CHAPTER III

Transcriptome-Based SNP Discovery Revealed SNPs Associated with Sex Determination and with RNA Editing Involved in Sex Differentiation in Snakeskin Gourami (*Trichopodus pectoralis*)

3.1 Abstract

Sex-related markers in fish would provide a useful tool for both ecological and aquaculture research, particularly when information on sex chromosomes in fish remains limited. This study aimed to develop sex-related markers in the snakeskin gourami (*Trichopodus pectoralis*), which is an important freshwater fish showing female-biased growth. Transcriptome-based SNP discovery in ovaries and testes of juvenile and adult individuals was performed, and 21 candidate sex-specific RNA SNPs were identified. Using both Sanger sequencing and PACE SNP analysis in cDNA samples, 9 nucleotide variants were observed in *ift172*, *man2a1*, *crk*, *ezh2*, *cs*, *cast*, and *kap3l,* which exhibited sex-specific RNA SNPs with heterozygosity in males and homozygosity in females. In addition, *ift172*, *man2a1*, and *kap3l* demonstrated male-biased expression while *crk*, *ezh2*, and *cs* were female-biased transcripts, suggesting their involvement in gonadal differentiation and maturation. Three SNPs (SNP_19, SNP_20, and SNP_21) showed consistent DNA variants in the genome, providing useful markers for genotypic sex identification. The other 6 RNA SNPs (SNP_2, SNP_3, SNP_4, SNP_5, SNP_12, and SNP_15) were not detected in the genome. In addition, male-biased expression of *adar* was observed, which suggested that RNA editing may contribute to sex differentiation in the snakeskin gourami. A mismatch-sensitive T7EI assay validated SNP_19 and SNP_20, which provided a reliable and cost-effective approach to sex-specific DNA marker development. Overall, this study demonstrated and validated the sex-specific SNPs that could offer a useful tool for further application of sex identification in the snakeskin gourami.

Keywords: Sex determination, SNP, RNA editing, Transcriptome, Snakeskin gourami (*Trichopodus pectoralis*).

3.2 Introduction

Among vertebrates, teleost fish exhibit a wide variety of sex determination mechanisms, which may be influenced by genetic sex determination (GSD), environmental sex determination (ESD), or their interaction (Devlin & Nagahama, 2002). While several fish species are sexually dimorphic in size, body colour, or behaviour, others display no observable differences between males and females (Mesnick & Ralls, 2018; Uba, 2019; Han et al., 2020). Although the effects of ESD on sex differentiation cannot be excluded, genetic differences between males and females under GSD may serve as valuable tools in both ecological research and aquaculture investigations (Rajendiran et al., 2021). However, identification of sex chromosomes or sex-determining genes in fish remains limited (Kobayashi et al., 2013; Wang et al., 2024). As an alternative, sex-related markers have been widely investigated as practical tools in ecological and aquaculture research (Martinez et al., 2026; L. Zhang et al., 2025).

In species lacking external sexual dimorphism, molecular sex markers are especially useful for monitoring sex ratios in natural populations, thereby supporting ecological and conservation efforts (Vu et al., 2019). Temperature-dependent sex determination (TSD) may lead to male-biased populations under climate change, potentially disrupting population dynamics and threatening ecological balance and biodiversity (Geffroy & Wedekind, 2020; Ospina-Álvarez & Piferrer, 2008). Thus, sex-related markers can be used to track population-level sex ratios and assess ecological responses to thermal stress. In aquaculture, where external sexual dimorphism is often absent, sex-related markers facilitate the identification of sex, aiding artificial breeding and selective breeding programs (Abaho et al., 2025; Curzon et al., 2022). Conversely, in species that display sexual dimorphism at maturity (Kabpha et al., 2023; Mu et al., 2024), these markers are valuable for early-stage sex identification. Particularly in indirect methods of sex reversal, neomales (XX individuals phenotypically male) can be used as broodstock to produce all-female offspring (Luckenbach et al., 2017; Mu et al., 2024). In such cases, a sex-related marker is essential for identifying genotypic sex beyond relying solely on progeny sex ratios. Therefore, sex-related markers are indispensable tools in both ecological research and aquaculture, particularly where information on fish sex chromosomes remains limited.

Advances in bioinformatics and molecular genetics have facilitated the discovery of sex-specific DNA markers in the fish genome (Chen et al., 2007, 2015; Hua et al., 2023;

Vale et al., 2014; L. Zhang et al., 2025). However, other regulatory mechanisms—such as RNA editing, epigenetics, and alternative splicing—may also influence sex determination and differentiation by altering transcript-level nucleotide variants and expression levels (Lu et al., 2022; Wang et al., 2020).

Transcriptome analysis using high-throughput RNA sequencing (RNA-Seq) has emerged as a powerful method for quantifying and identifying RNA molecules in biological samples. RNA-Seq also enables identification of single-nucleotide polymorphisms (SNPs), which are valuable for discovering genetic markers (Zhang et al., 2024). Notably, transcriptome-based SNP discovery can yield putative RNA markers without requiring prior genomic sequence information. In addition, SNPs are bi-allelic, reproducible, and readily incorporated into high-throughput genotyping platforms such as PCR Allele Competitive Extension (PACE), allowing for scalable and accurate analyses (Torres et al., 2025). Furthermore, transcriptomic datasets have successfully enabled SNP marker development in species such as the black-spotted frog (*Pelophylax nigromaculatus*), black-faced blenny (*Tripterygion delaisi*), and various catfish species (Schunter et al., 2014; Wang et al., 2013; Zhang et al., 2024).

The snakeskin gourami (*Trichopodus pectoralis*) is an economically important freshwater fish, widely distributed and cultured across Southeast Asia. Females are generally larger than males; thus, feminisation using 17β-estradiol-supplemented diets has been employed to promote all-female production (Kabpha et al., 2023). Alternatively, indirect all-female production using neomales may offer a sustainable and food-safe approach for improved productivity. Despite its economic significance, the sex-determination mechanism in the snakeskin gourami remains poorly understood, and no sex-specific DNA markers have been reported to date. Previous transcriptomic analyses using RNA-Seq have identified differentially expressed genes (DEGs) between testes and ovaries, shedding light on molecular pathways involved in sex differentiation and maturation in both juvenile and adult stages (Boonanuntanasarn et al., 2020; Jantawongsri et al., 2025). However, no sex-specific molecular markers have yet been identified. In this study, we applied a transcriptome-based SNP discovery approach using juvenile and adult fish to identify and validate candidate sex-specific RNA SNPs. RT-PCR followed by Sanger sequencing and PACE SNP analysis was used for marker validation. In addition, by integrating transcriptomic and genomic data, we characterised sex-associated nucleotide variants at both RNA and DNA levels. We further demonstrated the utility of a mismatch-

sensitive T7 Endonuclease I (T7EI) assay to validate DNA-level SNP genotypes. Moreover, we provide the first evidence that RNA editing may play a role in the sex determination mechanism of the snakeskin gourami.

3.3 Materials and methods

3.3.1 Ethics statement

The study protocols for fish care and experimental procedures were reviewed and approved by the Kasetsart University Institutional Animal Care and Use Committee (ACKU 61-FIS-004) and conducted in accordance with institutional and national guidelines.

3.3.2 Experimental design, experimental fish, and sample collection

To discover sex-specific SNP markers in snakeskin gourami, RNA sequencing was performed to determine SNPs associated with sex differences in juvenile (sex differentiation stage) and adult (maturation stage). The transcriptomic resources for comparison between the testes and ovary in juvenile fish and adult fish used in this study were obtained from previously published datasets by Boonanuntanasarn et al. (2020) and Jantawongsri et al. (2025), respectively. Snakeskin gourami (*T. pectoralis*) were collected from eight commercial farms (Juvenile; 4 farms (replications), adult; 4 farms (replications) located in Samutr Sakhon, Thailand. Juvenile males (body weight, 19.22 ± 0.23 g; gonad weight, 0.01 ± 0.001 g; gonadosomatic index [GSI], 0.05 ± 0.003 %), females (body weight, 18.68 ± 0.39 g; GSI, 0.73 ± 0.03%) were used for sampling of testes (n = 50/replication, 4 replications) and ovary (n = 25/replication, 4 replications), respectively. In adult fish (harvestable size), testes (n = 10/replication, 4 replications) and ovary (n = 10/replication, 4 replications) were collected from males (body weight, 217.17 ± 9.89 g; GSI, 0.05 ± 0.005%) and females (body weight, 226.94 ± 7.69 g; GSI, 4.17 ± 0.31%). All fish were euthanized using tricaine methanesulfonate (MS-222; 150 mg/L) before dissection. Histological examination was performed to confirm gonadal sex (Boonanuntanasarn et al. 2020; Jantawongsri et al. 2025).

3.3.3 RNA sequencing, read pre-processing, *de novo* transcriptome assembly, and SNP Identification

Total RNA was isolated from approximately 100 mg of gonadal tissue (testis or ovary) using TRIzol™ reagent (Invitrogen, Carlsbad, CA, USA) and digested with

RNase-free DNase (Promega, Madison, WI, USA) according to the manufacturer's instructions. The concentration and purity of extracted RNA were evaluated using a NanoDrop™ ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and RNA integrity was assessed by 1% agarose gel electrophoresis. Further quality control was performed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) to determine the RNA integrity number (RIN). For downstream transcriptome analysis, testicular RNA samples with RIN ≥ 7.0 were selected. In contrast, ovarian RNA samples with RIN ≥ 4.5 were accepted due to the naturally high abundance of low-molecular-weight 5S rRNA observed in juvenile ovary tissues.

For transcriptome analysis in juvenile fish, four replications of testicular RNA (pooled equal amounts of RNA from fifty testes in one replication) and ovarian RNA (pooled equal amounts of RNA from twenty-five ovaries in one replication) were used for transcriptome analysis. In adult fish, four replications of testicular RNA (pooled equal amounts of RNA from ten testes) and ovarian RNA (pooled equal amounts of RNA from ten ovaries) were used. For each replicate, 2 µg of total RNA was used to construct each RNA library for juvenile ovaries (JO1, JO2, JO3, and JO4) juvenile testis (JT1, JT2, JT3, and JT4), adult ovaries (AO1, AO2, AO3, and AO4) or adult testis (AT1, AT2, AT3, and AT4) as described previously (Boonanuntanasarn et al. 2020; Jantawongsri et al. 2025). All libraries were sequenced on an Illumina HiSeq 2500 platform at the Novogene Bioinformatics Institute (Beijing, China) following the manufacturer's instructions.

Raw Illumina HiSeq 2500 sequencing data were base-called into FASTQ files and subjected to quality filtering. Clean reads were assembled *de novo* using Trinity (version 2.0.6 for juvenile and version r2014-04-13p1 for adult) with default parameters (k-mer = 25; minimum contig length = 200 bp) (Grabherr et al., 2011). Redundant transcripts were hierarchically clustered using Corset v1.05 (Davidson and Oshlack, 2014). The longest transcript from each cluster was designated as a unigene.

The *de novo* transcriptome assembly is provided as a reference for reading mapping. The clean reads were aligned back to the assembled transcriptome using Bowtie2 (Langmead and Salzberg, 2012). Read sorting and screening were performed with SAMtools and Picard, followed by SNP calling using the Genome Analysis Toolkit (GATK v3). In Juvenile fish, sex associated nucleotide variants between JO1-4 and JT1-4 were explored. In addition, nucleotide variants between AO1-4 and

AT1-4 were screened using transcriptome information in adult fish. SNPs exhibiting consistent sex-linked polymorphisms across juvenile and adult groups were identified and designated as putative sex-specific RNA SNP markers for subsequent validation.

