRATERLERLQEFSALPTMEQAA
O.niloticus       VCHGELALSAALLEELRQRLDQTLVQMTEIIREEEVSLGAKESLGRLKELSALQEKEHAT
P.scalare         VCHGELALSPALLEELRQRLEQTLVQMTYIIREENVAQGAKKSLGRLKDLSALQKDEHDT
A.schlegelii      VHRRELALSPALAEELGQRLEQTMMQITEVIREEDVGHRATERLGRLRELSAFPKEEPAT
L.japonicus       VHHGELALSPALLEEVRQRLEQTVVQIMEVIRGEEMGHRSTVRLGRLKELSAFPEKEPAA
                  *   :**:* ** **: *** :  :*:  :** :::.  :.  * **:::**:   *  :
```

```
                                                                TGFB→
S.lalandi           GRSQYCAFLLLKALQTVAREYELQRGLRATRADTNNPVRGSVCSVRSLTVNLERRLVGPS
O.niloticus         GGSQFRVFLLLKALQTVAQTYDAQRKLRATRADPSSSVRGGVCGLKALTVSLTKLLVGPS
P.scalare           GGSQFRAFLLLKALQTVAQTYDMQRKLRATRANPSSSEMRHICGLRTLTVSLAKLLVGPP
A.schlegelii        GESQYRAFLLLRALQTVARTYVLKRGQRTARAGPSNPGGANICGLRSLTVSLERHLLAPN
L.japonicus         GESQYRAFLLLKALQTVARAYEVQRGLRATRAGPNNPVRGHVCGLRSLTVSLERRLVGPN
                    * **: .****:******: *  :*  *::**  ..     :*.:::***.* : *:.*

S.lalandi           AANIKNCHGSCAFPLTNANDHVVLLHFHIESG-----NMDERAPCCVPVAYEPLEVVGLN
O.niloticus         SANINNCHGSCTFPLTNGNNHAILLNSHIETG-----NADERSPCCVPVAYEALEVVDWN
P.scalare           SANINSCHGSCAFPLANANNHAIMLNSHIESG-----NVDVRAPCCVPVAYDPLEVMDWN
A.schlegelii        TANIDNCQGTCPVPLVNPTNHAVLLNSHIESERAASRSADERPPCCVPVAYEALEVVDFN
L.japonicus         TANINNCHGSCAFPLVNANNHAVLLNSHIESG-----NVDERAPCCVPVAYEALEVVDLN
                    :***..*:*:* .**.* .:*.::*. ***:      . * * ********: ***:  *

S.lalandi           EHGTYLSMIPDMVAKECGCR
O.niloticus         ADGTFISIKPDAVARECGCR
P.scalare           AEETFISIKPDVIARECGCR
A.schlegelii        EDLTQVRPFENVVVKECGCR
L.japonicus         PHGTYLSIKPDMVAKECGCR
                     . * :    : :.:*****

  [*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 53. Alignment of the *S. lalandi* partial AMH amino acid sequence with four other AMH sequences in Perciformes. The conserved AMH_N region is highlighted in blue, while the TGFβ domain is highlighted in green.

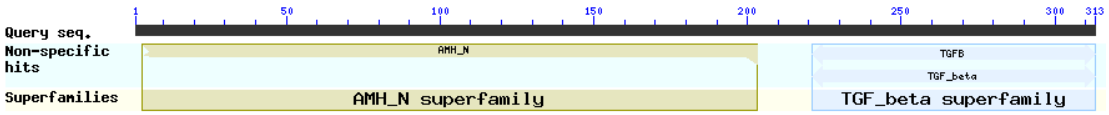


Figure 54. Conserved domains of AMH in *S. lalandi*, generated by BLAST searching the partial amino acid sequence.

To confirm the identity of the sequence, a phylogenetic tree was constructed using the *S. lalandi* partial AMH amino acid sequence and other known AMH sequences in vertebrates obtained from the UniProt database (Figure 55).

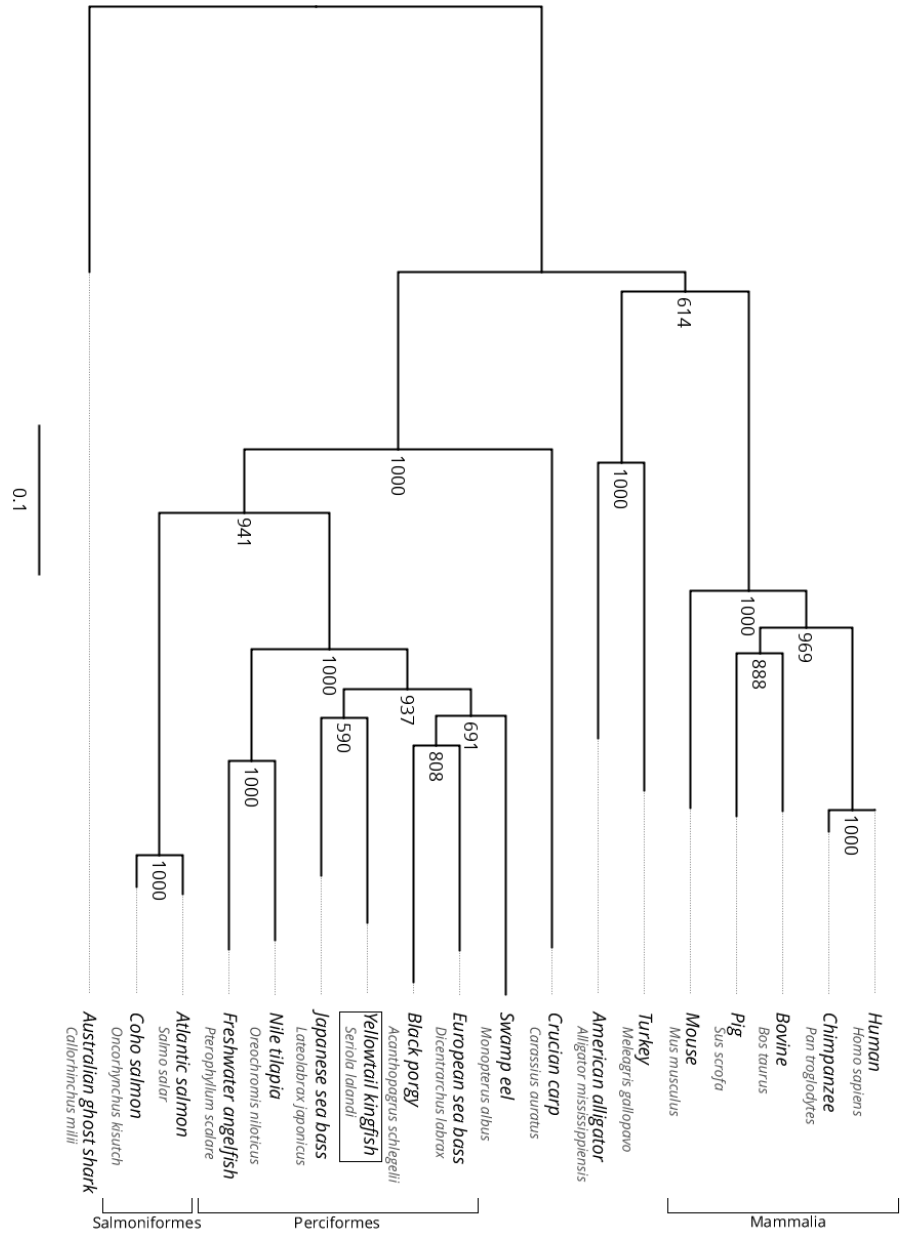


Figure 55. Phylogenetic tree constructed using the neighbour joining method from an alignment in ClustalX 2.1, showing the relationship between AMH proteins across a range of vertebrate species. Black lines represent genetic distance. Grey lines are species labels. In addition to the *S. lalandi* partial amino acid sequence (boxed), sequences used were: *H. sapiens* (P03971), *P. troglodytes* (H2QEW1), *B. taurus* (P03972), *S. scrofa* (P79295), *M. musculus* (P27106), *M. gallopavo* (G1MT41), *A. mississippiensis* (Q9PUJ9), *C. auratus* (K4PJL6), *S. salar* (Q5XZF0), *O. kisutch* (E7EEK9), *O. niloticus* (A1XPM9), *P. scalare* (W0G6F5), *L. japonicus* (H6WS80), *D. labrax* (Q0VYL2), *A. schlegelii* (D2XUR9), and *M. albus* (W0FQQ7). The sequence for *C. milli* (V9KJS5) was used as an outgroup.

3.5.3 *Cyp19a1a*

The *S. lalandi* partial gonadal aromatase amino acid sequence was aligned to other available aromatase amino acid sequences in Perciformes using Clustal Omega (Figure 56). Conserved regions and important structural characteristics were identified using InterProScan and BLAST, which revealed the entire *S. lalandi* partial sequence matches the cytochrome P450 domain (Figure 57).

```
S.lalandi       ------------------------------------------------------------
O.niloticus     MDLISACEQAMNPVGLDAVVARSLCD---LKCHPIDGISMATRTLILLVCLLLVAWSHTD
P.major         MDLISACELVMPQVGLDTAVADLVSMSHNATAVGTPGISVATRTLILLVCLLLVAWNHTE
S.aurata        MDLISACERVMPQVGLDTTAADLVPMSHNASAVGAPGISVVTRTFILLICLLLVAWNSME
D.labrax        MDLISACERAMTPVGLDTIVADLVSTSPNATAVGSPGISVATITLILLVCLLLVAWSHTD

                                                          P450→
S.lalandi       ---------------------------------------------WINGEETLIISRPSAV
O.niloticus     KKIVPGPSFCLGLGPLLSYLRFIWTGIGTASNYYNNKYGDIVRVWINGEETLILSRSSAV
P.major         KKSVPGPSFCLGLGPLLSYLRFIWTGIGTASNYYNKTYGDIVRVWINGEETLILSRASAV
S.aurata        KKSVPGPSFCLGLGPLLSYLRFSWTGIGTASNYYNKKYGDMVRVWINGEETLILSRASAV
D.labrax        KNTVPGPSFCLGLGPLLSYLRFIWTGIGTASNYYNNKYGDIVRVWINGEETLILSRASVV
                                                         *********:** *.*

S.lalandi       HHVLKNGLYCSRFGSKQGLSCVGMNERGLIFNNNVTLWKKMRTYFTKALTGPGLQKTVEV
O.niloticus     HHVLKNGNYTSRFGSIQGLSYLGMNERGIIFNNNVTLWKKIRTYFAKALTGPNLQQTVDV
P.major         HHVLKNGQYTSRFGSRQGLSCIGMNERGIIFNNNVTLWKKIRTYFTKALTGPSLQQTVEI
S.aurata        HHVLKSGQYTSRFGSRQGLSCIGMNERGIIFNNNVTLWKKIRTYFTKALTGPSLQQTVEI
D.labrax        HHVLKNGHYTSRFGSKQGLSCMGMYERGIIFNNNVTLWKQIRNYFSKALTGPGLQQTVEV
                *****.* * ***** **** :** ***:**********::*.**:****** **:**::

S.lalandi       CVSSTQTHLDDLQSLSHVDVLGLLRCIVVDISNRLFLGVPVDEKELLVKIHKYFDTWQSV
O.niloticus     CVSSIQAHLDHLDSLGHVDVLNLLRCTVLDISNRLFLNVPLNEKELMLKIQKYFHTWQDV
P.major         CVSSTQTHLDNLDTLDHVDVLSLLRCTVVDISNRLFLGIPVDEKELLLKIQKYFDTWQTV
S.aurata        CVSSTQTHLDNLDVLDQVDVLSLLRCTVVDISNRLFLDTPVDEKELLLKIQKYFDTWQTV
D.labrax        CASSTQTHLDDLDKLDNVDVLSLLRCTVVDISNRLFLGVPVNEKELLLKIQKYFDTWQTV
                *.** *:***.*: * .**** **** *:******** *::****::**:***.*** *

S.lalandi       LIKPDIYFKLDWIHRRHKTAAQELRDAIESLVEQKRRDVEQADKLDNINFTTELIFAQNH
O.niloticus     LIKPDIYFKFRWIHHRHKTATQELQDAIKRLVDQKRKNMEQADKLDNINFTAELIFAQNH
P.major         LIKPDIYFKFGWIHQRHKTAAQELQDAIESLVEQKRRDMEQADKLDNINFTAELIFAQNH
S.aurata        LIKPDIYFKFGWIHQRHKAAAQELQDAIESLVEQKRRDMEQADKLDNINFTAELIFAQNH
D.labrax        LIKPDIYFKFDWIHQRHKTAAQELQDAIESLVEQKRRDMEQADKLD-INFTADLIFAQNR
                ********* ***:***:*:***:***: **:***:::******* ****::******:

S.lalandi       GELSAENVLQCVLEM---------------------------------------------
O.niloticus     GELSAENVTQCVLEMVIAAPDTLSLSLFFMLLLLKQNPHVEPQLLQEIDAVVGERQLQNQ
P.major         GELSAENVRQCVLEMVIAAPDTLSISLFFMLLLLKQNPDVELQLLQEIDTVIGDRQLQNR
S.aurata        GELSAENVRQCVLEMVIAAPDTLSISLFFMLLLLKQHPDVELQLLQEIDTVIGERQLQNG
D.labrax        GELTAENVRQCVLEMVIAAPDTLSVSLFFMLLLLKQNPDVELQLLQEIDTVVGERQLQNG
                ***:**** ******

S.lalandi       ------------------------------------------------------------
O.niloticus     DLHKLQVMESFIYECLSFHPVVDFTMRRALSDDIIEGYRISKGTNIILNTGRMHRTEFFL
P.major         DPQKLHVLESFINECLRFHPVVDFTMRRALSDDIIDGYRVPKGTNIILNTGHMHRTEFFN
S.aurata        DPQKLHVLESFINECLRFHPVVDFTMRRALSDDIIDGYRVPKGTNIILNTGHMHRTEFFN
D.labrax        DLQRLQVLESFVNECLRFHPVVDFTMRRALSDDIIDGYRVPKGTNIILNTGHMHRTEFFL
```

```
S.lalandi       ------------------------------------------------------------
O.niloticus     KGNQFNLEHFENNVPRPPTFQPFGSGPRACIGKHMAMVMMKSILVTLLSQYSVCTHEGPI
P.major         KPDEFRLENFEKNAPRR-YFQPFGSGPRACVGKHIAMVMMKSILVTLLSQYSVCPHEGLT
S.aurata        KPDEFRLENFERTAPRR-YFQPFGSGPRACVGKHIAMVMMKSVLVTLLSQYSVCPHEGLT
D.labrax        KPNEFNLDNFKKNAPRR-YFQPFGSGPRACVGKHIAMVMMKSILVTLLSQYSVCPHKGLT


S.lalandi       ---------------------------------------------
O.niloticus     LDCLPQTNNLSQQPVEHQQAETEHLHMRFLPRQGSSCQTLKDPNL
P.major         LDCLPQTNNLSQQPVEHQQEAESHLSMRFLPRQRSSWQTL-----
S.aurata        LDCLPQTNNLSQQPVEHQQEAESHLSMRFLPRQRGSWQTL-----
D.labrax        LDCLPQTNNLSQQPVEHQQEA-EHLSMRFLSRQRGSWKTL-----

[*]=Conserved sequence [:]=Conservative mutation [.]=Semi-conservative mutation
```

Figure 56. Alignment of the *S. lalandi* partial gonadal aromatase amino acid sequence with four other aromatase sequences in Perciformes. The conserved cytochrome P450 domain is highlighted in blue.

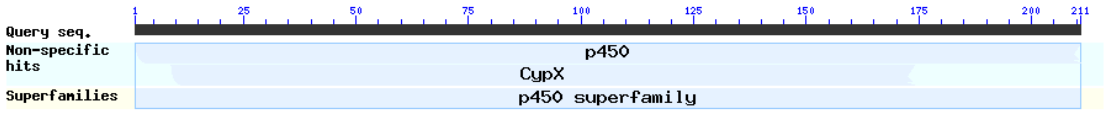


Figure 57. Conserved domains of gonadal aromatase in *S. lalandi*, generated by BLAST searching the partial amino acid sequence.

To confirm the identity of the sequence, a phylogenetic tree was constructed using the *S. lalandi* partial gonadal aromatase amino acid sequence and other known aromatase sequences in vertebrates obtained from the UniProt database (Figure 58).

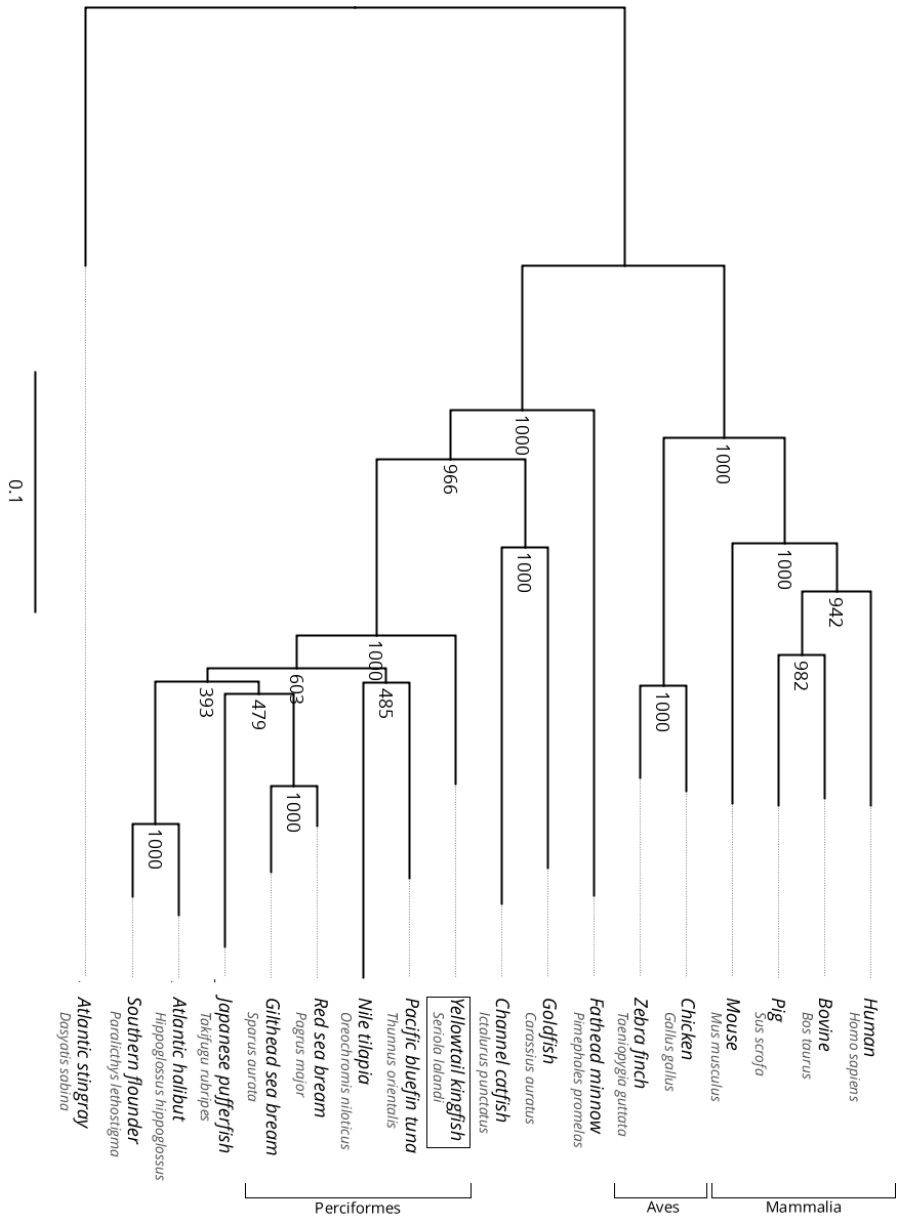


Figure 58. Phylogenetic tree constructed in FigTree 1.4.2 using the neighbour joining method from an alignment in ClustalX 2.1, showing the relationship between gonadal aromatase proteins across a range of vertebrate species. Black lines represent genetic distance. Grey lines are species labels. In addition to the *S. lalandi* partial amino acid sequence (I1SV64; boxed), sequences used were: *H. sapiens* (P11511), *B. taurus* (P46194), *S. scrofa* (Q29624), *M. musculus* (P28649), *G. gallus* (P19098), *T. guttata* (Q92112), *P. promelas* (Q90W65), *C. auratus* (H2DKW3), *I. punctatus* (Q92111), *T. orientalis* (E7FLX6), *O. niloticus* (P70091), *H. hippoglossus* (Q98SH1), *P. lethostigma* (Q3ZEL4), *T. rubripes* (A9EE90), *P. major* (Q8UWB5) and *S. aurata* (Q90XD5). The sequence for *D. sabina* (Q9PVG1) was used as an outgroup.

3.6 Real-Time PCR

3.6.1 Primer efficiencies

Due to this study being the first time that the expression of *Vasa*, *Amh* and *Cyp19a1a* has been looked at in this species, and the newly-designed primers had not been used before, it was important to work out the primer efficiency of each primer pair. Perfect primers with an efficiency of 100% would double the quantity of DNA after each cycle. However, it is relatively common for primers to deviate from this ideal. Calculating the exact efficiency can be used to determine whether these primers can be used; if the actual efficiency is lower than 100%, it can be used in any subsequent analysis. This investigation determined amplification efficiency by generating a calibration curve using serial dilutions of template cDNA of known concentration. Below is a demonstration of what was done to calculate this, using the housekeeping gene β-actin as an example.

β-actin primers were run in duplicate for each serial dilution of cDNA template that was prepared from *S. lalandi* testes. The amplification plot (Figure 59) was used to calculate the C_T values by placing a threshold line on the Y-axis, at a point where a detectable amount of amplicon product could be measured from each amplification during the early exponential phase. In addition, the melt curve (Figure 60) and gel electrophoresis analysis (Figure 61) of this real-time reaction confirmed that a single product was amplified within each reaction, and that no product was amplified in the negative control.

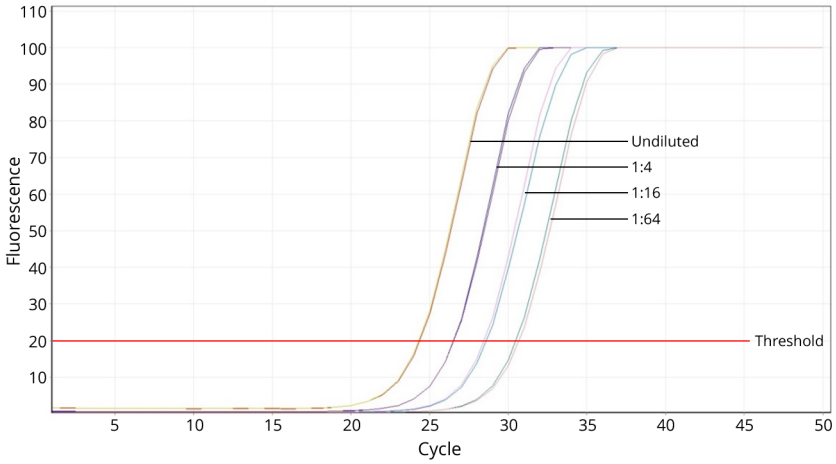


Figure 59. Amplification curves obtained for β-actin primers, using different concentrations (Undiluted, 1:4, 1:16 and 1:64) of template testes cDNA. This figure is an example of where the threshold can be set to obtain the corresponding C_T values.

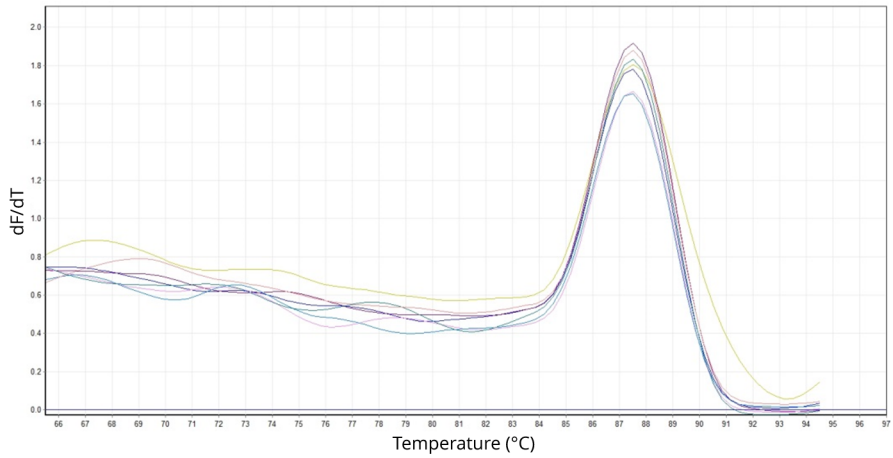


Figure 60. Melt curve obtained using the β-actin primers, which indicated the generation of one distinct product with no non-specific amplification.

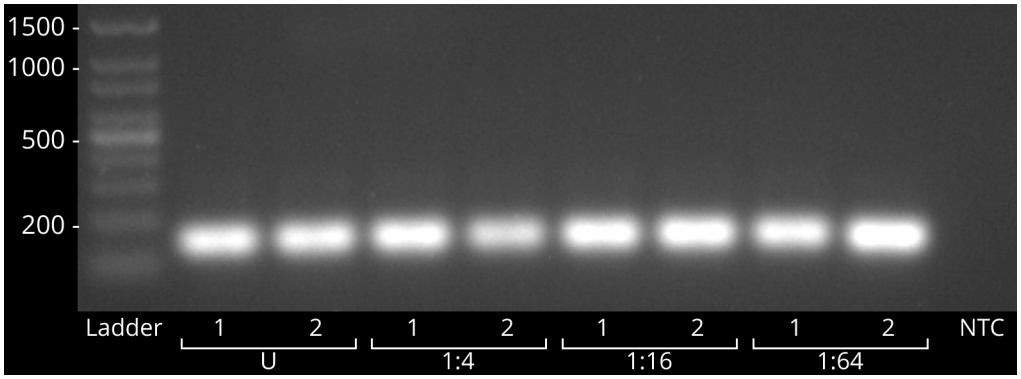


Figure 61. Gel electrophoresis analysis of real-time PCR products amplified using different concentrations of template testes cDNA with the β-actin primers. Each template generated a single band, with no band in the no template control (NTC). Lane #1 contains a 100 bp ladder. Lane #10 contains NTC. The remaining lanes contain paired replicates of undiluted template (U), 1:4 diluted template, 1:16 diluted template, and 1:64 diluted template.

The β-actin C_T values were plotted against the initial amount of input material on a semi-$\log_{10}$ plot (Figure 62). Using these points, a trend line was drawn which generated an R^2 value of 0.99, and an equation of the line:

$$y = -3.3718\,x + 27.74$$

This equation shows the line has a slope of -3.3718. The closer the slope is to -3.33, the closer the amplification efficiency is to 100%. Using the amplification efficiency calculation, this was calculated for β-actin as follows:

$$Efficiency\ (E) = 10^{\left(\frac{-1}{-3.3718}\right)} - 1 = 0.98$$

This equation shows the amplification efficiency was 0.98, or 98%. This value was required for the $\Delta\Delta C_T$ method to analyse real-time PCR data for this gene.

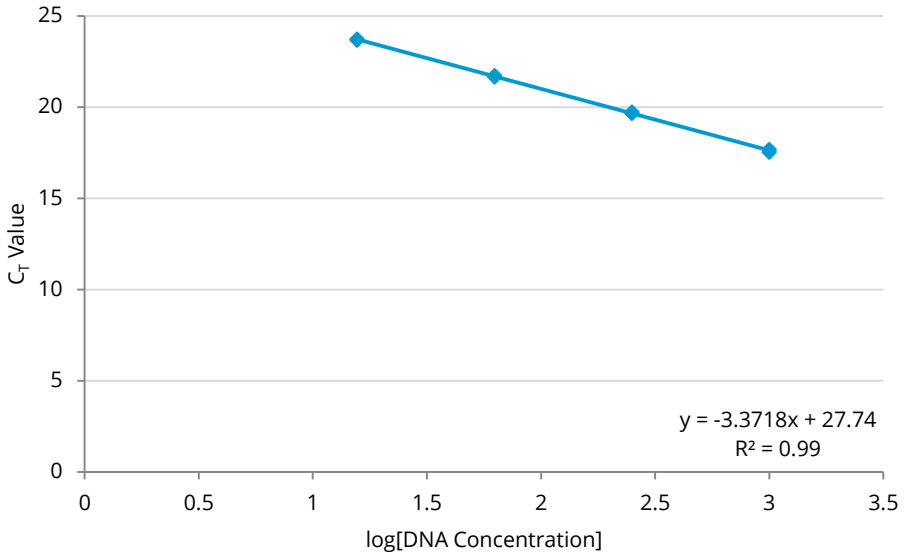


Figure 62. Standard curve plotting C_T values obtained using β-actin primers against the concentration of cDNA template. Equation of the line and R^2 are shown.

