Investigation of candidate sex determination and sex differentiation genes in sea lamprey, *Petromyzon marinus*, and Pacific lamprey, *Entosphenus tridentatus*

By

Arfa A. Khan

A thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
In partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Biological Sciences
University of Manitoba
Winnipeg, Manitoba

Copyright © 2017 by Arfa Khan
Abstract

Genes associated with sex differentiation (the process by which undifferentiated gonads develop into testes or ovaries) are relatively conserved among vertebrates, but those associated with sex determination (the genetic and/or environmental “switch” that activates differentiation) are more variable, particularly in fishes. Virtually nothing is known regarding the genetic basis of these processes in lampreys, one of the two extant groups of jawless fishes. Therefore, I tested whether 19 candidate sex determination genes identified from other vertebrates showed sex specific sequence differences in sea lamprey, *Petromyzon marinus*, and Pacific lamprey, *Entosphenus tridentatus*; 11 and seven genes amplified in sea and Pacific lampreys, respectively, but none showed sex specific differences. I also used qRT-PCR to measure expression of seven candidate sex differentiation genes (e.g., *SOX9, DMRT1, WT1, DAZAPI*) prior to and during testicular differentiation in sea lamprey, and found that expression patterns were consistent with their presumed role in other vertebrates.
Acknowledgements

I would like to thank my advisor, Dr. Margaret Docker, for all her support. I couldn’t have done it without her mentoring and passion for learning more about the mystery that is lampreys. You have been influential in my growth as a student during my time at the University of Manitoba and have motivated me every step of the way.

I would also like to thank my committee, Drs. Alberto Civetta and Colin Garroway, for their comments and suggestions that have contributed greatly to the development of my thesis.

I thank all my past and present lab mates for their help along the way. I am especially grateful to Erin Spice for teaching me lab techniques and her constant support. Special thanks to Nisha Ajmani for teaching me how to navigate through the sea lamprey genome and how to use different Ensembl features and tools.

A very special thanks to our collaborators, Drs. Richard Manzon (University of Regina), Michael Wilkie and Jonathan Wilson and their students (Wilfrid Laurier University), John Hume (Michigan State University) and Dr. Jon Hess (Columbia River Inter-Tribal Fish Commission), for lamprey sampling.

This research was made possible by funding from the University of Manitoba Graduate Enhancement of Tri-council Stipends (GETS) and Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant.

Finally and most importantly, I would like to thank my family, husband and friends for their encouragement and patience. I am very thankful for my parents’ courage leaving the comfort of their families and home, and immigrating with my brother, sister and me to Canada 17 years
ago. We have been blessed with wonderful opportunities which we could have only dreamt of; thank you for everything Mama and Papa.
Table of Contents
Abstract .................................................................................................................. ii
Acknowledgements .................................................................................................. iii
Table of Contents .................................................................................................... v
List of Tables ............................................................................................................ vii
List of Figures .......................................................................................................... ix

1. Chapter 1: General introduction ................................................................. 1
   1.1. Lamprey biology: life-cycle .................................................................... 2
   1.2. Lamprey biology: conservation and control .......................................... 6
       1.2.1. Control of the invasive sea lamprey in the Great Lakes ..................... 6
       1.2.2. Conservation of the Pacific lamprey ................................................. 7
       1.2.3. Conservation of the anadromous sea lamprey in their native range ... 9
   1.3. Sex determination in vertebrates ............................................................. 12
       1.3.1. Genetic basis of sex determination of model vertebrates ............... 14
       1.3.2. Sex determination in fish ................................................................. 16
       1.3.3. Sex determination in lampreys ......................................................... 19
   1.4. Genes involved in sex differentiation ..................................................... 21
       1.4.1. Sex differentiation in model vertebrates ............................................ 21
       1.4.2. Sex differentiation in fish ................................................................. 22
       1.4.3. Sex differentiation in lampreys ......................................................... 25
   1.5. Testicular differentiation ......................................................................... 28
       1.5.1. Spermatogenesis ............................................................................ 28
       1.5.2. Testicular differentiation in lampreys ............................................... 29
   1.6. Candidate genes for sex determination in vertebrates and conserved sex differentiation genes in lampreys ................................................................. 32
   1.7. Thesis objectives ..................................................................................... 42
   1.8. Significance ............................................................................................. 43
   1.9. Literature cited ........................................................................................ 46

   2.1. Abstract .................................................................................................... 76
   2.2. Introduction .............................................................................................. 78
   2.3. Methods .................................................................................................... 81
       2.3.1. Sample collection ............................................................................ 81
       2.3.2. DNA isolation and RNase elimination ............................................. 82
       2.3.3. Identifying target genes and PCR .................................................... 83
       2.3.4. Sequencing preparation and analysis .............................................. 85
           2.3.4.1. Propanol precipitate, cycle sequencing, and ethanol precipitation . 85
           2.3.4.2. Sequencing alignment and analysis ......................................... 87
   2.4. Results ...................................................................................................... 88
   2.5. Discussion ............................................................................................... 90
   2.6. Literature cited ........................................................................................ 103
3. Chapter 3: Gene expression in sea lamprey, *Petromyzon marinus* during testis differentiation and development .................................................. 187

3.1. Abstract .................................................................................................. 187
3.2. Introduction ............................................................................................. 188
  3.2.1. Stages of testicular differentiation and development in lampreys .......... 190
  3.2.2. Gene expression during testicular differentiation .............................. 194
3.3. Methods .................................................................................................. 197
  3.3.1. Sample collection and preservation ................................................ 197
  3.3.2. Identification of conserved testis differentiation genes and reference genes .......................................................... 199
  3.3.3. RNA isolation and cDNA synthesis ................................................ 201
  3.3.4. Quantitative reverse-transcriptase polymerase chain reaction qRT-PCR ...................................................... 202
  3.3.5. Gene expression analysis ............................................................... 202
  3.3.6. Histological sampling and preparation .......................................... 203
  3.3.7. Histological analysis ...................................................................... 205
3.4. Results .................................................................................................... 206
3.5. Discussion .............................................................................................. 209
3.6. Literature cited ...................................................................................... 219
3.7. Tables and Figures .................................................................................. 229

Appendices ..................................................................................................... 255

4. Chapter 4: Conclusion and future directions ............................................ 267

4.1. Contributions of this study .................................................................... 267
4.2. Study limitations .................................................................................... 269
4.3. Future directions .................................................................................... 271
4.4. Conclusion ............................................................................................. 274
4.5. Literature cited ...................................................................................... 276
List of Tables

Table 2.1. Sampling location for Pacific lamprey and sea lamprey used for sequencing (111)

Table 2.2. List of DMRT1 amino acid sequences used for constructing comparison alignment (113)

Table 2.3. List of DMRTA2 amino acid sequences used for constructing comparison alignments (114)

Table 2.4. List of DMRT2A amino acid sequences used for constructing comparison alignments (115)

Table 2.5. List of FEM1c amino acid sequences used for constructing comparison alignments (116)

Table 2.6. List of FOXL2 amino acid sequences used for constructing comparison alignments (117)

Table 2.7. List of Rspo1 amino acid sequences used for constructing comparison alignments (118)

Table 2.8. List of NR5A1 amino acid sequences used for constructing comparison alignments (119)

Table 2.9. List of SOX2 and SOX3 amino acid sequences used for constructing comparison alignments (121)

Table 2.10. List of SOX8 amino acid sequences used for constructing comparison alignments (123)

Table 2.11. List of SOX9 amino acid sequences used for constructing comparison alignments (124)

Table 2.12. List of Hsp90b1 or TRA1 amino acid sequences used for constructing comparison alignments (125)

Table 2.13. List of Wnt3 amino acid sequences used for constructing comparison alignments (126)

Table 2.14. List of Wnt4 and Wnt3 amino acid sequences used for constructing comparison alignments (127)

Table 2.15. List of Wnt4 and Wnt5 amino acid sequences used for constructing comparison alignments (129)

Table 2.16. List of Wnt5 amino acid sequences used for constructing comparison alignments (131)

Table 2.17. List of Wnt3, 4, and 5 amino acid sequences used for constructing comparison alignments (133)
Table 2.18. List of WT1 amino acid sequences used for constructing comparison alignments (135)

Table 2.19. Primers and PCR annealing temperatures for successfully amplified sex determination genes (136)

Table 2.20. Details regarding sex determining genes successfully amplified (137)

Table 2.21. Primers and details for genes that did not amplify in the sea lamprey and Pacific lamprey (140)

Table 2.22. Genes that were successfully amplified and that implicated roles in vertebrates (141)

Table 3.1. Reference and seven candidate sex differentiation genes for qPCR (229)

Table 3.2. Reference and target gene primer information used for qPCR (231)

Table 3.3. Summary of male and female sea lamprey gonadal development stages used in this study (232)

Table 3.4. Details of gonadal development in the male sea lamprey (233)

Table 3.5. Description of stages of testicular development in larval and metamorphosing male sea lamprey (234)

Table 3.6. Total number of larval, metamorphosing and adult lamprey samples available for this study (235)

Table 3.7. Sampling location information for larval, metamorphosing and adult sea lamprey (236)

Table 3.8. P-values reported for statistical analyses for gene expression in DAZAPI, DMRTA2, GCL, SF1, SOX8, SOX9 and WT1 (237)

Appendix A3.1. List of all male and female sea lamprey samples used for qPCR (255)

Appendix A3.2. List of genes for which initial attempts at qPCR were unsuccessful (260)

Appendix A3.3-3.10. Calibration curves for reference and target genes used for quantitative reverse-transcriptase PCR (262)

Appendix A3.11. Mean Ct values from gene expression data (266)
List of Figures

Figure 1.1. Overview of sea lamprey life cycle..........................................................74
Figure 1.2. Developmental process of spermatogenesis in the sea lamprey....................75
Figure 2.1. Alignment of DMRT1 amino acid sequences..............................................143
Figure 2.2. Neighbour-joining tree comparing DMRT1 amino acid sequences.............144
Figure 2.3. Alignment of DMRTA2 amino acid sequences........................................145
Figure 2.4. Neighbour-joining tree comparing DMRTA2 amino acid sequences..........146
Figure 2.5. Alignment of DMRTA amino acid sequences..........................................147
Figure 2.6. Neighbour-joining tree comparing DMRTA2 amino acid sequences.........148
Figure 2.7. Alignment of FEM1c amino acid sequences............................................149
Figure 2.8. Neighbour-joining tree comparing FEM1c amino acid sequences.............150
Figure 2.9. Alignment of FOXL2 amino acid sequences...........................................151
Figure 2.10. Neighbour-joining tree comparing FOXL2 amino acid sequences............152
Figure 2.11. Alignment of RSPO1 amino acid sequences..........................................153
Figure 2.12. Neighbour-joining tree comparing RSPO1 amino acid sequences..........154
Figure 2.13. Alignment of NR5A1 amino acid sequences.........................................155
Figure 2.14. Neighbour-joining tree comparing NR5A1 amino acid sequences..........156
Figure 2.15. Alignment of SOX2 and SOX3 amino acid sequences.............................157
Figure 2.16. Neighbour-joining tree comparing SOX2 and SOX3 amino acid sequences...158
Figure 2.17. Alignment of SOX8 amino acid sequences...........................................159
Figure 2.18. Neighbour-joining tree comparing SOX8 amino acid sequences.............160
Figure 2.19. Alignment of SOX9 amino acid sequences...........................................161
Figure 2.20. Neighbour-joining tree comparing SOX9 amino acid sequences.............162
Figure 2.21. Alignment of HSP90b1 or TRA1 amino acid sequences..........................163
Figure 2.22. Neighbour-joining tree comparing HSP90b1 or TRA1 amino acid sequences...164
Figure 2.23. Alignment of WNT3 amino acid sequences.........................................165
Figure 2.24. Neighbour-joining tree comparing WNT3 amino acid sequences.............166
Figure 2.25. Alignment of WNT4 and WNT3 amino acid sequences........................................167
Figure 2.26. Neighbour-joining tree comparing WNT4 amino acid sequences.........................168
Figure 2.27. Alignment of WNT4 and WNT5 amino acid sequences........................................169
Figure 2.28. Neighbour-joining tree comparing WNT4 and WNT5 amino acid sequences............170
Figure 2.29. Alignment of WNT5 amino acid sequences..........................................................171
Figure 2.30. Neighbour-joining tree comparing WNT5 amino acid sequences..........................172
Figure 2.31. Neighbour-joining tree comparing WNT3, 4, and 5 amino acid sequences..............173
Figure 2.32. Alignment of WT1 amino acid sequences.............................................................174
Figure 2.33. Neighbour-joining tree comparing WT1 amino acid sequences.............................175
Figure 2.34. Neighbour-joining tree comparing sequenced DMRTA2 nucleotide sequences....176
Figure 2.35. Neighbour-joining tree comparing sequenced FEM1c nucleotide sequences.........177
Figure 2.36. Neighbour-joining tree comparing sequenced FOXL2 nucleotide sequences.........178
Figure 2.37. Neighbour-joining tree comparing sequenced RSPO1 nucleotide sequences......179
Figure 2.38. Neighbour-joining tree comparing sequenced NR5A1 nucleotide sequences......180
Figure 2.39. Neighbour-joining tree comparing sequenced SOX2 nucleotide sequences...........181
Figure 2.40. Neighbour-joining tree comparing sequenced SOX8 nucleotide sequences.........182
Figure 2.41. Neighbour-joining tree comparing sequenced SOX9 nucleotide sequences.........183
Figure 2.42. Neighbour-joining tree comparing sequenced WNT3 nucleotide sequences........184
Figure 2.43. Neighbour-joining tree comparing sequenced WNT5 nucleotide sequences........185
Figure 2.44. Neighbour-joining tree comparing sequenced WT1 nucleotide sequences............186
Figure 3.1. Mean gene expression of SOX9 in all sea lamprey females a) and all males b) in developmental stages used in this study.................................................................238
Figure 3.2. Mean gene expression of SOX9 in sea lamprey from: a) stage 1 females compared to stage 2 males; b) stage 2 females compared to stage 2 males and c) stage 4 females compared to stage 4 males.................................................................239
Figure 3.3. Mean gene expression of WT1 in all sea lamprey females a) and all males b) in developmental stages used in this study.................................................................241
Figure 3.4. Mean gene expression of WT1 in sea lamprey from a) stage 1 females compared to stage 2 males and b) stage 4 females compared to stage 4 males..............................................242
Figure 3.5. Mean gene expression of DMRTA2 in all sea lamprey males from developmental stages used in this study. 243

Figure 3.6. Mean gene expression of DAZAP1 in all sea lamprey males from developmental stages used in this study. 244

Figure 3.7. Mean gene expression of SOX9 in all female and male sea lamprey. 245

Figure 3.8. (a-p) The processes of male testicular development during the late larval stage (prior to histological differentiation), metamorphosis (during which testicular differentiation occurs), and sexual maturity in the sea lamprey. 246
CHAPTER 1: General introduction

Sex determination is defined as the commitment of the gonad to a male or female fate, and sex differentiation is the diversification of undifferentiated tissue into tissues and phenotypes associated with the fate (Parma and Radi, 2012). Sex determination and sex differentiation are processes that have always intrigued biologists striving to decipher the mechanisms that select a male or female developmental path for an organism’s life. Through the advent of new technologies and discoveries, modes of sex determination have been elucidated for many invertebrates and vertebrates. Birds and mammals, for example, all exhibit genetic sex determination and the master sex-determining genes are conserved within each taxon, but sex determination mechanisms in other vertebrates are highly variable (Cutting et al., 2013). The mode of sex determination for a basal vertebrate, the lamprey, remains unknown. Lampreys are ancient agnathans found at the pivotal emergence point of vertebrates that can be traced back over 360 million years (Docker, 2006; Gess et al., 2006; Docker et al., 2015). Little is known about the mode of sex determination in lampreys; an environmental influence on sex determination has been suggested given observed correlations between sex ratio and population abundance (e.g., Docker and Beamish, 1994), but a genetic mode of sex determination has not been ruled out. Sequencing of the sea lamprey, Petromyzon marinus, genome has been a revolutionary breakthrough for molecular research into lamprey evolution and development and can possibly allow for the identification of homologous loci involved in sex determination from other species (Smith et al., 2013).

Mechanisms of sex determination can be genetic, environmental, or a combination of the two (Penman and Piferrer, 2008). There is great variability of sex determination mechanisms even among groups of closely related species (see Section 1.3). In contrast, genes involved in sex
differentiation in vertebrates are more conserved (Siegfried et al., 2010; see Section 1.4). This thesis aims to: 1) identify if genes involved in sex determination in other vertebrates are sex-linked in sea lamprey and Pacific lamprey, *Entosphenus tridentatus* (i.e., if these genes also represent candidate sex determining genes in lampreys); 2) to test if genes involved in sex differentiation in other vertebrates (i.e., candidate sex differentiation genes) show sex specific patterns of expression during testicular differentiation in sea lamprey; and 3) discuss the conservation of these genes among vertebrates. This thesis will include: a general introductory chapter (Chapter 1), which will provide an overview of lamprey biology, conservation and control of lampreys, review sex determination in invertebrates and model vertebrates, fish and lampreys, and review genes involved in sex differentiation in vertebrates; two research chapters (Chapters 2 and 3) related to aims 1) and 2) above; and a concluding chapter (Chapter 4) that will briefly discuss the significance of the findings from this thesis and implications on lamprey sex determination and male differentiation, the limitations of this study and possible future applications.

1.1. Lamprey biology: life-cycle

Lampreys are ancient agnathan vertebrates, with approximately 41 extant parasitic and non-parasitic species (Docker, 2006; Potter et al., 2015). Two traditionally and economically important parasitic species that can be found in North America are the sea and Pacific lamprey (Clemens et al., 2010). The sea lamprey is native to the Atlantic Ocean and adjacent Europe, where native populations are rapidly declining, yet the same species is a nuisance in the Great Lakes (Igoe et al., 2004). The construction of connecting waterways allowed for the access of this invasive species into the Great Lakes, and although sea lamprey is an example of one species that is a problematic parasitic lamprey, this is not the case for all parasitic lampreys. Such an
example is the Pacific lamprey, which has become of conservation concern in the Pacific Northwest due to anthropogenic impacts, particularly dam construction which blocks access to spawning habitat. Indigenous groups in the Columbia River basin have collected large tribal harvests of Pacific lamprey for generations as the fish migrate to spawning grounds starting in the spring and leading into fall, but the population is in severe decline (Clemens et al. 2013; Maitland et al. 2015). Other abiotic factors like mercury, PCBs, DDT and various other chemicals present in the environment are reducing their population size (Clemens et al., 2012).

The sea and Pacific lamprey belong to the family Petromyzontidae, which is found in the northern hemisphere. Some differences are evident between the landlocked and anadromous sea lamprey (where anadromous refers to fish that migrate from the sea to rivers to spawn), and anadromous Pacific lamprey. Sea lamprey are better able to cope in freshwater environments than the Pacific lamprey, yet the Pacific lamprey is better adapted to climbing vertical surfaces such as waterfalls and dams (Wallace and Ball, 1979; Beamish and Northcote, 1989; Clemens et al., 2010). Fecundity (i.e., total egg count in mature females) is directly proportional to body length (Beamish, 1980). The least fecund of these three lampreys is the Great Lakes sea lamprey at a maximum body length usually less than 700 mm. Higher egg counts are seen in the larger Pacific lamprey, at an intermediate length reaching up to 800 mm, and in the anadromous sea lamprey, in which body length can range from 800 to 900 mm (Beamish, 1980; Clemens et al., 2010). The total duration and distance of migration are shortest for the sea lamprey lasting a few months, and extremely long in the Pacific lamprey, lasting well over 12 months (Clemens et al., 2010). Although there are differences in body length, there are some commonalities between the sea and Pacific lamprey. The amount of time all lampreys spend as filter feeding larvae ranges from 3 to 8 years, and their time of maturation and reproduction, governed by external
environmental factors, are similar (Scott and Crossman, 1973; Dawson et al., 2015; Johnson et al. 2015).

The native habitat of the sea lamprey ranges between the northern latitude of 53 degrees by Newfoundland and Labrador and neighbouring Europe, to the southern-most limits of Florida, at a latitude of 30 degrees. The precise time of life events in lampreys can be difficult to pinpoint to a month, with an associated large degree of variability due to their developmental dependence on environmental factors such as latitude and temperature (Beamish, 1980). Although variability exists in the exact timing of events, approximate estimates are included to show the variability in modes of feeding and prolonged time leading up to sexual maturation.

The life of a sea lamprey begins as a filter feeding larva; the length of this stage ranges from 6 to 8 years in the anadromous form and averages 6 years in the landlocked population (Lowe et al., 1973; Beamish and Potter, 1975; Beamish, 1980). Larvae live a sedentary life, with growth largely dependent on temperature, and rely on algae and primarily on detritus for sustenance (Beamish, 1980; Sutton and Bowen, 1994; Dawson et al., 2015). Based on numerous studies of gonadal development and differentiation on both parasitic and non-parasitic lampreys, Hardisty (1969) found several differences in gonadal and sexual development during the larval stage between landlocked and anadromous populations of the same species: the sea lamprey. In the landlocked population, the highest activity in terms of gonadal development and growth takes place at lengths between 61 and 70 mm in age class two, whereas the same gonadal phase is seen later in the anadromous population, at lengths ranging from 81 to 110 mm, in individuals in age class three and four. Despite these differences, the size ranges at which anadromous and landlocked female larvae undergo gonadal development overlapped, and this was also consistent for males (Hardisty, 1969). Following the larval phase, the larva will undergo a dramatic change
known as metamorphosis at an average length of 130 mm, consisting of the lamprey undergoing both internal and external changes (Lowe et al., 1973; Potter et al., 1978; Manzon et al., 2015). Drastic external modifications consist of the formation of the circumoral disc with teeth, appearance of eyes, and changes in body colouration. Along with the external changes, internal remodeling of the gastrointestinal tract occurs, with the formation of a new esophagus, as well changes in the kidney and endostyle, to name a few (Youson et al., 1977; Manzon et al., 2015). These changes allow for the freshwater filter feeding larvae to transform in preparation for its new parasitic life at sea.

Lamprey growth is directly related to the temperature in the spring and fall seasons, with metamorphosis starting as early as May in some individuals, with completion by late fall in October or November of the same year (Manzon et al., 2015). Lampreys that have transformed or metamorphosed will migrate to feed in the fall, and the late spring transformers in the following season will migrate between March and May (Lowe et al., 1973; Beamish, 1980; Moser et al., 2015). The parasitic feeding phase, known as the juvenile phase, can typically last a year or more in the sea lamprey (Bergstedt and Swink, 1995). Extreme size variabilities can be found among lamprey in this group, as well as variation in the onset of feeding time (Beamish, 1980). Although it has been proposed that an earlier outmigration and onset of feeding (i.e., fall versus spring) might provide the lamprey with longer duration to feed before the migration back to spawn, Swink and Johnson (2014) did not find differences in size at spawning between fall and spring migrants.

Upon the completion of the trophic phase, lampreys will migrate upstream to spawning grounds. In the anadromous population, they can be seen residing in associated tributaries in early spring as they continue their final sexual maturation (Beamish, 1980; Johnson et al., 2015).
Leading up to the highly-synchronized fertilization event, males can be seen constructing nests (Harvey et al., 2008). Once sexually mature, the lamreys will undergo an energetically demanding external fertilization that takes place between June and August. Lamreys are semelparous, meaning they spawn during a single reproductive episode; thus, soon after the synchronous fertilization event, the lamreys will die in the streams (Lowe et al., 1973).

1.2. Lamprey biology: conservation and control

1.2.1. Control of the invasive sea lamprey in the Great Lakes

Both the Canadian and American governments have long been battling the sea lamprey in the Great Lakes, spending millions of dollars to control the pest (Fetterolf, 1980). Niagara Falls acted as a natural barrier and prevented the sea lamprey from access to Lake Erie and the upper Great Lakes until the formation of the Welland Canal to bypass the falls connecting Lake Ontario to Lake Erie in 1829 (Cuhel and Aguilar, 2013). Sea lamprey were blamed for the 1959 collapse of the lake trout, *Salvelinus namaycush*, fishery, an industry that profited an annual $5.5 million per 7000 tonnes (Smith and Tibbles, 1980). The Great Lakes Fishery Commission was created in 1955 to manage the Great Lakes, and since its creation until 1979, 50 million dollars had been spent on controlling the lamprey population, which is less than 5% of the monetary damages caused by the lamprey to the fishery industry up to that point (Fetterolf, 1980). Between 1998 and 2004, over 6 million dollars in total was used for control using lampricide treatment of the Great Lakes, which targets approximately half of the observed lamprey population in each respective lake (Irwin et al., 2012). The amount of money spent does not include the cost of restoration and the damage to fishery production. Sea lamprey damage on the fishery industry and the rehabilitation of fish stocks can cost up to $5 billion dollars per year (Marbuah et al., 2014). Numerous control measures to eradicate sea lamprey have been implemented, such as use
of the lampricide TFM (3-trifluoromethyl-4-nitrophenol), but unfortunately TFM is not always effective in killing sea lamprey larvae (Clifford et al., 2012), and it can have lethal effects on some non-target species (Boogaard et al., 2003). Some non-target species that are either of conservation concern or of economic importance such as lake sturgeon, *Acipenser fulvescens*, lake whitefish, *Coregonus clupeaformis*, Atlantic salmon, *Salmo salar*, and common shiner, *Luxilus cornutus*, display high sensitivity to TFM (Boogaard et al., 2003). Overall, however, current methods of controlling sea lamprey populations in the Great Lakes have been effective with a 90% population decrease since the peak of abundance (Siefkes, 2017). The decrease in the sea lamprey population has allowed the previously decimated lake trout population to thrive and the Great Lakes Fishery Commission reports the population has grown to almost 27 million fish/km/net at night, which is a shocking 1700% increase from the 1970s in Lake Superior (Siefkes, 2017).

### 1.2.2. Conservation of the Pacific lamprey

Although the landlocked sea lamprey is invasive and requires control, many other lamprey species are of conservation concern (Maitland et al. 2015). In particular, the Pacific lamprey, which is native to the west coast of North America, is of economic, cultural and medicinal value (Close et al., 2002). Pacific lamprey have long been used as religious food for the Indigenous peoples of the mid-Columbia River Plateau and still play an important role in cultural and ceremonial practices (Close et al., 2002). Willamette Falls in Oregon has long been the source of Pacific lamprey for Native American tribal harvest, but populations are in decline (Clemens et al., 2013). Awareness of the Pacific lamprey needs to be promoted as most people’s views on lampreys in North America are influenced by the sea lamprey and its devastation in the Great Lakes (Close et al., 2002). From the 1960s, in a period of just 7 years, the Pacific lamprey
adult population at Dalles Dam, Oregon, saw a 90% decline in the population, with a final population of 350,000 adults (Close et al., 1995). They have continued to decrease over the years and a similar pattern has been seen over a 35-year period, as the population had continued to decrease to 90% of those observed at McNary Dam, Oregon (Close et al., 1995). In Bonneville Dam, Oregon, between 1938 and 1969, and from 1997 to 2011, there has been on average a 63% decline in Pacific lamprey adult counts (Murauskas et al., 2013). The alarming rate of decline is due to a combination of anthropogenic factors including barriers such as dams, dredging, and chemicals (Maitland et al., 2015).

In the Columbia River basin, over 200 dams have been constructed; in some tributaries, such as the Umatilla River, Pacific lamprey are unable to pass beyond the dams (Gelfenbaum and Kaminsky, 2010; Jackson and Moser, 2012). Dams have the most dramatic and negative effect on upstream-migrating adults and downstream-migrating metamorphosed lampreys (Nunn and Cowx, 2002). If adult lamprey are unable to migrate upstream past dams, their grounds for spawning are substantially limited, leading to greater population loss (Moser and Close, 2003). As well, populations of Pacific lamprey stranded upstream of new dams may be extirpated (e.g., Beamish and Northcote 1989). There have been barriers found in over 4,100 streams in the coastal Oregon regions that would potentially block Pacific lamprey migration (Starcevich and Clements, 2013).

Excavation of waterways for human-navigation also have a negative impact on larval lampreys (Maitland et al., 2015). Dredging not only disrupts and decreases the natural habitat of larvae, but also causes the release of toxic compounds that are buried beneath the sediment (Erftemeijer et al., 2012). Nutrients that limit the growth of algae such as nitrogen and phosphorous are found in abundance in untreated water such as sewage, and lead to large-scale
blooms (Ryther and Dunstan, 1971). Water turbidity as a result of algal blooms and
eutrophication have shown to cause negative changes in fish—prey interactions, as well as
The large-scale impacts of excavation of land need to be further researched and how they impact Pacific lamprey.

The effect of other abiotic factors like mercury, PCBs, DDT and various other chemicals present in the environment may be reducing Pacific lamprey populations (Clemens et al., 2012). Larval Pacific lamprey have been found to have high concentrations of chemical pollutants in tissues that likely decrease survival and further decrease Pacific lamprey populations (Nilsen et al., 2015). Pentachlorophenol is a lipophilic chemical that can be found in water and exerts its action by uncoupling oxidative phosphorylation, preventing the production of ATP (McKim et al., 1987). Due to the lipophilic nature of pentachlorophenol, it can accumulate in the tissues of lampreys, being dangerous for the survival of the lamprey and for animals and people that rely on the lamprey for food. Mercury also is known to have detrimental effects in both its organic and inorganic form to fish development affecting their neural, skeletal, and cardiac systems to name a few negative impacts (Dong et al., 2016). The concentration of mercury in the sediment of the Columbia River basin is very high, and based on recent analysis, the concentration in the tissue of Pacific lamprey ammocoetes significantly exceeded that of the environment (Linley et al., 2016). The already declining population is adversely affected and this will just be perpetuated due to the negative effects of mercury.

1.2.3. **Conservation of the anadromous sea lamprey in their native range**

In medieval Europe, lampreys were considered food fit for royalty and since the 18th century an increase in harvesting has been seen in much of Europe (Renaud, 2011; Docker et al.,
The current status of sea lamprey in most of Europe is listed as vulnerable (Mateus et al., 2012). The sea lamprey is of economic importance and of conservation concern due to declining populations in much of Europe. The Garonne Basin in France has been a large supplier for sea lamprey harvest, but the population has seen a steady decline due to over-harvesting in many European countries (Beaulaton et al., 2008). In order to conserve the populations of the sea lamprey, habitat recovery and sustainable harvesting practices must be enforced to allow a long-term availability of sea lamprey in Portugal (Stratoudakis et al., 2016).

Conservation methods employed in North America for the Pacific lamprey can potentially be applied and used in Europe for the conservation of the sea lamprey. Similar obstacles that are causing a decline in the Pacific lamprey population pose as issues causing a decline in the sea lamprey native range, such as anthropogenic factors: dams, industrialization, over-harvesting and changes in temperature (Mateus et al., 2012; Maitland et al., 2015; Hansen et al., 2016). The most significant culprit leading to the decline in the sea lamprey population—as with the Pacific lamprey—is the construction of dams, leading to over an 80% decrease (Mateus et al., 2012). To enable passing of lamprey past the dams, vertical slot fish passes were installed along dams, in particular at one location, at the Coimbra Açude-Ponte Dam in the Modego River in Portugal. After surveillance of the sea lamprey population from 2011 to 2015, the number of sea lamprey larvae beyond the dam had seen a 30-fold increase, proving the success of fish passage installations (Pereira et al., 2017).

An alternative to building fish passages is the complete removal of barriers that prevent fish passage. Along the coast of southern France, dams were removed from the Scorff River and soon after the removal of the barriers, a significant increase in the number of sea lamprey spawning nests was seen in areas previously not colonized by the sea lamprey (Lasne et al.,
Likewise, in its native North American range, spawning runs of anadromous sea lamprey were blocked by dams constructed in the Penobscot River tributary in Maine in the 1800s. Upon the removal of the dams in 2009, anadromous sea lamprey were able to rapidly recolonize the previously restricted areas and a four times increase in the number of nesting sites was observed (Gardner et al., 2012; Hogg et al., 2013).

There have been discussions recently on the harvest and transfer of the invasive sea lamprey from the Great Lakes to the native range in the Atlantic and to neighbouring Europe, where the population is in rapid decline (Guo et al., 2017). Due to the accumulation of mercury in fish in the Great Lakes, there are many advisories given to the general public and sensitive public in terms of which species and sizes of fish should be avoided due to the higher prevalence on mercury and other contaminants (Bhasvar et al., 2011). Determining whether the concentration of mercury in lampreys is safe for consumption from different lakes will be crucial before any sea lamprey fisheries can be created and transferred to other regions.

Further employing new genetic technology can provide ways to conserve and manage lamprey populations. Using new genetic sequencing technologies and genetic analysis of Pacific lamprey, Hess et al. (2013) established that Pacific lamprey are able to maintain gene flow between populations within their range and a loss of genetic variability could potentially lead to localized extirpation. New and developing genetic and genomic technologies have important implications; for example RNA interference technology and other reverse genetic tools have the potential for application for control of invasive species (Heath et al., 2014; McCauley et al., 2015). Exploration of a molecular sex determination in lampreys could allow for the control of biased sex-ratios for species of both control and conservation (Hess et al., 2013; McCauley et al., 2015). In order to determine whether the methodology would be successful and species-specific,
a large number of embryos and larvae would be required before large-scale application. Technologies for culturing fish embryos for experimental aquaculture are available and can be used to rear embryos of lamprey species of conservation concern, as they can be reared in captivity and used for study without posing a threat to wild populations (Mitz et al., 2014; McCauley et al., 2015).

1.3. Sex determination in vertebrates

Sex determination is defined as the commitment of the gonad to a male or female fate, not to be confused with sex differentiation, which is the diversification of undifferentiated tissue into tissues and phenotypes associated with the fate (Parma and Radi, 2012). The gonad is the first organ that displays dimorphism, where it develops either into an ovary or testis from an undifferentiated gonad, triggered by various gene interactions in vertebrates (Siegfried, 2010). Birds and mammals all exhibit genetic sex determination, but sex determination mechanisms in other vertebrates are highly variable (Cutting et al., 2013). In other vertebrates, both genetic sex determination (GSD) and environmental sex determination (ESD) are known to occur (Bulmer, 1987; Takada et al., 2005; Heule et al., 2014) and mechanisms are variable. Variation in sex determination patterns, for example, can be present even between closely related species of lizards; in some cases, the sex is pre-determined by the genotype, and in others it is determined by environmental factors such as temperature (Bulmer, 1987). In several fish species, ESD and GSD have been shown to co-exist (e.g., Yamamoto et al., 2014; Palaiokostas et al., 2015).

Environmental sex determination is the determination of the sex of an individual at a critical period in embryonic development based on the environmental conditions, such as temperature (Charnov and Bull, 1977; Schwanz et al., 2016). Turtles display temperature-dependent sex determination (TSD), where males can be produced at low and females at higher
temperatures, or, females at low, males at intermediate, and females at high temperatures (Valenzuela, 2004). In the American alligator, *Alligator mississippiensis*, sex determination is highly sensitive to temperature changes, with a narrow range of 4°C determining male or female fate (Ferguson and Joanen, 1982). Alligators and crocodiles also have ESD, but male specific gene expression at critical embryo developing temperatures have been observed, with high expression of *Double-sex and mab-3 related transcription factor 1 (DMRT1)* in male alligators and *SRY*-related HMG-box transcription factor 9 (*SOX9*) expression only in male crocodile, *Crocodylus palustrius* (Smith et al., 1999; Agrawal et al., 2009; see Section 1.4 regarding genes involved in sex differentiation). Recently, further molecular studies have provided evidence to show male biased gene expression associated with the vanilloid receptor 1 protein channel activity, linking TSD with genetic patterns in the alligator (Yatsu et al., 2015). Some species with TSD, such as the painted turtle, *Chrysemys picta*, show gene expression patterns associated with male and female critical development temperatures. Genes that have shown to be up-regulated during male embryo development are *Anti-Müllerian hormone (AMH)*, Androgen receptor (*Ar*) and *GATA binding protein 4 (Gata4)* and female up-regulated genes are *Forkhead box L2 (FOXL2)*, *Aromatase* and β-*catenin*, to name a few (Radhakrishnan et al., 2017). Current evidence suggests that ESD is favoured when males and females have differences in age at maturation (Schwanz et al., 2016). Another mode of ESD is photoperiod-dependent sex determination, which is more reliable and consistent from year to year than temperature. This mode of sex determination as well as TSD is used by the California grunion, *Leuresthes tenuis*, with a higher prevalence of females in the population at long day lengths and cooler temperatures (Brown et al., 2014).

In species with GSD, the sex of the individual is determined upon fertilization. However,
genes involved in the molecular pathways of sex determination are variable and not conserved among all vertebrates due to the rapid action of evolutionary pressures, whereas the genes located downstream that are involved in sex differentiation are more conserved among species and retain higher gene sequence conservation (Cutting et al., 2013; Graves, 2013; Huele et al., 2014). It is also essential to point out that genes that are involved in sex determination are not exclusive to sex determination and can be involved in sex differentiation as well (Cutting et al., 2013). Genes that play vital roles in sex determination and differentiation that cannot easily be overridden by others will be preferred, and therefore conserved, regardless of their position within the gene cascade and can be expressed at a particular developmental stage, tissue specific, and sex exclusive manner (Graves and Peichel, 2010; Graves, 2013). These genes can be expressed in a male or female specific or a dosage sensitive manner.

1.3.1. Genetic basis of sex determination in model vertebrates

If sex determining genes become localized on chromosomes, these are referred to as sex chromosomes. In some species that have sex determining chromosomes, chromosomes are inherited in males and females as either ZZ/ZW (e.g., in birds) or XY/XX (e.g., in mammals), and the inheritance of the chromosomes at fertilization determines sex (Wallis et al., 2008; Cutting et al., 2013).

In birds, the female is the heterogametic individual ZW and the male is ZZ, and the master male sex determination gene is DMRT1 (Yano et al., 2012; Cutting et al., 2013). DMRT1 is found on the avian Z chromosome and is expressed in a dosage dependent manner, higher in developing male chickens, and required for testis development in all birds (Smith et al., 1999; Shetty et al., 2002). Snakes and geckoes also employ the ZZ/ZW chromosomal system, but DMRT1 is not the sex determining gene (Ohno, 1967; Trifinov et al., 2011). In geckoes, DMRT1
is an autosomal related gene located on both the Z and W chromosome, and is homologous to the avian Z chromosome (Kawai et al., 2009). What genes play a sex determining role in geckoes is unknown, but sex steroid hormones play an important role in the developing brain of the leopard gecko, *Eublepharis macularius* (Endo et al., 2008). The sex determining gene or genes for snakes also has yet to be discovered (Graves, 2013).

Sex determination in eutherian mammals is determined at conception by the X or Y chromosome from the sperm that leads to the female or male fate, respectively (i.e., females have XX chromosomes and males have XY). The male path is determined by the presence of the sex determining region Y (*SRY*) gene that is present on the Y chromosome in most therian mammals and initiates the *SOX9* cascade pathway to develop the undifferentiated gonad into a testis (Wallis et al., 2008; Parma and Radi, 2012). Prior to the discovery of the role of *SRY*, genes located up- and downstream of *SRY* were analyzed in model organisms including sex-reversed humans, and *SOX9* was revealed to be a close relative of *SRY* and one of the first genes expressed in a sex specific manner (Sinclair et al., 1990; Morais da Silva et al., 1996). Sex determination region on the Y chromosome (*sdY*) a homologue of *SRY*, is present in salmonid fishes, where it shares homology with an immune related gene, interferon regulatory factor 9 (*Irf9*) and is a conserved genetic region of sex determination in many species in the genus *Oncorhynchus* (Yano et al., 2012; Yano et al., 2013; see Section 1.3.2).

Even within mammals there is variability present in the mode of sex determination; notably the monotremes, the most ancient group of mammals, share greater homology in their chromosomal sex determination method with birds than with therian mammals (Veyrunes et al., 2008). The monotreme XY system bears similarity to the avian ZW system, indicating that this system perhaps arose in amniotes because geckoes, birds and monotremes retain the same
genomic area to serve as the sex chromosomes (Graves et al., 2006; Veyrunes et al., 2008; Graves, 2013). In both birds and mammals, sex determining genes are expressed in the gonads as well as non-gonadal tissue (Cutting et al., 2013).

