

RESEARCH ARTICLE

A Transcriptome Derived Female-Specific Marker from the Invasive Western Mosquitofish (*Gambusia affinis*)

Dunja K. Lamatsch^{1*}, Sofia Adolfsson², Alistair M. Senior^{3,4}, Guntram Christiansen¹, Maria Pichler¹, Yuichi Ozaki³, Linnea Smeds², Manfred Schartl⁵, Shinichi Nakagawa³

1 University of Innsbruck, Research Institute for Limnology, Mondsee, Austria, **2** Uppsala University, Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala, Sweden, **3** University of Otago, Department of Zoology, Dunedin, New Zealand, **4** University of Sydney, Charles Perkins Centre / School of Biological Sciences, Sydney, New South Wales, Australia, **5** University of Würzburg, Physiological Chemistry, Biocenter, Würzburg, Germany

* dunja.lamatsch@uibk.ac.at



OPEN ACCESS

Citation: Lamatsch DK, Adolfsson S, Senior AM, Christiansen G, Pichler M, Ozaki Y, et al. (2015) A Transcriptome Derived Female-Specific Marker from the Invasive Western Mosquitofish (*Gambusia affinis*). PLoS ONE 10(2): e0118214. doi:10.1371/journal.pone.0118214

Academic Editor: László Orbán, Temasek Life Sciences Laboratory, SINGAPORE

Received: February 20, 2013

Accepted: January 9, 2015

Published: February 23, 2015

Copyright: © 2015 Lamatsch et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: AS and SN were funded by the Marsden Fund (New Zealand, UOO0812). SA acknowledges funding from the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement n° 253511, and financial support from Hans Ellegren through a European Research Council Advanced Investigator Grant. Printing costs were partly covered by financial support of the Vice Rector for Research of the Leopold-Franzens-University of Innsbruck to DKL. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Sex-specific markers are a prerequisite for understanding reproductive biology, genetic factors involved in sex differences, mechanisms of sex determination, and ultimately the evolution of sex chromosomes. The Western mosquitofish, *Gambusia affinis*, may be considered a model species for sex-chromosome evolution, as it displays female heterogamety (ZW/ZZ), and is also ecologically interesting as a worldwide invasive species. Here, *de novo* RNA-sequencing on the gonads of sexually mature *G. affinis* was used to identify contigs that were highly transcribed in females but not in males (i.e., transcripts with ovary-specific expression). Subsequently, 129 primer pairs spanning 79 contigs were tested by PCR to identify sex-specific transcripts. Of those primer pairs, one female-specific DNA marker was identified, Sanger sequenced and subsequently validated in 115 fish. Sequence analyses revealed a high similarity between the identified sex-specific marker and the 3' UTR of the aminomethyl transferase (*amt*) gene of the closely related platyfish (*Xiphophorus maculatus*). This is the first time that RNA-seq has been used to successfully characterize a sex-specific marker in a fish species in the absence of a genome map. Additionally, the identified sex-specific marker represents one of only a handful of such markers in fishes.

Introduction

Systems of sex determination attract considerable scientific attention, partially due to the great variety of mechanisms that operate among different species. In general, the identification of sex-specific or sex-biased genes can shed light on sex determination, as well as other biological phenomena such as sexual dimorphism and sex-specific selection. In vertebrates, various chromosomal sex determination systems have evolved. The most extensively studied systems are male heterogamety (XX/XY system) in mammals and female heterogamety (ZZ/ZW system) in birds. These vertebrates generally have highly differentiated sex chromosomes, where the

Competing Interests: The authors have declared that no competing interests exist.

X and Z chromosomes are large and gene rich, whereas the Y and W chromosomes (those specific to the heterogametic sex) are smaller, highly heterochromatic and, for the most part contain only a few functional genes. This heteromorphism is thought to be due to degeneration, as result of non-recombination between sex-chromosomes in the heterogametic sex [1,2]. In contrast to mammals and most birds, many other vertebrates, have no cytogenetically distinct sex chromosomes (for an overview see Ellegren [3]), a factor that makes them valuable in evolutionary/genetic studies as they may represent species with young sex chromosomes (i.e. where degeneration of the sex-chromosome specific to the heterogametic sex has not yet occurred) or systems with halted Y/W degeneration.

Fish species are a particularly attractive group in which to study sex chromosomes because such taxa appear to have independently evolved a variety of sex determination systems [4]. Sex determination systems vary between closely related fish taxa [5–9] often without clear phylogenetic patterns [10,11], and may even vary within the same population [12]. Although most teleost species studied do not display differentiated sex chromosomes [13], an extreme diversity of sex determination systems can be found. Gonochoristic and hermaphroditic species are relatively common, and sperm-dependent parthenogens are also known to exist. The factors that initiate differentiation of phenotypic sex also vary highly, ranging from behavioural or environmental factors to strict genetic ones. Where genetic factors do determine sex in teleost fishes, those factors can involve monogenic or polygenic systems [14,15], as well as a variety of sex chromosome systems; e.g. single (XX/X0, XX/XY, ZZ/WZ) and multiple sex chromosomes ($X_1X_1X_2X_2/X_1X_2Y$, XX/XY_1Y_2 , ZZ/ZW_1W_2) [13,16–19].

The Western mosquitofish, (*Gambusia affinis*; Baird and Girard 1853), originates from North America but was distributed throughout the world for the biological control of mosquitos. However, the species is now largely regarded as pest in introduced locations [20–24]. *G. affinis* displays female heterogamety (ZW/ZZ), and is one of the few species where the W chromosome is the largest chromosome of the karyotype and hence, much larger than Z [25–27]. Its closely related sister taxa, *G. holbrooki*, is almost indistinguishable from *G. affinis* on the basis of morphology alone, but has homomorphic sex chromosomes with a contrasting XX/XY sex determination system [25]; i.e. male heterogamety.

Poeciliids, and the *Gambusia* species described above in particular, make excellent model systems in which to study the evolution of sex-determining systems, and sex chromosomes specifically. A key step in the study of sex-determination systems is the early identification of an individual's phenotypic sex. However, diagnosis of phenotypic sex in live early-stage embryos or fry on the basis of morphology is often not possible in (Poeciliid) fishes. Typically, males only develop secondary sexual characters such as the gonopodium (a highly specialized insemination apparatus modified from the anal fin) at onset of testosterone production after puberty [28]. Size is also not a reliable character with which to differentiate the sexes, due to individual variation in growth-rate and development. Thus, a sex-specific marker is required to identify sex in juveniles at early life-history stages (i.e. prior to morphological separation of the sexes). In addition to early identification of sex, markers that unequivocally indicate the genotypic sex of an individual (i.e. WZ vs ZZ) allow for the detection of naturally sex-reversed individuals, and the subsequent study of the causes of such aberrant sexual development.

The identification of sex-specific markers in fish has, however, proved problematic. Recombination between sex chromosomes is common in organisms that either lack heterogamety, or have sex chromosomes with limited differentiation (see [29–31]). Absence of recombination between heterogametic sex chromosomes leads to accumulation of repetitive DNA on the sex chromosome specific to the heterogametic sex (i.e. W or Y). This accumulation makes it difficult to find the few genes solely located on the W (or Y), even with the use of modern techniques (i.e. next generation sequencing). New approaches are therefore necessary to identify

sex chromosome specific sequences (see Chen et al. [32]). Here, we performed a non-targeted expression analysis using RNA-seq to identify female-biased loci in *G. affinis* potentially located on the W sex chromosome. This method was successful in identifying a female-specific molecular marker. This marker represents one of only a handful of such tools in non-model fish species.

