Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Transcriptomic insights into gametogenesis and steroidogenic enzyme gene regulation in Australasian snapper (*Chrysophrys auratus*) gonads during the reproductive cycle

Samuel J. Modern

sam.modern@plantandfood.co.nz

New Zealand Institute for Bioeconomy Science Limited

Matthew J. Wylie

New Zealand Institute for Bioeconomy Science Limited

Igor Ruza

New Zealand Institute for Bioeconomy Science Limited

Craig Radford

University of Auckland

Maren Wellenreuther

New Zealand Institute for Bioeconomy Science Limited

Research Article

Keywords: transcriptome, ovary, testis, bream, aquaculture, aromatase, sparid teleost

Posted Date: June 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9561420/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

1.1. Background

The Australasian snapper (*Chrysophrys auratus*) is a target for aquaculture development in New Zealand to diversify the sector. No studies have explored how gene expression and transcriptomic changes drive gametogenesis in snapper. This study presents the first exploration of transcriptome-wide changes in male and female snapper gonads throughout an entire spawning cycle. Samples of gonadal tissue were taken from 84 captive-bred snapper (42 females and 42 males) throughout their maiden spawning cycle.

1.2. Results

Transcriptomic sequencing reads were supplied to the BRAKER3 genome annotation pipeline to improve the snapper genome annotation, which now contains 26,418 annotated protein-coding genes. Gene expression was quantified using nf-core's rnaseq pipeline. In females, the expression of many key steroidogenic enzyme genes responsible for producing sex steroids increased as oogenesis progressed and peaked during spawning. The notable exception to this trend of expression in females was for cyp19a1a (aromatase), which did not vary significantly between gametogenic stages in the bulk gonad RNA-seq data. Transcriptome-wide differential expression analysis revealed 3,588 upregulated differentially expressed genes (DEGs) in females and 3,938 in males. Functional enrichment analysis revealed that ribosomal processes and metabolism-associated genes were upregulated in female gonads compared to males, while in males cell motility and morphogenesis-related genes were upregulated, and became increasingly upregulated as spermatogenesis progressed, likely reflecting spermiogenic processes.

1.3. Conclusions

The genome annotation provided by this study will aid further studies on snapper and will have utility beyond just reproductive studies, while the sex-biased gene expression revealed by transcriptome profiling may help future studies explore juvenile sexual development and sex inversion.

Highlights

- A total of 26,418 protein-coding genes were annotated in the snapper genome.
- Female expression of *cyp19a1a* remained constant throughout the reproductive cycle.
- 4,093 genes were upregulated in females and 8,474 were upregulated in males.
- Male gonads had the most upregulated genes during early spermatogenesis.
- Female gonads had the most upregulated genes during spawning.

1. Background

Aquaculture species diversification is an urgent global goal, driven by the need to enhance food security, reduce reliance on a limited number of farmed species, and build resilience against environmental and market pressures [1, 2]. In New Zealand, the sparid teleost Australasian snapper (*Chrysophrys auratus,* referred to as tāmure by the indigenous Māori people of New Zealand), hereafter referred to as snapper, is currently a target for domestication and aquaculture development. The species is common around the temperate and sub-tropical coastal waters of Australia and New Zealand, supporting commercial, customary, and recreational fisheries [3-6]. A close sparid relative of snapper, the red seabream (*Pagrus major*), has undergone over six decades of selective breeding efforts that have led to a large and prosperous industry in Japan [7, 8]. While snapper is not currently farmed commercially, efforts by the Plant & Food Research Group, New Zealand Institute for Bioeconomy Science (BSI) to develop the species for aquaculture are ongoing [9, 10]. Compared to the red seabream, snapper reproductive biology has been less intensely studied, especially in captivity.

Studying reproduction and gametogenesis is a vital part of developing new species for aquaculture, as it aids the development of methods to control reproductive processes and the timing of the spawning season [1, 11]. As demonstrated by prior work [12-14], the spawning season of wild snapper typically initiates when water temperatures reach 16° and concludes after the summer peak in temperature when water temperatures drop to 21°C. Mature individuals spawn daily in the wild, with synchronised ovulation within each aggregation peaking around midday [15]. Snapper are functionally gonochoristic, but display rudimentary juvenile hermaphroditism, wherein all juvenile fish initially start developing as females, with ovarian tissue [14, 16]. Following this, males undergo sexual differentiation in which their immature ovaries develop into testes. During this sex inversion the males undergo a period in which they possess both ovarian and testis tissue and their gonads are referred to as ovo-testes [16]. While the reproductive physiology of snapper has been well characterised, no studies have yet investigated gene expression in snapper gonads during the spawning cycle. Studies into how the expression of key genes drives reproductive processes in other species have been ongoing for decades [17]. Modern omics-based techniques have expanded this field, allowing the investigation of gene expression patterns across tissue specific transcriptomes, and can reveal insights which can help in aquaculture [18-22].

Reproductive timing is controlled by an endocrine signalling network involving multiple tissues, known as the brain-pituitary-gonad (BPG) axis. In this axis, the central nervous system integrates both external cues and internal feedback, directing processes related to sexual differentiation and gametogenesis via hormonal signalling [23]. Two important hormones which regulate gametogenic processes in vertebrates are the gonadotropin hormones follicle-stimulating hormone (FSH) and luteinizing hormone (LH), which are secreted by the pituitary into the bloodstream where they act on the gonads, regulating the expression of a multitude of genes [24]. Importantly, gonadotropin action triggers gonadal production of sex steroids such as the main teleost oestrogen, estradiol-17β (E2) and the main teleost androgen, 11-ketotestosterone (11-KT). These sex steroids drive the expression of key genes responsible for promoting sexual differentiation and gametogenesis [24-26]. Sex steroids can also enter the bloodstream and affect other tissues, for example, E2 stimulates the hepatic synthesis of vitellogenin, a vital egg yolk protein [27, 28].

In fish, FSH tends to promote early gametogenesis and vitellogenesis in females, while LH initiates the processes involved in gamete maturation [29]. However, not all fish fall into this generalised model of BPG-axis signalling. For example, in the sister species of snapper, the red seabream, FSH plays a less significant role in E2 production [30]. Also in the red seabream, expression of the FSH beta-subunit gene (*fshb*) in the pituitary was very low in females compared to males and did not vary significantly throughout oogenesis [31]. Taken together, these observations suggest that FSH plays a less important role in female red seabream than in males [31].

In males, synthesis of the sex steroids occurs in Leydig cells, while in females it occurs in the granulosa cells [32, 33]. Sex steroid synthesis is a complex, multifaceted process (Figure 1). The precursor molecule of steroid hormones is cholesterol, which is moved into mitochondria by the steroidogenic acute regulatory protein (StAR) [32]. Once in the mitochondria, the process of steroidogenesis is initiated. This process involves a chain of reactions catalysed by numerous steroidogenic enzymes including cytochrome P450 enzymes from the families 11, 17 and 19, as well as 11β-hydroxysteroid dehydrogenase enzymes, families 3 and 17 [32]. The enzyme cytochrome P450 family 19 subfamily A1A, encoded by the gene *cyp19a1a* catalyses the last step in the synthesis of E2 from testosterone (T) in the gonads [34, 35].

There is a need to close the gap in knowledge of the molecular controls of reproduction in snapper. To achieve this, the present study aims to build on prior work studying snapper reproduction in captivity by applying next-generation sequencing (NGS) techniques to the frozen gonad samples (n = 84) collected previously from maiden spawning (first-time spawners) snapper kept under ambient temperature and photoperiod conditions [14]. We aim to use these samples to investigate transcriptomic patterns across different stages of gametogenesis. First, this study will use these data to improve and expand upon the snapper genome annotation [36, 37]. Next, this work aims to gain insights into gametogenesis and hormonal signalling by characterising the expression dynamics of important genes involved in sex steroid biosynthesis across sexes and between developmental stages of gonads. These patterns will be compared to data previously obtained from snapper and other species. Due to its documented role in regulating the production of E2 and the expression profile seen in the red seabream [38], the gonadal aromatase gene *cyp19a1a* is expected to exhibit expression patterns similar to changes in plasma E2 concentrations reported previously in snapper [14]. Specifically, we expect that the expression of *cyp19a1a* will rise significantly throughout gametogenesis and peak during spawning. This study will also explore broader transcriptomic changes by analysing the wider transcriptome and applying differential expression analysis to investigate sex-biased gene expression and how gene expression varies throughout gametogenesis. Functional gene analysis will be applied to investigate the biological, cellular and molecular processes and pathways associated with these transcriptomic changes.

2. Materials and methods

2.1 Fish sampling

Methods detailing the collection of gonadal tissue samples are outlined by Modern, Ruza [14]. In brief, this study investigates 42 females and 42 males sampled throughout their maiden spawning cycle (Table 1) while maintained in captivity under ambient photoperiod and water temperature conditions. All individuals studied here were sampled by during this previous study, after the period of juvenile sex inversion, and thus could be conclusively sexed. Prior to each sampling event, fish were withheld food for 48 h, then mass-sedated using Aqui-S (8–20 mg/L) (Aqui-S NZ Ltd, Lower Hutt, NZ). Once lightly sedated, individual fish were netted from the tank at random and euthanized by Aqui-S overdose (50 mg/L). The previous work on these samples obtained three gonadal tissue samples from each of these fish. Two of these were subjected to histology and microscopy to stage the reproductive development of each sample based on criteria developed for snapper by Mackie, Jackson [39] and Modern, Ruza [14]. The third portion of gonadal tissue was snap-frozen on dry ice and stored at -80°C for bulk RNA-sequencing and transcriptomic analysis in the present study.

Females were classified into five stages of oogenesis based on the level of development of the most advanced oocyte within their ovaries [14]. These stages are labelled F1 (“immature”, previtellogenic, before maiden spawning; n = 6 fish), F2 (“resting post-spawning”; n = 7 fish), F3 (“developing”; n = 13 fish), F4 (“developed”, vitellogenic; n = 9 fish), and F5 (“spawning”; n = 7 fish). Males were classified into four stages based on the proportions of germ cells at the various spermatogenic stages within their testes (e.g. large aggregations of spermatozoa was indicative of the fourth stage), and whether the animal was releasing milt (spermiating) upon application of abdominal pressure. These stages were labelled M1 (“immature”, before maiden spawning; n = 0 fish due to sample limitations), M2 (“resting post-spawning”; n = 8 fish), M3 (“developing”; n = 16 fish), M4 (“developed”; n = 18 fish).

Table 1. Sampling number, date, mean body weight, and fork length of male and female snapper *(Chrysophrys auratus)* dissected at each time point during the annual reproductive cycle.

Sampling	Date	♀n	♀ Mean body weight (g ± SEM)	♀ Mean fork length (mm ± SEM)	♂n	♂ Mean body weight (g ± SEM)	♂ Mean fork length (mm ± SEM)
1	Dec-19	2	247.0 ± 6.0	317.5 ± 16.8	N/A	N/A	N/A
2	Mar-20	3	348.5 ± 17.0	258.3 ± 5.2	1	518 ± N/A	302 ± N/A
3	Jun-20	6	607.4 ± 22.9	298.0 ± 2.3	6	544.9 ± 21.9	297.7 ± 2.5
4	Jul-20	5	539.1 ± 17.3	295.2 ± 3.1	5	500.4 ± 22.2	278.4 ± 8.4
5	Aug-20	4	559.1 ± 13.1	298.3 ± 3.5	4	516 ± 67.5	292.8 ± 8.4
6	Oct-20	6	595.6 ± 20.2	301 ± 3.8	6	575.8 ± 7.8	301.3 ± 3.0
7	Nov-20	6	599.0 ± 39.7	298.5 ± 4.7	6	592.3 ± 21.1	297.8 ± 4.6
8	Dec-20	7	750.5 ± 44.6	327.1 ± 3.7	7	709.5 ± 37.9	314.9 ± 5.3
9	Feb-21	3	901.6 ± 82.2	338.0 ± 11.8	7	880.4 ± 47.4	343.7 ± 5.9
Total		**42**			**42**		

2.2 RNA extraction

RNA was extracted from gonadal tissue using the NucleoSpin® RNA isolation kit (Machery-Nagel, Düren, Germany) according to the manufacturer's instructions. For quality control (QC), the concentration of the RNA, suspended in RNase-free water, was estimated using a NanoDrop spectrophotometer (DeNovix® model DS-11 FX). Signs of RNA-degradation were investigated by checking for the presence of 18S and 28S rRNA-lines following gel electrophoresis. While some samples were subjected to repeat DNase treatment based on these QC results, no samples were removed from consideration at this stage.

2.3 Library preparation and RNA-sequencing

Further QC of the isolated RNA, complementary DNA (cDNA) library preparation, and sequencing were performed by Australian Genome Research Facility (AGRF) Ltd in Melbourne, Australia using their standard methodologies which involves calculating the RNA Integrity Number (RIN) for each sample. AGRF prepared cDNA libraries by reverse-transcribing the RNA, which were then sequenced to generate

150 base-pair (bp) paired-end reads in fastq format on the Illumina NovaSeq platform. Once generated, the reads from the tissue samples were transferred to BSI's computer cluster.

2.4 Genome annotation using BRAKER3, Liftoff and OrthoFinder

De-novo annotation of the sequences was undertaken using BRAKER3 version 3.0.8 with the BUSCO lineage set to Actinopterygii and using defaults for all other settings. BRAKER3 is a pipeline which uses GeneMark and AUGUSTUS to create high-quality gene predictions which are supported by both intrinsic evidence (the content of the genome itself) and extrinsic evidence (transcriptome and proteome data) [40-56]. BRAKER3's main inputs are the genome, RNA-sequencing data, and a database of proteins from related species. The RNA-sequencing data supplied to BRAKER in the current study was obtained from brain, pituitary and gonad samples from one male and female individual and was subjected to QC and trimmed using the same methodology as will be described in the next subsection of this paper. The protein evidence supplied to BRAKER was the OrthoDB Vertebrata partition created by the creators of BRAKER [57].