3.3.4 Validation of candidate RNA SNPs and their corresponding DNA variants using Sanger method

3.3.4.1 Sample collection and histological study

To validate the sex-specific RNA SNP markers, gonadal samples (testes and ovaries) from 10 juvenile fish including five females (body weight = 18.4 ± 2.3 g, gonad weight = 0.15 ± 0.1 g; GSI = 0.01 ± 0.01 %) and five males (body weight = 18.6 ± 2.6 g, gonad weight = 0.01 ± 0.01 g; GSI = 0.0008 ± 0.0004 %) and 10 adult fish including five females (body weight = 225.8 ± 24.22 g, gonad weight = 17.48 ± 7.35 g; GSI = 0.08 ± 0.04 %) and five males (body weight = 132.2 ± 9.73 g, gonad weight = 0.18 ± 0.05 g; GSI = 0.001 ± 0.0003 %) were collected from farms in Samut Sakhon, Thailand. Before sampling, all fish were euthanized with 150 mg/L of MS-222, and caudal fins were also collected for genomic DNA extraction. Gonadal tissues (ovary or testis) were collected and divided into two portions: one was immediately snap-frozen in liquid nitrogen and stored at −80°C for RNA extraction, while the other was fixed in Bouin's solution at 4°C for 16–18 h. Fixed tissues were then transferred to 80 % ethanol and stored at 4°C until histological processing for sex identification.

Histological analysis was performed to confirm sex and the presence of the developmental stage of testes or ovaries. Fixed gonadal tissues were dehydrated through a graded ethanol–xylene series, embedded in paraffin, sectioned transversely at 5 µm thickness, and stained with hematoxylin and eosin (H&E) according to the protocol described previously (Kabpha et al., 2023).

3.3.4.2 Total RNA extraction, reverse transcription PCR, and sequencing

To validate SNPs identified from RNA-seq, total RNA was extracted from approximately 50-100 mg of gonadal tissue (juvenile: 5 ovaries, 5 testes; adult: 5 ovaries, 5 testes; n = 20) using the same procedure described above for RNA isolation. RNA quantity and integrity were assessed as previously outlined. Subsequently, 2 µg of RNA was reverse-transcribed into first-strand cDNA using the SuperScript™ III Reverse Transcriptase Kit (Invitrogen) with random hexamer (Promega), following the manufacturer's instructions.

According to screening sex-specific SNPs in juvenile and adult fish, a total of 21 putative sex-specific SNP candidates consistently identified across juvenile and adult gonadal transcriptomes were selected for validation (Table 3.1). Primer pairs were designed using Primer3Plus (Table 3.2). Reverse transcription PCR (RT-PCR) amplification was performed using TaKaRa Ex Taq polymerase (TaKaRa Bio, Shiga, Japan) in a total reaction volume of 25 µL. Each reaction mixture contained 2.5 µL of 10× Ex Taq Buffer, 2 µL of dNTP Mixture (2.5 mM each), 0.125 µL of TaKaRa Ex Taq polymerase (5 U/µL), 2.5 µL of forward primer (2 µM), 2.5 µL of reverse primer (2 µM), 2.5 µL of 20× diluted cDNA template, and 12.875 µL of nuclease-free water. RT-PCR cycling was conducted under the following conditions: 95°C for 3 min; 35 cycles of 95°C for 30 s, annealing at the optimal temperature for each primer pair (55–65°C) for 30 s, and extension at 72°C for 25 s; followed by a final extension at 72°C for 10 min. The target PCR products were excised from agarose gels and purified using the QIAEX® II Gel Extraction Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol for Sanger sequencing. Only SNPs in which females consistently exhibited homozygous genotypes and males heterozygous genotypes in cDNA with transcriptome-based SNPs were retained for further validation using the PACE SNP genotyping assay. In addition, their functional annotation of the assembled transcriptome and the differentially expressed levels, presented as $\log_2$ fold change (male/female) values, was conducted following the procedures described in Boonanuntanasarn et al. (2020) and Jantawongsri et al. (2025).

Table 3.1 Sex-associated transcriptome-based SNPs in juvenile and adult snakeskin gourami.

SNPs	Clusters	Location in mRNA	Genotypes Female	Male
SNP_1	Cluster-22178.58277	3′-UTR	A/A	A/G
SNP_2	Cluster-22178.80501	3′ UTR	A/A	A/G
SNP_3	Cluster-22178.68092	3′ UTR	A/A	A/G
SNP_4	Cluster-22178.66311	3′ UTR	A/A	A/G
SNP_5	Cluster-22178.77480	CDS	A/A	A/G
SNP_6	Cluster-22178.67537	3′ UTR	A/A	A/G
SNP_7	Cluster-22178.36261	3′ UTR	A/A	A/G
SNP_8	Cluster-22178.64482	CDS	A/A	A/G
SNP_9	Cluster-22178.72264	3′ UTR	A/A	A/G

Table 3.1 (Cont.)

SNPs	Clusters	Location in mRNA	Genotypes Female	Male
SNP_10	Cluster-22178.76218	3′ UTR	C/C	C/A
SNP_11	Cluster-22178.86555	3′ UTR	A/A	A/G
SNP_12	Cluster-22178.67781	3′ UTR	A/A	A/G
SNP_13	Cluster-22178.82477	3′ UTR	A/A	A/G
SNP_14	Cluster-22178.67640	CDS	A/A	A/G
SNP_15	Cluster-22178.56828	3′ UTR	A/A	A/G
SNP_16	Cluster-22178.90734	CDS	A/A	A/G
SNP_17	Cluster-22178.75619	5′ UTR	T/T	T/A
SNP_18	Cluster-22178.91475	3′ UTR	G/G	G/C
SNP_19	Cluster-22178.57271	3′ UTR	G/G	G/T
SNP_20	Cluster-22178.57271	3′ UTR	C/C	C/T
SNP_21	Cluster-22178.57271	CDS	G/G	G/A

See nucleotide sequences in Appendix.

3.3.4.3 Genomic DNA extraction, DNA variant analysis using Sanger sequencing

Genomic DNA (gDNA) was extracted from approximately 100 mg of fin tissue collected from juvenile (5 females and 5 males) and adult (5 females and 5 males) snakeskin gourami using the standard phenol–chloroform extraction method. DNA concentration and purity were measured using a NanoDrop® spectrophotometer, and DNA integrity was verified by electrophoresis on a 0.7% agarose gel.

Candidate putative sex-specific RNA SNPs that were consistently detected in both juvenile and adult transcriptomes were selected for genomic validation. Since SNP_21 contained an intron part, another pair of primers was designed for Sanger sequencing and PACE SNPs genotyping assay (Table 3.2). A total of 20 individuals, identical to the cDNA validation sample (juveniles: 5 females and 5 males; adults: 5 females and 5 males), were tested using PCR. The PCR condition was performed as described above for cDNA validation. Subsequently, the target PCR products were excised from agarose gels and purified using the QIAEX® II Gel Extraction Kit (Qiagen) following the manufacturer's protocol for Sanger sequencing.

Table 3.2 Primers used for validating sex-specific SNP markers.

Name	Clusters	Primers (5′-3′)	Tm (°C)	Products (bp)	Application
SNP_1	Cluster-22178.58277	F: AATACGACTCACTATAGGGCGTGAAGGAACACGCGGAG R: TCAGCACATTTCCTGCCAGCTACA	65	517	Sanger sequencing
SNP_2	Cluster-22178.80501	F: CAGGCTTAGTGGTAGCTACACA R: GCATTGGTCGTTCAGTCCGA	60	308	Sanger sequencing
SNP_3	Cluster-22178.68092	F: CCGCATAGAGATCGACTGAAT R: TAAGTGTGGCCTTCCTGGTG	58	307	Sanger sequencing
SNP_4	Cluster-22178.66311	F: CCGGGTGAGGTTGTCAGTTT R: TCACAGTGCTCAGCTCTCAC	60	300	Sanger sequencing
SNP_5	Cluster-22178.77480	F: CAGCCCTGTGACTCCTCTTG R: AACCAGCCACGTCTGAAGG	60	289	Sanger sequencing
SNP_6	Cluster-22178.67537	F: AATACGACTCACTATAGGGGGTTTGGCTAACTTTGCA R: GTTCTGCAACACATGGCAAC	60	619	Sanger sequencing
SNP_7	Cluster-22178.36261	F: AATACGACTCACTATAGGGACCTGATCTGTTCAGTGGCCTCCC R: GTGCACCGAGTGAGCCTCCG	65	665	Sanger sequencing
SNP_8	Cluster-22178.64482	F: AATACGACTCACTATAGGGACGCCCTCGCTTTCCCCC R: GTCTGAGAAATCGACAGTCTGCTGGC	65	236	Sanger sequencing
SNP_9	Cluster-22178.72264	F: GGGAGTTTAAGAAACATATGGAAGCG R: CATGCTGCTGCTCCCCAG	60	289	Sanger sequencing

Table 3.2 (Cont.).

Name	Clusters	Primers (5′-3′)	Tm (°C)	Products (bp)	Application
SNP_10	Cluster-22178.76218	F: AATACGACTCACTATAGGGCAGAAGTCAAAGGTACCATACAGGCT R: TGTCAACATGTGTTACTGGTCCCAGAC	65	519	Sanger sequencing
SNP_11	Cluster-22178.86555	F: AATACGACTCACTATAGGGAGTCTTGAACTGTCACGCATGCTC R: CAGGTATGAGAGTGGTGTCTAGTCAGT	65	518	Sanger sequencing
SNP_12	Cluster-22178.67781	F: ACTCTATGAACACTCCACTTTAAAAGT R: TGTACATTACTGTTGCTGTACTTCA	58	315	Sanger sequencing
SNP_13	Cluster-22178.82477	F: TCCTGAATGCCTTTGTGTTTTT R: ACTGGCATCATGTGTGACCT	58	317	Sanger sequencing
SNP_14	Cluster-22178.67640	F: CAGCCTCCAGTTTCCTGTCC R: GGAGTAAACCCGTTCCGACC	60	319	Sanger sequencing
SNP_15	Cluster-22178.56828	F: CTGCTCACCTCCTTGCATGT R: TGTAAACCACGTGCTGTCCA	60	313	Sanger sequencing
SNP_16	Cluster-22178.90734	F: TGAAGAGGAAGCATCAGGATCC R: TTTTTCCCAAGGTTGCCACC	59	200	Sanger sequencing
SNP_17	Cluster-22178.75619	F: AATACGACTCACTATAGGGTGCCAGGCTCCCCTGTCTCC R: AAGCTTCCAGGAAACCGACAGTGA	65	518	Sanger sequencing
SNP_18	Cluster-22178.91475	F: AGCTGGCGGCAGATTTCTAA R: ACACCAGAAACACCTCAGGA	60	322	Sanger sequencing

Table 3.2 (Cont).

Name	Clusters	Primers (5′-3′)	Tm (°C)	Products (bp)	Application
SNP_19, 20	Cluster-22178.57271	F: AATACGACTCACTATAGGGCCAGGGCTGTAACTCTGGACT R: GCACCTTGATCAAAGCACTCTCCA	63	551	Sanger sequencing
SNP_19, 20	Cluster-22178.57271	F: AATACGACTCACTATAGGGCCAGGGCTGTAACTCTGGACT R: GCACCTTGATCAAAGCACTCTCCA	63	551	Sanger sequencing
SNP_21 (cDNA)	Cluster-22178.57271	F:AATACGACTCACTATAGTCTACAACTTGCCCGCAACC R:CCACCAGAACCAGAAGCTCC	60	520	Sanger sequencing
SNP_21 (gDNA)	Cluster-22178.57271	F: GAGACAGCTCTAGGGGCTCT R: TGCTCCTGTTTGGCCAGAAG	60	518	Sanger sequencing
SNP_2 (PACE)	Cluster-22178.80501	CRP:CTGTTGATGGATCTGCTATAGAACTTCAT ASP-A:GAAGGTGACCAAGTTCATGCTGTAATTATTATTACTGTCATCAAATCC**T** ASP-G:GAAGGTCGGAGTCAACGGATTCTGTAATTATTATTACTGTCATCAAATCC**C**	60	109 (ASP-A) 111 (ASP-G)	PACE genotyping
SNP_3 (PACE)	Cluster-22178.68092	CRP:GAAAGACATCCAGACACCATGACTGAA ASP-A:GAAGGTGACCAAGTTCATGCTCATCCACATGAGGTGATCCAGA**A** ASP-G:GAAGGTCGGAGTCAACGGATTCATCCACATGAGGTGATCCAGA**G**	60	73	PACE genotyping
SNP_4 (PACE)	Cluster-22178.66311	CRP:AAGGAACAGACTGACATTCTGGAAAAGAT ASP-A:GAAGGTGACCAAGTTCATGCTATTGTACTTATCTCACTCAAGAGCAAAT**A** ASP-G:GAAGGTCGGAGTCAACGGATTGTACTTATCTCACTCAAGAGCAAAT**G**	60	95 (ASP-A) 92 (ASP-G)	PACE genotyping

Table 3.2 (Cont).