Materials and Methods

Fish samples

Indigenous *G. affinis* (N = 44) from Mexico (Pena Blanca, Santa Cruz River system, north of Nogales, Sonora, Mexico; 25 females, 19 males) as well as introduced *G. affinis* (N = 71) from New Zealand, North Island (Chapel lake, Waikato University, Hamilton; 29 females, 42 males) were used for primer testing. Primers were also tested on *G. affinis*' sister species, *G. holbrooki*, from Leninskoe (North-East of Bishkek, Kyrgyzstan; 21 females, 7 males). *G. holbrooki* is also a common model organism and hence, the applicability of our marker to that species would likely be of wide interest.

Field studies (i.e. collections) did not involve endangered or protected species. All fish were caught as juveniles by hand netting, and transported back to laboratories in their respective countries (Dunedin, New Zealand, and Würzburg, Germany). Fish were then raised to maturity in temperature-controlled rooms, at an average of 25°C and under a 12:12 light:dark cycle. No specific collection permissions were required for Kyrgyzstan or New Zealand as *G. affinis* and *holbrooki* are introduced, invasive fish species. The *G. affinis* strain from Mexico is a long-established aquarium strain that was collected prior to the existence of regulations for fishing (i.e. decades ago). That strain was first kept for fish hobbyists, and only recently transitioned in to scientific use.

This study was carried out in strict accordance with the recommendations in the 'Guide for the Care and Use of Laboratory Animals' of the National Institutes of Health. All protocols were approved by the Animal Ethics Committee of the University of Otago (Permit Number: 87/08) and the Animal Protection Officer of the University of Würzburg from the Veterinary Office of the District Government of Lower Franconia, Germany. The number of fish killed or fin-clipped is reported yearly for each species (fin biopsy according to authorization 55.2–2531.01–49/08). Animals were terminated by cervical dislocation, and all efforts were made to minimize suffering.

Illumina HiSeq sequencing

Following the onset of maturity (i.e. sexual differentiation had occurred) 12 male and 12 female fish (*G. affinis*) from New Zealand were dissected and their gonads removed (testis from males and ovaries from females). Gonadal samples were stored in RNAlater (Ambion, Austin, Texas) following manufacturer's instructions to prevent RNA degradation, and transported to Uppsala University for RNA extraction. Total RNA was extracted from gonads using the RNeasy Mini Kit (Qiagen, Sollentuna, Sweden) following the supplier's recommendations. Before sequencing we pooled 12 male *G. affinis* into 6 groups each with two individuals, generating six 'male-expression' replicates. The same process was applied to 12 female *G. affinis*. Barcoded pools were then sequenced in two lanes of an Illumina HiSeq2000.

Sequencing libraries were prepared from 1–4 µg of total RNA according to the TruSeq RNA sample preparation guide #15008136 revA using reagents from the TruSeq RNA sample prep kit set A and set B v1 (Illumina, San Diego, CA). Briefly, poly-A containing mRNA was purified from 1.5 µg of total RNA using poly-T oligo attached magnetic beads, followed by fragmentation of the mRNA. First strand cDNA was synthesized using SuperScript III reverse

transcriptase (Invitrogen, Carlsbad, CA) and random hexamers, followed by second strand synthesis according to the manufacturer's reagents and protocols. The overhangs on the DNA fragments were end-repaired followed by purification using AMPure XP beads (Beckman Coulter, Brea, CA). An A-base was added to the blunt ends of the DNA fragments and adapters, and index tags for sequencing were ligated, followed by a new round of purification using AMPure XP beads. Libraries were amplified for 12–15 PCR cycles, followed by purification using AMPure XP beads. Library qualities were evaluated using the Agilent Technologies 2100 Bioanalyzer and a DNA 1000-kit. Adapter-ligated fragments were quantified by qPCR using the Library quantification kit for Illumina (KAPA Biosystems, Cambridge, MA) on a StepOne-Plus instrument (Applied Biosystems/Life technologies, Carlsbad, CA) prior to cluster generation and sequencing. A 6–10 pM solution of the pooled libraries (see below) was subjected to cluster generation on a cBot instrument (Illumina Inc.). Paired-end sequencing was performed for 100 cycles in one lane using a HiSeq2000 instrument (Illumina Inc), according to the manufacturer's protocols.

Base calling was performed on the instrument by RTA 1.10.36 and the resulting. bcl files were converted to Illumina qseq format with tools provided by OLB-1.9.0 (Illumina Inc.). To separate samples and PhiX control DNA sequenced in the same lane as the sample libraries, the qseq-files were de-multiplexed, allowing for one mismatch. Both de-multiplexing and mapping were done with CASAVA 1.7.0 (Illumina Inc.). Additional statistics on sequence quality were compiled from the base call files with an in-house script. Note that original raw reads have been deposited to NIH Short Read Archive, accession number SRP033398.

De novo assembly and differential expression analysis

Raw sequencing reads were filtered for unique pairs and trimmed, removing bases with quality scores <25, using ConDeTri v1.0 [33]. We then checked that there were no signs of contamination or sequence biases with FastQC v0.7.2 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Reads were assembled *de novo* with Oases v0.1.21 [34] defining a *k*-mer size of 33. An evaluation of *k*-mer 17–61 showed that this *k*-mer size optimized the relationship between contig N50, number of medaka (*Oryzias latipes*) genes (Ensembl 63) to which contigs align using reciprocal BLAST (in house script) and *G. affinis* contig coverage of medaka genes. The coverage was calculated as (medaka gene length + average UTR length) / contig length excluding N's. We used medaka genes to evaluate the *de novo* assembly, as this is the least divergent fully sequenced genome. Each pool was assembled separately. Contigs from all pools were then merged with Newbler v2.5.3 [35], which is designed to assemble longer reads.

We then mapped reads from each pool onto the contigs using BWA version 0.5.9 [36], not allowing for multiple hits and defining a maximum insert size of 250bp. Differential expression analysis was conducted with *baySeq* v1.6.0 (*R* package version 1.2.0; [37]) where we normalized over library size and gene length. This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GBAE00000000. The version described in this paper is the first version, GBAE01000000.

Female specific expression

Putative female specific contigs were identified based on expression profiles in males and females. We chose contigs for downstream analysis that were constructed exclusively from reads derived from female samples, that were >500bp and with a likelihood of differential expression of 1 (calculated in BaySeq).

Primer design

We then designed primers to test the female-only-expressed sequences identified as candidates for sex specific markers. Primers were designed for 79 contigs, excluding 7 contigs, which were confirmed to be subject to bacterial contamination.

All primers were designed with Primer3Plus [38] using default settings, except the following: primer T_m : min. 59, opt. 60, max. 61; max. T_m difference: 1. Advanced settings: Max Poly-x: 3; GC clamp: 1; product size: min. 480, opt. 500, max. 520. Restricting product size to 500bp seemed like a feasible approach that would cover introns that might enlarge the product manifold.

In a further approach, seven additional primers were designed for three *G. affinis* transcriptome contigs identified by BLAST to be sex-linked EST markers from *Oryzias hubbsi* clone br8179 (Genbank accession number AU171840), OLb06.11h (AB095500), and OLb22.11h (AV670414) [7]. *O. hubbsi* has, in similarity to *G. affinis*, a ZZ/ZW sex-chromosome system with a morphologically larger W than Z [7]. Each primer was tested simultaneously on three females and two males. A positive control was chosen from the transcriptome on the basis to be highly expressed in both sexes (contig15716X; S2 Table). This sequence refers to cathepsin K in the 5' UTR region and exon1 of *X. maculatus* (ENSEMBL). To avoid overlapping product sizes in the multiplex PCR, the primers for the positive control (15716_F:GGGGAACAAGGGT-TACGTCT, 15716_R:ACCACAGGAAGGGAGGAACT) were designed to result in a smaller product than all other products (i.e. 259bp).