In addition to BRAKER, Liftoff version 1.6.3 [58] was used to map gene annotations from the yellowfin seabream (*Acanthopagrus latus)* [59] to the snapper genome [36, 37]. Any yellowfin seabream gene detected in the genome which was not found in the BRAKER geneset was also added to the annotation. This was done to allow for the detection of any genes which may have been missed by BRAKER.

In combining the outputs of BRAKER and Liftoff, where a BRAKER gene overlapped with a liftoff gene, the BRAKER gene was kept because the BRAKER gene structure predictions are supported by extrinsic evidence and are more reliable than Liftoff gene structures. Next, the annotated coding sequence (CDS) regions were extracted and translated from each protein-coding gene (translations which introduced mid-sequence stop-codons were evidence of incorrect CDS reconstruction, and thus such genes were filtered out). These translated protein sequences, and the proteomes of yellowfin seabream, goldfish (*Carassius auratus*)*,* Atlantic cod (*Gadus morhua*)*,* three-spined stickleback (*Gasterosteus aculeatus aculeatus*)*,* medaka (*Oryzias latipes*), Atlantic salmon (*Salmo salar*) and zebrafish (*Danio rerio*) from the vertebrata OrthoDB partition were used as input to the program OrthoFinder version 2.5.5 [60, 61]. From these proteomes, OrthoFinder identified orthogroups of homologous genes across the species. For each identified orthogroup, OrthoFinder constructed rooted phylogenetic trees containing each gene from all species in the orthogroup. These phylogenetic trees were then used to determine the closest homologous gene in any of the species included in the analysis. The analysis in the current study used the zebrafish genome annotations as a reference, given the highly expansive and curated nature of the genome of this model species and also uses the zebrafish gene naming conventions and gene IDs to distinguish between genes.

2.5 QC, read trimming and quantification of gene expression

Each raw read file was subjected to QC via FastQC version 0.12.1. This revealed that the phred quality score of the first base on the 5' end was below 20 in every sample, so this base was removed in later processing steps. QC, adapter trimming, and transcript-level alignment and quantification were all done using nf-core's rnaseq pipeline version 3.17.0 [62] using nextflow version 24.04.4 [63] (though the gene-level counts were analysed in this study). This pipeline performed QC, trimming, mapping and gene counts quantification using a variety of tools including STAR version 2.7.11 [64] and Salmon version 1.10.1 [65]. The pipeline was supplied with the snapper genome [36, 37], annotation and a samplesheet containing the location of the RNAseq reads and read strandedness. Because the snapper genome assembly was supplied to the rnaseq pipeline using the "--fasta" flag, the pipeline ran the Salmon transcript quantifier's selective alignment algorithm using the entire genome as the decoy-aware transcriptome when creating the Salmon index (see rnaseq documentation for details [62]). As well the flag "--skip_biotype_qc", was used to instruct the pipeline to skip the additional featureCounts process for biotype QC (done to prevent errors due to a lack of the biotype field in the BRAKER gene annotations). Finally, "--extra_trimgalore_args", to instruct Trim Galore! to remove of the aforementioned low-quality bases at the 5' end which were not adequately removed using default settings (using the Trim Galore! arguments: "--clip_r1=1 --clip_r2=1"). The full command used to run the pipeline can be found in a repository listed in the data availability statement.

2.6 Differential gene expression analysis

In addition to investigating the role of key steroidogenic enzymes, differential expression analysis was performed across the entire transcriptome using the R package DESeq2 version 1.46.0 [66]. The raw gene counts from Salmon were then normalised using DESeq's median of ratios method. Prefiltering to keep only genes that had sufficiently large counts was performed (using the pre-normalised data) using the filterByExpr function in EdgeR version 4.4.2 [67] with the arguments "min.count" and "min.total.count" set to 10 and 15, respectively, which are the standard settings [68].

Following normalisation and prefiltering, a Wald test was performed to reveal differential expression between males and females. Further Wald tests were performed on differential expression between the different stages of gametogenesis in males and females separately (given that the male and female stages are not directly comparable). Following each stage-based Wald test, pairwise comparisons were used to test for differential expression between specific combinations of stages. For all tests, a significance threshold for p-values (adjusted for false-discovery rate using the Benjamini-Hochberg method [69]) of < 0.05 was used, and a log2 fold change of < -1 or > 1 was used to filter data for significance differential expression.

2.7 Functional gene analysis

The eggNOG mapper version 2.1.8, in the Galaxy community platform [47, 70-72], was used to annotate the genes predicted in the snapper annotation. Functional analysis using the Gene Ontology (GO) [73, 74] and Kyoto Encyclopedia of Genes and Genomes (KEGG) [75-77] databases were used for functional gene analysis. GO and KEGG gene set enrichment analyses (GSEA) were performed using the R package clusterProfiler version 4.14.4 [78]. For these tests, a gene universe comprising all the annotations for the genes found expressed in the snapper gonads (i.e. those not filtered out due to low expression) was used, as recommended for functional enrichment analyses [79]. In the GSEA analysis, genes were ranked using their Wald test statistic for each contrast. The Benjamini-Hochberg procedure to adjust the p-values for false discovery rate (FDR) was used, and a threshold for adjusted p-values of < 0.05 was used to determine significant upregulation and enrichment of GO terms or KEGG pathways.

2.8 Hormonal signalling associated gene expression

Steroidogenic genes involved in the biosynthesis of E2 and 11-KT were individually studied, as were genes associated with hormone reception. The steroidogenic enzyme genes selected for study were: steroidogenic acute regulatory protein (*star*); cytochrome P450 family 11 subfamily A member 1, tandem duplicate 2 (*cyp11a1.2*); cytochrome P450 family 17 subfamily A member 1 (*cyp17a1*); hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 1 (*hsd3b1*); gonad aromatase (*cyp19a1a*); forkhead box L2a (*foxl2a*); cytochrome P450 family 11 subfamily B member 1 (*cyp11b1*); and hydroxysteroid (11-beta) dehydrogenase 2 (*hsd11b2*) [32]. The hormone receptor genes selected were follicle-stimulating hormone receptor (*fshr*); luteinizing hormone/choriogonadotropin receptor (*lhcgr*); androgen receptor (*ar*); oestrogen receptor 1 (*esr1*); and oestrogen receptor 2a (*esr2a*).

For this key gene analysis, the gene-length scaled gene counts from Salmon were used as opposed to the raw counts to allow for comparisons of expression between different genes. This length-scaled gene expression was then normalised using the same procedure as used to normalise data during the DESeq analysis. The gene expression data were plotted using boxplots for each gametogenic stage in the order in which each gametogenic stage was seen during sampling (i.e. with F2 and M2 last). Statistical tests were also applied to this data. Parametric tests were used if data satisfied or could be transformed to satisfy the assumptions of parametric testing as verified by Shapiro-Wilk tests and residual QQ plots for normality, and Levene tests and residual vs fit plots for homogeneity of variances. The parametric test used was an analysis of variance (ANOVA), with post-hoc Tukey honestly significant difference (HSD) tests applied following a statistically significant ANOVA to determine significance in differences between groups. If the assumptions of parametric testing were not met, then a Kruskal-Wallis test was used, with post-hoc Dunn tests following a significant result to assess statistical significance between groups. For all tests, a threshold of $p < 0.05$ to determine statistical significance was applied. Since this analysis was concerned with specific genes rather than expression patterns in the wider transcriptome, transcriptome-wide p-value correction for multiple testing was not performed. However, Benjamini/Hochberg p-value

correction was performed over the results from the 28 statistical tests performed in both sexes over the panel of 14 genes to control the false discovery rate at 5%. Thus, the probability of obtaining a type I error remains 0.05 even across all our tests. Furthermore, Benjamini/Hochberg p-value correction is incorporated into the post-hoc testing functions to control for the multiple pairwise testing between the different groups. For these analyses, R version 4.2.2 and Python version 3.12.6 was used with the packages SciPy version 1.14.1 [80] and Statsmodels version 0.14.4 [81].

3. Results

3.1. RNA-extractions, sequencing yield and gene mapping

All samples had an RNA-integrity number (RIN) of greater than six (Table S1). Numerous samples had abnormally high 5S rRNA areas, which is to be expected from ovarian tissue samples [82], consequently no samples were removed from the analysis due to this. RNA sequencing of the 104 gonad samples yielded a total of 4.2 billion reads with an average GC content of 49% and an average sequencing error rate of 0.68%. From these reads, 3.8 billion were aligned to the reference genome, with a per-sample average unique alignment rate of 77%. Sample number 90 was removed from further analysis due to a sampling error in which non-gonadal tissue was collected instead of the gonad, as evidenced by the lack of gonadal tissue in histological micrographs [14]. Furthermore, no data for sample 55 were returned by AGRF, so this sample was also not considered in further analysis. This left 41 males and 41 females for subsequent analysis. More statistics including the full MultiQC report produced by the rnaseq pipeline can be accessed via the repository linked in the data availability statement.

3.2. Genome annotation

In the BRAKER3 annotation, BUSCO analysis found 3,564 complete BUSCOs (97.8%) in the annotation, with 46 fragmented and 32 missing BUSCOs. The final annotation file contained a total of 26,418 protein-coding genes. The gene predictions yielded 22,836 genes with CDS reconstructions which lacked mid-sequence stop-codons. From these reconstructed protein sequences, Orthofinder was able to assign 21,645 of the genes to orthogroups, of which 18,248 shared an orthogroup with at least one zebrafish gene. This allowed the identification of key genes, including genes involved in the production of steroidogenic sex hormones, which were all identified using the zebrafish orthologous genome annotations as determined via the gene trees constructed by Orthofinder [32].

3.3. Differentially expressed genes between sexes and gametogenic stages

PCA revealed that gonadal gene expression clustered into distinct male and female clusters (Figure S1). There were 3,588 differentially expressed genes (DEGs) upregulated in females, and 3,938 DEGs upregulated in males.

The number of DEGs upregulated in each stage of oogenesis (i.e. “set size”) tended to increase as oogenesis progressed from the early stages to spawning (F5) (Fig. 2a). The two early gametogenic stages had the lowest number of upregulated DEGs. Immature (F1) had 176 upregulated DEGs (when considering all of the pairwise comparisons which included this stage) while Resting post-spawning (F2) ovaries had 106 upregulated DEGs. Developing ovaries (F3) had 280 upregulated DEGs, developed (F4, vitellogenic) had 897, while spawning stage (F5) ovaries had 1466 upregulated DEGs, 997 of which were unique to this stage (Fig. 2a). These data reveal a general trend of a rising number of upregulated genes as oogenesis progressed.

In males, the number of DEGs in each stage did not vary much between the stages. The M2 stage had 2,671 upregulated DEGs, 1,976 of which were unique to this stage (Fig. 2b). Developing (M3) testes had the fewest upregulated DEGs, with 2,490. Developed (M4) testes had 2,684 upregulated DEGs. There were 1,761 DEGs which were common to both M3 and M4, which is far greater than the number of DEGs which were common to M2 and M3 (682) and M2 and M4 (3). M3 had the fewest unique DEGs, with only 37.

Figure 2. UpSet plots of the number of differentially expressed genes (DEGs) in the ovary at each stage of oogenesis in females (a.) and in the testis at each stage of spermatogenesis in males (b.). Each set represents the number of DEGs associated with each gametogenic stage from all of the pairwise comparisons of each stage. Horizontal bars indicate the total number of DEGs for each comparison, while vertical bars represent the size of intersections (shared DEGs) across multiple comparisons, as indicated by the connected dots below.

3.4. Functional gene analysis

Gene-set enrichment (GSE) revealed that the top significantly enriched GO terms and KEGG pathways in females tended to be associated with metabolism and ribosomal processes, whereas the top male terms and pathways tended to be associated with morphogenesis and cell motility (Fig. 3). In both males and females there were three different contrasts between stages which yielded significant GO terms or KEGG pathways (Table 2). In males these three contrasts were the only three contrasts which were tested (M2 vs M3, M2 vs M4, M3 vs M4). However, due to the higher number of gametogenic stages in females, there were 10 contrasts tested in females, meaning that seven of these contrasts yielded no significant GO terms or KEGG pathways.

Table 2

The top significant GO terms and KEGG pathways for each contrast of gametogenic stages. For each contrast the two most highly suppressed (lowest negative enrichment scores) and two most highly activated (highest positive enrichment scores) for both GO terms and KEGG pathways are listed. A negative enrichment score means that the term is upregulated significantly with the first stage of the contrast, while a positive enrichment score term is associated with the second stage in the contrast. Note: not all contrasts had significant enrichments.

Contrast	ID	n genes	Enrichment	Description
F1 vs F3	GO:0032989	10	0.60	cellular anatomical entity morphogenesis
F1 vs F3	GO:0048513	10	0.58	animal organ development
F2 vs F5	dre04820	57	-0.37	Cytoskeleton in muscle cells
F3 vs F5	dre04810	22	-0.46	Regulation of actin cytoskeleton
F3 vs F5	dre04510	32	-0.39	Focal adhesion
F3 vs F5	dre04820	41	-0.36	Cytoskeleton in muscle cells
M2 vs M3	GO:0006334	11	-0.69	nucleosome assembly
M2 vs M3	GO:0005871	12	-0.66	kinesin complex
M2 vs M3	GO:0050864	13	0.58	regulation of B cell activation
M2 vs M3	GO:0050871	11	0.57	positive regulation of B cell activation
M2 vs M3	dre00020	10	-0.61	Citrate cycle (TCA cycle)
M2 vs M3	dre03013	19	-0.48	Nucleocytoplasmic transport
M2 vs M3	dre04330	14	0.62	Notch signaling pathway
M2 vs M3	dre00310	10	0.59	Lysine degradation
M2 vs M4	GO:0030990	16	-0.79	intraciliary transport particle
M2 vs M4	GO:0030992	13	-0.79	intraciliary transport particle B
M2 vs M4	GO:0071526	12	0.67	semaphorin-plexin signaling pathway
M2 vs M4	GO:0034340	11	0.64	response to type I interferon
M2 vs M4	dre03013	30	-0.51	Nucleocytoplasmic transport
M2 vs M4	dre03015	33	-0.48	mRNA surveillance pathway
M2 vs M4	dre02010	10	0.57	ABC transporters
M2 vs M4	dre04060	28	0.51	Cytokine-cytokine receptor interaction
M3 vs M4	GO:0060541	14	0.63	respiratory system development
M3 vs M4	dre04514	10	0.59	Cell adhesion molecules

There was a large difference in the number of significant terms and pathways between males and females. The female contrast F1 vs F3 had seven significant GO terms and no significant KEGG pathways (Table S2). F2 vs F5 only had one significant KEGG pathway, while F3 vs F5 had only three significant KEGG pathways. Meanwhile in males, M2 vs M3 had 285 significant GO terms and 14 significant KEGG pathways. M2 vs M3 had 447 significant GO terms and 25 significant KEGG pathways, and M3 vs M4 only had one significant GO term and one significant KEGG pathway.