Name	Clusters	Primers (5′-3′)	Tm (°C)	Products (bp)	Application
SNP_5 (PACE)	Cluster-22178.77480	CRP:CACACTCCCTCACTGCCAGGTA ASP-A:GAAGGTGACCAAGTTCATGCTGTTGCAAAGCCCAGTGTAACACT**A** ASP-G:GAAGGTCGGAGTCAACGGATTGCAAAGCCCAGTGTAACACT**G**	60	81 (ASP-A) 78 (ASP-G)	PACE genotyping
SNP_12 (PACE)	Cluster-22178.67781	CRP:ACAACTTTGAAAAAAGGGAGACCAAAGATT ASP-A:GAAGGTGACCAAGTTCATGCTCCATAGTTAAAAGATACTGCCCCTTT**A** ASP-G:GAAGGTCGGAGTCAACGGATTCATAGTTAAAAGATACTGCCCCTTT**G**	60	78 (ASP-A) 77 (ASP-G)	PACE genotyping
SNP_15 (PACE)	Cluster-22178.56828	CRP:CAGAGATCATCTCAGCTGGCAGTTT ASP-A:GAAGGTGACCAAGTTCATGCTGCCATCTTTCCAACAAACCTCAC**T** ASP-G:GAAGGTCGGAGTCAACGGATTCCATCTTTCCAACAAACCTCAC**C**	60	70	PACE genotyping
SNP_19, 20 (PACE)	Cluster-22178.57271	CRP:CTACTAATGCATTAAATGTCCTAAGCTCCC ASP-GC:GAAGGTGACCAAGTTCATGCTAATATTGAGCATTTACAATACTGGACAT**GC** ASP-TT:GAAGGTCGGAGTCAACGGATTAATATTGAGCATTTACAATACTGGACAT**TT**	60	100	PACE genotyping
SNP_21 (cDNA, (PACE)	Cluster-22178.57271	CRP:CACCTATTTTATAATGACTCACCACTCCAT ASP-G:GAAGGTGACCAAGTTCATGCTTTCTGCTTCTCTAGTTTCTCCCA**G** ASP-A:GAAGGTCGGAGTCAACGGATTTTTCTGCTTCTCTAGTTTCTCCCA**A**	60	55	PACE genotyping

Table 3.2 (Cont).

Name	Clusters	Primers (5′-3′)	Tm (°C)	Products (bp)	Application
SNP_21 (gDNA, (PACE)	Cluster-22178.57271	CRP:ATAGTGTGGAGCTGGCCACCATTAT ASP-G:GAAGGTGACCAAGTTCATGCTTATAATGACTCACCACTCCATGAAA**C** ASP-A:GAAGGTCGGAGTCAACGGATTTTATAATGACTCACCACTCCATGAAA**T**	60	195 (ASP-G) 196 (ASP-A)	PACE genotyping

Abbreviation: Tm (°C) =Annealing temperature (°C); bp= base pair; ASP=Allele-specific forward primers; CRP=Common reverse primer; PACE=PCR Allele Competitive Extension.

Note: The underlined sequence (AATACGACTCACTATAGGG) corresponds to the T7 universal sequencing tag, which was added at the 5′ end of primers to facilitate sequencing. Underline indicates the universal FAM (GAAGGTGACCAAGTTCATGCT) and HEX (GAAGGTCGGAGTCAACGGATT) tails added to the 5′ end of allele-specific primers.

3.3.5 Validation of candidate RNA SNPs and their corresponding DNA variants using PACE SNP analysis

To strengthen the validation of sex-specific RNA SNP markers across a broader population, an additional cohort of 20 fish was collected from farms in Samut Sakorn and the Suranaree University of Technology Farm (SUT Farm), Thailand. The sample included 10 juveniles (5 females and 5 males) and 10 adults (5 females and 5 males); detailed morphometric and gonadosomatic index (GSI) data are summarized in Table 3.3. All fish were euthanized with an overdose of MS-222 (150 mg/L) before tissue collection. Caudal fin clips were collected from each individual and preserved for gDNA extraction. Gonadal tissues (ovaries or testes) were dissected, immediately snap-frozen in liquid nitrogen, and stored at -80°C until RNA extraction. Total DNA and RNA were extracted from the fin and gonad tissues, respectively, following previously described methods. Subsequently, cDNA was synthesized from the RNA extracts. For the PACE SNP analysis, allele-specific primers were designed using the PACE® Assay Design Template (3CR Bioscience Ltd., UK). Each assay consisted of two allele-specific forward primers, each carrying a universal FAM or HEX fluorescent tail, and one common reverse primer. Each assay consisted of two allele-specific forward primers (ASPs) carrying universal FAM or HEX fluorescent tails and one common reverse primer (CRP) (Table 3.2). SNP genotyping was conducted using the PACE® 2.0 Genotyping Master Mix (3CR Bioscience, Harlow, UK) following the manufacturer's instructions. The reaction (a total volume of 10 µL) consisted of 5 µL of 2× PACE 2.0 Genotyping Master Mix, 0.138 µL of custom-designed Assay Mix (containing two allele-specific forward primers (12 µM) and one common reverse primer (30 µM)), and 25 ng of cDNA template. PCR amplification was performed under the following cycling conditions: initial enzyme activation at 94°C for 15 min; followed by 10 cycles of 94°C for 20 s and 65–57°C for 60 s (decreasing 0.8°C per cycle); and then 30 cycles of 94°C for 20 s and 57°C for 60 s. After amplification, the fluorescence signals of FAM and HEX were recorded at or below 40°C in endpoint mode. The genotypes were determined based on the relative FAM and HEX fluorescence intensities, and cluster plots were generated for allele discrimination. The no-template control (NTC) wells showed no detectable fluorescence in either channel, confirming the absence of cross-contamination.

To further verify the sex-specific SNPs at the genomic level, the PACE SNP genotyping assay was performed using the same gDNA samples from 20 individuals. The PACE® reaction setup, primer composition, PCR cycling conditions, fluorescence signal detection, and genotype determination followed the procedures described for cDNA validation. Genotypic profiles were then compared between male and female fish to evaluate whether the SNPs identified in cDNA were also present at the genomic level.

Table 3.3 Morphometric and gonadal characteristics of the independent validation cohort.

Group	Sex	Body Weight (g)	Gonad Weight (g)	GSI (%)
Juvenile	Female (n=5)	14.4 ± 4.62	0.12 ± 0.03	0.01 ± 0.002
	Male (n=5)	18.6 ± 2.61	0.01 ± 0.01	0.0008 ± 0.0004
Adult	Female (n=5)	223.2 ± 25.09	18.67 ± 7.24	0.08 ± 0.04
	Male (n=5)	128.4 ± 27.19	0.19 ± 0.08	0.001 ± 0.0004

Note: Values are presented as mean ± SD (n = 5). GSI (gonadosomatic index) = (gonad weight/body weight) × 100. Juvenile females showed slightly higher GSI values than males, while adult females exhibited markedly higher gonad weight and GSI compared with males.

3.3.6 Validation of sex specific DNA SNPs in other fish populations

3.3.6.1 Sample collection and DNA variant analysis using Sanger sequencing

This study validated the sex-specific SNP markers (SNP_19, SNP_20, and SNP_21) in four fish populations (SUT1, SUT2, SUT3, and SUT4) that were obtained from the SUT farm. Adult fish that demonstrated obvious sexual dimorphism in body shape were anesthetized with 150 mg/L of MS-222. Caudal fin samples were collected from 100 adult fish, including fifty females (body weight: 211.84 ± 25.78 g; gonad weight: 18.43 ± 4.96 g; GSI: 0.09 ± 0.03 %) and fifty males (body weight: 124.24 ± 2.6 g; gonad weight: 0.17 ± 0.05 g; GSI: 0.001 ± 0.0003 %). Genomic DNA was extracted using fin tissues (100 mg) following the same procedure described above for gDNA

isolation. All 100 individuals (50 females and 50 males) were genotyped by PCR amplification and Sanger sequencing to confirm the SNP genotypes.

3.3.6.2 DNA variant genotyping using PACE SNP assay

To further validate the sex-specific SNP markers (SNP_19, SNP_20, and SNP_21) at the population level, genotyping was conducted using the PACE SNP assay. gDNA from 100 individuals (50 females and 50 males) previously analyzed by Sanger sequencing was used as templates. The PACE® reaction setup, primer composition, PCR cycling conditions, fluorescence detection, and genotype calling were performed following the same protocol used for cDNA validation. The NTC wells were included in each run to monitor potential contamination.

3.3.6.3 DNA variant genotyping using the T7 Endonuclease I assay

A PCR-based T7EI digestion for RNA SNP_19, SNP_20, and SNP_21 and their DNA variants was conducted in both male (n =10) and female (n =10). Purified PCR products (400 ng) from cDNA and gDNA (the same PCR products for sequencing) were used for analysis of SNP-induced heteroduplex formation using T7 Endonuclease I (New England Biolabs, USA) according to the manufacturer's instructions. The mixture was denatured at 95°C for 5 min and annealed using a two-step cooling protocol: from 95°C to 85°C at 2°C/s, then from 85°C to 25°C at 0.1°C/s, before being held at 4°C. Subsequently, 1 µL of T7EI enzyme was added to the reaction mixture (final volume = 20 µL) and incubated at 37°C for 15 min. The reaction was terminated by adding 1.5 µL of 0.25 M EDTA. Digested products were resolved on 1.5 % agarose gels by electrophoresis.

3.4 Results

3.4.1 Sex-specific SNPs identified from gonadal transcriptomes of juvenile and adult *T. pectoralis*

This study used transcriptome information of *T. pectoralis* comparing testes and ovaries between sexes in juvenile (JT1–4; 179,323,982 clean reads vs JO1–4; 190,792,984 clean reads) and adult stages (AT1–4; 174,826,234 clean reads vs AO1–4; 179,753,200 clean reads) (Table 3). *De novo* assembly and annotation yielded 50,517 unigenes for juveniles and 55,883 unigenes for adults, which were used as references for SNP discovery (Tabe 4). In total, 65,535 polymorphic SNPs associated with sex

differences were identified across 380 individuals (JO1-4: 100 fish, JT1-4: 200 fish, AO1-4: 40 fish, AT1-4: 40 fish) (Table S1; see Appendix). Comparative transcriptomic analysis revealed 24 putative sex-specific SNPs between JO1-4 and JT1-4 (Table S2; see Appendix) and 389 SNPs between AO1-4 and AT1-4 (Table S3; see Appendix). By integrating SNP information from both juvenile and adult stages, 21 candidate sex-specific RNA SNPs were obtained in *T. pectoralis*. These SNPs were distributed across different transcript clusters, and SNP positions were located within transcribed regions, including 5'-untranslated region (5'-UTR), coding DNA sequence (CDS), and 3'-untranslated region (3'-UTR) (Table 3.1).