All candidate sexing primers were tested by PCR amplification on genomic DNA. Primer pairs were scored based on their ability to produce bands from all female templates that differed from the bands produced from all male templates. Primer pairs with identical results on male and female templates were scored as non-specific. If a given primer pair amplified a different pattern in males and females it was considered sex-specific. Primers showing the slightest difference between male and female were tested again on 10 fishes from Mexico and New Zealand, respectively (5 females, 5 males) without positive control.

DNA extraction

DNA was extracted from fish organs (brain, liver, gills, kidney) or muscular tissue by DNeasy Blood&Tissue Kit (Qiagen, Vienna, Austria) and diluted to 50ng/ μ l prior to PCR amplification.

PCR conditions

For primer testing, a multiplex PCR Kit (Qiagen) was used following Kenta et al. [39] with minor adjustments. PCR was carried out in 10 μ l on a Mastercycler (Eppendorf, Vienna, Austria) with two primer pairs each. The PCR thermocycling conditions were identical for all multiplex sets: an initial denaturation step at 95°C for 15 min to activate the hot start Taq polymerase, followed by 10 touchdown cycles of denaturation at 94°C for 30 s, annealing at 60–51°C (decreasing by 1°C per cycle) for 90 s, and extension at 72°C for 90 s, followed by 40 subsequent similar cycles with annealing at 50°C for 90 s, finally followed by an extension at 60°C for 10 min. The PCR products were separated on 1.5% agarose gels, 0.5% TBE at 5V/cm, ethidium bromide stained and photographed under UV light. Amplification patterns were analysed by eye. For female specific PCR products the same conditions were used but without touchdown cycles ($T_a = 55^\circ\text{C}$) and with reduced number of cycles (i.e. 30) and a normal Taq (Dream Taq, Thermo Scientific, Vienna, Austria).

Table 1. PCR primers spanning the aminomethyl transferase (*amt*) gene of *X. maculatus* for sequencing in *G. affinis*.

Oligo name	Product length [bp]	Primer_F	Primer_R	males	females
Exon1	126	ATGTGGGCTCGGGTTACG	GGCACTGAAATCCCCTCTCT	+	+
Exon2	151	ACCACTCTCTTTGACTTCCACAG	GGAGCATGTGGCTGACATC	+	+
Exon2_Intron3	156	GACCAAATCCACGGAAAAG	GCTACAGATCATCTTGTGGAAATC	+	+
Exon2_3	358	ACTTCCACAGGAACAATGGTG	TCTGCGATATCTGCAACCAC	+	+
Intron3	1,386	GCTACAGATCATCTTGTGGAAATC	TTCATAAGAGCCGAGTCTTTGTC	+	+
Intron3_Exon4	200	CTCCCCACTCCTTTTCCTTC	CGATTCCAGGCCAGTCAG	+	+
Exon4_5	266	AGACCGACCAGGGTTACCTC	CTCCAGGTCCACATCAAACC	+	+
Exon6	104	CATGGCTCAGGTGCTTCAG	TCACCCTGCAGTCAGGAATAC	+	+
Exon7	172	TCTGTTCCCTCGTTCCAGAGTG	TCCAAACAAGACTGGCTTCC	+	+
Exon8	100	ATTGTGCCGCAAATCAAAG	CATCAGGGCTCAGGATGG	+	+
Exon9	105	AACGTCGCCATGGGTTAC	GGGCATCTTGCTGACAATG	+	+
Exon9_UTR	935	AACGTCGCCATGGGTTAC	CCCTCATATTTCAACCAATGTG	-	+
Gaf88	501	GGGACACTAGAGCCCACAAA	CAACCAATGTGGAGCATTTTC	-	+
Total	6,498				

Primers spanning from exon 9 into the 3'UTR and Gaf88 only amplified in females. T_m 60.0±1.0°C.

doi:10.1371/journal.pone.0118214.t001

Cloning, sequencing, and sequence analysis

Female-specific bands were cut from the gel, cleaned with the QIAquick Gelextraction Kit (QIAGEN) and sent for sequencing according to the sample submission guide for value read tubes (Eurofins MWG Operon, Ebersberg, Germany). Several bands from male PCR products were cloned into pGEM-T Easy Vector Systems (Promega, Mannheim, Germany), and transformed into competent cells of *E. coli* DH5 α strain (Invitrogen Life Technologies, Vienna, Austria) according to the manufacturer's instructions and sent for sequencing.

The putative aminomethyl transferase gene of *G. affinis* was amplified by different primers designed from the sequence information of *X. maculatus amt*-gene (ENSXMAT00000019396) and sent for sequencing (see primers in Table 1).

Sequence editing was performed using the computer program CodonCodeAligner 4.0 (Centerville, MA, USA). Sequences were subjected to BLASTN [40] searches at the National Center for Biotechnology Information (NCBI), using nucleotide collection (nr/nt) or BLAT [41] searches in ENSEMBL against the platyfish genome (*Xiphophorus maculatus*).

Species confirmation

To genetically confirm our specimens were indeed *G. affinis* and not the closely related *G. holbrooki*, we designed species-specific cytochrome oxidase subunit 1 (COI) primers: We downloaded the COI sequences (652bp) of 6 *G. affinis* and 10 *G. holbrooki* from Genbank, aligned them using Multalin (v 5.4.1; [42]) and identified the base positions where the sequences differed between the species. Primers were designed to cover regions with 2 and 3 nucleotide differences, respectively, between the two species with the 3' end of the primer ending on one of the nucleotide differences (in bold): COI_GafF: TAATTGGTGCCCCCGACATG; COI_GafR: GGAGGACAGCTGTAATTAGGACTGCTCAC (S1a Fig). With a T_m of 66°C and 68°C, respectively, the primers amplify a 327bp product at 66°C annealing temperature in *G. affinis* but not in *G. holbrooki* (T_m 60.4 / 67.4 °C) (S1b Fig).

The following amplification protocol was used: 50ng DNA, HotstarTaq (Qiagen) with 1.5mM MgCl₂ and 5 μ M of each primer. The PCR amplification was performed in a total

volume of 10 μ l for 15 min at 95°C, followed by 32 cycles of 30 s at 94°C for, 30 s at 66°C, and 45 s at 72°C, with a final elongation step of 10 min at 60°C. The PCR products were separated on 1.5% agarose gels/0.5x TBE at 5V/cm ethidium bromide stained and photographed under UV light.

To verify our species determination approach, we amplified the COI fragments at lower temperatures ([S1b Fig](#)) and sequenced products from both species. The resultant sequences were 100% concordant with the voucher sequences ([S1 Table](#)).

Species divergence

The species divergence time of *G. affinis* and *G. holbrooki* was estimated from mitochondrial DNA sequence difference values at the control region (acc. numbers: AY224097, GU188431) and cytochrome b gene (acc. numbers: EF017514, GU183104), respectively. Sequence difference values were 6 out of 396bp (1.52%) for control region, and 41 out of 876bp (4.68%) for cytochrome b (NCBI BLAST alignment, megablast; [\[41\]](#)).

To estimate the minimal and maximal divergence times of the two species, the sequence difference values were divided by the fastest and slowest rates of known calibrated molecular clocks for mitochondrial DNA in teleosts (i.e. 0.0076–0.0036 changes/site/Myr for cytochrome b, and 0.044–0.004 changes/site/Myr for control region) [\[43\]](#).