The immature female stage (F1) was associated with two morphogenesis-related GO terms and one related to organ development (Table 2). Multiple terms associated with the spawning stage (F5) mention cytoskeleton. In males there were numerous immune system-related terms associated with the immature stage (M2), while the cell motility associated terms which were seen upregulated in males as a whole were mainly associated with M4.

3.5. Expression of genes involved in production of steroidogenic sex hormones

In females, expression of the genes encoding enzymes involved in the synthesis of testosterone (*star; cyp11a1.2*, *cyp17a1*, *hsd3b1*; and hydroxysteroid (17-beta) dehydrogenase 1, *hsd17b1*) all increased significantly as gametogenesis progressed (Fig. 4a). However, the gene encoding the gonad aromatase enzyme, cytochrome P450, family 19, subfamily A, polypeptide 1a (*cyp19a1a*), which catalyses the reactions converting T or estrone (E1) to E2, did not vary between any stages of oogenesis (Fig. 5a). Despite this lack of significant variability, *cyp19a1a* was expressed at a comparatively high degree in the female gonads, with a mean normalised expression (141.7) higher than that of the other steroidogenic enzymes *cyp11a1.2* (53.3), *cyp17a1* (105.5), *hsd3b1* (2.7) and *hsd17b1* (21.3). Both enzymes responsible for the synthesis of 11-KT were expressed at similarly low levels in females, with average normalised expressions of 0.4 for *cyp11a1.2*, and 1.9 for *hsd11b2* (Fig. 5a). The only gene which had statistically different expression between the immature and resting post-spawning stages (F1 and F2) was cytochrome P450 family 11 subfamily B member 1 (*cyp11b1*), which was upregulated in immature females but showed very little expression in any other stage.

In males, the genes involved in synthesis of T had a less regular pattern of expression. The genes *star*, *cyp11a1.2* and *hsd3b1* were upregulated in stage M3, while *cyp17a1* and *hsd17b1* decreased from “resting post-spawning” (M2) to “developed” (M4) (Fig. 4b). The gene *cyp11b1*, involved in synthesis of 11-KT, was significantly downregulated in stage M4, but expression of this gene was high compared to most of the other enzymes throughout spermatogenesis (Fig. 5b). Meanwhile, *hsd11b2*, also involved in 11-KT synthesis, was significantly upregulated in the “resting post-spawning” stage (M2), and was expressed at a somewhat low degree, with an average expression of 9.0.

3.6. Expression of hormone receptor genes

The FSH receptor, *fshr*, was highly expressed in males. The mean normalised expression of *fshr* across all male stages was 184.6. This contrasts with females, where expression of the gene averaged across

all gametogenic stages was only 31.4. Expression of *fshr* was highest in the resting post-spawning stage (M2) in males, while in females it was slightly upregulated during the spawning stage (F5) (Fig. 6). Expression of *lhcgr* and *esr1* both increased in females throughout the spawning cycle and peaked during spawning (F5) before falling back to levels similar to those in the immature pre-spawning stage (F1). Expression of *lhcgr* displayed a different pattern in males as it was downregulated in late spermatogenesis (developed stage, M4) compared to early spermatogenesis (developing stage M3 and the resting post-spawning M2 stage). The *ar* and *esr2a* genes also displayed a similar pattern in males but neither varied significantly in females throughout gametogenesis, though the androgen receptor was expressed to a higher degree in females overall, with an average expression of 134.9 in females and 9.0 in males.

4. Discussion

Broodstock management for aquaculture relies on a thorough understanding of reproductive processes in captivity [11, 83]. To expand knowledge of the molecular controls of gametogenesis and the reproductive cycle in snapper, a target species for aquaculture development, we have improved the snapper genome annotation and investigated the transcriptomic profile during gametogenesis in male and female gonads. To annotate the genome, we leveraged RNA-sequencing data, combining this with protein sequences and annotations from related species to annotate over 26,000 protein-coding genes in the snapper genome. These annotations will aid further studies aiming to utilise modern genomic or transcriptomic techniques on snapper. This study also laid the groundwork for investigations into the molecular controls of sex determination and sexual differentiation in hatchery-reared snapper undergoing their maiden spawning reproductive cycle in captivity by highlighting sex-biased differential gene expression and important male- and female-specific genes. Further work may be able to expand upon this sex-biased gene set to develop a transcriptomic profile which can differentiate between sexes in snapper before the period of juvenile sex inversion [16], allowing further investigations to develop a greater understanding of juvenile sex inversion and sex determination in the species.

A greater number of upregulated DEGs were found in male testes than were found in female ovaries, which is a trend commonly seen in both teleosts and other vertebrates [84–90]. Future studies may be able to use the sex-biased expression gene set identified here to develop a transcriptomic profile able to distinguish the sex of undifferentiated juvenile individuals and ideally aid the development of a non-lethal molecular technique to determine the sex of juveniles. The smaller statistical power of the pairwise comparisons between stages in females (because there were smaller sample sizes in each pairwise comparison in females due to the larger number of female gametogenic stages) may be partly responsible for the smaller number of DEGs found in the pairwise comparisons between female stages of gametogenesis compared to males. This could in turn be responsible for the fewer GO terms or KEGG pathways which were found in the GSEA between stages in females.

The upset plots highlighted cases where two stages shared an unusually high degree of overlapping DEGs. The developing (M3) and developed (M4) stages were the two most similar stages in males, as

evidenced by the large number of upregulated genes shared between these stages, which mirrors the morphological similarities between these two stages of gametogenesis [14]. Functional enrichment analysis of sex-biased DEGs highlighted the large number of motor-proteins and cell motility GO terms upregulated during these two stages. In mammals, motor-proteins are also known to be crucial in supporting spermatogenesis via intra-cellular transport [91]. The process of spermiogenesis, which is very active during developing and developed (M3 and M4) stage testes [14], may also explain why motor protein and morphogenesis-related terms were upregulated in these stages. Indeed, a GO term which was upregulated in both stages compared to the post-spawning stage (M2) was the term “axoneme”; a vital component of the sperm flagellum [92]. Comparing the DEGs upregulated in stage M4 to the recently determined structure of the sperm axoneme of mice (*Mus musculus*) [93] reveals that all but two of the 39 proteins identified as part of the sperm axoneme were identified as upregulated DEGs in developed males (M4) (with seven more not having been found in the snapper genome). Two morphogenesis-related terms were also upregulated in the immature female stage (F1). Given there was also a term related to organ development upregulated in this stage, it could be that these terms may reflect the ongoing development in the ovaries as the fish prepare for maiden spawning.

Males expressed *fshr* to a much higher degree than females. FSH signalling is known to play a less significant role in gametogenesis in red seabream females than it does in males [38]. Sex-biased gene expression of *fshr* was also seen in the honeycomb grouper (*Epinephelus merra*) [94], a protogynous species in which FSH induce female-to-male sex change [95]. The upregulation of the FSH receptor in snapper indicates that further research could find that FSH also plays male-specific roles in gametogenesis or sexual development in snapper, though this cannot be concluded from the present results, in part because the FSH receptor is known to be cross-reactive and can bind to LH as well as FSH [96].

The expression of most steroidogenic enzyme genes in females followed a similar trend, increasing as oogenesis progressed and peaking during spawning (F5). Additionally, numerous important hormone receptor genes, including the *lhcgr*, *fshr* and *esr1* were identified as significantly upregulated in spawning females (F5), highlighting that this is a period of intense hormone activity in females. The notable exception to this trend was the gonad aromatase gene, *cyp19a1a*. In teleosts, this gene is essential for E2 production and promoting and regulating oogenesis [97, 98]. In other species, the expression of *cyp19a1a* generally rises throughout early oogenesis, usually peaking in the vitellogenic stage shortly before spawning, or just after the onset of spawning [18, 20, 98–102]. Thus, it was unexpected that *cyp19a1a* expression did not vary significantly between stages in female snapper, especially given the significant upregulation of *foxl2a* in stage F5, a transcription factor which has been found to upregulate *cyp19a1a* in other species [103, 104].

This lack of variation in *cyp19a1a* expression contrasts with previously reported E2 plasma hormone levels in snapper which did vary, peaking significantly during spawning [14]. This suggests that production of E2 throughout the spawning cycle in the snapper studied here may not be controlled by variation in the expression of the *cyp19a1a* gene. Generally, T is considered as the main precursor of E2

[32]. However, in the red seabream, E2 is primarily synthesised from E1 in a reaction catalysed by *hsd17b1*, though *cyp19a1a* is still required to catalyse the reaction which converts androstenedione to E1 [105]. In red seabream ovarian fragments, *cyp19a1a* expression increased in response to the *in vivo* administration of LH [30] and an agonist of gonadotropin-releasing hormone (GnRHa) in slow-release pellets *in vivo* [106] - and expression of the *cyp19a1a* gene varied significantly throughout the spawning cycle in red seabream [38]. Our data show a steady expression of *cyp19a1a* throughout oogenesis and thus hint at differing regulatory mechanisms in the control E2 synthesis in snapper compared to the red seabream. Instead of being regulated by expression of *cyp19a1a*, in the fish studied here, E2 synthesis may have been regulated by the availability of the precursor of E2, be it E1, as in the red seabream, or T. If E2 is mainly synthesised from E1, the rate-limiting step could be the synthesis of E1 from androstenedione, which may be limited by the expression of *hsd17b1*, which encodes the enzyme responsible for catalysing the conversion of E1 to E2 [32]. This is supported by the expression patterns of *hsd17b1* expression, which were similar to changes in the plasma E2 levels seen in snapper [14]. Alternatively, the rate of E2 synthesis may be controlled by the availability and activity of the aromatase enzyme itself, rather than the expression of its gene, though the lack of variation in the aromatase gene expression seen here still contrasts with that seen in other species. It should be noted that the bulk-sequencing of ovarian tissue we performed here means that we cannot determine which cells were expressing *cyp19a1a* at any given time. It could be that the stable bulk expression seen here was masking shifts in the expression of the gene occurring in individual ovarian follicles at different stages of development. Regardless, the lack of variability is still intriguing and contrasts the significant changes in expression of the gene seen in red seabream ovaries [38].

The highly asynchronous mode of development in snapper spermatogenesis, wherein germ cells at most or all developmental stages are present in the testis at all times, makes sequencing the entire testis transcriptome, as done in this study, a relatively poor way to elucidate the molecular processes which govern the progression of spermatogenesis at each specific cell stage [107]. Indeed, the large degree of overlap in the DEGs between developing (M3) and developed (M4) testes indicates that little separates these two stages besides the spermiating status of the fish (release of milt upon application of abdominal pressure). Snapper females also present us with this issue due to their asynchronous mode of oocyte development, though the issue is less pronounced than in males, as advanced-stage germ cells such as vitellogenic or hydrated oocytes do not appear until later in the spawning cycle. Future studies on snapper could use different methods to those used here to obtain more informative results. For females, studies on other fish with asynchronous oocyte development have separated out ovarian follicles of different stages using sieves [108], or manually by using forceps [109]. Separating germ cells by stage in such ways for males is more challenging due to the small size of male germ cells. Therefore, future studies may instead consider single-cell RNA-sequencing techniques instead [110, 111]. Single-cell methods could also be applied to ovaries to distinguish gene expression in oocytes and the important somatic follicle cells surrounding the oocytes.

The resting post-spawning period (M2) is characterised by intense spermatogonial activity and is sometimes referred to as the “mitosis-dominated” period of spermatogenesis where spermatogonia

cells renew and start to differentiate [39, 107]. his M2 period displayed a similar number of upregulated DEGs compared to the later stages of spermatogenesis in males. This hints at a sustained level of activity in the testis throughout gametogenesis and contrasts with the early gametogenesis stages in females which displayed far lower number of upregulated DEGs compared to later stages. Instead, there was a general trend of increasing numbers of upregulated DEGs associated with each stage in females, hinting at a growing complexity in the processes within the ovaries as oogenesis progressed, with the number of DEGs peaking during the spawning season (F5). Similar trends of gene expression in the ovaries have also been seen in other species including zebrafish, spotted scat (*Scatophagus argus*) and greater amberjack (*Seriola dumerili*) [18, 19, 21].

5. Conclusion

This study presents the first exploration of transcriptome-wide changes in snapper gonads throughout their maiden spawning cycle. Consistent with findings in other species, large differences between male and female gene expression were seen [85–87, 112–114]. The number of differentially expressed genes in female gonads increased as oogenesis progressed, while in male gonads the most differentially expressed genes were found in the early developing testis. Functional analysis of differentially expressed genes was confounded in females due to the lack of any significantly activated GO or KEGG terms between oogenic stages. Expression patterns of key genes related to hormonal signalling and sex steroid biosynthesis revealed that in females, the expression of most steroidogenic enzymes involved in the production of E2 increased as oogenesis progressed, peaking during spawning, as did receptor genes for FSH, LH and E2. In contrast to this trend, and *cyp19a1a* expression from the red seabream [38], we did not find significant differences in the expression of *cyp19a1a* in snapper. By characterising the transcriptomic differences between the sexes and stages of gametogenesis, and by revealing important pathways using functional enrichment analysis, this study lays the groundwork for future studies developing an understanding of, and possibly methods to control, captive reproduction and sexual differentiation. Such studies could then explore how sex-biased gene expression during the sexual differentiation period might drive sex determination in snapper, which may also unlock the ability to control the sex of snapper broodstock for aquaculture.