1.3.2. Sex determination in fish

Sex determination mechanisms are more variable and less well known in fish (Siegfried, 2010). GSD systems in species with karyotypically distinguishable sex chromosomes receive more attention than other modes of sex determination because sex chromosomes are associated with a heterogametic and homogametic individual (Ferguson-Smith, 2007). Some fish can have multiple undifferentiated chromosomes that are similar in appearance, making it difficult to determine if there is a chromosomal basis for sex determination, and whether more than one pair of chromosomes might be involved (Ohno, 1974; Heule et al., 2014). In fishes that exhibit GSD, the associated sex determining genes have been found on chromosomes using sex linked traits, where the presence of a gene on a chromosome in one sex and the absence of that region in the opposite sex denotes the sex (Siegfried, 2010). In some fish species, the genotype can be regulated by both environmental and inherited factors (Takada et al., 2005; Heule et al., 2014; Yamamoto et al., 2014; Palaiokostas et al., 2015).

Some genes have been identified that lead to testis development in various fish species—although these are not necessarily the sex determination gene (i.e., the “master switch” that activates the male specific developmental cascade), they represent good candidate sex determination genes. For example, although fish do not possess Müllerian ducts, they express AMH, which plays a role in spermatogenesis by causing regression of Müllerian ducts. The presence of this hormone in fish indicates that the hormone arose before the appearance of Müllerian ducts (Kluver et al., 2007; Heule et al., 2014). AMH exerts its action by inhibiting the
production of aromatase, \textit{CYP19A1A}, which is essential for the production of estrogen. With the suppression of estrogen, testis development proceeds (Nagahama and Devlin, 2002). In at least one fish species (the Patagonian pejerrey, \textit{Odontesthes hatcheri}), \textit{amhy}, a paralogue that arose from the gene duplication of the ancestral gene \textit{AMH}, is the testis determining factor on the \textit{Y} chromosome. \textit{amhy} is expressed early in fertilization in the Patagonian pejerrey in the Sertoli cells surrounding the germ cells in the developing testis.

Sex determination genes in fishes can be highly variable, however, even among closely related species. In the example above, \textit{amhy} was found to be the testis determining in the Patagonian pejerrey, but not in medaka or Japanese ricefish, \textit{Oryzias latipes}, where \textit{AMH} is expressed in the gonads of both males and females (Kluver et al., 2007; Hattori et al., 2012; Cutting et al., 2013). Pejerrey and medaka are placed within the same superorder (Atherinea), albeit in different orders (Atheriniformes and Beloniformes; Nelson et al., 2004). Double-sex and mab-3 related-domain gene on the \textit{Y} chromosome (\textit{Dmy}) is a gene segment that arose from a gene duplication event in \textit{DMRT1}, and \textit{Dmy} acts as the testis determining gene in the medaka and a second species in the same genus, the Malabar ricefish, \textit{Oryzias curvinotus} (Matsuda et al., 2002; Nanda et al., 2002). However, in a third species in this genus, Luzon ricefish, \textit{Oryzias luzonensis}, males have 12 silent nucleotide substitutions within the gene gonadal somatic derived factor on the \textit{Y} chromosome (\textit{gsdfy}) relative to females and they use this gene as the master sex determination gene located in the upstream promoter region, where the promoter triggers higher expression in the male genital ridge (Myosho et al., 2012). Transformation experiments with \textit{O. luzonensis} and \textit{O. latipes} with \textit{gsdfy} led to genetic female \textit{XX} reversal to produce males. This is intriguing because although \textit{O. latipes} uses \textit{DMRT1} as the sex determining gene, \textit{gsdfy} was able to cause sex reversal. \textit{gsdfy} is located downstream of \textit{DMRT1} and may have a role in the master
sex determination cascade in *Oryzias* species (Myosho et al., 2012; Heule et al., 2014).

Conversely, sex determination genes may be shared among fish taxa that are not closely related. For example, *DMRT1*, which is the master male sex determination gene in birds (and its homologue *Dmy*, is the testis determining gene in medaka and Malabar ricefish), is a male specific gene that is expressed early during testicular differentiation in European eel, *Anguilla anguilla*, well before histological differentiation is observed (Nanda et al., 2002; Ijiri et al., 2008). A male specific role was also indicated for *DMRT1* in Nile tilapia, *Oreochromis niloticus* (Kobayashi et al., 2003). Eels are estimated to have diverged from the medaka and tilapia lineages more than 270 million years ago, and medakas and tilapias diverged from each other almost 120 million years (Kumar et al. 2017; http://www.timetree.org/).

In salmonids, the male specific *sdY* gene was found to be largely conserved throughout the genus *Oncorhynchus*, as well as in other species (e.g., brown trout, *Salmo trutta*) in the subfamily Salmoninae (Yano et al., 2012; Yano et al., 2013). *sdY* also appears to be male specific in most members of the other two salmonid subfamilies, Thymallinae and Coregoninae, but with some exceptions pertaining to the Coregoninae subfamily (Yano et al., 2012). In the European whitefish, *Coregonus lavaretus*, and the lake whitefish, *sdY* was found in both males and females (Yano et al., 2012), ruling out that *sdY* is a male specific locus in all salmonids. These species may have an *sdY* dosage dependent sex determination, or as a result of evolution, may have lost this region as the sex determining region and may rely on other genes (Smith et al., 1999; Yano et al., 2013). Since family Esocidae is considered to be one of the closely living relatives of the salmonids, Yano et al. (2013) searched for a similar region in the northern pike, *Esox lucius*, but were unsuccessful, thus restricting *sdY* to the family Salmonidae. Identification of these regions would provide for a method of molecular sex determination, along with better
control for sex determination in artificial selection of economically important species (Yano et al., 2013).

As mentioned above, fish can also exhibit environmental sex determination (Charnov and Bull, 1977). Charnov and Bull (1977) proposed that fish larvae will employ ESD when the larvae randomly disperse and are unable to choose their environment. ESD cues are commonly used by vertebrates, but this is not a good sex determining method when the length of the breeding season is short because similar environmental conditions can lead to biased sex ratios (Valenzuela and Lance, 2004; Brown et al., 2014). In the Atlantic silverside, *Menidia menidia*, the ESD factor is temperature and length of the growing season. Female silversides are produced during cool temperatures in early spring, and males develop with warmer environmental temperatures in the late summer (Baumann and Conover, 2011). A genetic mode of sex determination also exists in the Atlantic silverside, with variation in mode of sex determination observed among populations related to differences in the length of the growing season (Duffy et al., 2015). The level of TSD and GSD varies across the different populations, and no single population exclusively displays GSD or TSD (Duffy et al., 2015). For coastal Pacific species such as the California grunion, photoperiod is a large factor since there is little temperature variation (Brown et al., 2014).

### 1.3.3. Sex determination in lampreys

Despite decades of study, the factors that influence sex determination in lampreys continue to elude biologists (Docker, 1992; McCauley et al., 2015). At what point in the life cycle sex is determined is still unknown, and whether there is a genetic component has yet to be explored (Docker, 1992; Docker et al., 2003; Clemens et al., 2010; Smith et al., 2010).

ESD, specifically density-dependent sex determination, has been proposed in lampreys,
but results based on these studies have been inconclusive (Hardisty, 1965; Docker, 1992; Docker and Beamish, 1994; Dawson et al., 2015). There is perhaps an involvement of various factors that collectively determine the sex, such as genotype, abiotic factors, and population density, where over-crowding leads to higher male prevalence, a pattern that is similar between lampreys (Docker and Beamish, 1994) and eels (Davey and Jellyman, 2005; Geffroy and Bardonnet, 2016). Variability in sex ratios exists in larval and adult lampreys, and is especially pronounced in the Great Lakes sea lamprey. During the peak of sea lamprey abundance in the Great Lakes, the larval population was male-biased, but following the chemical application of TFM, a pronounced shift was seen leading to an excess of females (Wigley, 1959; Torblaa and Westman, 1980; Johnson et al., 2015). A recent study looking at the effect of growth rate of sea lamprey larvae based on the productivity of the environment suggested that growth rate and not density per se may be a cue related to sex determination. Sea lamprey larvae were tagged and released into productive and unproductive environments, and based on recapture of adults, unproductive environments were associated with slower growth rate and strongly male-biased sex ratios, whereas this bias was much reduced in the productive environment (Johnson et al., 2017).

Given the observational data suggesting density-dependent sex determination in lampreys, Docker (1992) conducted experiments to determine if sea lamprey larvae reared under conditions of different densities yielded different sex ratios. A slightly higher proportion of males was observed in the lower density tanks, but differences in sex ratio were not significant.

More recent studies have attempted to identify a genetic basis, or chromosomal basis, for sex determination in lampreys. Mateus et al. (2013) compared genomic DNA (gDNA) of male and female European brook lamprey, *Lampetra planeri* using reduced-genome-representation genotyping (i.e., Restriction site Associated DNA Sequencing, or RAD-Seq). Upon analysis of
the divergence between males and females using sex specific loci, the read coverage was analyzed using RAD loci. Graphs were plotted with lines depicting expected cluster areas for X-linked loci in a male heterogametic system, W-linked loci in a female heterogametic system and autosomal-linked loci. There was no deviation in the lamprey from the autosomal loci predicted line, indicating that if a genetic sex determination method is employed by lampreys, X or W-linked loci on chromosomes may not be involved. The same experiment was conducted for the threespine stickleback, *Gasterosteus aculeatus*, and a second cluster of loci was formed by the X-linked loci predicted line, which is absent in males. This indicates that the threespine stickleback have XX homogametic females (Mateus et al., 2013). This experiment provides further support for the threespine stickleback male Y chromosome that duplicated recently, approximately 13-16 million years ago from the X chromosome (White et al., 2015). The sea lamprey genome is composed of 84 pairs of chromosomes (2n=168 chromosomes) and sex specific loci differences may be unlikely due to their undifferentiated nature and based on results from *L. planeri*.

1.4. Genes involved in sex differentiation

1.4.1. Sex differentiation in model vertebrates

Although the master sex determining gene that triggers sex differentiation is highly variable in lower vertebrates (Section 1.3.2), the path vertebrates take during sex differentiation is very conserved even between fish and mammals (Siegfried, 2010; Graves, 2013). Genes that are involved in sex determination can be co-opted to have also roles in gonadal differentiation (Cutting et al., 2013). Male sexual differentiation takes place as male specific factors induce phenotypic changes in the undifferentiated gonad to develop into a testis. From a genetic perspective, these changes are a result of many genes and hormones such as the masculinization
hormone testosterone, which exert their actions by binding to transcription factors and activating gene cascades (Parma and Radi, 2012).

The SRY gene in mammals activates the downstream SOX9, which is thought to be conserved in the vertebrate lineage and is responsible for testis and Sertoli cell differentiation (Morais de Silva et al., 1996). In the absence of the therian male Y chromosome and thus the SRY gene, the female pathway leads to the development of the ovaries. In birds and some non-mammalian vertebrates, DMRT genes play the role of SRY, where DMRT-related genes activate SOX9 and subsequently AMH, which play a key role in testis production. AMH leads to the regression of Müllerian ducts in mice, Mus musculus, and although it is not required for testis differentiation, it indirectly promotes testis development by preventing female gonadal development (Cutting et al., 2013).

1.4.2. Sex differentiation in fish

In the Indonesian coelacanth, Latimeria menadoensis, a sarcopterygian (i.e., “lobe finned”) fish, some genes that are highly expressed during testis differentiation are SOX9, Sox (SRY-related HMG-box) transcription factor 10 (SOX10), Wilm’s tumour protein 1 (WT1), AMH, and Steroidogenic factor 1 (SF1) (Forconi et al., 2013). Genes that are involved in the sex differentiation process in other vertebrates, such as Sox (SRY-related HMG-box) transcription factor 8 (SOX8) and fibroblast growth family of genes, do not appear to play a role in coelacanth testis differentiation, indicating that other genes may have taken on the roles of these genes in early fishes. In another ancient fish lineage, the actinopterygian Siberian sturgeon, Acipenser baerii, there was no difference in the expression pattern of SOX9 in the post differentiated males and females (Berbejillo et al., 2012), but there was sexually dimorphic expression of several genes during differentiation: high expression of the androgen receptor AR, star protein,
aromatase, CYP19a1a, DMRT1, and SOX9 were associated with male testis differentiation. In a more recently derived fish lineage, the scorpaeniform sablefish, *Anoplopoma fimbria*, showed up-regulation of DMRT1, SOX9 and AMH during male differentiation and up-regulation of FOXL2 and aromatase during female differentiation (Smith et al., 2013).

During the important sex differentiation period from 35 to 70 dah (days after hatching) in the Nile tilapia, the expression of AMH in male testis substantially increased. The expression of SOX9 was similar in both sexes from 5 to 25 dah, but higher expression in males began at day 35 and was sustained (Ijiri et al., 2008). There are two transcripts of SOX9 in some fish species, where gene duplication of SOX9 has led to the evolution of a male SOX9a1 and female SOX9a2. In the yellow catfish, *Pelteobagrus fulvidraco*, genes associated with female differentiation are SOX9a2 and those associated with the male are SOX9a1, DMRT1 and androgen receptor (Chen et al., 2015). Similarly, two SOX9 transcripts are identified in the zebrafish, *Danio rerio*, as SOX9a, which is expressed during male differentiation along with DMRT1 and AMH, and the female differentiating variant SOX9B which is expressed along aromatase (Jorgensen et al., 2008).

The European eel has been an organism of interest to study sex determination and differentiation because of the unusual pattern and delay in sex differentiation, and similarities in the morphological differentiation of eels and lampreys (see Section 1.4.3, Chapter 3) makes this species of particular interest in this thesis. In the European eel, the undifferentiated gonad will develop into an ovary or an intersex organ known as the “Syrski organ” (Colombo and Grandi, 1996). The Syrski organ contains oocytes that are found within the developing testes, and the oocytes undergo apoptosis to produce the testis. The Syrski organ has only been able to produce a testis because it may have lost the ability to develop into an ovary (Geffroy and Bardonnet,
In the undifferentiated gonad of the eel, expression of genes such as aromatase, the enzyme that converts androgens into estrogens, is expressed at higher levels in females, and a down-regulation is required to allow testis development (Guiguen et al., 2010). In the European sea bass, *Dicentrachus labrax*, the aromatase promoter region in males has twice the amount of methylation than females, which would lead to a down-regulation, promoting testis development (Navarro-Martin et al., 2011).

Understanding genes involved in sex differentiation in fishes will shed light on the evolution and regulation of this process in vertebrates, and is also important for species conservation. In the endangered Iberian cyprinid, *Squalius pyrenaicus*, for example, differences in gene expression were observed during the non-breeding and breeding periods. Genes that are implicated in the female differentiation pathways, such as aromatase, the aromatase-activator *FOXL2*, and the estrogen receptor were found to be highly expressed in differentiating males as well as females (Machado et al., 2016). The expression of *FOXL2* in males is not commonly observed because *FOXL2* increases expression of aromatase and is a common *DMRT1* antagonist (Barrionuevo et al., 2016). Other highly expressed genes that had male biased expression were *Wingless-type MMTV integration site family member 4 (WNT4)* and *R-spondin 1 (RSPO1)*, but the expression of these female differentiating genes became female biased during the breeding period (Machado et al., 2016). Gene expression studies can help determine the pressure of selection on genes involved growth, physiology, and adaptability of fish to their environment (Nielsen et al., 2010). Changes in gene expression patterns are important indicators of the negative impact associated with anthropogenic factors and climate change, ultimately causing shifts in fish populations and sex ratios. Understanding sex specific gene expression patterns is essential for species valued for human consumption and biological conservation, and help
understand differences between behavioural and phenotypic plasticity and is a new approach to study behavioural ecology (Nielsen et al., 2010)

Sex control is also important to maximize productivity in aquaculture. For example, in the commercially important protandrous barramundi fish, *Lates calcarifer*, development into the female gonad does not take place for years and can pose problems for mating for a population of the same age group, as the population is all male for a prolonged period of time. The gene expression patterns of *aromatase CYP11b* and *DMRT1* were observed and used to determine male differentiation onset and completion, at 44-140 days post hatch with future applications for sex control (Banh et al., 2017). Candidate genes involved in sex determination and differentiation and their roles are reviewed in more detail in Section 1.6.

### 1.4.3. Sex differentiation in lampreys

In lampreys, sex differentiation in females occurs during the larval stage, several years in advance of testicular differentiation, which does not occur until metamorphosis (Hardisty, 1965; see Figure 1.1). That the lamprey gonad remains undifferentiated for several years makes them unusual relative to most other vertebrates (Spice et al., 2014), although what is observed in the eel is similar in some respects. In lampreys and eels, sex differentiation is delayed and the gonad first differentiates into an ovary or an intersex organ before eventual differentiation in a testis (Hardisty, 1965; Colombo and Grandi, 1996; see Section 1.4.2). Furthermore, given the phylogenetic placement of lampreys at the base of the vertebrate lineage, understanding more about ancient sex differentiation patterns and the genes involved will allow us to better understand how these patterns have evolved over hundreds of millions of years of vertebrate evolution.
As indicated, the gonad of developing male lampreys can contain both testis and oocytes, and possess the ability to differentiate into an ovary (Hardisty, 1965). Sex reversal in sea lamprey has even been observed where biopsies of males and females were taken at two time points; in some gonads, between time points, the entire stock of oocytes within the ovary underwent atresia and developed into a testis (Lowartz and Beamish, 2000). In lampreys, the presence of multiple germ cell types such as oocytes, undifferentiated germ cells and pre-spermatogonia are considered to be part of the early differentiation phase and is due to the bipotential capacity of the germ cells in the gonad; during subsequent testicular differentiation, the developing testis will then inhibit further oocyte development and induce oocyte apoptosis (Hardisty, 1965). A gene that has been implicated in female inhibition is DMRT1, which suppresses the female gonadal development genes FOXL2, RSPO1, WNT4, and aromatase (Graves 2013). Using quantitative real-time reverse transcriptase polymerase chain reaction (qRT-PCR), aromatase was not successfully amplified in northern brook lamprey, Ichthyomyzon fossor and chestnut lamprey, Ichthyomyzon castaneus, indicating that perhaps lampreys do not use aromatase (Spice, 2013). Through the availability of new genomic resources, we may be able to identify homologues and determine if aromatase is in fact present in lampreys and decipher a role it may play.

Previous work done by the Docker laboratory looked at ovarian differentiation in lampreys. Spice et al. (2014) used qRT-PCR and looked at the gene expression patterns during ovarian differentiation in the chestnut lamprey and the northern brook lamprey. The northern brook lamprey is non-parasitic following metamorphosis (i.e., does not feed at all) and undergoes rapid sexual maturation whereas the chestnut lamprey remains sexually immature until after the parasitic feeding phase; the northern brook lamprey also undergoes earlier ovarian differentiation
and attains a lower fecundity than the chestnut lamprey (Docker, 2009). Differences in gene expression were found between different age classes, with germ cell-less (gcl) expression significantly greater in differentiated northern brook lamprey ovaries. Comparing both species, expression of WT1 and cytochrome oxidase subunit III, was higher in the undifferentiated gonad, and conversely, expression of Deleted in azoospermia associated protein 1 (dazap1) and 17b-hydroxysteroid dehydrogenase was higher in the differentiating gonad. Ajmani (2017) expanded on work by Spice (2013) and looked at transcriptomic analysis of ovarian differentiation in chestnut and northern brook lampreys. She developed two pipelines for studying non-model species (one a genome-guided pipeline using the sea lamprey genome as a reference and the other was a de novo assembly pipeline) to identify suites of genes expressed during different stages of ovarian development (ranging from ovarian differentiation in larvae through metamorphosis) in these two species that show different developmental trajectories with respect to ovarian development. This study was able to identify genes that are up- and down-regulated during ovarian differentiation and maturation in these two species, and identify novel genes (i.e., those not previously characterized in the sea lamprey genome). However, in contrast to this growing body of information on ovarian differentiation in both parasitic and non-parasitic lamprey species, virtually nothing is known regarding the gene expression patterns associated with lamprey testicular differentiation.

Detection of lamprey testicular differentiation genes would provide more insight into gene evolution through the vertebrate lineage and would help determine, based on expression levels, when testicular sex differentiation occurs. However, genes such as AMH and DMRT1 are still unknown or not well characterized in lampreys. To understand whether the role of DMRT1 is similar to what is observed in other vertebrates, Mawaribuchi et al. (2017) used degenerate
primers to detect expression of $DMRT1$ in larvae and post-metamorphic Far Eastern brook lamprey, *Lethenteron reissneri*. They found that $DMRT1$ expression in male testes was significantly greater in post-metamorphic males than females, and in situ hybridization showed a high level of detection of $DMRT1$ in spermatogonial cysts of post-metamorphic males, with no detection in females (Mawaribuchi et al., 2017).

1.5. Testicular differentiation

1.5.1. Spermatogenesis

The focus of this thesis is primarily on testicular differentiation in lampreys; therefore, the process of spermatogenesis is described in detail here. The following brief discussion on the stages of development is summarized from a review paper on spermatogenesis in mammals by de Kretser et al. (1998). Fish species such as zebrafish, guppy, *Poecilia reticulata*, and rainbow trout, *Oncorhynchus mykiss*, can have considerably more spermatogonial generations in comparison to mammals, but the general pattern of development is consistent among vertebrates (Schulz et al., 2010; see Figure 1.1).

Spermatogenesis consists of three stages of divisions of undifferentiated primordial germ cells that lead to the production of mature spermatozoa (Figure 1.2). These divisions can further be separated into mitosis of spermatogonia, meiosis of spermatocytes into spermatids, and spermiogenesis to transform spermatids into motile spermatozoa (de Kretser et al., 1998).

Spermatogenesis is influenced by internal factors in animals that include hormonal interactions. Primordial germ cells become localized to gonadal tissue of the future testis and produce spermatogonia by mitotic divisions. Series of mitotic divisions produce spermatogonia such as type A and type B, with variable spermatogonial types in between various vertebrate species. Spermatogonia are named as type A if the nucleus of the spermatogonia does not contain
any heterochromatin, whereas type B nuclei contain heterochromatin and will appear darker when stained (de Rooij and Russell, 2000). Mitosis is completed once spermatogonia have undergone divisions to form type B spermatogonia.

Following mitosis, type B spermatogonia further divide to produce primary spermatocytes, and this is followed by meiosis. The diploid primary spermatocytes undergo meiosis I to produce two haploid secondary spermatocytes. The first division of meiosis can be further divided into stages that are determined based on the shape and size of the nuclei. The secondary spermatocytes undergo meiosis II to produce four haploid round spermatids.

The final stage of spermatogenesis is spermiogenesis, in which the spermatids that are produced by meiosis undergo drastic transformation to become flagellated spermatozoa. The rounded spermatid nucleus becomes localized to a polar region of the cell and an acrosome is formed and becomes attached to the nucleus. A key characteristic of the spermatozoa is the flagellum, which will provide the mature sperm with the ability to move and propel.

1.5.2. Testicular differentiation in lampreys

Male differentiation does not take place until or at the onset of metamorphosis, and up until this point, the gonad in presumptive males is undifferentiated (Hardisty, 1965). Histological changes during testicular differentiation have been described by Hardisty (1965, 1969, 1971). However, there has been little recent work focusing on testicular differentiation in lampreys. The following description of testicular differentiation in lampreys draws from the work of Hardisty (1965, 1969, 1971) as well as descriptions of testicular differentiation in the well-studied zebrafish, which also has a cystic pattern of gonadal development (Schulz et al., 2010). The gonadal development pattern of fish differs from that of amniote vertebrates, where the Sertoli cells retain the ability to grow and proliferate well into the adult phase of gonad
development, and the Sertoli cells in the gonad develop from a single spermatogonial germ cell (Schulz et al., 2010).

During the larval stage, or **stage 1**, the gonad is consistently larger in surface area when compared to other stages. The types of germ cells that are present are undifferentiated germ cells and some germ cells that appear to resemble type A spermatogonia. The criterion that is used to determine the changes in germ cell type is based on the size of the germ cell, where a change from type A undifferentiated to differentiated does not necessarily encompass a change in size of the gonad as a whole, but instead a more clearly defined outline to the germ cell and prominent, rounded nucleoli (Hardisty, 1971). At this stage, the gonad is starting to increase in size as the number of germ cells increase; however, some germ cells are expected to start degenerating when the larva reaches a length of about 90 mm (Figure 1.1). The degeneration of germ cells occurs in both males and females, with degeneration occurring at the same time as oocyte growth in females; however, this process is much more pronounced in males as there are no oocytes present to mask the extensive degeneration (Hardisty, 1965). In the current study (Chapter 3), samples from the larval stage, from age class 3 and 4 (as inferred from larval length) were included to account for when female differentiation occurs, which is during age 3 (starting at lengths of approximately 80-90 mm), with continuous oocyte growth up to lengths of 100 mm (Docker, 1992). By age classes 3 and 4, presumptive males could be distinguished from differentiated females.

**Stage 2** of testicular differentiation includes a transition of cell types seen in stage 1, as well as changes in the gonad unique to stage 2, accompanied by changes in the size and shape of the gonad. The majority of germ cells that are found in stage 1 and stage 2 of differentiation are of type A and B based on male lamprey histological observation, at the onset of metamorphosis,
and immediately following. Type A differentiated spermatogonia are found in cysts with 2-8 germ cells in each cyst and the gradual transition to type B in stage 2 shows a pronounced decrease in nuclear size, as well as a change in the distribution of chromatin and cytoplasm structure (Schulz et al., 2010). The amount of loose reticulum connective tissue has increased, which is indicative of testis differentiation (Hardisty, 1965). Type B spermatogonia occur in cysts with 16 or more germ cells that show a continuous decrease in cell volume and size and an increase in heterochromatin while continuing mitotic divisions (Schulz et al., 2010). The drastic decrease in the gonadal size during the larval phase is because of the decrease seen in the number of germ cells.

Mitosis is completed after type B spermatogonia have undergone their specified number of divisions, leading up to meiosis and the development of primary spermatocytes. The characteristics that are used to identify meiotic cells are based on the diameter, shape, and chromosome structure of the nucleus, with primary spermatocytes being more rounded than type B spermatogonia (Schulz et al., 2010). **Stage 3** of testicular differentiation, or the mid-differentiation phase, has germ cells that are of type B and primary spermatocytes. The structure of the testis has drastically changed from the early stages, as now the gonad has gained a lobed structure, with finger-like lobes housing clusters of germ cell nests (Hardisty, 1965, Hardisty, 1969). The size of the gonad is still quite small and continues to decrease as expected during metamorphosis (Hardisty, 1965). In meiosis I, the diploid homologous chromosomes in primary spermatocytes segregate into haploid secondary spermatocytes, and after the completion of meiosis II, sister chromatids segregate and result in four haploid spermatids.

In **stage 4**, the final stage of testicular differentiation, the majority of germ cells that were observed in zebrafish were primary spermatocytes with densely-pigmented nuclei (Schulz et al.,
2010). Compared to the last two stages, the lamprey gonad has increased in size, with a loss of lobation, a decrease in connective tissue, and an increase in the number of total germ cells that are present, as expected to occur during testis differentiation (Hardisty, 1965). The primary spermatocytes can be distinguished based on their dense staining and have large nuclei. Although the gonad does not appear too dissimilar in size from the undifferentiated gonad, one distinguishing feature that is used to differentiate them is the presence of large nuclei during later stages (Hardisty, 1965) in lampreys, between stages 2 and 4 during metamorphosis (Figure 1.1).

1.6. Candidate genes for sex determination in vertebrates and conserved sex differentiation genes in lampreys

Genes that have been shown to play a master role in vertebrate sex determination (in one or more species) are as follows: SOX2, SOX4, SOX8, SOX9, SOX17, sdY, AMH, AMHR2, WT1, FEM-1, SF1, DMRT1, FOXL2, GSDF, WNT4, GCL, RSPO1, and DAZAP1. Genes that are involved in male sex differentiation in other vertebrates are: SOX9, DMRT1, AMH, AMHR2, SF1, WT1. See Table 3.1 (Chapter 3) for more information on these genes.

SOX genes

_SRY_ box, sex-determining region on Y chromosome-type high-mobility group (HMG) box, or SOX genes are vital during organ development and have tissue specific expression (Wegner, 1998). The SOX family contains seven subgroups, A to G, that are determined based on their degree of similarity within the HMG domain between proteins and have a strong degree of conservation between different SOX genes. A high degree of conservation is observed not only within the HMG domain, but also in areas adjacent on the C or N-terminus or both (Wegner, 1998). These genes play a vital part in normal development, being expressed in tissue during at least one developmental stage, and mutations or deletions result in severe
developmental defects. SOX8, SOX9, and SOX10 belong to the group E SOX gene family and have roles in sex determination of mice and humans (Wegner, 1998). Genes that are known to determine sex in one organism may have a different function in another (Wegner, 1998). In chickens, Gallus gallus, SOX3, SOX4, SOX8, and SOX11 were not differentially expressed between male and female gonads; however, SOX9 was highly expressed in males (Takada et al., 2005).

Upon the cloning and characterization of SOX2 from the testis of Indian major carp, Labeo rohita, expression was detected in spermatogonial stem cells (Patra et al., 2015). SOX5 and SOX6 expression have been observed in high concentrations in rounded spermatids, indicating this occurs post-meiotically before spermiogenesis in male mice (Denny et al., 1992; Connor et al., 1995; see Section 1.5 for more information on the process of spermatogenesis).

The SOX8 gene is present in both male and female gonads of the red-eared slider turtle, Trachemys scripta elegans, and chicken, Gallus gallus domesticus, but expression of this gene in other vertebrates is strictly indicative of testis differentiation (Takada et al., 2005). In mice, SOX8 is expressed during gonadal development, where it activates AMH and SF1 and has also been associated with up-regulating AMH in males and down-regulating AMH in females (Schepers et al., 2003; Takada et al., 2005). In the mouse, SOX8 and SOX9 work synergistically to ensure testis development. In experiments with SOX9 ablation, SOX8 is able to temporarily compensate. Seminiferous tubules continue differentiation and testis development follows the expected trend, but mature males are sterile due to spermatogonial failure (Barrionuevo et al., 2009). In more ancient sarcopterygians, such as the African coelacanth, Latimeria chalumnae, SOX8 expression was not detected in the testis, indicating a testis-specific role later on in the vertebrate lineage (Forconi et al., 2013).
SOX9 has been linked to testis development as induced exposure of SOX9 in an XX mouse led to testis development and due to its known up-regulation of AMH in chick embryos (Vidal et al., 2001; Takada et al., 2005). In humans, SOX9 gene mutations hinder normal signaling pathways between genes during male development, which leads to XY sex reversal (Wagner et al., 1994). Of the three male human embryonic stem cell lines that were analyzed, all three layers expressed SOX9, highlighting the importance of this gene in normal development for all tissues (Kjartansdottir et al., 2015). High levels of AMH precede and up-regulate SOX9 expression in males and are linked to male sex determination in some species (Oréal et al., 1998; Takada et al., 2005; see Section 1.3 for more details). SOX9 is expressed in both males and females, but expression becomes exclusive to males upon expression of SRY, after which SOX9 is expressed all through testis development (Wegner, 1998). SOX9 and DMRT1 were expressed in male lambari fish, Astyanax altiparanae, but following spermiogenesis, expression was down-regulated, indicating their roles may be restricted to spermatogenesis and spermiogenesis (Adolfi et al., 2015). Male specific expression was also observed in chickens, where females had no expression at day 7 after hatch, a critical period for sexual dimorphism (Romanoff, 1960; Takada et al., 2005).

SOX17 is expressed in pre-meiotic spermatogonia in mice (Kanai et al., 1996). SOX17 has also been identified as a key regulator of primordial germ cell fate determination in humans and mice, determining whether the cells will go on to produce sperm or eggs (Irie et al., 2015). In the sea bass, Dicentrarchus labrax, a transcript of SOX17 was highly expressed in the gonad. The expression levels increased at day 150, which is at the same time as sex differentiation. The concentrations were the highest in females later in differentiation, indicating a potential role of SOX17 during early male differentiation (Navarro-Martin et al., 2009). In three male human
embryonic stem cell lines, *SOX17* was expressed in the male primordial germ cell that form spermatogonia, suggesting that it plays a role during spermatogenesis (Kjartansdottir et al., 2015). In the rice field fish, *Monopterus albus*, *SOX17* is expressed in female gonads, the undifferentiated gonad, and spermatogonial cells (Wang et al., 2003).

**sdY or SRY**

Sex determining region Y has been linked to sex determination processes in the rainbow trout (Yano et al., 2012). This is a relatively new gene in the realm of sex determination and it bears a great degree of resemblance to *SOX3* (Wegner, 1998). This indicates a recent evolution event between *SRY* and *SOX3*, although *SOX3* expression occurs within the cells of early urogenital ridge development in both male and female vertebrates, but has been lost in marsupials (Foster and Graves, 1994; Wallis et al., 2008). In the Indian ricefish, *Oryzias dancena*, *SOX3* is involved in testicular differentiation by up-regulation of *GSDF*, and loss of *SOX3* leads to XY reversal (Takehana et al., 2014). *SRY* protein is highly diverged in comparison to other SOX proteins where divergence can be determined by the observation of introns present in the ancestor, and those with additional introns indicate more recent divergence (Swain et al., 1998; Wegner et al., 1998).

**AMH and AMHR2**

*AMH* plays an important role in the development of the male urogenital tract by inducing the regression of Müllerian ducts, which would have given rise to the uterus, fallopian tubes, oviducts and vagina (Josso et al., 2001). Unlike vertebrates, fish do not have Müllerian ducts, but a homologue of *AMH* was identified in the Japanese eel, *Anguilla japonica* (Miura et al., 2002; Ijiri et al., 2008). Eel spermatogenesis related substances 21, *eSRS21*, amino acid sequence shares homology with *AMH* in mammals and chicken (Miura et al., 2002). Anti-Müllerian
hormone receptor 2 (AMHR2) has been determined to be one of the genes involved in sex determination in fugu, *Takifugu rupripes* (Myosho et al., 2012). In a natural population of fugu, the males had a single-nucleotide polymorphism, SNP, of GC within the kinase domain of *AMHR2*, and the females were homozygous for nucleotide C. The difference is sex specific and indicates that perhaps this missense mutation is responsible for sex determination (Kamiya et al., 2012). Mutations in the *AMHR2* have been reported to lead to sex reversals greater than 50% of the time in XY male medaka (Morinaga et al., 2007). *SOX* genes and Steroidogenic factor 1, *SF1*, induce up-regulation of *AMH* in chick embryos at day 6 and 7 (Josso et al., 2001; Takada et al., 2005).

**WT1**

*WT1* is known to be expressed during sex differentiation in mice (Pelletier et al., 1991). *WT1* was expressed at high levels in the catfish, *Clarias batrachus*, testis, and localisation indicated presence in spermatogonia, spermatocytes and Leydig cells (Murugananthkumar and Senthilkumaran, 2016). During gonadal development, *WT1* was up-regulated in presumptive male lampreys (Spice et al., 2014). Some females, however, had similar amplification of this gene, perhaps indication of expression in bipotential gonads. Further histological studies will help to distinguish presence or absence of spermatogonia in the gonad as this can be due to incomplete secondary development (Shapiro, 1987).

**FEM-1**

*Feminization 1 (FEM-1)* plays a masculinizing role in male nematode *Caenorhabditis elegans*, where mutant *FEM-1* mutants produce all females, highlighting the important role of *FEM-1* proteins to produce male specific tissues (Doniach and Hodgkin, 1984; Goodarzi et al., 2008). In the male sex determination pathway of *C. elegans*, *FEM* family of proteins, (*FEM1, 2*...
and 3) promote the male pathway by suppressing female sex determination genes and up-regulating genes required for spermatogenesis (Zanetti and Puoti, 2013). Ventura-Holman et al. (2003) have identified FEM1 protein homologues that are present from insects, mice, zebrafish and humans. On the contrary, FEM-1 is expressed in the eggs and embryos of the oriental river prawn, Macrobrachium nipponense, and is expected to play a role in ovarian differentiation as well as female sex determination (Ma et al., 2016). However, FEM-1 does not appear to play a sex determination or sex differentiation in mice and humans (Ma et al., 2016). The amino acid sequences for C. elegans, mice, zebrafish, and humans are roughly 80% identical, and the pathway that determines its role in the nematode has been understood (Ventura-Holman et al., 1998). This information can shed light on the signal pathways that determine cell-fate in lamprey gonadal differentiation as well.

**SF1**

SF1 is a transcription factor encoded by the nuclear receptor subfamily 5, group A, member 1 gene, expressed during sex differentiation in the developing testis of male mice (Ikeda et al., 1994; Takada et al., 2005). SF1 is found in Leydig and Sertoli cells and knockdown experiments show that loss of SF1 leads to the loss of gonadal structural integrity (Kato et al., 2012). In the presence of SF1, AMH is both activated and up-regulated in male gonad that causes regression of the Müllerian ducts (Josso et al., 2001; Wegner, 1998). In mammals, SF1 and SRY together activate SOX9 in the male gonad and are key for testis differentiation (Sekido and Lovell-Badge, 2008).

**DMRT1**

Double-sex and male abnormality-3, DM domain gene on the Y chromosome related transcription factor 1, belongs to the DMRT gene family and its role is equivalent to the SRY
genes in mammals (Kikuchi and Hamaguchi, 2013). The expression of DMRT1 has been detected in the male mouse as it develops and medaka is one of the few non-mammalian species in which sex determination processes involving DMRT1 have been found in the testis (Oréal et al., 2002; Matsuda, 2005; Kikuchi and Hamaguchi, 2013). In the loggerhead turtle, Caretta caretta, and red-eared slider turtle, sex can be experimentally determined during development by manipulation of incubation temperature, with an equal number of individuals being produced at an intermediate temperature and a higher prevalence of males and females being produced at low and high temperatures, respectively (Woolgar et al., 2013; Mork et al., 2014). In the red-eared slider turtle, there is temperature dependent expression of DMRT1 at low temperatures, which is known to produce males. DMRT1 is expressed more highly in males at low temperatures in the genital ridge of turtle embryos and promotes male development (Kettlewell et al., 2000). DMRT1 is one of the key sex determining genes in some fish (e.g., some but not all Oryzias species), being expressed in somatic cells around the testis during development (Matsuda et al., 2002; Kikuchi et al., 2013). During the larval stage in the lambari fish, there is no expression of DMRT1, but expression is detected in later stages of gonadal development (58 dah). During gonadal development and analysis of expression levels in both male and female lambari fish, both DMRT1 and SOX9 are expressed specifically within the testis (Adolfi et al., 2015). In the genetic female mouse, DMRT1 transgene introduction in the ovary leads to the silencing of FOXL2 and re-programming of the gonadal cell tissues to testis, even in the absence of other genes, SOX8 and SOX9, which are thought to be essential for masculinization (Lindeman et al., 2015).
**FOXL2**

*FOXL2* is thought to have ovary specific expression in medaka (Nakamoto et al., 2006). Similarly, *FOXL2* expression was consistently higher in female Nile tilapia than in males and is believed to act antagonistically with *DMRT1*, where an increase in one leads to a decrease in the other (Ijiri et al., 2008). *FOXL2* deficient females exhibited oocyte atresia and a decrease in aromatase activity, and some fish exhibited sex reversal due to the higher expression of *DMRT1* (Li et al., 2013). The opposite situation is observed in the male tilapia, where *DMRT1* deficiency leads to testis regression, degeneration of spermatogonia, and an increase in *FOXL2* expression (Li et al., 2013). In the mouse gonad, certain genes are required to remain active to maintain the integrity of the gonad, and in females with removal of *FOXL2*, reprogramming of the gonad is seen, with the ovary differentiating to a testis (Barrionuevo et al., 2016).

**GSDF**

*GSDF* is a secretory protein, part of the transforming growth factor-beta family (TGF-β) (Sawatari et al., 2007). In the rainbow trout, expression of *GSDF* is essential for proper gonadal development and can be detected in PGCs in the genital ridge, where it is required for germ cell proliferation, and also in Sertoli cells, for spermatogonial differentiation (Sawatari et al., 2007). *GSDF* is expressed early in development of fish such as medaka, where it is expressed exclusively in males in the undifferentiated gonads (Shibata et al., 2010). In medaka, it has been established that *dmy* is the sex determining gene, but instead of being directly involved, *dmy* recruits *GSDF*, and *DMRT1* which initiates testicular differentiation (Zhang et al., 2016). This would change the status of *GSDF* in medaka and classify it as a sex determining gene, similar to its role in the Luzon ricefish (Kamiya et al., 2012).
**WNT4**

WNT4 signaling pathways involve interactions with SOX genes (Wegner, 1998). WNT4 has been found in the undifferentiated gonad along with RSPO1, but they are significantly up-regulated when the WNT cascade is activated, leading to ovarian development in mammals (Chassot et al., 2012). WNT4 is believed to act as an “anti-male factor” by reducing serum androgen concentration to prohibit production of testosterone, and as a result, a loss of testis germ cells is seen in mice (Jordan et al., 2003). The role of WNT4 in fish has been poorly understood until recently, when cDNA of WNT4a and WNT4b was extracted from the mature ovary of the orange-spotted grouper, Epinephelus coioides. The two extracted sequences were highly expressed during early ovarian differentiation, and the sequences are also conserved among vertebrates, clustering with other WNT4 genes in mammals and teleost fishes (Chen et al., 2015).