Results

NGS data analysis

Per pool, the number of unique reads with quality >25 ranged from 14,658,731 to 47,081,412 (average 31,459,508). The number of contigs constructed ranged from 43,467 to 82,803 (average 64,734), total contig length from 27,788,480 bp to 63,525,980 bp (average 47,305,990 bp) and N50 from 961 bp to 1,515 bp (average 1,196 bp). Merging contigs with Newbler then resulted in 47,347 contigs with a total contig length of 63,648,638 bp and a N50 of 2,496 bp. These contigs were then analysed for differential expression.

Female specific expression

108 putative female-specific contigs were identified based on expression profiles in males and females. We excluded contigs which obviously showed contamination by bacteria according to Genbank (N = 7), and tested the remaining contigs (N = 79) until positive result. The supplementary [S2 Table](#) shows all contigs including positive control (contig15716X) and the three sex-linked EST markers from *Oryzias hubbsi*.

Search for sex-specific sequences

We tested 129 primer pairs from 79 contigs, covering a total of 61,763 bp, as well as 7 primer pairs derived from 3 sex-linked EST markers in *O. hubbsi* [\[7\]](#), which covered 3,202 bp. From a total of 136 tested primer pairs covering a total of 64,966 bp, we found one that differentially amplified male and female genomic DNA of *G. affinis*: Females showed a strong 500bp band, whereas males showed a multi-band profile ([Fig. 1](#)). The identified female-specific marker was termed Gaf88 and corresponds to contig23199X. This primer pair was tested on a total of 115 fishes: 25 females and 19 males from Mexico, and 29 females and 42 males from New Zealand. All but one of the tested individuals showed the banding pattern predicted by their phenotypic sex. When amplified with the same primers, the males and females of the sister species, *G. holbrooki*, gave a multi-band profile identical to that produced by male *G. affinis* (N = 7 and 21, males and females respectively; data not shown).

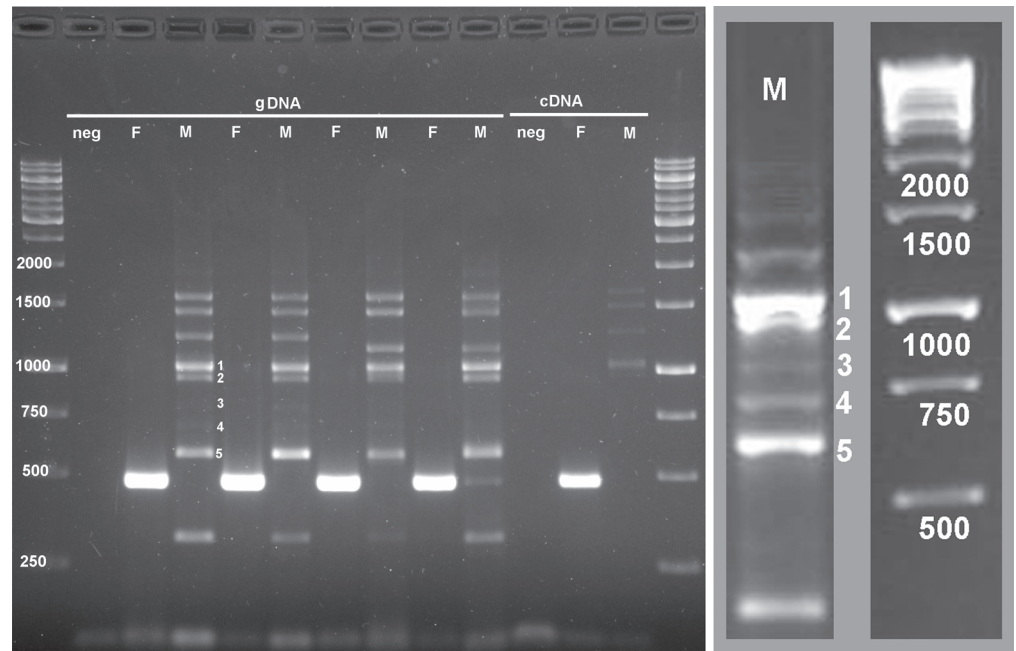


Fig 1. Sex-specific amplification of Gaf88. Sex-specific PCR amplification with primers specific to sequence contig23199X (Gaf88) from the transcriptome of *G. affinis*. Females (F) show a specific 500bp band identical to the original contig in genomic DNA (gDNA) as well as in cDNA, whereas males (M) do not show this band but a multiband-profile ranging from approx. 560–2000 bp. Male bands numbered 1–5 have been isolated and sequenced (enlargement). 1.5% agarose gel, 0.5%TBE, 5V/cm.

doi:10.1371/journal.pone.0118214.g001

Based on transcriptome reads (see [S2 Fig](#)), the sequence of Gaf88 was revealed to be a 779 bp contig. Sequencing of the female amplified products ([Fig. 1](#)) showed a 100% match with the original contig sequence (501bp, N = 2). No significant hits were found in BLASTN (NCBI), but a BLAT search against the platyfish genome (*Xiphophorus maculatus*) in ENSEMBL revealed on average 93.1% similarity with a predicted aminomethyl transferase gene (*amt*, ENSXMAT00000019396; scaffold JH556705.1: 1,171,505–1,178,951) (of 771 from 779 bp) ([S3 Fig, S3 Table](#)). The sequence match is in the 3' UTR of *Xma amt*.

The male sequences were mostly larger (approx. 560–2000bp) ([Fig. 1](#)). 31 cloned PCR products from two males were sequenced but gave no significant hits with either BLASTN (NCBI) or BLAT (ENSEMBL) (Genbank accession numbers KP179419–KP179449).

Primers designed to span the nine exons of *amt* from *X. maculatus* (ENSXMAT00000019396) amplified products in both, males and females, with no significant length differences. As expected, primers spanning from Exon 9 to the 3' UTR of *amt* (Exon9_UTR) as well as Gaf88 primers resulted in a product from females only (see [Table 1](#)).

Sequencing of all exons and introns from two males and two females resulted in a 6,498 bp consensus sequence (Genbank accession number KP113677), which showed 90% identity with *amt* from platy (93% query cover, E-value: 0) ([S3 Fig, Table 1](#)).

Discussion

The identification of sex-specific markers can be a key step in understanding reproductive biology, genetic factors involved in sexual dimorphisms, mechanisms of sex determination and the evolution of sex chromosomes within and between species. Here, we generated the

female-specific marker Gaf88 for the Western mosquitofish, *Gambusia affinis*, by screening sex-differentially expressed sequences from a transcriptome composed of pooled gonads.

To our knowledge, this is the first time that transcriptomes were successfully used to identify a sex-specific marker in a fish species. Although Hale et al. [44] attempted to discern a sex specific marker in sturgeon (*Acipenser fulvescens*) by massive parallel pyrosequencing of gonad transcriptomes, they ultimately failed to identify a sex-specific product from 73 candidate contigs. It seems that no method has yet been successful in identifying sex specific markers in sturgeon [45]. Given the falling price of transcriptomics many references can be found which describe the analysis of the transcriptomes of fish and list putative sex-related genes, but without diagnostic marker identification (in fishes e.g. Liu et al. [46] and Tao et al. [47] in tilapia; Shen et al. [48] in Asian arowana; Vidotto et al. [49] in Adriatic sturgeon; Sun et al. [50] in catfish).