Declarations

Ethics

Animal handling and manipulations were approved and conducted in accordance with the guidelines of the Nelson Marlborough Institute of Technology (NMIT) Animal Ethics Committee (Approval: AEC2020-PFR-01).

Consent for publication

All authors have consented to this publication.

Availability of data and materials

Snapper represents a taonga (cultural treasure) species in the indigenous Māori culture of Aoteroa New Zealand. We must adhere to the Guidelines for Genomic Research on Taonga Species. The result of this is that we must use a repository endorsed by these guidelines. Hence, our data has been uploaded to the Genomics Aoteroa repository, which ensures indigenous control and sovereignty over the data and that uses of the data benefit indigenous communities and aligns with values, applications, traditions and roles of these communities (Hudson et al., 2020; Jennings et al., 2023). The datasets generated and analysed during the current study are available upon request, including to reviewers, in the Genomics Aotearoa repository at the following digital object identifier link: https://doi.org/10.57748/9dc3-3b50

Competing interests

The authors declare no competing interests.

Funding

This study was funded by the New Zealand Government via the Ngā Pou Rangahau platform “Species Assessment for Open Ocean Aquaculture”, a research framework developed by BSI and supported by a Strategic Science Investment Fund grant from the Ministry of Business, Innovation and Employment.

Author’s contributions

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.	
Transcriptomic insights into gametogenesis and steroidogenic enzyme gene regulation in Australasian snapper (Chrysophrys auratus) gonads during the reproductive cycle Samuel J. Modern, Matthew J. Wylie, Igor Ruza, Craig A. Radford, Maren Wellenreuther	
Nature of contribution by PhD candidate	Lead author, principal investigator and data analyst
Extent of contribution by PhD candidate (%)	80%

CO-AUTHORS

Name	Nature of Contribution
Matthew J. Wylie	Sampling, supervision, reviewing, editing, conceptualisation
Igor Ruza	Sampling, RNA-extractions, reviewing
Craig A. Radford	Supervision, reviewing, editing
Maren Wellenreuther	Supervision, reviewing, editing

Acknowledgements

We thank the technical staff at Nelson Research Centre Finfish Facility for taking care of the fish and for assistance with sampling. We thank Julie Blommaertt, Jason Shiller, Charles David and Marcus Davy for their feedback on the final versions of this manuscript. This study was funded by the New Zealand Government via the Ngā Pou Rangahau platform “Species Assessment for Open Ocean Aquaculture”, a research framework developed by the Bioeconomy Science Institute (formerly Plant & Food Research) and supported by a Strategic Science Investment Fund grant from the Ministry of Business, Innovation and Employment

References

1. Camara, M.D. and J.E. Symonds. (2014). Genetic improvement of New Zealand aquaculture species: programmes, progress and prospects. New Zealand Journal of Marine and Freshwater Research, 48(3), 466-491. https://doi.org/10.1080/00288330.2014.932291
2. Cai, J., H.L. Chan, X. Yan, and P. Leung. (2023). A global assessment of species diversification in aquaculture. Aquaculture, 576, 739837. https://doi.org/10.1016/j.aquaculture.2023.739837
3. Leach, F. and J. Davidson. (2000). Pre-European Catches of Snapper (*Pagrus auratus*) in Northern New Zealand. Journal of Archaeological Science, 27(6), 509-522.

https://doi.org/10.1006/jasc.1999.0474

4. Parsons, D.M., M.A. Morrison, A.B. MacDiarmid, B. Stirling, P. Cleaver, I.W.G. Smith, and M. Butcher. (2009). Risks of shifting baselines highlighted by anecdotal accounts of New Zealand's snapper (*Pagrus auratus*) fishery. New Zealand Journal of Marine and Freshwater Research, 43(4), 965-983. https://doi.org/10.1080/00288330909510054
5. Bertram, A., J. Bell, C.J. Brauer, A. Fowler, P. Hamer, J. Sandoval-Castillo, J. Stewart, M. Wellenreuther, and L.B. Beheregaray. (2023). Biogeographic provinces and genomically delineated stocks are congruent in snapper (*Chrysophrys auratus*) from southeastern Australia. ICES Journal of Marine Science, 80(5), 1422-1430. https://doi.org/10.1093/icesjms/fsad068
6. Bertram, A., J. Bell, C. Brauer, D. Fairclough, P. Hamer, J. Sandoval-Castillo, M. Wellenreuther, and L.B. Beheregaray. (2024). Estimation of effective number of breeders and effective population size in an abundant and heavily exploited marine teleost. Evolutionary Applications, 17(7), e13758-n/a. https://doi.org/10.1111/eva.13758
7. Foscarini, R. (1988). A review: Intensive farming procedure for red sea bream (*Pagrus major*) in Japan. Aquaculture, 72(3), 191-246. https://doi.org/10.1016/0044-8486(88)90212-8
8. Kato, K. (2023). Breeding studies on red sea bream *Pagrus major*: mass selection to genome editing. Fisheries Science, 89(2), 103-119. https://doi.org/10.1007/s12562-022-01668-0
9. Samuels, G., L. Hegarty, W. Fantham, D. Ashton, J. Blommaert, M.J. Wylie, D. Moran, and M. Wellenreuther. (2024). Generational breeding gains in a new species for aquaculture, the Australasian snapper (*Chrysophrys auratus*). Aquaculture, 586. https://doi.org/10.1016/j.aquaculture.2024.740782
10. Moran, D., J. Schleyken, C. Flammensbeck, W. Fantham, D. Ashton, and M. Wellenreuther. (2023). Enhanced survival and growth in the selectively bred *Chrysophrys auratus* (Australasian snapper, tāmure). Aquaculture, 563, 738970. https://doi.org/10.1016/j.aquaculture.2022.738970
11. Mylonas, C.C., A. Fostier, and S. Zanuy. (2010). Broodstock management and hormonal manipulations of fish reproduction. General and Comparative Endocrinology, 165(3), 516-534. https://doi.org/10.1016/j.ygcen.2009.03.007
12. Scott, S.G. and N.W. Pankhurst. (1992). Interannual variation in the reproductive cycle of the New Zealand snapper *Pagrus auratus* (Bloch & Schneider) (Sparidae). Journal of Fish Biology, 41(5), 685-696. https://doi.org/10.1111/j.1095-8649.1992.tb02698.x
13. Sim-Smith, C.J., A.G. Jeffs, and C.A. Radford. (2012). Localised spawning omission in snapper, *Chrysophrys auratus* (Sparidae). Marine and Freshwater Research, 63(2), 150-159. https://doi.org/10.1071/MF11203
14. Modern, J.S., I. Ruza, C.K. Flammensbeck, C. Radford, M. Wellenreuther, and M.J. Wylie. (2025). Captive reproduction of Australasian snapper (Sparidae: *Chrysophrys auratus*) - an emerging bream species for New Zealand aquaculture. SSRN. https://doi.org/http://dx.doi.org/10.2139/ssrn.5639193

15. Scott, S.G., J.R. Zeldis, and N.W. Pankhurst. (1993). Evidence of daily spawning in natural populations of the New Zealand snapper *Pagrus auratus* (Sparidae). Environmental Biology of Fishes, 36(2), 149-156. https://doi.org/10.1007/BF00002794
16. Francis, M.P. and N.W. Pankhurst. (1988). Juvenile sex inversion in the New Zealand snapper *Chrysophrys auratus* (Bloch and Schneider, 1801) (Sparidae). Marine and Freshwater Research, 39(5), 625-631. https://doi.org/https://doi.org/10.1071/MF9880625
17. Richards, J.S. (1994). Hormonal Control of Gene Expression in the Ovary. Endocrine Reviews, 15(6), 725-751. https://doi.org/10.1210/edrv-15-6-725
18. Yang, J., X. Ru, Y. Huang, J. Wu, T. Yang, P. Chen, J. Li, K. Zhu, and C. Zhu. (2025). Transcriptomic Analysis Provides New Insights into Oocyte Growth and Maturation in Greater Amberjack (*Seriola dumerili*). Animals (Basel), 15(3), 333. https://doi.org/10.3390/ani15030333
19. Jiang, M.-Y., Y.-F. Zhou, H. Liu, Y.-X. Peng, Y.-Q. Huang, S.-P. Deng, Y. Huang, G. Shi, C.-H. Zhu, G.-L. Li, and D.-N. Jiang. (2023). Transcriptomic analysis provides new insights into the secondary follicle growth in spotted scat (*Scatophagus argus*). Frontiers in Marine Science, 10. https://doi.org/10.3389/fmars.2023.1114872
20. Meng, F., S. Sun, X. Xu, W. Yu, R. Gan, L. Zhang, and W. Zhang. (2022). Transcriptomic analysis provides insights into the growth and maturation of ovarian follicles in the ricefield eel (*Monopterus albus*). Aquaculture, 555, 738251. https://doi.org/10.1016/j.aquaculture.2022.738251
21. Bogoch, Y., A. Jamieson-Lucy, C.E. Vejnar, K. Levy, A.J. Giraldez, M.C. Mullins, and Y.M. Elkouby. (2022). Stage specific transcriptomic analysis and database for zebrafish oogenesis. Frontiers in Cell and Developmental Biology, 10, 826892. https://doi.org/10.3389/fcell.2022.826892
22. Wylie, M.J., J.E. Symonds, A.N. Setiawan, G.W. Irvine, H. Liu, A. Elizur, and P.M. Lokman. (2019). Transcriptomic Changes during Previtellogenic and Vitellogenic Stages of Ovarian Development in Wreckfish (Hāpuku), *Polyprion oxygeneios* (Perciformes). Fishes, 4(1), 16-0. https://doi.org/10.3390/fishes4010016
23. Borella, M.I., C. Chehade, F.G. Costa, L.W.O. de Jesus, M. Cassel, and S.R. Batlouni. (2020). Chapter 14 - The brain-pituitary-gonad axis and the gametogenesis. In B. Baldisserotto, E.C. Urbinati, and J.E.P. Cyrino (Eds.), *Biology and Physiology of Freshwater Neotropical Fish* (pp. 315-341). Academic Press. https://doi.org/10.1016/B978-0-12-815872-2.00014-2
24. Safian, D., J. Bogerd, and R.W. Schulz. (2019). Regulation of spermatogonial development by Fsh: The complementary roles of locally produced Igf and Wnt signaling molecules in adult zebrafish testis. General and Comparative Endocrinology, 284, 113244-113244. https://doi.org/10.1016/j.ygcen.2019.113244
25. Martinez-Bengochea, A., L. Doretto, I.F. Rosa, M.A. Oliveira, C. Silva, D.M.Z.A. Silva, G.R. Santos, J.S.F. Santos, M.M. Avelar, L.V. Silva, D. Lucianelli-Junior, E.R.B. Souza, R.C. Silva, A.B. Stewart, L.S.O. Nakaghi, F.N. Valentin, and R.H. Nóbrega. (2020). Effects of 17β-estradiol on early gonadal development and expression of genes implicated in sexual differentiation of a South American

teleost, *Astyanax altiparanae*. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, 248-249, 110467. https://doi.org/10.1016/j.cbpb.2020.110467

26. Wang, W., H. Zhu, Z. Tian, A. Sun, Y. Dong, T. Dong, and H. Hu. (2020). Effects of 11-Ketotestosterone on Development of the Previtellogenic Ovary in the Sterlet, *Acipenser ruthenus*. Frontiers in Endocrinology (Lausanne), 11, 115. https://doi.org/10.3389/fendo.2020.00115
27. Wallace, R.A. and L.W. Browder. (1985). Vitellogenesis and Oocyte Growth in Nonmammalian Vertebrates. In L.W. Browder (Ed.), *Developmental Biology* (Vol. 1, pp. 127-177). Springer US. https://doi.org/10.1007/978-1-4615-6814-8_3
28. Lubzens, E., G. Young, J. Bobe, and J. Cerdà. (2010). Oogenesis in teleosts: How fish eggs are formed. General and Comparative Endocrinology, 165(3), 367-389. https://doi.org/10.1016/j.ygcen.2009.05.022
29. Levavi-Sivan, B., J. Bogerd, E.L. Mañanós, A. Gómez, and J.J. Lareyre. (2010). Perspectives on fish gonadotropins and their receptors. General and Comparative Endocrinology, 165(3), 412-437. https://doi.org/10.1016/j.ygcen.2009.07.019
30. Kagawa, H., K. Gen, K. Okuzawa, and H. Tanaka. (2003). Effects of Luteinizing Hormone and Follicle-Stimulating Hormone and Insulin-Like Growth Factor-I on Aromatase Activity and P450 Aromatase Gene Expression in the Ovarian Follicles of Red Seabream, *Pagrus major*. Biology of Reproduction, 68(5), 1562-1568. https://doi.org/10.1095/biolreprod.102.008219
31. Gen, K., K. Okuzawa, B. Senthilkumaran, H. Tanaka, S. Moriyama, and H. Kagawa. (2000). Unique Expression of Gonadotropin-I and -II Subunit Genes in Male and Female Red Seabream (*Pagrus major*) During Sexual Maturation. Biology of Reproduction, 63(1), 308-319. https://doi.org/10.1095/biolreprod63.1.308
32. Rajakumar, A. and B. Senthilkumaran. (2020). Steroidogenesis and its regulation in teleost-a review. Fish Physiology and Biochemistry, 46(3), 803-818. https://doi.org/10.1007/s10695-019-00752-0
33. Payne, A.H. and G.L. Youngblood. (1995). Regulation of expression of steroidogenic enzymes in Leydig cells. Biology of Reproduction, 52(2), 217-225. https://doi.org/10.1095/biolreprod52.2.217
34. Chiang, E.F.-L., Y.-L. Yan, S.-K. Tong, P.-H. Hsiao, Y. Guiguen, J. Postlethwait, and B.-C. Chung. (2001). Characterization of duplicated zebrafish *cyp19* genes. Journal of Experimental Zoology, 290(7), 709-714. https://doi.org/https://doi.org/10.1002/jez.1121
35. Rashid, H., H. Kitano, K. Hoon Lee, S. Nii, T. Shigematsu, K. Kadomura, A. Yamaguchi, and M. Matsuyama. (2007). Fugu (*Takifugu rubripes*) Sexual Differentiation: CYP19 Regulation and Aromatase Inhibitor Induced Testicular Development. Sexual Development, 1(5), 311-322. https://doi.org/10.1159/000108935
36. Blommaert, J., J. Sandoval-Castillo, L.B. Beheregaray, and M. Wellenreuther. (2024). Peering into the gaps: Long-read sequencing illuminates structural variants and genomic evolution in the Australasian snapper. Genomics (San Diego, Calif.), 116(5), 110929. https://doi.org/10.1016/j.ygeno.2024.110929