**GCL**

Germ cell-less plays an important role in germ cell differentiation in Drosophila melanogaster, where it is essential for germ cell formation (Jongens et al., 1992). If a female D. melanogaster has reduced expression of GCL, she can either fail to have progeny or the offspring will be sterile at maturity (Jongens et al., 1992). Recently, the precise mode of action for GCL has been identified, where it is found in the nuclear envelope in Drosophila, based on its binding behaviour to a lamina-associated polypeptide 2 β, LAP2β (Lerit et al., 2017). The mouse homologue of Drosophila melanogaster germ-cell less 1 (mgcl-1) gene was knocked down in mice, and abnormalities in normal growth of the liver and pancreas were evident, with the most profound effect found in the testis (Kimura et al., 2003). The structure and function of the sperm were severely deformed with abnormalities in the acrosome, irregular chromatin condensation,
and abnormal expression of proteins associated with chromatin structure during spermatogenesis (Kimura et al., 2003). This highlights the importance of the role of GCL not only for normal sperm development but also for primordial germ cell growth and differentiation (Lerit et al., 2017). A homologue of GCL was found in medaka, and upon gene expression analysis, the ovary showed higher expression than the testis and tissues such as brain and liver. To further investigate the differences observed, in situ hybridization revealed expression in the medaka ovary but not the testis (Scholz et al., 2004). GCL was expressed in the female northern brook lamprey, with higher expression in age class 2 than in age class 1 (Spice, 2013). Further research on the role GCL plays in fish and, specifically, male lampreys needs to be investigated.

**RSPO1**

RSPO1 is a gene involved in multiple development and signaling pathways, and relatively conserved in vertebrates. In fish, it has been identified in zebrafish, where it is involved in the female differentiation β-catenin activation, and the induction of RSPO1 in the developing XY male medaka leads to male to female sex reversal (Sun and Zhang, 2012; Zhou et al., 2016). RSPO1 is a sex determination gene in mice, and is also expressed early in the undifferentiated mouse female gonad, and in synergism with WNT4, leads to ovarian differentiation (Garcia-Ortiz et al., 2009; Chassot et al., 2012). If mutations in the RSPO1 gene are found in developing males, abnormalities can lead to the loss of testis differentiation and sex reversal in mammals (Barrionuevo et al., 2012). Likewise, double RSPO1 mutants show female to male sex reversal (Parma et al., 2006; Parma and Radi, 2012).

**DAZAP1**

DAZAP1 is the protein that is associated with the Deleted in Azoospermia, DAZ, family of genes, which are found in gonads exclusively in a variety of animals including fruit fly,
roundworm, African clawed frog, *Xenopus laevis*, zebrafish, mice and humans (Fu et al., 2015). Expression of DAZ proteins has been largely restricted to specific cell types in the male germ cell line such as pro-spermatogonia and spermatogonia, indicating a role in pre-meiosis (Fu et al., 2015). Although in humans four DAZ proteins have been identified, and it was thought that all are essential for proper sperm production, evidence shows that even with the expression of one of four DAZ proteins, sperm production can still occur, although sperm count is lower than normal (Huang et al., 2008). The expression of DAZ and associated genes are also found to a lesser extent in ovaries of mice and humans, playing roles in reproductive cycles and ovary development (Pan et al., 2005). In mice, the highest expression of DAZAP1 is observed in individuals when spermatogenesis and spermiogenesis are completed. Deficiencies in DAZAP1 have been associated with abnormalities in spermatogenesis that lead to failure of sperm production (Yang and Yen, 2013).

1.7. Thesis objectives

The two main related objectives of this thesis were:

1. **Screen for candidate sex determining genes in the sea lamprey and Pacific lamprey genomes (Chapter 2).** Conserved genes involved in sex determination in other species (see Section 1.5), such as srY, sdY, AMH, AMHR2, SOX2, SOX8, SOX9, SOX10, SOX 17, DMRT1, SF-1, GSDF, TRA-1, RSPO1, WT1, WNT3/4/5, FOXL2, and FEM1, were examined. Primers for these candidate sex determination genes were designed using the sea lamprey genome as a reference and, using sea lamprey and Pacific lamprey genomic DNA, I tested for sex specific amplification of genes and nucleotide sequence differences between males and females. Genes displaying sex specific differences would represent genes with potential roles in genetic sex determination and can be further studied. However, considering the variability in sex
determining genes observed to date in non-amniotic vertebrates, and even among closely related species, it is very conceivable that none of these genes are involved in the sex determination process in the sea and Pacific lampreys or that the sex determination genes are not the same in these two lamprey species. However, it is equally important to rule out the sex specificity of these candidate sex determination genes in lampreys.

2. **Determine whether genes involved in testicular differentiation in other vertebrates are expressed in the sea lamprey in a sex specific and/or stage specific manner (Chapter 3).**

Some candidate genes involved in male testicular sex differentiation in vertebrates include: *SOX8, SOX9, DMRT*-related gene, *SF1, GCL, WT1*, and *DAZAP1* (see Section 1.5). Given that genes involved in sex differentiation are relatively conserved among vertebrates, I hypothesized that these genes will have similar roles in the sea lamprey during testicular differentiation. I designed primers for these genes using the sea lamprey genome as a reference, and measured gene expression in the testis or presumptive testis during the late larval stage (stage 1), throughout metamorphosis (stages 2-4), and in sexually mature males (Figure 1.1). Levels of gene expression were compared among stages of male differentiation, and between males and females, which undergo differentiation during the larval stage. Gene expression was measured using qRT-PCR, and a subset of samples were examined histologically to better characterize the processes that were happening in the testis at each stage of development and to relate these changes to gene expression.

### 1.8. Significance

Lampreys are one of the two surviving lineages of agnathans found at the pivotal emergence point of vertebrates and can be traced back over 360 million years, with over 41 extant lamprey species (Docker, 2006; Gess et al., 2006; Potter et al., 2015). Although an
immense amount of work has looked at lamprey physiology, development, phylogeny, and the histological process of sex differentiation, very little is known about the genetic basis of sex determination and differentiation in lampreys (McCauley et al., 2015). In terms of what is observed during sex differentiation in other fish, they undergo differentiation much sooner than in lampreys, which can be delayed for up to several years (Hardisty, 1965). Hormonal factors involved in lamprey sex differentiation are still unknown, and using sex steroids as identified from other vertebrates on lampreys for sex ratio manipulation has not proven to be successful, indicating the role of more ancient forms of sex steroids employed by vertebrates today (Docker, 1992). Other recent and exciting studies are also contributing to our understanding of sex determination and sex differentiation in vertebrates. For example, with the discovery of a species that was thought to be extinct, the African coelacanth, and the genes that are involved in the processes related to sex differentiation, it has been determined that genes involved in the sex differentiation processes in other vertebrates that were believed to not have arisen before this species and restricted to teleosts, were in fact present (Forconi et al., 2013). It is unknown if these shared genes are present in the more distantly related lampreys, but studying these genes in lampreys can further shed light onto the genetic basis of sex determination and differentiation in vertebrates. Due to the position of lampreys at the base of the vertebrate phylogeny, and with considerably more research to date on lampreys than on hagfishes, lampreys are an essential tool for understanding vertebrate evolution (Docker et al., 2015; McCauley et al., 2015). Using the available genomic information from the Arctic lamprey, Lethenteron camtschaticum, and the recently sequenced sea lamprey, an incredible advancement has taken place on the whole genome duplication events indicating that three rather than two whole genome duplication events may have occurred in lampreys (Mehta et al., 2013; Smith et al., 2013; McCauley et al.,
2015). Therefore, studying lampreys is not only fascinating in terms of their ancient roots and adaptations that have allowed them to persist for hundreds of millions of years, but also essential for understanding the evolution of processes in other vertebrates.

In addition to the evolutionary implications of this research, there are potential applications to conservation. The revolution of genetic technology and the ability to knock-down target genes that render organisms sterile (e.g., parasitic and invasive insects) provides a highly specific target compared to the conventional methods to control invasive species such as the parasitic sea lamprey (Franz and Robinson, 2011; Lee et al., 2013; Dong et al., 2016). A sterile-male release technique has been developed in sea lamprey, but the chemosterilant used has sterilizing effects on not only lampreys, but is harmful to many vertebrates, including humans (Docker et al., 2003; Bergstedt and Twohey, 2007). Further understanding of lamprey reproduction and the genes that are expressed during sex determination could allow us to sterilize or manipulate population sex ratios for the control of invasive sea lamprey, using lamprey specific genes (Docker et al., 2003; Sower et al., 2003). Of greatest immediate use for conservation and management would be genetic markers that allow for early identification of sex, since histological sex identification is not possible until individuals are several years old (Figure 1.1) and sex identification using external secondary sex characteristics is not possible until shortly before spawning (Johnson et al., 2015). See Chapter 4 for more details on future directions for lampreys in the genomic era.
1.9 Literature cited


de Rooij, D. G., & Russell, L. D. (2000). All you wanted to know about spermatogonia but were afraid to ask. *Journal of Andrology, 21*(6), 776–798.


Jepson, P. D., Deaville, R., Barber, J. L., Aguilar, À., Borrell, A., Murphy, S., … Law, R. J. (2016). PCB pollution continues to impact populations of orcas and other dolphins in European waters. *Scientific Reports, 6*, 1–17. [http://doi.org/10.1038/srep18573](http://doi.org/10.1038/srep18573)


72

Figure 1.1. Overview of the sea lamprey life cycle and timing of key events related to sex differentiation. Note that the gonad remains undifferentiated for several years and that the differentiation process is asynchronous in males and females; ovarian differentiation in sea lamprey occurs during the larval stage, generally at age 2–3 (dependent on size) whereas testicular differentiation does not occur until sometime just prior to or during metamorphosis. Approximate time spent in each stage is indicated below each life stage. Stages sampled from larval (1), metamorphosing (2, 3, 4), and sexually mature adults (mature) for gene expression analysis in Chapter 3 are indicated; stages 2–4 are thought to represent early, mid, and late stages of testicular differentiation. Cell types observed during testicular differentiation are included below each stage (see Chapter 3).
Figure 1.2. Developmental process of spermatogenesis in the sea lamprey (see Section 1.5.1). Images of spermatogonia, primary spermatocyte and spermatids are taken from histological images used in this study (see Figure 3.8).
CHAPTER 2: Candidate sex determination gene sequencing in lampreys: Pacific lamprey, *Entosphenus tridentatus* and sea lamprey, *Petromyzon marinus*

2.1. Abstract

Sex determination mechanisms are well understood in birds and mammals, but are far less understood in the non-amniotic vertebrates. In fish, sex determination is a malleable process shaped by various internal and external factors, where genetic factors, which can be highly variable among the same or closely related species, and the environment work synergistically to determine sex. Sex determination mechanisms in basal vertebrates are even more poorly understood. Previous studies on lampreys have suggested and explored density-dependent, environmental and genetic-chromosomal modes of sex determination, but results have been inconclusive. Therefore, this study has taken advantage of the newly available sea lamprey genome and revisited the question of the genetic basis of sex determination in two lamprey species: sea lamprey, *Petromyzon marinus*, and Pacific lamprey, *Entosphenus tridentatus*. Specifically, I evaluated whether 19 candidate sex determination genes identified from other vertebrates show sex specific presence or absence patterns or nucleotide differences in sea and Pacific lampreys. Of these 19 genes, six (*sRY, sDY, GSDF, TRA-1, AMH, AMHR2*) were found to not have homologues in the sea lamprey genome and two (*SOX10, SOX17*) had homologues but primers designed for them repeatedly failed to amplify in both lamprey species. The remaining 11 genes (*SOX2, SOX8, SOX9, WT1, FEM-1, SF1, DMRT1, FOXL2, WNT3, WNT5, and RSPO1*) were successfully amplified from sea lamprey genomic DNA; seven of these 11 also amplified in Pacific lamprey (*SOX8, SOX9, FEM-1, DMRT1, FOXL2, WNT3 and WNT5*). However, there were no sex specific differences in any of these candidate genes in either species. This study did, however, identify previously unannotated genes in the sea lamprey genome, and
presumptive identities were assigned to DMRTA2, FOXL2, NR5A1/ SF1, and WT1. Despite the results of this study, a genetic basis of sex determination cannot be ruled out for lampreys. Considering the variability in sex determining genes observed to date in other fishes, even among closely related species, it is conceivable that other genes are involved in the sex determination process in lampreys. Furthermore, the reference sea lamprey genome was constructed from a female, and only gene fragments were sequenced in the current study. Further whole-genome sequencing of males and females with accurate annotations will provide conclusive support for the absence or presence of genetic sex determination in lampreys.
2.2. Introduction

Sex determination is defined as the commitment of the gonad to a male or female fate that is influenced by various factors (Parma and Radi, 2012). Sexual dimorphism in adults is largely a result of hormone interactions, but increasing evidence in vertebrates shows the influence of genetically inherited factors leading to differences resulting in male and female sexes (Bermejo-Álvarez et al., 2011). Birds and mammals all exhibit genetic sex determination (GSD), but sex determination mechanisms in other vertebrates are highly variable (Cutting et al., 2013). Environmental sex determination (ESD) is a common mode of sex determination in some non-amniotic vertebrates (most notably many reptiles), where sexual fate is determined at a critical period in embryo development based on environmental factors such as photoperiod and temperature (Charnov and Bull, 1977; Schwanz et al., 2016; see Section 1.3). ESD, specifically temperature-dependent sex determination (TSD) has also been suggested in approximately 60 species of fish (e.g., Patagonian pejerrey, Odontesthes bonariensis; Atlantic silverside, Menidia menidia; and European sea bass, Dicentrarchus labrax; Yamamoto et al., 2014; Duffy et al., 2015; and Palaiokostas et al., 2015, respectively), although a recent analysis showed that TSD is not as common in fish as previously speculated (Ospina-Álvarez and Piferrer, 2008). Unlike reptiles, however, no fish species have been found that exclusively display ESD (Yamamoto et al., 2014; Duffy et al., 2015; Palaiokostas et al. 2015); ESD and GSD are not mutually exclusive. Even in fish species where GSD has been clearly demonstrated, the sex determining genes are often unknown and frequently variable among species (see Sections 1.3.2, 1.6). In birds and mammals, the master sex-determining genes (DMRT1 and SRY, respectively) are conserved within each taxon (Cutting et al., 2013), but many different genes have been identified that lead to testis development in various fish species. For example, certain salmonid fish species
have sex determination based on the presence of the sdY gene in males, which is not present in females (Yano et al., 2012; Yano et al., 2013), whereas in fugu or tiger pufferfish, *Takifugu rubripes*, sex specific differences in the *AMHR2* gene determines sex (Kamiya et al., 2012). *Dmy* acts as the testis determining gene in the medaka or Japanese ricefish, *Oryzias latipes*, and the Malabar ricefish, *Oryzias curvinotus* (Matsuda et al., 2002; Nanda et al., 2002), but *gsdfy* appears to be the master sex determination gene in a third species in this genus, Luzon ricefish, *Oryzias luzonensis*, (Myosho et al., 2012). In Nile tilapia, *Oreochromis niloticus*, *DMRT1* was found to have male specific role (Kobayashi et al., 2003).

Sex determination mechanisms in basal vertebrates, such as lampreys, are even more poorly understood. Lampreys are ancient fish that spend approximately 3-8 years in the larval phase (see Section 1.1), where sex may be determined by internal and external interactions, but at what time exactly sex is determined, or if a genetic component is involved in this process is still unknown (Docker, 1992; Docker et al., 2003; Clemens et al., 2010). Correlations between lamprey abundance or density and the proportion of males in a population have led to suggestions of density-dependent sex determination in lampreys (Torblaa and Westman, 1980; Docker and Beamish, 1994; see Section 1.3.3) and skewed sex ratios are not uncommon in lampreys (Hardisty, 1965). However, experimental results have been inconclusive; Docker (1992) found no significant differences in sex ratio of previously undifferentiated larval sea lamprey, *Petromyzon marinus*, reared for 3 years at four different densities (Docker, 1992). A recent field study suggested that unproductive environments were associated with slower growth rate and strongly male-biased sex ratios (Johnson et al., 2017) but, again, evidence for a direct environmental effect on sex determination is difficult to demonstrate.

A genetic mode of sex determination for lampreys has not been extensively explored
Mateus et al. (2013) analyzed the genomic DNA using RAD loci of both male and female European brook lamprey, *Lampetra planeri*, to observe any deviation in chromosome differences that would be indicative of physically extensive genomic differentiation between males and females. No such differences were found, but this does not rule out subtle genetic differences between the sexes. In fugu, a single nucleotide substitution in the *Anti-Müllerian hormone receptor 2 (AMHR2)* gene appears to be the only genetic polymorphism associated with sex (Kamiya et al., 2012).

This study will take advantage of the newly available sea lamprey genome (Smith et al., 2013) to further investigate the question of the genetic basis of sex determination in two lamprey species: sea lamprey and Pacific lamprey, *Entosphenus tridentatus*. It will evaluate whether 19 candidate master (*sRY* and *sDY*) and sex determination (i.e., *AMH, AMHR2, SOX2, SOX8, SOX9, SOX10, SOX17, WT1, FEM-1, SF1, DMRT1, FOXL2, GSDF, WNT3, WNT5, TRA-1 and RSPO1*) genes identified from other vertebrates show sex specific differences in amplification (i.e., related to presence or absence or fragment size) or nucleotide differences in these two lamprey species. Details about the roles of these genes are reviewed in Sections 1.3 and 1.4.

Both the sea lamprey and Pacific lamprey are North American parasitic species, and mitochondrial DNA (mtDNA) suggests that *Petromyzon* and *Entosphenus* diverged approximately 9-13 million years ago (Docker et al., 1999). Given the variability in sex determining genes observed to date in other fish and even among closely related species, I hypothesized that none of these genes will be candidate sex determining genes in lampreys or, if they are, that they are not necessarily the same in two species. Until recently, there had been relatively little research going into the study of the Pacific lamprey, and most of the information was applied based on what is known from the Great Lakes sea lamprey (Clemens et al., 2010).
order to conserve populations of Pacific lamprey (see Section 1.2.2) and identify sex specific trends, more research will need to be conducted on identifying genes involved in sex determination using the genome of the sea lamprey and other vertebrates.

2.3. Methods

2.3.1. Sample collection

Pacific lamprey were collected from Bonneville Dam, Columbia River main stem, Oregon or Washington from an artificial Pacific lamprey propagation project (Table 2.1). The 60 Pacific lamprey fin-clipping tissue samples (30 male and 30 female adults) were collected by Ralph Lampman and sent to the University of Manitoba by Dr. Jon Hess from the Columbia River Inter-Tribal Fish Commission (CRITFC).

Sea lamprey samples where the lamprey was at a stage of development where sex could be confidently determined were used. Sex was determined by internal inspection of gonads in individuals with a body length that exceeded 110 mm (see Figure 1.1). At this stage of gonad development, the ovary is much larger and takes up a much greater proportion of the body cavity than the testis (Docker, 1992). Larval and metamorphosing sea lamprey were collected in Richibucto River, New Brunswick, by collaborator Dr. Mike Wilkie (Wilfred Laurier University) and members of his laboratory; additional larval and adult samples were collected in the Great Lakes basin near the Hammond Bay Biological Station, Millersburg, Michigan, by Dr. John Hume (Michigan State University). Larval samples were collected from the Au Sable River in northern Michigan, and adult sea lamprey were collected from one of three rivers in northern Michigan: the Ocqueoc River or the Cheboygan River, which both drain into Lake Huron, or Carp Lake River which drains into Lake Michigan. The adult lampreys were collected by the US Fish and Wildlife Service (USFWS), and housed communally so exact sample collection date
and location information is not available. Samples were collected using electro-fishing and stored in RNA stabilization solution, were shipped by courier to the University of Manitoba, and kept frozen at -20°C until required for use. A 1 cm² epidermal sample was taken using a razor blade from the dorsal right hand side of the sea lamprey carcass, which was then immediately used in the appropriate DNA extraction protocol (Qiagen).

2.3.2. DNA isolation and RNase elimination

All the equipment that was to be used such as gloves, 1.5 mL Eppendorf tubes, tube holders and DNeasy spin columns were UV treated for 15 minutes prior to use and the extractions were done in a AC600 PCR workstation with ISO 5 HEPA-filtered air and UV light irradiation (AirClean® Systems).

Genomic DNA (gDNA) was extracted using the Qiagen DNeasy Blood and Tissue kit, using manufacturer’s instructions and the optional RNA elimination step, which used RNase A to ensure RNA-free gDNA. A modification to the protocol for the Pacific lamprey DNA isolation step was made, where instead of eluting the DNA into a total volume of 100 µL with Buffer AE, it was eluted with 50 µL to have more concentrated DNA.

DNA extractions for males and females for each species were done on separate days with UV treatment of equipment used between days to prevent cross-contamination. To ensure DNA was of good quality and not fragmented, extracted DNA was run on a 1.5% agarose gel for 40 minutes. Each sample was then diluted one part DNA to nine parts double-distilled molecular water (Qiagen), and used for subsequent PCR assays.
2.3.3. Identifying target genes and PCR

My goal was to test whether candidate master sex determining genes (sRY and sDY) and those implicated in the sex determination cascade (AMH, AMHR2, SOX2, SOX8, SOX9, SOX10, SOX17, WT1, FEM-1, SF1, DMRT1, FOXL2, GSDF, WNT4, TRA-1 and RSPO1) as identified from other species appear to have sex specific differences in sea and Pacific lampreys. These candidate genes were retrieved from various fishes and other model vertebrate species, and when available, an invertebrate outgroup was also included (Tables 2.2–2.18). Vertebrates and invertebrates that were used include: the common sea squirt, Ciona intestinalis; round worm, Caenorhabditis elegans; fruit fly, Drosophila melanogaster; chicken, Gallus gallus; medaka; mouse; Mus musculus; threespine stickleback, Gasterosteus aculeatus; zebrafish, Danio rerio; and fugu. Other fishes that have sex determination and differentiation genes that are annotated are: blind cave fish, Astyanax mexicanus; spotted gar, Lepisosteus oculatus; Atlantic cod, Gadus morhua; African coelacanth, Latimeria chalumnae; and Nile tilapia. Incorporating information from different organisms was used to compare the nucleotide and amino acid sequences (see Section 2.3.4.2 for details regarding subsequent analyses), and find conserved areas to design primers. Homologues of genes of interest were located in the sea lamprey genome by searching for the gene in the sea lamprey genome and seeing if Ensembl identifies the gene as a member of a gene family that includes other vertebrates, encoded for the same or similar genes (http://www.ensembl.org/Petromyzon_marinus/Info/Index Ensembl release 87: December 2016; Pmarinus_7.0).

The extracted DNA was used in PCR reactions using Taq polymerase (Promega) and GoTaq reagents (Invitrogen) and the reaction parameters were set at: 2 minutes of initial denaturation at 95°C, followed by 30 cycles of 30 second denaturation at 95°C, 30 second
annealing at primer specific temperatures and 30 seconds of elongation at 72°C (Table 2.19). This was followed by a final elongation at 72°C for 5 minutes. Primer specific annealing temperatures were set to 5°C less than the melting temperature of the primer with the lowest melting temperature. The Thermo Fisher Scientific website annealing temperature calculator was used for Taq polymerase to determine the ideal annealing temperatures for each pair of primers, diluted to 20 µM (https://www.thermofisher.com/us/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html).

The coding region for each gene from the sea lamprey genome was used to design primers and further used for sequencing. When primers designed for sea lamprey did not amplify the target regions in Pacific lamprey using GoTaq reagents (or where amplification was poor), various trouble-shooting methods were used. I tested different annealing temperatures, increased the amount of template and used Phire Hot Start II DNA Polymerase (Thermo Fisher Scientific), which is preferable for low quality DNA. Using Phire Hot Start II DNA Polymerase, the reaction parameters were set at: 5 minutes of initial denaturation at 98°C, followed by 30 cycles of 5 second denaturation at 98°C, 5 second annealing at 59°C and 20 seconds of elongation at 72°C. This was followed by a final elongation at 72°C for 1 minute. The annealing temperature was calculated using the Tm calculator on the Thermo Fisher Scientific website for Phire™ DNA polymerase. Details on the primer nucleotide sequences, annealing temperatures, amplicon size, and success of amplification are included in Tables 2.19–2.21. The roles of these candidate sex determination genes in other model vertebrates is summarized in Sections 1.3 and 1.4.

2.3.4. Sequencing preparation and analysis
2.3.4.1. Propanol precipitation, cycle sequencing, and ethanol precipitation

If the 10 µL reaction PCR amplified the target amplicon size when run alongside 1Kb Plus ladder (Invitrogen), the PCR reaction was repeated and the total reaction volume was increased to 30 µL. 5 µL from this reaction was then re-run on a gel to ensure that the amplicon was of the correct size with no non-specific amplification and the remaining PCR product was used for subsequent sequencing preparation.

The appropriate number of 0.6 mL tubes were labelled and to each, 13 µL of 3M sodium acetate (NaOAc), 38 µL of isopropanol (Fisher) and the PCR reaction product were added. The samples were gently inverted four times to ensure mixing of all solutions, and allowed to incubate on the bench top for 10 minutes. The samples were then centrifuged at 13,000 rotations per minute (rpm) using the Sorvall Legend RT Plus centrifuge (Thermo Scientific) at 4°C for 30 minutes. The samples were then carefully taken out of the centrifuge and the supernatant was pipetted out and discarded. 125 µL of cold 70% ethanol was added to each tube, ensuring no mixing occurred between the ethanol and precipitate. The tubes were then centrifuged for another 10 minutes at 13,000 rpm, and the supernatant was again carefully removed. Tubes at this point should have contained a clear pellet which was dried using a Savant DNA SpeedVac (Thermo Scientific) for 30 minutes. Vacuum dried precipitate samples were re-suspended by adding 6-8 µL of ddH2O, and allowed to re-suspend overnight at 4°C. Tubes were spun down using the Sorvall Legend Micro 21 Microcentrifuge (Thermo Scientific) and 1 µL of each propanol precipitate product and 4 µL of 10x BlueJuice Gel Loading Buffer (Invitrogen) were added to a gel loading plate. The samples were run on a gel to quantify the amount of DNA, alongside 1 µL of 100 base pair (bp) ladder (New England Biolabs).
The amount of the propanol precipitate product to be used for cycle sequencing was determined based on the amount of product that was present, determined by the brightness of the bands on the gel in comparison to the brightest 50 bp band. It was designated that if a band was of similar brightness, then 1 µL of the propanol precipitate would be used, and if a product band was fainter, then more product would be used. Cycle sequencing was done using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). The final reaction volume was set to 10.25 µL to account for any evaporation that took place during the thermal-cycling process.

BigDye cycle sequencing was performed using Veriti Thermal Cyclers (Applied Biosystems) and the BigDye kit, using manufacturer’s instructions for the PCR which consisted of an initial denaturation for 3 minutes at 94°C. This was followed by 25 cycles of: denaturation at 94°C for 30 seconds, annealing at 52°C for 20 seconds, and extension at 60°C for 30 seconds. The samples were then held at 4°C, or used immediately for the next step.

Cycle sequencing product was then spun down and kept covered with a paper towel to ensure it was not unnecessarily exposed to light to prevent any damage to light-sensitive dyes. The appropriate number of 0.6 mL Eppendorf tubes were labelled, and into each, 5 µL of 125 mM EDTA was added. Cycle sequencing product was added and I ensured that it made direct contact with the EDTA, and 63 µL of 95% ethanol was added to each tube. The reactions were then inverted four times to ensure proper mixing and allowed to incubate covered with a paper towel for 15 minutes. The samples were then spun at 13,000 rpm for 15 minutes at 4°C. The tubes are carefully taken out of the centrifuge, preventing tipping or knocking of tubes. At this point a round, cloudy white precipitate pellet may be visible at the bottom of the tube, and it is vital the precipitate doesn’t get dislodged. The supernatant was carefully pipetted off and discarded. 200 µL of 70% ethanol was added to rinse the pellet, and the tubes were spun at
13,000 rpm for 15 minutes at 4°C, after which the supernatant was carefully removed. The tubes containing the pellets were allowed to dry in the vacuum dryer for 20 minutes. Meanwhile Hi-Di formamide (Applied Biosystems) was taken out to thaw, covered with a paper towel. Once the samples had dried, 10 µL of the Hi-Di formamide was added to each tube and allowed to re-suspend covered on the bench top for 10 minutes. The samples were then briefly spun down and 10 µL was transferred to a sequencing plate. The plate was covered with thermal adhesive sealing film (Thermo Scientific) and spun down for 20 seconds using the Mini Plate Spinner 1000 (Labnet). The plate was then placed into the thermal cycler at 95°C for 5 minutes to denature. As soon as the samples were denatured, the tray was immediately placed on ice to cool, and spun down again using the plate spinner. The plate was then covered with the silicon plate cover, transferred to a sequence tray holder, and loaded into the sequencer. Samples were sequenced using the Applied Biosystems 3500 Genetic Analyzer and BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems).

2.3.4.2. Sequencing alignment and analysis

For each of the 19 candidate sex determination genes of interest, the amino acid sequence of the gene in sea lamprey, if available on Ensembl, was compared to that of other vertebrates (see Section 2.3.3). In addition to their value in designing primers for amplification of these genes in sea and Pacific lampreys, amino acid sequences were aligned to identify regions conserved among all vertebrates (and the invertebrate outgroup, when available) and used for phylogenetic analysis, to determine degree of similarity of the annotated gene in the sea lamprey genome to that of other vertebrates. For each gene, a neighbour-joining tree was constructed using the amino acid sequence for various vertebrates and an invertebrate outgroup, using the Poisson correction model and 1000 bootstrap replicates.
Of the targeted gene fragments, 11 genes were successfully sequenced in the sea lamprey and seven in the Pacific lamprey (see Section 2.5.1). For each of the candidate sex determination genes that were successfully sequenced in sea lamprey and Pacific lamprey, genomic nucleotide sequences, and the reference gene sequence that was used to design primers, were aligned using CodonCode Aligner V.6.0.2, and phylogenetic and molecular evolutionary analyses were conducted using Molecular Evolutionary Genetics Analysis (MEGA) version 6.06 (Tamura et al., 2013). Using the nucleotide sequence data, each gene was compared to the reference genome sequence to verify that the correct genomic region was sequenced and to determine if sex specific differences in nucleotide sequence were present within each species. The sequenced samples and the reference genome sequence were also used to construct neighbour-joining trees for each gene using the Kimura 2-parameter model and 1000 bootstrap replicates.

2.4. Results

Of the 19 candidate genes pursued in this study, six genes (sRY, sDY, GSDF, TRA-1, AMH, AMHR2) did not appear to have any homologous genes in the sea lamprey genome, and two (SOX10, SOX17) had homologues present, but the designed primer consistently failed to amplify the target of interest in both the sea and Pacific lamprey. All of the remaining 11 genes (SOX2, SOX8, SOX9, WT1, FEM-1, SF1, DMRT1, FOXL2, WNT3, WNT5, and RSPO1) were successfully amplified from sea lamprey genomic DNA and seven of these 11 also successfully amplified in Pacific lamprey (SOX8, SOX9, FEM-1, DMRT1, FOXL2, WNT3, WNT5; see Figures 2.1 to 2.33). Four of the 19 mentioned genes were novel genes and presumed identities were assigned (DMRT-related gene, FOXL2, NR5A1/SF1, WT1).

Novel genes, that is, those that have not been annotated in the sea lamprey genome on Ensembl, were identified by gene protein family trees and alignment and assigned tentative
identities based on similarities. Genes such as ENSPMAT0000003837 (DMRT-related gene; Figure 2.34) aligned with DMRTA2 or DMRT5, ENSPMAT00000011250 (FOXL2; Figure 2.36) aligned with FOXL2, ENSPMAT0000001607 (NR5A1 or SF1; Figure 2.38) aligned with NR5A1/SF1, and ENSPMAT0000008950 (WT1; Figure 2.43) sequence aligned with WT1 protein family.

Gene sequences that displayed slight differences between the gene fragments I sequenced in sea lamprey and the novel gene sequence available on Ensembl for the sea lamprey genome were: ENSPMAT0000001607 (NR5A1 or SF1; Figure 2.38; 0.5% different) and SOX9 (NCBI GenBank DQ136023.1; 1.4%). Difference between the sequenced sea lamprey samples and the annotated gene ENSPMAT0000003837 (DMRT family gene) was larger (Figure 2.3; 31.5%) and little sequence similarity was observed between my sequences and ENSPMAT0000008950 (WT1; Figure 2.43; 95.4% difference including uncalled region; 81% excluding) (see Section 2.5.1).

Eleven of the 19 candidate sex determining genes successfully amplified in the sea lamprey genome: DMRTA2, FEM1C, FOXL2, RSPO1, NR5A1/ SF1, SOX2, SOX8, SOX9, WNT3, WNT5 and WT1 (see Figure 2.34–2.43 for neighbour-joining trees for sequenced gDNA samples). Seven of the 19 genes were amplified and sequenced in the Pacific lamprey: DMRTA2, FEM1C, FOXL2, SOX8, SOX9, WNT3, and WNT5. There were no sex specific sequence differences in any of the sequenced genes in either the sea lamprey or Pacific lamprey. It is important, however, to point out that DMRTA2 amplified and was sequenced in both male and female sea lamprey but amplified successfully in 3/3 male Pacific lamprey but failed to amplify in 3/3 female Pacific lamprey. However, DNA quantity and quality of the Pacific lamprey samples was relatively poor and the number of males and females screened was small.
The available gDNA sequence, when translated into the complementary translated protein using the vertebrate mitochondrial model, showed that sea lamprey and Pacific lamprey sequences for the genes SOX (SRY-related HMG-box) transcription 8 (SOX8; Figure 2.39), Double-sex and mab-3 related transcription factor, (DMRT-related gene ENSPMAT0000003837; Figure 2.34), and Wingless-type MMTV integration site family member 3, (WNT3; Figure 2.41) were 1, 0.7, and 1.1% different, respectively. Genes that displayed differences in gDNA nucleotides between the two species, but had identical translated protein sequence, were the following genes: Feminization 1 homolog c (FEM1c; Figure 2.35), Forkhead box L2 (FOXL2; Figure 2.36) and Wingless-type MMTV integration site family member 5 (WNT5; Figure 2.42). SOX (SRY-related HMG-box) transcription 9 (SOX9) was the one gene that had no nucleotide base or protein sequence differences between the two species (Figure 2.40).

Sea lamprey samples that were sequenced and were identical to the reference annotated gene, with bootstrap support values given in parentheses, are as follows: FEM1c (96%), SOX8 (87%), WNT3 (97%), WNT5 (100%) and novel gene ENSPMAT0000011250 that forms a protein based family tree with other FOXL2 genes (63%). Differences that exist in the sequenced samples and reference gene sequences are primarily due to the portion of the genomic region selected to be sequenced, which did not cover the entire length of the gene. Gaps on either end of the sequenced fragment when compared to the reference were detected as being differences.

2.5. Discussion

2.5.1. Sequenced genes: lack of sex specific differences

Fragments of 11 candidate sex determination genes or genes involved in sex determination pathways were successfully sequenced in sea lamprey: DMRTA2, FEM1, FOXL2,
SF1, SOX2, SOX8, SOX9, RSPO1, WNT3, WNT5 and WT1 (Table 2.19). Of these genes, four were not successfully amplified in Pacific lamprey: RSPO1, NR5A1/ SF1, SOX2 and WT1, perhaps due to species specific sequence differences in the primer region due to the divergence of these two lamprey species approximately 13 million years ago (Docker et al., 1999). Genes that displayed slight sequence differences (0.5 and 1.4%) between the gene fragments I sequenced and the gene sequence available on Ensembl for the sea lamprey genome were found in NR5A1/ SF1 and SOX9 respectively, but differences were more pronounced in the DMRT-related family gene (31.5%) The suspected reason for the differences is due to the incorrect annotation of genes in the available sea lamprey genome. Novel genes that have not been annotated in the sea lamprey genome on Ensembl were identified in the current study by gene protein family trees and alignment and assigned tentative identities: DMRTA2, FOXL2, NR5A1/SF1 and WT1 (see Section 2.5.3). Although the exact role of these genes cannot be confirmed without further study, this preliminary designation of gene identities can help fill in the knowledge-gap of whether these genes are in fact present in lampreys and their roles, as well as the pattern of evolution of these genes in the vertebrate lineage. The substantially large difference (95.4%) between the sequenced and reference gene fragment of WT1 is due to 99 nucleotides of the 274 bp, or 36%, targeted gene fragment region containing uncalled base nucleotides in the reference on Ensembl; this will need to be re-examined. Excluding the 99 uncalled bases in the genomic sequence of WT1, my sequenced samples and the reference gene appears to be 81.1% different. Based on sequence base calling, regions that have uncalled regions can have discrepancies and background noise in downstream regions where multiple overlapping frequency peaks have been detected for a nucleotide position. The position of the reverse primer for WT1 is downstream of the 99-uncalled base region, and the downstream
nucleotides that have been called, may not be accurate if the portion immediately upstream has messy background sequence noise. If the downstream region is not correctly called, the sequence is not accurate and therefore primers that were designed may not be amplifying the target region. The large differences between the sequenced and available genome sequence cannot provide meaningful results for sex specific differences and require further work. Five gene fragments sequenced in sea lamprey in the current study were identical to the reference annotated gene: FEM1c, SOX8, WNT3, WNT5, and novel gene FOXL2. Sea and Pacific lamprey had identical nucleotide sequences for FEM1c, FOXL2, WNT5, and SOX9, and but showed differences ranging from 0.7 to 1.1% in SOX8, DMRT-related gene and WNT3 (see Section 2.4).

Significantly, given the objective of the current study, no gene fragments sequenced displayed any differences (i.e., nucleotide sequence differences, fragment length, or presence/absence) between males and female in either the sea lamprey or Pacific lamprey. However, DMRT-related gene, although it was successfully amplified and sequenced in both sexes in the sea lamprey, it only amplified successfully male Pacific lamprey. Amplification was weak in males, however, and DNA quality and quantity were generally poor in the Pacific lamprey samples; sex specific differences in amplification were likely a technical artifact. Moreover, sample sizes were small (three of each sex); although repeated attempts were made to amplify this gene fragment in females, it is worth pursuing for further study. This gene fragment should be explored in larger numbers of male and female Pacific lamprey, using high quality DNA.

2.5.2. Candidate sex determination and sex differentiation genes in lampreys

Eight genes that I proposed to look at were either not found in the sea lamprey genome or failed to amplify. This may be a result of the gene not being present in the genome, the gene
residing in the 20% region of DNA that is shed during programmed genome rearrangement during development (Smith et al., 2009; see below), the gene may be present but is not properly annotated or other genes take on similar roles for lampreys. Anti-Müllerian hormone (AMH), sex-determining region on the Y-chromosome (sRY), Transformer 1 (TRA-1), and gonadal somatic cell derived factor (GSDF) have been implicated in vertebrate sex developmental processes (see Sections 1.3 and 1.4), but they were not found in the sea lamprey genome. Two additional candidate genes (SOX10, SOX17) had homologues in the sea lamprey genome, but primers designed for them repeatedly failed to amplify in both lamprey species. Multiple degenerate primers were created for AMH using sequence from a number of other vertebrate genomes, and no regions within the sea lamprey beared similarity along with SOX10 and SOX17. Although the lamprey Testis-expressed 9, Tes9 and lamprey homologue of Zebrafish testis-expressed 38, Zie38, are not strictly sex determining genes, these genes appeared promising and the possibility of sex specific gene differences was pursued, but unsuccessful (Table 2.21). These genes may be present in the sea lamprey genome but not identified due to the assembly and annotation of the genome. However, although some of the genes involved in sex determination processes were not found in the sea lamprey, this does not conclusively show the absence of such genes, considering some of these genes were previously assumed absent in the coelacanth and recently disproved with the finding of GSDF. The Indonesian coelacanth, Latimeria menadoensis, was thought to be an extinct species, but upon the appearance of a male specimen, the testis and gDNA of the fish were saved and compared to the sequenced genome of the African coelacanth, Latimeria chalumnae. These two species of coelacanth are thought to have diverged 30–40 million years ago and upon analysis of genes expressed in the testis of the Indonesian coelacanth, researchers found GSDF, which was previously thought to be present
only within the teleost lineage, was present in both coelacanth species (Inoue et al., 2005; Forconi et al., 2013). Based on gene expression, Fibroblast growth factor 9 (FGF9) and SOX8 which also play roles in testis differentiation in most other vertebrates, were not expressed, indicating different roles for these genes in earlier vertebrate off-shoots, and perhaps lack of sex specific differences in lampreys as well (Forconi et al., 2013).