As well as the approach that we describe here, a string of other methods have also been successfully used to identify sex-specific markers in fishes. Those methods include, subtractive cloning (e.g. Nakayama et al. [51], in *Leporinus elongatus*), randomly amplified polymorphic DNA (RAPD; e.g. da Silva et al. [52], in *Brycon amazonicus*; Xia et al. [53], in *Paramisgurnus dabryanus*; Vale et al. [54], in turbot), representational difference analysis (RDA; e.g. Sato et al. [55], in *Oryzias*), amplified fragment length polymorphism (AFLP; e.g. Olmstead et al. [56], in the fathead minnow, *Pimephales promelas*; Cui et al. [57] in *Takifugu rubripes*; Chen et al. [58], in the tongue sole, *Cynoglossus semilaevis*; Brunelli and Thorgaard [59], in the Pacific salmon), Restriction-site Associated DNA (RAD) sequencing (e.g. Palaiokostas et al. [60], in the Atlantic halibut *Hippoglossus hippoglossus*), and genetic linkage map (Rondeau et al. [61], in sablefish *Anoplopoma fimbria*).

The female-specific marker we describe here identified sex in individuals from independent non-mixing populations (i.e. fish from Mexico and New Zealand). Among the 115 individuals subject to molecular sexing, we identified only one female that produced a negative amplification pattern following PCR with Gaf88. This fish was possibly a naturally feminized ZZ neo-female. Unfortunately, this individual was not available for cytogenetic analyses, as the presence or absence of W can easily be recognized in chromosomal metaphase spreads. In the future, our marker may be more widely applied to identify other such exceptional fish. Previous studies have suggested sex-determination to be relatively plastic in most teleosts, including *G. affinis* [62] (reviewed in Senior and Nakagawa [63] and Senior et al. [64]), thus naturally feminized or masculinized animals maybe widespread. In instances of sex-reversal identified by sex-specific marker, the karyotype may also be used to clarify the alternative hypothesis; namely that the sex-reversed fish was a recombinant and that the negative PCR result was the consequence of a W/Z sex chromosomal cross-over [65].

The sequence of Gaf88 shows a high similarity with the 3' UTR sequence of an ORF coding for an enzyme with homology to an aminomethyl transferase (*amt*) from a fish from the same family (Poeciliidae, *Xiphophorus maculatus*). This enzyme is a tetrameric protein of the "glycine cleavage" system. Glycine is not an essential amino acid but a neurotransmitter, and the breakdown of excess glycine is necessary for the normal development and function of nerve cells in the brain and spinal cord [66]. Due to its crucial biochemical role, it is not clear why the (likely) *amt*-gene should be differentially expressed in male and female gonads of *G. affinis*. The gene is present in males and females, as we have proven by sequencing, revealing a 90% identity with *amt* from *X. maculatus*. Based on these facts, two explanations for the lack of amplification of a product from male genomic DNA are identifiable to us: 1) differences in the primer binding sequence between W and Z or 2) a very large insertion in 3' UTR of the Z-copy, which yields a product size that cannot be amplified by conventional PCR.

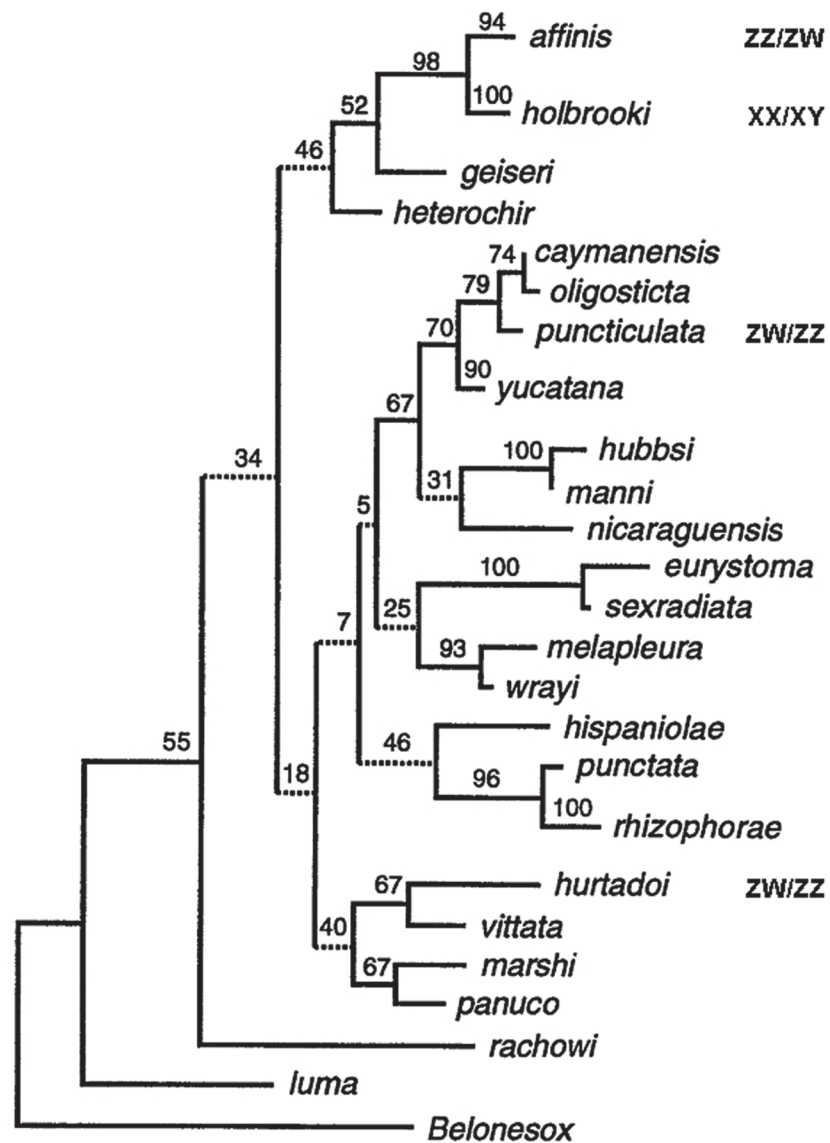


Fig 2. Phylogeny of *Gambusia*. A cladogram of the single most-parsimonious tree for *Gambusia* derived from up to 407bp of a segment of the mitochondrial cytochrome b gene. Where known, the sex determination mechanism is given. Oxford University Press grant permission for the requested material to be reused: [Fig. 1](#) from Lydeard et al. [69].

doi:10.1371/journal.pone.0118214.g002

According to Devlin and Nagahama [13] sex determination has been elucidated in only a few species of the genus *Gambusia*: an XX/XY system has been identified in *G. holbrooki*, whereas ZZ/ZW was found for *G. gaigei*, *G. puncticulata*, *G. hurtadoi*, *G. nobilis*, and of course *G. affinis* (Fig. 2; [26,67,68]). Since the sex determination system is not known for *G. heterochir* and *G. geiseri*, the sister clade to *G. affinis*/*G. holbrooki*, it is difficult to speculate about origin and evolution of the W chromosome in *Gambusia*. Testing Gaf88 widely within the genus may produce interesting insights into the evolution of sex chromosomes in this group. Unfortunately, perhaps the most interesting species to which Gaf88 might be applied (i.e. *G. heterochir* and *G. geiseri*) were not available to us as these species are currently of a conservation concern.

Here, we were only able to test our marker in the sister species of *G. affinis*, *G. holbrooki* (XX/XY). Although, we note that *G. holbrooki* is another common model organism, thus the outcome of the applicability of our marker to that species will likely be of some interest. Both, male and female *G. holbrooki* gave a banding pattern identical to that produced by male *G. affinis*, indicating that the female specific sequence is absent from *G. holbrooki*. It cannot be concluded whether: 1) the marker is specific to a newly derived W chromosome after the separation of the two sister species [69,70] or 2) whether there was an ancestral ZW/ZZ system in the group [(*affinis*, *holbrooki*) (*geiseri*, *heterochir*)], and *G. holbrooki* might have lost the W, developing a new XY system. A phylogenetic analysis in anurans suggests, however, that shifts from ZW to XY are more frequent than the reciprocal process (for a review see Bachtrog et al. [71]).