37. Catanach, A., R. Crowhurst, C. Deng, C. David, L. Bernatchez, and M. Wellenreuther. (2019). The genomic pool of standing structural variation outnumbers single nucleotide polymorphism by threefold in the marine teleost *Chrysophrys auratus*. Molecular Ecology, 28(6), 1210-1223. https://doi.org/10.1111/mec.15051
38. Gen, K., S. Yamaguchi, K. Okuzawa, N. Kumakura, H. Tanaka, and H. Kagawa. (2003). Physiological roles of FSH and LH in red seabream, *Pagrus major*. Fish Physiology and Biochemistry, 28(1-4), 77-80. https://doi.org/10.1023/B:FISH.0000030480.97947.ba
39. Mackie, M., G. Jackson, N. Tapp, J. Norriss, and A. Thomson, *Macroscopic and microscopic description of snapper (Pagrus auratus) gonads from Shark Bay, Western Australia*, D.o. Fisheries, Editor. 2009, Department of Fisheries, Western Australia.
40. Hoff, K.J., S. Lange, A. Lomsadze, M. Borodovsky, and M. Stanke. (2016). BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. Bioinformatics, 32(5), 767-769. https://doi.org/10.1093/bioinformatics/btv661
41. Brůna, T., K.J. Hoff, A. Lomsadze, M. Stanke, and M. Borodovsky. (2021). BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. NAR Genomics and Bioinformatics, 3(1), lqaa108-lqaa108. https://doi.org/10.1093/nargab/lqaa108
42. Hoff, K.J., A. Lomsadze, M. Borodovsky, M. Stanke, and M. Kollmar. (2019). Whole-Genome Annotation with BRAKER. Gene Prediction, 1962, 65-95. https://doi.org/10.1007/978-1-4939-9173-0_5
43. Simão, F.A., R.M. Waterhouse, P. Ioannidis, E.V. Kriventseva, and E.M. Zdobnov. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics, 31(19), 3210-3212. https://doi.org/10.1093/bioinformatics/btv351
44. Li, H. (2023). Protein-to-genome alignment with miniprot. Bioinformatics (Oxford, England), 39(1). https://doi.org/10.1093/bioinformatics/btad014
45. Huang, N. and H. Li. (2023). compleasm: a faster and more accurate reimplementation of BUSCO. Bioinformatics (Oxford, England), 39(10). https://doi.org/10.1093/bioinformatics/btad595
46. Brůna, T., L. Gabriel, and K.J. Hoff. (2024). Navigating eukaryotic genome annotation pipelines: A route map to BRAKER, Galba, and TSEBRA. https://doi.org/10.48550/arxiv.2403.19416
47. Buchfink, B., C. Xie, and D.H. Huson. (2015). Fast and sensitive protein alignment using DIAMOND. Nature Methods, 12(1), 59-60. https://doi.org/10.1038/nmeth.3176
48. Gotoh, O. (2008). Space-efficient and accurate method for mapping and aligning cDNA sequences onto genomic sequence. Nucleic Acids Research, 36(8), 2630-2638. https://doi.org/10.1093/nar/gkn105
49. Iwata, H. and O. Gotoh. (2012). Benchmarking spliced alignment programs including Spaln2, an extended version of Spaln that incorporates additional species-specific features. Nucleic Acids Research, 40(20), e161-e161. https://doi.org/10.1093/nar/gks708

50. Kovaka, S., A.V. Zimin, G.M. Pertea, R. Razaghi, S.L. Salzberg, and M. Pertea. (2019). Transcriptome assembly from long-read RNA-seq alignments with StringTie2. Genome Biology, 20(1), 278-278. https://doi.org/10.1186/s13059-019-1910-1
51. Pertea, G. and M. Pertea. (2020). GFF Utilities: GffRead and GffCompare [version 2; peer review: 3 approved]. F1000 research, 9, 304. https://doi.org/10.12688/f1000research.23297.2
52. Stanke, M., M. Diekhans, R. Baertsch, and D. Haussler. (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics, 24(5), 637-644. https://doi.org/10.1093/bioinformatics/btn013
53. Stanke, M., O. Schöffmann, B. Morgenstern, and S. Waack. (2006). Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC Bioinformatics, 7(1), 62-62. https://doi.org/10.1186/1471-2105-7-62
54. Brůna, T., A. Lomsadze, and M. Borodovsky. (2023). GeneMark-ETP: automatic gene finding in eukaryotic genomes in consistence with extrinsic data. bioRxiv. https://doi.org/10.1101/2023.01.13.524024
55. Gabriel, L., K.J. Hoff, T. Brůna, M. Borodovsky, and M. Stanke. (2021). TSEBRA: transcript selector for BRAKER. BMC bioinformatics, 22(1), 1-566. https://doi.org/10.1186/s12859-021-04482-0
56. Gabriel, L., T. Brůna, K.J. Hoff, M. Ebel, A. Lomsadze, M. Borodovsky, and M. Stanke. (2024). BRAKER3: Fully Automated Genome Annotation Using RNA-Seq and Protein Evidence with GeneMark-ETP, AUGUSTUS and TSEBRA (Updated February 29, 2024). Biotech Week, 76.
57. Tegenfeldt, F., D. Kuznetsov, M. Manni, M. Berkeley, E.M. Zdobnov, and E.V. Kriventseva. (2025). OrthoDB and BUSCO update: annotation of orthologs with wider sampling of genomes. Nucleic Acids Research, 53(D1), D516-D522. https://doi.org/10.1093/nar/gkae987
58. Shumate, A. and S.L. Salzberg. (2021). Liftoff: accurate mapping of gene annotations. Bioinformatics, 37(12), 1639-1643. https://doi.org/10.1093/bioinformatics/btaa1016
59. Lu, J., D. Gao, Y. Sims, W. Fang, J. Collins, J. Torrance, G. Lin, J. Xie, J. Liu, and K. Howe. (2022). Chromosome-level Genome Assembly of *Acanthopagrus latus* Provides Insights into Salinity Stress Adaptation of Sparidae. Marine Biotechnology (New York, N.Y.), 24(3), 655-660. https://doi.org/10.1007/s10126-022-10119-x
60. Emms, D.M. and S. Kelly. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biology, 16(1), 157-157. https://doi.org/10.1186/s13059-015-0721-2
61. Emms, D.M. and S. Kelly. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biology, 20(1), 238-238. https://doi.org/10.1186/s13059-019-1832-y
62. Patel, H., P. Ewels, J. Manning, M.U. Garcia, A. Peltzer, R. Hammarén, O. Botvinnik, A. Talbot, G. Sturm, M. Zepper, D. Moreno, P. Vemuri, M. Binzer-Panchal, E. Greenberg, silviamorins, L. Pantano, R. Syme, G. Kelly, F. Hanssen, J.A.F. Yates, J. Espinosa-Carrasco, rfenouil, L. Zappia, C. Cheshire, E. Miller, marchoeppner, P. Zhou, S. Guinchard, and G. Gabernet. (2024). *nf-core/rnaseq: nf-core/rnaseq v3.17.0 - Neon Newt*. In (Version 3.17.0) Zenodo.

63. Di Tommaso, P., M. Chatzou, E.W. Floden, P.P. Barja, E. Palumbo, and C. Notredame. (2017). Nextflow enables reproducible computational workflows. Nature Biotechnology, 35(4), 316-319. https://doi.org/10.1038/nbt.3820
64. Dobin, A., C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T.R. Gingeras. (2013). STAR: Ultrafast universal RNA-seq aligner. Bioinformatics, 29(1), 15-21. https://doi.org/10.1093/bioinformatics/bts635
65. Patro, R., G. Duggal, M.I. Love, R.A. Irizarry, and C. Kingsford. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods, 14(4), 417-419. https://doi.org/10.1038/nmeth.4197
66. Love, M.I., W. Huber, and S. Anders. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology (Online Edition), 15(12), 550-550. https://doi.org/10.1186/s13059-014-0550-8
67. Chen, Y., L. Chen, A.T.L. Lun, P.L. Baldoni, and G.K. Smyth, *edgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets*. 2024, Cold Spring Harbor Laboratory Press: Cold Spring Harbor.
68. Chen, Y., A.T.L. Lun, and G.K. Smyth. (2016). From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline [version 1; peer review: 5 approved]. F1000 research, 5, 1438. https://doi.org/10.12688/f1000research.8987.1
69. Benjamini, Y. and Y. Hochberg. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B, Methodological, 57(1), 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
70. Abueg, L.A.L., E. Afgan, O. Allart, A.H. Awan, W.A. Bacon, D. Baker, M. Bassetti, B. Batut, M. Bernt, D. Blankenberg, A. Bombarely, A. Bretaudeau, C.J. Bromhead, M.L. Burke, P.K. Capon, M. Čech, M. Chavero-Díez, J.M. Chilton, T.J. Collins, F. Coppens, N. Coraor, G. Cuccuru, F. Cumbo, J. Davis, P.F. De Geest, W. de Koning, M. Demko, A. DeSanto, J.M.D. Begines, M.A. Doyle, B. Droesbeke, A. Erxleben-Eggenhofer, M.C. Föll, G. Formenti, A. Fouilloux, R. Gangazhe, T. Genthon, J. Goecks, A.N.G. Beltran, N.A. Goonasekera, N. Goué, T.J. Griffin, B.A. Grüning, A. Guerler, S. Gundersen, O.J.R. Gustafsson, C. Hall, T.W. Harrop, H. Hecht, A. Heidari, T. Heisner, F. Heyl, S. Hiltemann, H.-R. Hotz, C.J. Hyde, P.D. Jagtap, J. Jakiela, J.E. Johnson, J. Joshi, M. Jossé, K. Jum'ah, M. Kalaš, K. Kamieniecka, T. Kayikcioglu, M. Konkol, L. Kostrykin, N. Kucher, A. Kumar, M. Kuntz, D. Lariviere, R. Lazarus, Y.L. Bras, G.L. Corguillé, J. Lee, S. Leo, L. Liborio, R. Libouban, D.L. Tabernero, L. Lopez-Delisle, L.S. Los, A. Mahmoud, I. Makunin, P. Marin, S. Mehta, W. Mok, P.A. Moreno, F. Morier-Genoud, S. Mosher, T. Müller, E. Nasr, A. Nekrutenko, T.M. Nelson, A.J. Oba, A. Ostrovsky, P.V. Polunina, K. Poterlowicz, E.J. Price, G.R. Price, H. Rasche, and B. Raubenolt. (2024). The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. Nucleic Acids Research, 52(W1), W83-W94. https://doi.org/10.1093/nar/gkae410
71. Cantalapiedra, C.P., A. Hernández-Plaza, I. Letunic, P. Bork, and J. Huerta-Cepas. (2021). eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the

metagenomic scale. Molecular Biology and Evolution, 38(12), 5825-5829. https://doi.org/10.1093/molbev/msab293

72. Huerta-Cepas, J., D. Szklarczyk, D. Heller, A. Hernández-Plaza, S.K. Forslund, H. Cook, D.R. Mende, I. Letunic, T. Rattei, Lars J. Jensen, C. von Mering, and P. Bork. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Research, 47(D1), D309-D314. https://doi.org/10.1093/nar/gky1085
73. Ashburner, M., C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, and G. Sherlock. (2000). Gene Ontology: tool for the unification of biology. Nature Genetics, 25(1), 25-29. https://doi.org/10.1038/75556
74. Aleksander, S.A., J. Balhoff, S. Carbon, J.M. Cherry, H.J. Drabkin, D. Ebert, N.L. Harris, S. Moxon, C.J. Mungall, P.W. Sternberg, P.D. Thomas, F. Quaglia, M. Giglio, G. Antonazzo, H. Attrill, G. dos Santos, C.J. Tabone, P. Zhou, S.H. Ahmed, P. Asanitthong, M.N. Erdol, M.C. Gage, M. Ali Kadhum, K.Y.C. Li, A. Michalak, A. Pritazahra, S.C.C. Saverimuttu, R. Su, K.E. Thurlow, R.C. Lovering, C. Logie, S. Oliferenko, J. Blake, L. Corbani, L. Ni, D. Sitnikov, C. Smith, J. Seager, L. Cooper, P. Jaiswal, S. Naithani, M. Lera-Ramirez, K. Rutherford, V. Wood, J.L. De Pons, M.R. Dwinell, G.T. Hayman, A.E. Kwitek, S.J.F. Laulederkind, M.A. Tutaj, M. Vedi, S.-J. Wang, P. D'Eustachio, K. Axelsen, S.A. Aleksander, S.R. Engel, K. Karra, S.R. Miyasato, R.S. Nash, S. Weng, E.D. Wong, E. Bakker, T.Z. Berardini, L. Reiser, A. Auchincloss, G. Argoud-Puy, M.-C. Blatter, E. Boutet, L. Breuza, C. Casals-Casas, N. Gruaz-Gumowski, C. Hulo, F. Jungo, P. Le Mercier, P. Masson, I. Pedruzzi, L. Pourcel, S. Poux, C. Rivoire, S. Sundaram, E. Bowler-Barnett, P. Denny, A. Ignatchenko, A. Lock, M. Magrane, M.J. Martin, P. Raposo, E. Speretta, N. Tyagi, K. Warner, R. Zaru, A.D. Diehl, R. Lee, J. Chan, M. Zarowiecki, M. Fisher, V. Ponferrada, A. Zorn, S. Ramachandran, and L. Ruzicka. (2023). The Gene Ontology knowledgebase in 2023. Genetics (Austin), 224(1). https://doi.org/10.1093/genetics/iyad031
75. Kanehisa, M., M. Furumichi, Y. Sato, Y. Matsuura, and M. Ishiguro-Watanabe. (2025). KEGG: biological systems database as a model of the real world. Nucleic Acids Research, 53(D1), D672-D677. https://doi.org/10.1093/nar/gkae909
76. Kanehisa, M. (2019). Toward understanding the origin and evolution of cellular organisms. Protein Science, 28(11), 1947-1951. https://doi.org/10.1002/pro.3715
77. Ogata, H., S. Goto, K. Sato, W. Fujibuchi, H. Bono, and M. Kanehisa. (1999). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research, 27(1), 29-34. https://doi.org/10.1093/nar/27.1.29
78. Wu, T., E. Hu, S. Xu, M. Chen, P. Guo, Z. Dai, T. Feng, L. Zhou, W. Tang, L. Zhan, X. Fu, S. Liu, X. Bo, and G. Yu. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (New York, NY), 2(3), 100141-100141. https://doi.org/10.1016/j.xinn.2021.100141
79. Wijesooriya, K., S.A. Jadaan, K.L. Perera, T. Kaur, M. Ziemann, and M.L. Kemp. (2022). Urgent need for consistent standards in functional enrichment analysis. PLoS Computational Biology, 18(3), e1009935-e1009935. https://doi.org/10.1371/journal.pcbi.1009935