This process was carried out for each primer set used for real-time PCR: a second housekeeping gene (*Gapdh*) and the sex differentiation genes *Vasa*, *Amh* and *Cyp19a1a*. The efficiency plots for each primer set are shown below (Figures 63, 65, 67, 69), with the calculated amplification efficiencies listed in Table 18. The melt curves for all primer sets showed that only one product was being isolated (Figures 64, 66, 68, 70).

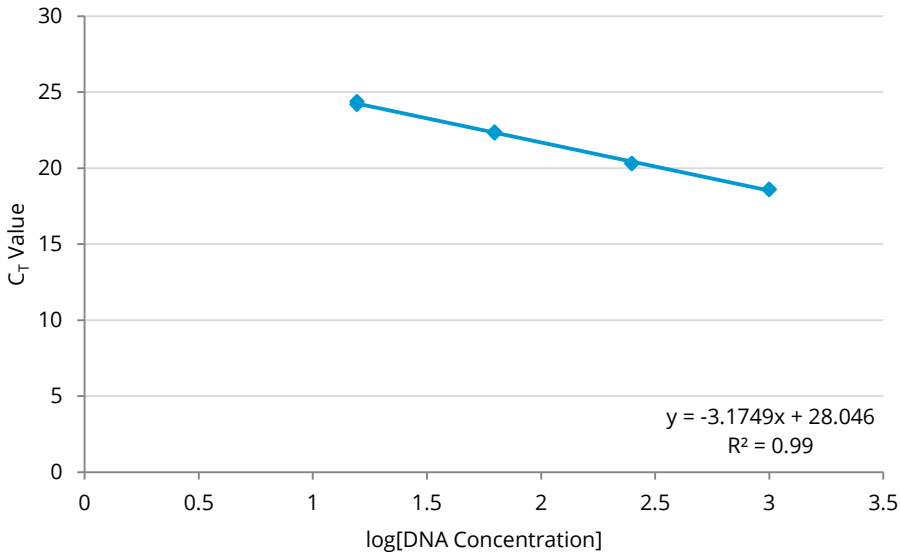


Figure 63. Standard curve plotting C_T values obtained using *Gapdh* primers against the concentration of cDNA template. Equation of the line and R^2 are shown.

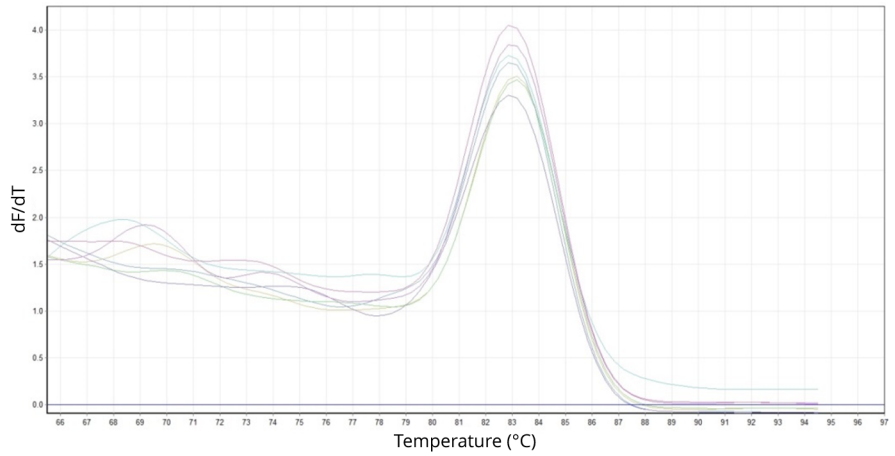


Figure 64. Melt curve obtained using the *Gapdh* primers, which indicated the generation of one distinct product with no non-specific amplification.

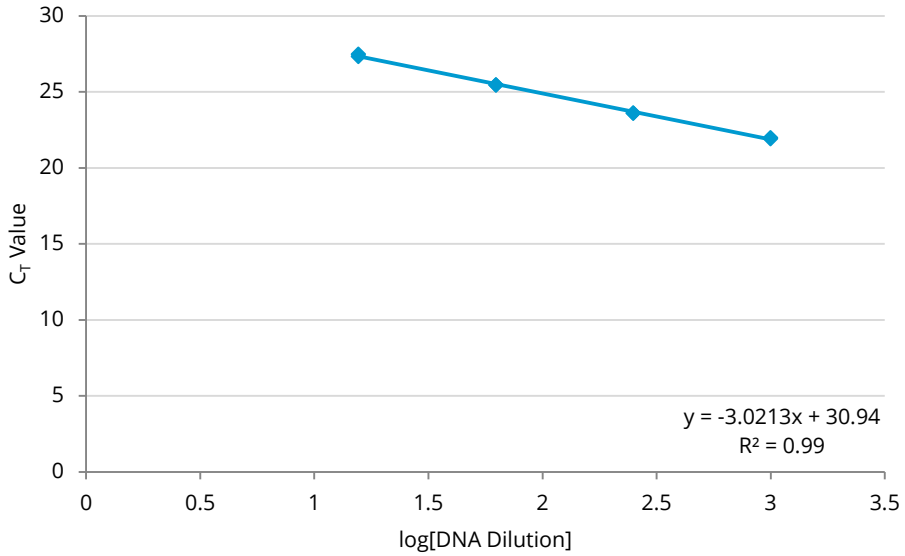


Figure 65. Standard curve plotting C_T values obtained using *Vasa* primers against the concentration of cDNA template. Equation of the line and R^2 are shown.

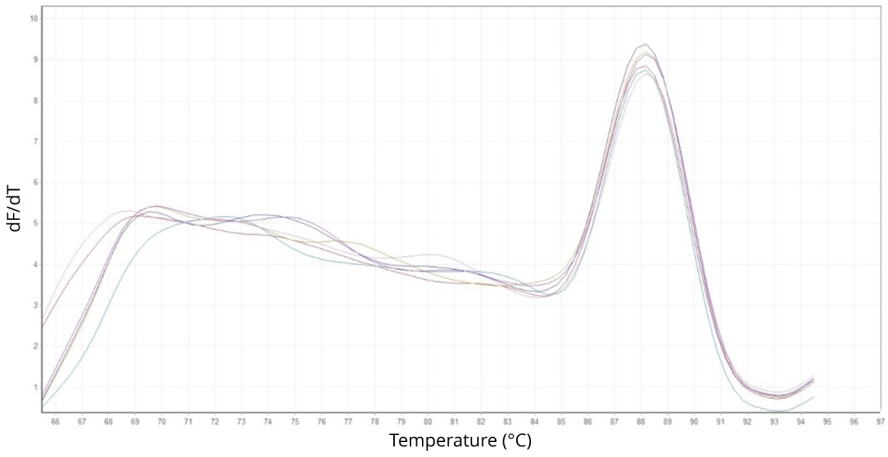


Figure 66. Melt curve obtained using the *Vasa* primers, which indicated the generation of one distinct product with no non-specific amplification.

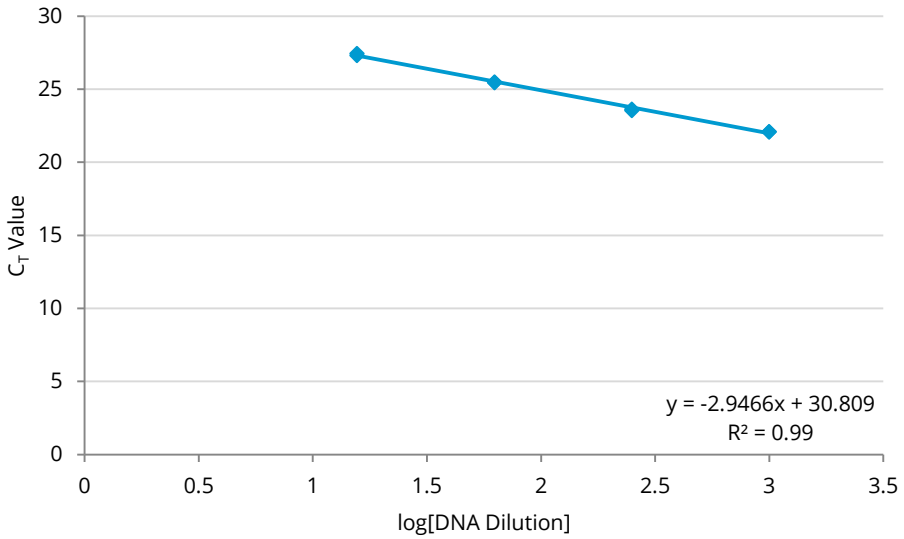


Figure 67. Standard curve plotting C_T values obtained using *Amh* primers against the concentration of cDNA template. Equation of the line and R^2 are shown.

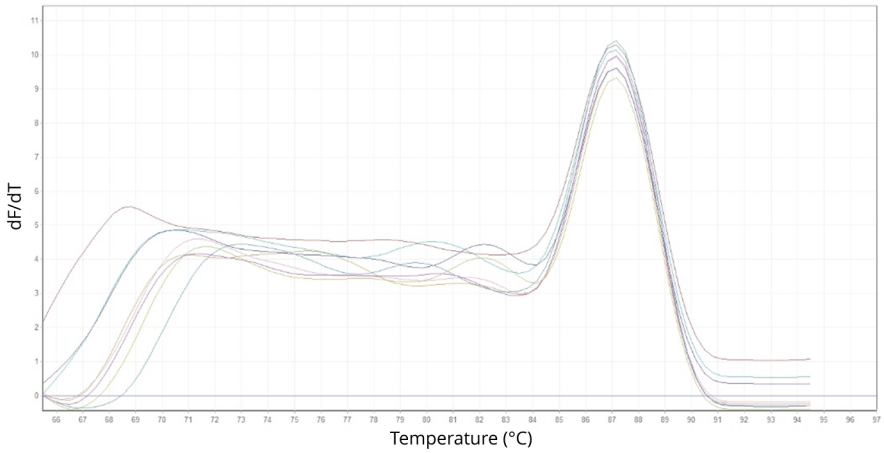


Figure 68. Melt curve obtained using the *Amh* primers, which indicated the generation of one distinct product with no non-specific amplification.

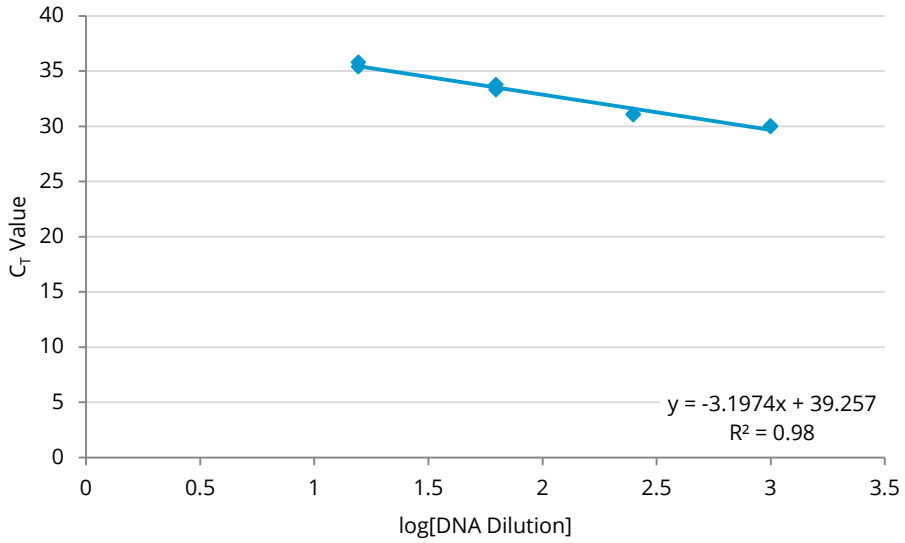


Figure 69. Standard curve plotting C_T values obtained using *Cyp19a1a* primers against the concentration of cDNA template. Equation of the line and R^2 are shown.

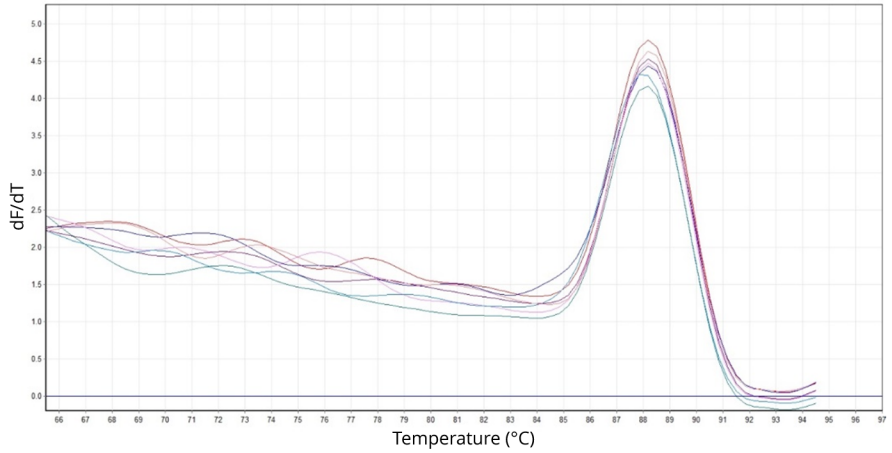


Figure 70. Melt curve obtained using the *Cyp19a1a* primers, which indicated the generation of one distinct product with no non-specific amplification.

Table 18. Amplification efficiencies of the real-time PCR primer sets used to target the housekeeping genes β-actin and *Gapdh*, and the sex differentiation genes *Vasa*. *Amh* and *Cyp19a1a*.

Gene target	Amplification efficiency (%)
Actb (β-actin)	98
Gapdh	106
Vasa	114
Amh	118
Cyp19a1a	105

3.6.2 Analysis of Real-Time Expression Data

As the amplification efficiencies of each primer set were close to 2, it was appropriate to use the $\Delta Cq = 2^{Cq(ref) - Cq(target)}$ method to evaluate gene expression. Each sex differentiation gene showed moderate expression at hatch, which decreased at 3 dph, increased at 12 dph and decreased again at 18 dph (Figure 71, 72). ANOVA analysis showed that only the *Cyp19a1a* data contained groups that were statistically different from each other. A closer examination at the groups using the Student's t-test showed it to be between expression at 3 dph and 18 dph. The ANOVA and Student's t-test statistics are detailed in Appendix V.

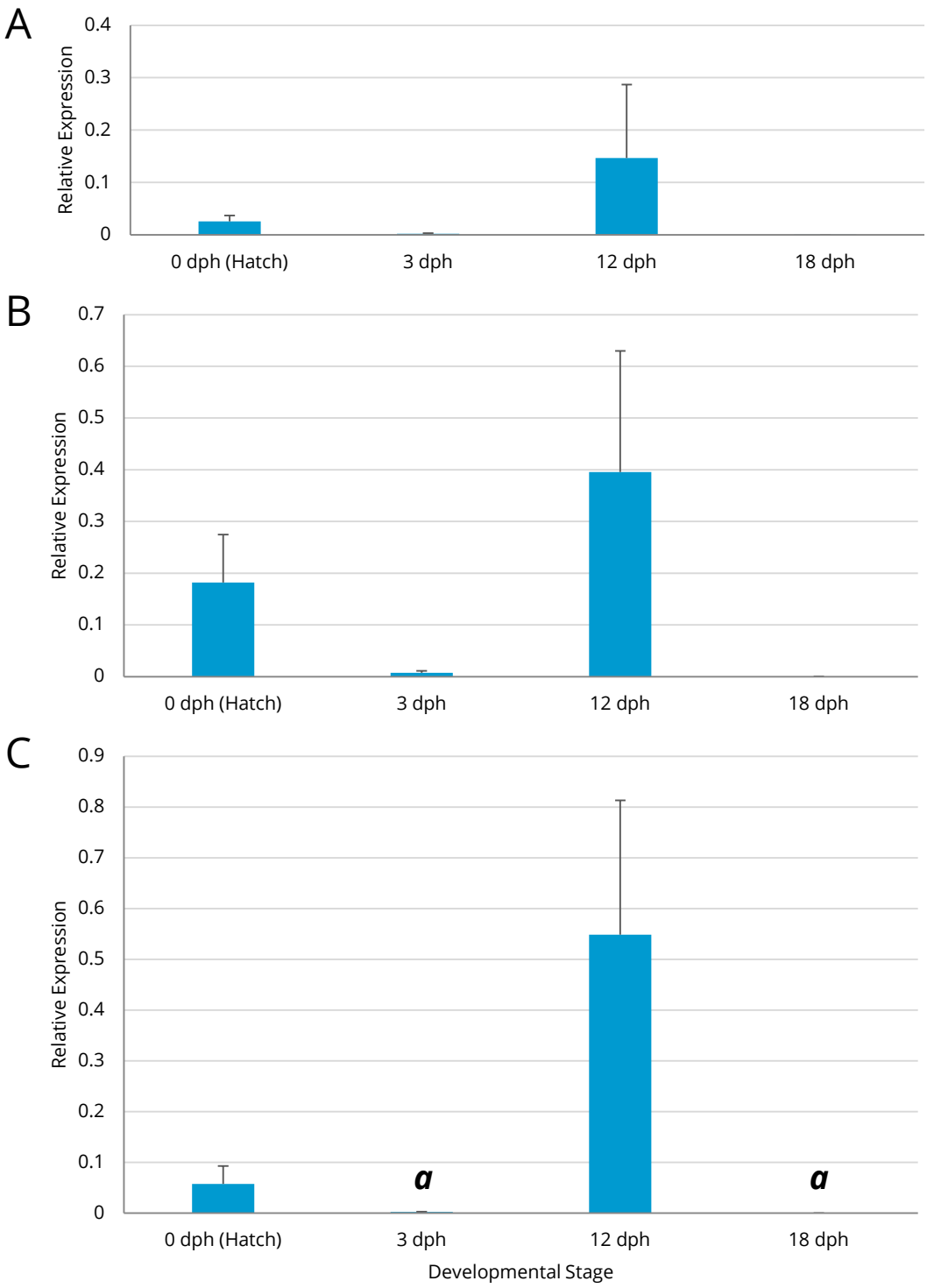


Figure 71. Relative expression of **A:** *Vasa,* **B:** *Amh* and **C:** *Cyp19a1a* within larvae at different stages of development. Data are means of four replicates (each containing 50 larvae) ± SE. Significantly different values ($p < 0.05$) are indicated by different letters above the bars.

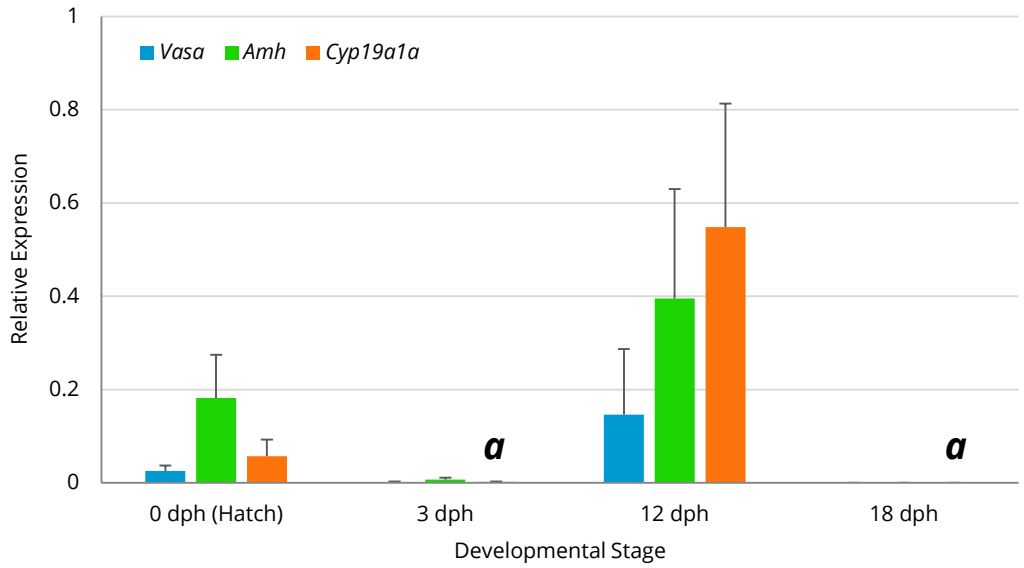


Figure 72. Comparison of the relative expression of *Vasa*, *Amh* and *Cyp19a1a* within larvae at different stages of development. Data are means of four replicates (each containing 50 larvae) ± SE. Significantly different values ($p < 0.05$) are indicated by different letters above the bars.

3.7 Histology

Sections of acceptable quality were obtained for larvae from 9 dph up to 60 dph, using a combination of paraffin and cryostat histology techniques. Identifiable or suspected organs and structures observed in these sections were labelled appropriately, referring to model sections as a guide (Genten *et al.*, 2009). A key detailing the abbreviations of organs and structures in this section is described below in Table 19.

The quality of paraffin sectioning improved sharply with the acquisition of fresh *p*-xylene for embedding preparation. However, due to time constraints, only four ages of samples (9, 15, 30 and 40 dph) were able to be processed and sectioned at this point. Prior to this, many samples became brittle during the embedding process and would crumble or tear during sectioning. The inner organs of some samples lost integrity during the embedding process and were unable to be sectioned. The

majority of paraffin-embedded samples could not be sectioned at lower than 6-8 μm thickness without tearing. Examples of poor sections are shown in Figure 73.

Cryostat embedding and sectioning was a significantly faster procedure than paraffin embedding and sectioning, but many sections were prone to curling, folding, shrinking and tearing (Figure 73). Nevertheless, reasonable quality sections were obtained for a small number of samples. The procedure was particularly effective for larger samples such as 60 dph larvae, which were oversized for the available paraffin embedding moulds and could not be processed using that method without being dissected. All samples sectioned on the cryostat could not be cut thinner than 10 μm without tearing. The resolution of these sections was poor at high magnification.

Table 19. Key to abbreviations used in all histology figures in this section.

AT	adipose tissue	M	buccal cavity (mouth)
An	anus	MD	mesonephric duct
B	brain	Me	epaxial muscle
E	eye	Mh	hypaxial muscle
Ei	iris	Ms	mesentery
El	lens	N	notochord
Er	retina	No	nostril
FF	finfold	Oe	oesophagus
Fp	pectoral fin	P	pancreatic tissue
Fv	pelvic fin	PW	peritoneal wall
Gi	gill	SB	swim bladder
H	heart	SC	spinal cord
I	intestine	St	stomach
Ip	posterior intestine (rectum)	To	tongue
Is	small intestine	U	urinary bladder
K	kidney	V	vertebra(e)
L	liver		

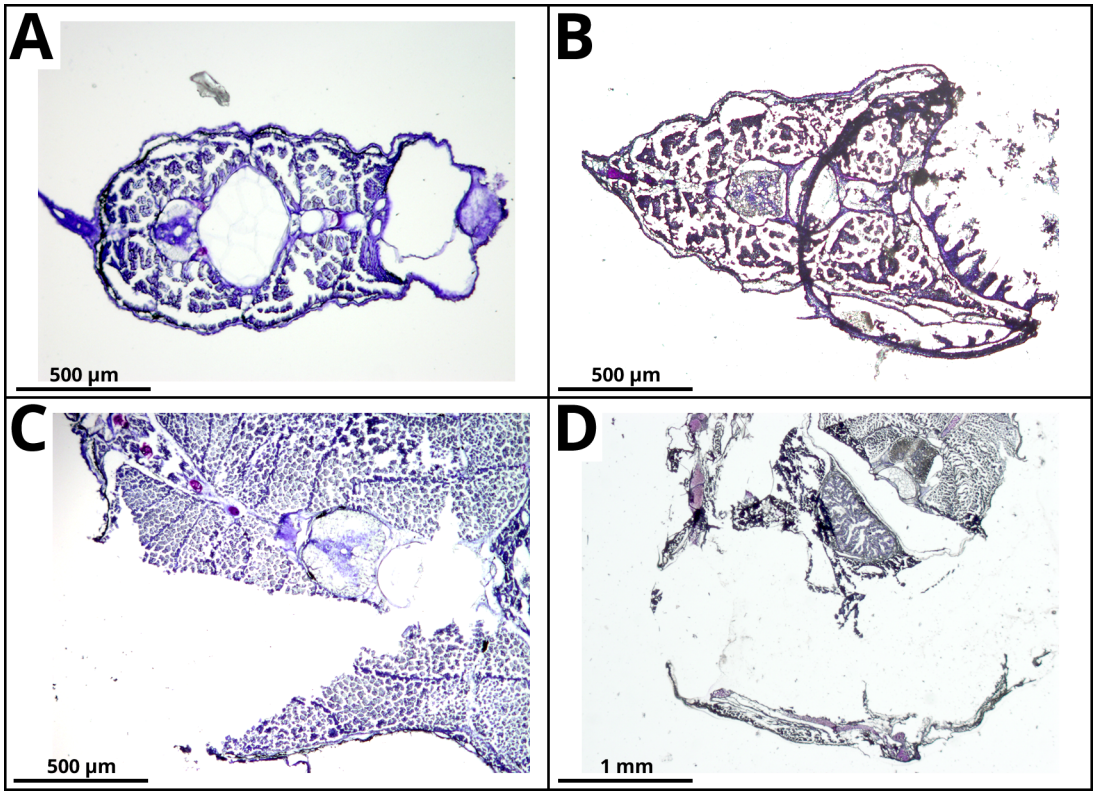


Figure 73. Photomicrographs of low quality 6 mm paraffin (A,C,D) and 10 mm cryostat (B) sections from kingfish larvae, stained using toluidine blue, captured at 25x (D) and 50x (A-C) zoom. **A:** Gut tissue missing from 15 dph sample. **B:** 25 dph section has folded, with tissue loss. **C:** 30 dph section torn through epaxial muscle. **D:** Poor integrity 35 dph sample, crumbled with significant tissue loss.

The youngest larvae successfully sectioned were 9 dph. A number of organs and structures were well defined at this age including the notochord and spinal cord, epaxial and hypaxial muscles, eyes, and rudimentary gut (Figure 74). Structures resembling kidneys (hematopoietic and excretory) and heart were present in these sections and were labelled accordingly. Unidentified but potentially gonad-related tissues were identified in Figure 75; specifically, the circled regions this figure show structures which may potentially be mesentery or presumptive genital ridge.