We estimated the divergence time of *G. affinis* and *G. holbrooki* using mitochondrial DNA sequences based on the fastest and slowest rates of known calibrated molecular clocks for mitochondrial DNA in teleosts [43]. The differences between cytochrome b sequences (0.0076–0.0036 changes/site/Myr) give a minimal age for the W-chromosome of *G. affinis* between 6.16 and 13 million years. The calculation for the control region (0.044–0.004 changes/site/Myr) gives a minimal age between 0.35 and 3.8 million years, always assuming the sex chromosome turnover between XX/XY and ZW/ZZ has evolved in parallel with the species divergence.

In contrast to most species where the sex-limited chromosome (W or Y) is smaller than the respective Z or X chromosome, the W-chromosome is the largest of the karyotype in *G. affinis*. This might indicate that genetic degeneration has hardly occurred; an assumption that is supported by an indifferent chromosome staining with DAPI or mithramycin (AT and CG-specific stain, respectively, for detection of highly repetitive DNA blocks; Schartl, Nanda, Schmid pers. comm.). A comparative genome hybridization (male and female DNA on female chromosomes) might indicate that the p-arm of the W is still recombining with the Z chromosome due to a balanced hybridization pattern. However, the q-arm of W shows an overrepresentation of female DNA sequences excluding recombination between W and Z (Lamatsch et al., in prep.). It is thus crucial to identify the chromosomal location of the female-specific marker in *G. affinis*.

Until only recently, the complete sequence of a W chromosome in any system of female heterogamety remained elusive; mostly because a large portion of the initial chicken W chromosome assembly was later discovered to be misassigned [32]. Comparison of the relatively young tongue sole sex chromosomes with those of birds and mammals, however, now provides important insights into ZW sex chromosome evolution [72,73]. Such sequence data will be integral to a better understanding the evolution of non-recombining sex chromosomes that are not subject to the potent forces of sexual selection (i.e. female specific chromosomes; [3]).

Therefore, in the future we plan to perform chromosome sorting and whole chromosome sequencing of the W chromosome of *Gambusia affinis*—a unique model species where (1) the sex chromosome has evolved as the largest chromosome of the karyotype [26], and (2) the closest relative has homomorphic chromosomes with an XX/XY sex determining system [25].

There remains a lack of knowledge concerning the roots of genetic sex determination, especially in lower vertebrates. As we have shown here, RNA-seq on transcriptomes may be a valuable tool to locate and isolate genetic markers for sex-specific regions of the genome.

Supporting Information

S1 Table. NCBI BLAST of amplified COI sequences for species confirmation.
(DOCX)

S2 Table. Information about 108 putative W-linked contigs (Genbank accession number GBAE01000000) from *G. affinis*, one positive control, and three sequences from *Oryzias*

hubbsi. Likelihood of differential expression (DE) calculated in Bayseq v1.6.0, length of contigs in bp and absolute read count for each sequenced pool (F1-M6, F = female and M = male). (DOCX)

S3 Table. ENSEMBL results of Gaf88 BLAT search against platyfish genome (*Xiphophorus maculatus*) sorted by E-value.

(DOCX)

S1 Fig. S1a: Primer design for species confirmation. Multalin (v 5.4.1; [42]) alignment 5' - 3' of the COI gene of *G. affinis* (Gaf: JN026704.1) and *G. holbrooki* (Gho: JN026706.1). The primers are marked in bold and underlined, sequence differences in red. Primers were chosen to give maximum melting temperature differences between both species (Gaf: 66.0/68.0°C, Gho: 60.4/67.4°C). Alignment parameters: Symbol comparison table: blosum62, Gap weight: 12, Gap length weight: 2. **S1b: Species-specific amplification COI primers.** PCR amplification of 327bp of the COI gene in *G. affinis* and *G. holbrooki* with Gaf primers (Gaf_F 66.0°C, Gaf_R 68.0°C) with a temperature gradient from 46–66°C. Due to the huge differences in T_m of the chosen primer sequences between both species (Gho: 60.4/67.4°C), there is hardly any product visible in *G. holbrooki* from 60°C upwards. 1.5% agarose gel, 0.5% TBE, 5V/cm.

(DOCX)

S2 Fig. NGS coverage of Gaf88. The number of female reads mapping to contig23199X (Gaf88). The window scale is 0–100 reads and the length of contig in base pairs is shown by the top scale bar. The blue lines indicate primer locations. Male coverage is 0 (not shown).

(DOCX)

S3 Fig. Alignment of the aminomethyl-transferase (*amt*) gene of *G. affinis* with *X. maculatus*. Multalin (v 5.4.1; [42]) alignment 5' - 3' of *Gambusia affinis* consensus sequence with aminomethyl-transferase (*amt*) gene of *Xiphophorus maculatus* (ENSXMAT00000019396) showing a query coverage of 93% and a sequence identity of 90%. The sequencing primers are marked in bold and underlined, sequence differences in red. Lilac = untranscribed regions (UTR), black = introns, blue = exons, light yellow indicates the sequence of contig23199X (Gaf88) from the transcriptome of *G. affinis* in the 3' UTR region of the *X. maculatus amt* gene. Alignment parameters: Symbol comparison table: blosum62, Gap weight: 12, Gap length weight: 2.

(DOCX)

Acknowledgments

We thank Losia Lagisz, Jiahui Nat Lim, and Petra Fischer for technical assistance. Matthias Stöck and Valery Eremchenko (Bishkek) kindly provided *G. holbrooki* from Kyrgyzstan. Illumina sequencing was performed at the SNP&SEQ Technology Platform of Uppsala University. Computational work was performed at the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) of Uppsala University, supported by the Swedish National Infrastructure for Computing (SNIC). Finally, we would like to thank Catherine E Grueber for her thoughts on the presentation of this work.

Author Contributions

Conceived and designed the experiments: DKL SN GC. Performed the experiments: DKL AS AMS MP YO MS SN. Analyzed the data: DKL SA LS MS MP. Contributed reagents/materials/analysis tools: DKL SA AMS MS SN. Wrote the paper: DKL SA SN.