Sex differentiation genes are more conserved than sex determination genes, and based on my results from sequencing, many of the genes involved in sex differentiation have been found based on partial gene fragment sequencing, although novel lamprey genes may also be present. Master sex determination genes sRY, sdY or Dmy have been identified in various fishes, amphibians, and mammals, but seeing as even closely related fish species can display extreme variability in modes of sex determination (see Sections 1.3.1 and 1.3.2), a genetic basis to sex determination in lampreys cannot be eliminated based on the results obtained in the current study. Lampreys diverged from the rest of the vertebrate lineage more than 500 million years ago Smith et al. (2013), and recent sequencing of the sea lamprey genome has revealed a few unusual characteristics about the genome. During embryogenesis, millions of base pairs, or a substantial 20% of the DNA from somatic cells, is shed and the genome undergoes dramatic rearrangement (Smith et al., 2009; Smith et al., 2013). Some genes of interest that I was not able to find may reside in the genomic portion that is shed, impairing our ability to identify these genes. The large extent of genome deletion and rearrangement would not be tolerated in other vertebrates, with almost 60% of the DNA consisting of long DNA repeats, compared to the human genome having 4.5% long repeats. A high number of repeat regions becomes an obstacle when assigning automated annotations (Smith et al., 2009; Smith et al., 2010). Lampreys are peculiar in the sense that their genome is two-thirds the size of the human genome, but dispersed among a very large
number of small chromosomes (in sea lamprey, $2n=164$), where the structure of the chromosomes is “acrocentric” or “dot-like” (Smith et al., 2010; Caputo et al., 2011). Not only that, there is a high proportion of GC repeats in centromeres and telomeres, which in turn leads to a biased and limited number of amino acids translated (Smith et al., 2013). When comparing the somatic genome of gills and germline genome of testes, the somatic genome contains less DNA with $2n=164$ and $2n=198$, respectively (Smith et al., 2010). Due to the difference in chromosome count between tissues, tissue specific differences exist and as a result, germline specific sequences are not present in the somatic tissues (Smith et al., 2009; Smith et al., 2010).

2.5.3. Novel genes annotated

Accurate genome annotation is essential for understanding the role of a gene, such as novel, original genes that have not been previously identified. Annotation methods have been used for genomic work and are essential in analysis of RNA-sequencing data to assign identities to novel genes, based on clustering patterns with annotated genomes (Roberts et al., 2011). Annotation of novel genes have important implications in terms of application. For example, expression of proteins by healthy and diseased individuals, and identification of the expressed proteins, hold invaluable information in terms of targets for treatment (Renner and Aszodi, 2000). Using one method to define the role of a gene is over-simplistic and generalizations can lead to improper gene annotations. To prevent incorrect predictions, it is essential to combine data for DNA, RNA, transcribed and translated sequences to allow for the most accurate identification of a gene, and to have large coverage of taxonomic groups (Klasberg et al., 2016).

Proper nomenclature for genes that have been identified is essential. In some of the genes that have been annotated in Ensembl, a search for the assigned synonym for SF1 does not give results for Steroidogenic factor 1, but instead for Splicing factor 1. Splicing factor 1 is essential
for intestinal epithelial cell differentiation and required for the alternative splicing of various RNA-binding protein complexes (Shitashige et al., 2007). Similarly, Tra-1 which is commonly used as an abbreviation for sex determining Transformer 1, was also used as an abbreviation for transcription associated protein 1 and TRA1, Tumour rejection antigen (Gp96) is used as an alias for Heat shock protein 90, beta member 1 (HSP90b1). HSP90b1 is a chaperone protein required for the proper folding of proteins and is responsible for gametogenesis failure in mammals because HSP90b1 deficient sperm are unable to fertilize oocytes (Audouard and Christians, 2011). The common gene name TRA-1 assigned to numerous genes should include more details as the function of a chaperone protein is very different from a sex determining gene for C. elegans (Hunter and Wood, 1990). Due to this discrepancy, the gene that was sequenced based on the TRA1 gene available on Ensembl appears to be the HSP90b1 gene (Table 2.12; Figures 2.21-2.22)

There were some unannotated novel genes that I was able to assign IDs to, such as DMRTA2 (ENSPMAT0000003837), Steroidogenic factor 1 (SF1) transcription factor that is encoded by the NR5A1 gene; ENSPMAT0000001250), and FOXL2 (ENSPMAT0000008950) based on sequence similarities to other vertebrate sequences. Although these genes show similarity to other reference genomes, further work must be conducted to understand whether they play similar roles to what is observed in other vertebrates. Some of the presumed gene annotations have been used to design primers and used for gene expression analysis during testicular differentiation (see Chapter 3 for more details) and are briefly discussed below.

**DMRT-related gene** – Upon analysis of the sex determining region on the Y chromosome related HMG-box genes, a close link has been identified between SOX and Double-
Sex and Mab-3 Related Transcription factor family of genes (Devlin and Nagahama, 2002). Although no sRY gene has been identified for the sea lamprey, a DMRT-related gene that is similar to DMRT1 has been uncovered and used for this study. DMRT1 is the sex determination gene in many fish species, and birds, and plays roles in male sex differentiation in mammals as well (Matsuda et al., 2002; Myosho et al., 2012; Yano et al., 2012; Cutting et al., 2013; Kikuchi et al., 2013; Heule et al., 2014). The sequence of the available gene aligns with DMRTA2 and DMRT5 based on protein alignment according to Ensembl. I did manual nucleotide alignment for this gene with other DMRT sequences such as DMRTA2, DMRT2A and DMRT1, and observed this gene aligns with the respective genes of other vertebrates. Alignment of DMRTA2 genes indicate that the sea lamprey gene aligns more closely with other vertebrates than with outgroup DMRT1 sequence of C. intestinalis (Figure 2.4). DMRT2A and DMRT1 neighbour-joining trees show that the sea lamprey forms a clade with the DMRT1 C. intestinalis sequence, with bootstrap support values of 99% and 100% (Figure 2.6; Figure 2.2). Sequencing of this gene was done to confirm the authenticity of the available sequence and if the primers were targeting the region of interest, although there were differences between the sequenced and reference gene used. Using this sequencing data, I subsequently used the available gene to design primers for understanding the potential function in gene expression during sea lamprey testicular differentiation (Figure. 2.34). Mawaribuchi et al. (2017) looked at the expression of DMRT1 related gene in the Far Eastern brook lamprey, Lethenteron reissneri, and their observations of the role DMRT1 plays was consistent with what has been reported in other vertebrates (see Chapter 3). Upon multiple attempts, this gene was successfully amplified in Pacific lamprey males, but not females, indicating that perhaps there may have species-specific sequence differences at the primer site, or the failure to amplify may have been due to poor DNA quality. This may be an artifact, where
DNA quantity and quality may not have been sufficient to allow amplification, or contaminants may be present in the sample preventing the primers from annealing to the target gDNA sample. Future studies can design primers upstream of the current primers, in hopes of amplifying the genomic region of DMRT-related gene in Pacific lamprey females as well, and sequencing a larger portion in both sexes and species, and this would provide more conclusive evidence.

NR5A1/SF1 – Nuclear receptor subfamily 5, group A, member 1 (NR5A1) is the gene that encodes the transcription factor Steroidogenic factor 1 (SF1) which is required for activation of downstream genes. It is believed to be a male sex determining gene in some mammals, where the loss of SF1 interferes with spermatogenesis and testicular development processes in mammals (Schepers et al. 2003; Takada et al., 2005; Sekido and Lovell-Badge, 2008). The putative gene in sea lamprey forms a clade with other vertebrate NR5A1 sequences and many protein coding regions are conserved in the selected vertebrates (Figure 2.13; Figure 2.14). The presumed SF1 gene was used to make primers to test the expression of this gene during testicular differentiation (Chapter 3).

WT1 – WT1 is a gene that is involved in the sex determination and sex differentiation pathways, where it opposes the action of female sex determination genes, and plays a role in the female-to-male sex-reversal process in some animals (Kalfa et al., 2008; Garcia-Ortiz et al., 2009). The novel gene sequence that is available on Ensembl has some discrepancies in the nucleotide bases when compared to what I sequenced, where bases are not called or are not similar to what was sequenced downstream of the forward primer. This has led to the sequenced samples being quite divergent from the sequenced gene for phylogeny interpretation (Figure 2.43). This gene was also sequenced to confirm the sequence was similar to what is observed in
other vertebrates, showing a large number of conserved protein coding regions among vertebrates (Figure 2.32-2.33) and subsequently used for qPCR confirmation (see Chapter 3).

**FOXl2** – *FOXl2* is implicated in female sex determination in some mammals such as mice and humans, where loss of *FOXl2* leads to female-to-male sex-reversal (Ottolenghi et al., 2007; Barrionuevo et al., 2016). The gene in the sea lamprey has not been annotated, but based on protein alignment, Ensembl identifies this unannotated gene as a member of the *FOXl2* protein family with other vertebrates. Based on the protein alignment and neighbour-joining tree, the invertebrate *C. intestinalis* gene appears to be more closely related to the *FOXl2* gene of other vertebrates than the presumed *FOXl2* gene of the sea lamprey, contradictory to what Ensembl protein alignments allude to (Figure 2.9-2.10). This may be due to only a portion of the sequence being available online, which contains many uncalled bases, leading to the inaccurate phylogeny interpretation. Furthermore, female *FOXl2* is found in testis tumour cells in children, where it opposes the actions of male sex determination genes *SF1* and *WT1*, and leads to granulosa cell development in the testis (Kalfa et al., 2008). *Wnt4* and *RSPO1* are genes recruited independently of *FOXl2*, working synergistically in sex determination and differentiation processes in females (Garcia-Ortiz et al., 2009). To determine if this presumed gene is in fact *FOXl2*, further analysis, such as qPCR on female and male gonads can be conducted to see what potential role it plays in sexual development.

**SOX8 and SOX9** – These genes have roles in sex determination in other vertebrates, with *SOX9* being the sex determination gene in many species, while also being essential for male development in most vertebrates. In the presence of both *sRY* and *SF1*, *SOX9* is targeted and up-regulated during the period of male sex determination (Sekido and Lovell-Badge, 2008; Wallis et al., 2008). Although there were not differences found in *SOX9*, this is consistent with *SOX9* gene
sequencing of other vertebrates, such as the European Atlantic sturgeon, Acipenser sturio, where no sex specific gene sequences were found (Hett et al., 2005). Despite the lack of evidence supporting a sex determination role for SOX9 in sturgeon, it plays a conserved role in male sex differentiation the Siberian sturgeon, Acipenser baerii (Berbejillo et al., 2012).

\( WNT4 \) – \( WNT4 \) is highly conserved among teleost fishes and mammals with roles involved in female development and is highly expressed during early ovarian differentiation (Chen et al., 2015). In the \( WNT \) family of genes, \( WNT1 \) to \( WNT7 \) arose in invertebrates, before the ancestor that gave rise to arthropods, with a round of gene duplication occurring in \( WNT3 \), 5, 7, and 10, in the jawed vertebrate-offshoot (Sidow 1992). This would restrict the number of \( WNT \) genes that are present in lampreys to those shared with arthropods. A genetic sex determination mechanism was unknown for the East African cichlid, Astatotilapia burtoni, but upon sequencing the genome Böhne et al. (2016) identified \( WNT4 \) is likely the male sex determination gene. The annotated \( WNT \) gene sequences available for the sea lamprey are \( WNT3 \) and \( WNT5 \), yet no sex specific regions were found. Based on phylogenetic analyses, sea lamprey \( WNT3 \) (Figure 2.23) and \( WNT5 \) (Figure 2.29) amino acid residues show a high degree of amino acid conservation and form clades with their respective gene families, and not with \( WNT4 \) (Figure 2.31).

2.5.4. Conclusion

Through the advancement of genetic tools, such as gene knock-down and mutant lines, genes can be targeted to observe the role of the genes and to determine if they are involved in processes such as sex determination (Chandler et al., 2009). Knowing the sex of fish at an early developmental stage provides early onset for control of population sex ratios and this has been of interest for aquaculture, and has important application benefits associated with management of
sex-dependent behaviours (Quéméré et al., 2014). A PCR-based method has been identified to
determine the sex of Atlantic salmon, Salmo salar and brown trout, Salmo trutta. Given its
presence in males only, amplification of a 500 bp segment of the sDY gene determines the
genetic sex of the fish (Quéméré et al., 2014).

Using molecular techniques, populations of lampreys can then be controlled in a sex
specific manner with applications for both control and conservation (McCauley et al., 2015). I
proposed to look at genes that are master sex determining genes and genes that have been
implicated in the vertebrate sex determination cascade, and identify homologous genes in Pacific
and sea lamprey. Any genes that would be used to target the sex development in the sea lamprey
will need to be species and sex specific to be harmless to lampreys of conservation concern.
Identification of the genetic basis of sex determination will require culturing of many embryos to
test the methodology, but these discoveries can be promising, where we can apply these practices
to rehabilitate threatened Pacific lamprey (McCauley et al., 2015). This will not only shed light
onto the unknown genetic mechanism involved in sex determination in lampreys, but
identification of these homologues would help understand how these genes have evolved in
vertebrates over hundreds of millions of years (Docker et al., 2015).

This study identified novel (i.e., previously unannotated) genes in the Ensembl sea
lamprey genome and assigned presumed identities based on amino acid sequence similarity to
genes in other vertebrates; future work can be conducted to confirm the role of these genes by
examining patterns of gene expression. Although no sex specific differences in gene sequence
were found within each species, species specific differences found may be useful to reconstruct
phylogenies, as most phylogenetic work on lampreys is done using mitochondrial DNA (e.g.,
Docker et al., 1999; Boguski et al., 2012).
Candidate and master sex determination genes are not as conserved among vertebrates, whereas sex differentiation genes show a greater conservation, but the position of the differentiation genes in ovarian and testicular differentiation cascades can vary among vertebrates. Based on this knowledge and the results from this study, it was not surprising that sex specific differences in candidate sex determination gene fragments, as identified from other vertebrates, were not found in lampreys. However, based on this preliminary study, it cannot be assumed that GSD is absent in lampreys, as only a handful of gene fragments were sequenced. ESD has been postulated for lampreys (Docker and Beamish, 1994), but ESD and GSD have been shown to coexist in other fish species; the Atlantic silverside, for example, can either use GSD or TSD in individuals in the same population (Duffy et al., 2015). Therefore, a further exploration of lamprey genomes is essential to understand if lampreys too use multiple sex determination modes.
2.6. Literature cited


### 2.7. Tables and Figures

Table 2.1. Sampling location for adult Pacific lamprey, *Entosphenus tridentatus*, and sea lamprey, *Petromyzon marinus*, used for sequencing candidate sex determination genes. Sample ID names and abbreviations are categorized by sex for both species and sample collection location. Developmental stage of each sample is included as adults of Met. = metamorphosing lamprey, along with length of individuals if available. GPS map coordinates and latitude (lat.) and longitude (long.) are included for each site.

<table>
<thead>
<tr>
<th>Species</th>
<th>Sex</th>
<th>Sample id</th>
<th>Sample id</th>
<th>Dev. Stage</th>
<th>Length (if available, in mm)</th>
<th>Collection location</th>
<th>Map coordinates</th>
<th>Lat., long. (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. tridentatus</em></td>
<td>Female</td>
<td>EtrPROP_YNsy13-0001</td>
<td>001</td>
<td>Adult</td>
<td>-</td>
<td>Bonneville Dam, Columbia River mainstem, Oregon/Washington</td>
<td>45° 38’ 39” (45.644284, -121° 56’ 26” 121.940637)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EtrPROP_YNsy13-0002</td>
<td>002</td>
<td>Adult</td>
<td>-</td>
<td>Bonneville Dam, Columbia River mainstem, Oregon/Washington</td>
<td>45° 38’ 39” (45.644284, -121° 56’ 26” 121.940637)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EtrPROP_YNsy13-0004</td>
<td>004</td>
<td>Adult</td>
<td>-</td>
<td>Bonneville Dam, Columbia River mainstem, Oregon/Washington</td>
<td>45° 38’ 39” (45.644284, -121° 56’ 26” 121.940637)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>EtrPROP_YNsy13-0003</td>
<td>003</td>
<td>Adult</td>
<td>-</td>
<td>Bonneville Dam, Columbia River mainstem, Oregon/Washington</td>
<td>45° 38’ 39” (45.644284, -121° 56’ 26” 121.940637)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EtrPROP_YNsy13-0006</td>
<td>006</td>
<td>Adult</td>
<td>-</td>
<td>Bonneville Dam, Columbia River mainstem, Oregon/Washington</td>
<td>45° 38’ 39” (45.644284, -121° 56’ 26” 121.940637)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EtrPROP_YNsy13-0007</td>
<td>007</td>
<td>Adult</td>
<td>-</td>
<td>Bonneville Dam, Columbia River mainstem, Oregon/Washington</td>
<td>45° 38’ 39” (45.644284, -121° 56’ 26” 121.940637)</td>
<td></td>
</tr>
<tr>
<td>P. marinus</td>
<td>Female</td>
<td>M12</td>
<td>M13</td>
<td>M14</td>
<td>M12</td>
<td>Met.</td>
<td>132</td>
<td>Hammond Bay Biological Station, MI</td>
</tr>
<tr>
<td>P. marinus</td>
<td>Male</td>
<td>M23</td>
<td>M27</td>
<td>M28</td>
<td>M29</td>
<td>M30</td>
<td>Met.</td>
<td>128</td>
</tr>
<tr>
<td>P. marinus</td>
<td></td>
<td>M23</td>
<td>M27</td>
<td>M28</td>
<td>M29</td>
<td>M30</td>
<td>Met.</td>
<td>132</td>
</tr>
</tbody>
</table>
Table 2.2. List of *Doublesex and mab-3 related transcription factor 1 (DMRT1)* amino acid sequences used for constructing comparison alignments (Figure 2.1) and neighbour-joining tree (Figure 2.2). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and the outgroup, vase tunicate, at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl id</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT00000012867</td>
<td>Doublesex and mab-3 related transcription factor 1</td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT00000015734</td>
<td>Doublesex and mab-3 related transcription factor 1</td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT00000019186</td>
<td>Doublesex and mab-3 related transcription factor 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENSGACT00000019187</td>
<td></td>
</tr>
<tr>
<td>Coelacanth</td>
<td><em>Latimeria chalumnae</em></td>
<td>ENSLACT00000015034</td>
<td>Doublesex and mab-3 related transcription factor 1</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT00000014125</td>
<td>Doublesex and mab-3 related transcription factor 1</td>
</tr>
<tr>
<td>Mouse</td>
<td><em>Mus musculus</em></td>
<td>ENSMUST00000025755</td>
<td>Doublesex and mab-3 related transcription factor 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENSMUST00000087525</td>
<td></td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT00000025382</td>
<td>Doublesex and mab-3 related transcription factor 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENSORLT00000025383</td>
<td></td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSTRUT00000040266</td>
<td>Doublesex and mab-3 related transcription factor 1</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT0000004009</td>
<td>Novel gene-consensus annotation aligns with</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Doublesex and mab-3 related transcription factor 5 based on protein family</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT0000003837</td>
<td>Novel gene-consensus annotation aligns with</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Doublesex and mab-3 related transcription factor 2 based on protein family</td>
</tr>
<tr>
<td>Vase tunicate</td>
<td><em>Ciona intestinalis</em></td>
<td>ENSCINT00000012177</td>
<td>Dmrt1 transcription factor protein</td>
</tr>
</tbody>
</table>
Table 2.3. List of *Doublesex and mab-3 related transcription factor-like family A2 (DMRTA2)* amino acid sequences used for constructing comparison alignments (Figure 2.3) and neighbour-joining tree (Figure 2.4). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup vase tunicate, at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT00000018983</td>
<td>DMRT-like family A2</td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000057590</td>
<td>DMRT-like family A2</td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT0000001067</td>
<td>DMRT-like family A2</td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT00000014768</td>
<td>DMRT-like family A2</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT00000011457</td>
<td>DMRT-like family A2</td>
</tr>
<tr>
<td>Tilapia</td>
<td><em>Oreochromis niloticus</em></td>
<td>ENSONIT0000002958</td>
<td>DMRT-like family A2</td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSTRUT0000004232</td>
<td>DMRT-like family A2</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT0000004009</td>
<td>Novel gene-consensus annotation aligns with Doublesex and mab-3 related transcription factor 5, based on protein family</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT0000003837</td>
<td>Novel gene-consensus annotation aligns with Doublesex and mab-3 related transcription factor 2, based on protein family</td>
</tr>
<tr>
<td>Vase tunicate</td>
<td><em>Ciona intestinalis</em></td>
<td>ENSCINT0000012177</td>
<td>Dmrt1 transcription factor protein</td>
</tr>
</tbody>
</table>
Table 2.4. List of *Doublesex and mab-3 related transcription factor 2A* (*DMRT2A*) amino acid sequences used for constructing comparison alignments (Figure 2.5) and neighbour-joining tree (Figure 2.6). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup vase tunicate, at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000010786</td>
<td>Doublesex and mab-3 related transcription factor 2 A</td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT00000019177</td>
<td>Doublesex and mab-3 related transcription factor 2 A</td>
</tr>
<tr>
<td>Tilapia</td>
<td><em>Oreochromis niloticus</em></td>
<td>ENSONIT00000017860</td>
<td>Doublesex and mab-3 related transcription factor 2 A</td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT00000025779</td>
<td>Doublesex and mab-3 related transcription factor 2 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENSORLT00000025780</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENSORLT00000025781</td>
<td></td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSTRUT00000040184</td>
<td>Doublesex and mab-3 related transcription factor 2 A</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT00000004009</td>
<td>Novel gene-consensus annotation aligns with Doublesex and mab-3 related transcription factor 5, based on protein family</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT0000003837</td>
<td>Novel gene-consensus annotation aligns with Doublesex and mab-3 related transcription factor 2, based on protein family</td>
</tr>
<tr>
<td>Vase tunicate</td>
<td><em>Ciona intestinalis</em></td>
<td>ENSCINT00000012177</td>
<td>Dmrt1 transcription factor protein</td>
</tr>
</tbody>
</table>
Table 2.5. List of *Feminization 1 homolog c (FEM1c)* amino acid sequences used for constructing comparison alignments (Figure 2.7) and neighbour-joining tree (Figure 2.8). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup roundworm, at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT00000013228</td>
<td>Fem-1 homolog c</td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000162614</td>
<td>Fem-1 homolog c</td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT0000006348</td>
<td>Fem-1 homolog c</td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT0000009855</td>
<td>Fem-1 homolog c</td>
</tr>
<tr>
<td>Coelacanth</td>
<td><em>Latimeria chalumnae</em></td>
<td>ENSLACT0000014417</td>
<td>Fem-1 homolog c</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT0000010913</td>
<td>Fem-1 homolog c</td>
</tr>
<tr>
<td>Tilapia</td>
<td><em>Oreochromis niloticus</em></td>
<td>ENSONIT0000018951</td>
<td>Fem-1 homolog c</td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT0000006228</td>
<td>Fem-1 homolog c</td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSTRUT0000030538</td>
<td>Fem-1 homolog c</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT0000011287</td>
<td>Fem-1 homolog c</td>
</tr>
<tr>
<td>Roundworm</td>
<td><em>Caenorhabditis elegans</em></td>
<td>F35D6.1a</td>
<td>Sex-determining protein fem-1</td>
</tr>
</tbody>
</table>
Table 2.6. List of *Forkhead box L2 (FOXL2)* amino acid sequences used for constructing comparison alignments (Figure 2.9) and neighbour-joining tree (Figure 2.10). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup vase tunicate, at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT00000018626</td>
<td>Forkhead box L2</td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART0000061828</td>
<td>Forkhead box L2</td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT0000022086</td>
<td>Forkhead box L2</td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT0000007467</td>
<td>Forkhead box L2</td>
</tr>
<tr>
<td>Coelacanth</td>
<td><em>Latimeria chalumnae</em></td>
<td>ENSLACT0000013755</td>
<td>Novel gene-consensus annotation aligns with</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Forkhead box L2, based on protein family</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT0000021531</td>
<td>Forkhead box L2</td>
</tr>
<tr>
<td>Tilapia</td>
<td><em>Oreochromis niloticus</em></td>
<td>ENSONIT0000026114</td>
<td>Forkhead box L2</td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT0000025059</td>
<td>Forkhead box L2</td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSTRUT00000019818</td>
<td>Forkhead box L2</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT0000011250</td>
<td>Novel gene-consensus annotation aligns with</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Forkhead box C2, based on protein family</td>
</tr>
<tr>
<td>Vase tunicate</td>
<td><em>Ciona intestinalis</em></td>
<td>ENSCINT0000009675</td>
<td>FoxL protein</td>
</tr>
</tbody>
</table>
Table 2.7. List of R-spondin 1 (RSPO1) amino acid sequences used for constructing comparison alignments (Figure 2.11) and neighbour-joining tree (Figure 2.12). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey which is located at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td>Astyanax mexicanus</td>
<td>ENSAMXT00000020906</td>
<td>R-spondin 1</td>
</tr>
<tr>
<td>Zebrafish</td>
<td>Danio rerio</td>
<td>ENSDART00000160671</td>
<td>R-spondin 1</td>
</tr>
<tr>
<td>Stickleback</td>
<td>Gasterosteus aculeatus</td>
<td>ENSGACT00000018124</td>
<td>R-spondin 1</td>
</tr>
<tr>
<td>Coelacanth</td>
<td>Latimeria chalumnae</td>
<td>ENSLACT00000019383</td>
<td>R-spondin 1</td>
</tr>
<tr>
<td>Spotted gar</td>
<td>Lepisosteus oculatus</td>
<td>ENSLOCT0000007205</td>
<td>R-spondin 1</td>
</tr>
<tr>
<td>Tilapia</td>
<td>Oreochromis niloticus</td>
<td>ENSONIT00000020541</td>
<td>R-spondin 1</td>
</tr>
<tr>
<td>Medaka</td>
<td>Oryzias latipes</td>
<td>ENSORLT00000018370</td>
<td>R-spondin precursor</td>
</tr>
<tr>
<td>Fugu</td>
<td>Takifugu rubripes</td>
<td>ENSTRUT00000024435</td>
<td>R-spondin 1</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td>Petromyzon marinus</td>
<td>ENSPMAT0000004846</td>
<td>R-spondin 1</td>
</tr>
</tbody>
</table>
Table 2.8. List of *Nuclear receptor subfamily 5, group A, member 1* (*NR5A1*) or *SF1* amino acid sequences used for constructing comparison alignments (Figure 2.13) and neighbour-joining tree (Figure 2.14). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup roundworm, located at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td>Astyanax mexicanus</td>
<td>ENSAMXT0000020361</td>
<td>Nuclear receptor subfamily 5, group A, member 1b</td>
</tr>
<tr>
<td>Zebrafish</td>
<td>Danio rerio</td>
<td>ENSDART0000028265</td>
<td>Nuclear receptor subfamily 5, group A, member 1b</td>
</tr>
<tr>
<td>Zebrafish</td>
<td>Danio rerio</td>
<td>ENSDART00000162004</td>
<td>Nuclear receptor subfamily 5, group A, member 1a</td>
</tr>
<tr>
<td>Cod</td>
<td>Gadus morhua</td>
<td>ENSGMOT0000010280</td>
<td>Nuclear receptor subfamily 5, group A, member 1a</td>
</tr>
<tr>
<td>Chicken</td>
<td>Gallus gallus</td>
<td>ENSGALT000001616</td>
<td>Nuclear receptor subfamily 5 group A member 1</td>
</tr>
<tr>
<td>Stickleback</td>
<td>Gasterosteus aculeatus</td>
<td>ENSGACT0000004669</td>
<td>Nuclear receptor subfamily 5, group A, member 1a</td>
</tr>
<tr>
<td>Stickleback</td>
<td>Gasterosteus aculeatus</td>
<td>ENSGACT0000024271</td>
<td>Nuclear receptor subfamily 5, group A, member 1b</td>
</tr>
<tr>
<td>Spotted gar</td>
<td>Lepisosteus oculatus</td>
<td>ENSLOCT0000006222</td>
<td>Nuclear receptor subfamily 5, group A, member 1a</td>
</tr>
<tr>
<td>Medaka</td>
<td>Oryzias latipes</td>
<td>ENSORLT0000020643</td>
<td>Novel gene- consensus annotation aligns with <em>Nuclear receptor subfamily 5</em> group A member 1, based on protein family</td>
</tr>
<tr>
<td>Medaka</td>
<td>Oryzias latipes</td>
<td>ENSORLT0000016541</td>
<td>Nuclear receptor subfamily 5, group A, member 1b</td>
</tr>
<tr>
<td>Rat</td>
<td>Rattus norvegicus</td>
<td>ENSRNOT0000017651</td>
<td>Nuclear receptor subfamily 5, group A, member 1</td>
</tr>
<tr>
<td>Western clawed frog</td>
<td>Xenopus tropicalis</td>
<td>ENSXETT0000025024</td>
<td>Nuclear receptor subfamily 5, group A, member 1</td>
</tr>
<tr>
<td>Organism</td>
<td>Scientific Name</td>
<td>GenBank ID</td>
<td>Function</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------</td>
<td>-------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT00000001607</td>
<td>Novel gene-consensus annotation aligns with Nuclear receptor subfamily 5 group A member 1, based on protein family</td>
</tr>
<tr>
<td>Roundworm</td>
<td><em>Caenorhabditis elegans</em></td>
<td>C27C7.8</td>
<td>Nuclear Hormone Receptor family</td>
</tr>
</tbody>
</table>
Table 2.9. List of SRY (sex determining region Y)-box 2 (SOX2) and SRY (sex determining region Y)-box 3 (SOX3) amino acid sequences used for constructing comparison alignments (Figure 2.15) and neighbour-joining tree (Figure 2.16). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup roundworm, located at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td>Astyanax mexicanus</td>
<td>ENSAMXT00000026772</td>
<td>SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>Cave fish</td>
<td>Astyanax mexicanus</td>
<td>ENSAMXT00000026094</td>
<td>SRY (sex determining region Y)-box 3</td>
</tr>
<tr>
<td>Zebrafish</td>
<td>Danio rerio</td>
<td>ENSDART00000104493</td>
<td>SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>Zebrafish</td>
<td>Danio rerio</td>
<td>ENSDART00000075617</td>
<td>SRY (sex determining region Y)-box 3</td>
</tr>
<tr>
<td>Cod</td>
<td>Gadus morhua</td>
<td>ENSGMOT00000021782</td>
<td>SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>Cod</td>
<td>Gadus morhua</td>
<td>ENSGMOT00000022195</td>
<td>SRY (sex determining region Y)-box 3</td>
</tr>
<tr>
<td>Stickleback</td>
<td>Gasterosteus aculeatus</td>
<td>ENSGACT00000008091</td>
<td>SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>Stickleback</td>
<td>Gasterosteus aculeatus</td>
<td>ENSGACT00000022734</td>
<td>SRY (sex determining region Y)-box 3</td>
</tr>
<tr>
<td>Coelacanths</td>
<td>Latimeria chalumnae</td>
<td>ENSLACT00000007556</td>
<td>SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>Spotted gar</td>
<td>Lepisosteus oculatus</td>
<td>ENSLOCT00000021488</td>
<td>SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>Spotted gar</td>
<td>Lepisosteus oculatus</td>
<td>ENSLOCT00000021635</td>
<td>SRY (sex determining region Y)-box 3</td>
</tr>
<tr>
<td>Tilapia</td>
<td>Oreochromis niloticus</td>
<td>ENSONIT00000026643</td>
<td>SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>Tilapia</td>
<td>Oreochromis niloticus</td>
<td>ENSONIT00000026187</td>
<td>SRY (sex determining region Y)-box 3</td>
</tr>
<tr>
<td>Medaka</td>
<td>Oryzias latipes</td>
<td>ENSORLT00000014648</td>
<td>Oryzias latipes SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>Species</td>
<td>Scientific Name</td>
<td>Accession Number</td>
<td>SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------</td>
<td>------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT00000002213</td>
<td>SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSTRUT0000010478</td>
<td>SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT0000010751</td>
<td>SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>Roundworm</td>
<td><em>Caenorhabditis elegans</em></td>
<td>F40E10.2</td>
<td>SOX (mammalian SRY box) family- consensus annotation aligns with SRY (sex determining region Y)-box 2, based on protein family</td>
</tr>
</tbody>
</table>
Table 2.10. List of *SRY (sex determining region Y)-box 8 (SOX8)* amino acid sequences used for constructing comparison alignments (Figure 2.17) and neighbour-joining tree (Figure 2.18). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup roundworm, located at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT00000019537</td>
<td>SRY-box 8</td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000169936</td>
<td>SRY (sex determining region Y)-box 8a</td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000055055</td>
<td>SRY (sex determining region Y)-box 8b</td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT00000010122</td>
<td>SRY-box 8</td>
</tr>
<tr>
<td>Chicken</td>
<td><em>Gallus gallus</em></td>
<td>ENSGALT00000008449</td>
<td>SRY-box 8</td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT00000014850</td>
<td>SRY-box 8</td>
</tr>
<tr>
<td>Human</td>
<td><em>Homo sapiens</em></td>
<td>ENST00000293894</td>
<td>SRY-box 8</td>
</tr>
<tr>
<td>Coelacanth</td>
<td><em>Latimeria chalumnae</em></td>
<td>ENSLACT00000019016</td>
<td>SRY-box 8</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT0000002111</td>
<td>SRY-box 8</td>
</tr>
<tr>
<td>Tilapia</td>
<td><em>Oreochromis niloticus</em></td>
<td>ENSONIT0000006616</td>
<td>SRY-box 8</td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT00000010029</td>
<td>SRY-box containing gene 8</td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSTRUT00000023299</td>
<td><em>Takifugu rubripes</em> DNA-binding transcription factor Sox8</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT00000011398</td>
<td>SRY (sex determining region Y)-box 8b</td>
</tr>
<tr>
<td>Roundworm</td>
<td><em>Caenorhabditis elegans</em></td>
<td>F40E10.2</td>
<td>SOX (mammalian SRY box) family-consensus annotation aligns with SRY (sex determining region Y)-box 2, based on protein family</td>
</tr>
</tbody>
</table>
Table 2.11. List of \textit{SRY (sex determining region Y)-box 9 (SOX9)} amino acid sequences used for constructing comparison alignments (Figure 2.19) and neighbour-joining tree (Figure 2.20). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup roundworm, located at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td>\textit{Astyanax mexicanus}</td>
<td>ENSAMXT00000007169</td>
<td>SRY-box 9</td>
</tr>
<tr>
<td>Zebrafish</td>
<td>\textit{Danio rerio}</td>
<td>ENSDART00000005676</td>
<td>SRY (sex determining region Y)-box 9a</td>
</tr>
<tr>
<td>Zebrafish</td>
<td>\textit{Danio rerio}</td>
<td>ENSDART00000064500</td>
<td>SRY (sex determining region Y)-box 9b</td>
</tr>
<tr>
<td>Cod</td>
<td>\textit{Gadus morhua}</td>
<td>ENSGMOT00000010155</td>
<td>SRY (sex determining region Y)-box 9a</td>
</tr>
<tr>
<td>Chicken</td>
<td>\textit{Gallus gallus}</td>
<td>ENSGALT00000038513</td>
<td>Transcription factor SOX-9</td>
</tr>
<tr>
<td>Stickleback</td>
<td>\textit{Gasterosteus aculeatus}</td>
<td>ENSGACT00000014819</td>
<td>SRY-box 9</td>
</tr>
<tr>
<td>Human</td>
<td>\textit{Homo sapiens}</td>
<td>ENST00000245479</td>
<td>SRY-box 9</td>
</tr>
<tr>
<td>Coelacanth</td>
<td>\textit{Latimeria chalumnae}</td>
<td>ENSLACT00000021484</td>
<td>SRY-box 9</td>
</tr>
<tr>
<td>Spotted gar</td>
<td>\textit{Lepisosteus oculatus}</td>
<td>ENSLOCT00000015413</td>
<td>SRY (sex determining region Y)-box 9a</td>
</tr>
<tr>
<td>Mouse</td>
<td>\textit{Mus musculus}</td>
<td>ENSMUST0000000579</td>
<td>SRY (sex determining region Y)-box 9</td>
</tr>
<tr>
<td>Medaka</td>
<td>\textit{Oryzias latipes}</td>
<td>ENSORLT0000009986</td>
<td>Oryzias latipes Sry-box containing protein 9b</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td>\textit{Petromyzon marinus}</td>
<td>DQ136023.1</td>
<td>Sox9 partial mRNA</td>
</tr>
<tr>
<td>Roundworm</td>
<td>\textit{Caenorhabditis elegans}</td>
<td>F40E10.2</td>
<td>SOX (mammalian SRY box) family- consensus</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>annotation aligns with SRY (sex determining</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>region Y)-box 2, based on protein family</td>
</tr>
</tbody>
</table>
Table 2.12. List of *Heat shock protein 90, beta (grp94 (HSP90b1)* or *TRA1* amino acid sequences used for constructing comparison alignments (Figure 2.21) and neighbour-joining tree (Figure 2.22). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup roundworm, located at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT00000010375</td>
<td>Heat shock protein 90, beta (grp94), member 1</td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000004879</td>
<td>Heat shock protein 90, beta (grp94), member 1</td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT00000015182</td>
<td>Heat shock protein 90, beta (grp94), member 1</td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT00000013325</td>
<td>Heat shock protein 90, beta (grp94), member 1</td>
</tr>
<tr>
<td>Coelacanth</td>
<td><em>Latimeria chalumnae</em></td>
<td>ENSLACT00000005884</td>
<td>Heat shock protein 90 beta family member 1</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT00000018524</td>
<td>Heat shock protein 90, beta (grp94), member 1</td>
</tr>
<tr>
<td>Tilapia</td>
<td><em>Oreochromis niloticus</em></td>
<td>ENSONIT00000010731</td>
<td>Heat shock protein 90, beta (grp94), member 1</td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT00000005468</td>
<td>Heat shock protein 90, beta (grp94), member 1</td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSTRUT00000043644</td>
<td>Heat shock protein 90, beta (grp94), member 1</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT0000001245</td>
<td>Heat shock protein 90, beta (grp94), member 1</td>
</tr>
<tr>
<td>Roundworm</td>
<td><em>Caenorhabditis elegans</em></td>
<td>F38E11.1</td>
<td>Heat shock protein</td>
</tr>
</tbody>
</table>
Table 2.13. List of *Wingless-type MMTV integration site family, member 3* (*WNT3*) amino acid sequences used for constructing comparison alignments (Figure 2.23) and neighbour-joining tree (Figure 2.24). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup vase tunicate, located at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT00000020427</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000131175</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT00000021171</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT00000004974</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
<tr>
<td>Coelacanth</td>
<td><em>Latimeria chalumnae</em></td>
<td>ENSLACT00000008860</td>
<td>Wnt family member 3</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT00000016035</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
<tr>
<td>Tilapia</td>
<td><em>Oreochromis niloticus</em></td>
<td>ENSONIT00000022996</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT00000015059</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSRUTF00000008413</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT00000007828</td>
<td>Wingless-type MMTV integration site family, member 3A</td>
</tr>
<tr>
<td>Vase tunicate</td>
<td><em>Ciona intestinalis</em></td>
<td>ENSCINT00000024195</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
</tbody>
</table>
Table 2.14. List of **Wingless-type MMTV integration site family, member 4 (WNT4)** to sea lamprey **Wingless-type MMTV integration site family, member 3 (WNT3)** amino acid sequences used for constructing comparison alignments (Figure 2.25) and neighbour-joining tree (Figure 2.26). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup vase tunicate, located at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td>Astyanax mexicanus</td>
<td>ENSAMXT00000003672</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Cave fish</td>
<td>Astyanax mexicanus</td>
<td>ENSAMXT00000003674</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Zebrafish</td>
<td>Danio rerio</td>
<td>ENSDART000000105133</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Cod</td>
<td>Gadus morhua</td>
<td>ENSGMOT00000017892</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Stickleback</td>
<td>Gasterosteus aculeatus</td>
<td>ENSGACT00000014452</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Coelacanth</td>
<td>Latimeria chalumnae</td>
<td>ENSLACT00000017139</td>
<td>Wnt family member 4</td>
</tr>
<tr>
<td>Spotted gar</td>
<td>Lepisosteus oculatus</td>
<td>ENSLOCT0000001279</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Spotted gar</td>
<td>Lepisosteus oculatus</td>
<td>ENSLOCT0000001293</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Tilapia</td>
<td>Oreochromis niloticus</td>
<td>ENSONIT00000025518</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENSONIT00000025519</td>
<td></td>
</tr>
<tr>
<td>Organism</td>
<td>Scientific Name</td>
<td>Ensembl ID</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------</td>
<td>---------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT00000015061</td>
<td>Oryzias latipes wingless-type MMTV integration site family member 4a (wnt4a)</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT00000007828</td>
<td>Wingless-type MMTV integration site family, member 3A</td>
</tr>
<tr>
<td>Vase tunicate</td>
<td><em>Ciona intestinalis</em></td>
<td>ENSCINT00000024195</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
</tbody>
</table>
Table 2.15. List of *Wingless-type MMTV integration site family, member 4 (WNT4)* to sea lamprey *Wingless-type MMTV integration site family, member 5 (WNT5)* amino acid sequences used for constructing comparison alignments (Figure 2.27) and neighbour-joining tree (Figure 2.28). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup vase tunicate, located at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT0000003672</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT0000003674</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000105133</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT0000017892</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT0000014452</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Coelacanth</td>
<td><em>Latimeria chalumnae</em></td>
<td>ENSLACT0000017139</td>
<td>Wnt family member 4</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT000001279</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT000001293</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td>Tilapia</td>
<td><em>Oreochromis niloticus</em></td>
<td>ENSONIT0000025518</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENSONIT0000025519</td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Genus</td>
<td>Scientific Name</td>
<td>Gene Accession Number</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------</td>
<td>--------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Medaka</td>
<td>Oryzias latipes</td>
<td>ENSORLT00000015061</td>
<td>Oryzias latipes wingless-type MMTV integration site family member 4a (wnt4a)</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td>Petromyzon marinus</td>
<td>ENSPMAT00000010914</td>
<td>Wingless-type MMTV integration site family, member 5b</td>
</tr>
<tr>
<td>Vase tunicate</td>
<td>Ciona intestinalis</td>
<td>ENSCINT00000024195</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
</tbody>
</table>
Table 2.16. List of Wingless-type MMTV integration site family, member 5 (WNT5) amino acid sequences used for constructing comparison alignments (Figure 2.29) and neighbour-joining tree (Figure 2.30). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup vase tunicate, located at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT0000015358</td>
<td>Wingless-type MMTV integration site family, member 5a</td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000164739</td>
<td>Wingless-type MMTV integration site family, member 5a</td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000172249</td>
<td>Wingless-type MMTV integration site family, member 5a</td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT0000012299</td>
<td>Wingless-type MMTV integration site family, member 5a</td>
</tr>
<tr>
<td>Coelacanth</td>
<td><em>Latimeria chalumnae</em></td>
<td>ENSLACT00000006648</td>
<td>Wnt family member 5A</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT0000017511</td>
<td>Wingless-type MMTV integration site family, member 5a</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT0000017512</td>
<td>Wingless-type MMTV integration site family, member 5a</td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT0000011064</td>
<td>Wingless-type MMTV integration site family, member 5a</td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSTRUT0000021691</td>
<td>Wingless-type MMTV integration site family, member 5a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENSTRUT0000021692</td>
<td></td>
</tr>
<tr>
<td>Organism</td>
<td>Scientific Name</td>
<td>Gene Accession</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------</td>
<td>----------------</td>
<td>-------------------------------------------------------------------</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT00000010914</td>
<td>Wingless-type MMTV integration site family, member 5b</td>
</tr>
<tr>
<td>Vase tunicate</td>
<td><em>Ciona intestinalis</em></td>
<td>ENSCINT00000024195</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
</tbody>
</table>
Table 2.17. List of *Wingless-type MMTV integration site family, member 3, 4 and 5* (WNT3, WNT4, and WNT5, respectively) amino acid sequences used for constructing comparison alignments and neighbour-joining tree (Figure 2.31). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey, which is located at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT00000020427</td>
<td><em>Wingless-type MMTV integration site family, member 3</em></td>
</tr>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT0000003672 ENSAMXT0000003674</td>
<td><em>Wingless-type MMTV integration site family, member 4a</em></td>
</tr>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT0000015358</td>
<td><em>Wingless-type MMTV integration site family, member 5a</em></td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000131175</td>
<td><em>Wingless-type MMTV integration site family, member 3</em></td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000105133</td>
<td><em>Wingless-type MMTV integration site family, member 4a</em></td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART00000164739 ENSDART00000172249</td>
<td><em>Wingless-type MMTV integration site family, member 5a</em></td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT0000021171</td>
<td><em>Wingless-type MMTV integration site family, member 3</em></td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT0000017892</td>
<td><em>Wingless-type MMTV integration site family, member 4a</em></td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT0000012299</td>
<td><em>Wingless-type MMTV integration site family, member 5a</em></td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT00000004974</td>
<td><em>Wingless-type MMTV integration site family, member 3</em></td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT0000014452</td>
<td><em>Wingless-type MMTV integration site family, member 4a</em></td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT0000014455</td>
<td><em>Wingless-type MMTV integration site family, member 3</em></td>
</tr>
<tr>
<td>Coelacanth</td>
<td><em>Latimeria chalumnae</em></td>
<td>ENSLACT0000008860</td>
<td><em>Wnt family member 3</em></td>
</tr>
<tr>
<td>Species</td>
<td>Scientific Name</td>
<td>Accession Numbers</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------------------</td>
<td>-------------------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Coelacanth</td>
<td><em>Latimeria chalumnae</em></td>
<td>ENSLACT00000017139</td>
<td>Wnt family member 4</td>
</tr>
<tr>
<td>Coelacanth</td>
<td><em>Latimeria chalumnae</em></td>
<td>ENSLACT00000006648</td>
<td>Wnt family member 5A</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT00000016035</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT0000001279</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENSLOCT0000001293</td>
<td></td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT00000017511</td>
<td>Wingless-type MMTV integration site family, member 5a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENSLOCT00000017512</td>
<td></td>
</tr>
<tr>
<td>Tilapia</td>
<td><em>Oreochromis niloticus</em></td>
<td>ENSONIT00000022996</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
<tr>
<td>Tilapia</td>
<td><em>Oreochromis niloticus</em></td>
<td>ENSONIT00000025518</td>
<td>Wingless-type MMTV integration site family, member 4a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENSONIT00000025519</td>
<td></td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT00000015059</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT00000015061</td>
<td><em>Oryzias latipes</em> wingless-type MMTV integration site family member 4a (wnt4a)</td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT00000011064</td>
<td>Wingless-type MMTV integration site family, member 5a</td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSTRUT0000008413</td>
<td>Wingless-type MMTV integration site family, member 3</td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSTRUT00000021691</td>
<td>Wingless-type MMTV integration site family, member 5a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENSTRUT00000021692</td>
<td></td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT0000007828</td>
<td>Wingless-type MMTV integration site family, member 3A</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT00000010914</td>
<td>Wingless-type MMTV integration site family, member 5b</td>
</tr>
</tbody>
</table>
Table 2.18. List of *Wilm’s tumour 1 (WT1)* amino acid sequences used for constructing comparison alignments (Figure 2.32) and neighbour-joining tree (Figure 2.33). Ensemble ID indicates peptide coding sequence for the respective species’ genomes. Species are organized in alphabetical order by scientific name, except for sea lamprey and outgroup solitary sea squirt, which are located at the bottom.