3.4.2 Validation of candidate sex-specific RNA SNPs

This study validated the candidate sex-specific RNA SNPs using testicular and ovarian cDNA from juvenile (5 males and 5 females) and adult fish (5 males and 5 females), and a histological study was conducted to confirm testicular and ovarian tissues. Figure 3.1 shows the representative of immature ovary and testis in juvenile and mature ovary and testis in adult fish. In juveniles, ovaries contained oogonia (OG) and previtellogenic oocytes (PVO) (Figure 3.1A and 3.1B). In adults, ovaries exhibited a full spectrum of oocyte development, and there are various developmental stages of oocytes, including OG, PVO, early vitellogenic oocytes (EVO), and mature oocytes (MO) (Figure 3.1C and 3.1D). Juvenile testes displayed the typical lobular testes containing spermatogonia (SG), primary spermatocytes (PSC), secondary spermatocytes (SSC), and a low proportion of spermatid (ST) (Figure 3.1E and 3.1F). In adults, testes exhibited developed spermatogenesis, with lobules characterized by the presence of SG, PSC, SSC, and a high proportion of ST (Figure 3.1G and 3.1H). These ovaries and testes were used to validate the candidate sex-specific RNA SNPs.

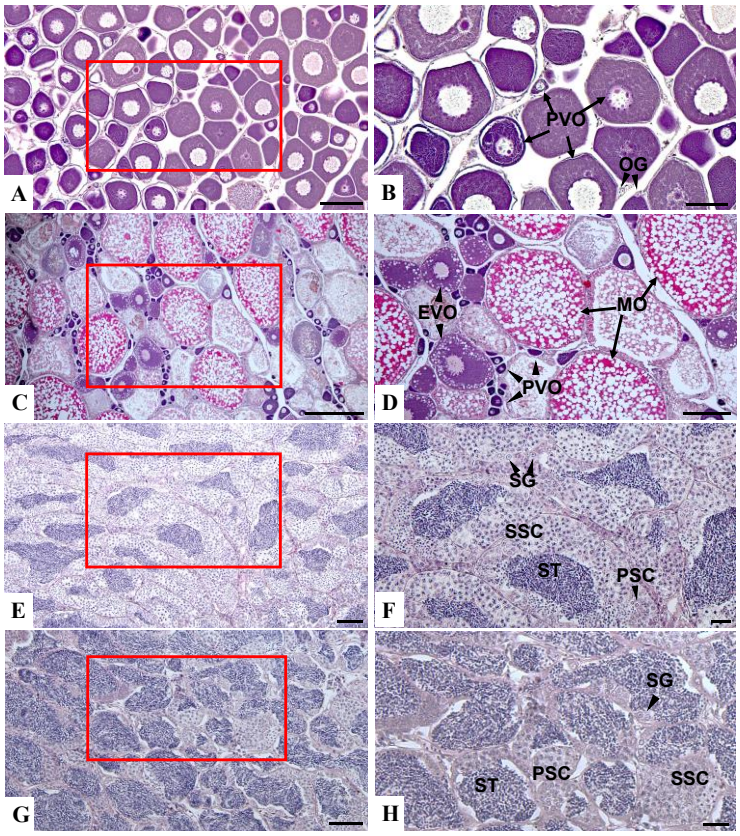

Figure 3.1 Histological characterization of juvenile and adult gonads in snakeskin gourami. The left shows representative histological sections of juvenile ovary (A), adult ovary (C), juvenile testis (E), and adult testis (G). Panels of right (B, D, F, H) represent higher magnification views of the boxed regions in A, C, E, G, respectively. Ovarian cell types observed include oogonia (OG), previtellogenic oocyte (PVO), early vitellogenic oocytes (EVO), and maturing oocyte (MO). Testicular cell types include spermatogonia (SG), primary spermatocytes (PSC), secondary spermatocytes (SSC), and spermatids (ST). Scale bars: 100 μm (A, C), 50 μm (B, D, E, G), and 20 μm (F, H).

The RT-PCR followed by Sanger sequencing showed that among 21 candidate sex-specific RNA SNPs, RNA variants of nine SNPs, including SNP_2, SNP_3, SNP_4, SNP_5, SNP_12, and SNP_15, SNP_19, SNP_20, and SNP_21, exhibited concordant genotypes with those of transcriptome-based SNPs (Table 6; Figure 3.2). Males were heterozygous A/G, A/G, A/G, A/G, A/G, A/G, G/T, C/T, and G/A, respectively, whereas female samples presented homozygous A/A, A/A, A/A, A/A, A/A, A/A, G/G, C/C, and G/G, respectively

(Figure 3.2). Using the PACE assay, nine candidate SNPs (SNP_2, SNP_3, SNP_4, SNP_5, SNP_12, SNP_15, SNP_19, SNP_20, and SNP_21) exhibited genotype patterns consistent with sex, showing predominantly homozygosity in females and heterozygosity in males (Figure 3.3). Distinct fluorescence clusters were clearly separated for most loci (e.g., SNP_2, SNP_3, SNP_4, SNP_19, and SNP_20), while partial overlaps between male and female genotypes were observed at SNP_5, SNP_12, and SNP_15. These results were overall in agreement with those obtained from RT-PCR followed by Sanger sequencing.

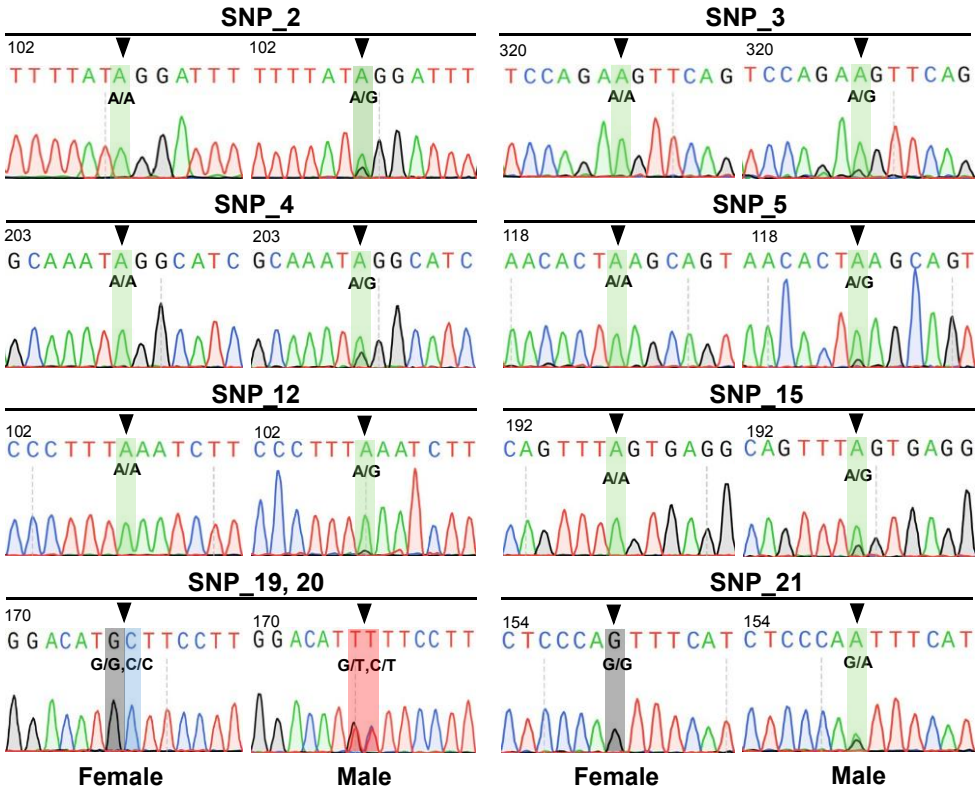

Figure 3.2 Validation of sex-associated SNPs at the RNA level by Sanger sequencing in snakeskin gourami. Representative chromatograms from RT-PCR amplified cDNA show sex-specific genotypes of SNP_2, SNP_3, SNP_4, SNP_5, SNP_12, and SNP_15 in which females consistently exhibited homozygous A/A genotypes, whereas males were heterozygous A/G. For SNP_19, SNP_20, and SNP_21, male chromatograms displayed overlapping peaks corresponding to heterozygous genotypes (e.g., G/T, C/T, or G/A), while female chromatograms showed uniform homozygous bases (e.g., G/G or C/C).

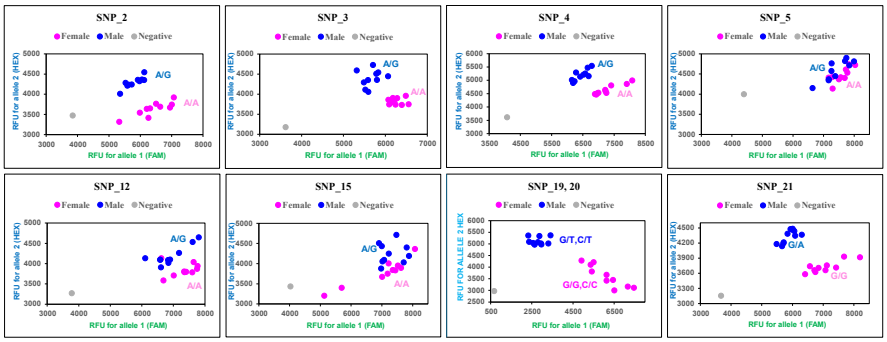

Figure 3.3 PACE genotyping analysis of nine SNPs using cDNA in snakeskin gourami. Scatter plots show fluorescence signal intensities (RFU) for allele 1 (FAM, x-axis) and allele 2 (HEX, y-axis). Each point represents an individual fish (pink = female; blue = male; gray = negative control). SNP_2, SNP_3, SNP_4, SNP_5, SNP_12, and SNP_15 exhibited female homozygosity (A/A) and male heterozygosity (A/G), whereas SNP_19 and SNP_20 showed homozygous G/G, C/C (female) and heterozygous G/T, C/T (male) genotypes, and SNP_21 displayed homozygous G/G (female) and heterozygous G/A (male) genotypes.

3.4.3 Differential expression and transcript annotation of 9 sex-associated SNPs in juvenile and adult gonads of *T. pectoralis*

Transcript annotation revealed that the nine validated SNPs (SNP_2, SNP_3, SNP_4, SNP_5, SNP_12, SNP_15, SNP_19, SNP_20, and SNP_21) were distributed across seven genes: *ift172*, *man2a1*, *crk*, *ezh2*, *cs*, *cast*, and *kap3l* (Table 3.4). Among these, SNP_5 and SNP_21 were located in CDS, whereas the others were situated in 3′UTRs (Table 3.4; Figure S1; see Appendix). Specifically, SNP_5 in *ezh2* resulted in a non-synonymous substitution (AAG⟶GAG; K⟶E), while SNP_21 in *kap3* lwas synonymous. The remaining 3′UTR SNPs—SNP_2 (*ift172*), SNP_3 (*man2a1*), SNP_4 (*crk*), SNP_12 (*cs*), SNP_15 (*cast*), SNP_19, and SNP_20 (*kap3l*) (Table 3.4).

Differential expression analysis revealed distinct sex-biased transcriptional patterns among six genes (Table 3.4). For instance, *ift172* (both juvenile and adult), *man2a1* (both juvenile and adult), and *kap3l* (adult only) exhibited male-biased expression. In contrast, *crk* (both juvenile and adult), *cs* (adult only), and *ezh2* (adult only) showed significantly higher expression in females than in males. Note that the *cast* showed no significantly different expression level (Table 3.4).

Table 3.4 Validated sex-associated SNPs and their differential expression in gonadal tissues of juvenile and adult snakeskin gourami.