References

1. Bachtrog D (2013) Evolution of sex chromosomes. In: Losos JB, Baum DA, Futuyma DJ, Hoekstra HE, Lenski RE et al., editors. *The Princeton Guide to Evolution*. Princeton, USA: Princeton University Press. pp. 387–396.
2. Graves JM (2014) Avian sex, sex chromosomes, and dosage compensation in the age of genomics. *Chromosome Research*: 1–13. doi: [10.1007/s10577-014-9413-0](https://doi.org/10.1007/s10577-014-9413-0) PMID: [24700106](https://pubmed.ncbi.nlm.nih.gov/24700106/)
3. Ellegren H (2011) Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. *Nat Rev Genet* 12: 157–166. doi: [10.1038/nrg2948](https://doi.org/10.1038/nrg2948) PMID: [21301475](https://pubmed.ncbi.nlm.nih.gov/21301475/)
4. Kikuchi K, Hamaguchi S (2013) Novel sex-determining genes in fish and sex chromosome evolution. *Developmental Dynamics* 242: 339–353. doi: [10.1002/dvdy.23927](https://doi.org/10.1002/dvdy.23927) PMID: [23335327](https://pubmed.ncbi.nlm.nih.gov/23335327/)
5. Ross JA, Urton JR, Boland J, Shapiro MD, Peichel CL (2009) Turnover of sex chromosomes in the Stickleback fishes (Gasterosteidae). *PLoS Genetics* 5.
6. Tripathi N, Hoffmann M, Weigel D, Dreyer C (2009) Linkage analysis reveals the independent origin of Poeciliid sex chromosomes and a case of atypical sex inheritance in the guppy (*Poecilia reticulata*). *Genetics* 182: 365–374. doi: [10.1534/genetics.108.098541](https://doi.org/10.1534/genetics.108.098541) PMID: [19299341](https://pubmed.ncbi.nlm.nih.gov/19299341/)
7. Takehana Y, Naruse K, Hamaguchi S, Sakaizumi M (2007) Evolution of ZZ/ZW and XX/XY sex-determination systems in the closely related medaka species, *Oryzias hubbsi* and *O. dancena*. *Chromosoma* 116: 463–470. PMID: [17882464](https://pubmed.ncbi.nlm.nih.gov/17882464/)
8. Peichel CL, Ross JA, Matson CK, Dickson M, Grimwood J, et al. (2004) The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. *Current Biology* 14: 1416–1424. PMID: [15324658](https://pubmed.ncbi.nlm.nih.gov/15324658/)
9. Woram RA, Gharbi K, Sakamoto T, Hoyheim B, Holm L-E, et al. (2003) Comparative genome analysis of the primary sex-determining locus in salmonid fishes. *Genome research* 13: 272–280. PMID: [12566405](https://pubmed.ncbi.nlm.nih.gov/12566405/)
10. Charlesworth D, Mank JE (2010) The birds and the bees and the flowers and the trees: lessons from genetic mapping of sex determination in plants and animals. *Genetics* 186: 9–31. doi: [10.1534/genetics.110.117697](https://doi.org/10.1534/genetics.110.117697) PMID: [20855574](https://pubmed.ncbi.nlm.nih.gov/20855574/)
11. Mank JE, Avise JC (2009) Evolutionary diversity and turn-over of sex determination in teleost fishes. *Sexual Development* 3: 60–67. doi: [10.1159/000223071](https://doi.org/10.1159/000223071) PMID: [19684451](https://pubmed.ncbi.nlm.nih.gov/19684451/)
12. Voff JN, Schartl M (2001) Variability of genetic sex determination in poeciliid fishes. *Genetica* 111: 101–110. PMID: [11841158](https://pubmed.ncbi.nlm.nih.gov/11841158/)
13. Devlin RH, Nagahama Y (2002) Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture* 208: 191–364.
14. Kosswig C (1964) Polygenic sex determination. *Experientia* 20: 190–199. PMID: [5322616](https://pubmed.ncbi.nlm.nih.gov/5322616/)
15. Liew WC, Bartfai R, Lim Z, Sreenivasan R, Siegfried KR, et al. (2012) Polygenic sex determination system in Zebrafish. *PLoS ONE* 7: e34397. doi: [10.1371/journal.pone.0034397](https://doi.org/10.1371/journal.pone.0034397) PMID: [22506019](https://pubmed.ncbi.nlm.nih.gov/22506019/)
16. Schartl M, Galiana-Arnoux D, Schultheis C, Böhne A, Voff J (2011) A primer of sex determination. In: Evans J, Pilastro A, I S, editors. *Ecology and Evolution of Poeciliid Fishes*. Chicago: The University of Chicago Press.
17. Ezaz T, Stiglec R, Veyrunes F, Marshall Graves JA (2006) Relationships between vertebrate ZW and XY sex chromosome systems. *Current Biology* 16: R736–R743. PMID: [16950100](https://pubmed.ncbi.nlm.nih.gov/16950100/)
18. Kazianis S (2005) Sex-determination in platyfishes and swordtails. In: Uribe MC, Grier H, editors. *Viviparous Fishes*. Holmstead Florida: New Life Publications. pp. 381–400.
19. De Souza Valentim F, Porto J, Bertollo L, Gross M, Feldberg E (2013) XX/XO, a rare sex chromosome system in Potamotrygon freshwater stingray from the Amazon Basin, Brazil. *Genetica* 141: 381–387. doi: [10.1007/s10709-013-9737-2](https://doi.org/10.1007/s10709-013-9737-2) PMID: [24068425](https://pubmed.ncbi.nlm.nih.gov/24068425/)
20. Pyke GH (2008) Plague Minnow or Mosquito Fish? A Review of the Biology and Impacts of Introduced *Gambusia* Species. *Annual Review of Ecology Evolution and Systematics* 39: 171–191.
21. Ling N (2004) *Gambusia* in New Zealand: really bad or just misunderstood? *New Zealand Journal of Marine and Freshwater Research* 38: 473–480.
22. Komak S, Crossland MR (2000) An assessment of the introduced mosquitofish (*Gambusia affinis holbrooki*) as a predator of eggs, hatchlings and tadpoles of native and non-native anurans. *Wildlife Research* 27: 185–189.
23. Haynes J, Cashner R (1995) Life history and population dynamics of the western mosquitofish: a comparison of natural and introduced populations. *Journal of Fish Biology* 46: 1026–1041.
24. Arthington AH, Lloyed LN (1989) Introduced poeciliids in Australia and New Zealand. In: Meffe GK, Snelson FF, editors. *Ecology and evolution of livebearing fishes (Poeciliidae)*. New Jersey: Prentice Hall. pp. 333–348.