<table>
<thead>
<tr>
<th>Common name</th>
<th>Species name</th>
<th>Ensembl transcript ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cave fish</td>
<td><em>Astyanax mexicanus</em></td>
<td>ENSAMXT00000013387</td>
<td>Wilm’s tumour 1a</td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART000000146543</td>
<td>Wilm’s tumour 1b</td>
</tr>
<tr>
<td>Zebrafish</td>
<td><em>Danio rerio</em></td>
<td>ENSDART000000163794</td>
<td>Wilm’s tumour 1b</td>
</tr>
<tr>
<td>Cod</td>
<td><em>Gadus morhua</em></td>
<td>ENSGMOT00000013510</td>
<td>Wilm’s tumour 1b</td>
</tr>
<tr>
<td>Stickleback</td>
<td><em>Gasterosteus aculeatus</em></td>
<td>ENSGACT00000004329</td>
<td>Wilm’s tumour 1a</td>
</tr>
<tr>
<td>Coelacanth</td>
<td><em>Latimeria chalumnae</em></td>
<td>ENSLACT00000018732</td>
<td>Wilm’s tumour 1</td>
</tr>
<tr>
<td>Spotted gar</td>
<td><em>Lepisosteus oculatus</em></td>
<td>ENSLOCT00000005435, ENSLOCT00000005447</td>
<td>Wilm’s tumour 1a</td>
</tr>
<tr>
<td>Medaka</td>
<td><em>Oryzias latipes</em></td>
<td>ENSORLT00000012371</td>
<td>Oryzias latipes Wilm’s, tumour suppressor 1b (wt1b)</td>
</tr>
<tr>
<td>Fugu</td>
<td><em>Takifugu rubripes</em></td>
<td>ENSTRUT00000016171, ENSTRUT00000016172</td>
<td>Wilm’s tumour 1b</td>
</tr>
<tr>
<td>Sea lamprey</td>
<td><em>Petromyzon marinus</em></td>
<td>ENSPMAT00000009895</td>
<td>Novel gene-consensus annotation aligns with Wilm’s tumour homolog, based on protein family</td>
</tr>
<tr>
<td>Solitary sea</td>
<td><em>Ciona savignyi</em></td>
<td>SINCSAVT0000000152</td>
<td>WT1 protein</td>
</tr>
<tr>
<td>Squirt</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.19. Primers and PCR annealing temperatures ($T_a$) for the 11 candidate sex determination genes that successfully amplified from sea lamprey genomic DNA; PL indicates that this gene also amplified in Pacific lamprey. In some cases, multiple primer pairs were designed to amplify larger fragment of the gene. See Table 2.20 for details regarding primer positions within each gene, amplicon size, and amplification success in Pacific lamprey.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene symbol</th>
<th>Forward sequence</th>
<th>Reverse sequence</th>
<th>Primer $T_a$ °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRY (sex determining region y)-box 2</td>
<td>SOX2</td>
<td>GGCAACAGAGTCAGCAGAT</td>
<td>CAGAGAGTACTTGTCTCTTCTTCAG</td>
<td>55.1</td>
</tr>
<tr>
<td>SRY (sex determining region y)-box 8</td>
<td>SOX8 PL</td>
<td>1F- ACTCGGTATGACGACGA</td>
<td>1R- CAGAGTCTTTGCTCAGGCT</td>
<td>54.5</td>
</tr>
<tr>
<td>SRY (sex determining region y)-box 9</td>
<td>SOX9 PL</td>
<td>1F- GATTTCAAGAGTGACGACATAC</td>
<td>2R- TCGTCGGCCATACCGGAGCT</td>
<td>55.2</td>
</tr>
<tr>
<td>Double-sex and mab-3 related transcription factor 2</td>
<td>DMRT2 PL</td>
<td>GACGAAGCCACGATTTCA</td>
<td>TCGCTTGAGAGAGCTACA</td>
<td>54.7</td>
</tr>
<tr>
<td>R-spondin 1</td>
<td>RSPO1</td>
<td>CACAGAAACGTCTGAGGA</td>
<td>TTGCTATTCTGGGAGGACAG</td>
<td>53.6</td>
</tr>
<tr>
<td>Wilms tumour protein 1</td>
<td>WT1</td>
<td>CGAGTGTCCTGTTGTTGT</td>
<td>TGATCGGAGCGGAGAGAAT</td>
<td>54.1</td>
</tr>
<tr>
<td>Wingless-type MMTV integration site family member 3</td>
<td>WNT3 PL</td>
<td>CGTCTCTATCTCTGGACTTCA</td>
<td>AAGACGCAGTTGCACTTC</td>
<td>54.7</td>
</tr>
<tr>
<td>Wingless-type MMTV integration site family member 5</td>
<td>WNT5 PL</td>
<td>GGCAAGTTTCTCAAGAGGAGAG</td>
<td>TTCACTTGCACACGAC.COM</td>
<td>56.2</td>
</tr>
<tr>
<td>Forkhead box l2</td>
<td>FOXL2 PL</td>
<td>GCATCTACGAGTACCATCTC</td>
<td>TCGACGTATCGTGCCCCT</td>
<td>53.3</td>
</tr>
<tr>
<td>Steroidogenic factor 1</td>
<td>SF1</td>
<td>1F- GGTGAGCAGATGATGAGAC</td>
<td>1R- CCCGTGACCAGGAGGAT</td>
<td>55.5</td>
</tr>
<tr>
<td>ENSPMAG00000001450 Feminization 1</td>
<td>FEM1 PL</td>
<td>1F- AACCGCAGAGTGTCAGAG</td>
<td>1F- CGGTACAGTACCTCTTTTC</td>
<td>54.3</td>
</tr>
</tbody>
</table>
Table 2.20. Details regarding candidate sex determining genes that successfully amplified from sea lamprey genomic DNA and, where applicable, that also amplified in Pacific lamprey. Total number of nucleotides in transcript (i.e., excluding introns), forward (F) and reverse (R) primer positions in gene, expected amplicon size, and size of fragment sequenced are indicated along with if the gene amplified and sequenced successfully in one or both lamprey species. In cases where more than one set of primers was made, the primer that was used in sequencing is **bolded**. * = indicates that the gene is novel (i.e., not annotated in Ensembl) and a presumptive identity has been given.

<table>
<thead>
<tr>
<th>Gene of interest</th>
<th>Gene in sea lamprey genome / Ensembl transcript ID</th>
<th>Transcript length (bp)</th>
<th>Corresponding primer positions in gene</th>
<th>Expected amplicon size (bp)</th>
<th>Successfully sequenced (bp)</th>
<th>Species amplified in</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOX2</td>
<td>SOX2 ENSPMAT0000010751</td>
<td>1,236</td>
<td>F 167–185 (intron 1-2) R 436–459 (exon 2)</td>
<td>293</td>
<td>250</td>
<td>Sea lamprey only</td>
</tr>
<tr>
<td>SOX8</td>
<td>SOX8 ENSPMAT0000011398</td>
<td>489</td>
<td>Sox8-1 F 238–255 (exon 1) R 457–475 (exon 1) Sox8-2 F 19–36 R 237–255</td>
<td>237</td>
<td>195</td>
<td>Both</td>
</tr>
<tr>
<td>DMRT1</td>
<td>DMRTA2* ENSPMAT00000003837</td>
<td>657</td>
<td>DmrtA2-1 (exon 1) F 31–49 (exon 3) R 365–383</td>
<td>352</td>
<td>270</td>
<td>All sea lamprey and only Pacific lamprey males¹</td>
</tr>
<tr>
<td>Gene</td>
<td>Gene*</td>
<td>RefSeq</td>
<td>F, R Coordinates</td>
<td>Transcripts</td>
<td>Start Codon</td>
<td>Stop Codon</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>--------</td>
<td>---------------------------------------</td>
<td>--------------------------------------------------</td>
<td>-------------</td>
<td>------------</td>
</tr>
<tr>
<td>RSPO1</td>
<td>RSPO1</td>
<td>204</td>
<td>Rspl  F 70–88 (GL477522: 81879-81896)</td>
<td>238  R 290–308 (intron 1-2 and exon 2)</td>
<td>238</td>
<td>188</td>
</tr>
<tr>
<td>NR5A1</td>
<td>NR5A1/SF1*</td>
<td>564</td>
<td>F 2–22 (exon 1)</td>
<td>503  R 484–506 (exon 1)</td>
<td>503</td>
<td>458</td>
</tr>
<tr>
<td>WT1</td>
<td>WT1*</td>
<td>660</td>
<td>Trans. 1: WT1t.1-2 F 235–253 R 735–755</td>
<td>383 Trans. 1: WT1t.1-2 F 268–286 (exon 3)</td>
<td>383</td>
<td>270</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Trans. 2: WT1t.2-1 F 269–289 R 630–648</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WT1t.2-2 F 109–126 R 505–524</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WNT3</td>
<td>WNT3</td>
<td>1,068</td>
<td>Wnt3-1 F 585–605 (exon 3) R 984–1002</td>
<td>417 Wnt3-1 F 585–605 (exon 3) R 984–1002</td>
<td>417</td>
<td>374</td>
</tr>
<tr>
<td>WNT5</td>
<td>WNT5</td>
<td>1,122</td>
<td>Wnt5b-1 F 177–195 R 633–651</td>
<td>474 Wnt5b-1 F 177–195 R 633–651 Wnt5b-2 F 403–425 (exon 5)</td>
<td>474</td>
<td>312</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wnt5b-2 F 403–425 (exon 5) R 742–762 (exon 5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wnt5b-3 F 411–433 R 754–774</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOXL2</td>
<td>FOXL2*</td>
<td>360</td>
<td>FoxL2-1</td>
<td>200 FoxL2-1</td>
<td>200</td>
<td>154</td>
</tr>
</tbody>
</table>
DMRTA2 was sequenced in both male and female sea lamprey, but only amplified in Pacific lamprey males; 3 females and males of each species were sequenced, and amplification of Pacific lamprey females was attempted multiple times. It is noteworthy to point out that bands were weak in Pacific lamprey males for this gene, and DNA quantity and quality were an issue in this species.
Table 2.21. Primers and other details for the candidate sex determination genes for which homologues were found or genes appeared sufficiently homologous (bolded) but did not amplify in sea lamprey or Pacific lamprey genomic DNA. Problems associated with primers for these genes indicated as not successful in amplifying correct gene with non-target amplification of no amplification. * denotes degenerate primers were designed for gene of interest which did not appear to have any homologous genes.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene symbol</th>
<th>Forward sequence</th>
<th>Reverse sequence</th>
<th>Intended fragment size deduced from:</th>
<th>Fragment size</th>
<th>Primer problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sox (SRY-related HMG-box) transcription factor 10</td>
<td>SOX10</td>
<td>1- GCAGCATAAGAGGCAC 2- GTCGGATACCAATCAGG</td>
<td>1- ATGCCGATGGTCTGGAA 2- GGTGGTCTTTATGTG</td>
<td>P. marinus genome</td>
<td>1- 366 2- 527</td>
<td>Non-specific amplification</td>
</tr>
<tr>
<td>Sox (SRY-related HMG-box) transcription factor 17</td>
<td>SOX17</td>
<td>GAGAATACGCAGACCCAT GAA</td>
<td>TGGCAGAGGTGGAGACA</td>
<td>Lethenteron camtschaticum transcriptome</td>
<td>370</td>
<td>No amplification</td>
</tr>
<tr>
<td>Anti-Müllerian hormone</td>
<td>AMH*</td>
<td>1- YST TYC TK C TGM WRG CYC TDC AGA 2- AAC CAC GCC RTS STG MTC WWY KM 3- YST TYC TK C TGM WRG CYC TDC AGA</td>
<td>1- KMR WWG AKC ASS AYG GCG TGG TT 2- TCG TAK GCC ACS GGC ACR CAG CA 3- TCG TAK GCC ACS GGC ACR CAG CA</td>
<td>Multiple vertebrate species*</td>
<td>1- 218 2- 60 3- 278</td>
<td>No amplification</td>
</tr>
<tr>
<td>Tes9</td>
<td>GAGAGCAACAAGAGCTTC TC</td>
<td>CATCATCTGCAGGATTCTT</td>
<td>P. marinus genome</td>
<td>249</td>
<td>Non-target sequenced</td>
<td></td>
</tr>
<tr>
<td>Zebrasfish testis-expressed 38</td>
<td>Zte38</td>
<td>GCTCATCCGCAAGTTTGG</td>
<td>TCATCGAGTGAAGGTCG AG</td>
<td>Danio rerio genome</td>
<td>212</td>
<td>Non-target sequenced</td>
</tr>
</tbody>
</table>
Table 2.22. Genes that were successfully amplified and involved in the sex determination cascade as known from other vertebrates and their implicated roles (see Section 1.3 and 1.4). This list includes the 11 genes that amplified from sea lamprey genomic DNA; $^\text{PL}$ indicates that this gene also amplified in Pacific lamprey (see Section 2.4). * denotes that gene was not previously annotated in Ensembl, but was identified based on similarity to respective gene in other vertebrates and invertebrates.

<table>
<thead>
<tr>
<th>GENE NAME</th>
<th>Gene symbol</th>
<th>Role</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sox (SRY-related HMG-box) transcription factor 2</td>
<td>SOX2</td>
<td>Implicated in sex determination pathway in lake sturgeon, Japanese flounder</td>
<td>Foster and Graves, 1994; Hale et al., 2010; Gao et al., 2014</td>
</tr>
<tr>
<td>SRY (sex determining region y)-box 8</td>
<td>SOX8 $^\text{PL}$</td>
<td>Role in sex determination pathway in mice, chicken, humans, expressed in both sexes in both sexes in turtles</td>
<td>Takada et al., 2004; Takada et al., 2005; Barrionuevo et al., 2009; Jiang et al., 2013</td>
</tr>
<tr>
<td>SRY (sex determining region y)-box 9</td>
<td>SOX9 $^\text{PL}$</td>
<td>Expression required during sex determination in mice, precedes expression of other genes; also roles in sex determination in fish, mammals, and birds</td>
<td>Morais da Silva, 1996; Oréal et al., 1998; Vernetti et al., 2013</td>
</tr>
<tr>
<td>Double-sex and mab-3 related transcription factor 2</td>
<td>DMRT2 $^\text{PL}$</td>
<td>DMRT family of genes with roles in sex determination/differentiation in most vertebrates; master sex determination role in some birds, fish and frogs</td>
<td>Herpin and Schartl, 2011; Forconi et al., 2013</td>
</tr>
<tr>
<td>R-spondin 1</td>
<td>RSPO1</td>
<td>Involved in sex determination pathway</td>
<td>Wallis et al., 2008</td>
</tr>
<tr>
<td>Wilms tumour protein 1</td>
<td>WT1</td>
<td>Controls transcription of sRY and is necessary for gonad development</td>
<td>Matsuzawa-Watanabe et al., 2003</td>
</tr>
<tr>
<td>Wingless-type MMTV integration site family member 3/5</td>
<td>WNT3/5 $^\text{PL}$</td>
<td>Implicated in female sex determination, WNT family of genes before genome duplication</td>
<td>Sidow et al., 1992; Trukhina et al., 2013</td>
</tr>
<tr>
<td>Forkhead box 12</td>
<td>FOXL2 $^\text{PL}$</td>
<td>Role in female sex determination pathway in medaka and mice</td>
<td>Cutting et al., 2013; Trukhina et al., 2013</td>
</tr>
<tr>
<td>Steroidogenic factor 1</td>
<td>$SF1^*$</td>
<td>Sex determination role in humans, chicken and platypus</td>
<td>Wallis et al., 2008</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------</td>
<td>--------------------------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>ENSPMAG00000001450</td>
<td>$FEM1^{PL}$</td>
<td>Role in female pathway in nematodes</td>
<td>Ventura-Holman et al., 1998</td>
</tr>
</tbody>
</table>
Figure 2.1. Alignment of *Drosophila and mab-3 related transcription factor 1 (DMRT1)* amino acid sequence from various vertebrates and invertebrate *Ciona* (Table 2.2) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.2 and sequences are used to construct a neighbour-joining tree (Figure 2.2).
Figure 2.2. Neighbour-joining tree comparing *Doublesex and mab-3 related transcription factor 1 (DMRT1)* amino acid sequences from various vertebrates and invertebrate *Ciona* (Table 2.2). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name.
| 1. Astyanax_mexicanus_DMRTA2 | LLQVPLPTAECAGSHEVYALGKHVGKCKCMCAKLLIAASVYMAR |
| 2. Ciona_intestinalis_DMRT1 | FLTASDPYPHCAFCHSVVATLGHKTYCKKCLCPCTLLIARCTYMAA |
| 3. Danio_rerio_DMRTA2 | LLAQEPYPHHCAFCHSVVATLGHKTYCKKCLCPCTLLIARCTYMAA |
| 4. Gadus_morhua_DMRTA2 | LLASYPYPHHCAFCHSVVATLGHKTYCKKCLCPCTLLIARCTYMAA |
| 5. Gasterosteus_aculeatus_DMRTA2 | LLASDPYPHCAFCHSVVATLGHKTYCKKCLCPCTLLIARCTYMAA |
| 6. Lepisosteus_oculatus_DMRTA2 | LLASDPYPHCAFCHSVVATLGHKTYCKKCLCPCTLLIARCTYMAA |
| 7. Oreochromis_niloticus_DMRTA2 | LLTADYPYPHCAFCHSVVATLGHKTYCKKCLCPCTLLIARCTYMAA |
| 8. Petromyzon_marinus_ENSMAT000000004009 | LLTADYPYPHCAFCHSVVATLGHKTYCKKCLCPCTLLIARCTYMAA |
| 9. Takifugu_rubripes_DMRTA2 | LLGQYPYPHCAFCHSVVATLGHKTYCKKCLCPCTLLIARCTYMAA |

Figure 2.3. Alignment of *Doublesex and mab-3 related transcription factor-like family A2 (DMRTA2)* amino acid sequence from various vertebrates and invertebrate *Ciona* (Table 2.3) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.3 and sequences are used to construct neighbour-joining tree (Figure 2.4).
Figure 2.4. Neighbour-joining tree comparing *Doublesex and mab-3 related transcription factor-like family A2* (DMRTA2) amino acid sequences from various vertebrates and invertebrate *Ciona* (Table 2.3). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name.
Figure 2.5. Alignment of *Doublesex and mab-3 related transcription factor 2A* (*DMRT2A*) amino acid sequence from various vertebrates and invertebrate *Ciona* (Table 2.4) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name, except for sea lamprey and outgroup vase tunicate, which are given at the bottom. Regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.4 and sequences are used to construct neighbour-joining tree (Figure 2.6).
Figure 2.6. Neighbour-joining tree comparing *Doublesex and mab-3 related transcription factor 2A* (*DMRT2A*) amino acid sequences from various vertebrates and invertebrate *Ciona* (Table 2.4). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name.
Figure 2.7. Alignment of *Feminization 1 homolog c (FEM1c)* amino acid sequence from various vertebrates and invertebrate roundworm (Table 2.5) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.5 and sequences are used to construct neighbour-joining tree (Figure 2.8).
Figure 2.8. Neighbour-joining tree comparing *Feminization 1 homolog c* (*FEM1c*) amino acid sequences from various vertebrates and invertebrate roundworm (Table 2.5). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name. *FEM1c* was annotated in the sea lamprey genome and the gene name is included above.
Figure 2.9. Alignment of *Forkhead box L2 (FOXL2)* amino acid sequence from various vertebrates and invertebrate *Ciona* (Table 2.6) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.6 and sequences are used to construct neighbour-joining tree (Figure 2.10).
Figure 2.10. Neighbour-joining tree comparing *Forkhead box L2* (*FOXL2*) amino acid sequences from various vertebrates and invertebrate *Ciona* (Table 2.6). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The *Petromyzon marinus* novel gene corresponds to the unannotated gene ENSPMAT00000011250 that share sequence homology with *FOXL2* sequences from other vertebrates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name.
Figure 2.11. Alignment of *R-spondin 1 (RSPO1)* amino acid sequence from various vertebrates (Table 2.7) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.7 and sequences are used to construct neighbour-joining tree (Figure 2.12).
Figure 2.12. Neighbour-joining tree comparing R-spondin 1 (RSPO1) amino acid sequences from various vertebrates (Table 2.7). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name. RSPO1 was annotated in the sea lamprey genome and the gene name is included above.
1. Astyanax mexicanus NR5A1B
2. Danio rerio NR5A1A
3. Danio rerio NR5A1B
4. Gadus morhua NR5A1A
5. Gallus gallus NR5A1A
6. gasterosteus aculeatus NR5A1A
7. gasterosteus aculeatus NR5A1B
8. Lepisosteus oculatus NR5A1
9. Oryzias latipes NR5A1
10. Oryzias latipes NR5A1B
11. Petromyzon marinus novel
12. Rattus norvegicus NR5A1
13. Xenopus tropicalis NR5A1

1. Astyanax mexicanus NR5A1B
2. Danio rerio NR5A1A
3. Danio rerio NR5A1B
4. Gadus morhua NR5A1A
5. Gallus gallus NR5A1A
6. Gasterosteus aculeatus NR5A1A
7. Gasterosteus aculeatus NR5A1B
8. Lepisosteus oculatus NR5A1
9. Oryzias latipes NR5A1
10. Oryzias latipes NR5A1B
11. Petromyzon marinus novel
12. Rattus norvegicus NR5A1
13. Xenopus tropicalis NR5A1

Figure 2.13. Alignment of Nuclear receptor subfamily 5, group A, member 1 (NR5A1) amino acid sequence from various vertebrates (Table 2.8) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.8 and sequences are used to construct neighbour-joining tree (Figure 2.14).
Figure 2.14. Neighbour-joining tree comparing *Nuclear receptor subfamily 5, group A, member 1 (NR5A1)* or amino acid sequences from various vertebrates and invertebrate roundworm (Table 2.8). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name.
Figure 2.15. Alignment of SRY (sex determining region Y)-box 2 (SOX2) and SRY (sex determining region Y)-box 3 (SOX3) amino acid sequence from various vertebrates and invertebrate roundworm (Table 2.9) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.9 and sequences are used to construct neighbour-joining tree (Figure 2.16).
Figure 2.16. Neighbour-joining tree comparing SRY (sex determining region Y)-box 2 (SOX2) and SRY (sex determining region Y)-box 3 (SOX3) amino acid sequences from various vertebrates and invertebrate roundworm (Table 2.9). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name. SOX2 was annotated in the sea lamprey genome and the gene name is included above.
Figure 2.17. Alignment of SRY (sex determining region Y)-box 8 (SOX8) amino acid sequence from various vertebrates and invertebrate roundworm (Table 2.10) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.10 and sequences are used to construct neighbour-joining tree (Figure 2.18).
Figure 2.18. Neighbour-joining tree comparing *SRY* (*sex determining region Y*)-box 8 (*SOX8*) amino acid sequences from various vertebrates and invertebrate roundworm (Table 2.10). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name. *SOX8* was annotated in the sea lamprey genome and the gene name is included above.
Figure 2.19. Alignment of SRY (sex determining region Y)-box 9 (SOX9) amino acid sequence from various vertebrates and invertebrate roundworm (Table 2.11) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.11 and sequences are used to construct neighbour-joining tree (Figure 2.20).
Figure 2.20. Neighbour-joining tree comparing *SRY* (*sex determining region Y*)-box 9 (*SOX9*) amino acid sequences from various vertebrates and invertebrate roundworm (Table 2.11). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name. *SOX9* was annotated in the sea lamprey genome and the gene name is included above.
Figure 2.21. Alignment of *Heat shock protein 90, beta (grp94 (HSP90b1)) or TRA1* amino acid sequence from various vertebrates and invertebrate roundworm (Table 2.12) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.12 and sequences are used to construct neighbour-joining tree (Figure 2.22).
Figure 2.22. Neighbour-joining tree comparing *Heat shock protein 90, beta (grp94 (HSP90b1))* or *TRA1* amino acid sequences from various vertebrates and invertebrate roundworm (Table 2.12). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site.
Figure 2.23. Alignment of **Wingless-type MMTV integration site family, member 3 (WNT3)** amino acid sequence from various vertebrates and invertebrate *Ciona* (Table 2.13) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.13 and sequences are used to construct neighbour-joining tree (Figure 2.24).
Figure 2.24. Neighbour-joining tree comparing *Wingless-type MMTV integration site family, member 3* (*WNT3*) amino acid sequences from various vertebrates and invertebrate *Ciona* (Table 2.13). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name. *WNT3* was annotated in the sea lamprey genome and the gene name is included above.
Figure 2.25. Alignment of *Wingless*-type MMTV integration site family, member 4 (*WNT4*) and sea lamprey *WNT3* amino acid sequence from various vertebrates and invertebrate *Ciona* (Table 2.14) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.14 and sequences are used to construct neighbour-joining tree (Figure 2.26).
Figure 2.26. Neighbour-joining tree comparing *Wingless*-type MMTV integration site family, member 4 (WNT4) with sea lamprey WNT3 amino acid sequences from various vertebrates and invertebrate *Ciona* (Table 2.14). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name. *WNT3* was annotated in the sea lamprey genome and the gene name is included above.
Figure 2.27. Alignment of Wingless-type MMTV integration site family, member 4 (WNT4) and sea lamprey WNT5 amino acid sequence from various vertebrates and invertebrate Ciona (Table 2.15) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.15 and sequences are used to construct neighbour-joining tree (Figure 2.28)
Figure 2.28. Neighbour-joining tree comparing Wingless-type MMTV integration site family, member 4 (WNT4) with sea lamprey WNT5 amino acid sequences from various vertebrates and invertebrate Ciona (Table 2.15). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name. WNT5 was annotated in the sea lamprey genome and the gene name is included above.
Figure 2.29. Alignment of Wingless-type MMTV integration site family, member 5 (WNT5) amino acid sequence from various vertebrates and invertebrate Ciona (Table 2.16) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.16 and sequences are used to construct neighbour-joining tree (Figure 2.30).
Figure 2.30. Neighbour-joining tree comparing Wingless-type MMTV integration site family, member 5 (WNT5) amino acid sequences from various vertebrates and invertebrate Ciona (Table 2.16). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name. WNT5 was annotated in the sea lamprey genome and the gene name is included above.
Figure 2.31. Neighbour-joining tree comparing Wingless-type MMTV integration site family, member 3, 4 and 5 (WNT3, WNT4 and WNT5) amino acid sequences from various vertebrates and invertebrate Ciona (Table 2.17). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name. WNT3 and WNT5 were annotated in the sea lamprey genome and the gene name is included above.
Figure 2.32. Alignment of *Wilm’s tumour 1 (WT1)* amino acid sequence from various vertebrates and invertebrate *Ciona* (Table 2.18) aligned in MEGA 6.06. Sequences are arranged in alphabetical order by scientific name and regions with 100% conserved amino acid sites are highlighted for each column. Ensembl ID’s for species are given in Table 2.18 and sequences are used to construct neighbour-joining tree (Figure 2.33).
Figure 2.33. Neighbour-joining tree comparing *Wilms's tumour 1 (WT1)* amino acid sequences from various vertebrates and vertebrate *Ciona* (Table 2.18). This tree was constructed in MEGA 6.06 using the Poisson correction method model, with 1000 bootstrap replicates. The scale bar represents the number of amino acid substitutions per site. Sea lamprey genes were retrieved from Ensembl, and if a gene was not annotated, the given Ensembl ID was used instead of the presumptive gene name.
Figure 2.34. Neighbour-joining tree comparing *Doublesex and mab-3 related transcription factor-like family A2 (DMRTA2)* nucleotide sequences (270 bp fragment) from *P. marinus* and *E. tridentatus* males and females (Table 2.21). This tree was constructed in MEGA 6.06 using the Kimura 2-parameter method, with 1000 bootstrap replicates. The tree is drawn to scale, with branch lengths next to each branch in evolutionary distance units that were used to infer the phylogenetic tree. Primers for this gene were designed using novel gene ENSPMAT00000003837.
Figure 2.35. Neighbour-joining tree comparing *Feminization 1 homolog c* (*FEM1c*) nucleotide sequences (269 bp fragment) from *P. marinus* and *E. tridentatus* males and females (Table 2.21). This tree was constructed in MEGA 6.06 using the Kimura 2-parameter method, with 1000 bootstrap replicates. The tree is drawn to scale, with branch lengths next to each branch in evolutionary distance units that were used to infer the phylogenetic tree. Primers for this gene were designed using the annotated *FEM1c* gene ENSPMAT00000011287.
Figure 2.36. Neighbour-joining tree comparing *Forkhead box L2* (FOXL2) nucleotide sequences (154 bp fragment) from *P. marinus* and *E. tridentatus* males and females (Table 2.21). This tree was constructed in MEGA 6.06 using the Kimura 2-parameter method, with 1000 bootstrap replicates. The tree is drawn to scale, with branch lengths next to each branch in evolutionary distance units that were used to infer the phylogenetic tree. Primer for this gene was designed using the novel FOXL2 gene ENSPMAT0000011250 as template.
Figure 2.37. Neighbour-joining tree comparing R-spondin 1 (RSPO1) nucleotide sequences (188 bp fragment) from *P. marinus* males and females (Table 2.21). This tree was constructed in MEGA 6.06 using the Kimura 2-parameter method, with 1000 bootstrap replicates. The tree is drawn to scale, with branch lengths next to each branch in evolutionary distance units that were used to infer the phylogenetic tree. Primer for this gene was designed using the annotated RSPO1 gene ENSPMAT00000004846 as template.
Figure 2.38. Neighbour-joining tree comparing *Nuclear receptor subfamily 5, group A, member 1 (NR5A1)* or *SF1* nucleotide sequences (488 bp fragment) from *P. marinus* males and females (Table 2.21). This tree was constructed in MEGA 6.06 using the Kimura 2-parameter method, with 1000 bootstrap replicates. The tree is drawn to scale, with branch lengths next to each branch in evolutionary distance units that were used to infer the phylogenetic tree. Primer for this gene was designed using the novel gene ENSPMAT00000001607 as template.
Figure 2.39. Neighbour-joining tree comparing *SRY* (*sex determining region Y*)-box 2 (*SOX2*) and nucleotide sequences (250 bp fragment) from *P. marinus* males and females (Table 2.21). This tree was constructed in MEGA 6.06 using the Kimura 2-parameter method, with 1000 bootstrap replicates. The tree is drawn to scale, with branch lengths next to each branch in evolutionary distance units that were used to infer the phylogenetic tree. Primer for this gene was designed using the annotated gene ENSPMAT00000010751 as template.
Figure 2.40. Neighbour-joining tree comparing *SRY* (*sex determining region Y*)-box 8 (*SOX8*) and nucleotide sequences (195 bp fragment) from *P. marinus* and *E. tridentatus* males and females (Table 2.21). This tree was constructed in MEGA 6.06 using the Kimura 2-parameter method, with 1000 bootstrap replicates. The tree is drawn to scale, with branch lengths next to each branch in evolutionary distance units that were used to infer the phylogenetic tree. Primer for this gene was designed using the annotated gene ENSPMAT0000011398 as template.
Figure 2.41. Neighbour-joining tree comparing SRY (sex determining region Y)-box 9 (SOX9) and nucleotide sequences (516 bp fragment) from Petromyzon marinus and Entosphenus tridentatus males and females (Table 2.21). This tree was constructed in MEGA 6.06 using the Kimura 2-parameter method, with 1000 bootstrap replicates. The tree is drawn to scale, with branch lengths next to each branch in evolutionary distance units that were used to infer the phylogenetic tree. Primer for this gene was designed using the annotated gene DQ136023.1 as template.
Figure 2.42. Neighbour-joining tree comparing *Wingless-type MMTV integration site family, member 3* (WNT3) and nucleotide sequences (374 bp fragment) from *P. marinus* and *E. tridentatus* males and females (Table 2.21). This tree was constructed in MEGA 6.06 using the Kimura 2-parameter method, with 1000 bootstrap replicates. The tree is drawn to scale, with branch lengths next to each branch in evolutionary distance units that were used to infer the phylogenetic tree. Primer for this gene was designed using the annotated gene ENSPMAT00000007828 as template.
Figure 2.43. Neighbour-joining tree comparing Wingless-type MMTV integration site family, member 5 (WNT5) and nucleotide sequences (312 bp fragment) from P. marinus and E. tridentatus males and females (Table 2.21). This tree was constructed in MEGA 6.06 using the Kimura 2-parameter method, with 1000 bootstrap replicates. The tree is drawn to scale, with branch lengths next to each branch in evolutionary distance units that were used to infer the phylogenetic tree. Primer for this gene was designed using the annotated gene ENSPMAT00000010914 as template.
Figure 2.44. Neighbour-joining tree comparing *Wilm’s tumour 1 (WT1)* and nucleotide sequences (270 bp fragment) from *P. marinus* males and females (Table 2.21). This tree was constructed in MEGA 6.06 using the Kimura 2-parameter method, with 1000 bootstrap replicates. The tree is drawn to scale, with branch lengths next to each branch in evolutionary distance units that were used to infer the phylogenetic tree. Primer for this gene was designed using the annotated gene ENSPMAT0000008950 as template.
CHAPTER 3: Gene expression in sea lamprey, *Petromyzon marinus*, during testis differentiation and development

3.1. Abstract

Sex differentiation is the diversification of undifferentiated tissue into tissues and phenotypes associated with the male or female fate. In lampreys, ovarian differentiation occurs during the larval stage but testicular differentiation appears to be delayed until metamorphosis. Some of the genes involved in ovarian differentiation in lampreys have recently been identified, but nothing is known regarding the genetic factors involved in testicular differentiation. I therefore tested whether candidate sex differentiation genes, specifically key Sertoli cell transcription factors involved in testicular differentiation and spermatogenesis in other vertebrates (e.g., *SOX9*, *DMRT1*, *WT1*, and *DAZAP1*), show differential gene expression prior to and during testicular differentiation in sea lamprey, *Petromyzon marinus*. Using qRT-PCR, I compared gene expression in ovaries and testes from larval, metamorphosing, and adult sea lamprey (i.e., between males and females and among stages of gonadal development in males), and performed histological analysis on all testes. An increase in germ cells in the testes coincided with increased expression of *SOX9*, *DMRT1* and *WT1*, and higher *DAZAP1* expression was observed during spermatogenesis and spermiogenesis. The increase in *SOX9* expression was followed by an increase in *DMRTA2*, with significantly higher expression in male larvae relative to female larvae. Thus, genes involved in testicular differentiation appear to be conserved across vertebrates, although the involvement of lamprey-specific genes cannot yet be ruled out.
3.2. **Introduction**

Lamprey sex differentiation is both size and age dependent, with low population densities appearing to favour a female biased population sex ratio (Docker and Beamish, 1994). Lampreys are unusual relative to most other fishes in the sense that they spend approximately 3-8 years in the larval phase, and undergo very little development in the gonad during the first year or two of life. Female gonadal differentiation in sea lamprey, *Petromyzon marinus*, for example, starts between year 2 and 3, at a body length of approximately 71-80 mm (Hardisty, 1969), with ovarian differentiation completed at 3-4 years or 90-100 mm (Hardisty, 1971; Docker, 1992). Male differentiation is even further delayed, not occurring until the onset of metamorphosis at the age of approximately 5–7 years (Hardisty, 1971). In other fishes, such as the Nile tilapia, *Oreochromis niloticus*, the critical period of differentiation is 5-10 days-after hatch (dah), and the first signs of ovarian differentiation are visible 23-26 dah (Nakamura et al., 1998). Like lampreys, spermatogenesis generally occurs after ovarian differentiation in other fishes, but at 50-70 dah in tilapia and at approximately 40 dah in the Japanese rice-fish, *Oryzias latipes*. Male differentiation is similarly delayed until after ovarian differentiation is complete in the zebrafish, *Danio rerio*, 20-60 days post fertilization (Siegfried, 2010).