SNPs	Gene name	Gene description	Genotypes		Juvenile			Adult		
			Female	Male	Sex presenting upregulation	Log2 fold change	q-value	Sex presenting upregulation	Log2 fold change	q-value
SNP_2	*ift172*	Intraflagellar transport protein 172	A/A	A/G	Male	4.323	<0.001	Male	3.829	<0.001
SNP_3	*man2a1*	mannosidase, alpha, class 2A, member 1	A/A	A/G	Male	1.024	0.035	Male	2.096	<0.001
SNP_4	*crk*	v-crk avian sarcoma virus CT10 oncogene	A/A	A/G	Female	1.158	<0.001	Female	0.953	<0.001
SNP_5	*ezh2*	enhancer of zeste 2 polycomb repressive complex 2 subunit	A/A	A/G	ns	-	-	Female	0.676	0.039
SNP_12	*cs*	citrate synthase	A/A	A/G	ns	-	-	Female	1.193	<0.001
SNP_15	*cast*	calpastatin	A/A	A/G	ns	-	-	ns	-	-
SNP_19			G/G	G/T	ns	-	-	Male	1.191	<0.001
SNP_20	*kap3l*	kinesin-associated protein 3	C/C	C/T	ns	-	-	Male	1.191	<0.001
SNP_21			G/G	G/A	ns	-	-	Male	1.191	<0.001

ns = non-significance.

See nucleotide sequence in Appendix.

3.4.4 Validation of candidate sex-specific DNA SNPs in the genome using Sanger sequencing, PACE SNP analysis

Following cDNA validation, genotyping of gDNA from 10 females and 10 males by PCR and Sanger sequencing demonstrated, among 9 sex-specific RNA SNPs, three SNPs, SNP_19 (G/T), SNP_20 (C/T), and SNP_21 (G/A) showed consistent sex-specific genotypes at the genomic level: all male individuals were heterozygous, whereas all female individuals were homozygous (Figure 3.4). Six SNPs (SNP_2, SNP_3, SNP_4, SNP_5, SNP_12, and SNP_15) exhibited no sex-specific SNPs in the genome, as all individuals were homozygous (A/A) (Figure 3.4). Consistently, PACE assays using gDNA also showed that SNP_19, SNP_20, and SNP_21 displayed male heterozygosity and female homozygosity, while SNP_2, SNP_3, SNP_4, SNP_5, SNP_12, and SNP_15 were homozygous in both sexes (Figure 3.5).

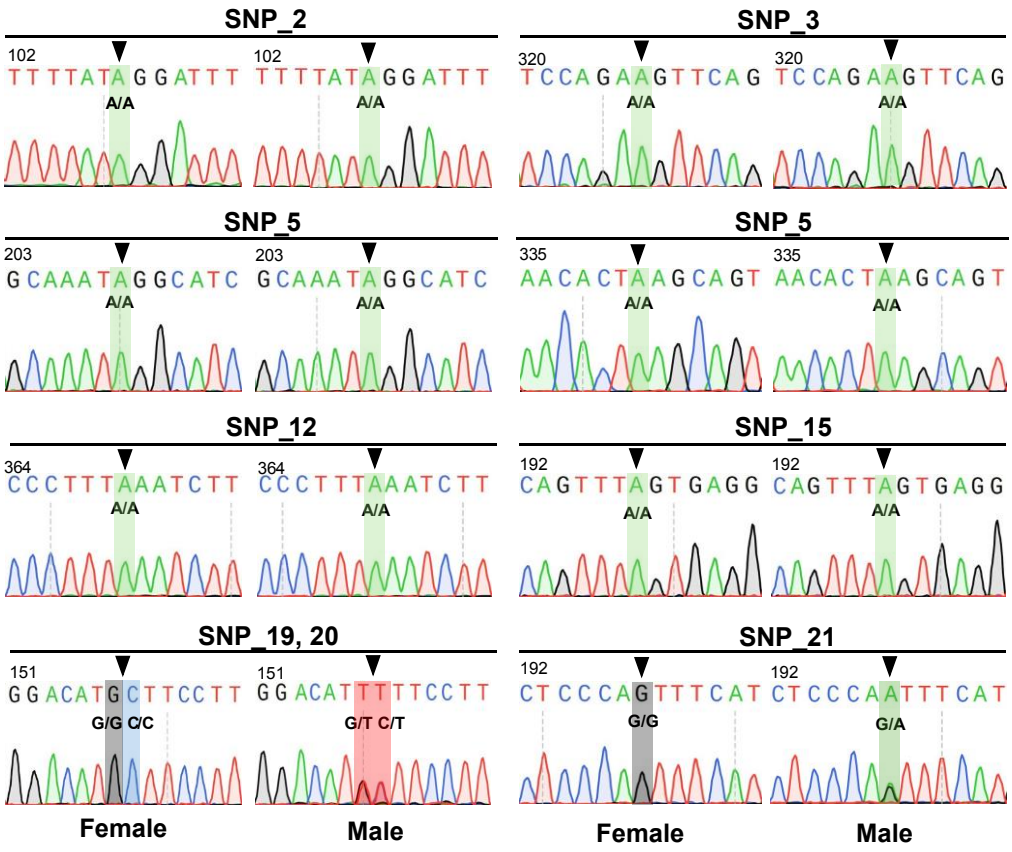

Figure 3.4 Validation of sex-specific SNPs in gDNA in snakeskin gourami using Sanger sequencing. Representative chromatograms are shown for female (left) and

male (right) individuals. SNP_2, SNP_3, SNP_4, SNP_5, SNP_12, and SNP_15 were consistently homozygous (A/A) in female and male genomes. In contrast, males exhibited heterozygosity at SNP_19 (G/T), SNP_20 (C/T), and SNP_21 (G/A), whereas females exhibited homozygosity at SNP_19 (G/G), SNP_20 (C/C), and SNP_21 (G/G).

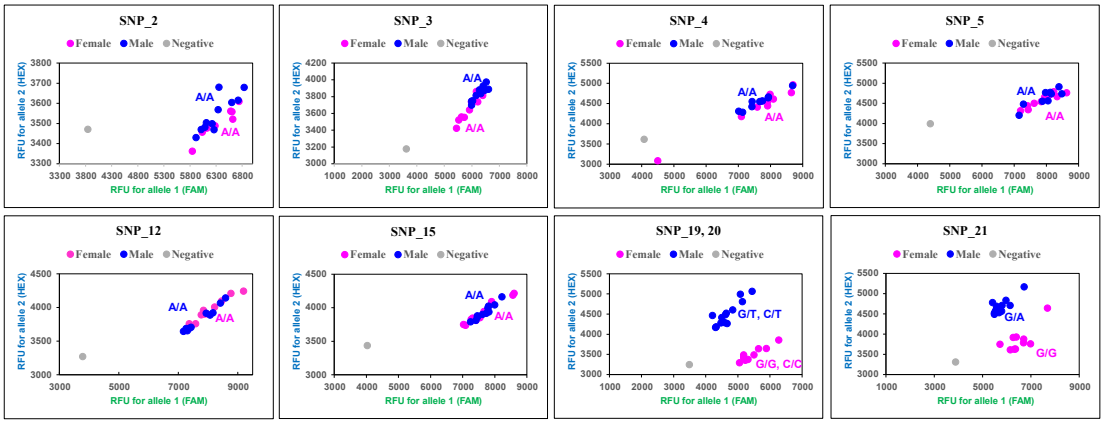

Figure 3.5 PACE genotyping of nine SNPs using gDNA from snakeskin gourami. Scatter plots show fluorescence signal intensities (RFU) for allele 1 (FAM, x-axis) and allele 2 (HEX, y-axis). Each point represents an individual fish (pink = female; blue = male; gray = negative control). For SNP_2, SNP_3, SNP_4, SNP_5, SNP_12, and SNP_15, both sexes clustered as A/A homozygotes. In contrast, SNP_19 and SNP_20 segregated by sex (females G/G, C/C, males G/T, C/T), and SNP_21 showed females G/G and males G/A.

3.4.5 Differential expression in *adar* might be associated with RNA editing in males

Six RNA SNPs (SNP_2, SNP_3, SNP_4, SNP_5, SNP_12, and SNP_15) showed no sex-specific variation in gDNA, as all individuals were homozygous (A/A), while their mRNA sequences displayed heterozygous (A/G). This incongruence between transcriptomic and genomic genotypes indicates post-transcriptional RNA editing. Supporting this interpretation, our RNA-seq analysis showed that *adar* encoding the adenosine deaminase acting on RNA was significantly upregulated in males compared to females at both developmental stages (Figure 3.6). Specifically, *adar* expression was

4.67-fold higher in juvenile males ($\log_2$ fold change = 4.6716, q = 8.19×10^{-71}) and 3.79-fold higher in adult males ($\log_2$ fold change = 3.7861, q = 1.59×10^{-64}).

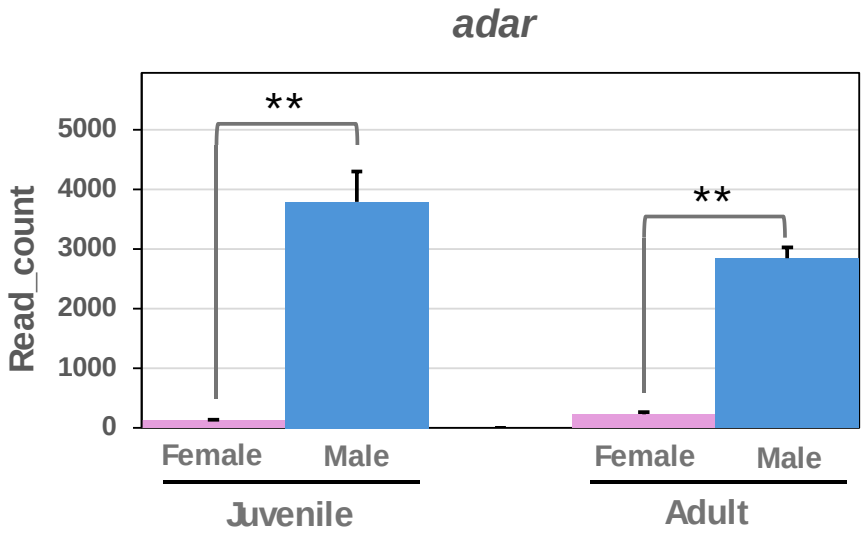

Figure 3.6 Expression levels of *adar* in juvenile and adult snakeskin gourami. Bar plots represent normalized read counts (mean ± SD) for males and females. In juveniles, *adar* expression was significantly higher in males than in females ($\log_2$ fold change = 4.6716, q-value = 8.19×10^{-71}). In adults, *adar* also exhibited male-biased expression ($\log_2$ fold change = 3.7861, q-value = 1.59×10^{-64}). Asterisks indicate statistically significant differences between sexes ($P < 0.01$).

3.4.6 Validation of sex specific SNPs in SNP_19, SNP_20, and SNP_21 in other fish populations by Sanger sequencing and PACE assay

The sex-specific SNP markers SNP_19, SNP_20, and SNP_21 were further validated in four fish populations (SUT1–SUT4) from the SUT Farm using both Sanger sequencing and the PACE SNP genotyping assay. Genotyping of gDNA from 50 females and 50 males consistently revealed sex-linked patterns across all populations. In both assays, all males exhibited heterozygous genotypes, SNP_19 (G/T), SNP_20 (C/T), and SNP_21 (G/A), whereas all females were homozygous for SNP_19 (G/G), SNP_20 (C/C), and SNP_21 (G/G) (Figure 3.4; Table 3.5).

Table 3.5 Validation of sex-specific SNPs in different fish populations.

Population	No. of fish	SNP_19 and SNP_20		SNP_21	
		Female (GC/GC)	Male (GC/TT)	Female (G/G)	Male (G/A)
SUT1	4	2	2	2	2
SUT2	20	10	10	10	10
SUT3	60	30	30	30	30
SUT4	16	8	8	8	8

See nucleotide sequences in Appendix.