80. Virtanen, P., R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, İ. Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, and P. van Mulbregt. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods, 17(3), 261-272. https://doi.org/10.1038/s41592-019-0686-2
81. Seabold, S. and J. Perktold, *Statsmodels: Econometric and Statistical Modeling with Python*, in *Proceedings of the 9th Python in Science Conference*, S. van der Walt and J. Millman, Editors. 2010. p. 92-96.
82. Diaz de Cerio, O., I. Rojo-Bartolomé, C. Bizarro, M. Ortiz-Zarragoitia, and I. Cancio. (2012). 5S rRNA and Accompanying Proteins in Gonads: Powerful Markers to Identify Sex and Reproductive Endocrine Disruption in Fish. Environmental Science & Technology, 46(14), 7763-7771. https://doi.org/10.1021/es301132b
83. Teletchea, F. (2015). Domestication of Marine Fish Species: Update and Perspectives. Journal of marine science and engineering, 3(4), 1227-1243. https://doi.org/10.3390/jmse3041227
84. Nyuji, M., Y. Hongo, M. Yoneda, and M. Nakamura. (2020). Transcriptome characterization of BPG axis and expression profiles of ovarian steroidogenesis-related genes in the Japanese sardine. BMC Genomics, 21(1), 668-668. https://doi.org/10.1186/s12864-020-07080-1
85. Zhao, C., C. Bian, X. Mu, X. Zhang, and Q. Shi. (2024). Gonadal transcriptome sequencing reveals sexual dimorphism in expression profiling of sex-related genes in Asian arowana (*Scleropages formosus*). Frontiers in Genetics, 15, 1381832-1381832. https://doi.org/10.3389/fgene.2024.1381832
86. Li, S., G. Lin, W. Fang, P. Huang, D. Gao, J. Huang, J. Xie, and J. Lu. (2020). Gonadal Transcriptome Analysis of Sex-Related Genes in the Protandrous Yellowfin Seabream (*Acanthopagrus latus*). Frontiers in Genetics, 11, 709-709. https://doi.org/10.3389/fgene.2020.00709
87. Li, C., R. Cai, W. Shi, H. Zhang, Z. Liu, F. Xie, Y. Chen, and Q. Hong. (2024). Comparative transcriptome analysis of ovaries and testes reveals sex-biased genes and pathways in zebrafish. Gene, 901, 148176-148176. https://doi.org/10.1016/j.gene.2024.148176
88. Li, X., B. Lin, X. Zhang, X. Shen, H. Ouyang, Z. Wu, Y. Tian, L. Fang, and Y. Huang. (2022). Comparative transcriptomics in the hypothalamic-pituitary-gonad axis of mammals and poultry. Genomics (San Diego, Calif.), 114(4), 110396-110396. https://doi.org/10.1016/j.ygeno.2022.110396
89. Yang, D., C. Yin, Y. Chang, Y. Dou, Z. Hao, and J. Ding. (2016). Transcriptome analysis of male and female mature gonads of Japanese scallop *Patinopecten yessonsis*. Genes & Genomics, 38(11), 1041-1052. https://doi.org/10.1007/s13258-016-0449-8
90. González-Castellano, I., C. Manfrin, A. Pallavicini, and A. Martínez-Lage. (2019). De novo gonad transcriptome analysis of the common littoral shrimp *Palaemon serratus*: novel insights into sex-related genes. BMC Genomics, 20(1), 757-757. https://doi.org/10.1186/s12864-019-6157-4

91. Wang, L., T. Bu, X. Wu, L. Li, F. Sun, and C.Y. Cheng. (2024). Motor proteins, spermatogenesis and testis function. Advances in protein chemistry and structural biology, 141, 381. https://doi.org/10.1016/bs.apcsb.2024.03.003
92. Linck, R.W., H. Chemes, and D.F. Albertini, *The axoneme: the propulsive engine of spermatozoa and cilia and associated ciliopathies leading to infertility*, in *Journal of Assisted Reproduction and Genetics*. 2016, Springer US: New York. p. 141-156.
93. Zhu, Y., G. Yin, L. Tai, B. Wang, Y. Lu, Z. Zhang, S. Chen, and F. Sun. (2024). In-cell structural insight into the asymmetric assembly of central apparatus in mammalian sperm axoneme. bioRxiv, 2024.08.06.606614. https://doi.org/10.1101/2024.08.06.606614
94. Alam, M.A., Y. Kobayashi, T. Hirai, and M. Nakamura. (2010). Isolation, characterization and expression analyses of FSH receptor in protogynous grouper. Comparative biochemistry and physiology. Part A, Molecular & Integrative Physiology, 156(3), 364-371. https://doi.org/10.1016/j.cbpa.2010.03.001
95. Kobayashi, Y., M.A. Alam, R. Horiguchi, A. Shimizu, and M. Nakamura. (2010). Sexually Dimorphic Expression of Gonadotropin Subunits in the Pituitary of Protogynous Honeycomb Grouper (*Epinephelus merra*): Evidence That Follicle-Stimulating Hormone (FSH) Induces Gonadal Sex Change. Biology of Reproduction, 82(6), 1030-1036. https://doi.org/10.1095/biolreprod.109.080986
96. Liguang, Y.A.N., P. Swanson, and W.W. Dickhoff. (1992). A two-receptor model for salmon gonadotropins (GTH I and GTH II). Biology of Reproduction, 47(3), 418-427. https://doi.org/10.1095/biolreprod47.3.418
97. Lau, E.S.-W., Z. Zhang, M. Qin, and W. Ge. (2016). Knockout of Zebrafish Ovarian Aromatase Gene (*cyp19a1a*) by TALEN and CRISPR/Cas9 Leads to All-male Offspring Due to Failed Ovarian Differentiation. Scientific Reports, 6(1), 37357. https://doi.org/10.1038/srep37357
98. Guo, C., K. Zhang, C. Li, R. Xing, S. Xu, D. Wang, and X. Wang. (2024). Cyp19a1a Promotes Ovarian Maturation through Regulating E2 Synthesis with Estrogen Receptor 2a in *Pampus argenteus* (Euphrasen, 1788). International Journal of Molecular Sciences, 25(3), 1583. https://doi.org/10.3390/ijms25031583
99. Goto-Kazeto, R., K.E. Kight, Y. Zohar, A.R. Place, and J.M. Trant. (2004). Localization and expression of aromatase mRNA in adult zebrafish. General and Comparative Endocrinology, 139(1), 72-84. https://doi.org/10.1016/j.ygcen.2004.07.003
100. Zapater, C., F. Chauvigne, A.P. Scott, A. Gomez, I. Katsiadaki, and J. Cerda. (2012). Piscine Follicle-Stimulating Hormone Triggers Progestin Production in Gilthead Seabream Primary Ovarian Follicles. Biology of Reproduction, 87(5), 111-111. https://doi.org/10.1095/biolreprod.112.102533
101. Roy Moulik, S., P. Pal, S. Majumder, B. Mallick, S. Gupta, P. Guha, S. Roy, and D. Mukherjee. (2016). Gonadotropin and sf-1 regulation of *cyp19a1a* gene and aromatase activity during oocyte development in the rohu, *L. rohita*. Comparative Biochemistry and Physiology. Part A, Molecular & integrative physiology, 196, 1-10. https://doi.org/10.1016/j.cbpa.2016.02.004

102. Wylie, M.J., A.N. Setiawan, G.W. Irvine, J.E. Symonds, A. Elizur, M. Dos Santos, and P.M. Lokman. (2018). Ovarian development of captive F1 wreckfish (hāpuku) *Polyprion oxygeneios* under constant and varying temperature regimes – Implications for broodstock management. General and Comparative Endocrinology, 257, 86-96. https://doi.org/10.1016/j.ygcen.2017.08.022

103. Zhang, X., M. Li, H. Ma, X. Liu, H. Shi, M. Li, and D. Wang. (2017). Mutation of *foxl2* or *cyp19a1a* Results in Female to Male Sex Reversal in XX Nile Tilapia. Endocrinology (Philadelphia), 158(8), 2634-2647. https://doi.org/10.1210/en.2017-00127

104. Shen, Z.-G., N. Eissa, H. Yao, Z.-G. Xie, and H.-P. Wang. (2018). Effects of Temperature on the Expression of Two Ovarian Differentiation-Related Genes *foxl2* and *cyp19a1a*. Frontiers in Physiology, 9, 1208. https://doi.org/10.3389/fphys.2018.01208

105. Ohta, K., S. Yamaguchi, A. Yamaguchi, K. Okuzawa, K. Gen, H. Kagawa, and M. Matsuyama. (2002). Biosynthesis of estradiol-17β in the ovarian follicles of the red seabream *Pagrus major* during vitellogenesis. Fisheries Science, 68(3), 680-687. https://doi.org/10.1046/j.1444-2906.2002.00477.x

106. Gen, K., K. Okuzawa, N. Kumakura, S. Yamaguchi, and H. Kagawa. (2001). Correlation Between Messenger RNA Expression of Cytochrome P450 Aromatase and Its Enzyme Activity During Oocyte Development in the Red Seabream (*Pagrus major*). Biology of Reproduction, 65(4), 1186-1194. https://doi.org/10.1095/biolreprod65.4.1186

107. Grier, H. and M. Aranzábal. (2009). The Testis and Spermatogenesis in Teleosts. In B. Jamieson (Ed.), *Reproductive Biology and Phylogeny of Fishes (Agnathans and Bony Fishes)* (pp. 133-156). CRC Press. https://doi.org/10.1201/9781482280609-10

108. Qu, J., R. Li, Y. Xie, Y. Liu, J. Liu, and Q. Zhang. (2022). Differential transcriptomic profiling provides new insights into oocyte development and lipid droplet formation in Japanese flounder (*Paralichthys olivaceus*). Aquaculture, 550, 737843. https://doi.org/https://doi.org/10.1016/j.aquaculture.2021.737843

109. García-López, Á., M.I. Sánchez-Amaya, and F. Prat. (2011). Targeted gene expression profiling in European sea bass (*Dicentrarchus labrax*, L.) follicles from primary growth to late vitellogenesis. Comparative biochemistry and physiology. Part A, Molecular & Integrative Physiology, 160(3), 374-380. https://doi.org/10.1016/j.cbpa.2011.07.006

110. Yang, Y., Y. Zhou, G. Wessel, W. Hu, and D. Xu. (2024). Single-cell transcriptomes reveal spermatogonial stem cells and the dynamic heterogeneity of spermatogenesis in a seasonal breeding teleost. Development (Cambridge), 151(22). https://doi.org/10.1242/dev.203142

111. Qian, P., J. Kang, D. Liu, and G. Xie. (2022). Single Cell Transcriptome Sequencing of Zebrafish Testis Revealed Novel Spermatogenesis Marker Genes and Stronger Leydig-Germ Cell Paracrine Interactions. Frontiers in Genetics, 13, 851719-851719. https://doi.org/10.3389/fgene.2022.851719

112. Guo, J., E.J. Grow, C. Yi, H. Mlcochova, G.J. Maher, C. Lindskog, P.J. Murphy, C.L. Wike, D.T. Carrell, A. Goriely, J.M. Hotaling, and B.R. Cairns. (2017). Chromatin and Single-Cell RNA-Seq Profiling Reveal Dynamic Signaling and Metabolic Transitions during Human Spermatogonial Stem Cell Development. Cell Stem Cell, 21(4), 533-546.e6. https://doi.org/10.1016/j.stem.2017.09.003

113. Hermann, B.P., K. Cheng, A. Singh, L. Roa-De La Cruz, K.N. Mutoji, I.C. Chen, H. Gildersleeve, J.D. Lehle, M. Mayo, B. Westernströer, N.C. Law, M.J. Oatley, E.K. Velte, B.A. Niedenberger, D. Fritze, S. Silber, C.B. Geyer, J.M. Oatley, and J.R. McCarrey. (2018). The Mammalian Spermatogenesis Single-Cell Transcriptome, from Spermatogonial Stem Cells to Spermatids. Cell Reports (Cambridge), 25(6), 1650-1667.e8. https://doi.org/10.1016/j.celrep.2018.10.026

114. Sohni, A., K. Tan, H.-W. Song, D. Burow, D. de Rooij, L. Laurent, T.-C. Hsieh, R. Rabah, S. Hammoud, E. Vicini, and M. Wilkinson. (2019). The neonatal and adult human testis defined at the single-cell level. The Journal of Urology, 201(Supplement 4). https://doi.org/10.1097/01.JU.0000555269.58885.62

115. Hudson, M., A. Thompson, P. Wilcox, J.P. Mika, C. Battershill, M. Stott, R.T.r. Brooks, L. Warbrick, and i.b. Te Kotahi Research Institute. (2021). Te Nohonga Kaitiaki guidelines for genomic research on taonga species (with background). University of Waikato.