When the first signs of sex differentiation are evident in the sea lamprey, there are two apparent cyst types: small and large cysts that form the male and female gonad, respectively. All larval lamprey gonads (i.e., both future males and future females) contain oocytes (Hardisty, 1969). A decline in the number of undifferentiated germ cells, and an increase in oocytes, occurs in females at lengths of approximately 91-100 mm, and is indicative of ovarian differentiation, but the regression of cysts can be masked by the rapid growth of the oocytes. At this length, approximately 45% of the developing gonad in the presumptive female is composed of oocytes,
with the remainder of the gonad consisting of undifferentiated germ cells, and ovarian
differentiation is not completed until a minimum of 100 mm body length (Hardisty, 1971).

Beyond this point, larvae with undifferentiated gonads are presumed to be future males
(Docker, 1992). The number of individuals with oocytes, and oocyte count starts to decrease in
males at lengths of 90-110 mm, indicating that gradual regression of oocytes has begun in the
gonad of the future male. Large numbers of oocytes and cysts containing undifferentiated germ
cells in the males start to regress years prior to and during differentiation (Hardisty, 1965;
Hardisty, 1969). Some oocytes persist and continue growth, and the undifferentiated germ cells
that have survived (in clusters or alone) slowly proliferate and will give rise to the future testis
(Hardisty, 1965; Hardisty, 1969). At lengths of 115-120 mm, the gonad of a future male sea
lamprey resembles the undifferentiated gonad of a 60-mm larva in size and appearance, but the
testis has undergone many changes. The criterion that can be used to distinguish the
undifferentiated gonad from a presumptive testis is the presence of basophilic oocytes and loose
reticulum in the developing testis. Based on the observations made by Docker (1992) oocytes
were present in males of all sizes, but did not increase in diameter between the different larval
stages. The number of oocytes was no more than 6 oocytes per cross section in males, and
prevalence decreased as body length increased (Docker, 1992). During differentiation, groups of
germ cells start to become pinched apart by follicle cells, that alter the outline of the testis,
making it lobe-shaped. Testicular differentiation commences at the onset of metamorphosis,
where a considerable increase in mitotic activity is observed, and the testis starts to gain more the
finger-line extensions that house germ cells (Hardisty, 1971).

The process of ovarian differentiation has been reasonably well studied in lampreys—
both histologically (e.g., Hardisty, 1965; Hardisty, 1971; Docker, 1992) and in terms of gene
expression (Spice et al., 2014)—given the relative ease with which larval lampreys can be collected. Comparatively less is known about testicular differentiation in lampreys, particularly with respect to the genetic factors involved, because of the relative scarcity of metamorphosing lampreys, which are available only for a limited time each year. The current study, however, examined whether genes involved in testicular differentiation in other vertebrates are expressed in the sea lamprey in a sex specific or stage specific manner, and used histological analysis to determine what was happening at the cellular level within the sea lamprey testis at each stage.

3.2.1. Stages of testicular differentiation and development in lampreys

Unless indicated otherwise, the stages of testicular differentiation in lampreys described below are summarized from Hardisty (1965; see also Section 1.4.3).

Testicular differentiation – late larval stage and metamorphosis

During sexual differentiation, cells begin to slowly lose their pluripotency and become committed to an irreversible fate and in lampreys, but not restricted to lampreys, germ cell clusters undergo extensive degeneration but individual germ cells can survive and still retain the ability to differentiate at a later point in the individual’s life (Hardisty, 1965; Arnold, 2017; see Figure 1.1). Testicular differentiation in lamprey commences at the onset of metamorphosis (Hardisty, 1971).

All lampreys undergo metamorphosis, where profound developmental changes occur within the larval lamprey, transforming an immature larva through various intermediate stages into parasitic or non-parasitic juveniles (Youson, 1980; Manzon et al., 2015). During metamorphosis, lampreys are unable to feed because their mouth structures and intestine are undergoing changes to prepare them for their parasitic trophic phase (Reis-Santos et al., 2008; Manzon et al., 2015). There are various internal and endogenous factors that determine
metamorphosis such as size, lipid accumulation and condition factor, and the process is highly synchronized within a population, typically beginning in the summer and lasting for 3-4 months (Manzon et al., 2015). The body length at which metamorphosis commences in sea lamprey ranges from 127 to 180 mm, slightly after 5 years of age, with differences in size of metamorphosis evident among landlocked and anadromous populations (Hardisty, 1965; Manzon et al., 2015).

At the same time that lampreys are experiencing drastic external remodelling, there is an increase in pro-spermatogonial mitotic activity in male gonads and development in oocytes that persist (see Section 1.5; Figure 1.2). Later stages of testicular differentiation are defined by the progressive loss of lobation, and an increase in the size of nuclei of somatic cells that are found within germ cells. At this time, the testis becomes permeated with clusters of spermatogonia that start to appear, generally in most vertebrates under the influence of anterior pituitary glycoproteins: growth hormone (GH), follicle-stimulating hormone (FSH), and luteinizing hormone (LH) (Shapiro, 1987; Schulz et al., 2010). Both LH and FSH are vital in regulating testicular physiology in mammals, where they act with their respective receptors to regulate the production of Leydig cells and Sertoli cells (Schulz et al., 2010). Lampreys, however, do not have LH or FSH, but they do have gonadotropin-β, a precursor that underwent gene duplication to produce LH and FSH in the jawed vertebrates (Sower et al., 2006; Roch et al., 2011, Sower, 2015).

A progressive increase in the abundance of spermatogonia is seen throughout metamorphosis. Metamorphosis have been divided into seven stages by Youson and Potter (1980), and well-studied over the years (Manzon et al., 2015). The stage of the individual can be identified from external appearance (e.g., eye and oral disc development), and my focus will be
on testis differentiation during this period of dramatic changes. Ideally, given the synchronized nature of metamorphosis, I hope to be able to assign the observed changes happening in the gonad to their respective stage of metamorphosis (Youson 1980). As metamorphosis is completed, parasitic individuals are ready to migrate downstream and begin the trophic phase of their lifecycle.

**Testicular development – trophic phase**

Although parasitic lampreys are still sexually immature at the completion of metamorphosis and reach maturity much later than non-parasitic species, they have the advantage of feeding for a greater time, which is rewarded by considerably larger body size and a higher fecundity (Hardisty, 1971). In parasitic lampreys, much of the testis development occurs during or after the trophic phase. Compared to the size of the testis during metamorphosis, the pouched lamprey, *Geotria australis*, has a 20-fold increase in size following this phase, and continue to undergo maturation during their migration to the spawning areas (Potter and Beamish, 1977; Potter and Robinson, 1991). At this time, the spermatogonia undergo meiosis I and II to produce primary and secondary spermatocytes and meiosis is completed with the production of spermatids that continue maturing during the lamprey upstream migration (Sower, 2003).

**Final spermatogenesis and sexual maturation – upstream migration and spawning**

Meiosis is completed during the lengthy upstream migration, with the production of spermatids. The spermatids continue maturing in the cysts for a period of approximately 2 months in the pouched lamprey, and sperm are released when the cyst walls degenerate, although species-specific differences in timing of maturation are present (Potter and Robinson, 1991). There is not an abundance of information available looking at the development of the testis
during the trophic phase, and most lampreys that are examined are collected during their upstream migration. This is mainly due to the difficulty in obtaining individuals when they are feeding at sea or in large lakes, but spermatogenesis still continues during their upstream migration.

The sea lamprey continues gonadal maturation during its upstream migration. Spermatids continue maturing into sperm during the spawning run, that typically lasts up to 4 months ranging from May to August (Potter and Beamish, 1977; Sower, 2003). During the upstream migration of the landlocked sea lamprey, lamprey were captured from May to June, and the gonadal histology was examined in those individuals (Fahien and Sower, 1990). Additional lamprey were caught and reared until July, to ensure histological examination of spermiation. Schreiner and Schreiner (1904), which was translated by Brodal and Fange (1963), described the maturational changes that take place in gametes during sperm production in Atlantic hagfish, *Myxine glutinosa*. Fahien and Sower (1990) categorized these changes into seven stages; *one*: presence of primary spermatocytes; *two*: primary and dividing spermatocytes, *three*: spermatocytes and spermatids, *four*: spermatids and immature sperm, *five*: immature sperm, *six*: immature and mature sperm, and *seven*: mature sperm. Based on these categories, lampreys captured in mid-May ranged from stages one to three depicting maturation from spermatocytes to spermatids. In the latter half of May and early June, during stages four to seven, spermatids through mature sperm were observed. In July, all individuals were at stage seven, where the testis was composed of mature sperm (Fahien and Sower, 1990). These seven stages of sperm production are not to be confused with the seven stages of metamorphosis, nor are they to be confused with the stages of testicular differentiation referred to in the current study (Section 3.1.4).
3.2.2. Gene expression during testicular differentiation

The testis is composed of two types of somatic tissues: interlobular tissue and lobule wall cells, which are mostly found in the vascular strands that segregate neighbouring lobules, and connective tissue (Hardisty, 1971). Interlobular tissue is composed of polygonal cells that appear similar to Leydig cells, due to their agranular reticulum structure and the presence of lipid and cholesterol deposits. The presence of these inclusions within the interlobular cells increase in later spermatogenesis and indicate a potential role during the degradation of lobules to release sperm (Hardisty, 1971). Lobule cell walls have rapid development of cells during the later stages of spermatogenesis and bear some resemblance to Sertoli cells. Both of these cell types have been implicated in steroid secretion.

Steroid 3-ol-dehyrdogenase is a steroid enzyme that determines the sites of cellular steroidogenesis (Hardisty and Barnes, 1968). Tests were conducted to detect concentrations of the steroid in lipid and cholesterol droplets during late spermatogonial development. Low concentrations were detected during the primary spermatocyte development stage, and concentrations rapidly increased as spermatids differentiated (Hardisty and Barnes, 1968). An abundance of work since the 1960s has looked at steroidogenic enzymes and sex steroids in lampreys (Yun et al., 2003; Bryan et al., 2008; Ahmadi et al., 2011; Farrokhnejad et al., 2014). An increase in the number of positive reactions for the presence of steroid biosynthesis were detected as secondary sexual characteristics developed (Hardisty, 1971). Steroid biosynthesis in the interlobular tissue of the testis is homologous to Leydig cells (Hardisty and Barnes, 1968).

Knowing this, we can see what genes are being expressed in the Leydig cells of other well-studied vertebrates and detect if they are present in maturing lampreys. For example, the mouse has expression of the genes Anti-Müllerian hormone (AMH) and SOX (SRY-related
HMG-box) transcription factor 9 (SOX9) that may be involved in sex differentiation as well as sex determination, and we can try to identify regions that may be homologous or conserved in lampreys (Takada et al., 2005). AMH is regulated by SOX9 and is required for testis development; upon over-expression of AMH, male chickens had disrupted expression of SOX9, males and females were incapable of gonadal steroidogenesis and mature individuals displayed an external female phenotype, emphasizing the role of AMH not only in testis development, but also its impact on steroidogenesis (Lambeth et al., 2016). More details on genes involved in sex determination and sex differentiation are included in Section 1.6.

In other vertebrates, an increase in germ cells in the testes has been found to coincide with increased expression of SOX (SRY-related HMG-box) transcription factor 8 (SOX8), SOX9, Steroidogenic factor 1 (SF1), Doublesex- and mab-3-related transcription factor 1 (DMRT1) or DMRT-family gene and Wilm’s tumour protein 1 (WT1) (Nachtigal et al., 1998; Marchand et al., 2000; Vizziano et al., 2007; Mawaribuchi et al., 2017) and higher Deleted in Azoospermia, DAZ, associated protein 1 (DAZAPI) expression was observed during spermatogenesis and spermiogenesis (Huang et al., 2008). The increase in SOX9 expression was followed by an increase in the expression of DMRT1 or DMRT-related genes during male differentiation (Marchand et al., 2000). I therefore wanted to test whether key Sertoli cell transcription factors involved in testicular differentiation and spermatogenesis in other vertebrates (SOX8, SOX9, SF1, DMRT1, WT1, and DAZAPI) show differential gene expression prior to and during testicular differentiation in sea lamprey.

No studies have examined gene expression during testicular differentiation, but gene expression has been investigated during ovarian differentiation in lampreys by Spice (2013) and Spice et al. (2014). Germ cell-less (GCL), which is expressed during testis development in
zebrafish, was found to be expressed at higher levels in age class 2 chestnut lamprey, *Ichthyomyzon castaneus*, and northern brook lamprey, *Ichthyomyzon fossor*, compared to other age class groups (Li et al., 2006; Spice et al., 2014), suggesting that it may play a role in early differentiation. *WT1*, which acts as a Müllerian inhibiting factor that prohibits the formation of Müllerian ducts, was more highly expressed in presumptive males than in females (Spice et al., 2014). Findings from these previous studies in lampreys thus also helped guide the current research.

Therefore, the objective of this study was to determine whether the genes identified above (*SOX8, SOX9, DMRT*-related gene, *SF1, GCL, WT1*, and *DAZAP1*) as key genes involved in testicular differentiation in other vertebrates are expressed in the sea lamprey in a sex specific and or a stage specific manner. Given that sex differentiation genes are relatively conserved among vertebrates (see Section 1.6), I predicted that these male differentiation genes will have similar roles during sea lamprey testicular differentiation. Primers were designed using the recently sequenced sea lamprey genome as a reference (Smith et al., 2013; see Chapter 2) and the level of gene expression prior to, during and after testicular differentiation was measured. Stages were assigned as followed based on developmental stages: stage 1) larval; 2) early metamorphosis; 3) mid-metamorphosis; 4) late metamorphosis; and 5) sexually mature. Gene expression was measured using qRT-PCR and compared among males from the different developmental stages, and compared similarly among the same stages in females. Females and males were assigned to stages based on developmental stage, and not stages of ovarian and testicular differentiation. In males, stage 1 represents presumptive but still undifferentiated males sampled after the point at which ovarian differentiation has occurred in the females, and stages 2-
4 are very active periods in terms of male differentiation, whereas metamorphosis in the sea lamprey female is a period of very little change (Lewis and McMillan, 1965).

This research will help determine whether the genes that are expressed during testicular differentiation in one of the two extant lineages of jawless vertebrates are the same involved in this process in other vertebrates and will further our understanding of gene conservation and evolution of sex differentiation in the vertebrate lineage. This work may also be able to determine when key molecular events that initiate testicular differentiation occur in lampreys and, based on gene expression profiles that are female or male biased, we can potentially differentiate the sex of an individual well before histological evidence shows signs of ovary or testis development. In European eel, *Anguilla anguilla*, sexually dimorphic gene expression patterns were used to estimate sex ratios prior to histological differentiation (Geffroy et al., 2016). European eel is another species in which the gonad remains undifferentiated for an extended period of time (Section 1.4.2); molecular markers for early sex identification would be a valuable tool in lampreys as well.

3.3. Methods

3.3.1. Sample collection and preservation

More than 300 sea lamprey larvae were collected from an anadromous population in Richibucto River, New Brunswick, using pulsed direct-current electrofishing. Larvae were collected in June 2015 and June 2016, and transported live in aerated 40L coolers to Wilfrid Laurier University in Waterloo, ON. The landlocked (Great Lakes) population was sampled using the identical protocol by US Fish and Wildlife Service (USFWS)/US Geological Survey (USGS) personnel from the Hammond Bay Biological Station (HBBS), Great Lakes Science Centre, Millersburg, MI. Landlocked sea lamprey larvae were kept at HBBS for approximately
4-6 weeks, before being shipped live to Wilfrid Laurier University in plastic bags with oxygen-saturated water, where they were subsequently maintained in well aerated aquaria (pH 7.8-8.0, temperature 16-18°C), provided with 4-5 cm of sand for burrowing. All sampling was performed under WLU Animal Care Committee Animal Utilization Protocol (AUP) R16005.

Samples were maintained at Wilfrid Laurier University and monitored for signs of metamorphosis (Manzon et al., 2015), at which point a subsample of individuals were collected during early, mid- and late metamorphosis. Larvae were sampled after the point at which metamorphosis was initiated in the rest of the population. Sea lamprey were anaesthetized with an overdose of MS222 (1.5 g/L plus 3.0 g/L NaHCO₃) and sacrificed by decapitating posterior to the last branchiopore. A homemade RNAlater recipe was developed by Jonathan Wilson and used to preserve tissues collected from Wilfred Laurier University. The RNAlater was made by combining 40 mL 0.5 M EDTA, 25 ml 1M sodium citrate, 700 g ammonium sulfate and 935 ml of sterile distilled water. The solution was stirred on a hot plate stirrer on low heat until the ammonium sulfate was completely dissolved. If the pH needed to be adjusted to pH of 5.2, it was titrated with 1M H₂SO₄. The solution was then transferred to a container and kept at room temperature or refrigerated. RNAlater was injected into the gut using a gauge needle to perfuse the intestine and preserve the RNA integrity of the internal organs. A vertical incision was made along the ventral midline from posterior to the last branchiopore to the cloaca. Tissues that remained in the carcass were the intestine, liver, gallbladder, kidneys and gonad. The carcass was then placed in a 10 mL Falcon tube and filled with RNAlater to fully saturate the tissue, and kept frozen at −20°C. Other tissues were collected to enable investigation of gene expression in other tissues during metamorphosis. The brain was also removed and placed in a 1.5 mL centrifuge tube with 1 mL of RNAlater. The tail was collected by sectioning posterior to the cloaca and
preserving in absolute ethyl alcohol, should high quality DNA be desired from these specimens in the future. The procedures up to this point were conducted at Wilfred Laurier University by Drs. Michael Wilkie and Jonathan Wilson and members of their laboratories, and the samples were subsequently shipped by courier to the University of Manitoba on dry ice. See Appendix A3.1 for information on collection location, sex, stage, date of sampling, length, weight, and details on histology use for each lamprey used in this study.

Once the samples arrived at the University of Manitoba, the carcass was removed from the Falcon tube using a pair of forceps dipped in ethyl alcohol and placed on a paper towel on a clean surface. Two forceps were used to pry the body wall away from the ventral-midline. The testis was nested between the kidneys, located posterior to the gut, running down the body length. The testis was pulled away from the body and sectioned into two using a clean razor blade. Each tissue sample was placed in a 1.5 mL centrifuge tube with 1 mL RNAlater, one for RNA gene expression analysis and the other sample for histology. Samples were kept at –20°C. Ovaries were kept frozen in RNAlater for further studies.

Sexually mature male and female sea lamprey were collected by Dr. John Hume (Michigan State University) from the Ocqueoc River or Cheboygan River, MI, in July 2016, shortly before spawning. Lampreys were measured for length and weight, and two subsamples of the gonads were taken for RNA and histological analysis and preserved in RNAlater and 10% formalin, respectively, prior to shipment to the University of Manitoba (see Section 3.2.6).

3.3.2. Identification of conserved testis differentiation genes and reference genes

Genes that are found to be expressed during testicular differentiation in other vertebrates were identified as potential candidate genes for the current study. Genes that showed consistent patterns of expression from all vertebrates and with a potential role identified were particularly
targeted, and sequences for the gene were obtained using Ensembl (Yates et al., 2016). The sea lamprey genome was sequenced by Smith et al. (2013) and annotated sequences are available on Ensembl under the model ‘lamprey’; Ensembl release 87: December 2016; Pmarinus_7.0 was used in the current study. Although the Arctic lamprey, Lethenteron camtschaticum genome is also available, it is not well annotated, and there is no other sequenced genome for a vertebrate species that would be closely related to lamprey. Whenever possible, genes for the sequences of interest were taken from other fish species. Mammals were less frequently included since they showed significant differences in gene sequence relative to fishes and have genes that evolved after more whole genome wide duplication events.

Seven candidate genes (WT1, DAZAP1, SOX8, SOX9, GCL, DMRTA2 and SF1) and one reference gene (RPL10a; see Section 3.2.4) were selected. If the gene of interest was found to already be annotated in the sea lamprey genome, the coding sequence was used to design primers for qPCR (Tables 3.1 and 3.2). Integrated Device Technology IDT PrimerQuest tool was used to design primers ideally with 50% GC, and a melting temperature (Tm) of 60⁰C, ensuring qPCR primer protocols were followed (Jozefczuk and Adjaye, 2011). The primer sequences that were generated were then scanned for primer-dimers, self-dimers and hairpins using the IDT OligoAnalyzer Tool. Any sequences that had a Tm that was similar to the temperatures used in PCR, or with a ΔG that was more negative than −9 kcal/mole were not used, as this increases the likelihood of self-dimerization and hetero-dimerization. The primer sequences were then checked against the sea lamprey genome on Ensembl using the BLAST/BLAT tool to ensure that both the forward and the reverse primers will target and amplify the correct gene of interest. In the case that a gene of interest was not annotated in the sea lamprey genome, primers were designed based on the sequences from other vertebrates. This was achieved by aligning
sequences from select other vertebrates using the Molecular Evolutionary Genetics Analysis (MEGA) software version 6.06 (Tamura et al., 2013; see Section 2.3.3). Areas that showed a high degree of conservation were identified as good sites for primer design. In order to target a region for designing primers, at least six consecutive nucleotides had to be similar in both the forward and reverse primer sequence, but ideally more than six nucleotides were conserved to reduce the possibility of non-specific amplification. These primer sequences were then searched for in the sea lamprey genome to identify the presumed genes of interest; Ensembl displays the percent similarity of the primer input sequence to the reference genome, and this was conducted to ensure the primers only matched the gene of interest, with a higher percent similarity than to other genes. If there was no other gene that appeared to be similar, the primers were selected for future testing of specificity, following the primer design procedure outlined above. Genes for which primers pairs were unsuccessful in amplifying the desired target are listed in Appendix A3.2.

3.3.3. RNA isolation and cDNA synthesis

A mortar and pestle apparatus was made using a 1.5 mL Eppendorf tube and P1000 (Fisher) pipette tip, where the point of the pipette tip was placed into the flame of a Bunsen burner and immediately placed into the tube, ensuring the flamed portion of the tip has met the bottom of the tube. The tip was then gently twisted out of the tube, and the opening of the tip was replaced by a smooth rounded edge to help homogenize the tissue. Approximately 30 mg of tissue was used to extract total RNA using the Qiagen RNeasy Mini Kit as per manufacturer’s instructions. An optional elimination of genomic DNA step, using the RNase-free DNase set (Qiagen), was incorporated into the protocol as indicated in the kit handbook. RNA quantity and
quality of 1 µL was determined using the NanoVue Plus spectrophotometer. Samples were always kept on ice and, following RNA isolation, immediately frozen at –80°C.

Complementary DNA (cDNA) was synthesized using the manufacturer’s instructions for the QuantiTect Reverse Transcription Kit (Qiagen) and using approximately 100 ng of isolated RNA. The cDNA concentration was measured using the NanoVue Plus spectrophotometer. Working stock concentration that was made had a final volume of 50 µL with approximately 100 ng of cDNA per mL. For qPCR assays, the samples were diluted to 1:24 using RNase-free water. For calibration curve purposes, the dilution series was made using 1:9, 1:24, 1:49, 1:99, 1:249 and 1:499.

3.3.4. Quantitative reverse-transcriptase polymerase chain reaction qRT-PCR

The Life Technologies StepOnePlus Real-time PCR system using Life (Applied Biosystems) SYBR Green PCR mix for assays was used to determine if the expression of the candidate sex differentiating genes is different between the sexes and among stages. Reference genes, beta-actin, tubulin and ribosomal protein RPL10a were created, but the most consistent results were obtained from the ribosomal protein gene, and RPL10a was used throughout as the reference gene. qPCR assays were conducted in duplicate for each individual for RPL10a and for all other target genes, using the default protocol using StepOnePlus™ Real-Time PCR System and StepOnePlus™ Software v2.3 (Applied Biosystems).

A calibration curve was made with a 5-fold dilution series of 1:9, 1:24, 1:49, 1:99 and 1:249. Calibration curves are included for the reference and target genes in the Appendix (see A3.3 to A3.10).
3.3.5. Gene expression analysis

A line of best fit was estimated from the calibration curve, and the slope of the lines was then used to determine the efficiency, (E) values, of the primer calibration curves (Jozefczuk and Adjaye, 2011). E values are ideally 100% for PCR reactions, where the product is expected to double with each cycle, and values were determined by using the following equation $E = 10^{1 - \frac{1}{\text{slope}}}$. E values were compared to the control sample, M3, a female larva that had already completed ovarian differentiation and was not yet undergoing sexual maturation, to allow for easier visualization and unmasking of genes involved in sex differentiation in males. Gene expression analysis was done using the Pfaffl (2001) method; E values were used to calculate the expression ratio (R), where $R = (E_{\text{target}})^{\Delta \text{CP}_{\text{Target}} (\text{control} - \text{sample})} / (E_{\text{reference RPL10a}})^{\Delta \text{CP}_{\text{ref}} (\text{control} - \text{sample})}$ (Pfaffl, 2001). The significance of the differences would be based on the fold difference expression of the gene at that stage, in the individual, to the expression of the base level reference gene $RPL10a$. Mean Ct values for gene expression during developmental stages 1 to 5 and mature males and females are included as averages in the Appendix (see A3.9.).

GraphPad Instat 3 (GraphPad Software, San Diego, California) was used to compare gene expression between developmental stages in males and females. Relative gene expression analysis was conducted using the non-parametric Kruskal-Wallis test with a Dunn’s post-test, with a $p$-value level of significance set at 0.05 (Figures 3.1-3.7). This test compared the expression of the target gene, relative to the reference gene, for developmental stages 1-5 separately for males and females. Secondly, pairwise comparisons between relative gene expression at each gonadal stage in males versus females were conducted using a Mann-Whitney test. An addition comparison was conducted between stage 1 females (late larval) and stage 2 males (early metamorphosis) to more explicitly compare gene expression in females which have
just completed ovarian differentiation to males which are beginning to undergo testicular differentiation. Recall that developmental stages in this study have been based on life cycle stage, and that strictly comparable stages of gonadal differentiation in males versus females are not being compared (Figures 1.1, Table 3.3), although stage 5 represents sexual maturity in both males and females.

3.3.6. **Histological sampling and preparation**

Because lampreys were sampled by colleagues at Wilfrid Laurier University and used for studying various tissues and organs, we were unable to take a cross-section through the body wall for histological analysis for these samples. Instead, the gonad was excised from the body after preservation in RNAlater and, at the time of harvest for RNA extraction from one half of the testis, the other half was transferred to 10% formalin. For larval and mature sea lamprey from Hammond Bay Biological Station (Dr. John Hume), a 5-mm cross-section was taken through the mid-point of the trunk of the animal and immediately placed into 10% formalin (see Appendix A3.1 for sample details).

Histological processing was done by the McGill Goodman Cancer Research Centre Histology Facility. Tissue samples were kept in 10% formalin until they were ready to be processed using the Sakura Tissue-Tek VIP 6 Vacuum Infiltration Processor system. Tissue samples were removed from the formalin solution and placed into 70% ethanol for 10 minutes, and transferred to 80% ethanol for 25 minutes at room temperature. The samples were then transferred into 95% ethanol for 25 minutes and repeated, and the same step was repeated twice with 100% ethanol for 25 minutes each. Following this step, the samples were then transferred into three different xylene solution jars, for 25 minutes in each. The samples were then
embedded into paraffin at 60°C for 45 minutes, and this was repeated for a total of four times. Paraffin embedded samples were then kept overnight.

Histology sections were made with a microtome at a thickness of 8 µm. Hematoxylin and eosin staining was done on a Leica ST5020 Multistainer, using the Leica ST Infinity H&E Staining system. The steps for slide staining were done according the manufacturer’s instructions for the kit. Slides were processed through two xylene solutions, for 5 minutes in each, transferred to 100% ethanol for 2 minutes, and repeated twice more in fresh solution. Slides were placed in 80% ethanol for 2 minutes, washed with running tap water for 2 minutes, and dipped into Hemalast (Leica Biosystems) for 30 seconds. Slides were stained with hematoxylin for 2 minutes and 30 seconds, washed with tap water for 2 minutes to remove excess stain, before being dipped into the differentiator for 45 seconds, rinsed for 2 minutes, and then immersed into bluing agent for 1 minute. After rinsing off excess stain with tap water for 2 minutes, transferred to 80% ethanol, the slides were dipped into eosin stain for 30 seconds, washed with 100% ethanol for 2 minutes, and repeated three times. Samples were placed into two solutions of xylene for 2 minutes each before being allowed to dry and cover-slipped.

3.3.7. Histological analysis

Images were captured using the Zeiss Axio Imager.Z1 microscope with Zeiss Axio Cam 105 colour camera. Images were processed using the Zeiss ZEN 2.3 blue edition software. The images were analyzed using ImageJ (National Institutes for Health).

The statistical test that was used for analysis of the histology sections was a Kruskal-Wallis test, which is a non-parametric test, and a Dunn’s post-test (GraphPad Software Inc). The entire cross-sectional area of the gonad within the body wall was measured and the number of germ cells in the entire cross-sectional area of the gonad was counted for each individual.
3.4. Results

3.4.1. Histology

A qualitative description of the processes and cell types that are present in each stage in males and females is given in Table 3.3; quantitative information (i.e., regarding gonad size, germ cell count) and other sample details are included in Tables 3.4 and 3.5. All of the stages of testicular differentiation mentioned in Tables 3.4 and 3.5 are illustrated in Figure 3.8. No obvious differences were observed between males from the anadromous and landlocked populations, although sample sizes were insufficient to test quantitatively. Over all samples, the maximum number of germ cells (62-458) were found in presumptive males during the larval stage (stage 1; e.g., Figure 3.8a to h) and the minimum number (2-66) was observed immediately thereafter (i.e., during early metamorphosis, stage 2; e.g., Figure 3.8i, j and k) (Table 3.4). The number of germ cells per cross-section was significantly different among stages 1-4 ($p = 0.0082$). The Dunn’s post test showed a significant 11.95-fold difference between the number of germ cells between stage 2 versus stage 4 (mean germ cell difference = 163, $p < 0.01$; Figure 3.8i, Figure 3.8n). There was no significant difference observed between other stages of gonadal development, including between stages 1 and 2, given the variation observed within each of these stages.

The mean area from the larval stage to late metamorphosis decreased from 82,723 $\mu m^2$ to 21,291 $\mu m^2$, respectively (Table 3.4), and was significantly different among stages ($p = 0.003$; Table 3.3). As with germ cell count, the greatest decrease was observed between stages 1 and 2.

3.4.2. Sex ratios

Metamorphosing sea lamprey from the anadromous population in the Richibucto River ($n = 72$) were highly female biased (80.5% female, 19.4% male). Metamorphosing landlocked sea
lamprey from the Great Lakes (n = 37) showed only a slight excess of females (59% female, 41% male) (Table 3.6). Mature adults collected from the Great Lakes basin showed equal sex ratios, although sample size was small (n = 8).

3.4.3. Gene expression

The total number of samples that were captured and available for use are included in Table 3.6, but only samples for which sex could be unambiguously identified were used for qPCR (Table 3.7). During analysis, individuals from landlocked and anadromous populations were combined (see Section 3.3.1), and the total number of each sex and stage used qPCR analysis were as follows:

<table>
<thead>
<tr>
<th>Stage</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>5 (Mature)</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Calibration curves for the reference gene and seven candidate genes expressed during testicular differentiation in sea lamprey are found in the Appendix (A3.3 to A3.10), and mean Ct values for males and females for gene expression is in A3.11. All target gene expression was compared to the relative gene expression of the reference gene RPL10a.

There were no significant differences ($p > 0.05$) in gene expression of GCL and SF1 between males and females at each developmental stage (Table 3.8), and no significant expression patterns were observed between stages in males ($GCL p = 0.2969; SF1 p = 0.2131$) or females ($GCL p = 0.3700; SF1 p = 0.2354$). No significant sex or stage specific differences in SOX8 expression were observed (Table 3.8), but gene expression differences approached significance in the two following comparisons: between stage 1 females and stage 2 males ($p =$
0.0727, potentially with higher expression in males) and stage 2 females and stage 2 males (p = 0.0831, again with potentially higher expression in males).

SOX9 expression displayed significant trends in males and females of various stages. Among females, very significant differences in SOX9 expression were observed among stages (p = 0.0013, Figure 3.1a). Stage 1 females had significantly greater expression than stage 2 females (17.90-fold) and stage 4 females (19.44-fold) (p < 0.05, Figure 3.1a). Males displayed extremely significant gene expression when compared to males of different stages of differentiation (p = 0.0001, Figure 3.1b). Stage 1 males had the greatest expression of all the stages, with a 26.12-fold higher expression than stage 4 (p < 0.001). Mature males had a 17.67-fold higher expression than stage 4 males (p < 0.05, Figure 3.1b). Based on gene expression comparisons between males and females, stage 1 females had a significantly higher expression than stage 2 males (p = 0.0040, Figure 3.2a) Among stage 2 individuals, males had a greater expression in SOX9 than females (p = 0.0062, Figure 3.2b). Stage 4 males compared to stage 4 females had significantly higher expression (p = 0.0349, Figure 3.2c).

There were also significant sex and stage specific differences in gene expression of WT1. Among females, sexually mature adults had higher expression of WT1 than in other stages of females (p = 0.0135, Figure 3.3a). Mature males also showed significantly higher expression relative to developing males (p = 0.0073). Among developing males, WT1 expression was 20.71-fold higher in stage 1 relative to stage 2 (Figure 3.3b). Stage 1 females had significantly higher expression than stage 2 males (p = 0.0137, Figure 3.4a), and stage 4 males had significantly higher expression than stage 4 females (p = 0.0212, Figure 3.4b).

There were no significant differences in gene expression patterns for DMRTA2 in pairwise comparisons between males versus females of the same stages. Within males,
significant differences in gene expression were observed between stages \((p = 0.0066)\). Stage 1 males had the greatest expression in \textit{DMRTA2}, being 25.5-fold greater than in stage 3 males \((p = 0.0066, \text{Figure 3.5})\).

Significant differences in expression of \textit{DAZAP1} were seen among stages in males \((p = 0.0012, \text{Figure 3.6})\), where mature males had 20.41, 25.43 and 17.50-fold higher gene expression than stage 2, 3 and 4 males, respectively \((p < 0.05)\). Gene expression comparisons between males and females were not significant (Table 3.8), although two differences approached significance: stage 1 males versus stage 1 females and stage 4 males versus stage 4 females \((p = 0.0768\) and \(p = 0.0836\), respectively), with \textit{DAZAP1} expression appearing slightly higher in males.

Pooling all the females from different developmental stages into one group and males into another group, gene expression was compared from the two groups to see if there were any sex specific general patterns. The only gene that showed significant difference between males and females was \textit{SOX9}, with males exhibiting slightly greater overall expression than females \((p = 0.0407, \text{Figure 3.7})\).

\textbf{3.5. Discussion}

Testicular differentiation takes place soon before or at the onset of metamorphosis in the sea lamprey (Hardisty, 1965). The current study showed that histological signs of male differentiation start to become apparent during the late larval period, at lengths as small as 92-101 mm. Most significantly, this is the first study to look at gene expression during testicular differentiation in lampreys. Following the gene expression profile of seven candidate genes identified from other vertebrates across stages of male gonadal differentiation (and compared to gene expression patterns in ovaries), this study suggests that \textit{WT1}, \textit{SOX9}, \textit{DAZAP1} and \textit{DMRTA2}
up-regulation might precede and predict testicular differentiation in sea lamprey, thus possibly serving as bio-markers for lamprey testicular differentiation.

*DAZAP1* is associated with the *DAZ* family of proteins that are exclusively expressed in testes during all stages of development (Fu et al., 2015). We observed high expression of *DAZAP1* in males at stage 1 (i.e., during the late larval stage, prior to obvious histological signs of testicular differentiation; see Table 3.3) and in sexually mature males just prior to spawning (Figure 3.6). Similar patterns are expressed in mice, *Mus musculus*, where the highest expression has been observed in individuals undergoing spermatogenesis and/or spermiogenesis (Huang et al., 2008). As expected, the increase in *DAZAP1* expression in sea lamprey at the beginning of the differentiation process at stage 1 corresponded with occurrence of a large number of germ cells starting to undergo differentiation and degeneration and type A spermatogonia starting to grow (Huang et al., 2008; Figure 3.8 A-H). Expression of *DAZAP1* was also significantly higher in mature males that were undergoing late spermiogenesis, with a large number of spermatozoa permeating the gonad, which is consistent with what is known about expression in the mouse gonad (Yan and Yen, 2013; Figure 3.8p). We expected expression to be high in stage 1 and mature stage male sea lamprey, in order to have successful production of spermatozoa at this crucial point, since deficiencies in *DAZAP1* are associated with failure in spermatogenesis (Yang and Yen, 2013).