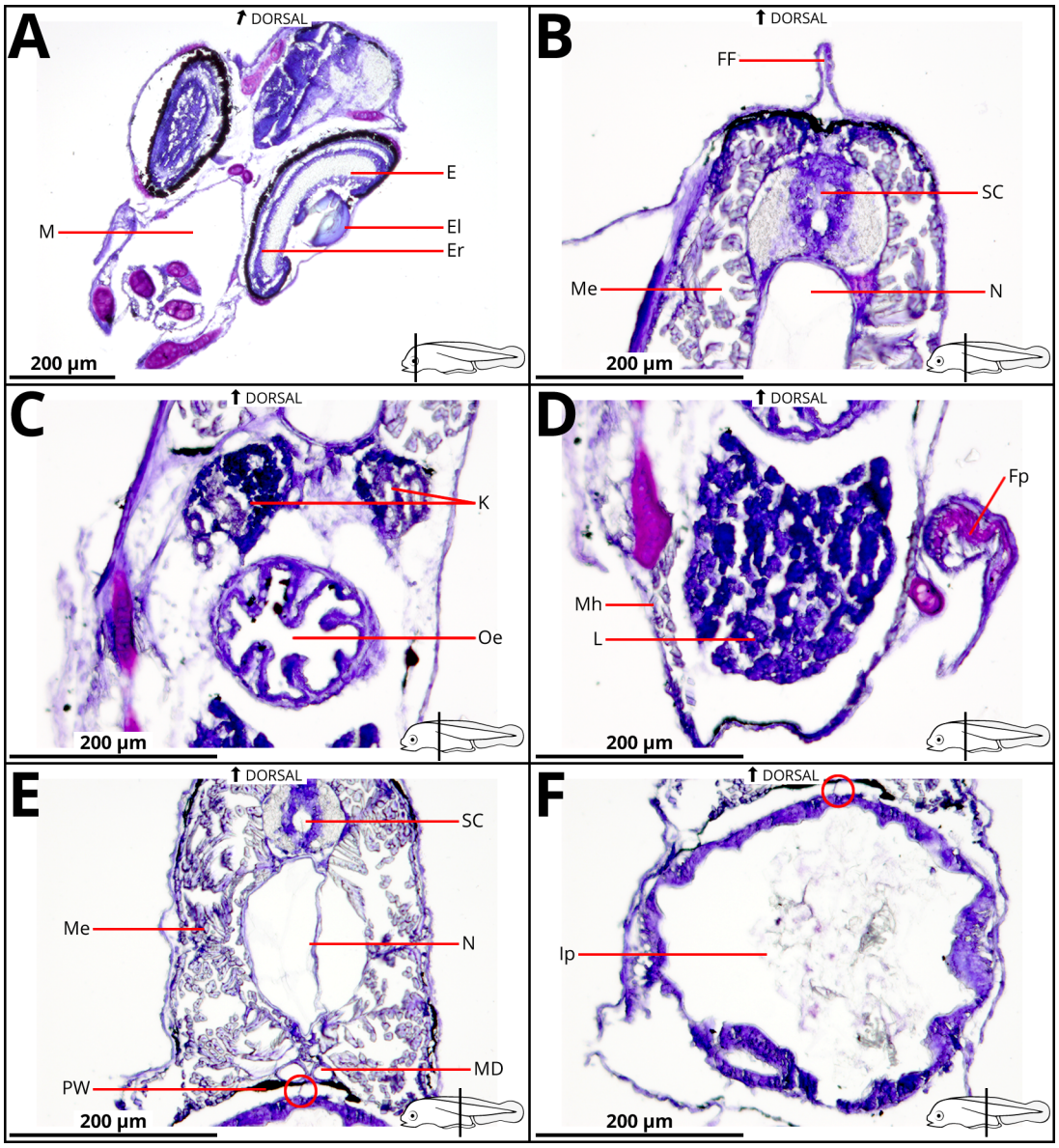


Figure 74. Photomicrographs of 6 mm paraffin sections from 9 dph kingfish larvae, stained using toluidine blue, captured at 100x (A) and 200x (B-F) zoom. **A:** Head. **B-D:** Dorsal, middle and ventral regions of the anterior trunk. **E-F:** Dorsal and ventral posterior trunk. The circled area shows an unidentified structure, which may be mesentery or presumptive genital ridge.

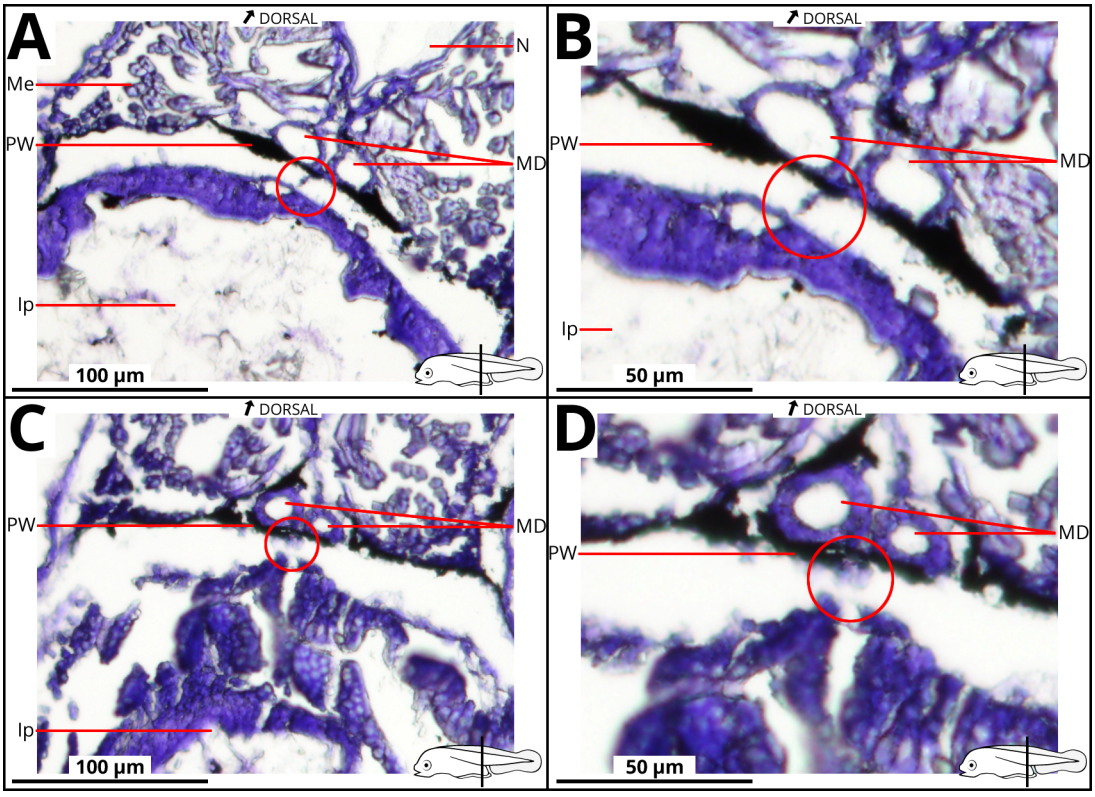


Figure 75. Photomicrographs of 6 mm sections from 9 dph kingfish larvae, stained using toluidine blue, captured at 200x zoom. **A:** Anterior trunk. Circled area shows potential mesentery or presumptive genital ridge. **B:** Image A digitally enlarged to 200% size. **C:** Anterior trunk. The circled area shows blurry unidentified region, potentially PGCs. **D:** Image C digitally enlarged to 200%.

At 15 dph, the gut appeared more complex with coiling visible, and a structure resembling liver is also present (Figure 76). However, many sections for this age exhibited tissue loss, particularly around the posterior gut. This is especially evident in images C and D of this figure. A structure suspected to be gall bladder was observed adjacent to the intestine and ventral to the swim bladder in these images. An area of interest was observed in images E and F, where a region of the peritoneal wall or mesentery appeared round. This may represent PGCs or the presumptive genital ridge.

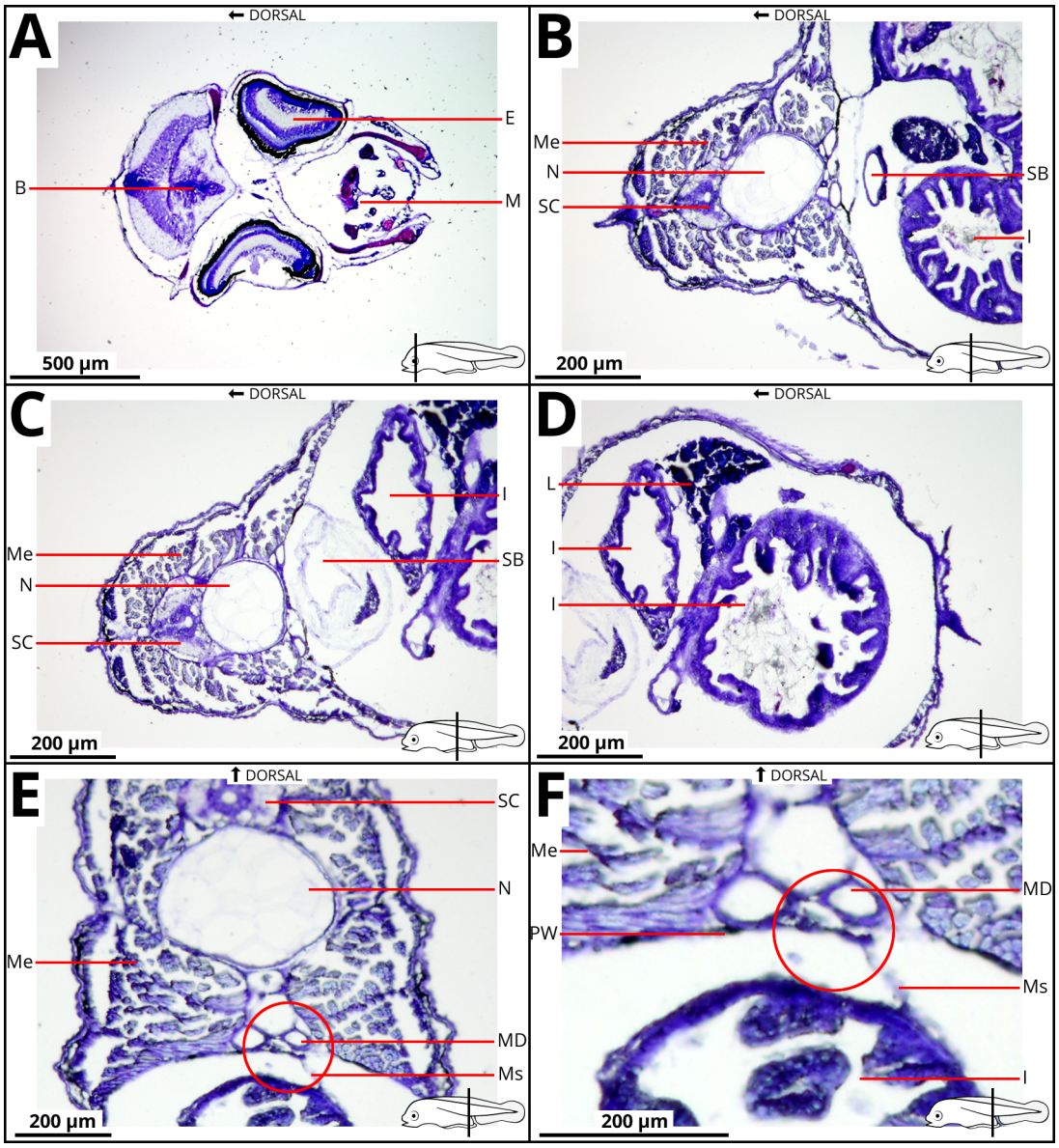


Figure 76. Photomicrographs of 6 mm paraffin sections from 15 dph kingfish larvae, stained using toluidine blue, at 50x (A) and 100x (B-F) zoom. **A:** Head. **B:** Dorsal anterior trunk. **C-D:** Mid trunk. **E:** Dorsal posterior trunk. The circled area indicates potential PGCs or presumptive genital ridge. **F:** Image E digitally enlarged to 200%.

Transverse sectioning was attempted for 21 dph larvae, but was unsuccessful during the time available. However, longitudinal sections were obtained for this age. A number of microphotographs were digitally overlapped to create merged whole-mount images (Figure 77). Although no presumptive gonad tissues or cells were identified, a comprehensive snapshot of general organ development was obtained.

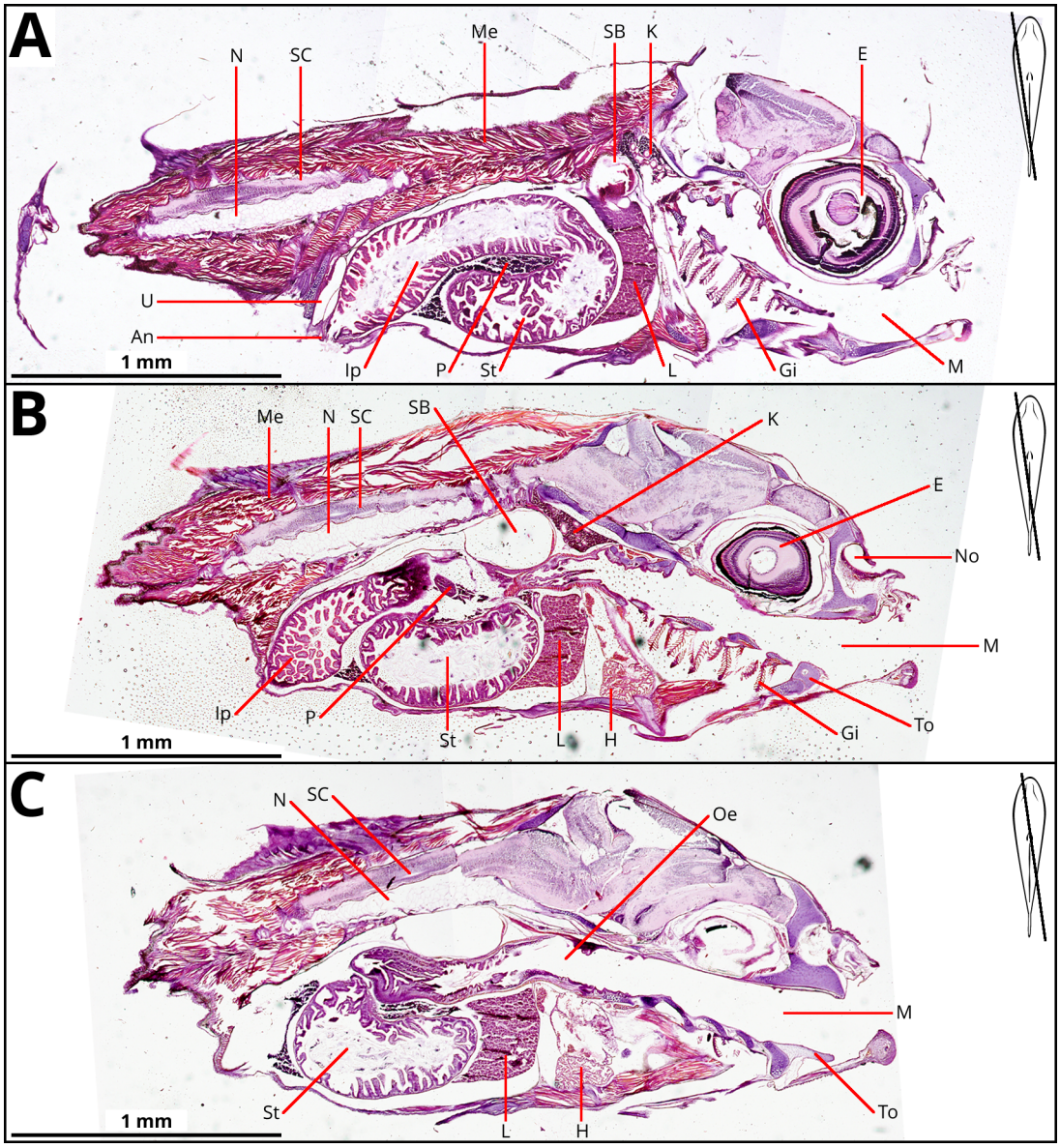


Figure 77. Composite photomicrographs of 6 mm paraffin longitudinal sections from 21 dph kingfish larvae, stained using hematoxylin and eosin, at 50x zoom. **A:** Section angled through eye to tail. **B:** Section angled from inner eye through spine to end of trunk on opposite lateral side. **C:** Section angled from behind eye, through spine to opposite lateral side behind anus.

At 25 dph, gut complexity continued to increase, with a number of other structures visible surrounding the intestinal loops (Figure 78). Tissue loss was a problem for early paraffin sectioning (images A and B of this figure). Due to time constraints, paraffin embedding was not revisited following the acquisition of fresh *p*-xylene. The cryostat sections were more complete but showed distortion, particularly towards the anterior trunk (image C of this figure). An unidentified structure of potential interest was circled and enlarged in imagine F.

At 30 dph, a region of unidentified cells was observed beneath the swim bladder and above the intestine (Figure 79). At 35 dph, large cells in were observed in this region among what appeared to be tightly packed somatic cells (Figure 80). This is an interesting landmark near where the presumptive gonad is thought to occur, as detailed for 40 dph below. Only a small number of good quality sections were obtained, with the majority tearing or crumbling (*e.g.* image D in Figure 73).

General organ development and positioning was well detailed at 40 dph, but no presumptive gonad tissues were positively detected. A number of photomicrographs were digitally overlapped to create merged whole-mount images (Figure 81). Potential sites of interest to gonad development are shown at greater magnification in Figure 82, notably adjacent to the peritoneal wall, swim bladder, and intestines.

Comprehensive whole-mount transverse sections were obtained by cryostat for 60 dph larvae. A number of photomicrographs were digitally overlapped to create a merged whole-mount image (Figure 83). The upper dorsal region (primarily epaxial muscle) was missing as an artifact of sectioning. No presumptive gonad tissues were able to be identified. The mid-trunk region between the swim bladder, liver and intestine showed an area of unidentified tissue, but it could not be ascertained whether it was presumptive gonad tissue. A number of dense tissues were also observed among the small intestines. These regions are shown at 100x magnification in images B and C of Figure 83.

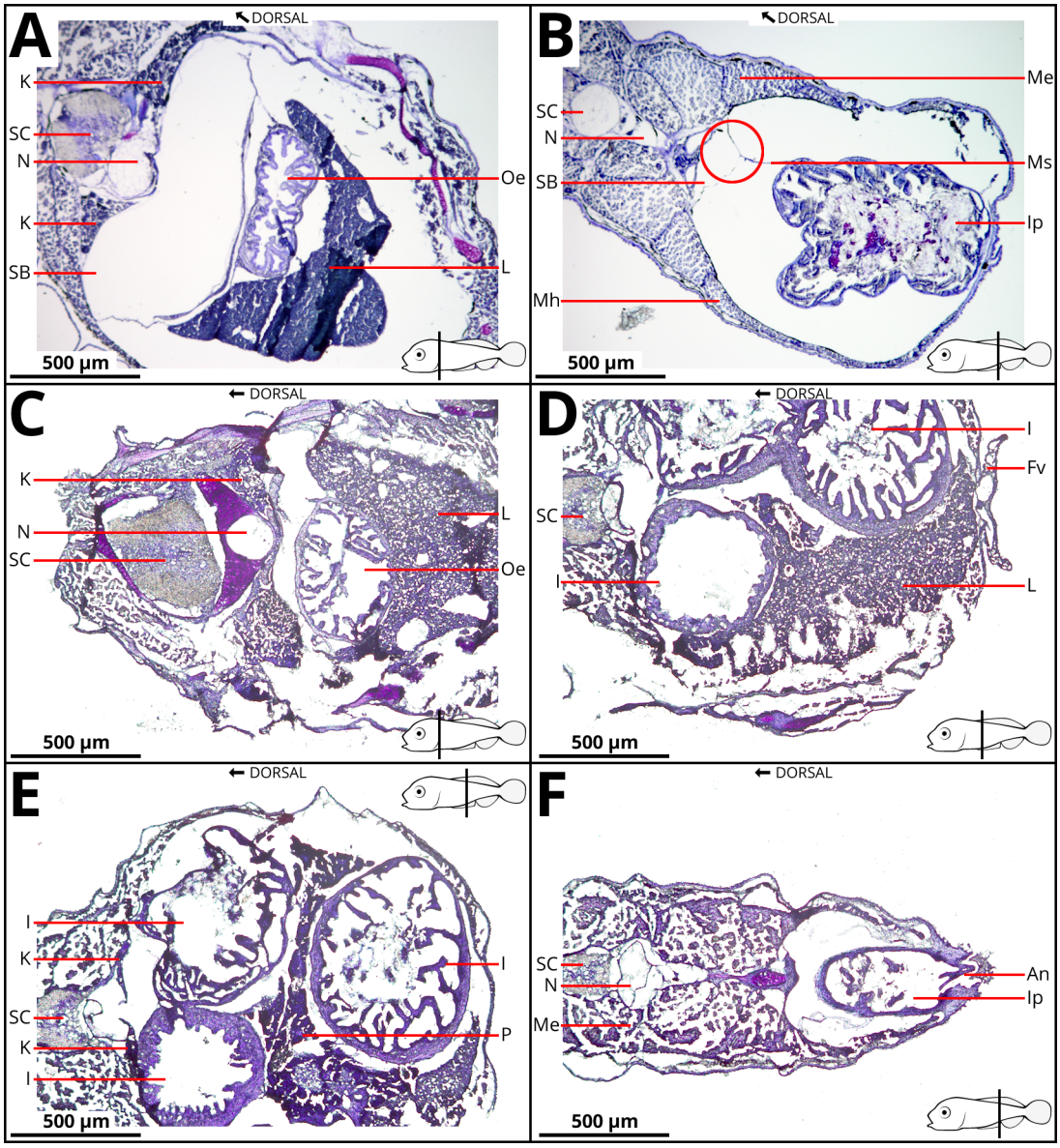


Figure 78. Photomicrographs of 6 mm paraffin (A-B) and 10 mm cryostat (C-F) sections from 25 dph kingfish larvae, stained using toluidine blue, at 50x zoom. **A:** Anterior trunk, with folding of heart tissue and loss of surrounding tissues. **B:** Posterior trunk, showing compression of gut tissue and loss of surrounding tissues. **C:** Dorsal anterior trunk. **D:** Ventral mid trunk. **E:** Ventral posterior trunk. **F:** Ventral posterior trunk. The circled region is shown digitally enlarged (inset).

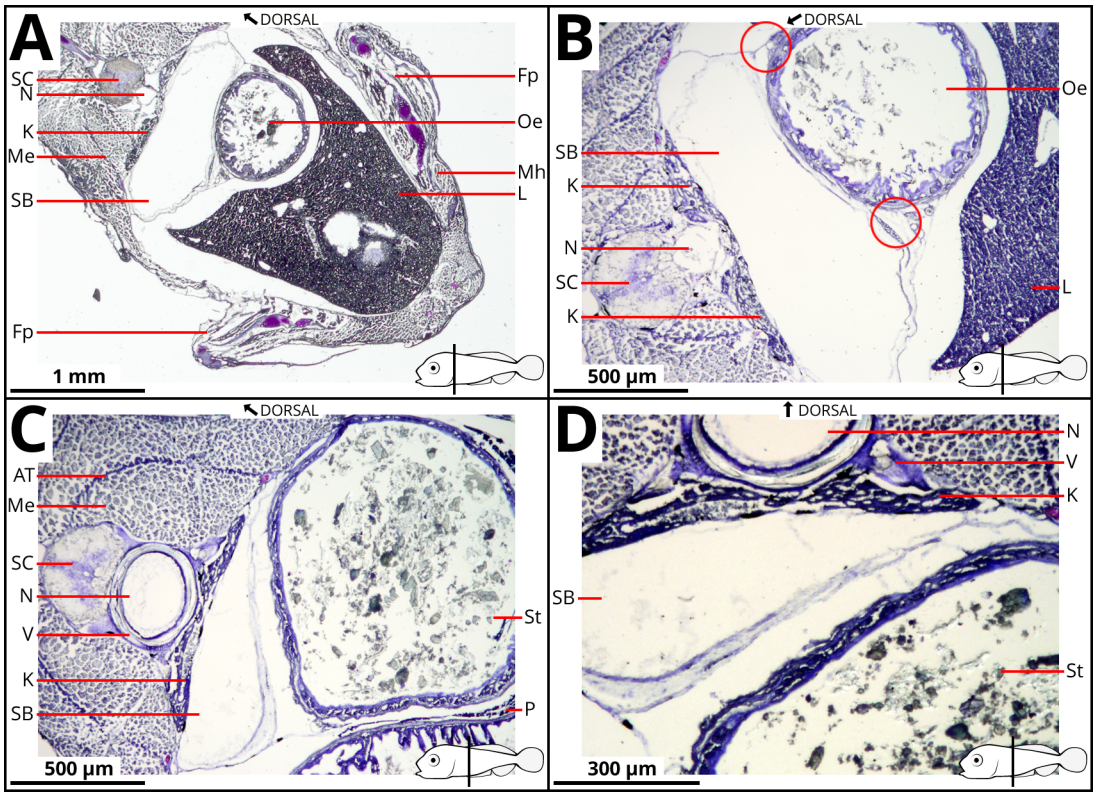


Figure 79. Photomicrographs of 6 mm paraffin sections from 30 dph kingfish larvae, stained using toluidine blue, at 25x (A) and 50x (B-D) zoom. **A:** Anterior trunk. **B:** Anterior trunk. A region of unidentified cells is circled. **C:** Mid trunk, showing notochord, kidneys and intestine. **D:** Image C, rotated and enlarged to 200% with increased contrast.

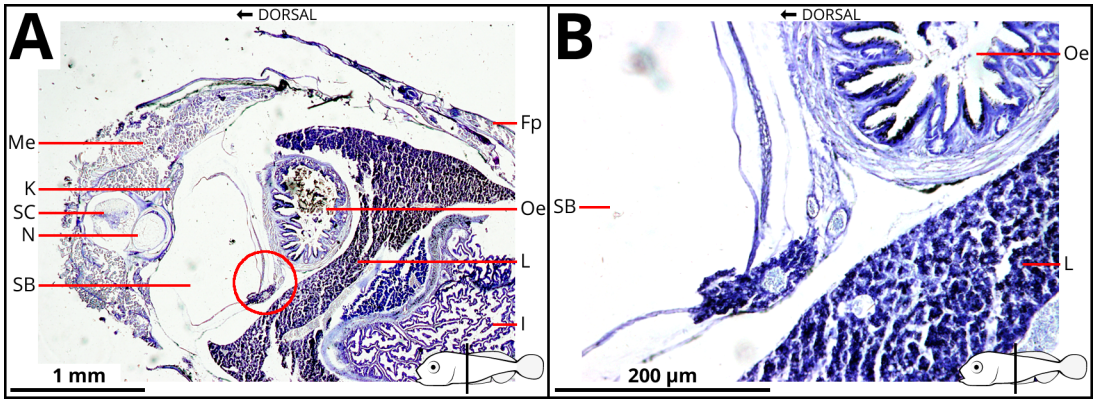


Figure 80. Photomicrographs of 6 mm paraffin sections from 35 dph kingfish larvae, stained using toluidine blue, at 25x (A) and 200x (B) zoom. **A:** Dorsal mid trunk, with unidentified region of large cells circled. **B:** Close up of suspected blood vessels in image A.

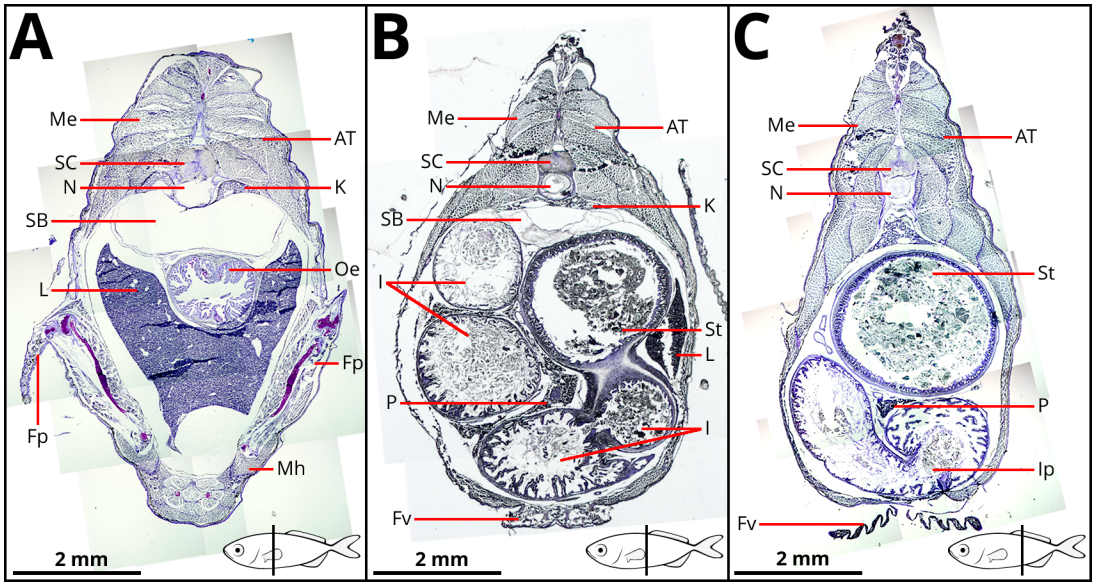


Figure 81. Composite photomicrographs of 6 mm paraffin sections from 40 dph kingfish larvae, stained using toluidine blue. Individual photomicrographs were captured at 25x (B) and 50x (A,C) zoom. **A:** Anterior trunk. **B:** Mid trunk. **C:** Posterior trunk.