Figures

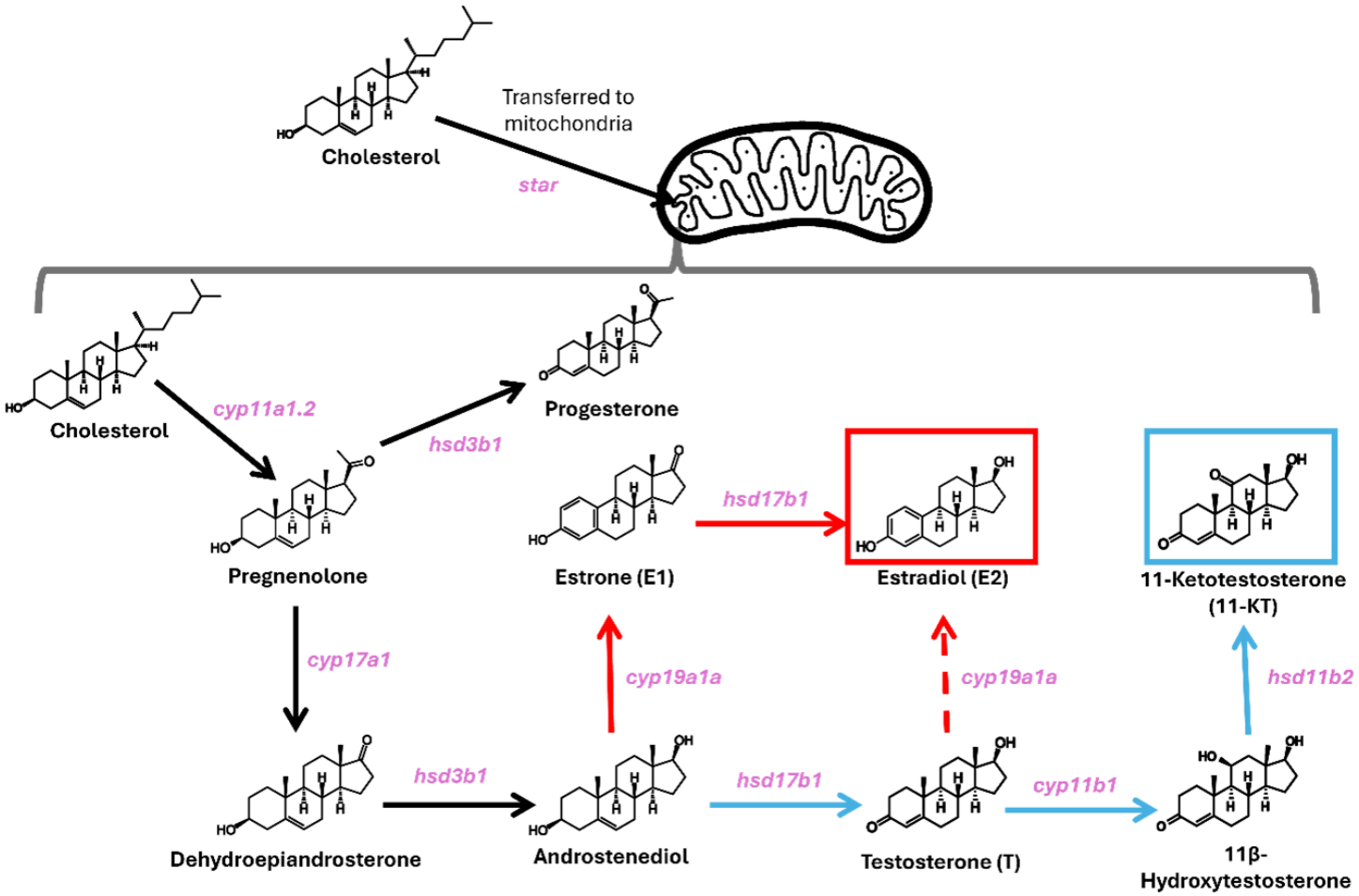


Figure 1

The steroid hormone biosynthesis pathways responsible for the production of the major androgen (11-KT, blue box) and oestrogen (E2, red box) in teleosts. Stages along the processes are denoted using arrows. Black arrows show steps common to the synthesis of both E2 and 11-KT, blue arrows denote

steps responsible for the production of 11-KT while red arrows denote stages responsible for the production of E2. The dashed red arrow represents an alternate pathway for E2 production which is not considered to be the primary pathway for E2 production in the red seabream (*Pagrus major*), the sister species of snapper [105]. The name of each gene encoding the enzyme / protein responsible for each step is given in pink text.

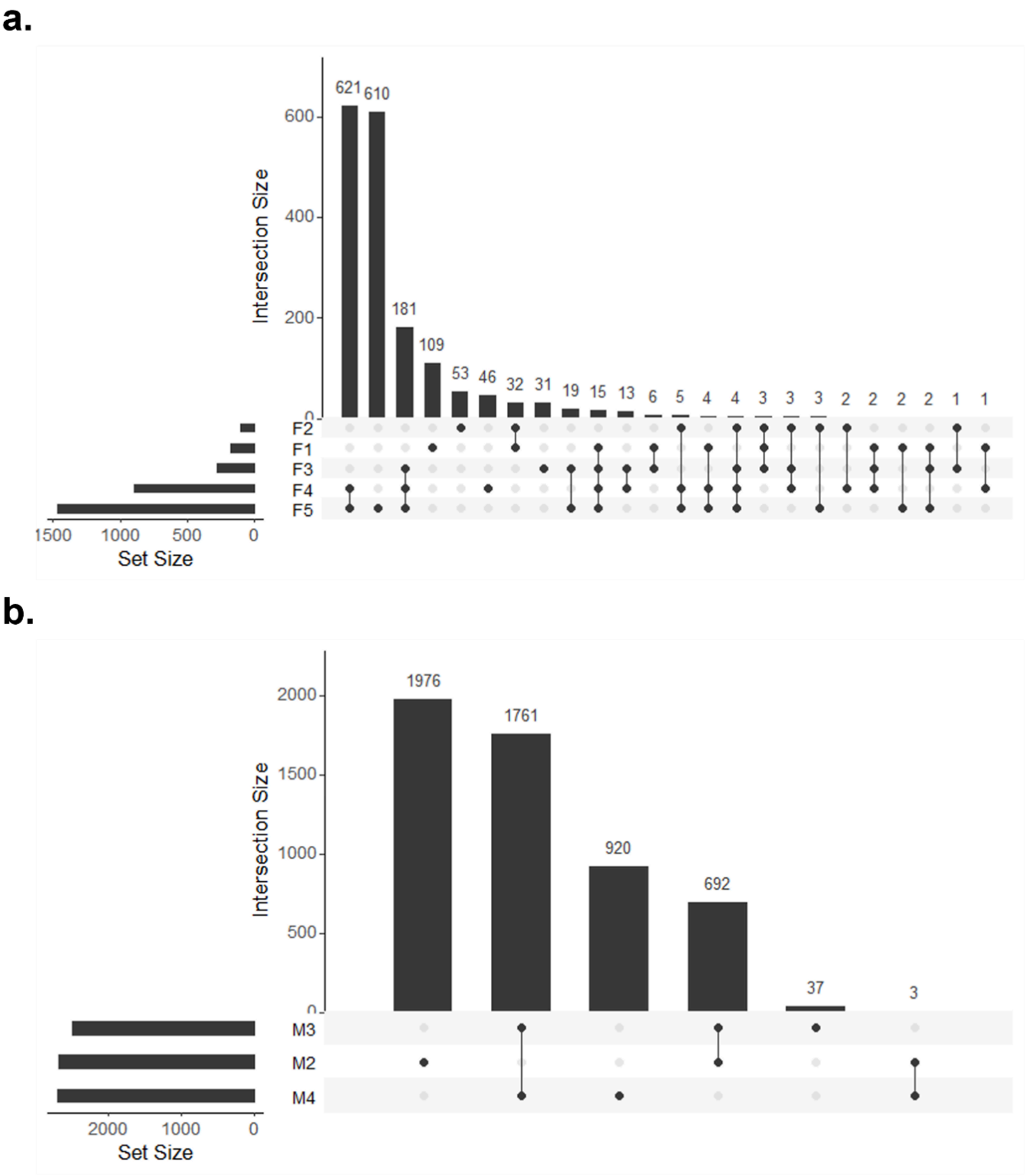


Figure 2

UpSet plots of the number of differentially expressed genes (DEGs) in the ovary at each stage of oogenesis in females (a.) and in the testis at each stage of spermatogenesis in males (b.). Each set represents the number of DEGs associated with each gametogenic stage from all of the pairwise comparisons of each stage. Horizontal bars indicate the total number of DEGs for each comparison, while vertical bars represent the size of intersections (shared DEGs) across multiple comparisons, as indicated by the connected dots below.

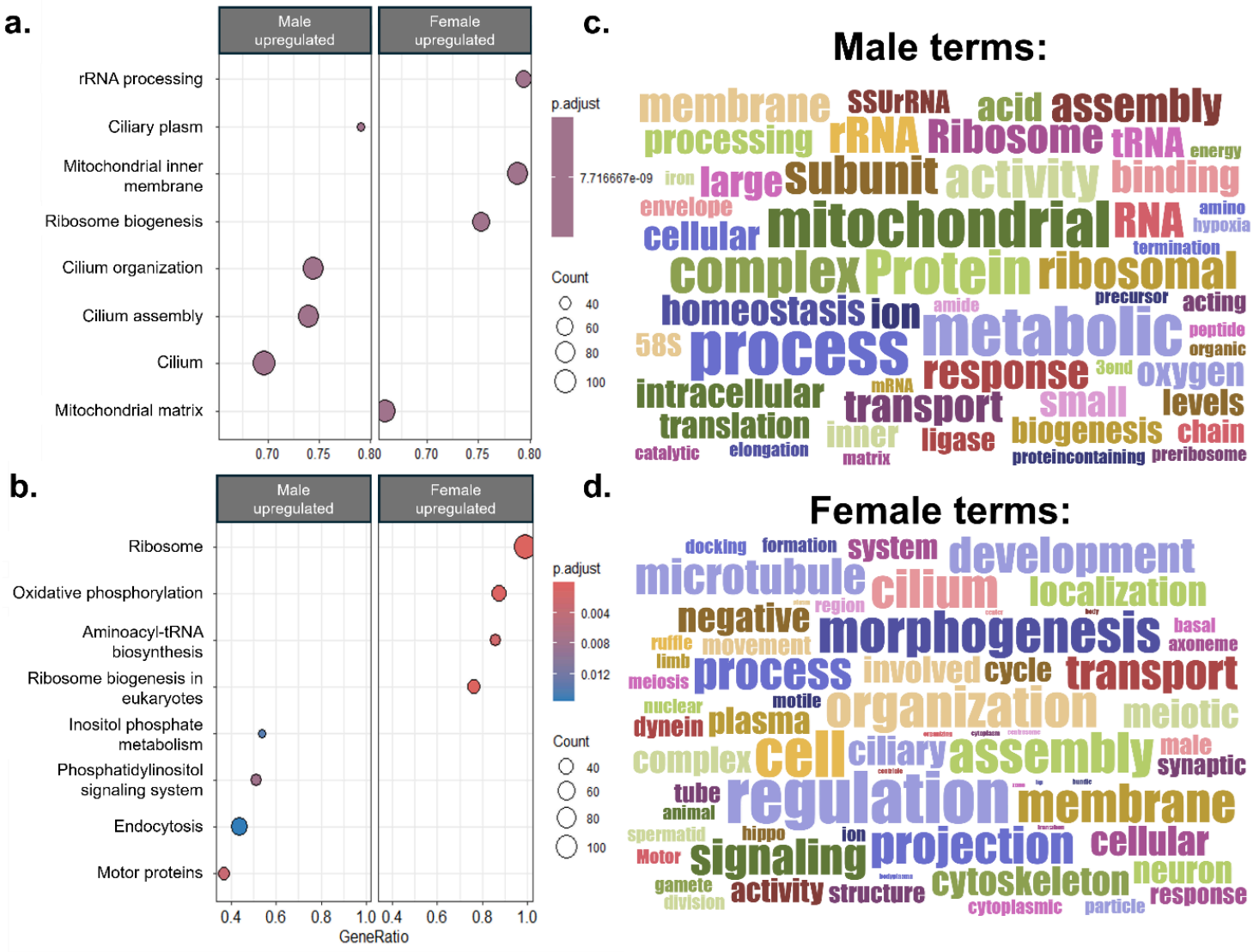


Figure 3

Dotplots based on gene set enrichment analysis (GSEA) results (a. and b.), showing the top eight enriched terms for a. gene ontology (GO) terms, and b. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. In both plots, each dot represents a GO term / KEGG pathway, with its position along the x-axis indicating the direction of enrichment: terms on the left are upregulated in male gene sets, while those on the right are upregulated in females. Dot size corresponds to the number of genes contributing to each term, and colour reflects the significance level (adjusted p-value). Only terms meeting the enrichment significance threshold (adjusted $p < 0.05$) are shown. Note, there were 202

enriched GO terms between both males and females. Therefore, only the top eight terms for both males and females are shown, all of which had very low adjusted p-values. To the right of these plots are word clouds displaying the top 100 most common words contained in the significantly activated GO terms of males (c.) and females (d.). The relative size of each word relates to the relative frequency that word occurred in the enriched GO terms, with larger words relating to a higher frequency.