*SOX9* has been linked to male testicular differentiation in humans, mice, chicken, *Gallus gallus*, lambari fish, *Astyanax altiparanae*, Nile tilapia, a number of other birds and non-mammalian vertebrates, and mammals (Morais de Silva et al., 1996; Wagner et al., 1994; Vidal et al., 2001; Takada et al., 2005; Adolff et al., 2015; Ijiri et al., 2008; Cutting et al., 2013; Sekido and Lovell-Badge, 2008). Our results for what kind of expression pattern to expect from *SOX9*,
with an increase observed in males, matches what is in the literature. \textit{SOX9} was revealed to be one of the first genes identified to be related to sRY, and having sex specific expression, becoming male specific during differentiation (Sinclair et al., 1990; Wegner, 1998). In sea lamprey males, \textit{SOX9} was up-regulated was in stage 1, stage 2 and stage 4 (Figure 3.1b, Figure 3.2b, c). \textit{SOX9} gene expression was greatest at stage 1, coinciding with the first signs of spermatogenesis, when many germ cells and type A spermatogonia appear in the gonad (Figure 3.1b; Table 3.4). During stage 2, although there has been a decrease in the gonadal area and in number of undifferentiated germ cells, the number of type A and B spermatogonia start to increase, coinciding with the increase in \textit{SOX9} expression at this stage (Figure 3.2b, Figure 3.8i-k).

Although it can be difficult to extrapolate information for gene expression from vertebrates with an earlier differentiation period than the sea lamprey, the Siberian sturgeon, \textit{Acipenser baerii}, is similar to the sea lamprey in many respects. In both species, there are no known sex-linked markers, the males have an immature gonad for a period of years, followed by a very slow differentiation process where the gonad continues to grow slowly (Hardisty 1965, Hardisty 1969, Berbejillo et al., 2012). A similar pattern of \textit{SOX9} gene expression is observed in both of these species, with expression in the sea lamprey decreasing in subsequent stages of differentiation at stage 3, and then rising in mature males, and this is similar to what is observed in the Siberian sturgeon. In stage 4 sea lamprey males, \textit{SOX9} expression increased, which histologically coincides with a drastic increase in the number of germ cells, with most of the connective tissue in the gonad being replaced by an abundance of darkly stained round cells, which are primary spermatocytes (Figure 3.2c; Figure 3.8n-o; Table 3.4). \textit{SOX9} was expressed in both male and female lamprey and this is the case for many other fish species, such as the
walking catfish, *Clarias batrachus*, where the highest expression has been observed in the gonads and brain, and a marked decrease was observed after spawning (Bhat et al., 2016).

Although expression of *SOX9* in the male lambari fish is very high during spermatogenesis and significantly decreases during spermiogenesis, this is not the case for what was observed in our mature lamprey, due to spermiation not yet being completed in the mature lamprey (Adolfi et al., 2015; Figure 3.1b). The mature male lamprey gonad samples consist of maturing spermatozoa which continue to mature until fertilization, and marking the end of spermiogenesis.

Doublesex and mab-3-related transcription factor A2, *DMRTA2*, is the lamprey gene that has been identified to share a great degree of similarity to *DMRT1* gene sequences that are available for other vertebrate species, although two *DMRT* transcription factors are available, for the sea lamprey, *DMRTA2* and *DMRT2A* (see Section 2.5.3). In birds, the master sex determining gene is Doublesex and mab-3-related transcription factor 1, *DMRT1*, and it is required in a dosage dependent manner for all birds (Yano et al., 2012; Shetty et al., 2002; Smith et al., 1999). Double-sex and mab-3 related-domain gene on the Y chromosome, DMY, is the testis determining factor in the Japanese rice-fish, and the Malabar ricefish, *Oryzias curvinotus* (Matsuda et al., 2002; Nanda et al., 2002). *DMRT1* is thought to play a role in the sex determination cascade in *Oryzias* species (Myosho et al., 2012; Heule et al., 2014). Based on our expression data, *DMRT1* is very highly expressed during stage 1 in late larval sea lamprey males (Figure 3.6). Expression of *DMRT1* in mice is detected well before any histological signs of male differentiation are apparent, as I observed when comparing gene expression and histological data (Nanda et al. 2002; Ijiri et al., 2008). Following this stage during which *DMRT1* was highly expressed, stage 2 male sea lamprey exhibited a dramatic increase in the number of germ cells, with germ cells of type A and B spermatogonia found in clusters (Figure 3.8i-k;
Figure 3.6). The expression of DMRT1 in male sea lamprey is coordinated with the increase in expression of SOX9 at stage 1, and continuing to be expressed into stage 2, signifying cross-talk between genes. In the mouse gonad, DMRT1 is required for testis maintenance, and works in concert with SOX8 and SOX9 to maintain the integrity of the testis; upon deletion of SOX9, DMRT1 is not expressed in males, resulting in reprogramming of the testis and testis-to-ovary reversal (Barrionuevo et al., 2016). Based on this, SOX9 expression is required at the early stages of lamprey testicular differentiation, and is needed to ensure DMRT1 is expressed in the developing male gonad.

WT1 is expressed in the gonads of both male and female sea lamprey (Figure 3.3). In the male sea lamprey, an increase in expression was observed during stage 1, during the late larval stage when germ cells are first apparent; WT1 then stayed at a relatively constant low level but increased again in stage 4 males, during late metamorphosis when the gonad consists of primary spermatocytes (Figure 3.4b). My observations of WT1 expression broadly agree with results in chestnut lamprey and northern brook lamprey, where higher expression was observed in presumptive males prior to testicular differentiation, as well as prior to ovarian differentiation (Spice, 2013; Spice et al., 2014). In other vertebrates, WT1 has been found to aid in Sertoli cell and spermatogonia differentiation, and deficiencies in mice led to spermatogenic arrest, preventing male differentiation (Zheng et al., 2014). Based on this, I expected that WT1 would be up-regulated during the onset of differentiation, and when spermatogonia are present in abundance, and I did indeed find that WT1 is up-regulated in stage 1 when histological analysis showed the germ cells starting to differentiate into type A spermatogonia (Figure 3.3b; Figure 3.8a-h). The highest expression of WT1 was observed at stage 4, when the gonad is permeated with a large number of primary spermatocytes (Figure 3.4b; Figure 3.8n-o). SF1 is one factor
that *WT1* is reliant on, for constant expression to activate transcription, leading to *WT1* transcription factors bind to the DNA and RNA binding site of *sRY* in mammals (Kato et al., 2012; Matsuzawa-Watanabe et al., 2003). Mutations in *WT1* have profound effects that don’t allow the *WT1*-sRY complex to be formed, leading to gonadal dysgenesis (Matsuzawa-Watanabe et al., 2003).

Although I did not find any significant up- or down-regulation in Steroidogenic factor 1 (*SF1*), this may be because expression at consistent levels is required to ensure proper expression of other genes. Loss of *SF1* can not only lead to the loss of structural integrity of the gonad, but it also leads to a decrease in the expression of other important Sertoli cell factors *WT1, SOX9* and the androgen receptor (Kato et al., 2012). *WT1* and *SOX9* had specific male-biased expression patterns during testicular differentiation in sea lamprey, and any loss of *SF1* expression would interfere with spermatogenesis and testicular development processes (Figures 3.1-3.4). *SF1* is also known to up-regulate Anti-Müllerian hormone (*AMH*) in male mice and its regulative role is known in other instances; for example, in the mammalian testis, it up-regulates *sRY* and *SOX9*, which is essential for testis differentiation (Schepers et al., 2003; Takada et al., 2005; Sekido and Lovell-Badge, 2008). Despite not finding any significant expression of *SF1* at different developmental stages in either male or female sea lamprey, we were able to determine that this presumptive *SF1* gene showed close homology to the *SF1* gene of other annotated genes in vertebrates (see Section 2.4; Figure 2.13-2.14).

Similarly to *SF1*, *SOX8* did not show significant changes during differentiation in male and female sea lamprey, with consistent expression during all stages of development and no sex specific expression. *SOX8* is essential for testis differentiation and works in conjunction with *SF1* to up-regulate *AMH* in mice (Schepers et al., 2003). *SOX8* expression, relative to the
reference gene \textit{RPL10a}, was not statistically significant in the sea lamprey gonad at any stage, but neared significance at stage 1 and 2 when compared to other males, which coincides with the stages where \textit{SOX9} was up-regulated. In experiments in mice with \textit{SOX9} removal, \textit{SOX8} was able to temporarily compensate for the loss of \textit{SOX9}: seminiferous tubules continued differentiation and testis development followed the expected trend, although mature male mice were still sterile due to eventual spermatogenic failure (Barrionuevo et al., 2009). This suggests that constant expression of \textit{SOX8} may be required to ensure that other transcription factors and genes are activated and expressed.

Based on histology and gene expression in sea lamprey, it thus appears that expression of \textit{SOX9, DMRTA2, WT1} and \textit{DAZAP1} are associated with the initial testicular differentiation period at both stage 1 and stage 2 (i.e., during the late larval stage and early metamorphosis). Lamprey testis lobule cell walls appear similar in structure to Sertoli cells (Hardisty and Barnes, 1968; Hardisty, 1971), and transcription factors that are associated with Sertoli cells during spermatogenesis are \textit{SOX9, DMRTA2} and \textit{WT1} (Berbejillo et al., 2012; Kato et al., 2012). \textit{SOX9} and \textit{WT1} showed an increase in expression again upon the appearance of primary spermatocytes, and \textit{DAZAP1} was highly expressed in mature males when spermatids are present.

Samples from stage 1 were larvae that, based on total length, presumably belong to age-classes 3 and 4, which were grouped together due to the small number available for each cohort. In this group, two individuals displayed a large variation in the number of oocytes present in the presumed males, ranging from 1 oocyte to 71 oocytes per cross-section of the entire gonad. More meaningful results can be obtained by having a larger sample size from the intermediate stages, as results can be very skewed based on small sample sizes (Table 3.4; Figure 3.8).
Similarly, although sample sizes were small, female biased sex ratios were observed among the metamorphosing sea lamprey studied here: the anadromous and landlocked populations were 80% and 59% female, respectively. In the Great Lakes, a preponderance of female larvae was observed after the initiation of sea lamprey control measures, leading to suggestions of density-dependent sex determination (see Section 1.3.3). Conversely, in times of high larval density in Lewis Creek, Vermont, high density of sea lamprey larvae resulted in male biased sex ratios (Zerrenner and Marsden, 2005), and a similar pattern was observed in the least brook lamprey, *Lampetra aepyptera*. The sex ratios observed in the current study thus suggest that population density in the Richibucto River is lower than in the Great Lakes. However, the sea ratio of upstream migrating lampreys is often male biased; for example, 3 males to 1 female sea lamprey were observed in the Carp River by Applegate and Thomas (1965) and, in upstream migrating sea lamprey in Lake Erie, an excess of males was observed between the years 2007 and 2008 (Hansen et al., 2016). Within the anadromous population in Europe, data from the Garonne Basin in France shows a slightly greater proportion of female sea lamprey (Beaulaton et al., 2008). Observation of sex ratios during metamorphosis are thus important to determine whether there might be differences in sea ratio among different stages, potentially as the result of sex specific differences in rates of metamorphosis or rates of mortality following metamorphosis. For the purposes of the current study, the female biased sex ratios observed in the two populations sampled resulted in a small number of males available for gene expression and histological analysis at each stage.

The samples that we were able to obtain successful histological sections for were metamorphosing lamprey mostly from the anadromous population (see Appendix A3.1). Although populations of the same species which show similar patterns of growth, Hardisty
(1965; Hardisty, 1969) suggested that there were marked differences in the time of
differentiation between the landlocked and anadromous population, with the anadromous
population showing a slower proliferation rate with a higher number of undifferentiated germ
cells well into the later stages of development, and accelerated gonadal growth and development
in the landlocked population. In the current study, no obvious histological differences were
observed and samples were combined for analysis of gene expression, but comparisons in the
future should systematically compare male differentiation between these two populations, both
histologically and in terms of gene expression.

As the first study looking at gene expression during testicular differentiation in lampreys,
this work provides exciting results and support for the conservation of genes involved in sex
differentiation as identified from other vertebrates. In sea lamprey DMRT-related gene, SOX9,
WT1 and DAZAPI expression was associated with the onset of histological male differentiation.
This study identified a higher expression of SOX9 in males during the late larval stage and at
sexual maturity (corresponding with an increase in germ cells), significantly greater expression
of SOX9 in males than in females at the onset of metamorphosis (i.e., during the early stages of
testicular differentiation compared to females which have already undergone ovarian
differentiation), higher WT1 expression in the late larval stage (i.e., at presumed onset of spermatogenesis). DAZAPI was expressed at all stages of male development with greatest
expression at periods associated with the onset of spermatogenesis and spermiogenesis, and
higher expression also in mature males that contained spermatozoa. DMRT-related gene
expression was highest in late larval stage males, and no significant expression of DMRT-related
gene and DAZAPI was found in females. Although much work remains to be done to fully
elucidate the sex determination and sex differentiation genes in lampreys, and ways in which this
research might have applications for conservation and control (see Chapter 4), this study showed that many genes involved in testicular differentiation appear to be conserved across vertebrates, although the involvement of lamprey-specific genes cannot yet be ruled out.
3.6. Literature cited


### 3.7. Tables and Figures

Table 3.1. Reference and seven candidate genes for qPCR, including their presumptive role during testicular differentiation and source of sequence for primer design. * designates the gene is not annotated, and it was presumed that the gene was the indicated gene based on homology.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene Symbol</th>
<th>Primer Type</th>
<th>Role</th>
<th>Source</th>
<th>Ensembl ID/ GenBank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribosomal Protein 110a</td>
<td><em>RPL10A</em></td>
<td>Reference</td>
<td>Suitable for qPCR validation(^1)</td>
<td><em>P. marinus</em></td>
<td>ENSPMAG00000008233</td>
</tr>
<tr>
<td>Wilms tumour protein</td>
<td><em>WT1</em></td>
<td>Target</td>
<td>Important in testis development(^2)</td>
<td><em>P. marinus</em></td>
<td>ENSPMAT00000008950</td>
</tr>
<tr>
<td>Deleted in azoospermia-associated protein 1</td>
<td><em>DAZAP1</em></td>
<td>Target</td>
<td>Highly expressed in testes, deficiency leads to spermatogenic arrest(^3)</td>
<td><em>Ichthyomyzon fossor</em> transcriptome</td>
<td>See Spice (2013)</td>
</tr>
<tr>
<td>SOX (SRY-related HMG-box) transcription factor 8</td>
<td><em>SOX8</em></td>
<td>Target</td>
<td>Essential for testis differentiation(^4)</td>
<td><em>P. marinus</em></td>
<td>ENSPMAT00000011398</td>
</tr>
<tr>
<td>SOX (SRY-related HMG-box) transcription factor 9</td>
<td><em>SOX9</em></td>
<td>Target</td>
<td>Regulate testis development(^5)</td>
<td><em>P. marinus</em></td>
<td>DQ136023.1</td>
</tr>
<tr>
<td>Germ cell-less</td>
<td><em>GCL</em></td>
<td>Target</td>
<td>Promotes germ cell development,(^6) spermatogenesis(^7)</td>
<td><em>Ichthyomyzon fossor</em> transcriptome</td>
<td>See Spice (2013)</td>
</tr>
<tr>
<td>Doublesex and mab-3-related transcription factor a2</td>
<td><em>DMRTA2</em></td>
<td>Target</td>
<td>Testis-differentiation(^8)</td>
<td><em>P. marinus</em></td>
<td>ENSPMAT00000004009</td>
</tr>
<tr>
<td>Steroidogenic factor 1</td>
<td><em>SF1</em></td>
<td>Target</td>
<td>Essential for spermatogenesis(^9)</td>
<td><em>P. marinus</em></td>
<td>ENSPMAG0000001450</td>
</tr>
</tbody>
</table>
1. Chen et al., 2011
2. Matsuzawa-Watanabe et al., 2003
3. Yang and Yen, 2013 et al., 2009
4. Berbejillo et al., 2012
5. Lerit et al., 2017
6. Kimura et al., 2003
7. Adolfini et al., 2015
8. Kato et al., 2012
9. Takada et al., 2005
Table 3.2. Primers for reference and target genes used for qPCR, size of gene transcript length, location of forward and reverse primers within the gene sequence, amplicon size, and optimal PCR annealing temperature for each primer pair.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>Primer Type</th>
<th>Forward sequence</th>
<th>Reverse sequence</th>
<th>Transcript length (bp)</th>
<th>Start and stop location</th>
<th>Fragment size amplified (bp)</th>
<th>Annealing temperature °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribosomal Protein 110a</td>
<td>RPL10A</td>
<td>Ref.</td>
<td>GTTGCTGTTGGG CATGTA</td>
<td>CGTTCTGCC AGTTCTTCTT</td>
<td>651</td>
<td>F: 496-514 R: 580-599</td>
<td>103</td>
<td>54.1</td>
</tr>
<tr>
<td>Wilms tumour protein</td>
<td>WT1</td>
<td>Target</td>
<td>CCGTCTGATCAC TTGAAGAC</td>
<td>ACTTCTTCTGTGCAGTTTGG</td>
<td>903</td>
<td>F: 733-753 R: 808-827</td>
<td>94</td>
<td>54.1</td>
</tr>
<tr>
<td>Deleted in azoospermi-associate protein 1</td>
<td>DAZAP1</td>
<td>Target</td>
<td>GGGATTACTTTA CAGAGGTGG</td>
<td>TGGCCTTCA ACTGACTCT</td>
<td>309</td>
<td>F: 125-147 R: 230-248</td>
<td>123</td>
<td>53.9</td>
</tr>
<tr>
<td>SOX (SRY-related HMG-box) transcription factor 8</td>
<td>SOX8</td>
<td>Target</td>
<td>ACTCGGTATGAC GACGA</td>
<td>GCTTGCACTT GGAGGAG</td>
<td>489</td>
<td>F: 238-255 R: 348-365</td>
<td>127</td>
<td>54.5</td>
</tr>
<tr>
<td>SOX (SRY-related HMG-box) transcription factor 9</td>
<td>SOX9</td>
<td>Target</td>
<td>AGCCACTACAG CGATCAAC</td>
<td>CTGGTGTGTGTGAGGGCGAATG</td>
<td>1,307</td>
<td>F: 859-878 R: 916-937</td>
<td>78</td>
<td>55.7</td>
</tr>
<tr>
<td>Germ cell-less</td>
<td>GCL</td>
<td>Target</td>
<td>TCCACCACAACT TGATGAC</td>
<td>AACAGCCACT TCTTGAGC</td>
<td>-</td>
<td>-</td>
<td>145</td>
<td>53.8</td>
</tr>
<tr>
<td>Doublesex and mab-3-related transcription factor A2</td>
<td>DMRTA2</td>
<td>Target</td>
<td>CTACAAACTCCCG GTGTCG</td>
<td>TCTTAGCAG GAGGCCTGT</td>
<td>432</td>
<td>F: 25-43 R: 113-130</td>
<td>105</td>
<td>54.7</td>
</tr>
<tr>
<td>Steroidogenic factor 1</td>
<td>SF1</td>
<td>Target</td>
<td>GGTGGACGATC AGATGAGC</td>
<td>CCCGTGACC AGAGGAGT</td>
<td>564</td>
<td>F: 3-23 R: 100-117</td>
<td>114</td>
<td>57.7</td>
</tr>
</tbody>
</table>
Table 3.3. Summary of male and female gonadal development at stage categories used in this study: 1 = late larval stage; 2, 3, 4 = early, mid and late metamorphosis; 5 = sexually mature. Male development images were taken for this thesis and female development image for stage 1 was originally published in Hardisty (1965) © 1965 John Wiley and Sons. The female ovarian development figures for stages 2-5 were originally published in Lewis and McMillan (1965) © 2005 John Wiley and Sons. Figures were reproduced with permission. Figure 1.1 depicts when stages occur in the lamprey life cycle, and detailed summaries of gonad size and cell types are included in Table 3.4 and 3.5.

<table>
<thead>
<tr>
<th>Developmental stage</th>
<th>Male development</th>
<th>Gonadal development and cell type</th>
<th>Female development</th>
<th>Gonadal development and cell type</th>
</tr>
</thead>
</table>
| Larval stage 1      | ![Image](https://via.placeholder.com/150) | Germ cells – Type A spermatagonia  
> gonad resembles the undifferentiated gonad  
> oocytes may still be found but fewer in number and undergoing degeneration  
> germ cells present in clusters  
> structure bulky and lobulated | ![Image](https://via.placeholder.com/150) | Primordial germ cells – primary oocytes  
> many cell nests  
> increase in number of oocytes  
> oocyte growth and enlargement  
> further lobulation of ovary structure  
> few densely stained/basophilic oocytes |
| Early metamorphosis 2 | ![Image](https://via.placeholder.com/150) | Type A/B spermatagonia  
> decrease in gonad size  
> decrease in loose reticulum connective tissue, indicative of testis differentiation  
> increase in the number of germ cell cysts | ![Image](https://via.placeholder.com/150) | Primary oocytes  
> full complement of oocytes present before metamorphosis  
> metamorphosis is ovary “resting” phase  
> absence of cell nests |
| Mid-metamorphosis 3  | ![Image](https://via.placeholder.com/150) | Type B spermatagonia and primary spermatocytes  
> testis structure starts to change and becomes separated in lobes permeated with germ cell cysts  
> amount of connective tissue and gonad size continues to decrease  
> with a simultaneous increase in the number of germ cells. | ![Image](https://via.placeholder.com/150) | Primary oocytes  
> no cell nests  
> oocytes at same developmental stage  
> large oocytes with irregular nucleus |
| Late-metamorphosis 4 | ![Image](https://via.placeholder.com/150) | Primary spermatocytes  
> increase in size of gonad compared to early and mid-differentiation stage  
> testis is starting to lose lobation- characteristic or early differentiation  
> progressive decrease in connective tissue.  
> large number of germ cells  
> primary spermatocytes with dense-stained cells and large nuclei | ![Image](https://via.placeholder.com/150) | Primary oocytes  
> no cell nests  
> vacuolated nucleoli  
> basophilic oocytes |
| Mature 5             | ![Image](https://via.placeholder.com/150) | Spermatozoa  
> spermatozoa in gonad of mature sea lamprey  
> tremendous increase in gonad following trophic phase  
> final maturation during upstream migration until spawning in summer | ![Image](https://via.placeholder.com/150) | Ovulation  
> rapid growth and increase in oocyte diameter  
> yolk platelets and vitelline membrane increase  
> released oocyte after ovulation |
Table 3.4. Details of gonadal development in male sea lamprey *Petromyzon marinus*, at stage categories used in this study: stage 1 (presumptive male larvae, age-class 3-4), stage 2 (early metamorphosis), stage 3 (mid-metamorphosis), and stage 4 (late metamorphosis). Range of sample size length, area of gonad, average area of gonad with standard deviation, and germ cell count (range and average) are given.

<table>
<thead>
<tr>
<th>Developmental stage</th>
<th>Gonadal stage</th>
<th>Length (mm)</th>
<th>N</th>
<th>Area of gonad (µm²)</th>
<th>Average area</th>
<th>Number of oocytes</th>
<th>Germ cell / spermatogonia count</th>
<th>Average germ cell count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Larvae age-class 3-4</td>
<td>Stage 1</td>
<td>92 – 101</td>
<td>7</td>
<td>41,509 – 148,898</td>
<td>82,723 ± 37,945</td>
<td>1 – 71</td>
<td>62 – 458</td>
<td>137</td>
</tr>
<tr>
<td>Early metamorphosis</td>
<td>Stage 2</td>
<td>112 – 132</td>
<td>4</td>
<td>1,066 – 24,892</td>
<td>11,630 ± 10,995</td>
<td>0</td>
<td>2 – 66</td>
<td>35</td>
</tr>
<tr>
<td>Mid-metamorphosis</td>
<td>Stage 3</td>
<td>125 – 133</td>
<td>2</td>
<td>11,453 – 17,580</td>
<td>14,516 ± 4,332</td>
<td>0</td>
<td>111 – 138</td>
<td>124</td>
</tr>
<tr>
<td>Late metamorphosis</td>
<td>Stage 4</td>
<td>118 – 127</td>
<td>5</td>
<td>14,612 – 38,772</td>
<td>21,291 ± 10,203</td>
<td>0</td>
<td>125 – 300</td>
<td>194</td>
</tr>
</tbody>
</table>
Table 3.5. Description of stages of testicular development in larval and metamorphosing sea lamprey *Petromyzon marinus*. Stage of development was identified based on presence of more than half of cell type present in the differentiating testis based on testis development stages in the sea lamprey (Hardisty, 1965) and zebrafish *Danio rerio* (Schulz et al., 2010). Table 3.4 and Figure 3.8 correspond with this table.

<table>
<thead>
<tr>
<th>Stage of development of presumptive males</th>
<th>Stage</th>
<th>Gonad features</th>
<th>Gonad development description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Larval</td>
<td>1</td>
<td>Germ cells – Type A spermatagonia</td>
<td>The gonad is largest in area at this stage. Germ cells are present that resemble the state of the indifferent gonad. Oocytes may still be found but are fewer in number and undergoing degeneration, indicative of the male gonad. The number of germ cells starts to increase, resembling type A differentiating spermatagonia, and are present in clusters, with the number of germ cells increasing in number as the size of the gonad increases. As the gonad continues to grow, the structure starts to change to become more bulky and lobulated. The number of germ cells starts to decrease at body lengths of approximately 90 mm in males and females, but this decrease is more obvious in males as oocyte growth in females masks degeneration of undifferentiated germ cells.</td>
</tr>
<tr>
<td>Early differentiation</td>
<td>2</td>
<td>Type A/ B spermatagonia</td>
<td>There is a significant increase in the amount of loose reticulum connective tissue, indicative of testis differentiation with no oocytes persisting. Dramatic increase in the number of germ cell cysts that range from type A spermatagonia to type B spermatagonia and large decrease in size of gonad.</td>
</tr>
<tr>
<td>Mid-differentiation</td>
<td>3</td>
<td>Type B spermatagonia and primary spermatocytes</td>
<td>The testis structure starts to change and becomes separated in lobes permeated with germ cell cysts that are starting to develop primary spermatocytes. The amount of connective tissue and gonad size continues to decrease with a simultaneous increase in the number of germ cells.</td>
</tr>
<tr>
<td>Late differentiation</td>
<td>4</td>
<td>Primary spermatocytes</td>
<td>Dramatic increase in size of gonad compared to early and mid-differentiation stage. The testis is starting to lose lobation which is a characteristic or early differentiation, and progressive decrease in connective tissue. There is a large increase in the number of germ cells; most germ cells at this stage will be primary spermatocytes, with dense-stained cells. The whole gonad consists of packed germ cell nests, and although the germ cells don’t appear too different from undifferentiated germ cells, the defining characteristic of transforming male gonad is the presence of larger nuclei.</td>
</tr>
<tr>
<td>Mature</td>
<td>5</td>
<td>Spermatids</td>
<td>A lack number of densely packed rounded spermatids developing in the testicular tissue is evident in maturing males. Spermiogenesis is completed during the final period leading up to fertilization, where the spermatids continue maturing and will become more elongated in shape, and into motile sperm or spermatozoa.</td>
</tr>
</tbody>
</table>
Table 3.6. Total number of larval, metamorphosing and mature adult sea lamprey available for this study, *Petromyzon marinus*, although not all the samples were used for qPCR (see Table 3.7 for samples used). Percentage of females and males collected from each population is indicated. * = specific sampling information is not available for mature sea lamprey that were collected by the United States Fish and Wildlife Service from Ocqueoc River or Cheboygan River that drain into Lake Huron, or Carp Lake River that drains into Lake Michigan. ** = additional information of positively identified samples as male or female is found in Table 3.7. Coordinates for sampling location indicated in Table 3.7.

<table>
<thead>
<tr>
<th>Location</th>
<th>Life stage</th>
<th>n</th>
<th>% Females</th>
<th>% Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richibucto River, NB- Anadromous</td>
<td>Metamorphosing</td>
<td>72</td>
<td>80.5</td>
<td>19.4</td>
</tr>
<tr>
<td>Hammond Bay Biological Station, MI- Great Lakes Landlocked</td>
<td>Metamorphosing</td>
<td>37</td>
<td>59</td>
<td>41</td>
</tr>
<tr>
<td>Hammond Bay Biological Station</td>
<td>Larvae</td>
<td>70</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Lake Huron or Lake Michigan*</td>
<td>Mature adults</td>
<td>8</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>
Table 3.7. Sampling locations for larval, metamorphosing and mature sea lamprey, *Petromyzon marinus*. Sampling location includes the larger water body associated with sampling locale and, if known, the exact sampling location along with map coordinates. * = only larval samples where the sex of the larvae was positively determined were included. *N/A = indicates sampling site information is not available; samples were collected from either the Ocqueoc River or Cheboygan River that drain into Lake Huron, or Carp Lake River that drains into Lake Michigan. Table includes samples that were used for experiments based on information given in Table 3.6.

<table>
<thead>
<tr>
<th>Name of water body</th>
<th>Sampling location</th>
<th>Stage</th>
<th>N</th>
<th>Females</th>
<th>Males</th>
<th>Map coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Au Sable River</td>
<td>Michigan</td>
<td>Larval*</td>
<td>19</td>
<td>9</td>
<td>10</td>
<td>44° 47' N 51.61&quot; N 84° 45' W 28.64&quot; W 45° 29' 28.21&quot; N 84° 02' 2.72&quot; W 46° 49' 16.73&quot; N - 64° 56' 20.33&quot; W</td>
</tr>
<tr>
<td>Hammond Bay Biological Station</td>
<td>Michigan</td>
<td>Metamorphosing</td>
<td>24</td>
<td>10</td>
<td>14</td>
<td>45º 29' N 84º 02' 2.72&quot; W 46° 49' 16.73&quot; N - 64° 56' 20.33&quot; W</td>
</tr>
<tr>
<td>Richibucto River</td>
<td>New Brunswick</td>
<td>Metamorphosing</td>
<td>25</td>
<td>15</td>
<td>10</td>
<td>45º 29' N 84º 02' 2.72&quot; W 46° 49' 16.73&quot; N - 64° 56' 20.33&quot; W</td>
</tr>
<tr>
<td>Hammond Bay Biological Station</td>
<td>Michigan</td>
<td>Post-metamorphic</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>45º 29' N 84º 02' 2.72&quot; W 46° 49' 16.73&quot; N - 64° 56' 20.33&quot; W</td>
</tr>
<tr>
<td>Lake Huron or Lake Michigan</td>
<td>Northern Michigan</td>
<td>Mature adults</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>N/A*</td>
</tr>
</tbody>
</table>

*N/A*
Table 3.8. *P*-values reported for statistical analyses for gene expression in *DAZAP1*, *DMRTA2*, *GCL*, *SF1*, *SOX8*, *SOX9* and *WT1*. The Kruskal-Wallis (non-parametric ANOVA) test was conducted for all females and all males (i.e., across developmental stages in females and across developmental stages in males), and if values were significant, a Dunn’s multiple comparisons test was done. All pairwise comparisons were conducted using the Mann-Whitney test. Females from each developmental stage were compared to males of the same developmental stage; also, stage 1 females (i.e., in the late larval stage but which have recently undergone ovarian differentiation) were compared to males undergoing early testicular differentiation in early metamorphosis (stage 2). Values are broadly categorized as: es = extremely significant (*p* < 0.001), vs = very significant (*p* = 0.001-0.01), s = significant (*p* = 0.01-0.05), and nqs = not quite significant (*p* = 0.06-0.10); all other values were not significant even at *p* = 0.10.

<table>
<thead>
<tr>
<th>Comparison of developmental stages</th>
<th><em>DAZAP1</em></th>
<th><em>DMRTA2</em></th>
<th><em>GCL</em></th>
<th><em>SF1</em></th>
<th><em>SOX8</em></th>
<th><em>SOX9</em></th>
<th><em>WT1</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>All females</td>
<td>0.1177</td>
<td>0.8654</td>
<td>0.3700</td>
<td>0.2131</td>
<td>0.3133</td>
<td>0.0013</td>
<td>0.0135</td>
</tr>
<tr>
<td>All males</td>
<td>0.0012</td>
<td>0.0066</td>
<td>0.2969</td>
<td>0.2354</td>
<td>0.2321</td>
<td>0.0001</td>
<td>0.0073</td>
</tr>
<tr>
<td>Stage 1 females vs. males</td>
<td>0.0768</td>
<td>0.1419</td>
<td>0.6354</td>
<td>0.1878</td>
<td>0.9451</td>
<td>0.2398</td>
<td>0.8392</td>
</tr>
<tr>
<td>Stage 2 females vs. males</td>
<td>0.447</td>
<td>0.6334</td>
<td>0.6334</td>
<td>0.1011</td>
<td>0.0831</td>
<td>0.0062</td>
<td>0.8241</td>
</tr>
<tr>
<td>Stage 3 females vs. males</td>
<td>0.4534</td>
<td>0.2500</td>
<td>0.9999</td>
<td>0.5714</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.7857</td>
</tr>
<tr>
<td>Stage 4 females vs. males</td>
<td>0.0836</td>
<td>0.3786</td>
<td>0.4820</td>
<td>0.1227</td>
<td>0.1553</td>
<td>0.0349</td>
<td>0.0212</td>
</tr>
<tr>
<td>Stage 5 females vs. males</td>
<td>0.7922</td>
<td>0.9307</td>
<td>0.4286</td>
<td>0.9999</td>
<td>0.9307</td>
<td>0.9999</td>
<td>0.7922</td>
</tr>
<tr>
<td>Stage 1 females vs. stage 2 males</td>
<td>0.7986</td>
<td>0.9999</td>
<td>0.8081</td>
<td>0.2141</td>
<td>0.0727</td>
<td>0.0040</td>
<td>0.0137</td>
</tr>
</tbody>
</table>
Figure 3.1. Mean (+ SE) gene expression (relative to the reference gene RPL10a) of SOX9, SOX (SRY-related HMG-box) transcription factor 9, in sea lamprey, *Petromyzon marinus*, from: a) females, showing very significant \((p = 0.0013)\) differences in gene expression with stage; and b) males, showing extremely significant \((p = 0.0001)\) differences with stage. For pairwise comparisons, significant fold-differences \((p < 0.05)\) between stages are indicated. Sample sizes for each sex and stage are given in Section 3.4.3; details of gonadal development at each stage are given in Tables 3.4 and 3.5.
Expression ratio

Developmental stages

Stage 1 Female

Stage 2 Male

Stage 2 Female

Stage 2 Male

Stage 4 Female

Stage 4 Male

\[\text{SOX9} \quad p = 0.0040\]

\[\text{SOX9} \quad p = 0.0062\]

\[\text{SOX9} \quad p = 0.0349\]
Figure 3.2. Mean (± SE) gene expression (relative to the reference gene RPL10a) of SOX9, SOX (SRY-related HMG-box) transcription factor 9, in sea lamprey, Petromyzon marinus, from: a) stage 1 females compared to stage 2 males (p = 0.0040, very significant); b) stage 2 females compared to stage 2 males (p = 0.0062, very significant); and c) stage 4 females compared to stage 4 males (p = 0.0349, significant). Sample sizes for each sex and stage are given in Section 3.4.3; details of gonadal development at each stage are given in Tables 3.4 and 3.5. P-values of non-significant comparisons are presented in Table 3.8.
Figure 3.3. Mean (± SE) gene expression (relative to the reference gene RPL10a) of WT1, Wilms tumour protein, in sea lamprey, *Petromyzon marinus*, from: a) females, showing significant (*p* = 0.0135) differences in gene expression with stage; and b) males, showing very significant (*p* = 0.0073) differences with stage. For pairwise comparisons, significant fold-differences (*p* < 0.05) between stages are indicated. Sample sizes for each sex and stage are given in Section 3.4.3; details of gonadal development at each stage are given in Tables 3.4 and 3.5.
Figure 3.4. Mean (± SE) gene expression (relative to the reference gene *RPL10a*) of *WT1*, Wilms tumour protein, in sea lamprey, *Petromyzon marinus*, from: a) stage 1 females compared to stage 2 males (*p* = 0.0137, significant); and b) stage 4 females compared to stage 4 males (*p* = 0.0212, significant). Sample sizes for each sex and stage are given in Section 3.4.3; details of gonadal development at each stage are given in Tables 3.4 and 3.5. *P*-values of non-significant comparisons are presented in Table 3.8.
Figure 3.5. Mean (+ SE) gene expression (relative to the reference gene \textit{RPL10a}) of \textit{DMRTA2} Doublesex and mab-3-related transcription factor A2, in male sea lamprey, \textit{Petromyzon marinus}, showing very significant ($p = 0.0066$) differences with stage. Sample sizes for each stage are given in Section 3.4.3; details of gonadal development at each stage are given in Tables 3.4 and 3.5. \textit{P}-values of non-significant comparisons are presented in Table 3.8.
Figure 3.6. Mean (± SE) gene expression (relative to the reference gene RPL10a) of DAZAP1 Deleted in azoospermia-associated protein 1 in male sea lamprey, Petromyzon marinus, showing very significant ($p = 0.0012$) differences with stage. Sample sizes for each stage are given in Section 3.4.3; details of gonadal development at each stage are given in Tables 3.4 and 3.5. $P$-values of non-significant comparisons are presented in Table 3.8.
Figure 3.7. Mean (± SE) gene expression (relative to the reference gene \textit{RPL10a}) of \textit{SOX9}, SOX (SRY-related HMG-box) transcription factor 9, in all female (n=36) and male (n=41) sea lamprey, \textit{Petromyzon marinus}, showing significant ($p = 0.0407$) differences with sex.
a) Stage 1- Cross section through gonad of presumptive male larva, 92 mm length, and estimated age-class 3-4. DC pointing to degenerating cyst in gonad. Germ cells are present in clusters and resemble undifferentiated germ cells and type A undifferentiated spermatagonia. Germ cell nest peripheries are not well-defined, and are starting to permeate the lobed gonad and showing signs of degeneration. 20X magnification (LV 55).

b) Stage 1- Cross section through gonad of presumptive male larva, 98 mm length and estimated age-class 3-4. Germ cells are found in more obvious clusters resembling type A undifferentiated spermatagonia with progressive increase in loose reticulum connective tissue. 20X magnification (LV 57).
c) Stage 1 - Cross section through gonad of presumptive male larva, 92 mm length, and estimated age-class 3-4. Gonad continues to increase in size as the number of germ cells increase and are found throughout the gonad, in nests are randomly dispersed. Gonad structure is starting to become more bulky and lobed and the appearance of Type A undifferentiated spermatogonia. 20X magnification (LV 58).

d) Stage 1 - Cross section through gonad of presumptive male larva, 99 mm length and estimated age-class 4. Dramatic increase in the number of germ cells and gonad size. Type A spermatogonia germ cell type. Increase in the amount of loose reticulum connective tissue. 20X magnification (LV 59)
c) Stage 1 - Image through same gonad cross section as in d (see above). Oocytes can be expected to persist up to lengths of 10 mm in males. The gonad otherwise resembles an undifferentiated gonad, but the presence of oocytes helps to distinguish it as differentiating. DC is pointing to one of the many germ cells undergoing degeneration, and O is pointing to an oocyte 40X magnification (LV 59)

f) Stage 1 - Cross section through gonad of presumptive atypical male larva, 95 mm length and estimated age-class 3-4. Atypical presumptive male. Gonad is starting to become more bulky and lobed, permeated with oocytes which is atypical from what has been observed in other presumptive males of the same stage. All the oocytes are basophilic which precedes degeneration, and more than half of the oocytes are undergoing degeneration. Germ cells mostly segregated from oocytes, but some are found among oocytes and are differentiating into type A spermatogonia. DO designates a cluster of degenerating oocytes, which is characterized by the breakdown of cell inclusions, and GC pointing to germ cell clusters. 20X magnification. (LV 60)
g) Stage 1 - Cross section through gonad of presumptive male larva, 101 mm length and estimated age-class 4. Pronounced lobed structure of the gonad with fewer germ cells as degeneration continues, but the persisting germ cells are larger in size. Connective tissue is present in the centre and periphery of the gonad. 20X magnification (LV 61).

h) Stage 1 - Image through same gonad cross section as in g (see above). Germ cells are increasing in size and are of type A spermatogonia, and degeneration continues. DC is pointing to one of many degenerating cysts. 40X magnification (LV 61).
i) Stage 2 - Cross section through gonad of early metamorphosing male, 112 mm in length. Very small gonad with only a few germ cells present along with loose reticulum connective tissue. Dramatic decrease in gonad size from larval stage. Large amount of germ cells have undergone degeneration, and the few that are present are type A spermatogonia and are lining the periphery of the gonad. 40X magnification (16-49).

j) Stage 2 - Cross section through gonad of early metamorphosing male, 123 mm in length. This size, shape, and presence of densely-stained cells and granular elements are typical of early testicular differentiation. Decrease in the number of germ cells, but germ cells are differentiating as type A and B spermatogonia. 40X magnification (16-57).
k) Stage 2 - Cross section through gonad of early metamorphosing male, 132 mm in length. The structure of the gonad is starting to change in shape as it continues developing. Large number of densely-stained cells. Germ cells still resemble type A and B spermatogonia. GC is pointing to an isolated cluster of germ cells. 40X magnification (16-47).

l) Stage 3 - Cross section through gonad of mid metamorphosing male, 133 mm in length. The size and the shape has become larger and more finger-like lobes housing groups of germ cells. Germ cells range from type B spermatogonia with signs of development of early primary spermatocytes, with a large number of densely stained and packed germ
m) Stage 3- Cross section through gonad of mid metamorphosing male, 125 mm in length. Dramatic development of the gonad, in size, shape and an increase in the number of germ cells. Germ cells consist of early primary spermatocytes and are found in the developed lobes. 40X magnification (16-76).

n) Stage 4- Cross section through the gonad of late metamorphosing male, 127 mm in length. Gonad has increased in size and has lost the lobes associated with earlier stages. Primary spermatocytes are found in cysts throughout the gonad and very little connective tissue is left between them and mostly restricted to the central portion of the gonad. 40X magnification (16-96).
o) Stage 4- Cross section through the gonad of late metamorphosing male, 127 mm in length. Gonad showing extensive lobation and increase in size. Lobes will continue to regress as the gonad differentiates. Lobes encompass cysts of primary spermatocytes that will continue development into the trophic phase. 40X magnification (16-95).

p) Stage 5- Cross section through testis showing spermatozoa of mature male prior to the completion of spermiogenesis, 454 mm in length. This image included to show the dramatic development of the gonad during the trophic phase, downstream and upstream migration. 20X magnification (SLM 1).
Figure 3.8. The processes of male testicular development during the late larval stage (prior to histological differentiation), metamorphosis (during which testicular differentiation occurs), and sexual maturity in the sea lamprey *Petromyzon marinus*. Scale bars are present on the bottom right or top right of each image. Following paraffin embedding, 8 µm sections were stained with hematoxylin and eosin and images were captured using the Zeiss Axio Imager.Z1 microscope with Zeiss Axio Cam 105 colour camera. Images were processed using the Zeiss ZEN 2.3 blue edition software, and enhanced using ImageJ. Sample numbers for each individual pictured above are given in parentheses, and can be matched with sample details presented in Appendix A3.1.
### Appendices

A3.1. List of differentiating males, males, female and undifferentiated sea lamprey, *Petromyzon marinus*, samples used for qPCR. Details on life stage (larval, anadromous or landlocked metamorphosing transformers, or mature adults), developmental/gonadal stage used for comparisons in this study (as categorized in Section 1.5.2 and Tables 3.4 and 3.5), larval age class (AC= age-class, as estimated by length), stage of metamorphosis in transformers (Amm.= ammocoete/larva; Manzon et al., 2015), or post-metamorphic stage (Juv.= juvenile, i.e., following stage 7 but prior to sexual maturation; Mature). Length of lamprey (mm) and weight (grams) are included along with collection date and location (HBBS = Hammond Bay Biological Station), date sacrificed for sampling, and whether histology was done on samples. All bolded sample ID names have corresponding histology images in Figure 3.8.