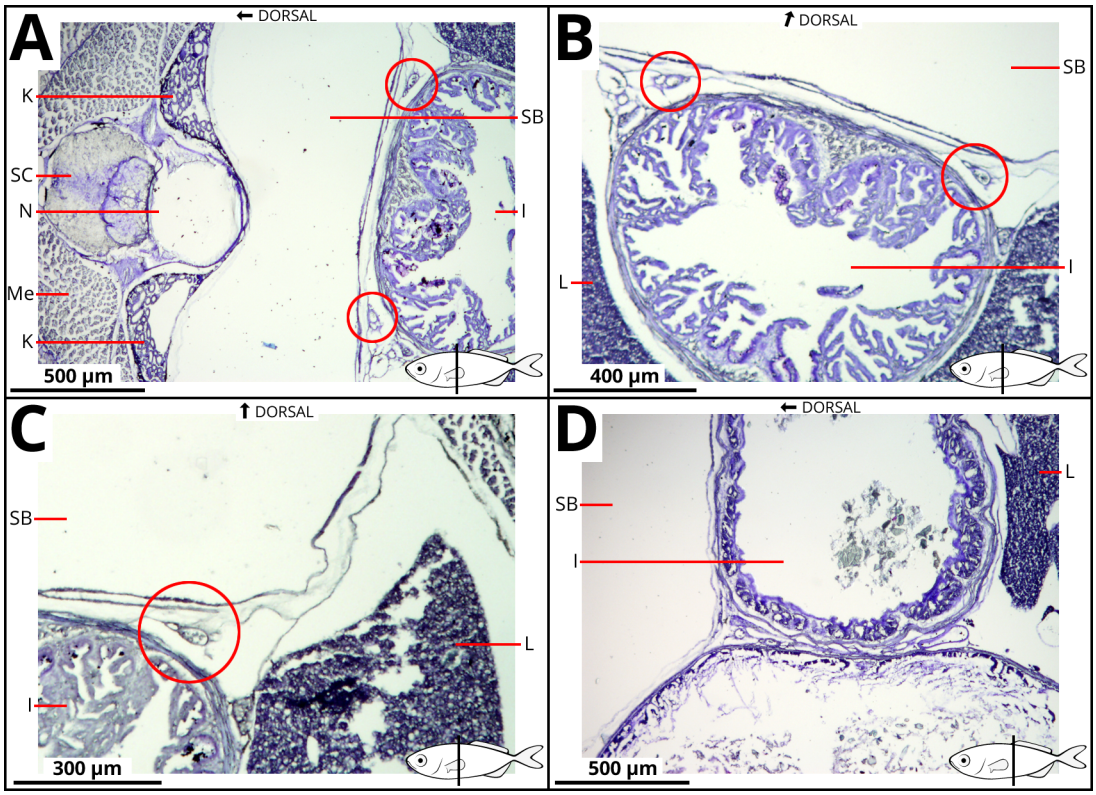


Figure 82. Photomicrographs of 6 mm paraffin sections from 40 dph kingfish larvae, stained with toluidine blue, captured at 50x zoom. Circled regions show sites of potential interest to presumptive gonad development. **A:** Dorsal anterior trunk. **B:** Ventral anterior trunk. **C:** Mid trunk, showing region between intestines and swim bladder. **D:** Posterior trunk, showing pancreatic region between intestines.

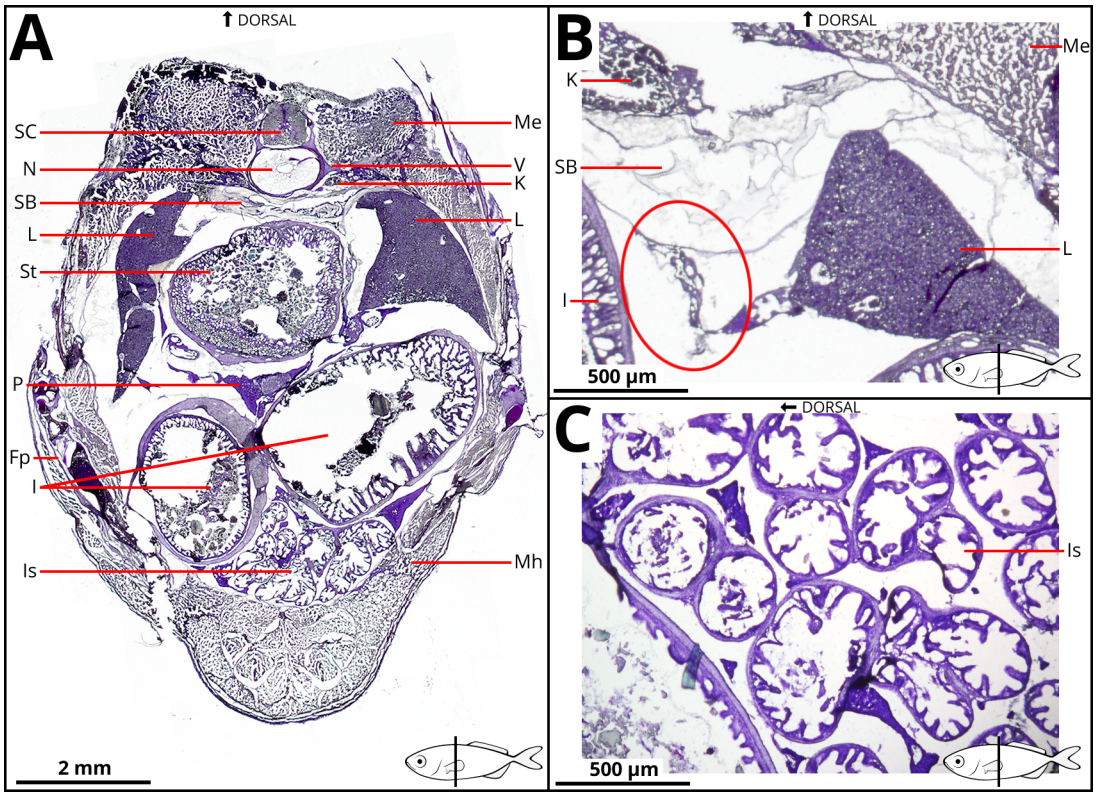


Figure 83. Photomicrographs of 10 mm cryostat sections from 60 dph kingfish larvae, stained with toluidine blue, captured at 25x (A-B) and 50x (C) zoom. **A:** Merged photomicrographs of mid trunk. The missing dorsal tissue in this image is an artifact of cryostat sectioning. **B:** Close up of mid trunk region near liver and intestine, with an unidentified region of interest circled. **C:** Close up of mid-trunk small intestine.

Chapter 4
Discussion

4.1 Rationale

Fully understanding the life cycle is of paramount importance for any aquaculture species. Currently, very little is known about sex differentiation within the yellowtail kingfish, which is a target species for NZ aquaculture and is already being farmed within Australia. This study sought to develop a number of genetic tools to help identify the window of sex differentiation in *S. lalandi*. Understanding this will impact farming production efficiency and give us a better understanding of this process within this species. The initial identification of key sex differentiation gene sequences in *S. lalandi* was a fundamental aim, which enabled the use of molecular techniques for characterising the expression profiles of these genes during larval development. Two approaches to monitor this were taken. The first approach involved the design of primers to targeted nucleotide sequences for use in real-time PCR to measure the expression of gene transcripts. The second approach aimed to develop histological techniques to obtain larval sections and relate expression profiles to the actual structural development of presumptive gonad tissue during larval development. The outcomes from this would help deepen the understanding of sex differentiation, develop informative molecular tools for future use and provide a broad information base for further work with this species.

The process of sex differentiation in fish is a complex process, which has been shown to occur at different times within every species investigated and involves the expression and function of a number of key genes (Devlin & Nagahama, 2002; Piferrer & Guiguen, 2008). Some of these genes have been characterised in fish, but very little has currently been described in *S. lalandi*. From the available literature, a

number of potential gene targets were identified as suitable markers to discover within the kingfish, to help with studies on their sex differentiation. The characterisation of each of these genes within *S. lalandi* was achievable due to the availability of an RNA-Seq transcriptome, which had been produced from the testes, ovary and pituitary tissues of wild-caught *S. lalandi*. Using sequences of previously characterised sex differentiation genes from other Perciformes, allowed the identification of partial nucleotide sequences in *S. lalandi* for a number of genes of interest. *Vasa*, *Amh*, and *Cyp19a1a* were three genes selected and identified in this study as primary candidates, due to their central role in sex differentiation and potential application as molecular markers. *Vasa* is integral to PGC migration and could be utilised for *in-situ* hybridisation to detect presumptive gonad tissues, while *Amh* and *Cyp19a1a* exhibit sex-specific expression in some species and may therefore assist in early detection of sex differentiation. A further three genes involved in sex differentiation were also explored in *S. lalandi*: *Dmrt1*, *Dax1* and *Sox9b*. However, these genes were not characterised further in this study due to time and resource constraints, but could be valuable for future work.

4.2 Transcriptomic Technologies

The '-omic' technologies, which include genomic, transcriptomic and proteomic approaches, adopt a holistic view of the molecules that make up a cell, tissue or organism. They are primarily aimed at the detection of genes (genomics), gene expression (transcriptomics) or protein expression (proteomics) in a specific biological sample in a non-targeted and non-biased manner. Their use has been limited due to the costs involved in the application of newly developed technology. However, with the development of cheaper approaches, recent studies have applied these approaches to aquaculture species for gene characterisation and expression studies (Yufera *et al.*, 2012; von Schalburg *et al.*, 2014; Fernández *et al.*, 2015). Complete genome sequences have been produced for numerous farmed fish,

including Atlantic salmon, Atlantic cod and European sea bass (W. S. Davidson *et al.*, 2010; Star *et al.*, 2011; Tine *et al.*, 2014). This has allowed investigations into gene discovery and how physiological processes have evolved within vertebrates (Martins *et al.*, 2013; Austbo *et al.*, 2014; Huang *et al.*, 2014). In addition, transcriptomics, which offers a snapshot of gene expression in a particular tissue or cell, has been utilised for studies of farmed fish. Studies have examined transcriptomic responses to functional feeds and bacterial infections in Atlantic salmon (Tacchi *et al.*, 2011a; Tacchi *et al.*, 2011b). In addition, in crucian carp, the transcriptome sequencing of brain, muscle, liver and kidneys has provided a valuable resource for studies into selective breeding and functional or evolutionary studies of genes involving hypoxia tolerance (Liao *et al.*, 2013). Within *Seriola* species, very little has been done at the genetic level and only a few recent investigations have taken advantage of some of these approaches. A partial genomic library (150 to 700 bp) has been constructed for isolating microsatellite markers in *S. lalandi* and *S. quinqueradiata* (Ohara *et al.*, 2005) and a limited transcriptomic approach has been used to profile the expression of immune genes in *S. quinqueradiata* kidney tissues (Darawiroj *et al.*, 2008). However, in many other aquaculture species, the application of these molecular approaches within studies are increasing (Qian *et al.*, 2014) and it is only a matter of time before they are routinely used within kingfish research.

4.3 Gene Discovery and Characterisation

This study utilised a previously constructed RNA-Seq transcriptome library produced from testes, ovary and pituitary tissues of wild-caught adult *S. lalandi*. As key components in the hypothalamic-pituitary-gonadal (HPG) axis, it was expected that transcripts for the sex differentiation genes of interest would reliably be expressed within these tissues and would therefore be detectable in the RNA-Seq library. BLAST searching the RNA-Seq library using previously characterised nucleotide sequences from other Perciformes led to the detection of transcripts for a

number of important genes known to be involved in sex differentiation. From these, *Vasa*, *Amh* and *Cyp19a1a* were selected to be further characterised in this study. PCR was initially used to confirm the cDNA sequence of these transcripts, due the presence of errors that are accumulated in sequences generated within the RNA-Seq library. As the entire coding DNA sequence (CDS) could not be determined from the RNA-Seq library for any of the genes investigated, this necessitated the use of RACE PCR in an attempt to obtain the 3′ and 5′ ends of the sequences. Due to time constraints, 3′ RACE PCR was only successfully accomplished for *Amh*, and 5′ RACE PCR was not pursued for any genes.

In the amplification of gene sequences, there were a number of approaches that were employed to ensure the correct sequence was obtained. When designing primers, high quality regions of consensus transcript sequences from the RNA-Seq library (*i.e.* areas with a large number of overlapping sequences) were targeted in order to minimise the risk of using areas which contained sequencing errors. Furthermore, several different forward and reverse primers were designed to allow a nested PCR approach, helping to reduce the chance of non-specific product amplification. This was particularly beneficial when RACE PCR was used, since each round only utilised a single gene-specific primer together with a generic adapter primer, and typically generated non-specific products in the first round which were then excluded in the second round. In addition, it was very uncommon to achieve the amplification of one specific band when carrying out PCR for the first time and each PCR required substantial optimisation. To help increase primer binding specificity, the annealing temperature was increased in small increments and extension times were adjusted according to the expected product size, if known. If time had permitted, there was much more that could have been attempted in order to obtain the full sequences of each gene, but what was obtained was enough to confirm the identity of the gene and use in expression analysis.

None of the problems encountered with amplification appeared to be due to the quality of the cDNA produced from the tissues used in this investigation. For RNA

extraction, there was a changeover of the method used, with TRIzol® reagent being replaced by the R&A-Blue™ reagent. However, the quality and yield of RNA was equally satisfactory for each method. In addition, two different methods were also used for cDNA synthesis: the qScript™ Flex cDNA Synthesis Kit and the Tetro cDNA Synthesis Kit, but when the housekeeping gene *Actb* (β-actin) was utilised to validate cDNA quality, a strong band was seen in each case. This gene is part of the actin family of proteins which are highly conserved and found in all eukaryotes, being the most abundant intracellular protein in eukaryotic cells (Gibbs *et al.*, 2003; Artman *et al.*, 2014). Owing to its central role in cell motility, the β-actin isoform is expressed in most cell types and was an ideal positive control for this application. The β-actin primers used in this study were actually specific to *S. lalandi* and worked reliably for initial PCRs as a positive control when testing cDNA.

In addition to problems with amplification, issues were also encountered with the transformation and cloning steps, before the amplified products were sequenced. In many cases, after plating transformed cells onto LB agar plates spread with IPTG and Xgal, colonies would grow very slowly over several days with unreliable results. Typically very few colonies grew, but occasionally a great number of very small colonies would cover the plate. This could have been due to an unforeseen issue with the chemically competent cells that were not being efficiently transformed by the plasmid containing the PCR product. It was also possible that some components of the ligation reactions may have interfered with transformation efficiency; for this reason, an extra step was included before transformation, which heated the ligation reaction to inactivate the enzyme. Eventually, when a reasonable number of colonies grew adequately, blue-white colony screening was generally reliable with a small number of blue colonies interspersed among a majority of white colonies. Generally, these white colonies contained the correct sized ligated product when screened by PCR. There were also instances of colonies which appeared to contain the desired plasmid insert when screened by PCR, but were not able to be grown in LB media. It was unclear what may have caused this. One possible explanation could have been

the presence of bacteriophages, but this would need to be validated. Regardless of each of these hurdles, sequences were cloned and sequenced for some of the targeted genes and the results were analysed, along with any predicted sequence obtained from the transcriptome. Primers were then designed for analysing their expression in larvae at different stages.

4.3.1 Characterisation of *Vasa*

This study analysed a 1670 bp fragment of the *Vasa* nucleotide sequence in *S. lalandi*, of which 900 bp was confirmed by sequencing. Conserved regions in the predicted amino acid sequence were searched for using the BLAST search tool and InterProScan. The sequence was shown to contain the helicase superfamily C-terminal and DEAD-box helicase domains characteristic of VASA (Fujiwara *et al*., 1994; Blazquez *et al*., 2011; Presslauer *et al*., 2012). The DEAD-box helicases are a diverse family of proteins important for ATP-dependent RNA unwinding, with roles in RNA degradation, ribosome biogenesis, splicing, nucleocytoplasmic transport and organellar gene expression. The helicase C-terminal region is found in proteins belonging to the helicase superfamilies SF1 and SF2, and is closely linked with DEAD/DEAH RNA helicases (Caruthers *et al*., 2000). The presence of these regions confirmed the identity of the cDNA sequence cloned from *S. lalandi*. After performing an alignment with other known vertebrate VASA sequences, the phylogenetic tree revealed that the *S. lalandi* amino acid sequence showed greatest similarity to other Perciformes, grouping strongly with the closely related species *S. quinqueradiata* and also to *T. japonicus*. The tree conformed to what had been shown previously when *Vasa* was analysed across different taxonomic groups (Blazquez *et al*., 2011), that mammals, Perciformes and Salmoniformes formed distinctly separate branches. If they had been included, additional sequences from other vertebrate classes (*e.g.* Aves, Reptilia, Chondrichthyes) would have improved the definition of these groups. However, it was very clear that the gene identified within *S. lalandi* was a homologue of *Vasa*.

It is well documented that *Vasa* is expressed strongly in germ cells of both sexes in a wide range of species (Fujiwara *et al.*, 1994; Tohru Kobayashi *et al.*, 2000; Devlin & Nagahama, 2002; Huang *et al.*, 2014; Yuan *et al.*, 2014), but may also be expressed weakly in somatic tissues. Somatic cell expression of *Vasa* has been observed in tissues such as the kidney and muscle of frog (*Rana nigromaculata*) (Jia *et al.*, 2009), and the heart and brain of *O. mykiss* (Yoshizaki *et al*., 2000). In European sea bass, Blazquez *et al.* (2011) found *Vasa* was highly expressed in ovary and testes, and weakly expressed in some somatic tissues. It was also noted that females of this species exhibited higher *Vasa* expression levels as early as 100 dph, with a stimulatory effect of estrogen and highly sexually dimorphic growth suggested as possible contributing factors. In this study, *Vasa* was found by PCR to be expressed strongly in testes, but was virtually undetectable in ovary. The cause of this is currently unknown, but it is speculated that an error during RNA extraction or cDNA synthesis may have contributed to this. To overcome this, a solution could be to prepare cDNA from three males and three females, and repeat the PCRs. For this study however, this discrepancy was inconsequential as expression in one tissue type was sufficient for cloning and sequencing.

The elucidation of the 3′ end of the *Vasa* nucleotide sequence was not able to be achieved in the time available for this study. Many of the problems encountered with 3′ RACE PCR may have been a consequence of primer design. It was challenging to accurately predict the size of RACE PCR products because the nucleotide sequence was unknown. However, using the alignment of the transcriptome sequence with known Perciformes sequences, it is likely that the 3′ RACE PCR product of the slVASA-3′F1/F2 primers would have been in excess of 1500 bp. The risk of non-specific binding or secondary structure formation increases with product length, which may have been a contributing factor to the problems encountered.

4.3.2 Characterisation of *Amh*

This study analysed a 1275 bp fragment of the *Amh* nucleotide sequence in *S. lalandi*, of which 1023 bp was confirmed by PCR, cloning and sequencing. The elucidation of the complete 3′ end of the *Amh* nucleotide sequence by 3′ RACE PCR, including the UTR and poly(A) tail, was successful. The products amplified using the slAMH-3′F1/F2 primers were ~685 bp. This is in contrast to the unsuccessful >1500 bp *Vasa* 3′ RACE PCR products, further suggesting that a smaller target sequence length is more suitable for this technique. Conserved regions in the predicted amino acid sequence obtained from this confirmed nucleotide sequence were searched for using the BLAST search tool and InterProScan. The sequence was shown to contain the distinctive Anti-Müllerian hormone N terminal region (AMH_N) and a C terminal transforming growth factor β (TGFβ) domain, both of which are characteristic of AMH (Imbeaud *et al.*, 1994; Halm *et al.*, 2007; Kumar *et al.*, 2010) and confirm the identity of the cDNA sequence cloned from *S. lalandi*. *Amh* is a member of the transforming growth factor-β superfamily of growth and differentiation factors (Josso *et al.*, 1993). Cleavage of the N-terminal, results in release of the TGFβ_C domain, which is the active part of the Amh protein (Wilson *et al.*, 1993).

After performing an alignment with other known vertebrate AMH sequences, the phylogenetic tree revealed that the *S. lalandi* amino acid sequence showed greatest similarity to other Perciformes. Mammals formed a distinct cluster, grouping closer to reptiles and birds than to fish. The Asian swamp eel *Monopterus alba* occurred in the cluster of Perciformes, despite belonging to the order Synbranchiformes. It is likely that this is due to the absence of other members of this order in the phylogenetic tree. If they had been included, additional sequences from other vertebrate classes would have improved the definition of these groups. However, it was very clear that the gene identified within *S. lalandi* was a homologue of *Amh*.

Amh is strongly associated with male gonadal tissues. In mammals, expression of *Amh* causes regression of female Müllerian ducts during early development. *Amh* expression persists after male reproductive development, but commences again in females in ovarian granulosa cells during gestation; the same pattern is also observed in birds and reptiles (Halm et al., 2007). In fish, the exact of mode action of *Amh* varies between species. In juvenile and adult sablefish, expression of *Amh* was significantly higher in testes relative to ovaries (Smith *et al.*, 2013). Similarly in adult Japanese flounder, expression of *Amh* was exclusive to testes, with no signal detected from ovary or a range of somatic tissues (Yoshinaga *et al.*, 2004). However, some species do not show any sexual dimorphism of *Amh*, such as medaka (Piferrer & Guiguen, 2008). Adult zebrafish showed *Amh* expression in gonads both sexes, specifically in presumptive Sertoli cells of testes and granulosa cells in ovaries (Rodriguez-Mari *et al.*, 2005). The same pattern is observed in the Cyprinid *Squalis pyrenaicus* and the related hybrid complex *S. alburnoides* (Pala *et al.*, 2008). European sea bass express high levels of *Amh* in both juvenile testes and ovary, although lower in the latter, with expression also detected in somatic tissues (Halm et al., 2007). In this study, *Amh* was found by PCR to be expressed strongly in testes, but was virtually undetectable in ovary. This is in line with most of the species described above, and was suitable for cloning and sequencing.

4.3.3 Characterisation of *Cyp19a1a*

A 633 bp fragment of the *Cyp19a1a* nucleotide sequence from *S. lalandi* was found on the UniProt database. Conserved regions in the predicted amino acid sequence were searched for using the BLAST search tool and InterProScan. The sequence was confirmed to belong to the cytochrome P450 superfamily, reinforcing the identity of the sequence. Cytochrome P450 enzymes are found in all kingdoms of life with a variety of roles, including oxidation of steroids and fatty acids, hormone synthesis and vitamin metabolism. *Cyp19a1a* encodes the gonadal isoform of the aromatase

enzyme, responsible for catalysing the conversion of androgens to estrogens. Expression of both the brain and gonadal isoforms of P450 aromatase is well conserved among teleosts. In goldfish (*C. auratus*), zebrafish, and killifish (*Fundulus heteroclitus*), *Cyp19a1a* has been found to be expressed exclusively in ovary tissues, while *Cyp19a1b* was expressed highly in neural tissues and at a low level in ovary (Callard *et al.*, 2001; Tchoudakova *et al.*, 2001; Greytak *et al.*, 2005; Sawyer *et al.*, 2006).

After performing an alignment with other known vertebrate aromatase sequences, the phylogenetic tree revealed that the *S. lalandi* amino acid sequence showed greatest similarity to other Perciformes. Mammals and birds formed distinct clusters, grouping separately to all of the fish sequences. The *S. lalandi* sequence grouped closely with other Perciformes and marine teleosts. If they had been included, additional sequences from other vertebrate classes would have improved the definition of these groups. However, it was very clear that the *S. lalandi* aromatase sequence was a homologue of *Cyp19a1a*. As with *Vasa*, 3′ RACE PCR was unsuccessful at elucidating the 3′ end of the *Cyp19a1a* nucleotide sequence in the time available for this study. Again, this may have been a consequence of primer design. Using the alignment of the transcriptome sequence with known Perciformes sequences, it is likely that the 3′ RACE PCR product of the slCYP-3′F1/F2 primers would have been in excess of 1500 bp. The risk of non-specific binding or secondary structure formation increases with product length, which may have been a contributing factor to the problems encountered. However, it is worth noting that 3′ RACE PCR of testes cDNA did generate sequences showing homology to *ZP* in other Perciformes. *ZP* is a diverse domain typically occurring in glycoproteins, which are the major constituents of the egg envelope in vertebrates (specifically the zona pellucida in mammals). However, in most vertebrates, *ZP* glycoproteins are only expressed in developing oocytes and liver (Kanamori *et al.*, 2003). Because *ZP* was not expected to be expressed in testes and was not a target sex differentiation gene, the origin of these *ZP*-like sequences in the *S. lalandi* testes cDNA remains unknown.

4.4 Gene Expression during Larval Development

A number of aspects of the real-time gene expression analysis in this study could be improved on. Due to time and resource constraints of this study, only selected times during development were analysed by real-time PCR. Analysing the remaining time points and including the results would generate a better profile of gene expression during development. Only 50 larvae were pooled for each time point; increasing this to 100 would increase the volume of RNA available and improve the representation of both sexes in the pool. Using a number of other housekeeping genes to normalise the real-time data would be very beneficial, as it is currently unclear if the expression of the β-actin and *Gapdh* genes change in expression during development. Choosing the right housekeeping gene has been a recurring topic of debate (de Kok *et al.*, 2005; Caradec *et al.*, 2010), but as in other studies, there are ways to screen for variability in the expression of housekeeping genes, particularly if they are being used during development as was done in this investigation. A number of recent studies have attempted to determine the best reference genes to use in real-time analysis for developing fish (Fernandes *et al.*, 2008; Infante *et al.*, 2008; Zhong *et al.*, 2008; Mitter *et al.*, 2009; Øvergård *et al.*, 2010; Dang & Sun, 2011; De Santis *et al.*, 2011; C. G. Yang *et al.*, 2013; Liu *et al.*, 2014; Mo *et al.*, 2014). However, it appears from these results that no single set of reference genes are universal and that different fish species will require a different selection of housekeeping genes. Some of the studies go as far as to show that different housekeeping genes may be better candidates for different tissues within one species. However, there was not enough time available to do this during this study. The results presented are a preliminary look at a selection of the samples collected, to give some insight into what may be happening with gene expression of sex differentiation genes during development. It is recommended that more housekeeping genes are selected and used in the analysis of this data, to make sure the best possible targets are chosen.

4.4.1 Expression of *Vasa* during Larval Development

The real-time analysis of *Vasa* expression in *S. lalandi* did not reveal any statistically significant difference in expression between hatch, 3 dph, 12 dph and 18 dph. A notable increase in expression was observed at 12 dph, but it is likely that this expression pattern was a due to differential expression of the housekeeping genes β-actin and *Gapdh* during larval development, as the same pattern was also observed for *Amh* and *Cyp19a1a*. Future work will need to focus on the selection of suitable housekeeping genes and their expression during larval development. Increasing the number of larvae pooled per time point and performing real-time PCR on the remaining time points would be valuable to improving accuracy and completing the expression profile of *Vasa* during early development.

The expression of *Vasa* in developing larvae has been documented in a number of fish species. In most cases, expression of *Vasa* decreased from hatch before plateauing and increasing until gonadal sex differentiation. In European sea bass, *Vasa* was expressed strongly in fertilised eggs and at hatch, decreasing exponentially until 30 dpf. *Vasa* expression then increased again exponentially until 150 dpf, at which point sex differentiation could be detected histologically (Blazquez *et al.*, 2011). In half-smooth tongue sole (*Cynoglossus semilaevis*), *Vasa* expression followed a similar pattern but with some sexual dimorphism. Expression in both sexes gradually decreased from hatching, reaching the lowest level at 16 dph in females and 46 dph in males, before continuously increasing in both sexes until 150 dph (Huang *et al.*, 2014). In Nile tilapia, two isoforms of *Vasa* are expressed during larval development: *Vas* and *Vas-s*, the latter of which misses part of the *Vas* N-terminal region. Both isoforms were expressed in embryos and newly hatched larvae. By 5 dph, only *Vas-s* continued to be expressed, coinciding with the localisation of PGCs to the site of the presumptive gonad. This regime continued until gonadal sex differentiation at 30 dph, when *Vas-s* expression decreased and *Vas* expression increased (Tohru Kobayashi *et al.*, 2002).