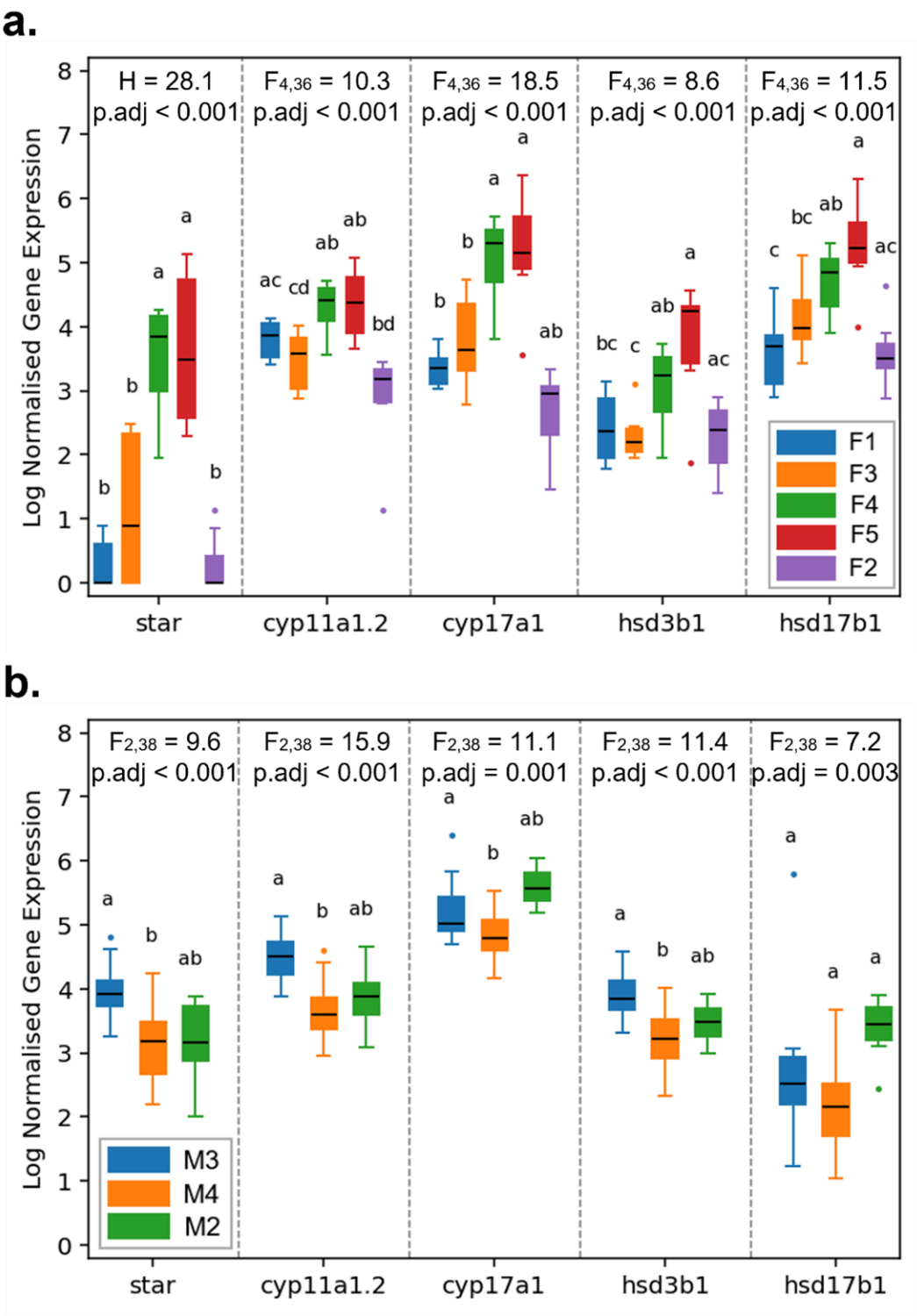


Figure 4

Gene expression box plots for steroidogenic enzymes responsible for the production of testosterone in females (a.), and males (b.). Bars are grouped by gene and the colour of the bars relates to the gametogenic stage: red = “immature” (F1) for females and “developing” (M3) for males, blue = “developing” (F3) for females and “developed” (M4) for males, orange = “developed” (F4, vitellogenic) for females and “resting post-spawning” (M2) for males, green = “spawning” (F5) in females and purple = “resting post-spawning” (F2) for females. The post Benjamini–Hochberg p-value correction results of ANOVA or Kruskal-Wallis tests (depending on whether the assumptions of an ANOVA could be satisfied) are given above each bar plot. Provided a significant ANOVA or Kruskal-Wallis test result was obtained, the significance between groups as determined by post-hoc testing is indicated by letters above the bars. Any two bars sharing the same letter are not significantly different to each other.

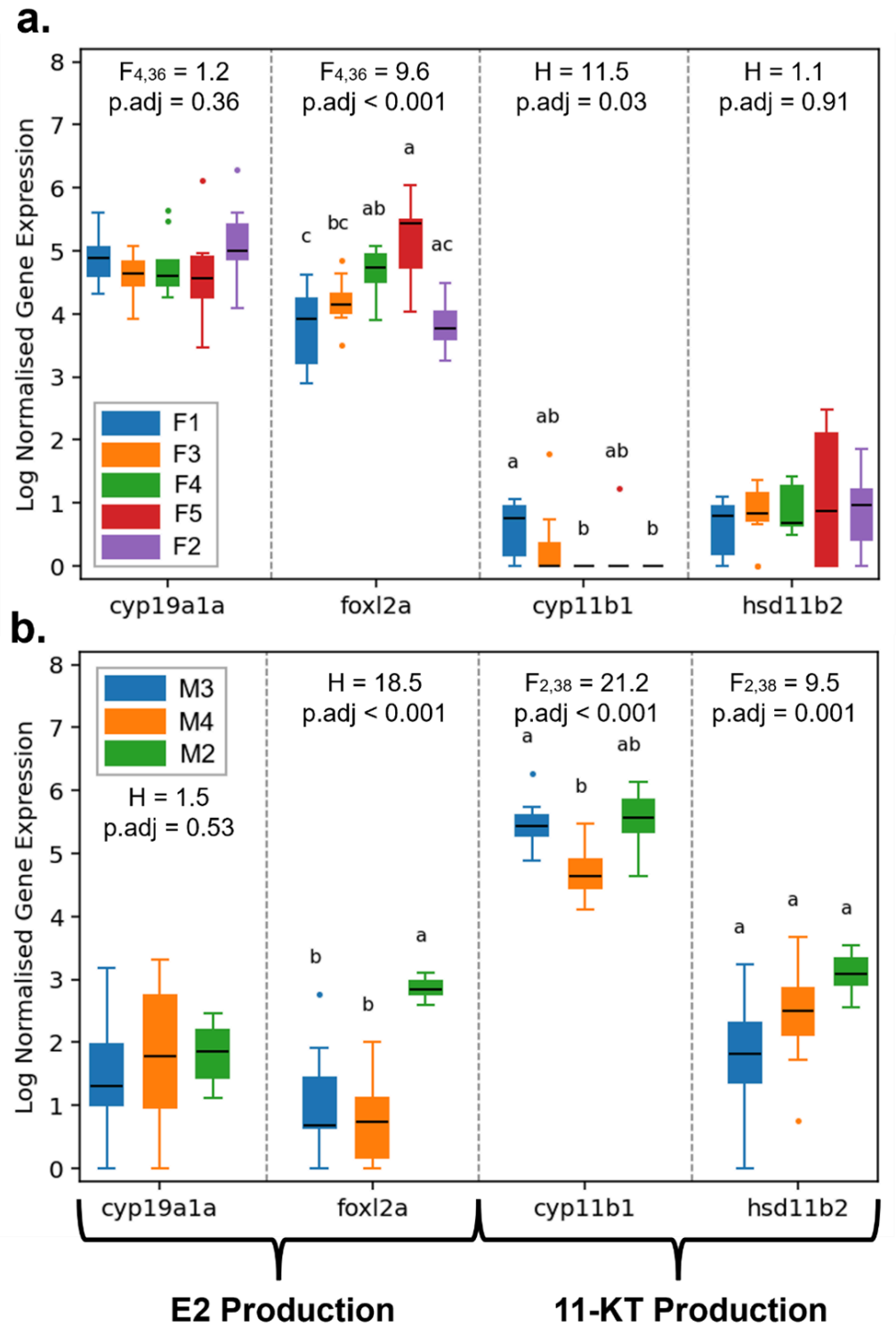


Figure 5

Gene expression box plots for steroidogenic enzymes associated with the production of E2 (*cyp19a1a* and *foxl2a*) and 11-KT (*cyp11b1* and *hsd11b2*) in females (a.), and males (b.). Bars are grouped by gene and the colour of the bars relates to the gametogenic stage: red = “immature” (F1) for females and “developing” (M3) for males, blue = “developing” (F3) for females and “developed” (M4) for males, orange = “developed” (F4, vitellogenic) for females and “resting post-spawning” (M2) for males, green =

“spawning” (F5) in females and purple = “resting post-spawning” (F2) for females. The post Benjamini–Hochberg p-value correction results of ANOVA or Kruskal-Wallis tests (depending on whether the assumptions of an ANOVA could be satisfied) are given above each bar plot. Provided a significant ANOVA or Kruskal-Wallis test result was obtained, the significance between groups as determined by post-hoc testing is indicated by letters above the bars. Any two bars sharing the same letter are not significantly different to each other.

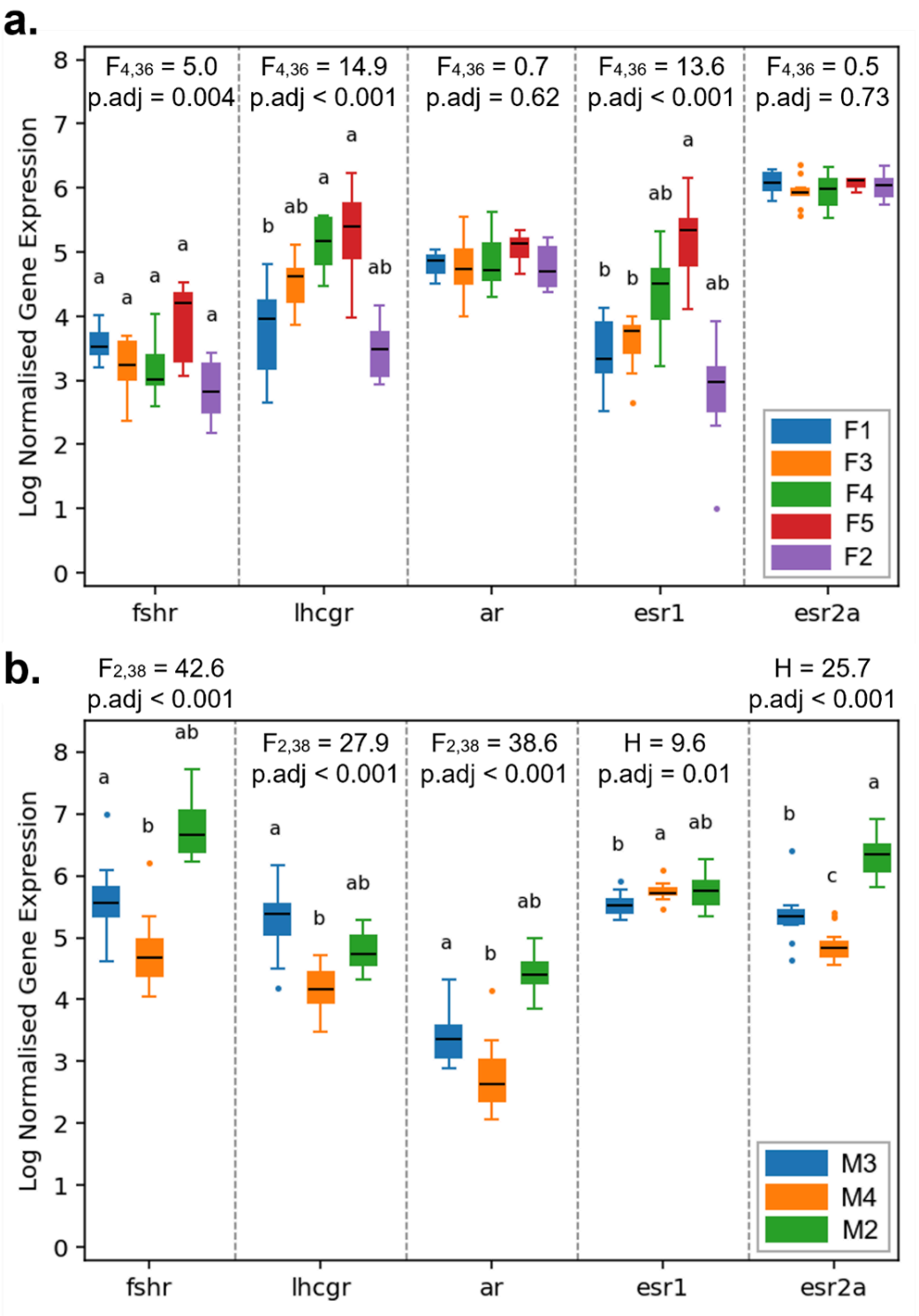


Figure 6

Gene expression box plots for genes associated with hormone reception in the gonads in females (a.), and males (b.). Bars are grouped by gene and the colour of the bars relates to the gametogenic stage: red = “immature” (F1) for females and “developing” (M3) for males, blue = “developing” (F3) for females and “developed” (M4) for males, orange = “developed” (F4, vitellogenic) for females and “resting post-spawning” (M2) for males, green = “spawning” (F5) in females and purple = “resting post-spawning” (F2) for females. The post Benjamini–Hochberg p-value correction results of ANOVA or Kruskal-Wallis tests (depending on whether the assumptions of an ANOVA could be satisfied) are given above each bar plot. Provided a significant ANOVA or Kruskal-Wallis test result was obtained, the significance between groups as determined by post-hoc testing is indicated by letters above the bars. Any two bars sharing the same letter are not significantly different to each other.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplimentary.zip