3.4.7 Identification of RNA SNPs of SNP_19, SNP_20, and SNP_21 and their DNA variants using T7EI

This study performed a PCR-based T7EI digestion for RNA SNP_19, SNP_20, and SNP_21 in both male (10 fish) and female (10 fish). Using the same RT-PCR and PCR products for sequencing, the amplified fragments of 551 bp of SNP_19 and SNP_20 were obtained. After T7EI digestion and analysis on a 2% agarose gel, male-specific SNP markers were clearly distinguished, as evidenced by the cleavage of amplicons into fragments of 175 bp and 374 bp (SNP_19, SNP_20) according to two-mismatch DNA base-pair cleavage (Figure 3.7). In contrast, female samples showed no digestion and retained intact products of 551 bp (Figure 3.7). In addition, when using both cDNA and gDNA templates, PCR amplification of SNP_21 yielded fragments of 520 bp (cDNA) and 518 bp (gDNA) in both sexes. Following T7EI digestion and electrophoresis on a 2% agarose gel, male-specific SNP markers were identified, as the amplicons were cleaved into fragments of 177 bp and 343 bp (cDNA) and 198 bp and 320 bp (gDNA) (Figure 3.7).

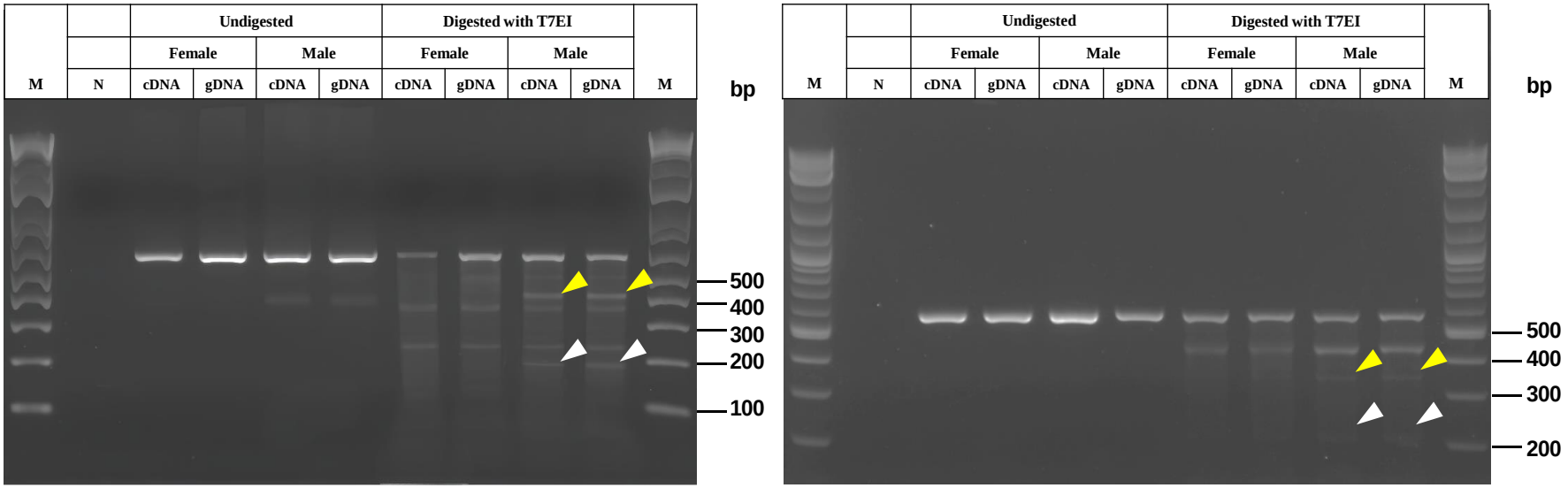

Figure 3.7 T7EI heteroduplex cleavage assay of three SNP-containing loci (SNP_19, SNP_20, and SNP_21) in snakeskin gourami. PCR products were amplified from purified cDNA and gDNA of female and male individuals and either left undigested or digested with T7EI. For SNP_19, 20 (left panel), both sexes exhibited a single intact band in the undigested samples, whereas specific cleavage fragments were detected only in male samples after T7EI digestion (indicated by yellow and white arrowheads), while female samples remained undigested. In contrast, the T7EI assay for SNP_21 (right panel) did not produce clear cleavage fragments (white arrowheads), indicating that the single-nucleotide mismatch at this locus could not be efficiently recognized by the enzyme.

Abbreviation: M = DNA marker; N = negative control.

3.5 Discussion

SNP-associated traits have been widely investigated for biological markers to identify variable traits displayed from genetic polymorphism, and transcriptome technology is a powerful tool for large-scale discovery of SNPs. Particularly in fish, when reference whole-genome sequencing data are limited, high-throughput RNA-seq analysis could be applied to survey a large number of SNPs at relatively low costs. For example, in yellow catfish (*Pelteobagrus fulvidraco*), RNA-seq analysis to identify differentially expressed genes (DEGs) between the ovary and testes also revealed 26,450 putative SNPs from approximately 1.2 million reads (Chen et al., 2015). In addition, in the green mud crab (*Scylla paramamosain*), transcriptome sequencing identified 13,271 SNPs from testis (171,962 reads) and ovary (193,154 reads) libraries (Gao et al., 2014). Likewise, in chum salmon (*Oncorhynchus keta*), 56,020 putative SNPs were obtained from 187,836 unigenes (Li et al., 2022). In this study, no reference genome is currently available for snakeskin gourami, whereas transcriptome data are available. Indeed, RNA-seq was employed to reveal sex-biased gene expression between testes and ovary in both juvenile and adult stages (Boonanuntanasarn et al., 2020; Jantawongsri et al., 2025). Using these databases, a total of 65,535 putative RNA SNPs associated with sex differences were discovered from juvenile (4 pools of 200 testes [JT1-4] vs 100 ovaries [JO1-4]) and adult (4 pools of 40 male [AT1-4] vs 40 female [AO1-4]), which generated 370,116,966 and 354,579,434 reads, respectively. According to a large number of reads and annotated unigenes (juvenile; 50,517 unigenes, adult; 55,883 unigenes), our transcriptome SNPs information was considered acceptable for further screening of sex-determination associated SNPs.

Practically, conventional Sanger sequencing has been implemented to validate and confirm SNP variants. For example, in the black-spotted frog (*Pelophylax nigromaculatus*), transcriptome analysis in testes and ovary revealed 35,406 RNA SNPs. Among these SNPs, 134 RNA SNPs which were in DEGs and consistent with male heterozygosity and female homozygosity were discovered. Subsequent DNA sequencing in gonadal tissues confirmed 83 sex-specific SNPs (Zhang et al., 2024). In addition, in the Chinese soft-shelled turtle (*Pelodiscus sinensis*), the gonadal RNA-seq revealed 35 novel polymorphic SNP markers from the comparative transcriptome of male and female (Li et al., 2022). Among these candidate SNPs, two were significantly

associated with sex phenotypes, with females being homozygous and males being heterozygous (Li et al., 2022). Furthermore, restriction-site associated DNA sequencing (RAD-seq) in mud crabs (*Scylla paramamosain*) identified 20 sex-specific SNPs (Shi et al., 2018). Subsequently, 10 SNPs were confirmed as heterozygous in females and homozygous in males by PCR with Sanger sequencing, demonstrating evidence for the ZW/ZZ sex determination system (Shi et al., 2018). In this study, filtering putative RNA SNPs using combination of databases from juvenile (JO vs JT SNPs) and adult (AO vs AT SNPs), 21 SNPs showed sex-specific RNA SNPs. Subsequently, RT-PCR with Sanger sequencing revealed 9 sex-specific RNA SNPs, with homozygous genotypes in females and heterozygous genotypes in males, providing a hypothesis for the XX/XY sex determination system in snakeskin gourami. Our findings were consistent with the previous findings, indicating an XX/XY mechanism in the snakeskin gourami (Panthum et al., 2021).

For large-scale SNP genotyping, PACE SNP assays provide a rapid, accurate, and high-throughput method for DNA variant genotyping assays, and thereby integrating into selective breeding programs (Torres et al., 2025). For example, in crop breeding, the utility of the PACE assay was demonstrated, as shown in the development of a robust SNP marker set for polyploid roses (Patzer et al., 2024). In Amazon Tambaqui (*Colossoma macropomum*), the practicality of PACE was applied to validate four potentially sex-specific SNPs from 14,805 SNPs in double digest restriction site-associated DNA (ddRAD) sequencing information, which showed 90.0 – 96.7 % association with the phenotypic sex (Varela et al., 2021). In line with these findings, the present study applied PACE genotyping assays to snakeskin gourami and revealed 9 sex-specific RNA SNPs, consistently showing homozygous and heterozygous genotypes across females and males, respectively. These results were in full concordance with Sanger sequencing, thereby confirming the reliability of PACE and highlighting its potential as a rapid and reliable approach for sex determination in fish populations.

Our *de novo* transcriptome assembly and annotation revealed that among nine RNA SNPs associated with sex determination, two SNPs are located in CDS (SNP_5 and SNP_21) of genes whose functions appeared to be crucial in sex differentiation. The nucleotide variant of SNP_5 is located in the CDS of *ezh2* and demonstrates a non-synonymous substitution with a change from lysine (K) to glutamic acid (E). This finding suggested that a change in an allele of AAG (K) to GAG (E) could be associated with sex

determination. In addition, *ezh2* showed significantly higher expression in adult females than in adult males, suggesting that its expression level might be involved in sex maturation. Consistent with this notion, studies in mammals have demonstrated that Ezh1/Ezh2-catalyzed H3K27-trimethylation regulates sex-biased gene expression in the liver, where loss of Ezh1/Ezh2 leads to the upregulation of female-biased genes in males and a concomitant disruption of sex-dependent transcriptional programs. Such differences may reflect lineage-specific adaptations of *ezh2*-mediated chromatin remodeling to reproductive development (Lau-Corona et al., 2020). These findings suggest a conserved role of *ezh2* in establishing sex differentiation across vertebrates, although the direction of regulation remains to be investigated in fish. The nucleotide variant of SNP_21 is found in the *kap3* CDS, although it demonstrates synonymous mutations. Indeed, a synonymous mutation was revealed to modulate translational efficiency and protein folding, thereby leading to affecting phenotype (Komar et al., 2024). In fact, *kap3* showed significant male-biased DEG in adult fish. A recent report has highlighted the critical role of kinesin motor proteins in spermatogenesis, particularly in the remodeling of Sertoli cell junctions, the restructuring of the blood–testis barrier, and the transport of spermatids and other organelles along the seminiferous epithelium in mammals (Yao et al., 2022). These processes are highly dependent on microtubule-based dynamics, for which kinesins act as key mediators and coordinators of germ cell movement and cell cycle progression. Therefore, our findings on sex-associated SNPs in *kap3l*, together with its male-biased expression, point to an important functional link between kinesin-associated proteins and sex differentiation in teleosts. Taken together, RNA variants in CDS together with differentially expressed levels in adults might be related to sex differentiation and sex maturation, respectively.

The 3′UTR of an mRNA molecule plays a crucial role in post-transcriptional gene regulation by controlling mRNA stability, translation efficiency, and subcellular localization (Hong and Jeong, 2023). Given these diverse regulatory roles, polymorphisms in 3′UTRs (SNP_2, SNP_3, SNP_19, and SNP_20) may drive sex-biased expression patterns and contribute to the molecular basis of sex differentiation. For example, SNP_2 is found at the 3’UTR of putative *ift172,* which produces a component of the intraflagellar transport complex. It was reported that *ift172* is essential for ciliary function, gamete motility, and sensory signaling in mice and humans (Zheng et al.,

2023). A variant of SNP_3 was found in the 3'UTR of *man2a1,* which encodes an α-mannosidase localized in the Golgi apparatus, which is critical for protein folding, stability, and function (Chu et al., 2012; Treccarichi et al., 2025). Its function was revealed to be important for germ cell survival, sperm maturation, and follicle development in oogenesis and spermatogenesis in mammals (Akintayo and Stanley, 2019). In addition, the nucleotide variants of SNP_19 and SNP_20 are found in the 3'UTR of *kap3,* which was revealed to play an important role in the transport of spermatids along the seminiferous epithelium in mammals (Yao et al., 2022). Note that nucleotide variants in *kap3* were observed in both the CDS and the 3'UTR. According to DEG analysis, *ift172, man2a1,* and *kap3* (adult only) exhibited a male-biased pattern. These findings could suggest that not only RNA variants, but also differential expression levels may contribute to male-biased reproductive processes.