4.4.2 Expression of *Amh* during Larval Development

The real-time analysis of *Amh* expression in *S. lalandi* did not reveal any statistically significant difference in expression between hatch, 3 dph, 12 dph and 18 dph. As described for *Vasa* previously, the observed expression pattern was possibly due to differential expression of the housekeeping genes β-actin and *Gapdh* during larval development. Future work will need to focus on the selection of suitable housekeeping genes and their expression during larval development. Increasing the number of larvae pooled per time point and performing real-time PCR on the remaining time points would be valuable to improving accuracy and completing the expression profile of *Amh* during early development.

The expression of *Amh* in developing larvae has been documented in a number of fish species. *Amh* has shown diversification during its evolution; in the majority of cases, expression of *Amh* was exclusive to testes and undifferentiated gonadal tissue. In zebrafish, *Amh* expression first occurred while gonads were undifferentiated, but became sexually dimorphic by 31 dpf with expression occurring in testes but not ovaries (Rodriguez-Mari *et al.*, 2005). In Nile tilapia, *Amh* was expressed at low levels in both sexes prior to differentiation, but an increase occurred in testes from 15 dph (Siegfried, 2010).

4.4.3 Expression of *Cyp19a1a* during Larval Development

A statistically significant difference in *Cyp19a1a* expression in *S. lalandi* was revealed by real-time analysis between 3 dph and 18 dph. However, these were the lowest points of expression observed for this gene. As with *Vasa* and *Amh*, the observed expression pattern was possibly due to differential expression of the housekeeping genes β-actin and *Gapdh* during larval development. Future work will need to focus on the selection of suitable housekeeping genes and their expression during larval development. Increasing the number of larvae pooled per time point and performing

real-time PCR on the remaining time points would be valuable to improving accuracy and completing the expression profile of *Cyp19a1a* during early development. Once this has been achieved, a greater range of fish sizes and ages can be examined.

The expression of *Cyp19a1a* has been documented in a number of fish species, but specific details of expression profiles in developing larvae remain scarce. In zebrafish, *Cyp19a1a* expression was found to increase steadily from fertilisation, with strong expression in undifferentiated gonads at 17 dpf. Zebrafish gonads had differentiated into testes and ovaries by 31 dpf; at this point, *Cyp19a1a* expression was detectable only in granulosa cells surrounding oocytes, and was undetectable in testes (Rodriguez-Mari *et al.*, 2005). In Atlantic halibut and European sea bass, increased expression of *Cyp19a1a* has been measured well before histologically detectable ovarian differentiation occurs (Blazquez & Piferrer, 2004; Matsuoka *et al.*, 2006).

4.5 Larval Physiology and Development

One of the goals pursued in this study was to compare the expression profiles of key sex differentiation genes to the structural development of gonad tissues of *S. lalandi* during early development. This could then be related to gonadal development. This required the histological processing of tissues from fish of different ages and developmental stages, including relatively newly hatched larvae with a simple body plan through to settled juveniles which had initiated bone ossification and possessed a spectrum of tissue densities. Fish from 3 dph to 60 dph were sampled as there were no previous studies focusing on early gonad development in this species. Fernández *et al.* (2015) have since published a study on the migration of PGCs to the genital ridge, elucidating the early formation of the gonad in this species. While the main purpose for histological processing in the current study was to examine the

development and potential sexual differentiation of the gonad tissues, this was not achieved without considerable technical challenges.

One of the most significant issues encountered with paraffin embedding and sectioning was tearing or crumbling. In addition to practicing and improving proficiency at operating the microtome, this was extensively troubleshot by modifying sectioning technique, altering the alcohol and xylene wash protocol, and increasing the liquid paraffin infiltration time. However, sectioning immediately and dramatically improved in quality following the acquisition of a fresh stock of *p*-xylene, replacing the depleted mixed isomer drum-grade xylene. Although it was not possible to quantitatively assess the purity of the two solutions, it is suspected that the mixed isomer xylene was inefficient as a clearing agent. This would mean that some of the dehydrating ethanol would remain at the time of paraffin infiltration, thereby weakening the integrity of the tissue during the sectioning process and resulting in cracked, poor quality sections. One potential explanation for the difference between the xylene solutions could relate to age and possible hygroscopic contamination from moisture. Time constraints did not permit a great range of high quality sections to be produced following this change. Had further processing of different aged larvae been possible with these operational changes, it is likely to have yielded greater information regarding the development of the early gonad.

The resolution of the sections produced using the different processing techniques varied. OCT-embedded cryostat samples could not be sectioned thinner than 10 μm without tearing or crumbling. The minimum thickness of paraffin sections was 6 μm before similar issues were encountered. It is possible that 6-10 μm was too thick to provide the resolution required to identify PGCs, as thin sections typically provide greater resolution and clarity of cells. However, 6 μm paraffin sections have been successfully used to identify PGCs in larvae of yellowfin tuna (*Thunnus albacares*) (Toru Kobayashi *et al.*, 2015).

Plastic resin embedding is a technique which utilises a harder matrix than paraffin or OCT embedding. Using an ultramicrotome, sections can potentially be cut from resin thinner than 0.5 μm (Eastham & Essex, 1969) and are less prone to tissue shrinkage (Verhoef *et al.*, 2013), which was a problem encountered with the cryostat. Resin embedding has been used successfully with embryonic and larval fish such as sturgeon, grey mullet (*Mugil cephalus*) and grass goby (*Zosterisessor ophiocephalus*), with sections cut as fine as 3 μm (Corsi *et al.*, 2003; Saito *et al.*, 2014). A comparison between paraffin and resin embedding of mature albacore (*Thunnus alalunga*) ovaries determined that plastic embedding was superior, showing less distortion of 'free space' and connective tissue (Saber *et al.*, 2015). Plastic resin embedding was not pursued in this study primarily due to lack of equipment availability, but would likely be valuable to improving section quality and resolution.

Two biological staining procedures were used in this study: toluidine blue, and hematoxylin and eosin (H&E). Toluidine blue is a basic stain which stains nucleic acids blue and polysaccharides purple. It can also produce metachromasia which stain proteoglycans red. Hematoxylin is basic and stains cell nuclei blue, while the acidic eosin counterstains eosinophilic structures (*e.g.* cell cytoplasm, collagen, erythrocytes) in shades of red, pink and orange (Anjum, 2013). Toluidine blue was effective at staining larval sections and was an efficient procedure for quickly processing a large number of slides, making it an ideal choice for staining preliminary sections with uncertain quality. The H&E method was very effective and produced superior contrast of tissues to toluidine blue. However, the procedure was tedious for staining a large number of slides. For this reason, H&E was used exclusively for sections known to be of reasonable to high quality. Due to the ongoing issues with section quality throughout this study, only sections of 21 dph were stained using H&E. Time constraints did not permit H&E staining of the sections produced following the acquisition of *p*-xylene.

4.5.1 Primordial Germ Cells

The role and development of PGCs have been examined in a broad range of organisms, including fish such as the model species zebrafish and medaka, and farmed species such as Atlantic cod, European sea bass, longtooth grouper (*Epinephelus bruneus*) and sturgeon (*Acipenser* sp.) (Koprunner *et al.*, 2001; Li *et al.*, 2009; Blazquez *et al.*, 2011; Haugen *et al.*, 2011; Presslauer *et al.*, 2012; Sao *et al.*, 2012; Saito *et al.*, 2014; Psenicka *et al.*, 2015). The PGCs migrate to the genital ridge under the influence of cell signalling during embryological development (Devlin & Nagahama, 2002; Presslauer *et al.*, 2012). However, to date there have been only a few studies of PGC migration, proliferation and differentiation in *Seriola* species. In *S. quinqueradiata*, primordial gonad differentiation was examined to determine an optimal developmental stage for intraperitoneal spermatogonial transplantation, with a view to produce allogenic larvae for developing surrogate broodstock technology (Morita *et al.*, 2012). Fernández *et al.* (2015) recently characterised the migration of PGCs in *S. lalandi*, determining key time points during larval development. It is therefore timely that the present study attempts to characterise molecular markers to identify sexual differentiation, a key stage in the gonadal development following the migration and proliferation of PGCs. This also involves the identification of the presumptive gonad. Collectively this information is essential for characterising sex differentiation in *S. lalandi*.

PGCs often first occur embryonically or very early in development, before migrating to the site of the presumptive gonad (Devlin & Nagahama, 2002; Presslauer *et al.*, 2012). These PGCs are often visible using light microscopy and have been found in the dorsal mesentery in medaka, rockfish (*Sebastes schlegeli*) and *Ditrema temmincki*, and between the gut and mesonephric duct in tilapia and flounder (*P. olivaceus*) (Devlin & Nagahama, 2002). PGCs were observed to have settled at the site of the presumptive gonad in *S. lalandi* by 7-10 dph (Fernández *et al.*, 2015). In the present study, landmark structures such as the mesonephric ducts and peritoneal wall,

which occur near the expected site of the presumptive gonad, were clearly visible at 9 and 15 dph. At 9 dph, unidentified structures were observed adjacent to the peritoneal wall, immediately below the mesonephric ducts and above the intestine (Figure 75). It was impossible to determine whether this was simply connective mesentery or PGCs gathered to form the genital ridge. Interestingly, however, this tissue coincided with the location of the PGCs and presumptive genital ridge in 10 dph *S. lalandi* identified by Fernández *et al.* (2015). At 15 dph, connective mesentery was visible between the peritoneal wall and intestine (Figure 76: images E & F). Fernández *et al.* (2015) found that, by 15 dph, the PGCs were enclosed in somatic cells at the genital ridge, forming an early gonadal framework. A potentially analogous structure was visible in the present study near the posterior of the gut cavity (Figure 84). A small dark stained region was evident below the mesonephric ducts. This was intersected by a thread of mesentery running from the top of the intestine toward the peritoneal wall. As early gonadal development tends to be evident at the posterior of the gut cavity (Sao *et al.*, 2012), it is easiest to observe development in this region. However, the quality of these sections was not high enough to provide certainty.

Many of the sections for the older larvae and juvenile fish were from the mid gut region. The paired lobes of the gonads are likely to have spatial separation as the girth of the fish increases. At 40 dph, two structures were observed adjacent to the swim bladder, and in bilaterally symmetrical locations either side of the intestine (circled in Figure 82). Both structures had an oval shape with a smaller ring-like structure attached on the outer edge. The latter was speculated to be a possible blood vessel. Other studies have also observed the addition of a gonadal blood supply as the early gonad develops (Haugen *et al.*, 2011; Sao *et al.*, 2012). There was an organised arrangement of cells with tube structures attached to the intestine. This lay adjacent to one of the presumptive gonadal lobes. This arrangement was surmised to be the gall bladder and appeared in a similar location in many of the mid gut sections. The sections obtained from the 30 dph fish did not clearly reveal

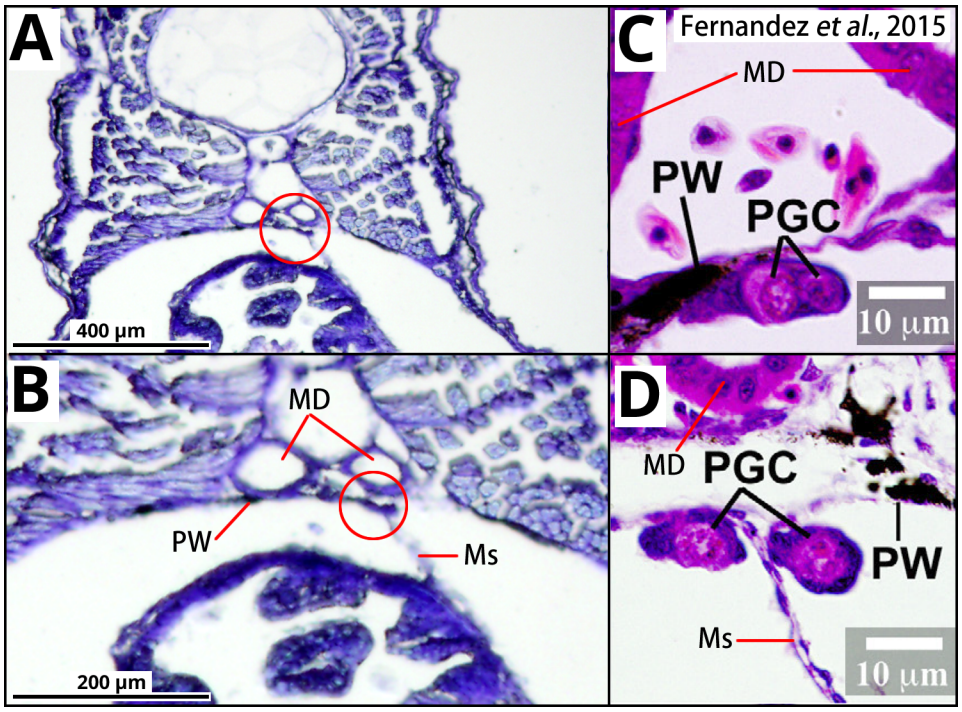


Figure 84. Comparison of the unidentified structure in a 15 dph larva from Figure 76, with the PGCs and presumptive genital ridge identified in the same region by Fernández *et al.* (2015). **A:** Unidentified structure between the peritoneal wall, mesonephric ducts and mesentery, captured at 100x zoom. **B:** Image A digitally enlarged to 200% size. **C-D:** Colour-adjusted images from Fernández *et al.* (2015) of 15 dph and 18 dph *S. lalandi* larvae respectively, showing PGCs adjacent to the peritoneal wall, mesonephric ducts and mesentery.

the oval structures described above; however, the other landmark tissues were evident (Figure 79). The fact that structures representing the presumptive gonad are not clearly evident is likely to be an artifact of the histological processing. At 60 dph, a structure was observed between the swim bladder, intestines and liver (Figure 83, image B) in the same region as the oval structures at 40 dph. Trunk samples at this age could not be embedded whole in paraffin as they were too large for the embedding moulds. OCT embedding and cryostat sectioning was successful for this age, but it was not possible to attain high quality section resolution with this technique. However, the large size of these samples allowed the generation of detailed composite photomicrograph images, which profiled general organ

development at this age. As many of the histological sections in the current study were plagued with resolution issues, it is difficult to identify the PGCs and developing gonad with certainty. However, the evidence indicates that the presumptive gonad may well have been located in many of the larval stages from 9 dph to 60 dph.

Finally, while outside the scope of this study, the images represent a comprehensive visual description of general tissue development in larval *S. lalandi*. This information may be useful for future studies on the larval life stage of this species, such as early digestive or immune system development.

4.6 Conclusion

This study successfully identified transcripts of numerous key genes involved in sex differentiation in the *S. lalandi* RNA-Seq transcriptome library. The nucleotide sequences for two of these genes; *Vasa* and *Amh*; were confirmed in adult kingfish by sequencing. A preliminary look at the expression of *Vasa*, *Amh* and *Cyp19a1a* at hatch, 3 dph, 12 dph and 18 dph using real-time PCR presented insight into the possible functioning and expression of these genes during early development. However, with the exception of *Cyp19a1a* at 3 dph and 18 dph, no significant changes in expression were detected. There were no obvious indications of sex differentiation during this time period. The use of histological approaches highlighted structures likely to be linked to the early gonad. However, these could not be identified due to section quality. These processing issues were largely overcome during later stages of the project, providing a pathway forward for further work. Collectively this work has established a platform for future studies of sex differentiation in *S. lalandi*. This has been achieved through the creation of a suite of molecular tools and refinement of appropriate histological techniques for the early life stages of this species. This will assist in the development of tools for the improvement of *S. lalandi* aquaculture, for applications such as monosex culture and surrogate broodstock technologies.

4.7 Future Recommendations

This study has developed a broad foundation for further molecular and histological studies of sex differentiation in *S. lalandi*. The elucidation of the complete nucleotide sequences of *Amh* by 5′ RACE PCR, and for *Vasa* and *Cyp19a1a* by both 3′ and 5′ RACE PCR, remains necessary for the complete characterisation of these genes in *S. lalandi*. Conducting real-time PCR on the remaining time points would help to complete the profile of gene expression during development, and increasing the number of pooled larvae to 100 for each time point would increase RNA yield and improve representation of both sexes. The uncertainty surrounding the expression of the housekeeping genes β-actin and *Gapdh* must be addressed for future studies using real-time PCR in this species, potentially selecting other candidate housekeeping genes which may be more suitable for this purpose. Recent advances in transcriptomic technology means that studies in new species, where no genetic information previously exists, can be carried out with relative ease, permitting a full transcriptome analysis instead of assessing individual genes one by one (Qian *et al.*, 2014). Future work could employ RNA-Seq to look at gene expression during larval development, providing a thorough depiction of sex differentiation where all genes important in this process could be identified and analysed simultaneously. A limitation of transcriptomics is that gene expression is not always indicative of protein expression. The application of proteomics may provide a more accurate profile of sex differentiation protein expression during early development of *S. lalandi*, as has been demonstrated for other species (Ziv *et al.*, 2008; Wright *et al.*, 2012; Groh *et al.*, 2013), and thereby resolve some of the uncertainty regarding gene expression.

Although the quality of the histological sections improved with refinement of the paraffin embedding process, resolution remained lower than desired. This is likely to relate to the thickness of the sections. It would be recommended to either produce thinner sections or use plastic resin embedding to produce semi-thin sections of

S. lalandi larvae. To conclusively identify PGCs in fish, many studies have utilised *in-situ* hybridisation or immunohistochemistry, using probes or antibodies designed for *Vasa* (Tohru Kobayashi *et al.*, 2000; Yoshizaki *et al.*, 2000; Tanaka *et al.*, 2001; Tohru Kobayashi *et al.*, 2002; Li *et al.*, 2009; Presslauer *et al.*, 2012; Huang *et al.*, 2014; Saito *et al.*, 2014; Yuan *et al.*, 2014; Toru Kobayashi *et al.*, 2015). Both approaches are highly effective at detecting PGCs in embryonic and larval fish. One of the objectives of the present study was to elucidate the nucleotide sequence for *Vasa* in *S. lalandi*; this achievement will now enable the development of RNA probes to conclusively identify the location of PGCs, and therefore the developing early gonad. Although not a primary objective of this study, this would have been attempted if time had permitted. However, the design and optimisation of probes is notoriously time consuming, and it was not feasible to undertake this in the time available. This would be a valuable next step for future research for characterising PGCs in *S. lalandi*.

Once these recommendations for optimising target gene expression and histological processing have been achieved, then an expanded time-series of juvenile *S. lalandi* can be analysed to successfully describe sex differentiation in this species.

References

Ali, S.E., Thoen, E., Vralstad, T., Kristensen, R., Evensen, O., Skaar, I. (2013). Development and reproduction of *Saprolegnia* species in biofilms. *Vet Microbiol 163*(1-2), 133-141.

Anjum, N. (2013). The merits of using alternate staining with toluidine blue and hematoxylin-eosin during Mohs micrographic surgery for basal cell carcinomas. *Journal of the American Academy of Dermatology 68*(4), AB88.

Aquaculture New Zealand. (2012). New Zealand Aquaculture: A sector overview with key facts, statistics and trends. Aquaculture New Zealand. New Zealand.

Artman, L., Dormoy-Raclet, V., von Roretz, C., Gallouzi, I.E. (2014). Planning your every move: the role of β-actin and its post-transcriptional regulation in cell motility. *Semin Cell Dev Biol 34*, 33-43.

Asche, F., Guttormsen, A.G., Nielsen, R. (2013). Future challenges for the maturing Norwegian salmon aquaculture industry: An analysis of total factor productivity change from 1996 to 2008. *Aquaculture 396-399*, 43-50.

Austbo, L., Aas, I.B., Konig, M., Weli, S.C., Syed, M., Falk, K., Koppang, E.O. (2014). Transcriptional response of immune genes in gills and the interbranchial lymphoid tissue of Atlantic salmon challenged with infectious salmon anaemia virus. *Dev Comp Immunol 45*(1), 107-114.

Babiak, J., Babiak, I., van Nes, S., Harboe, T., Haugen, T., Norberg, B. (2012). Induced sex reversal using an aromatase inhibitor, Fadrozole, in Atlantic halibut (*Hippoglossus hippoglossus* L.). *Aquaculture 324-325*, 276-280.

Bain, P.A., Schuller, K.A. (2012). Molecular cloning of glutathione peroxidase cDNAs from *Seriola lalandi* and analysis of changes in expression in cultured fibroblast-like cells in response to tert-butyl hydroquinone. *Aquaculture 324-325*, 182-193.

Bairagi, A., Sarkar Ghosh, A., Sen, S.K., Ray, A.K. (2004). Evaluation of the nutritive value of *Leucaena leucocephala* leaf meal, inoculated with fish intestinal bacteria *Bacillus subtilis* and *Bacillus circulans* in formulated diets for rohu, *Labeo rohita* (Hamilton) fingerlings. *Aquaculture Research 35*, 436-446.

Baroiller, J.F., D'Cotta, H.D. (2001). Environmental and sex determination in farmed fish. *Comparative Biochemistry and Physiology Part C 130*, 399-409.

Barrington, M. (2012). No word on fish farm revival plan. *The Northern Advocate.* The New Zealand Herald. New Zealand.

Beardmore, J.A., Mair, G.C., Lewis, R.I. (2001). Monosex male production in finfish as exemplified by tilapia: applications, problems and prospects. *Aquaculture 197*, 283-301.

Bertani, G. (1951). Studies on lysogenesis I. The mode of phage liberation by lysogenic *Escherichia coli*. *Journal of Bacteriology 62*(3), 293-300.

Blazquez, M., Gonzalez, A., Mylonas, C.C., Piferrer, F. (2011). Cloning and sequence analysis of a vasa homolog in the European sea bass (*Dicentrarchus labrax*): tissue distribution and mRNA expression levels during early development and sex differentiation. *Gen Comp Endocrinol 170*(2), 322-333.

Blazquez, M., Piferrer, F. (2004). Cloning, sequence analysis, tissue distribution, and sex-specific expression of the neural form of P450 aromatase in juvenile sea bass (*Dicentrarchus labrax*). *Mol Cell Endocrinol 219*(1-2), 83-94.

Bondad-Reantaso, M.G., Subasinghe, R.P., Arthur, J.R., Ogawa, K., Chinabut, S., Adlard, R., . . . Shariff, M. (2005). Disease and health management in Asian aquaculture. *Vet Parasitol 132*(3-4), 249-272.

Bostock, J., McAndrew, B., Richards, R., Jauncey, K., Telfer, T., Lorenzen, K., . . . Corner, R. (2010). Aquaculture: global status and trends. *Philosophical Transactions of the Royal Society B: Biological Sciences 365*(1554), 2897-2912.

Bricknell, I., Dalmo, R.A. (2005). The use of immunostimulants in fish larval aquaculture. *Fish Shellfish Immunol 19*(5), 457-472.

Burrell, M., Meehan, L., Munro, S. (2006). The New Zealand Aquaculture Strategy. New Zealand Aquaculture Council. New Zealand.

Bustos, P.A., Young, N.D., Rozas, M.A., Bohle, H.M., Ildefonso, R.S., Morrison, R.N., Nowak, B.F. (2011). Amoebic gill disease (AGD) in Atlantic salmon (*Salmo salar*) farmed in Chile. *Aquaculture 310*(3-4), 281-288.

Bwanika, G.N., Murie, D.J., Chapman, L.J. (2007). Comparative age and growth of Nile tilapia (*Oreochromis niloticus* L.) in lakes Nabugabo and Wamala, Uganda. *Hydrobiologia 589*(1), 287-301.

Callard, G.V., Tchoudakova, A.V., Kishida, M., Wood, E. (2001). Differential tissue distribution, developmental programming, estrogen regulation and promoter characteristics of *cyp19* genes in teleost fish. *Journal of Steroid Biochemistry & Molecular Biology 79*, 305-314.

Campanella, J.J., Bitincka, L., Smalley, J. (2003). MatGAT: An application that generates similarity/identity matrices using protein or DNA sequences. *BMC Bioinformatics 4*.

Caradec, J., Sirab, N., Keumeugni, C., Moutereau, S., Chimingqi, M., Matar, C., . . . Loric, S. (2010). 'Desperate house genes': the dramatic example of hypoxia. *Br J Cancer 102*(6), 1037-1043.

Caruthers, J.M., Johnson, E.R., McKay, D.B. (2000). Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *PNAS 97*(24), 13080-13085.

Chakraborty, S.B., Banerjee, S. (2009). Culture of Nile Tilapia under different traditional and non-traditional methods in India. *World Journal of Fish and Marine Sciences 1*(3), 212-217.

Corsi, I., Mariottini, M., Sensini, C., Lancini, L., Focardi, S. (2003). Cytochrome P450, acetylcholinesterase and gonadal histology for evaluating contaminant exposure levels in fishes from a highly eutrophic brackish system: the Orbetello Lagoon, Italy. *Marine Pollution Bulletin 46*, 203-212.

Covello, J.M., Friend, S.E., Purcell, S.L., Burka, J.F., Markham, R.J.F., Donkin, A.W., . . . Fast, M.D. (2012). Effects of orally administered immunostimulants on inflammatory gene expression and sea lice (*Lepeophtheirus salmonis*) burdens on Atlantic salmon (*Salmo salar*). *Aquaculture 366-367*, 9-16.

Dang, W., Sun, L. (2011). Determination of internal controls for quantitative real time RT-PCR analysis of the effect of *Edwardsiella tarda* infection on gene expression in turbot (*Scophthalmus maximus*). *Fish Shellfish Immunol 30*(2), 720-728.

Darawiroj, D., Kondo, H., Hirono, I., Aoki, T. (2008). Immune-related gene expression profiling of yellowtail (*Seriola quinqueradiata*) kidney cells stimulated with ConA and LPS using microarray analysis. *Fish Shellfish Immunol 24*(2), 260-266.

Davidson, J., Good, C., Barrows, F.T., Welsh, C., Kenney, P.B., Summerfelt, S.T. (2013). Comparing the effects of feeding a grain- or a fish meal-based diet on water quality, waste production, and rainbow trout *Oncorhynchus mykiss* performance within low exchange water recirculating aquaculture systems. *Aquacultural Engineering 52*, 45-57.

Davidson, W.S., Koop, B.F., Jones, S.J.M., Iturra, P., Vidal, R., Maass, A., . . . Omholt, S.W. (2010). Sequencing the genome of the Atlantic salmon (*Salmo salar*). *Genome Biology 11*(9), 403-409.

de Kok, J.B., Roelofs, R.W., Giesendorf, B.A., Pennings, J.L., Waas, E.T., Feuth, T., . . . Span, P.N. (2005). Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. *Lab Invest 85*(1), 154-159.

De Santis, C., Smith-Keune, C., Jerry, D.R. (2011). Normalizing RT-qPCR data: are we getting the right answers? An appraisal of normalization approaches and internal reference genes from a case study in the finfish *Lates calcarifer*. *Mar Biotechnol (NY) 13*(2), 170-180.

Desprez, D., Mélard, C., Hoareau, M.C., Bellemène, Y., Bosc, P., Baroiller, J.F. (2003). Inheritance of sex in two ZZ pseudofemale lines of tilapia *Oreochromis aureus*. *Aquaculture 218*(1-4), 131-140.

Deutsch, L., Gräslund, S., Folke, C., Troell, M., Huitric, M., Kautsky, N., Lebel, L. (2007). Feeding aquaculture growth through globalization: Exploitation of marine ecosystems for fishmeal. *Global Environmental Change 17*(2), 238-249.

Devlin, R.H., Nagahama, Y. (2002). Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture 208*, 191-364.

Díaz, N., Ribas, L., Piferrer, F. (2013). The relationship between growth and sex differentiation in the European sea bass (*Dicentrarchus labrax*). *Aquaculture 408-409*, 191-202.

Dittami, S.M., Riisberg, I., John, U., Orr, R.J., Jakobsen, K.S., Edvardsen, B. (2012). Analysis of expressed sequence tags from the marine microalga *Pseudochattonella farcimen* (Dictyochophyceae). *Protist 163*(1), 143-161.

Doupé, R., Recher, H. (1999). Gene pool management of hatchery Barramundi *Lates calcarifer* for production and stock augmentation programmes. *Pacific Conservation Biology 5*, 73-75.