<table>
<thead>
<tr>
<th>Sex</th>
<th>Life stage</th>
<th>Developmental/gonadal stage</th>
<th>Larval age class (AC), stage of metamorphosis (1-7) or post-metamorphosis</th>
<th>Length (mm)</th>
<th>Mass (g)</th>
<th>Location</th>
<th>Collection date</th>
<th>Sampling date</th>
<th>Histology</th>
<th>Sample ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diff.</td>
<td>Larval</td>
<td>1</td>
<td>AC 3-4</td>
<td>92</td>
<td>--</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>Yes</td>
<td>LV 55</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff.</td>
<td>Larval</td>
<td>1</td>
<td>AC 3-4</td>
<td>97</td>
<td>--</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>Yes</td>
<td>LV 56</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff.</td>
<td>Larval</td>
<td>1</td>
<td>AC 3-4</td>
<td>98</td>
<td>--</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>Yes</td>
<td>LV 57</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff.</td>
<td>Larval</td>
<td>1</td>
<td>AC 3-4</td>
<td>92</td>
<td>--</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>Yes</td>
<td>LV 58</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff.</td>
<td>Larval</td>
<td>1</td>
<td>AC 3-4</td>
<td>95</td>
<td>--</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>Yes</td>
<td>LV 60</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff.</td>
<td>Larval</td>
<td>1</td>
<td>AC 4</td>
<td>99</td>
<td>--</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>Yes</td>
<td>LV 59</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff.</td>
<td>Larval</td>
<td>1</td>
<td>AC 4</td>
<td>101</td>
<td>--</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>Yes</td>
<td>LV 61</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff.</td>
<td>Larval</td>
<td>1</td>
<td>AC 4</td>
<td>96</td>
<td>--</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>--</td>
<td>LV 51</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff.</td>
<td>Larval</td>
<td>1</td>
<td>AC 4</td>
<td>94</td>
<td>--</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>--</td>
<td>LV 53</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff.</td>
<td>Larval</td>
<td>1</td>
<td>AC 4</td>
<td>95</td>
<td>--</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>--</td>
<td>LV 54</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (anadromous)</td>
<td>2</td>
<td>Amm/1</td>
<td>128</td>
<td>3.28</td>
<td>Richibucto River</td>
<td>June, 2015</td>
<td>August 26th, 2015</td>
<td>--</td>
<td>M23 M</td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (anadromous)</td>
<td>2</td>
<td>1</td>
<td>112</td>
<td>2.1</td>
<td>Richibucto River</td>
<td>June, 2015</td>
<td>August 19th, 2016</td>
<td>Yes</td>
<td>16-49 M</td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (anadromous)</td>
<td>2</td>
<td>1</td>
<td>117</td>
<td>2.5</td>
<td>Richibucto River</td>
<td>June, 2015</td>
<td>July 26th, 2016</td>
<td>Yes</td>
<td>16-27 M</td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (anadromous)</td>
<td>2</td>
<td>1</td>
<td>122</td>
<td>2.14</td>
<td>Richibucto River</td>
<td>June, 2015</td>
<td>August 6th, 2015</td>
<td>Yes</td>
<td>M8 M</td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (anadromous)</td>
<td>2</td>
<td>1</td>
<td>123</td>
<td>2.3</td>
<td>Richibucto River</td>
<td>June, 2015</td>
<td>August 19th, 2016</td>
<td>Yes</td>
<td>16-57 M</td>
</tr>
<tr>
<td>Sex</td>
<td>Transformers (anadromous)</td>
<td>Size</td>
<td>Length</td>
<td>River</td>
<td>Date</td>
<td>Expected</td>
<td>Yes/No</td>
<td>Size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>---------------------------</td>
<td>------</td>
<td>--------</td>
<td>-------</td>
<td>------</td>
<td>----------</td>
<td>--------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>2</td>
<td>1</td>
<td>Ricibucto River</td>
<td>6/13</td>
<td>June, 2015</td>
<td>Yes</td>
<td>16-47 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>2</td>
<td>2</td>
<td>HBBS</td>
<td>11/1</td>
<td>November 27th, 2015</td>
<td>--</td>
<td>M36 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>3</td>
<td>3</td>
<td>Richibucto River</td>
<td>7/26</td>
<td>June, 2015</td>
<td>Yes</td>
<td>16-32 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>3</td>
<td>5</td>
<td>Richibucto River</td>
<td>8/13</td>
<td>June, 2015</td>
<td>Yes</td>
<td>16-76 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>3</td>
<td>3 to 4</td>
<td>HBBS</td>
<td>6/13</td>
<td>August 6th, 2015</td>
<td>Yes</td>
<td>M10 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (anadromous)</td>
<td>4</td>
<td>6</td>
<td>Ricibucto River</td>
<td>6/13</td>
<td>June, 2015</td>
<td>Yes</td>
<td>16-88 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (anadromous)</td>
<td>4</td>
<td>7</td>
<td>Richibucto River</td>
<td>8/13</td>
<td>August 6th, 2015</td>
<td>Yes</td>
<td>16-89 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (anadromous)</td>
<td>4</td>
<td>7</td>
<td>Richibucto River</td>
<td>8/13</td>
<td>October 6th, 2016</td>
<td>Yes</td>
<td>16-95 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (anadromous)</td>
<td>4</td>
<td>7</td>
<td>Richibucto River</td>
<td>8/13</td>
<td>October 6th, 2016</td>
<td>Yes</td>
<td>16-96 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>4</td>
<td>7</td>
<td>HBBS</td>
<td>6/13</td>
<td>October 5th, 2015</td>
<td>Yes</td>
<td>M29 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>4</td>
<td>7</td>
<td>HBBS</td>
<td>6/13</td>
<td>October 5th, 2015</td>
<td>Yes</td>
<td>M26 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>4</td>
<td>7</td>
<td>HBBS</td>
<td>6/13</td>
<td>November 27th, 2015</td>
<td>Yes</td>
<td>M49 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>4</td>
<td>7</td>
<td>HBBS</td>
<td>6/13</td>
<td>November 27th, 2015</td>
<td>--</td>
<td>M40 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>4</td>
<td>7</td>
<td>HBBS</td>
<td>6/13</td>
<td>October 5th, 2015</td>
<td>Yes</td>
<td>M27 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>4</td>
<td>6 to 7</td>
<td>HBBS</td>
<td>6/13</td>
<td>October 5th, 2015</td>
<td>Yes</td>
<td>M30 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>4</td>
<td>6 to 7</td>
<td>HBBS</td>
<td>6/13</td>
<td>November 27th, 2015</td>
<td>--</td>
<td>M42 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>4</td>
<td>6 to 7</td>
<td>HBBS</td>
<td>6/13</td>
<td>October 5th, 2015</td>
<td>Yes</td>
<td>M28 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>4</td>
<td>6 to 7</td>
<td>HBBS</td>
<td>6/13</td>
<td>November 27th, 2015</td>
<td>--</td>
<td>M47 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>4</td>
<td>6 to 7</td>
<td>HBBS</td>
<td>6/13</td>
<td>November 27th, 2015</td>
<td>--</td>
<td>M46 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>4</td>
<td>6 to 7</td>
<td>HBBS</td>
<td>6/13</td>
<td>November 27th, 2015</td>
<td>--</td>
<td>M44 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Transformers (landlocked)</td>
<td>4</td>
<td>6 to 7</td>
<td>HBBS</td>
<td>6/13</td>
<td>March 24th, 2016</td>
<td>Yes</td>
<td>M52 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Stage</td>
<td>Age</td>
<td>Length</td>
<td>River</td>
<td>Date 1</td>
<td>Date 2</td>
<td>Yes/No</td>
<td>Code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>------</td>
<td>--------</td>
<td>-------------</td>
<td>----------</td>
<td>----------</td>
<td>--------</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Mature</td>
<td>5</td>
<td>425</td>
<td>Ocqueoc River or Cheboygan River</td>
<td>July 21st, 2016</td>
<td>July 21st, 2016</td>
<td>Yes</td>
<td>SLM 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Mature</td>
<td>5</td>
<td>454</td>
<td>Ocqueoc River or Cheboygan River</td>
<td>July 21st, 2016</td>
<td>July 21st, 2016</td>
<td>Yes</td>
<td>SLM 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Mature</td>
<td>5</td>
<td>486</td>
<td>Ocqueoc River or Cheboygan River</td>
<td>July 21st, 2016</td>
<td>July 21st, 2016</td>
<td>Yes</td>
<td>SLM 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Mature</td>
<td>5</td>
<td>497</td>
<td>Ocqueoc River or Cheboygan River</td>
<td>July 21st, 2016</td>
<td>July 21st, 2016</td>
<td>Yes</td>
<td>SLM 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Larval</td>
<td>1</td>
<td>90</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>--</td>
<td>LV 44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Larval</td>
<td>1</td>
<td>93</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>--</td>
<td>LV 46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Larval</td>
<td>1</td>
<td>91</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>--</td>
<td>LV 48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Larval</td>
<td>1</td>
<td>80</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>--</td>
<td>LV 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Larval</td>
<td>1</td>
<td>80</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>--</td>
<td>LV 19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Larval</td>
<td>1</td>
<td>80</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>--</td>
<td>LV 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Larval</td>
<td>1</td>
<td>80</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>--</td>
<td>LV 23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Larval</td>
<td>1</td>
<td>100</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
<td>July 22nd, 2016</td>
<td>--</td>
<td>LV 42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Transformers (anadromous)</td>
<td>2</td>
<td>1</td>
<td>Richibucto River</td>
<td>June, 2015</td>
<td>August 6th, 2015</td>
<td>--</td>
<td>M6 F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Transformers (anadromous)</td>
<td>2</td>
<td>1</td>
<td>Richibucto River</td>
<td>June, 2015</td>
<td>August 6th, 2015</td>
<td>--</td>
<td>M2 F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Transformers (anadromous)</td>
<td>2</td>
<td>1</td>
<td>Richibucto River</td>
<td>June, 2015</td>
<td>August 26th, 2015</td>
<td>--</td>
<td>M18 F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Transformers (anadromous)</td>
<td>2</td>
<td>1</td>
<td>Richibucto River</td>
<td>June, 2015</td>
<td>August 6th, 2015</td>
<td>--</td>
<td>M3 F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Transformers (anadromous)</td>
<td>2</td>
<td>1</td>
<td>Richibucto River</td>
<td>June, 2015</td>
<td>August 26th, 2015</td>
<td>--</td>
<td>M22 F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Transformers (anadromous)</td>
<td>2</td>
<td>1</td>
<td>Richibucto River</td>
<td>June, 2015</td>
<td>August 26th, 2015</td>
<td>--</td>
<td>M21 F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Description</td>
<td>Age</td>
<td>Weight</td>
<td>Location</td>
<td>Date</td>
<td>Code</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------------------</td>
<td>-----</td>
<td>--------</td>
<td>----------</td>
<td>------------------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Female   | Transformers (landlocked)                  | 2   | 132    | HBBS     | August 6th, 2015 | M12 F
| Female   | Transformers (landlocked)                  | 2   | 134    | HBBS     | November 27th, 2015 | M33 F
| Female   | Transformers (landlocked)                  | 2   | 136    | HBBS     | November 27th, 2015 | M35 F
| Female   | Transformers (landlocked)                  | 2   | 132    | HBBS     | August 6th, 2015 | M15 F
| Female   | Transformers (anadromous)                  | 3   | 130    | Richibucto River | June, 2015 | 16-48 F
| Female   | Transformers (anadromous)                  | 3   | 143    | Richibucto River | June, 2015 | M4 F
| Female   | Transformers (anadromous)                  | 3   | 123    | Richibucto River | June, 2015 | 16-50 F
| Female   | Transformers (anadromous)                  | 3   | 131    | Richibucto River | June, 2015 | 16-39 F
| Female   | Transformers (anadromous)                  | 3   | 131    | Richibucto River | June, 2015 | 16-61 F
| Female   | Transformers (anadromous)                  | 4   | 129    | Richibucto River | June, 2015 | M41 F
| Female   | Transformers (anadromous)                  | 4   | 114    | Richibucto River | June, 2015 | M50 F
| Female   | Transformers (landlocked)                  | 4   | 122    | HBBS     | October 5th, 2015 | M31 F
| Female   | Transformers (landlocked)                  | 4   | 129    | HBBS     | November 27th, 2015 | M48 F
| Female   | Transformers (landlocked)                  | 4   | 132    | HBBS     | November 27th, 2015 | M39 F
| Female   | Transformers (landlocked)                  | 4   | 136    | HBBS     | October 5th, 2015 | M25 F
| Female   | Transformers (landlocked)                  | 4   | 125    | HBBS     | October 5th, 2015 | M32 F
| Female   | Transformers (landlocked)                  | 4   | 134    | HBBS     | November 27th, 2015 | M45 F
| Female   | Transformers (landlocked)                  | 4   | 140    | HBBS     | November 27th, 2015 | M43 F
| Female   | Post-metamorphic juveniles                 | 4   | 125    | HBBS     | March 24th, 2017 | M53 F
| Female   | Post-metamorphic juveniles                 | 4   | 134    | HBBS     | March 24th, 2018 | M51 F
| Female   | Mature adults                              | 5   | 437    | Ocqueoc River or | July 21st, 2016 | SLF 5

258
<table>
<thead>
<tr>
<th>Sex</th>
<th>Life Stage</th>
<th>ID</th>
<th>Quantity</th>
<th>Performance</th>
<th>Date 1</th>
<th>Date 2</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ocqueoc River or Cheboygan River</td>
<td>SLF 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ocqueoc River or Cheboygan River</td>
<td>SLF 4</td>
</tr>
<tr>
<td>Female</td>
<td>Mature adults</td>
<td>5</td>
<td>Mature</td>
<td>476</td>
<td>246</td>
<td>July 21st, 2016</td>
<td>July 21st, 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ocqueoc River or Cheboygan River</td>
<td>SLF 2</td>
</tr>
<tr>
<td>Undiff</td>
<td>Larval</td>
<td>1</td>
<td>AC 2</td>
<td>69</td>
<td>--</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
</tr>
<tr>
<td>Undiff</td>
<td>Larval</td>
<td>1</td>
<td>AC 2-3</td>
<td>79</td>
<td>--</td>
<td>Au Sable River</td>
<td>July 22nd, 2016</td>
</tr>
</tbody>
</table>

Au Sable River
A3.2. List of genes for which initial attempts at qPCR were unsuccessful, due to non-specific amplification, no amplification, or the results were not consistent. * indicates genes that were not annotated in the sea lamprey genome, and **bolded** gene symbol indicate genes that were successfully amplified with different primer pairs.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene Symbol</th>
<th>Primer Type</th>
<th>Role</th>
<th>Source</th>
<th>Primer pairs made</th>
<th>Primer problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOX (SRY-related HMG-box) transcriptation factor 9</td>
<td><strong>SOX9</strong></td>
<td>Target</td>
<td>Regulate testis development¹</td>
<td><em>P. marinus</em> genome</td>
<td>2</td>
<td>Inconsistent results two pairs made, one worked</td>
</tr>
<tr>
<td>SOX (SRY-related HMG-box) transcriptation factor 8</td>
<td><strong>SOX8</strong></td>
<td>Target</td>
<td>Essential for testis differentiation²</td>
<td><em>P. marinus</em> genome</td>
<td>2</td>
<td>Non-specific amplification – two pairs made, one had consistent results</td>
</tr>
<tr>
<td>Wilms tumour protein 1</td>
<td><strong>WT1</strong></td>
<td>Target</td>
<td>Important in testis development³</td>
<td><em>P. marinus</em> genome</td>
<td>6</td>
<td>Non-specific amplification – six pairs made, two had successful amplification</td>
</tr>
<tr>
<td>Daz interacting protein</td>
<td><strong>DZIP</strong></td>
<td>Target</td>
<td>Expressed in meiotic male germ cells⁴</td>
<td><em>P. marinus</em> genome</td>
<td>1</td>
<td>No amplification</td>
</tr>
<tr>
<td>Double-sex and mab-3 related transcriptation factor 2</td>
<td><strong>DMRT2</strong></td>
<td>Target</td>
<td>Testis differentiation⁵</td>
<td><em>P. marinus</em> genome</td>
<td>1</td>
<td>No amplification</td>
</tr>
<tr>
<td>Male germ-cell associated kinase</td>
<td><strong>MAK</strong></td>
<td>Target</td>
<td>Expressed in male germ cells⁶</td>
<td><em>P. marinus</em> genome</td>
<td>1</td>
<td>No amplification</td>
</tr>
<tr>
<td>Steroidogenic factor 1*</td>
<td><strong>SF1</strong></td>
<td>Target</td>
<td>Essential for spermatogenesis⁷</td>
<td><em>P. marinus</em> genome*</td>
<td>2</td>
<td>No amplification with one primer pair, second primer pair worked</td>
</tr>
<tr>
<td>Anti Müllerian hormone *</td>
<td><strong>AMH</strong></td>
<td>Target</td>
<td>Role in spermatogenesis⁸; Müllerian duct regression⁹</td>
<td><em>P. marinus</em> genome*</td>
<td>1</td>
<td>No amplification</td>
</tr>
<tr>
<td>Zebrafish testis-expressed</td>
<td><strong>ZTE38</strong></td>
<td>Target</td>
<td>Enhanced in zebrafish testis¹⁰</td>
<td><em>P. marinus</em> genome*</td>
<td>1</td>
<td>No amplification</td>
</tr>
<tr>
<td>Lamprey testis-expressed 9</td>
<td><strong>TEX9</strong></td>
<td>Target</td>
<td>Expression in spermatogenesis</td>
<td><em>P. marinus</em> genome</td>
<td>1</td>
<td>No amplification</td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------</td>
<td>--------</td>
<td>-------------------------------</td>
<td>---------------------</td>
<td>---</td>
<td>------------------</td>
</tr>
<tr>
<td>Beta-actin</td>
<td><strong>B-ACTIN</strong></td>
<td>Reference</td>
<td>Suitable for qPCR validation</td>
<td><em>P. marinus</em> genome</td>
<td>1</td>
<td>Inconsistent results</td>
</tr>
<tr>
<td>Tubulin gamma 1 protein</td>
<td><strong>TUBG1</strong></td>
<td>Reference</td>
<td>Suitable for qPCR validation</td>
<td><em>P. marinus</em> genome</td>
<td>1</td>
<td>Inconsistent results</td>
</tr>
</tbody>
</table>

1. Berbejillo et al., 2012
2. Barrionuevo et al., 2009
3. Matsuzawa-Watanabe et al., 2003
4. Moore et al., 2004
5. Adolfi et al., 2015
6. Koji et al., 1992
7. Kato et al., 2012
8. Kluver et al., 2007
9. Heule et al., 2014
10. Sreenivasan et al., 2008
11. Kim et al., 2015
12. Chen et al., 2011
13. Chen et al., 2011

---

*Danio rerio* genome

*P. marinus* genome
Figure A3.3. Calibration curve for reference gene RPL10a, ribosomal protein L10a. Slope of linear trend line used to determine efficiency, $E = 10^{[-1/(\text{slope})]}$.

Figure A3.4. Calibration curve for WT1, Wilms tumour protein 1. Slope of the linear trend line used to determine efficiency, $E = 10^{[-1/(\text{slope})]}$. 
Figure A3.5. Calibration curve for \textit{DAZAP1}, Deleted in azoospermia-associated protein 1. Slope of trend line used to determine efficiency, $E = 10^{\frac{1}{1/\text{slope}}}$. 

Figure A3.6. Calibration curve of \textit{SOX8}, SOX (SRY-related HMG-box) transcription factor 8. Slope of trend line used to determine efficiency, $E = 10^{\frac{1}{1/\text{slope}}}$.
Figure A3.7. Calibration curve of SOX9, SOX (SRY-related HMG-box) transcription factor 9. Slope of trend line used to determine efficiency, $E = 10^{-\frac{1}{1/(\text{slope})}}$.

Figure A3.8. Calibration curve of GCL, Germ cell-less. Slope of trend line used to determine efficiency, $E = 10^{-\frac{1}{1/(\text{slope})}}$. 
Figure A3.9. Calibration curve of *DMRTA2*, Doublesex- and mab-3-related transcription factor A2. Slope of trend line used to determine efficiency, $E = 10^{-\frac{1}{1/(\text{slope})}}$.

Figure A3.10. Calibration curve of *SF1*, Steroidogenic factor 1. Slope of trend line used to determine efficiency, $E = 10^{-\frac{1}{1/(\text{slope})}}$. 
A3.11. Mean Ct values from gene expression data from candidate genes expressed during testicular differentiation from male and female sea lamprey, *Petromyzon marinus*. Gene expression for candidate genes include *SF1* = steroidogenic factor-1, *DAZAP1* = Deleted in azoospermia associated protein 1, *GCL* = Germ cell-less, *DMRTA2* = Doublesex and mab-3-related transcription factor A2, *SOX8* = SOX (SRY-related HMG-box) transcription factor 8, *SOX9* = SOX (SRY-related HMG-box) transcription factor 9, *WT1* = Wilms tumour protein 1 and reference gene, *RPL10a* = ribosomal protein L10a. Stages of gonadal development for males correspond with Table 3.4, Table 3.5 and Figure 3.8, and mature includes upstream migrating mature adults. Females from the same developmental stage were assigned same stages as males for comparison and to account for genes expressed during similar life stag

<table>
<thead>
<tr>
<th>Genes</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>MAT</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>MAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPL10a</td>
<td>34.07</td>
<td>35.33</td>
<td>34.86</td>
<td>32.25</td>
<td>34.45</td>
<td>32.69</td>
<td>30.65</td>
<td>34.62</td>
<td>29.9</td>
<td>34.17</td>
</tr>
<tr>
<td>SF1</td>
<td>34.39</td>
<td>36.21</td>
<td>37.78</td>
<td>34.94</td>
<td>36.47</td>
<td>36.47</td>
<td>34.07</td>
<td>36.18</td>
<td>34.82</td>
<td>35.32</td>
</tr>
<tr>
<td>DAZAP1</td>
<td>32.35</td>
<td>37.84</td>
<td>39.05</td>
<td>34.3</td>
<td>31.25</td>
<td>33.17</td>
<td>32.47</td>
<td>35.8</td>
<td>34.48</td>
<td>31.09</td>
</tr>
<tr>
<td>GCL</td>
<td>36.98</td>
<td>37.8</td>
<td>36.45</td>
<td>37.95</td>
<td>37.17</td>
<td>38.23</td>
<td>35.89</td>
<td>36.62</td>
<td>36.81</td>
<td>38.54</td>
</tr>
<tr>
<td>DMRTA2</td>
<td>27.42</td>
<td>33.11</td>
<td>36.32</td>
<td>31.37</td>
<td>32.75</td>
<td>31.16</td>
<td>31.31</td>
<td>33.58</td>
<td>30.55</td>
<td>32.96</td>
</tr>
<tr>
<td>SOX8</td>
<td>35.5</td>
<td>34.44</td>
<td>34.19</td>
<td>32.56</td>
<td>34.65</td>
<td>34.95</td>
<td>31.74</td>
<td>34.45</td>
<td>31.69</td>
<td>34.64</td>
</tr>
<tr>
<td>SOX9</td>
<td>26.64</td>
<td>31.83</td>
<td>32.03</td>
<td>31.97</td>
<td>29.55</td>
<td>27.75</td>
<td>31.97</td>
<td>31.61</td>
<td>31.82</td>
<td>29.67</td>
</tr>
<tr>
<td>WT1</td>
<td>32.35</td>
<td>37.04</td>
<td>33.5</td>
<td>33.18</td>
<td>33.55</td>
<td>32.38</td>
<td>31.1</td>
<td>34.4</td>
<td>33.94</td>
<td>33.12</td>
</tr>
</tbody>
</table>
CHAPTER 4: Conclusion and future directions

4.1. Contributions of this study

The genetic basis of sex determination and sex differentiation are reasonably well understood in birds and mammals, but are far less understood in the non-amniotic vertebrates. This is particularly true in lampreys, ancient agnathans found at the pivotal emergence point of vertebrates. Previous studies on lampreys have suggested and explored environmental and genetic-chromosomal modes of sex determination (e.g., Docker, 1992; Docker and Beamish, 1994; Mateus et al., 2012), but results have been inconclusive (see Section 2.2). Focusing on sex differentiation, other studies have begun exploring gene expression during ovarian (but not testicular) differentiation in lampreys (Spice, 2013; Spice et al., 2014; Ajmani 2017). The current thesis has expanded our current understanding of the genetic basis of sex determination and testicular differentiation in lampreys, using the recently sequenced genome of the sea lamprey, *Petromyzon marinus*, as a reference to obtain genes in both processes that have been identified from other vertebrates (Smith et al., 2013).

In Chapter 2, I searched the sea lamprey genome for homologues of 19 candidate sex determination genes and, of those which amplified from genomic DNA, tested for sex specific differences in sea lamprey (11 gene fragments; Table 2.22) and Pacific lamprey, *Entosphenus tridentatus* (7 gene fragments). Six of the candidate genes (*sRY, sDY, GSDF, TRA-1, AMH, AMHR2*) were found to not have homologues in the sea lamprey genome (which was derived from a female sea lamprey; see Section 4.2), and two (*SOX10, SOX17*) had homologues but primers designed for them repeatedly failed to amplify in both lamprey species. This study did, however, identify previously unannotated genes in the sea lamprey genome, and presumptive identities were assigned to *DMRT*-related gene, *FOXL2, SF1/NR5A1*, and *WT1*. The genes
identified in Chapter 2 were also used as potential candidate genes for testicular differentiation and further explored in Chapter 3.

In Chapter 3, I characterized gene expression patterns of seven candidate sex differentiation genes (WT1, DAZAP1, SOX8, SOX9, GCL, DMRTA2 and SF1) prior to and during testicular differentiation and at sexual maturity in male sea lamprey, comparing differences among stages of testicular development and between male and female larval, metamorphosing and adult sea lamprey (see Table 3.3). In lampreys, sex differentiation in females occurs during the larval stage, several years in advance of testicular differentiation, which does not occur until metamorphosis (Hardisty, 1965; Figure 1.1). Histological analyses were also conducted on all testes to determine if up- or down-regulation of certain genes preceded development of specific histological features, in order to infer the role these genes play in male differentiation. Molecular evidence from this study provided support for three Sertoli cell transcription factors: SOX9, DMRTA2 (DMRT-related gene), and WT1; and DAZAP1 as having roles in spermatogenesis and spermiogenesis, as identified from other vertebrates (Nachtigal et al., 1998; Huang et al., 2008; Berbejillo et al., 2012). There was significantly greater expression of Sertoli cell transcription factors in males relative to females, and I was able to identify that high expression of these factors was observed in larval lampreys ranging from 92 to 101 mm. This is before the size at which clear histological signs of testicular differentiation are observed (Docker, 1992; Hardisty, 1965, Hardisty, 1971). An increase in germ cells, oocyte degeneration and the first appearance of pre-spermatogonia was correlated with the significantly higher expression of SOX9, DMRTA2, WT1 and DAZAP1 in late-stage presumptive male larvae (i.e., in male gonad stage 1). The high expression of these genes and factors preceded the first signs of spermatogonia and the dramatic decrease in the size of the gonad and increase in the amount of connective tissue. This suggests
that the still undifferentiated gonad is preparing for testicular differentiation well in advance of metamorphosis, and that sex specific gene expression profiles can potentially identify the future sex of individuals. In European eel, *Anguilla anguilla*, sexually dimorphic gene expression patterns were seen to be early molecular indicators of sex differentiation and were used to estimate sex ratios prior to histological differentiation (Geffroy et al., 2016).

My results are consistent with previous observations that genes involved in sex determination show variable patterns of gene expression among vertebrates, but genes involved in sex differentiation are relatively conserved (Graves, 2013). Sex differentiation genes may be co-opted for sex determination processes as well (Cutting et al., 2013), making it difficult in non-amniotic vertebrates to distinguish between sex determination and sex differentiation genes.

### 4.2. Study limitations

Although no sex specific differences were observed in the candidate sex determination genes surveyed, this does not rule out a genetic mode of sex determination in lampreys. Only gene fragments were sequenced in the current study and, notably, the genome that is available for the sea lamprey was sequenced from a female; male specific candidate sex determination genes (e.g., *sRY, sDY*) would not be expected in this female. Furthermore, considering the variability in sex determining genes observed to date in other fishes, it is possible that other genes entirely are involved in the sex determination process in lampreys (see Section 4.3). Also, although the sea lamprey genome is a great resource for research of this type, it is still a “work in progress.” Not all of the genes in the sea lamprey genome have been annotated and that the annotations on Ensembl are regularly being updated. A constant recurrence during this study was that a gene that was annotated on Ensembl would not be annotated at a later time, would be annotated differently, or the sequence would be unavailable altogether. This was either due to
incorrect identification of the annotated gene, or incorrect sequencing of a target gene. For example, *Slo3* is a sperm specific potassium channel that was believed to have evolved in mammals, but upon further analysis of the presence in other non-mammalian lineages, the gene was present in many other vertebrates, and two novel genes that resemble the mammalian *Slo1* and *Slo3* are present in sea lamprey genome (Vicens et al., 2017). These two sperm specific potassium channel genes were not comparable to other vertebrates due to partial availability of the gene sequences from the sea lamprey genome, but would be important for molecular comparisons and to determine the precise evolution and role of such genes.

In terms of my research to characterize gene expression during testicular differentiation, limitations were related to the long life cycle of the sea lamprey and samples that were not always available from the desired stages. Metamorphosis only happens at one precise time of the year and is synchronized among individuals in a population (Manzon et al., 2015); waiting for samples to reach a specific stage can take months to years. Sea lamprey do not occur in Manitoba and, given their pest status in the Great Lakes, cannot be imported live into Manitoba. Previous work on gene expression related to ovarian differentiation focused on two lamprey species native to Manitoba (chestnut and northern brook lampreys, *Ichthyomyzon castaneus* and *I. fossor*, respectively), but they are relatively rare (greatly limiting sample sizes, particularly during metamorphosis) and sequence differences between these species and the sea lamprey make use of the sea lamprey genome even more problematic (Ajmani, 2017). Sea lamprey thus needed to be collected elsewhere and sent to us already dissected, which also presented limitations as to when sampling could be conducted. This study made use of samples available from collaborators from two sea lamprey populations: the landlocked Great Lakes population and the Atlantic coast anadromous population. There have been differences observed in terms of timing of testicular
differentiation by Hardisty (1969), but my study looked at testicular differentiation from a molecular perspective. Samples were collected with the intention of equally representing the various stages of metamorphosis but, due to the female biased sex ratios, I was unable to obtain a large sample size of males and to get a good representation of males from each stage of development. Future studies should use larger sample sizes from a period of over two seasons to have a good representation from the various stages. This is essential to ensure an outlier does not account for the expression or observation for the entire stage, which can be problematic when working with a small sample size. Ideally, I would also have examined the landlocked and anadromous populations separately, but sample sizes were insufficient. Nevertheless, there were no obvious differences observed in the histological development of testes in landlocked versus anadromous sea lamprey, and pooling samples by stage of development appeared reasonable under the circumstances.

4.3. Future directions

My research in Chapter 2 failed to find sex specific differences in candidate sex determination genes identified from other vertebrates. Although their role in lamprey sex determination cannot be entirely ruled out (see Section 4.2), it is possible—considering the variability in sex determining genes observed to date in other fishes and their phylogenetic position—that other genes are involved in the sex determination process in lampreys. Whole genome sequencing of males and females with accurate annotations will be required to provide conclusive support for the absence or presence of genetic sex determination in lampreys. Since the sea lamprey genome is high in GC repeats (Smith et al., 2013), which can be a problem when using next-generation sequencing techniques, there have been profound advances particularly tackling problems associated with this. For example, the genome of *Streptomyces* sp. PAMC
26508 is composed of greater than 70% GC repeats. Using next-generation Illumina sequencing, large gaps were present in high repeat regions. The promoter regions of over 40 transcription factors that are essential for various developmental processes were successfully analyzed despite being very GC-rich (Long et al., 2011). There were several genes and transcription factors that I was unable to find in the sequenced sea lamprey genome, perhaps due to similar sequencing limitations. Using third-generation or single-molecule real-time sequencing technology (SMRT), greater coverage and resolution of the genome has been achieved in other organisms (Shin et al., 2013). Similar techniques can be employed to re-sequence the sea lamprey genome to cover a greater portion of the genome. Although Illumina sequencing methods provide fast and relatively inexpensive genome sequencing, read lengths of sequences are short, which can hinder the assembly process, especially if GC-repeat rich regions are detected. Therefore, SMRT sequencing is a better alternative for increasing sequence read length and is also better at reading repetitive reads (Burby et al., 2017). Considering that the sequenced sea lamprey genome is that of a female, and used as a reference for designing primers, it is not surprising that gene sequences that are unique and found only in males in some species (sRY, sDY, AMH, AMHR2 and GSDF) were not able to be sequenced. Anti-Müllerian hormone (AMH) is involved in sex determination and differentiation in multiple vertebrates in comparison to sRY/sDY, and therefore of interest. Upon removal of essential markers for male development such as SOX8 and SOX9, AMH continues to cause the regression of Müllerian ducts, highlighting its independent role during testis differentiation (Barriónuevo et al., 2009). If this gene is found in a non-sequence GC rich region, using different techniques could allow for the identification of this gene in lampreys as well.
More research into gene expression during sex differentiation (Chapter 3) is also needed. Genes that show a male-biased pattern during differentiation could be further studied over a greater range of developmental stages to identify if differentiation can be identified from an even earlier stage. However, the limitation of using samples from earlier stages (i.e., at lengths smaller than 90 mm) is that it can be difficult to distinguish the undifferentiated gonad of a future male from a slowly developing ovary (i.e., presumptive males are identified as those individuals which are not female by approximately 90 mm total length; see Figure 1.1). In some fish species (e.g., rainbow trout, *Oncorhynchus mykiss*), monosex populations have been produced, permitting study of gene expression in gonads of known fate long before any histological signs of differentiation (Baron et al., 2005). An addition to this would be conducting a study similar to what was done by Lowartz and Beamish (2000), where gonadal biopsies were taken from sea lamprey larvae for histological analysis and periodically re-sampled. Gonadal biopsy samples could be collected for gene expression and histological analysis, so that gene expression could be monitored in a single individual through the differentiation process.

In addition to gene expression in the gonad itself, another tissue that would be interesting to examine for sex specific differences in gene expression would be brain, which is expected to show changes in expression of genes during sex differentiation (e.g., in the walking catfish, *Clarias batrachus*; Bhat et al., 2016). Metamorphosing sea lamprey samples that were collected had liver, kidney, tail, carcass and brain preserved, and future studies can look at the expression of sex differentiation genes in brain tissue and identify sex specific differences of annotated and novel genes.

My research in Chapter 3 inferred the role genes play in testicular differentiation by correlating level of expression with development of specific histological features. The next step
in understanding the role of these genes would be through gene-knockdown. The SOX E gene family consists of SOX8, 9 and 10, and morpholino knockdown of these genes showed that this family is essential for neural crest development that produces cartilage and branchial basket formation in lampreys (Lakiza et al., 2011). Heath et al. (2014) have shown the effectiveness of RNA interference (RNAi) technology as another method to knockdown genes in sea lamprey embryos and larvae. They showed that RNAi can be applied to the sea lamprey, with higher mortality observed following treatment with short interfering RNAs (siRNAs) of key genes such as actinin, calmodulin and splicing factor. While these genes are not unique to lampreys, this provides a preliminary proof of concept (Heath et al., 2014). Rearing sea lamprey larvae with morpholino or RNAi gene-knockdown of specific male-differentiation genes and observing the effects on the sex ratios of the population would be very interesting. However, due to the lengthy larval period—and particularly the lengthy period of sexual indeterminacy—this would be a difficult undertaking. Nevertheless, if unique species and sex specific genes for knockdown can be identified to target, and with evidence that this treatment is indeed effective, control of invasive sea lamprey in the Great Lakes and restoration of species of conservation concern (e.g., Pacific lamprey) may be possible.

### 4.4. Conclusion

The results of my thesis are consistent with previous observations that genes involved in sex determination are variable among vertebrates, but genes involved in sex differentiation are relatively conserved. No sex specific differences in any candidate sex determination genes were identified, but my research emphasizes the importance of more lamprey genome sequencing (from a number of males and females) and accurate annotations. Sequencing of the sea lamprey genome has been instrumental in research lamprey and vertebrate genomic research (Smith et al.,
2013; McCauley et al., 2015), and additional sequencing taking advantage of new technologies can be used to further improve the genomic resources for lampreys.

Even though my results suggest that expression of some genes are conserved in testicular differentiation among vertebrates, there may be unique lamprey specific genes, and using qPCR alone to screen and detect genes that are conserved may not be sufficient to determine this. Nevertheless, using qPCR, I was able to determine and gain confirmation that similar expression results were obtained from comparable studies looking at the expression of genes such as WTI and DMRT1 in other lamprey species (Spice, 2013; Spice et al., 2014; Mawaribuchi et al., 2017). Certainly, further research is required to assess similar gene expression patterns in other lampreys before any control mechanisms can be established in order to prevent damage to non-target native lampreys. With the increasing cost-effectiveness of sequencing and increase in genomic and transcriptomic studies, identification of sex determination and sex differentiation genes can further our understanding into sex specific genes, and whether genetic sex determination is present in the one of the longest surviving and successful vertebrate lineage.
4.5. Literature cited