Our results showed other RNA variants in SNP_4, SNP_12, and SNP_15, which were assembled and annotated as *crk*, *cs*, and *cast*, respectively. Among them, *crk* and *cs* (adult only) showed higher expression in the ovary than in the testes. The adaptor protein *crk* plays a conserved role in regulating cytoskeletal dynamics and cell fate (Shi et al., 2023), although its function related to sex differentiation remains to be elucidated. Citrate synthase (CS), a core enzyme of the tricarboxylic acid (TCA) cycle, has distinct roles in reproduction beyond its mitochondrial function. In females, *cs* is important in the *de novo* synthesis of fatty acids and membrane lipids (Verschueren et al., 2019). Differential expression of *cs* may act as a sex-biased metabolic regulator, potentially supporting the high energy demands of oogenesis. In males, the extra-mitochondrial isoform (*eCS*), abundantly localized in the sperm head, is crucial for initiating the first spike of Ca^{2+} oscillation in the initial calcium spike for activating the egg for proper fertilization (Kang et al., 2020). Although the *cast* did not show differential expression between males and females, its function was crucial in both male and female reproductive systems, particularly in sperm and oocytes. In the calpain-calpastatin system, calpastatin contributes to the calcium-dependent acrosome reaction, which is a critical step for sperm to penetrate the oocyte in mammals (Yudin et al., 2000). In oocytes, the calpain-calpastatin system plays a role in calcium-mediated processes during oocyte activation in humans (Ben-Aharon et al., 2005). In livestock, SNPs in the *cast* were linked to several reproductive traits (Gábor et

al., 2012). Taken together, nucleotide variants in the 3'UTR is several genes that are crucial in the reproductive system could have an important role in sex determination.

In addition to RNA-level variation, three of the nine sex-specific SNPs (SNP_19, SNP_20, and SNP_21) were consistently validated at the gDNA level, including PCR followed by Sanger sequencing, PACE genotyping assays, and the T7EI mismatch cleavage assay. The fish population of SUT1, SUT2, SUT3, and SUT4 was originally collected from commercial farms and natural water sources in the Central and Northeast parts of Thailand. The presence of sex-related SNP_19, SNP_20, and SNP_21 in multiple fish populations suggests that these markers are not confined to sex identification in the experimental population alone. Nevertheless, additional population-level validations are necessary to confirm their utility as sex-specific markers. Note that these 3 SNPs are located in *kap3l*, which was revealed to play a critical role in spermatogenesis and the transport of spermatids in mammals (Yao et al., 2022). The Sanger sequencing demonstrated complete concordance and exhibited a clear genotypic dimorphism: homozygous in females and heterozygous in males in the snakeskin gourami. Their presence in both transcriptomic and genomic datasets confirms their concordance with sex-specific SNP markers. In addition, PACE SNP assays exhibited male heterozygous and female homozygous genotypes, providing high-throughput, accuracy, reliability, and versatility for further marker-assisted sex identification applications in fish populations. T7EI is a mismatch-sensitive nuclease that cleaves heteroduplex DNA at sites of base-pair mismatches or small insertions/deletions, making it a cost-effective and widely useful method for SNP and mutation validation (Vouillot et al., 2015). This study employed the T7EI assay to detect DNA heteroduplex according to male heterozygous genotypes of SNP_19 with SNP_20 and SNP_21 for both cDNA and gDNA, confirming the reliability of the T7EI assay in discriminating between genotypic sexes in snakeskin gourami. However, the T7EI assay for the single-nucleotide mismatch of SNP_21 could not give clear digested bands. It was revealed that the T7EI assay is capable of identifying indels larger than 2 bp, while it has limited sensitivity to single-nucleotide mismatches (Bhattacharya and Van Meir, 2019). Collectively, the consistent outcomes across three independent methodologies and various fish populations strongly support that *kap3l*-linked SNPs represent gDNA-based SNPs as reliable sex-linked markers, demonstrating their utility for genetic sex identification in snakeskin gourami.

Our results showed that, using both the Sanger sequencing and PACE analysis of DNA variants in gDNA, six SNPs (SNP_2, SNP_3, SNP_4, SNP_5, SNP_12, and SNP_15) did not exhibit a heterozygous genotype in males. These findings suggested that there might be post-transcriptional processes during sex differentiation in male fish. Although Sanger sequencing of cDNA revealed mixed peaks (e.g., A/G) at SNPs_5, SNPs_12, and SNPs_15, the subsequent PACE genotyping on genomic DNA did not consistently yield discrete male/female cluster separation. This discrepancy can be attributed to the nature of A-to-I RNA editing, which commonly occurs at partial editing frequencies rather than complete conversion. For instance, several human and mammalian studies report editing efficiencies of < 20% at the majority of sites (Chalk et al. 2019) and only ~10–50% even at relatively efficient loci (Daniel et al. 2017). In addition, the editing index across transcriptomes often reveals substantial numbers of partially edited transcripts (< 1% in some contexts; Bazak et al. 2014). Because the PACE assay is optimized to distinguish discrete allele-specific genotypes (i.e., homozygous vs heterozygous) rather than finely quantify editing ratios, samples in which only ~10–60 % of transcripts carry the edited form will produce intermediate fluorescence signals and result in cluster overlap or unclear segregation. These overlapping clusters thus reflect incomplete editing rather than assay failure. In the context of our male-biased editing loci, the higher ADAR activity in males may increase but not guarantee 100 % editing, resulting in a spectrum of edited transcripts that complicates allele discrimination by PACE. Ultimately, these findings underscore that PACE genotyping of cDNA/edited transcripts should be interpreted with caution when editing rates are moderate or variable, and that Sanger or other quantitative assays like deep sequencing may better capture editing proportions.

To further interpret this observation at the molecular level, one plausible mechanism involves adenosine-to-inosine (A-to-I) RNA editing, a well-documented post-transcriptional modification catalyzed by adenosine deaminases acting on RNA (ADARs), which convert adenosine residues into inosine that is interpreted as guanosine during translation (Nishikura, 2016; Snyder et al., 2017). Our results showed that *adar* exhibited upregulation in the testes compared with that in the ovary in both juveniles and adults. The male-biased *adar* expression would elevate RNA editing activity in males, which may contribute to male-specific RNA variants, even in the absence of corresponding genomic polymorphisms. Previous studies in the olive flounder (*Paralichthys olivaceus*) revealed

extensive A-to-I RNA editing, predominantly occurring in the 3′UTRs of transcripts, with higher editing frequencies in testes than ovaries, consistent with elevated *adar1*, *adar2*, and *adar3* expression in the testis (Wang et al., 2020). Moreover, A-to-I editing is known to be highly active during early embryogenesis, when development depends on maternal mRNAs and proteins, as demonstrated in zebrafish (Buchumenski et al., 2021). Together, these findings suggest that RNA editing represents a key regulatory mechanism underlying sexual differentiation, particularly in heterogametic males, and may also explain the discrepancies observed between RNA and DNA variants.

3.6 Conclusion

In summary, transcriptome-based SNP discovery in juvenile and adult snakeskin gourami identified nine RNA SNPs located in *ift172, man2a1, crk, ezh2, cs, cast,* and *kap3l*, with differential expression between testes and ovaries, suggesting that both nucleotide variants and expression levels are associated with sex differentiation. These RNA SNPs exhibited heterozygosity in males and homozygosity in females, supporting an XY sex-determining system in *T. pectoralis*. Among the nine RNA SNPs, three exhibited concordant nucleotide variants at the genomic level, demonstrating their potential as reliable markers for genotypic sex identification. The absence of genomic variants in the remaining six RNA SNPs, together with male-biased expression of *adar,* indicates that RNA editing may play a regulatory role in the process of sex differentiation. Furthermore, a mismatch-sensitive T7EI assay confirmed the sex-associated nucleotide variants at SNP_19 and SNP_20. Collectively, these findings provide novel molecular evidence for the genetic and post-transcriptional regulation of sex differentiation and offer promising markers for sex identification and selective breeding in the snakeskin gourami.

3.7 Acknowledgements

This work was supported by (i) Suranaree University of Technology (SUT), (ii) Thailand Science Research and Innovation (TSRI), and (iii) National Science, Research and Innovation Fund (NSRF) (FF3-303-68-12-11(F)). The authors also express their gratitude for the funding support from the Agricultural Research Development Agency (Public Organisation) (PRP6705031040).

3.8 References

Abaho, I., Kwikiriza, G., Atukwatse, F., Izaara, A. A., Ekwangu, J., Baguma, S. D., ... & Kasozi, N. (2025). Selective breeding for genetic improvement of Nile tilapia (Oreochromis niloticus Linnaeus, 1758) in Uganda: current status, challenges, and future perspectives. **Animals**, 15(2), 142.

Akintayo, A., & Stanley, P. (2019). Roles for Golgi glycans in oogenesis and spermatogenesis. **Frontiers in cell and developmental biology**, 7, 98.

Bazak, L., Haviv, A., Barak, M., Jacob-Hirsch, J., Deng, P., Zhang, R., ... & Levanon, E. Y. (2014). A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. **Genome research**, 24(3), 365-376.

Ben-Aharon, I., Ben-Yosef, D., Amit, A., & Shalgi, R. (2005). Expression and immunolocalization of the calpain-calpastatin system in the human oocyte. **Fertility and sterility**, 83(6), 1807-1813.

Bhattacharya, D., & Van Meir, E. G. (2019). A simple genotyping method to detect small CRISPR-Cas9 induced indels by agarose gel electrophoresis. **Scientific reports**, 9(1), 4437.

Boonanuntanasarn, S., Jangprai, A., & Na-Nakorn, U. (2020). Transcriptomic analysis of female and male gonads in juvenile snakeskin gourami (Trichopodus pectoralis). **Scientific reports**, 10(1), 5240.

Buchumenski, I., Holler, K., Appelbaum, L., Eisenberg, E., Junker, J. P., & Levanon, E. Y. (2021). Systematic identification of A-to-I RNA editing in zebrafish development and adult organs. **Nucleic Acids Research**, 49(8), 4325-4337.

Chalk, A. M., Taylor, S., Heraud-Farlow, J. E., & Walkley, C. R. (2019). The majority of A-to-I RNA editing is not required for mammalian homeostasis. **Genome biology**, 20(1), 268.

Chen, S. L., Li, J., Deng, S. P., Tian, Y. S., Wang, Q. Y., Zhuang, Z. M., ... & Xu, J. Y. (2007). Isolation of Female-Specific AFLP Markers and Molecular Identification of Genetic Sex in Half-Smooth Tongue Sole (Cynoglossus semilaevis) Song-Lin Chen et al. **Marine Biotechnology**, 9(2), 273-280.

Chen, X., Mei, J., Wu, J., Jing, J., Ma, W., Zhang, J., ... & Gui, J. F. (2015). A comprehensive transcriptome provides candidate genes for sex determination/differentiation and SSR/SNP markers in yellow catfish. **Marine Biotechnology**, 17(2), 190-198.

Chu, J., Mir, A., Gao, N., Rosa, S., Monson, C., Sharma, V., ... & Sadler, K. C. (2013). A zebrafish model of congenital disorders of glycosylation with phosphomannose isomerase deficiency reveals an early opportunity for corrective mannose supplementation. **Disease Models & Mechanisms**, 6(1), 95-105.

Curzon, A. Y., Shirak, A., Meerson, A., Degani, G., Hurvitz, A., Ben‐Naim, N., ... & Seroussi, E. (2022). Cross‐species conservation of a transposase‐linked element enables genetic sexing of commercial populations of Russian sturgeon (Acipenser gueldenstaedtii). **Animal genetics**, 53(3), 441-446.

Daniel, C., Widmark, A., Rigardt, D., & Öhman, M. (2017). Editing inducer elements increases A-to-I editing efficiency in the mammalian transcriptome. **Genome biology**, 18(1), 195.