Eastham, W.N., Essex, W.B. (1969). Use of tissues embedded in epoxy resin for routine histological examination of renal biopsies. *J. Clin. Path. 22*, 99-106.

FAO. (2012). The world state of fisheries and aquaculture. Food and Agriculture Organization of the United Nations. Rome.

Fernandes, J.M., Mommens, M., Hagen, O., Babiak, I., Solberg, C. (2008). Selection of suitable reference genes for real-time PCR studies of Atlantic halibut development. *Comp Biochem Physiol B Biochem Mol Biol 150*(1), 23-32.

Fernández-Palacios, H., Schuchardt, D., Roo, J., Hernández-Cruz, C., Izquierdo, M. (2014). Spawn quality and GnRHa induction efficiency in longfin yellowtail (*Seriola rivoliana*) broodstock kept in captivity. *Aquaculture 435*, 167-172.

Fernández, J.A., Bubner, E.J., Takeuchi, Y., Yoshizaki, G., Wang, T., Cummins, S.F., Elizur, A. (2015). Primordial germ cell migration in the yellowtail kingfish (*Seriola lalandi*) and identification of stromal cell-derived factor 1. *General and Comparative Endocrinology*.

Fitzpatrick, J.L., Henry, J.C., Liley, N.R., Devlin, R.H. (2005). Sperm characteristics and fertilization success of masculinized coho salmon (*Oncorhynchus kisutch*). *Aquaculture 249*(1-4), 459-468.

Florent, R.L., Becker, J.A., Powell, M.D. (2007a). Efficacy of bithionol as an oral treatment for amoebic gill disease in Atlantic salmon *Salmo salar* (L.). *Aquaculture 270*(1-4), 15-22.

Florent, R.L., Becker, J.A., Powell, M.D. (2007b). Evaluation of bithionol as a bath treatment for amoebic gill disease caused by *Neoparamoeba* spp. *Vet Parasitol 144*(3-4), 197-207.

Frost, L.A., Evans, B.S., Jerry, D.R. (2006). Loss of genetic diversity due to hatchery culture practices in barramundi (*Lates calcarifer*). *Aquaculture 261*(3), 1056-1064.

Fugelstad, J., Bouzenzana, J., Djerbi, S., Guerriero, G., Ezcurra, I., Teeri, T.T., . . . Bulone, V. (2009). Identification of the cellulose synthase genes from the Oomycete *Saprolegnia monoica* and effect of cellulose synthesis inhibitors on gene expression and enzyme activity. *Fungal Genet Biol 46*(10), 759-767.

Fuji, K., Yoshida, K., Hattori, K., Ozaki, A., Araki, K., Okauchi, M., . . . Sakamoto, T. (2010). Identification of the sex-linked locus in yellowtail, *Seriola quinqueradiata*. *Aquaculture 308*, S51-S55.

Fujiwara, Y., Komiya, T., Kawabata, H., Sato, M., Fujimoto, H., Furusawa, M., Noce, T. (1994). Isolation of a DEAD-family protein gene that encodes a murine homolog of *Drosophila* vasa and its specific expression in germ cell lineage. *Proc Natl Acad Sci U S A 91*, 12258-12262.

Garcia-Segura, L.M. (2008). Aromatase in the brain: not just for reproduction anymore. *J Neuroendocrinol 20*(6), 705-712.

Gatlin, D.M., Barrows, F.T., Brown, P., Dabrowski, K., Gaylord, T.G., Hardy, R.W., . . . Wurtele, E. (2007). Expanding the utilization of sustainable plant products in aquafeeds: a review. *Aquaculture Research 38*(6), 551-579.

Gaylord, T.G., Barrows, F.T. (2009). Multiple amino acid supplementations to reduce dietary protein in plant-based rainbow trout, *Oncorhynchus mykiss*, feeds. *Aquaculture 287*(1-2), 180-184.

Genten, F., Terwinghe, E., Danguy, A. (2009). Atlas of Fish Histology. Science Publishers, Enfield (NH) Jersey Plymouth.

Gibbs, P.J., Cameron, C., Tan, L.C., Sadek, S.A., Howell, W.M. (2003). House keeping genes and gene expression analysis in transplant recipients: a note of caution. *Transplant Immunology 12*(1), 89-97.

Glenn, T.C. (2011). Field guide to next-generation DNA sequencers. *Mol Ecol Resour 11*(5), 759-769.

Goudie, C.A., Simco, B.A., Davis, K.B., Carmichael, G.J. (1994). Growth of channel catfish in mixed sex and monosex pond culture. *Aquaculture 128*(1-2), 97-104.

Greytak, S.R., Champlin, D., Callard, G.V. (2005). Isolation and characterization of two cytochrome P450 aromatase forms in killifish (*Fundulus heteroclitus*): differential expression in fish from polluted and unpolluted environments. *Aquat Toxicol 71*(4), 371-389.

Groh, K.J., Schonenberger, R., Eggen, R.I., Segner, H., Suter, M.J. (2013). Analysis of protein expression in zebrafish during gonad differentiation by targeted proteomics. *Gen Comp Endocrinol 193*, 210-220.

Guiguen, Y., Fostier, A., Piferrer, F., Chang, C.F. (2010). Ovarian aromatase and estrogens: a pivotal role for gonadal sex differentiation and sex change in fish. *Gen Comp Endocrinol 165*(3), 352-366.

Hafeez-Ur-Rehman, M., Ahmed, I., Ashraf, M., Khan, N., Rasool, F. (2008). The culture performance of mono-sex and mixed-sex Tilapia in fertilized ponds. *International Journal of Agriculture and Biology 10*(3), 352-354.

Haffray, P., Bruant, J.-S., Facqueur, J.-M., Fostier, A. (2005). Gonad development, growth, survival and quality traits in triploids of the protandrous hermaphrodyte gilthead seabream *Sparus aurata* (L.). *Aquaculture 247*(1-4), 107-117.

Haffray, P., Lebègue, E., Jeu, S., Guennoc, M., Guiguen, Y., Baroiller, J.F., Fostier, A. (2009). Genetic determination and temperature effects on turbot *Scophthalmus maximus* sex differentiation: An investigation using steroid sex-inverted males and females. *Aquaculture 294*(1-2), 30-36.

Halm, S., Rocha, A., Miura, T., Prat, F., Zanuy, S. (2007). Anti-Müllerian hormone (*AMH*/AMH) in the European sea bass: its gene structure, regulatory elements, and the expression of alternatively-spliced isoforms. *Gene 388*(1-2), 148-158.

Harikrishnan, R., Balasundaram, C., Heo, M.-S. (2010). Molecular studies, disease status and prophylactic measures in grouper aquaculture: Economic importance, diseases and immunology. *Aquaculture 309*(1-4), 1-14.

Harikrishnan, R., Balasundaram, C., Heo, M.-S. (2011). Fish health aspects in grouper aquaculture. *Aquaculture 320*(1-2), 1-21.

Haugen, T., Andersson, E., Norberg, B., Taranger, G.L. (2011). The production of hermaphrodites of Atlantic cod (*Gadus morhua*) by masculinization with orally administered 17-α-methyltestosterone, and subsequent production of all-female cod populations. *Aquaculture 311*(1-4), 248-254.

Hauptman, B.S., Barrows, F.T., Block, S.S., Gibson Gaylord, T., Paterson, J.A., Rawles, S.D., Sealey, W.M. (2014). Evaluation of grain distillers dried yeast as a fish meal substitute in practical-type diets of juvenile rainbow trout, *Oncorhynchus mykiss*. *Aquaculture 432*, 7-14.

Huang, J., Chen, S., Liu, Y., Shao, C., Lin, F., Wang, N., Hu, Q. (2014). Molecular characterization, sexually dimorphic expression, and functional analysis of 3'-untranslated region of vasa gene in half-smooth tongue sole (*Cynoglossus semilaevis*). *Theriogenology 82*(2), 213-224.

Hutson, K.S., Ernst, I., Whittington, I.D. (2007). Risk assessment for metazoan parasites of yellowtail kingfish *Seriola lalandi* (Perciformes: Carangidae) in South Australian sea-cage aquaculture. *Aquaculture 271*(1-4), 85-99.

Imbeaud, S., Carré-Eusèbe, D., Rey, R., Belville, C., Josso, N., Picard, J.Y. (1994). Molecular genetics of the persistent müllerian duct syndrome: a study of 19 families. *Human Molecular Genetics 3*(1), 125-131.

Imsland, A.K., Folkvord, A., Grung, G.L., Stefansson, S.O., Taranger, G.L. (1997). Sexual dimorphism in growth and maturation of turbot, *Scophthalmus maximus* (Rafinesque, 1810). *Aquaculture Research 28*, 101-114.

Infante, C., Matsuoka, M.P., Asensio, E., Canavate, J.P., Reith, M., Manchado, M. (2008). Selection of housekeeping genes for gene expression studies in larvae from flatfish using real-time PCR. *BMC Mol Biol 9*, 28.

Jia, R., Nie, L.W., Wang, N., Wang, J. (2009). Molecular cloning and expression patterns of the *Vasa* gene from *Rana nigromaculata* (Amphibia: Anura). *Zoologia 26*(2), 316-322.

Johnsen, H., Andersen, O. (2012). Sex dimorphic expression of five *dmrt* genes identified in the Atlantic cod genome. The fish-specific *dmrt2b* diverged from *dmrt2a* before the fish whole-genome duplication. *Gene 505*(2), 221-232.

Josso, N., Cate, R.L., Picard, J.Y., Vigier, B., di Clemente, N., Wilson, C., . . . Carré-Eusèbe, D. (1993). Anti-Müllerian hormone: the Jost factor. *Rec Prog Horm 48*, 1-59.

Kanamori, A., Naruse, K., Mitani, H., Shima, A., Hori, H. (2003). Genomic organization of ZP domain containing egg envelope genes in medaka (*Oryzias latipes*). *Gene 305*(1), 35-45.

Kavumpurath, S., Pandian, T.J. (1994). Masculinization of fighting fish, *Betta splendens* Regan, using synthetic or natural androgens. *Aquaculture Research 25*(4), 373-381.

Kendon, P. (2012). No go for Katikati fish farm. *Sunlive.* Sun Media Limited. Bay of Plenty, New Zealand.

Kim, D.H., Austin, B. (2006). Cytokine expression in leucocytes and gut cells of rainbow trout, *Oncorhynchus mykiss* Walbaum, induced by probiotics. *Vet Immunol Immunopathol 114*(3-4), 297-304.

Kobayashi, T., Honryo, T., Agawa, Y., Sawada, Y., Tapia, I., Macìas, K.A., . . . Yagishita, N. (2015). Gonadogenesis and slow proliferation of germ cells in juveniles of cultured yellowfin tuna, *Thunnus albacares*. *Reproductive Biology*.

Kobayashi, T., Kaijura-Kobayashi, H., Nagahama, Y. (2000). Differential expression of *vasa* homologue gene in the germ cells during oogenesis and spermatogenesis in a teleost fish, tilapia, *Oreochromis niloticus*. *Mechanisms of Development 99*, 139-142.

Kobayashi, T., Kajiura-Kobayashi, H., Nagahama, Y. (2002). Two isoforms of *vasa* homologs in a teleost fish: their differential expression during germ cell differentiation. *Mech Dev 111*, 167-171.

Koprunner, M., Thisse, C., Thisse, B., Raz, E. (2001). A zebrafish *nanos*-related gene is essential for the development of primordial germ cells. *Genes Dev 15*(21), 2877-2885.

Kumar, A., Kalra, B., Patel, A., McDavid, L., Roudebush, W.E. (2010). Development of a second generation anti-Müllerian hormone (AMH) ELISA. *J Immunol Methods 362*(1-2), 51-59.

Kwon, J.Y., Kim, J. (2013). Differential expression of two distinct aromatase genes (*cyp19a1a* and *cyp19a1b*) during vitellogenesis and gestation in the viviparous black rockfish *Sebastes schlegelii*. *Animal Cells and Systems 17*(2), 88-98.

Lazado, C.C., Caipang, C.M.A. (2014). Atlantic cod in the dynamic probiotics research in aquaculture. *Aquaculture 424-425*, 53-62.

Leninger, A.L., Nelson, D.L., Cox, M.M. (1975). Lenninger's principles of biochemistry. W. H. Freeman and Co, New York, NY.

Li, M., Hong, N., Xu, H., Yi, M., Li, C., Gui, J., Hong, Y. (2009). Medaka *vasa* is required for migration but not survival of primordial germ cells. *Mech Dev 126*(5-6), 366-381.

Li, M., Wang, L., Wang, H., Liang, H., Zheng, Y., Qin, F., . . . Wang, Z. (2013). Molecular cloning and characterization of *amh*, *dax1* and *cyp19a1a* genes and their response to 17alpha-methyltestosterone in Pengze crucian carp. *Comp Biochem Physiol C Toxicol Pharmacol 157*(4), 372-381.

Liao, X., Cheng, L., Xu, P., Lu, G., Wachholtz, M., Sun, X., Chen, S. (2013). Transcriptome analysis of Crucian carp (*Carassius auratus*), an important aquaculture and hypoxia-tolerant species. *PLOS One 8*(4), 1-11.

Liu, C., Xin, N., Jiang, L., Zhai, J., Zhang, Q., Jie, Q. (2014). Reference gene selection for quantitative real-time RT-PCR normalization in the half-smooth tongue sole (*Cynoglossus semilaevis*) at different development stages, in various tissue types and on exposure to chemicals. *PLOS One 9*(3), e91715.

Livak, K.J., Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-ΔΔCT) Method. *Methods 25*(4), 402-408.

Ludbrook, L.M., Harley, V.R. (2004). Sex determination: a 'window' of DAX1 activity. *Trends Endocrinol Metab 15*(3), 116-121.

MacKenzie, L.A., Smith, K.F., Rhodes, L.L., Brown, A., Langi, V., Edgar, M., . . . Preece, M. (2011). Mortalities of sea-cage salmon (*Oncorhynchus tshawytscha*) due to a bloom of *Pseudochattonella verruculosa* (Dictyochophyceae) in Queen Charlotte Sound, New Zealand. *Harmful Algae 11*, 45-53.

Mardones, F.O., Perez, A.M., Valdes-Donoso, P., Carpenter, T.E. (2011). Farm-level reproduction number during an epidemic of infectious salmon anemia virus in southern Chile in 2007-2009. *Prev Vet Med 102*(3), 175-184.

Martins, R.S., Power, D.M., Fuentes, J., Deloffre, L.A., Canario, A.V. (2013). *DAX1* regulatory networks unveil conserved and potentially new functions. *Gene 530*(1), 66-74.

Matsuoka, M.P., van Nes, S., Andersen, O., Benfey, T.J., Reith, M. (2006). Real-time PCR analysis of ovary- and brain-type aromatase gene expression during Atlantic halibut (*Hippoglossus hippoglossus*) development. *Comp Biochem Physiol B Biochem Mol Biol 144*(1), 128-135.

MBIE. (2013). iFAB 2013 seafood review. Ministry of Business, Innovation and Employment. New Zealand.

McGettigan, P.A. (2013). Transcriptomics in the RNA-seq era. *Curr Opin Chem Biol 17*(1), 4-11.

Menuet, A., Pellegrini, E., Brion, F., Gueguen, M.M., Anglade, I., Pakdel, F., Kah, O. (2005). Expression and estrogen-dependent regulation of the zebrafish brain aromatase gene. *J Comp Neurol 485*(4), 304-320.

Militz, T.A., Southgate, P.C., Carton, A.G., Hutson, K.S. (2013). Dietary supplementation of garlic (*Allium sativum*) to prevent monogenean infection in aquaculture. *Aquaculture 408-409*, 95-99.

Mitter, K., Kotoulas, G., Magoulas, A., Mulero, V., Sepulcre, P., Figueras, A., . . . Sarropoulou, E. (2009). Evaluation of candidate reference genes for QPCR during ontogenesis and of immune-relevant tissues of European seabass (*Dicentrarchus labrax*). *Comp Biochem Physiol B Biochem Mol Biol 153*(4), 340-347.

Mo, F., Zhao, J., Liu, N., Cao, L., Jiang, S. (2014). Validation of reference genes for RT-qPCR analysis of CYP4T expression in crucian carp. *Genetics and Molecular Biology 37*(2), 500-507.

Moran, D., Pether, S.J., Lee, P.S. (2009). Growth, feed conversion and faecal discharge of yellowtail kingfish (*Seriola lalandi*) fed three commercial diets. *New Zealand Journal of Marine and Freshwater Research 43*(4), 917-927.

Moran, D., Smith, C.K., Gara, B., Poortenaar, C.W. (2007). Reproductive behaviour and early development in yellowtail kingfish (*Seriola lalandi* Valenciennes 1833). *Aquaculture 262*(1), 95-104.

Morita, T., Kumakura, N., Morishima, K., Mitsuboshi, T., Ishida, M., Hara, T., . . . Yoshizaki, G. (2012). Production of donor-derived offspring by allogeneic transplantation of spermatogonia in the yellowtail (*Seriola quinqueradiata*). *Biol Reprod 86*(6), 176.

Mouse Genome Informatics. (2014). Guidelines for Nomenclature of Genes, Genetic Markers, Alleles, and Mutations in Mouse and Rat

Muncaster, S., Norberg, B., Andersson, E. (2013). Natural sex change in the temperate protogynous Ballan wrasse *Labrus bergylta*. *J Fish Biol 82*(6), 1858-1870.

Mussely, H., Goodwin, E. (2012). Report No. 2094: Feasibility of land-based aquaculture in New Zealand. Cawthron Institute. Nelson, New Zealand.

Mylonas, C.C. (2013). Diversify: Enhancing the European aquaculture production by removing production bottlenecks of emerging species, producing new products and accessing new markets. Hellenic Center for Marine Research, Institute of Marine Biology, Biotechnology and Aquaculture. Crete, Greece.

Nakada, M. (2008). Capture-based aquaculture of yellowtail. In A. Lovatelli & P. F. Holthus (Eds.), *Capture-based aquaculture: Global overview. FAO Fisheries Technical Paper. No. 58.* (pp. 199-215). Rome: FAO.

Norberg, B., Valkner, V., Huse, J., Karlsen, I., Grung, G.L. (1991). Ovulatory rhythms and egg viability in the Atlantic halibut (*Hippoglossus hippoglossus*). *Aquaculture 97*(4), 365-371.

Ohara, E., Nishimura, T., Nagakura, Y., Sakamoto, T., Mushiake, K., Okamoto, N. (2005). Genetic linkage maps of two yellowtails (*Seriola quinqueradiata* and *Seriola lalandi*). *Aquaculture 244*(1-4), 41-48.

Okubo, K., Takeuchi, A., Chaube, R., Paul-Prasanth, B., Kanda, S., Oka, Y., Nagahama, Y. (2011). Sex differences in aromatase gene expression in the medaka brain. *J Neuroendocrinol 23*(5), 412-423.

Olsen, R.L., Hasan, M.R. (2012). A limited supply of fishmeal: Impact on future increases in global aquaculture production. *Trends in Food Science & Technology 27*(2), 120-128.

Orellana, J., Waller, U., Wecker, B. (2014). Culture of yellowtail kingfish (*Seriola lalandi*) in a marine recirculating aquaculture system (RAS) with artificial seawater. *Aquacultural Engineering 58*, 20-28.

Øvergård, A.C., Nerland, A.H., Patel, S. (2010). Evaluation of potential reference genes for real time RT-PCR studies in Atlantic halibut (*Hippoglossus Hippoglossus* L.); during development, in tissues of healthy and NNV-injected fish, and in anterior kidney leucocytes. *BMC Mol Biol 11*, 36.

Pala, I., Kluver, N., Thorsteinsdottir, S., Schartl, M., Coelho, M.M. (2008). Expression pattern of anti-Müllerian hormone (*amh*) in the hybrid fish complex of *Squalius alburnoides*. *Gene 410*(2), 249-258.

Palstra, A.P., Beltran, S., Burgerhout, E., Brittijn, S.A., Magnoni, L.J., Henkel, C.V., . . . Planas, J.V. (2013). Deep RNA sequencing of the skeletal muscle transcriptome in swimming fish. *PLoS One 8*(1), e53171.

Penman, D.J., Piferrer, F. (2008). Fish Gonadogenesis. Part I: Genetic and Environmental Mechanisms of Sex Determination. *Reviews in Fisheries Science 16*(sup1), 16-34.

Péron, G., François Mittaine, J., Le Gallic, B. (2010). Where do fishmeal and fish oil products come from? An analysis of the conversion ratios in the global fishmeal industry. *Marine Policy 34*(4), 815-820.

Piferrer, F., Guiguen, Y. (2008). Fish Gonadogenesis. Part II: Molecular Biology and Genomics of Sex Differentiation. *Reviews in Fisheries Science 16*(sup1), 35-55.

Piferrer, F., Ribas, L., Diaz, N. (2012). Genomic approaches to study genetic and environmental influences on fish sex determination and differentiation. *Mar Biotechnol (NY) 14*(5), 591-604.

Poortenaar, C.W., Hooker, S.H., Sharp, N. (2001). Assessment of yellowtail kingfish (*Seriola lalandi lalandi*) reproductive physiology, as a basis for aquaculture development. *Aquaculture 201*, 271-286.

Presslauer, C., Nagasawa, K., Fernandes, J.M., Babiak, I. (2012). Expression of *vasa* and *nanos3* during primordial germ cell formation and migration in Atlantic cod (*Gadus morhua* L.). *Theriogenology 78*(6), 1262-1277.

Psenicka, M., Saito, T., Linhartova, Z., Gazo, I. (2015). Isolation and transplantation of sturgeon early-stage germ cells. *Theriogenology 83*(6), 1085-1092.

Qian, X., Ba, Y., Zhuang, Q., Zhong, G. (2014). RNA-Seq technology and its application in fish transcriptomics. *OMICS 18*(2), 98-110.

Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., . . . Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics 13*, 341.

Read, E.S., Barrows, F.T., Gibson Gaylord, T., Paterson, J., Petersen, M.K., Sealey, W.M. (2014). Investigation of the effects of dietary protein source on copper and zinc bioavailability in fishmeal and plant-based diets for rainbow trout. *Aquaculture 432*, 97-105.

Rodriguez-Mari, A., Yan, Y.L., Bremiller, R.A., Wilson, C., Canestro, C., Postlethwait, J.H. (2005). Characterization and expression pattern of zebrafish Anti-Müllerian hormone (*amh*) relative to *sox9a*, *sox9b*, and *cyp19a1a*, during gonad development. *Gene Expr Patterns 5*(5), 655-667.

Rohde, K. (2005). Marine Parasitology. CSIRO Publishing, Collingwood.

Rougeot, C., Jacobs, B., Kestemont, P., Melard, C. (2002). Sex control and sex determinism study in Eurasian perch, *Perca fluviatilis*, by use of hormonally sex-reversed male breeders. *Aquaculture 211*, 81-89.

Rougeot, C., Prignon, C., Ngouana Kengne, C.V., Mélard, C. (2008). Effect of high temperature during embryogenesis on the sex differentiation process in the Nile tilapia, *Oreochromis niloticus*. *Aquaculture 276*(1-4), 205-208.

Rozen, S., Skaletsky, H.J. (2000). Primer3 on the WWW for general users and for biologist programmers. In S. Krawetz & S. Misener (Eds.), *Bioinformatics methods and protocols: Methods in molecular biology*. Totowa, NJ: Humana Press.

Ruksana, S., Pandit, N.P., Nakamura, M. (2010). Efficacy of exemestane, a new generation of aromatase inhibitor, on sex differentiation in a gonochoristic fish. *Comp Biochem Physiol C Toxicol Pharmacol 152*(1), 69-74.

Saber, S., Macías, D., Ortiz de Urbina, J., Kjesbu, O.S. (2015). Stereological comparison of oocyte recruitment and batch fecundity estimates from paraffin and resin sections using spawning albacore (*Thunnus alalunga*) ovaries as a case study. *Journal of Sea Research 95*, 226-238.

Sadovy, Y., Colin, P.L. (1995). Sexual development and sexuality in the Nassau grouper. *Journal of Fish Biology 46*(6), 961-976.

Saillant, E., Fostier, A., Menu, B., Haffray, P., Chatain, B. (2001). Sexual growth dimorphism in sea bass *Dicentrarchus labrax*. *Aquaculture 202*, 371-387.

Saito, T., Psenicka, M., Goto, R., Adachi, S., Inoue, K., Arai, K., Yamaha, E. (2014). The origin and migration of primordial germ cells in sturgeons. *PLoS One 9*(2), e86861.

Sanchez-Pla, A., Reverter, F., Ruiz de Villa, M.C., Comabella, M. (2012). Transcriptomics: mRNA and alternative splicing. *J Neuroimmunol 248*(1-2), 23-31.

Sánchez, P., Ambrosio, P.P., Flos, R. (2010). Stocking density and sex influence individual growth of Senegalese sole (*Solea senegalensis*). *Aquaculture 300*(1-4), 93-101.

Sao, P.N., Hur, S.W., Lee, C.H., Lee, Y.D. (2012). Gonadal sex differentiation of hatchery-reared longtooth grouper (*Epinephelus bruneus*). *Dev. Reprod. 16*(3), 185-193.

Sarropoulou, E., Fernandes, J.M. (2011). Comparative genomics in teleost species: Knowledge transfer by linking the genomes of model and non-model fish species. *Comp Biochem Physiol Part D Genomics Proteomics 6*(1), 92-102.

Sawyer, S.J., Gerstner, K.A., Callard, G.V. (2006). Real-time PCR analysis of cytochrome P450 aromatase expression in zebrafish: gene specific tissue distribution, sex differences, developmental programming, and estrogen regulation. *Gen Comp Endocrinol 147*(2), 108-117.

Schulz, R.W., de Franca, L.R., Lareyre, J.J., Le Gac, F., Chiarini-Garcia, H., Nobrega, R.H., Miura, T. (2010). Spermatogenesis in fish. *Gen Comp Endocrinol 165*(3), 390-411.

Selim, K.M., Shinomiya, A., Otake, H., Hamaguchi, S., Sakaizumi, M. (2009). Effects of high temperature on sex differentiation and germ cell population in medaka, *Oryzias latipes*. *Aquaculture 289*(3-4), 340-349.

Sepúlveda, F.A., González, M.T. (2014). Molecular and morphological analyses reveal that the pathogen Benedenia seriolae (Monogenea: Capsalidae) is a complex species: Implications for yellowtail *Seriola* spp. aquaculture. *Aquaculture 418-419*, 94-100.

Sharp, N.J., Diggles, B.K., Poortenaar, C.W., Willis, T.J. (2004). Efficacy of Aqui-S, formalin and praziquantel against the monogeneans, *Benedenia seriolae* and *Zeuxapta seriolae*, infecting yellowtail kingfish *Seriola lalandi lalandi* in New Zealand. *Aquaculture 236*(1-4), 67-83.

Siegfried, K.R. (2010). In search of determinants: gene expression during gonadal sex differentiation. *J Fish Biol 76*(8), 1879-1902.

Silver, N., Best, S., Jiang, J., Thein, S.L. (2006). Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol Biol 7*, 33.

Singh, A.K. (2013). Introduction of modern endocrine techniques for the production of monosex population of fishes. *Gen Comp Endocrinol 181*, 146-155.

Skjelbred, B., Horsberg, T.E., Tollefsen, K.E., Andersen, T., Edvardsen, B. (2011). Toxicity of the ichthyotoxic marine flagellate *Pseudochattonella* (Dictyochophyceae, Heterokonta) assessed by six bioassays. *Harmful Algae 10*(2), 144-154.

Smith, E.K., Guzman, J.M., Luckenbach, J.A. (2013). Molecular cloning, characterization, and sexually dimorphic expression of five major sex differentiation-related genes in a Scorpaeniform fish, sablefish (*Anoplopoma fimbria*). *Comp Biochem Physiol B Biochem Mol Biol 165*(2), 125-137.

Snyder, G.S., Gaylord, T.G., Barrows, F.T., Overturf, K., Cain, K.D., Hill, R.A., Hardy, R.W. (2012). Effects of carnosine supplementation to an all-plant protein diet for rainbow trout (*Oncorhynchus mykiss*). *Aquaculture 338-341*, 72-81.

Stanley, J. (1976). A review of methods for obtaining monosex fish and progress report on production of monosex white amur. *Journal of Aquatic Plant Management 15*, 68-70.

Star, B., Nederbragt, A.J., Jentoft, S., Grimholt, U., Malmstrom, M., Gregers, T.F., . . . Jakobsen, K.S. (2011). The genome sequence of Atlantic cod reveals a unique immune system. *Nature 477*(7363), 207-210.

Strobl-Mazzulla, P.H., Lethimonier, C., Gueguen, M.M., Karube, M., Fernandino, J.I., Yoshizaki, G., . . . Somoza, G.M. (2008). Brain aromatase (Cyp19A2) and estrogen receptors, in larvae and adult pejerrey fish *Odontesthes bonariensis*: Neuroanatomical and functional relations. *Gen Comp Endocrinol 158*(2), 191-201.

Su, B., Peatman, E., Shang, M., Thresher, R., Grewe, P., Patil, J., . . . Dunham, R.A. (2013). Expression and knockdown of primordial germ cell genes, *vasa*, *nanos* and *dead end* in common carp (*Cyprinus carpio*) embryos for transgenic sterilization and reduced sexual maturity. *Aquaculture*.

Symonds, J.E., Walker, S.P., Pether, S., Gublin, Y., McQueen, D., King, A., . . . Bruce, M. (2014). Developing yellowtail kingfish (*Seriola lalandi*) and hāpuku (*Polyprion oxygeneios*) for New Zealand aquaculture. *New Zealand Journal of Marine and Freshwater Research 48*(3), 371-384.

Tacchi, L., Bickerdike, R., Douglas, A., Secombes, C.J., Martin, S.A. (2011a). Transcriptomic responses to functional feeds in Atlantic salmon (*Salmo salar*). *Fish Shellfish Immunol 31*(5), 704-715.

Tacchi, L., Bron, J.E., Taggart, J.B., Secombes, J., Bickerdike, R., Adler, M.A., . . . Martin, S.A.M. (2011b). Multiple tissue transcriptomic responses to *Piscirickettsia salmonis* in Atlantic salmon (*Salmo salar*). *Physiol Genomics 43*, 1241-1254.

Tafalla, C., Bogwald, J., Dalmo, R.A. (2013). Adjuvants and immunostimulants in fish vaccines: current knowledge and future perspectives. *Fish Shellfish Immunol 35*(6), 1740-1750.

Tanaka, M., Kinoshita, M., Kobayashi, D., Nagahama, Y. (2001). Establishment of medaka (*Oryzias latipes*) transgenic lines with the expression of green fluorescent protein fluorescence exclusively in germ cells: a useful model to monitor germ cells in a live vertebrate. *Proc Natl Acad Sci U S A 98*(5), 2544-2549.

Taranger, G.L., Carrillo, M., Schulz, R.W., Fontaine, P., Zanuy, S., Felip, A., . . . Hansen, T. (2010). Control of puberty in farmed fish. *Gen Comp Endocrinol 165*(3), 483-515.

Tchoudakova, A., Kishida, M., Wood, E., Callard, G.V. (2001). Promoter characteristics of two *cyp19* genes differentially expressed in the brain and ovary of teleost fish. *Journal of Steroid Biochemistry & Molecular Biology 78*, 427-439.

Thakur, N.L., Jain, R., Natalio, F., Hamer, B., Thakur, A.N., Muller, W.E. (2008). Marine molecular biology: an emerging field of biological sciences. *Biotechnol Adv 26*(3), 233-245.

Tine, M., Kuhl, H., Gagnaire, P.A., Louro, B., Desmarais, E., Martins, R.S., . . . Reinhardt, R. (2014). European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nat Commun 5*, 5770.

Trip, E.D., Clements, K.D., Raubenheimer, D., Choat, J.H. (2011). Reproductive biology of an odacine labrid, *Odax pullus*. *J Fish Biol 78*(3), 741-761.

Tubbs, L.A., Poortenaar, C.W., Sewell, M.A., Diggles, B.K. (2005). Effects of temperature on fecundity in vitro, egg hatching and reproductive development of *Benedenia seriolae* and *Zeuxapta seriolae* (Monogenea) parasitic on yellowtail kingfish *Seriola lalandi*. *Int J Parasitol 35*(3), 315-327.

van den Berg, A.H., McLaggan, D., Diéguez-Uribeondo, J., van West, P. (2013). The impact of the water moulds *Saprolegnia diclina* and *Saprolegnia parasitica* on natural ecosystems and the aquaculture industry. *Fungal Biology Reviews 27*(2), 33-42.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology 3*(7), 0034.0031-0034.0011.

Verhoef, S.P., van Dijk, P., Westerterp, K.R. (2013). Relative shrinkage of adipocytes by paraffin in proportion to plastic embedding in human adipose tissue before and after weight loss. *Obes Res Clin Pract 7*(1), e8-13.

Vike, S., Duesund, H., Andersen, L., Nylund, A. (2014). Release and survival of infectious salmon anaemia (ISA) virus during decomposition of Atlantic salmon (*Salmo salar* L.). *Aquaculture 420-421*, 119-125.

Viñas, J., Asensio, E., Cañavate, J.P., Piferrer, F. (2013). Gonadal sex differentiation in the Senegalese sole (*Solea senegalensis*) and first data on the experimental manipulation of its sex ratios. *Aquaculture 384-387*, 74-81.

von Schalburg, K.R., Gowen, B.E., Messmer, A.M., Davidson, W.S., Koop, B.F. (2014). Sex-specific expression and localization of aromatase and its regulators during embryonic and larval development of Atlantic salmon. *Comp Biochem Physiol B Biochem Mol Biol 168*, 33-44.

Wassilieff, M. (2012). 'Aquaculture - Salmon', Te Ara - the Encyclopedia of New Zealand. From *http://www.teara.govt.nz/en/aquaculture/page-2*

Watts, M., Pankhurst, N.W., King, H.R. (2004). Maintenance of Atlantic salmon (*Salmo salar*) at elevated temperature inhibits cytochrome P450 aromatase activity in isolated ovarian follicles. *General and Comparative Endocrinology 135*(3), 381-390.

Weltzien, F.A., Andersson, E., Andersen, O., Shalchian-Tabrizi, K., Norberg, B. (2004). The brain-pituitary-gonad axis in male teleosts, with special emphasis on flatfish (*Pleuronectiformes*). *Comp Biochem Physiol A Mol Integr Physiol 137*(3), 447-477.

Wilson, C.A., Di Clemente, N., Ehrenfels, C., Pepinsky, R.B., Josso, N., Vigier, B., Cate, R.L. (1993). Müllerian inhibiting substance requires its N-terminal domain for maintenance of biological activity, a novel finding within the transforming growth factor-β superfamily. *Mol Endocrinol 7*, 247-257.

Wright, P.C., Noirel, J., Ow, S.Y., Fazeli, A. (2012). A review of current proteomics technologies with a survey on their widespread use in reproductive biology investigations. *Theriogenology 77*(4), 738-765 e752.

Xiang, L.X., He, D., Dong, W.R., Zhang, Y.W., Shao, J.Z. (2010). Deep sequencing-based transcriptome profiling analysis of bacteria-challenged *Lateolabrax japonicus* reveals insight into the immune-relevant genes in marine fish. *BMC Genomics 11*, 472.

Yajima, Y., Vestergaard, M.C., Takagi, M. (2012). Damage in brain development by morpholino knockdown of zebrafish *dax1*. *J Biosci Bioeng 113*(6), 683-688.

Yamamoto, S., Shirakashi, S., Morimoto, S., Ishimaru, K., Murata, O. (2011). Efficacy of oral praziquantel treatment against the skin fluke infection of cultured chub mackerel, *Scomber japonicus*. *Aquaculture 319*(1-2), 53-57.

Yang, C.G., Wang, X.L., Tian, J., Liu, W., Wu, F., Jiang, M., Wen, H. (2013). Evaluation of reference genes for quantitative real-time RT-PCR analysis of gene expression in Nile tilapia (*Oreochromis niloticus*). *Gene 527*(1), 183-192.

Yang, X., Scheffler, B.E., Weston, L.A. (2006). Recent developments in primer design for DNA polymorphism and mRNA profiling in higher plants. *Plant Methods 2*(1), 4.

Yoshinaga, N., Shiraishi, E., Yamamoto, T., Iguchi, T., Abe, S., Kitano, T. (2004). Sexually dimorphic expression of a teleost homologue of Müllerian inhibiting substance during gonadal sex differentiation in Japanese flounder, *Paralichthys olivaceus*. *Biochem Biophys Res Commun 322*(2), 508-513.

Yoshizaki, G., Takeuchi, Y., Sakatani, S., Takeuchi, T. (2000). Germ cell-specific expression of green flurorescent protein in transgenic rainbow trout under control of the rainbow trout *vasa*-like gene promoter. *Int. J. Dev. Biol. 44*, 323-326.

Yuan, Y., Li, M., Hong, Y. (2014). Light and electron microscopic analyses of Vasa expression in adult germ cells of the fish medaka. *Gene 545*(1), 15-22.

Yufera, M., Halm, S., Beltran, S., Fuste, B., Planas, J.V., Martinez-Rodriguez, G. (2012). Transcriptomic characterization of the larval stage in gilthead seabream (*Sparus aurata*) by 454 pyrosequencing. *Mar Biotechnol (NY) 14*(4), 423-435.

Zhao, Y., Yang, Z., Phelan, J.K., Wheeler, D.A., Lin, S., McCabe, E.R. (2006). Zebrafish *dax1* is required for development of the interrenal organ, the adrenal cortex equivalent. *Mol Endocrinol 20*(11), 2630-2640.

Zhong, Q., Zhang, Q., Wang, Z., Qi, J., Chen, Y., Li, S., . . . Lan, X. (2008). Expression profiling and validation of potential reference genes during *Paralichthys olivaceus* embryogenesis. *Mar Biotechnol (NY) 10*(3), 310-318.

Ziv, T., Gattegno, T., Chapovetsky, V., Wolf, H., Barnea, E., Lubzens, E., Admon, A. (2008). Comparative proteomics of the developing fish (zebrafish and gilthead seabream) oocytes. *Comp Biochem Physiol Part D Genomics Proteomics 3*(1), 12-35.

Appendix I
Buffers and Solutions

0.1% DEPC H_2O

2 ml	DEPC
Make up to 2 L	mqH_2O

1X PBS : Phosphate buffered saline

8 g	NaCl
0.25 g	KCl
0.2 g	KH_2PO_4
1.15 g	Na_2HPO_4
Make up to 1 L	mQH_2O

50X TAE : Tris-acetate EDTA buffer

242 g	Tris base
57.1 ml	Glacial acetic acid
100 ml	0.5 M EDTA
Make up to 1 L	mQH_2O

1X TAE : Tris-acetate EDTA buffer

40 ml	50X TAE Buffer
1960 ml	mQH_2O

Lysogeny broth (LB) media

12.5 g	LB
Make up to 500 ml	mQH_2O

Mix solution before autoclaving. Store at 4°C.

Lysogeny broth (LB) agar (plates)

12.5 g	LB
7.5 g	Agar
Make up to 500 ml	mQH_2O

Mix solution before autoclaving. Incubate broth at 60°C for one hour. Add 5 µl of ampicillin (100 mg/ml) and mix by inversion. Pour LB agar into plates and allow to set.

Toluidine blue solution (Paraffin)

0.1 g	Toluidine blue dye
10 ml	H_2O

Toluidine blue solution (Cryostat)

0.28 g	Toluidine blue dye
0.40 g	Urea
28 ml	Ethanol (neat)
12 ml	mQH_2O

Dissolve urea in H_2O. Slowly mix in ethanol. Add toluidine blue and mix. Filter sterilise.

Hematoxylin solution

50 g	Alum
1 g	Hematoxylin dye
0.2 g	Sodium iodate
20 ml	Glacial acetic acid
1000 ml	mQH_2O

Dissolve alum in H_2O. Add hematoxylin until dissolved. Add sodium iodate and acetic acid. Bring to boil and cool. Filter sterilise.

Eosin 1% solution

1 g	Eosin Y
20 ml	mQH2O
80 ml	ethanol

6X Gel loading buffer

6 ml	Glycerol
25 mg	Bromophenol Blue
20 µl	Xylene Cyanole
Make up to 10 ml	mQH_2O

Appendix II
Nucleotide Sequences

This appendix contains nucleotide sequences of *S. lalandi*. All are consensus sequences constructed by BLAST searching the *S. lalandi* gonad transcriptome Ion Torrent RNA-Seq library, with the exception of *Cyp19a1a* which was sourced from the UniProt database, and *Actb* which was designed by Dr. Steve Bird prior to this study. Primer binding sites are highlighted in grey. Overlapping sites are highlighted in dark grey.

Vasa

GGAGGAGATGATGGAGTGTGTGAAAATGGGTTTAGAGGAGGAAGCCGAGGAGGACGAGGCAGCAGGGG
AGGACGAGGCTTCAGACAAGGTGGAGACCAGGGAGGCAGAGGAGGCTTTGGAGGAGGTTACAGAGGAA
AAGATGAACAGATCTTTGCTCGAGGAGAAGATAAAGATCCAGAGAAGAAGGATGGGAACGATGGTGAC
AGACCAAAGGTCACCTACGTCCCTCCAACCCTCCCTGAGGATGAGGACTCCATCTTTGCCCACTATGA
GTCAGGCATCAACTTCAACAAGTATGACGACATTCTGGTTGACGTCAGTGGAACCAACCCACCACAGG
CCATCATGACCTTTGCTGAGGCAGCACTGTGCGAGTCCCTGAGTAAAAACGTCAGTAAATCTGGTTAC
GTGAAGCCGACTCCTGTGCAGAAGCACGGCATCCCCATCATCTCTGCCGGCAGAGATCTCATGGCCTG
TGCCCAGACTGGATCTGGTAAAACGGCTGCGTTCCTGCTGCCCATCCTACAGCAGCTGATGGCAGACG
GCGTGGCAGCCAGTCAGTTCAGTGAACTGCAGGAGCCTGAAGCAATCATCGTCGCTCCAACCAGGGAG
CTCATCAACCAGATCTACCTGGAGGCCAGGAAGTTCGCCTTTGGGACCTGTGTGCGCCCAGTGGTGGT
TTATGGTGGAGTCAGCACTGGACACCAGATCAGGGAAATATGCAGAGGATGCAACGTCCTGTGTGGAA
CTCCAGGGAGACTGCTGGATGTGATTGGACGAGGAAAGGTGGGGCTCCACAAGCTGCGCTACCTGGTG
CTGGATGAGGCCGACCGTATGCTGGATATGGGCTTTGAGCCTGACATGCGCCGCCTGGTGGGCTCCCC
TGGAATGCCCTCCAAAGAGCAGCGTCAGACCCTGATGTTCAGCGCAACCTACCCCGAGGACATCCAGA
GGATGGCAGCTGACTTCCTCAAGACCGACTACCTGTTCTTAGCTGTGGGCGTGGTGGGCGGAGCTTGC
AGTGATGTGGAGCAGAAGTTTATTGAAGTAACTAAGTTCTCCAAGAGGGAGCAGCTTCTTGACATCCT
GAAGACAACAGGAACAGAGCGCACCATGGTGTTTGTGGAGACCAAGAGACAGGCTGATTTTATTGCCA
CATTCTTGTGCCAGGAGAAGGTTCCAACTACCAGCATCCATGGGGACCGGGAGCAGCGGGAACGGGAG
CAGGCCCTGGCAGACTTCCGCTCTGGGAGATGTCCAGTCCTGGTGGCAACCTCTGTTGCTGCCCGTGG
TCTGGATATTCCAGATGTTCAGCATGTGGTGAACTTTGACCTCCCCAACAATATAGACGAATACGTCC
ACCGTATTGGGAGAACTGGCCGCTGTGGAAACACTGGAAGGGCGGTGTCGTTCTATGACCCCGACAAT
GATGGACAGCTGGCTGGGTCCCTTGTCAGCATCCTGTCCAAGGCCCAGCAGGAAGTGCCCTCCTGGTT
AGAGGAGTGTGTGTTCAGCGGCTCAGGTGTGAACCCCTCCAGGAGGACCTTTGCGTCCACAGACTCCA
GGAAGGGTCCACAGGGCAGCTCCTTTCAGGACAGCAGTATGACGAGCCAGCCGGCTGTTCCTGCTGCA
GCTGATGATGAAGACTGGGAG

Amh

ATGAAGAAGAGCCTAAAAGACCTCCTCATTGGTGAAAAAACAGGAAGTACAATCAGCATAACTCTACT
TCTGCTTTTCTCTGGGGGAACGGGAACTGATACAAGACAAACACACGTTTCAGGTTCATCTCTGGCCT
CTTCACAGACCTTCTCCTTCCTCTGTGAGCTGAAACGGTTCCTAGGTGAAGTCCTGCCTCAGYACCAC
CCTGAGTCCCCTCCATTCCAGCTGGAGTCCTTACCCTCCATGCCACCCCTGTCACTGGGCTTATCCTC
CAGTGAGACCCTGCTGGCAGGACTGATCAACTCCTCTTTCCTTACTGTCTTCTCCTTCGCTAGATGGG
GCTCCATTTCTCAGGTGCATCCTGGACAGCTGGCCATGTCTCCTGCACTGGTGGAGGAGGTCAGGCAG
AGATTGGCAGATNTGNGGATGCAGATAATGGAAGTAATAAGGGAGAAGGAGGTGGGTCACAGGGCCAC
GGAGAGGCTGGAGAGGCTTCAAGAATTCAGTGCATTACCGACGATGGAACAAGCAGCAGGACGGAGCC
AGTACTGTGCGTTTCTTCTGCTGAAGGCCCTGCAGACAGTGGCCCGTGAGTACGAGCTGCAGAGAGGA
CTGCGAGCCACCAGAGCGGACACCAATAACCCGGTGAGGGGCAGTGTCTGCAGTGTGAGAAGCCTCAC
CGTCAACCTTGAAAGGCGTCTTGTGGGTCCAAGCGCTGCCAACATCAAAAACTGCCATGGCTCTTGTG
CTTTCCCCCTGACCAATGCAAACGACCATGTTGTCCTGCTCCACTTCCACATTGAGAGCGGGAATATG
GATGAGCGGGCTCCGTGCTGTGTGCCTGTTGCCTATGAACCCCTTGAGGTGGTGGGTTTGAATGAACA
TGGGACTTACCTCTCCATGATACCAGATATGGTAGCAAAGGAGTGTGGATGCCGC

Dmrt1

CCCCCAAGGATGCCCAAGTGCTCCCGCTGTAGGAACCACGGCTACGTGTCTCCCCTGAAGGGACACAA
GCGCTTCTGCAACTGGAAGGACTGCCAGTGTCCCAAGTGTAAACTGATAGCCGAGAGGCAGAGAGTCA
TGGCGGCCCAGGTCGCCTTGAGGAGGCAGCAGGCTCAGGAGGAGGAGCTCGGGATTTGTAGTCCGATG
ACTCTGTCGGTCCCTGAGGTGCTAGTGAAGAATGAAGCTGGAGCGGACTGCCTGTTCTCTGTGGAGGG
ACGATCCCTGACACCCACCAGCATATCCACTTCTGCTGTTGCTGTCACAGGGAGTCGCTCGGCATCCT
CCTCCAGCCCGTCAGCCGCCGCCAGGGCTCACACCGAGGGAGCGTCCGACCTGCTGATGGAAACCTCC
TATTACAATTTCTACCAGCCTCCACGCTACCCYTGCCTACTACAGCAACCTCTACAACTACCAGCAAT
ACCAGCAGGTGCCCCATGGTGATGGCCGACTGTCCGGCCACAACATGTCCTCTCCGTACCGCATGCAT
TCCTACTACCCAGCCAGTGTSKMYTCTGAYKYAGGGCCTGGGCTCCTCCACCTGTGTGCCGCCCCTCT
TCAGCCTGGAGGACAACAGTTACAACAACAACAACAACAACAACAACAACAATTGCTNCTGAAACCAT
GGCAGCCTCTTTCTCACCCAGCAGCATCACCGGCGCTCACGACTCCACCATGACCTGCGGGTCCATCA
TCTCCCTGGTTAACTCTGTCGTCAAGCCTGAGGCCAGCGACGAGACGGCCAACTTCTCCAACTCCAT

Dax1

CGGCTCCACACGGCGCCGAGGTGTCCTTCGCTCCCCGCAGGTGACGTGCAAAGCCGCGTCGGCGGTTC
TGGTGAAGACGCTGCGGTTCGTGNAAAAACGTCCCCCTGCTTTCGCGAGTTGCCGGAGGACGACCAAC
TGATGCTGATCCGAAGTGGCTGGGCACCCTTGCTCGTACTGGGGCTCGCACAGGACCGAGTGGACTTT
GAGACCACGGAGACCGTGGAGCCCAGCATGCTGCAGCGCATCCTCACGGGTTTACCGGACAGGCAGAG
CGAGGCTCTGGCCGGCCAGAGCAGCAGAGGAGCAGTCGGGGTCTCTGTCGTAGATATCGAAGCTATCA
AAGCTTTCCTGAAGAAGTGCTGGAGTGTGGATATCAGTACGAAGGAGTACGCGTATCTGAAAGGAGCC
GTGCTGTTCAACCCAGACCTGGAGGGCTTGNCGCTGCCTCCACTACATCCAGTCTCTGCGTCGGGAGG
CGCACCAGGCTCTGAACGAGCATGTCAGGCTGATCCATCGCGAAGACACAACACGGTTCGCCAAGCTG
CTCATAGCTCTGTCCATGCTGAGGGCCATCAGCCCGCCGGTGGTCGCCCAGCTCTTCTTCAGACCCGT
CAT

Sox9b

ATGGATAAAGTGAAGCGGCCCATGAACGCGTTCATGGTCTGGTCCCGGGGTCAGCGGCGCAAGATGGC
TCAAGAAAACCCCAAAATGCACAACTCGGAGATCAGTAAGCGGCTGGGGGCCGAGTGGAAGCTGCTGA
CGGACGCCGAGAAGCGGCCNCTTCATCGACGAGGCCAAGNCGGCTGCGGGCCGTCCACATGAAGGAGT
AYCCSCGAYTACAAGTACARGCCGCGCCGCAAGACAAANGCCGCTGCTCAAGAAGGACGCTCAGGTGG
CCAAGTACCCGCTGTCAGCGGGGAACCTGCTGGCGGCGGCGGCGCAGGGCCAGGGTGGTAGTCCGAGG
ATGGAGA

Cyp19a1a

ENA Accession Number: HQ449733

TGGATCAATGGAGAGGAGACCCTAATAATCAGCAGGCCATCAGCAGTGCACCATGTACTGAAGAATGG
ACTTTATTGTTCACGTTTTGGGAGCAAGCAGGGACTCAGCTGTGTTGGTATGAATGAGAGAGGCCTCA
TATTTAACAACAATGTCACTTTGTGGAAAAAGATGCGCACCTATTTCACCAAAGCCCTGACAGGTCCA
GGTTTGCAGAAGACAGTGGAGGTTTGCGTCTCCTCCACACAGACTCACCTGGACGACCTCCAGAGTCT
GAGTCACGTGGATGTCCTCGGCTTGCTGCGCTGCATCGTGGTCGACATCTCCAACAGACTCTTCCTGG
GGGTACCTGTCGATGAGAAAGAGCTGTTGGTGAAGATTCACAAGTATTTTGACACGTGGCAGAGTGTG
CTGATCAAACCGGACATCTACTTCAAGTTGGACTGGATTCACCGGCGGCACAAGACTGCAGCCCAGGA
GCTGCGAGATGCCATAGAGAGCCTTGTGGAGCAGAAGAGGAGGGATGTGGAGCAGGCAGATAAACTGG
ACAACATCAACTTCACCACAGAGCTCATATTTGCACAGAACCACGGCGAGCTGTCTGCTGAGAACGTG
TTGCAGTGTGTGTTGGAGATG

Actb

ATGTGCAAAGCCGGATTCGCCGGAGACGACGCCCCTCGTGCTGTCTTCCCCTCCATCGTTGGTCGCCC
CAGGCATCAGGGAGTGATGGTGGGTATGGGCCAGAAGGACAGCTACGTTGGTGATGAAGCCCAGAGCA
AGAGAGGTATCCTGACCCTGAAGTACCCCATCGAGCACGGTATTGTGACCAACTGGGATGACATGGAG
AAGATCTGGCATCACACCTTCTACAACGAGCTGAGAGTTGCCCCTGAGGAGCACCCTGTCCTGCTCAC
AGAGGCCCCCCTGAACCCYAARGCCAACAGGGAGAAGATGACCCAGATCATGTTCGAGACCTTCAACA
CCCCYGCCATGTACGTTGCCATCCAGGCTGTGCTGTCCCTGTATGCCTCTGGTCGTACCACYGGTATT
GTCATGGACTCYGGTGATGGTGTGACCCACACAGTGCCCATCTACGAGGCTACGCCCTGCCCCNACGC
CATCCTGNCGTCTGGACTTGGCCGGCMGCGACCTCACAGACTACCTCATGAAGATCCTGACAGAGCGT
GGNCTACTCCTTCACCACCACAGCYNGAGMGKGARATYGTGCGTGRCMWWRRSGAGAAGCTGTGCTAC
GTCGCCCTGGACTTCGAGCAGGAGATGGGTACCGCTGCCTCCTCCTCMTCCMTGGAGAAGAGCTACGA
GCTGCCTGACGGACAGGTCATCACCATCGGCAATGAGAGGTTCCGTTGCCCAGAGGCCCTCTTCCAGC
CTTCCTTCCTCGGTATGGAGTCNNCGTGCGGAATCCACGAGACCACCTACAACAGCATCATGAAGTGC
GACGTCGACATCCGTAAGGACCTGTACGCCAACACCGTGCTGTCTGGAGGTACCACCATGTACCCCGG
CATCGCTGACAGGATGCAGAAGGAGATCACAGCCCTGGCCCCATCCACCATGAAGATCAAGATCATTG
CCCCACCAGAGCGTAAATACTCTGTCTGGATCGGAGGCTCCATCCTGGCTTCCCTGTCCACCTTCCAG
CAGATGTGGATCAGCAAGCAGGAGTACGATGAGTCCGGYCCCTNCCATCGTTGTGTTTTTGAGTTGCT
TTAGTATGGGGGAGAGC

Appendix III
Ladders and Vector Maps

This appendix contains details of the DNA ladders and pLUG-Prime® cloning vector used in this study.