Davidson, N. M., & Oshlack, A. (2014). Corset: enabling differential gene expression analysis for de novo assembled transcriptomes. **Genome biology**, 15(7), 410.

Devlin, R. H., & Nagahama, Y. (2002). Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. **Aquaculture**, 208(3-4), 191-364.

Gábor, M., Trakovická, A., & Miluchová, M. (2012). The analysis of the MspI polymorphism within exon 3 of the calpastatin gene in Slovak Simmental cattle. **Scientific Papers Animal Science and Biotechnologies**, 45(1), 160-160.

Gao, J., Wang, X., Zou, Z., Jia, X., Wang, Y., & Zhang, Z. (2014). Transcriptome analysis of the differences in gene expression between testis and ovary in green mud crab (Scylla paramamosain). **BMC genomics**, 15(1), 585.

Geffroy, B., & Wedekind, C. (2020). Effects of global warming on sex ratios in fishes. **Journal of Fish Biology**, 97(3), 596-606.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., ... & Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. **Nature biotechnology**, 29(7), 644-652.

Han, C., Zhu, Q., Zhou, X., Ouyang, H., Han, L., Chen, J., ... & Zhang, Y. (2020). A PCR-based genetic sex identification method in spotted mandarin fish (Siniperca scherzeri) and big eye mandarin fish (Siniperca kneri). **Aquaculture Reports**, 18, 100552.

Hong, D., & Jeong, S. (2023). 3'UTR diversity: expanding repertoire of RNA alterations in human mRNAs. **Molecules and cells**, 46(1), 48-56.

Hua, J., Qiang, J., Tao, Y., Li, Y., Lu, S., & Bing, X. (2023). Development and validation of a PCR-RFLP/TaqMan MGB probe method for rapid sex identification of largemouth bass (Micropterus salmoides). **Aquaculture Reports**, 30, 101593.

Jantawongsri, K., Phonsiri, K., Jangprai, A., Na-Nakorn, U., & Boonanuntanasarn, S. (2025). Transcriptomic analysis based on RNA-seq reveals differential gene expression patterns in gonads of adult snakeskin gourami (Trichopodus pectoralis). **Aquaculture**, 604, 742361.

Kabpha, A., Phonsiri, K., Pasomboon, P., & Boonanuntanasarn, S. (2023). Effects of dietary supplementation of estradiol-17β during fry stage on growth, physiological and immune parameters and gonadal gene expression in adult snakeskin gourami. **animal**, 17(9), 100950.

Kang, W., Harada, Y., Yamatoya, K., Kawano, N., Kanai, S., Miyamoto, Y., ... & Miyado, K. (2020). Extra-mitochondrial citrate synthase initiates calcium oscillation and suppresses age-dependent sperm dysfunction. **Laboratory investigation**, 100(4), 583-595.

Kobayashi, Y., Nagahama, Y., & Nakamura, M. J. S. D. (2012). Diversity and plasticity of sex determination and differentiation in fishes. **Sexual Development**, 7(1-3), 115-125.

Komar, A. A., Samatova, E., & Rodnina, M. V. (2024). Translation rates and protein folding. **Journal of Molecular Biology**, 436(14), 168384.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. **Nature methods**, 9(4), 357-359.

Lau-Corona, D., Bae, W. K., Hennighausen, L., & Waxman, D. J. (2020). Sex-biased genetic programs in liver metabolism and liver fibrosis are controlled by EZH1 and EZH2. **PLoS genetics**, 16(5), e1008796.

Li, P., Liu, W., Lu, W., & Wang, J. (2022). Biochemical indices, gene expression, and SNPs associated with salinity adaptation in juvenile chum salmon (Oncorhynchus keta) as determined by comparative transcriptome analysis. **PeerJ**, 10, e13585.

Li, W., Zhu, J., Chen, C., Wang, Y., Lei, L., Geng, L., ... & Zhu, X. (2022). Isolation and characterization of sex-linked SNP markers from transcriptomic sequences of

the Chinese soft-shelled turtle (Pelodiscus sinensis). **Conservation Genetics Resources**, 14(2), 131-136.

Lu, Y. F., Liu, Q., Liu, K. Q., Wang, H. Y., Li, C. H., Wang, Q., & Shao, C. W. (2022). Identification of global alternative splicing and sex-specific splicing via comparative transcriptome analysis of gonads of Chinese tongue sole (Cynoglossus semilaevis). **Zoological Research**, 43(3), 319.

Luckenbach, J. A., Fairgrieve, W. T., & Hayman, E. S. (2017). Establishment of monosex female production of sablefish (Anoplopoma fimbria) through direct and indirect sex control. **Aquaculture**, 479, 285-296.

Martinez, V., Mondaca, C., Dorner, J., Zamorano, P., Palomino, J., & Galarce, N. (2025). Validation of sex determination genes in the Coho salmon (O. kisutch) genome and development of a rapid diagnostic test using CRISPR/Cas12. **Aquaculture**, 743150.

Mesnick, S., & Ralls, K. (2018). Sexual dimorphism. In **Encyclopedia of marine mammals** (pp. 848-853). Academic Press.

Mu, J., Hu, W., Chen, R., Yang, Y., Li, H., Li, W., ... & Xu, D. (2024). Production of neomale and neofemale large yellow croaker (Larimichthys crocea) and establishment of all-female populations. **Aquaculture**, 590, 741010.

Nishikura, K. (2016). A-to-I editing of coding and non-coding RNAs by ADARs. **Nature reviews Molecular cell biology**, 17(2), 83-96.

Ospina-Alvarez, N., & Piferrer, F. (2008). Temperature-dependent sex determination in fish revisited: prevalence, a single sex ratio response pattern, and possible effects of climate change. **PloS one**, 3(7), e2837.

Panthum, T., Laopichienpong, N., Kraichak, E., Singchat, W., Ho My Nguyen, D., Ariyaraphong, N., ... & Srikulnath, K. (2021). The snakeskin gourami (Trichopodus pectoralis) tends to exhibit XX/XY sex determination. **Fishes**, 6(4), 43.

Patzer, L., Thomsen, T., Wamhoff, D., Schulz, D. F., Linde, M., & Debener, T. (2024). Development of a robust SNP marker set for genotyping diverse gene bank collections of polyploid roses. **BMC Plant Biology**, 24(1), 1076.

Rajendiran, P., Jaafar, F., Kar, S., Sudhakumari, C., Senthilkumaran, B., & Parhar, I. S. (2021). Sex determination and differentiation in teleost: roles of genetics, environment, and brain. **Biology**, 10(10), 973.

Schunter, C., Garza, J. C., Macpherson, E., & Pascual, M. (2014). SNP development from RNA–seq data in a nonmodel fish: How many individuals are needed for accurate allele frequency prediction?. **Molecular Ecology Resources**, 14(1), 157-165.

Shi, L., Song, H., Zhou, B., & Morrow, B. E. (2023). Crk/Crkl regulates early angiogenesis in mouse embryos by accelerating endothelial cell maturation. **bioRxiv**, 2023-07.

Shi, X., Waiho, K., Li, X., Ikhwanuddin, M., Miao, G., Lin, F., ... & Ma, H. (2018). Female-specific SNP markers provide insights into a WZ/ZZ sex determination system for mud crabs Scylla paramamosain, S. tranquebarica and S. serrata with a rapid method for genetic sex identification. **BMC genomics**, 19(1), 981.

Snyder, E. M., McCarty, C., Mehalow, A., Svenson, K. L., Murray, S. A., Korstanje, R., & Braun, R. E. (2017). APOBEC1 complementation factor (A1CF) is dispensable for C-to-U RNA editing in vivo. **Rna**, 23(4), 457-465.

Torres, A., Argyris, J., & Gaudino, R. (2025). PCR Allele Competitive Extension (PACE) Workflow for Genotyping by SNP Identification. **In Genotyping: Methods and Protocols** (pp. 143-155). New York, NY: Springer US.

Treccarichi, S., Vinci, M., Cirnigliaro, L., Messina, A., Palmigiano, A., Pettinato, F., ... & Barone, R. (2025). MAN2A2-related glycosylation defects in autism and cognitive delay. **Scientific Reports**, 15(1), 24471.

Uba, K. I. N. (2019). Sexual shape dimorphism in the monomorphic fish Decapterus macrosoma (Teleostei: Carangidae). **Computational Ecology and Software**, 9(4), 134.

Vale, L., Dieguez, R., Sánchez, L., Martínez, P., & Vinas, A. (2014). A sex-associated sequence identified by RAPD screening in gynogenetic individuals of turbot (Scophthalmus maximus). **Molecular biology reports**, 41(3), 1501-1509.

Varela, E. S., Bekaert, M., Ganeco-Kirschnik, L. N., Torati, L. S., Shiotsuki, L., de Almeida, F. L., ... & Migaud, H. (2021). A high-density linkage map and sex-linked markers for the Amazon Tambaqui Colossoma macropomum. **BMC genomics**, 22(1), 709.

Verschueren, K. H., Blanchet, C., Felix, J., Dansercoer, A., De Vos, D., Bloch, Y., ... & Verstraete, K. (2019). Structure of ATP citrate lyase and the origin of citrate synthase in the Krebs cycle. **Nature**, 568(7753), 571-575.

Vouillot, L., Thélie, A., & Pollet, N. (2015). Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. **G3: Genes, Genomes, Genetics**, 5(3), 407-415.

Vu, N. V., Eardley, D. L., Delomas, T. A., & Campbell, M. R. (2019). Identification of sex-specific SNPS in burbot Lota lota using RAD sequencing: conservation and management applications. **Fisheries and Aquatic Sciences**, 22(1), 18.

Wang, J., Tao, W., Kocher, T. D., & Wang, D. (2024). Sex chromosome turnover and biodiversity in fishes. **Journal of Genetics and Genomics**, 51(12), 1351-1360.

Wang, L., Wu, Z., Zou, C., Liang, S., Zou, Y., Liu, Y., & You, F. (2020). Sex-dependent RNA editing and N6-adenosine RNA methylation profiling in the gonads of a fish, the olive flounder (Paralichthys olivaceus). **Frontiers in Cell and Developmental Biology**, 8, 751.

Wang, R., Sun, L., Bao, L., Zhang, J., Jiang, Y., Yao, J., ... & Liu, Z. (2013). Bulk segregant RNA-seq reveals expression and positional candidate genes and allele-specific expression for disease resistance against enteric septicemia of catfish. **BMC genomics**, 14(1), 929.

Yao, M., Qu, H., Han, Y., Cheng, C. Y., & Xiao, X. (2022). Kinesins in mammalian spermatogenesis and germ cell transport. **Frontiers in Cell and Developmental Biology**, 10, 837542.

YUDIN, A. I., GOLDBERG, E., ROBERTSON, K. R., & OVERSTREET, J. W. (2000). Calpain and calpastatin are located between the plasma membrane and outer acrosomal membrane of cynomolgus macaque spermatozoa. **Journal of andrology**, 21(5), 721-729.

Zhang, F., Mu, M., Wang, Z., Zhang, H., Song, Y., & Xiao, R. (2025). The Characteristics and Functions of SSRs and SNPs Based on the Transcriptome of Tuta absoluta Exposed to Different Concentrations of Abamectin and Chlorantraniliprole. **Insects**, 16(5), 446.

Zhang, X., Wan, G., Li, Z., Wu, Q., Xiong, S., Wang, X., ... & Hu, Y. (2024). Integrated transcriptome and SNP analysis reveals sex-related genetic insights in the black-

spotted frog (Pelophylax nigromaculatus). **Aquaculture International**, 32(5), 6227-6249.

Zheng, N. X., Miao, Y. T., Zhang, X., Huang, M. Z., Jahangir, M., Luo, S., & Lang, B. (2023). Primary cilia-associated protein IFT172 in ciliopathies. ***Frontiers in Cell and Developmental Biology***, *11*, 1074880.