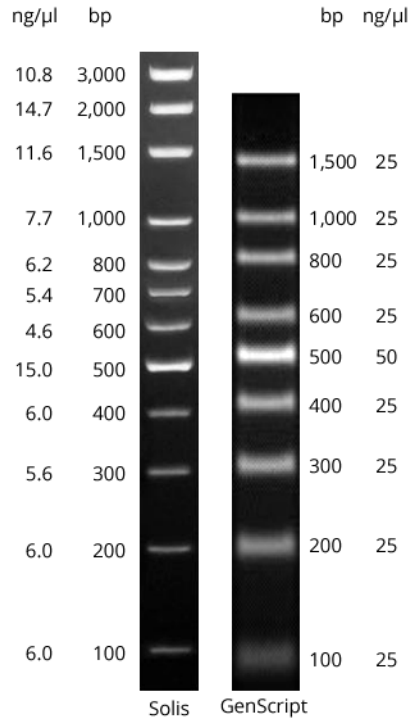
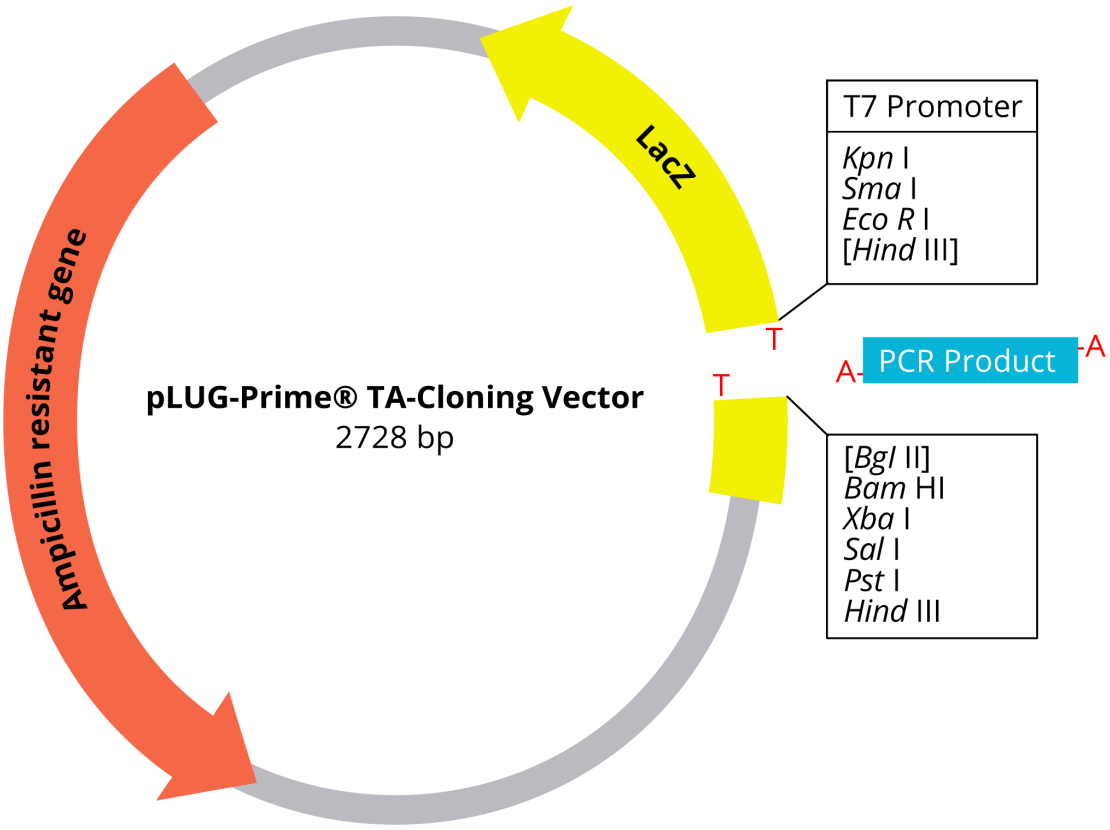


Figure 85. Solis Biodyne (left) and GenScript (right) 100 bp DNA ladders used in this study. Original images were retrieved from *http://www.genscript.com/dna_ladders.html* and *https://www.sbd.ee/EN/products/dna_ladders/100_bp_dna_ladder_*




```
301  TACGCCAGCT   GGCGAAAGGG   GGATGTGCTG   CAAGGCGATT   AAGTTGGGTA
     ATGCGGTCGA   CCGCTTTCCC   CCTACACGAC   GTTCCGCTAA   TTCAACCCAT

                M13 Forward Primer(-40)    M13 Forward Primer(-20)
351  ACGCCAGGGT   TTTCCCAGTC   ACGACGTTGT   AAAACGACGG   CCAGTGAATT
     TGCGGTCCCA   AAAGGGTCAG   TGCTGCAACA   TTTTGCTGCC   GGTCACTTAA

                          T7 Promoter          KpnI      SmaI      EcoRI     HindIII
401  GTAATACGAC   TCACTATAGG   GCGAGCTCGG   TACCCGGGCG   AATTCCAAGC
     CATTATGCTG   AGTGATATCC   CGCTCGAGCC   ATGGGCCCGC   TTAAGGTTCG

                        BglII      BamHI      XbaI       SalI       PstI
451  TT  Insert-  AGATCTGGAT   CCCCTCTAGA   GTCGACCTGC   AGGCATGCAA
     AA  - DNA    TCTAGACCTA   GGGGAGATCT   CAGCTGGACG   TCCGTACGTT

     HindIII
493  CGTTGGCGTA   ATCATGGTCA   TAGCTGTTTC   CTGTGTGAAA   TTGTTATCCG
     GCAACCGCAT   TAGTACCAGT   ATCGACAAAG   GACACACTTT   AACAATAGGC
                M13 Reverse Primer
```

Figure 86. Simplified map and sequence reference points of the pLUG Prime® TA-Cloning Vector. This figure was recreated from the iNtRON Biotechnology protocol for this kit, available online at *http://www.labotaq.com/protocolo/11061.pdf*.

Appendix IV
Productivity Bottlenecks within Aquaculture

IV.1 Diseases

Poor fish health and infectious diseases are major threats for many fish farms, with the risk of transmission increasing significantly when stocking densities are high. The susceptibility to particular diseases varies across species and geographic location. Furthermore, disease can take a number of forms including viral, bacterial, fungal and parasitic. Intensive culture methods in aquaculture frequently lead to outbreaks of disease and present a suite of issues.

For instance, the global Atlantic salmon farming industry has been stifled by an epidemic of infectious viruses and parasites. Infectious salmon anaemia virus (ISAV) has been particularly problematic in every country which farms salmon, including Australia and New Zealand, with many strains causing internal bleeding which results in substantial fish mortalities (Mardones *et al.*, 2011; Austbo *et al.*, 2014). Although research into this disease continues, there is currently no effective treatment for infected fish other than prompt removal of mortalities to decrease the spreading rate (Vike *et al.*, 2014).

Similarly, amoebic gill disease (AGD) has been a notorious issue for salmon aquaculture, exacerbated by co-infection alongside sea lice (*Caligus rogercresseyi*) (Bustos *et al.*, 2011). Freshwater and chemical baths have proven effective for short term treatment, but have associated problems including handling-induced stress, bath solution toxicity, and increased susceptibility to re-infection post-bathing (Florent *et al.*, 2007b; Yamamoto *et al.*, 2011). Options remain limited for long term

control, although orally delivered parasiticides (*e.g.* bithionol) have been shown to reduce the intensity of the disease (Florent *et al.*, 2007a).

Another particularly problematic pathogen for salmon aquaculture is the oomycete *Saprolegnia*, a fungi-like organism which causes saprolegniosis in a variety of fish. The disease is characterised in adult fish by destruction of epidermal tissue, lethargy and eventually respiratory failure. Eggs are also susceptible to the disease, as shown in Figure 87 (van den Berg *et al.*, 2013). Unlike viruses and bacteria, oomycetes like *Saprolegnia* are eukaryotic; this can make it difficult to develop treatments which are both effective against the pathogen but do not affect the host. Malachite Green was an effective treatment for destroying *Saprolegnia* for some time, but was banned in 2002 due to its carcinogenic properties (Fugelstad *et al.*, 2009). Formaldehyde has also been widely used and is quite effective, but it is highly toxic to fish and humans. In addition, even following treatment, *Saprolegnia* has been demonstrated to persist as biofilms and infections can reoccur (Ali *et al.*, 2013). It has been purported that formaldehyde is also likely to be banned (van den Berg *et al.*, 2013), stressing the need to develop new treatment methods.

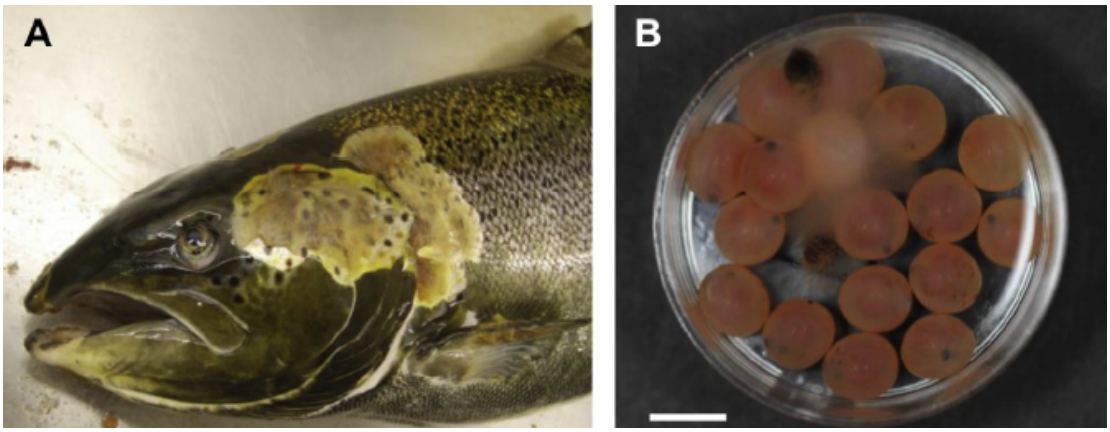


Figure 87. Photographs of *Saprolegnia* infection, adapted from van den Berg *et al.* (2013). (A) shows an Atlantic salmon with late stage saprolegniosis. (B) shows hyphal growth also visible on Atlantic salmon eggs.

There are numerous indirect effects of disease which bolster the impetus to develop more effective treatments. Mass mortality can occur in hatcheries by intracohort cannibalism when larval and juvenile fish are afflicted by diseases; this has been observed in groupers afflicted by viral necrosis in Southeast Asia (Harikrishnan *et al.*, 2010). Furthermore, non-lethal effects of diseases may affect the cosmetic appearance, rendering surviving fish unmarketable (Hutson *et al.*, 2007). For species reared in cages, the potential of disease transmission between farmed and wild fish stocks presents a risk of ongoing re-infection and can also foster animosity from conservation groups, which can lead to bureaucratic resistance for both new and existing farm operators (Bondad-Reantaso *et al.*, 2005). Treating and preventing diseases in aquaculture is therefore essential to keep operations viable.

IV.2 Genetic Diversity

The viability of hatchery selective breeding programmes can be threatened by some management practices. In addition to the inherent reduction of genetic diversity which results from the typically small size of broodstock populations relative to wild populations, some practices can further restrict genetic diversity and lead to inbreeding depression in later generations (Doupé & Recher, 1999; Frost *et al.*, 2006). For example, barramundi hatcheries in Australia routinely grade fish by size, culling both large and small fish to retain a relatively even distribution of intermediate-sized fish. Although this practice dramatically reduces intra-specific cannibalism and aggressive dominance, thereby resulting in more even growth rates, it can affect the representation of families in the gene pool and alter the outcomes of selected breeding programmes (Frost *et al.*, 2006). Furthermore, the same study revealed that inputs to mass spawning events could be strongly skewed, with certain individuals dominating in genetic contribution. Frost *et al.* (2006) postulated that this may be due to both behavioural differences during mating, and also sperm competition as a result of sperm quality differences between individual males. Particular

consideration must be given to such factors when managing selective breeding programmes if genetic diversity is to be conserved in subsequent generations of broodstock.

IV.3 Diets

Many commonly cultured fish species are carnivorous and rely on a diet containing fishmeal derived from capture fisheries. Although a proportion of fishmeal is created from commercial bycatch and the byproducts of fish processing, generally it is derived from small pelagic fish species (Péron *et al.*, 2010). Many of these fisheries are transitioning away from fishmeal and towards direct human consumption; for example, the majority of Atlantic herring (*Clupea harengus*) caught in Norway and Chilean jack mackerel (*Trachurus murphyi*) caught in Chile are now sold for human consumption, where traditionally they were used for fishmeal and fish oil production (Olsen & Hasan, 2012). Of the annual seafood yield of capture fisheries globally, approximately 23% becomes fishmeal, of which an estimated 70% is used for commercial aquaculture feeds (Deutsch *et al.*, 2007). As aquaculture is expected to continue growing rapidly into the future, sustainable solutions will need to be found to address increases in fishmeal demand.

Alternative protein sources to fishmeal have been a subject of research for a number of years, with significant progress made towards plant-based ingredients. In order to be a viable option for aquaculture feeds, plant-based ingredients must be widely available, competitively priced, easily transported and easily processed (Gatlin *et al.*, 2007). A number of studies on rainbow trout (*Oncorhynchus mykiss*) have determined that completely plant-based feeds can yield equal or even greater growth performance (Gaylord & Barrows, 2009; Snyder *et al.*, 2012; J. Davidson *et al.*, 2013; Hauptman *et al.*, 2014; Read *et al.*, 2014). However, the nutritional differences between plant-based and fishmeal-based feeds affected the suitability of the former

for some species, particularly those which are strongly adapted for carnivorous diets (Olsen & Hasan, 2012). The profiles of nutrients and anti-nutrients in plant-based ingredients are generally not ideal for aquaculture feeds. Fibre and starch content must be low, while protein content, digestibility and palatability must be high (Gatlin *et al.*, 2007). Some anti-nutritional compounds present in plant-based feeds can be problematic; lectins, for example, are bioactive proteins found in soybeans and a range of plants which can have numerous physiological effects. Lectins can interfere with nutrient absorption by binding with intestinal glycoproteins, and can disrupt cell membranes to trigger immune responses which can cause cell mortality (Gatlin *et al.*, 2007). Anti-nutrients and their effects need to be thoroughly researched and understood if plant-based diets are to be successful in aquaculture.

There have been gains made in the field of nutritional augmentation of plant-based ingredients. Selective breeding and genetic modification have yielded notable advancements in plant traits to improve their suitability for aquaculture feeds, such as low phytic acid content (an unusable form of phosphorous for monogastric animals), high lysine (an amino acid often limited in plants) and oil content, both low and high β-glucan levels, and a range of starch structures to suit the needs of various diets (Gatlin *et al.*, 2007). However, these approaches are often unfavourable due to public resistance towards genetic manipulation. Processing technologies can also alter and improve the composition of feeds. Microbial fermentation, enzyme treatments and bioactive chemical supplementation are powerful tools for increasing protein, adding desired nutrients and removing antinutrients (Bairagi *et al.*, 2004; Olsen & Hasan, 2012). Air classification is a mechanical modification technique which separates proteins from carbohydrates using air pressure; this has been shown to increase protein levels of rice and barley, as much as doubling in the latter (Gatlin *et al.*, 2007). These technologies will become increasingly important as the demand for plant-based aquaculture feeds grows.

Providing a diet which closely matches the nutritional requirements of any domesticated species is a basic principle of animal husbandry. The inclusion of

performance-boosting components outside of the basic requirements for survival and growth is referred to as a functional feed. Common examples of additives include vitamins, nucleotides, prebiotics, probiotics, and immunostimulants as mentioned previously. Probiotics typically consist of bacteria which improve gastrointestinal performance and can have a notable effect on fish immunity, as has been evidenced in Atlantic cod by increased immune response gene expression (Lazado & Caipang, 2014). Increases of inflammatory cytokine activity in response to the probiotic *Carnobacterium maltaromaticum* have been observed in rainbow trout (Kim & Austin, 2006). Additionally, interactions have been noted between probiotics and various immune cells including monocytes, macrophages and neutrophils (Tacchi *et al.*, 2011a). This activity can be further enhanced with oligosaccharide prebiotics which are metabolised by probiotics and other health-promoting bacteria. Functional feeds can greatly improve productivity and have become an important part of modern aquaculture.

Understanding how fish respond to functional feeds on the molecular level is important for optimising feed composition. Tacchi *et al.* (2011a) conducted a transcriptome analysis of Atlantic salmon to investigate the response to a diet supplemented with a range of nucleotides, prebiotics and vitamins. Improvements in feed conversion ratio were detected, as were significant changes to liver processes. No changes to gene expression were detected in white skeletal muscle. Gene ontogeny analysis revealed changes to processes related to immune parameters and oxygen transport, effectively down regulating the immune system and redirecting energy towards growth. Other studies have begun to have successful results with immunostimulants, such as the successful reduction of sea lice burden on Atlantic salmon (Covello *et al.*, 2012). This study linked two immunostimulants with changes in expression of inflammatory genes, but complete characterisation of the signalling pathways remains elusive. Such pathways should be a strong focus of future molecular research to gain the fine-scale information to quantify vaccine potency and efficacy (Tafalla *et al.*, 2013).

Appendix V
Real-Time PCR Results & Statistics

This appendix details the real-time PCR results obtained for the target housekeeping and sex differentiation genes in this study, and the statistical tests used to analyse the significance of the results.

Table 20. Real-time PCR results for the housekeeping genes β-actin and *Gapdh*, showing raw and mean C_T values.

β-actin	Name	C_T	Mean	*Gapdh*	Name	C_T	Mean
1	H1	23.55		1	H1	33.65	
2	H1	23.82	23.685	2	H1	34.37	34.01
3	H2	35.82		3	H2	25.9	
4	H2	34.46	35.14	4	H2	25.84	25.87
5	H3	28.77		5	H3	32.73	
6	H3	29.01	28.89	6	H3	32.61	32.67
7	H4	32.49		7	H4	30.89	
8	H4	32.25	32.37	8	H4	30.73	30.81
9	3.1	23.15		9	3.1	23.02	
10	3.1	23.3	23.225	10	3.1	23.15	23.085
11	3.2	22.04		11	3.2	24.48	
12	3.2	22.07	22.055	12	3.2	24.46	24.47
13	3.3	22.59		13	3.3	23.46	
14	3.3	22.82	22.705	14	3.3	23.45	23.455
15	3.4	28.99		15	3.4	29.28	
16	3.4	28.94	28.965	16	3.4	29.39	29.335
17	12.1	33.28		17	12.1	32.54	
18	12.1	33.08	33.18	18	12.1	32.58	32.56
19	12.2	34.42		19	12.2	32.04	
20	12.2	34.85	34.635	20	12.2	34.11	33.075
21	12.3	35.56		21	12.3	26.7	
22	12.3	35.09	35.325	22	12.3	26.82	26.76
23	12.4	27.65		23	12.4	25.52	
24	12.4	27.84	27.745	24	12.4	25.53	25.525
25	18.1	17.8		25	18.1	21.76	

26	18.1	17.96	17.88	26	18.1	21.56	21.66
27	18.2	15.94		27	18.2	20.36	
28	18.2	16.12	16.03	28	18.2	20.23	20.295
29	18.3	16.26		29	18.3	19.75	
30	18.3	16.02	16.14	30	18.3	19.72	19.735
31	18.4	16.1		31	18.4	20.06	
32	18.4	16.19	16.145	32	18.4	20.2	20.455

Table 21. Real-time PCR results for the sex differentiation genes *Vasa* and *Amh*, showing raw and mean C_T values.

Vasa	Name	C_T	Mean	*Amh*	Name	C_T	Mean
1	H1	32.76		1	H1	33.56	
2	H1	32.59	32.675	2	H1	32.69	33.125
3	H2	37		3	H2	31.3	
4	H2	37.61	37.305	4	H2	33.86	32.58
5	H3	34.95		5	H3	32.39	
6	H3	35.85	35.4	6	H3	31.41	31.9
7	H4	36.87		7	H4	36.09	
8	H4	41.6	39.235	8	H4	35.05	35.57
9	3.1	32.48		9	3.1	32.72	
10	3.1	31.89	32.185	10	3.1	33.14	32.93
11	3.2	34.66		11	3.2	32.69	
12	3.2	34.68	34.67	12	3.2	32.6	32.645
13	3.3	33.23		13	3.3	30.1	
14	3.3	33.39	33.31	14	3.3	30.05	30.075
15	3.4	37.11		15	3.4	34.91	
16	3.4	36.7	36.905	16	3.4	34.96	34.935
17	12.1	33.24		17	12.1	33.29	
18	12.1	34.13	33.685	18	12.1	34.87	34.08
19	12.2	39.94		19	12.2	34.13	
20	12.2	40.04	39.99	20	12.2	33.44	33.785
21	12.3	38.93		21	12.3	34.42	
22	12.3	38.77	38.85	22	12.3	33.55	33.985
23	12.4	35.99		23	12.4	36.16	
24	12.4	38.3	37.145	24	12.4	36.41	36.285
25	18.1	32.72		25	18.1	31.57	
26	18.1	34.51	33.615	26	18.1	30.8	31.185
27	18.2	31.72		27	18.2	28.95	
28	18.2	31.44	31.58	28	18.2	29.23	29.09
29	18.3	30.87		29	18.3	29.09	
30	18.3	31.44	31.155	30	18.3	29.04	29.065

31	18.4	31.81		31	18.4	29.29	
32	18.4	31.62	31.715	32	18.4	29.04	29.62625

Table 22. Real-time PCR results for the sex differentiation gene *Cyp19a1a*, showing raw and mean C_T values.

Cyp19a1a	Name	C_T	Mean
1	H1	34.29	
2	H1	32.52	33.405
3	H2	35.64	
4	H2	35	35.32
5	H3	36.78	
6	H3	38.28	37.53
7	H4	34.24	
8	H4	34.17	34.205
9	3.1	32.63	
10	3.1	32.21	32.42
11	3.2	33.79	
12	3.2	33.83	33.81
13	3.3	32.16	
14	3.3	31.73	31.945
15	3.4	37.49	
16	3.4	36.65	37.07
17	12.1	32.87	
18	12.1	32.55	32.71
19	12.2	34.73	
20	12.2	33.37	34.05
21	12.3	33.7	
22	12.3	32.29	32.995
23	12.4	40.39	
24	12.4	40.25	40.32
25	18.1	32.35	
26	18.1	32.99	32.67
27	18.2	32.75	
28	18.2	32.68	32.715
29	18.3	31.35	
30	18.3	31.51	31.43
31	18.4	33.67	
32	18.4	33.1	33.385

Table 23. Summary of average C_T values obtained from Table 20, 21 and 22.

Name	β-actin	*Gapdh*	Mean	*Amh*	*Cyp19a1a*	*Vasa*
H1	23.685	34.01	28.3818	33.125	33.405	32.675
H2	35.14	25.87	30.15082	32.58	35.32	37.305
H3	28.89	32.67	30.72192	31.9	37.53	35.4
H4	32.37	30.81	31.58037	35.57	34.205	39.235
3.1	23.225	23.085	23.15489	32.93	32.42	32.185
3.2	22.055	24.47	23.23114	32.645	33.81	34.67
3.3	22.705	23.455	23.07695	30.075	31.945	33.31
3.4	28.965	29.335	29.14941	34.935	37.07	36.905
12.1	33.18	32.56	32.86854	34.08	32.71	33.685
12.2	34.635	33.075	33.84601	33.785	34.05	39.99
12.3	35.325	26.76	30.74568	33.985	32.995	38.85
12.4	27.745	25.525	26.61186	36.285	40.32	37.145
18.1	17.88	21.66	19.67945	31.185	32.67	33.615
18.2	16.03	20.295	18.03687	29.09	32.715	31.58
18.3	16.14	19.735	17.84721	29.065	31.43	31.155
18.4	16.145	20.455	18.17267	29.62625	33.385	31.715

Table 24. ΔCq values obtained using values from Table 23.

Name	*Amh*	*Cyp19a1a*	*Vasa*
H1	0.037338	0.030752	0.051006
H2	0.185671	0.027792	0.007021
H3	0.441939	0.008924	0.039062
H4	0.062951	0.162146	0.004963
Average	**0.181975**	**0.057404**	**0.025513**
Standard Error	**0.092503**	**0.035247**	**0.011539**
3.1	0.001141	0.001625	0.001913
3.2	0.001466	0.000654	0.00036
3.3	0.007823	0.00214	0.000831
3.4	0.018129	0.004127	0.004627
Average	**0.00714**	**0.002137**	**0.001933**
Standard Error	**0.003973**	**0.000732**	**0.000955**
12.1	0.431831	1.116156	0.567833
12.2	1.043198	0.868148	0.014141
12.3	0.105893	0.210324	0.003634
12.4	0.001225	7.47E-05	0.000675
Average	**0.395537**	**0.548676**	**0.146571**

Standard Error	**0.234549**	**0.264503**	**0.14045**
18.1	0.000344	0.000123	6.38E-05
18.2	0.000471	3.81E-05	8.38E-05
18.3	0.00042	8.15E-05	9.86E-05
18.4	0.000357	2.63E-05	8.38E-05
Average	**0.000398**	**6.72E-05**	**8.25E-05**
Standard Error	**2.94E-05**	**2.2E-05**	**7.14E-06**

Table 25. Summary of ΔCq values for *Vasa*.

Groups	Count	Sum	Average	Variance
H	4	0.102051	0.025513	0.000533
Day 3	4	0.007731	0.001933	3.65E-06
Day 12	4	0.586282	0.146571	0.078905
Day 18	4	0.00033	8.25E-05	2.04E-10

Table 26. ANOVA of ΔCq values for *Vasa*.

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.05824	3	0.019413	0.977489	**0.435614**	3.490295
Within Groups	0.238325	12	0.01986			
Total	0.296565	15				

Table 27. Summary of ΔCq values for *Amh*.

Groups	Count	Sum	Average	Variance
H	4	0.727899	0.181975	0.034227
Day 3	4	0.028559	0.00714	6.31E-05
Day 12	4	1.582147	0.395537	0.220053
Day 18	4	0.001591	0.000398	3.46E-09

Table 28. ANOVA of ΔCq values for *Amh*.

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.416179	3	0.138726	2.18172	**0.143121**	3.490295
Within Groups	0.763029	12	0.063586			
Total	1.179208	15				

Table 29. Summary of ΔCq values for *Cyp19a1a*.

Groups	Count	Sum	Average	Variance
H	4	0.229614	0.057404	0.004969
Day 3	4	0.008547	0.002137	2.14E-06
Day 12	4	2.194702	0.548676	0.279847
Day 18	4	0.000269	6.72E-05	1.94E-09

Table 30. ANOVA of ΔCq values for *Cyp19a1a*.

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.84737	3	0.282457	3.966828	**0.035407**	3.490295
Within Groups	0.854456	12	0.071205			
Total	1.701827	15				

Table 31. Students t-test (two-sample assuming equal variances) for *Cyp19a1a*, between hatch (H) and 3 dph.

	H	3 dph
Mean	0.057404	0.002137
Variance	0.004969	2.14E-06
Observations	4	4
Pooled Variance	0.002486	
Hypothesized Mean Difference	0	
df	6	
t Stat	1.567636	
P(T<=t) one-tail	0.084006	
t Critical one-tail	1.94318	
P(T<=t) two-tail	**0.168012**	
t Critical two-tail	2.446912	

Table 32. Students t-test (two-sample assuming equal variances) for *Cyp19a1a*, between hatch (H) and 12 dph.

	H	**12 dph**
Mean	0.057404	0.548676
Variance	0.004969	0.279847
Observations	4	4
Pooled Variance	0.142408	
Hypothesized Mean Difference	0	
df	6	
t Stat	-1.84107	
P(T<=t) one-tail	0.057604	
t Critical one-tail	1.94318	
P(T<=t) two-tail	**0.115208**	
t Critical two-tail	2.446912	

Table 33. Students t-test (two-sample assuming equal variances) for *Cyp19a1a*, between hatch (H) and 18 dph.

	H	**18 dph**
Mean	0.057404	6.72E-05
Variance	0.004969	1.94E-09
Observations	4	4
Pooled Variance	0.002485	
Hypothesized Mean Difference	0	
df	6	
t Stat	1.626685	
P(T<=t) one-tail	0.077464	
t Critical one-tail	1.94318	
P(T<=t) two-tail	**0.154928**	
t Critical two-tail	2.446912	

Table 34. Students t-test (two-sample assuming equal variances) for *Cyp19a1a*, between 3 dph and 12 dph.

	3 dph	**12 dph**
Mean	0.002137	0.548676
Variance	2.14E-06	0.279847
Observations	4	4
Pooled Variance	0.139925	
Hypothesized Mean Difference	0	
df	6	
t Stat	-2.06628	
P(T<=t) one-tail	0.04216	
t Critical one-tail	1.94318	
P(T<=t) two-tail	**0.08432**	
t Critical two-tail	2.446912	

Table 35. Students t-test (two-sample assuming equal variances) for *Cyp19a1a*, between 3 dph and 18 dph.

	3 dph	**18 dph**
Mean	0.002137	6.72E-05
Variance	2.14E-06	1.94E-09
Observations	4	4
Pooled Variance	1.07E-06	
Hypothesized Mean Difference	0	
df	6	
t Stat	2.827308	
P(T<=t) one-tail	0.015032	
t Critical one-tail	1.94318	
P(T<=t) two-tail	**0.030064**	
t Critical two-tail	2.446912	

Table 36. Students t-test (two-sample assuming equal variances) for *Cyp19a1a,* between 12 dph and 18 dph.

	12 dph	**18 dph**
Mean	0.548676	6.72E-05
Variance	0.279847	1.94E-09
Observations	4	4
Pooled Variance	0.139924	
Hypothesized Mean Difference	0	
df	6	
t Stat	2.074111	
P(T<=t) one-tail	0.041706	
t Critical one-tail	1.94318	
P(T<=t) two-tail	**0.083412**	
t Critical two-tail	2.446912	