25. Black DA, Howell WM (1979) The North American mosquitofish, *Gambusia affinis*: a unique case in sex chromosome evolution. *Copeia* 1979: 509–513.
26. Chen T, Ebeling A (1968) Karyological evidence of female heterogamety in the mosquitofish, *Gambusia affinis*. *Copeia*: 70–75.
27. Zhuang Z, Wu D, Zhang S, Pang Q, Wang C, et al. (2006) G-banding patterns of the chromosomes of tonguefish *Cynoglossus semilaevis* Günther, 1873. *Journal of Applied Ichthyology* 22: 437–440.
28. Lampert KP, Schmidt C, Fischer P, Voff J-N, Hoffmann C, et al. (2010) Determination of onset of sexual maturation and mating behavior by melanocortin receptor 4 polymorphisms. *Current Biology* 20: 1729–1734. doi: [10.1016/j.cub.2010.08.029](https://doi.org/10.1016/j.cub.2010.08.029) PMID: [20869245](https://pubmed.ncbi.nlm.nih.gov/20869245/)
29. Bull J (1983) *Evolution of sex determining mechanisms* London: Benjamin/Cummings Publishing Company.
30. Ohno S (1967) *Sex chromosomes and sex linked genes*. Berlin: Springer.
31. van Doorn GS, Kirkpatrick M (2010) Transitions between male and female heterogamety caused by sex-antagonistic selection. *Genetics* 186: 629–645. doi: [10.1534/genetics.110.118596](https://doi.org/10.1534/genetics.110.118596) PMID: [20628036](https://pubmed.ncbi.nlm.nih.gov/20628036/)
32. Chen N, Bellott D, Page D, Clark A (2012) Identification of avian W-linked contigs by short-read sequencing. *BMC Genomics* 13: 183. doi: [10.1186/1471-2164-13-183](https://doi.org/10.1186/1471-2164-13-183) PMID: [22583744](https://pubmed.ncbi.nlm.nih.gov/22583744/)
33. Smeds L, Künstner A (2011) ConDeTri—a content dependent read trimmer for Illumina data. *PloS one* 6: e26314. doi: [10.1371/journal.pone.0026314](https://doi.org/10.1371/journal.pone.0026314) PMID: [22039460](https://pubmed.ncbi.nlm.nih.gov/22039460/)
34. Schulz MH, Zerbino DR, Vingron M, Birney E (2012) Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28: 1086–1092. doi: [10.1093/bioinformatics/bts094](https://doi.org/10.1093/bioinformatics/bts094) PMID: [22368243](https://pubmed.ncbi.nlm.nih.gov/22368243/)
35. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376–380. PMID: [16056220](https://pubmed.ncbi.nlm.nih.gov/16056220/)
36. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25: 1754–1760. doi: [10.1093/bioinformatics/btp324](https://doi.org/10.1093/bioinformatics/btp324) PMID: [19451168](https://pubmed.ncbi.nlm.nih.gov/19451168/)
37. Hardcastle TJ, Kelly KA (2010) baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. *BMC bioinformatics* 11: 422. doi: [10.1186/1471-2105-11-422](https://doi.org/10.1186/1471-2105-11-422) PMID: [20698981](https://pubmed.ncbi.nlm.nih.gov/20698981/)
38. Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, et al. (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic acids research* 35: W71–W74. PMID: [17485472](https://pubmed.ncbi.nlm.nih.gov/17485472/)
39. Kenta T, Gratten J, Haigh N, Hinten G, Slate J, et al. (2008) Multiplex SNP-SCALE: a cost-effective medium-throughput single nucleotide polymorphism genotyping method. *Molecular ecology resources* 8: 1230–1238. doi: [10.1111/j.1755-0998.2008.02190.x](https://doi.org/10.1111/j.1755-0998.2008.02190.x) PMID: [21586010](https://pubmed.ncbi.nlm.nih.gov/21586010/)
40. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of molecular biology* 215: 403–410. PMID: [2231712](https://pubmed.ncbi.nlm.nih.gov/2231712/)
41. Kent WJ (2002) BLAT—the BLAST-like alignment tool. *Genome research* 12: 656–664. PMID: [11932250](https://pubmed.ncbi.nlm.nih.gov/11932250/)
42. Corpet F (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic acids research* 16: 10881–10890. PMID: [2849754](https://pubmed.ncbi.nlm.nih.gov/2849754/)
43. Burrige CP, Craw D, Jack DC, King TM, Waters JM (2008) Does fish ecology predict dispersal across a river drainage divide? *Evolution* 62: 1484–1499. doi: [10.1111/j.1558-5646.2008.00377.x](https://doi.org/10.1111/j.1558-5646.2008.00377.x) PMID: [18363866](https://pubmed.ncbi.nlm.nih.gov/18363866/)
44. Hale MC, Jackson JR, DeWoody JA (2010) Discovery and evaluation of candidate sex-determining genes and xenobiotics in the gonads of lake sturgeon (*Acipenser fulvescens*). *Genetica* 138: 745–756. doi: [10.1007/s10709-010-9455-y](https://doi.org/10.1007/s10709-010-9455-y) PMID: [20386959](https://pubmed.ncbi.nlm.nih.gov/20386959/)
45. Keyvanshokoh S, Gharaei A (2010) A review of sex determination and searches for sex-specific markers in sturgeon. *Aquaculture research* 41: e1–e7.
46. Liu F, Sun F, Li J, Xia JH, Lin G, et al. (2013) A microsatellite-based linkage map of salt tolerant tilapia (*Oreochromis mossambicus* x *Oreochromis spp.*) and mapping of sex-determining loci. *BMC genomics* 14: 58. doi: [10.1186/1471-2164-14-58](https://doi.org/10.1186/1471-2164-14-58) PMID: [23356773](https://pubmed.ncbi.nlm.nih.gov/23356773/)
47. Tao W, Yuan J, Zhou L, Sun L, Sun Y, et al. (2013) Characterization of gonadal transcriptomes from Nile tilapia (*Oreochromis niloticus*) reveals differentially expressed genes. *PLoS ONE* 8: e63604. doi: [10.1371/journal.pone.0063604](https://doi.org/10.1371/journal.pone.0063604) PMID: [23658843](https://pubmed.ncbi.nlm.nih.gov/23658843/)
48. Shen XY, Kwan HY, Thevasagayam NM, Prakki SRS, Kuznetsova IS, et al. (2014) The first transcriptome and genetic linkage map for Asian arowana. *Molecular Ecology Resources* 14: 622–635. doi: [10.1111/1755-0998.12212](https://doi.org/10.1111/1755-0998.12212) PMID: [24354690](https://pubmed.ncbi.nlm.nih.gov/24354690/)

49. Vidotto M, Grapputo A, Boscari E, Barbisan F, Coppe A, et al. (2013) Transcriptome sequencing and de novo annotation of the critically endangered Adriatic sturgeon. *BMC genomics* 14: 407. doi: [10.1186/1471-2164-14-407](https://doi.org/10.1186/1471-2164-14-407) PMID: [23773438](https://pubmed.ncbi.nlm.nih.gov/23773438/)
50. Sun F, Liu S, Gao X, Jiang Y, Perera D, et al. (2013) Male-biased genes in catfish as revealed by RNA-seq analysis of the testis transcriptome. *PLoS ONE* 8: e68452. doi: [10.1371/journal.pone.0068452](https://doi.org/10.1371/journal.pone.0068452) PMID: [23874634](https://pubmed.ncbi.nlm.nih.gov/23874634/)
51. Nakayama I, Foresti F, Tewari R, Scharl M, Chourrout D (1994) Sex chromosome polymorphism and heterogametic males revealed by two cloned DNA probes in the ZW/ZZ fish *Leporinus elongatus*. *Chromosoma* 103: 31–39. PMID: [8013252](https://pubmed.ncbi.nlm.nih.gov/8013252/)
52. da Silva EM, Wong MSL, Martins C, Wasko AP (2012) Screening and characterization of sex-specific DNA fragments in the freshwater fish matrinhã, *Brycon amazonicus* (Teleostei: Characiformes: Characidae). *Fish physiology and biochemistry* 38: 1487–1496. doi: [10.1007/s10695-012-9638-9](https://doi.org/10.1007/s10695-012-9638-9) PMID: [22527611](https://pubmed.ncbi.nlm.nih.gov/22527611/)
53. Xia X, Zhao J, Du Q, Zhi J, Chang Z (2011) Cloning and identification of a female-specific DNA marker in *Paramisgurnus dabryanus*. *Fish physiology and biochemistry* 37: 53–59. doi: [10.1007/s10695-010-9415-6](https://doi.org/10.1007/s10695-010-9415-6) PMID: [20607392](https://pubmed.ncbi.nlm.nih.gov/20607392/)
54. Vale L, Dieguez R, Sánchez L, Martínez P, Viñas A (2014) A sex-associated sequence identified by RAPD screening in gynogenetic individuals of turbot (*Scophthalmus maximus*). *Molecular Biology Reports* 41: 1501–1509. doi: [10.1007/s11033-013-2995-3](https://doi.org/10.1007/s11033-013-2995-3) PMID: [24415295](https://pubmed.ncbi.nlm.nih.gov/24415295/)
55. Sato T, Yokomizo S, Matsuda M, Hamaguchi S, Sakaizumi M (2001) Gene-centromere mapping of medaka sex chromosomes using triploid hybrids between *Oryzias latipes* and *O. luzonensis*. *Genetica* 111: 71–75. PMID: [11841190](https://pubmed.ncbi.nlm.nih.gov/11841190/)
56. Olmstead AW, Villeneuve DL, Ankley GT, Cavallin JE, Lindberg-Livingston A, et al. (2011) A method for the determination of genetic sex in the fathead minnow, *Pimephales promelas*, to support testing of endocrine-active chemicals. *Environmental Science & Technology* 45: 3090–3095.
57. Cui J-Z, Shen X-Y, Gong Q-L, Yang G-P, Gu Q-Q (2006) Identification of sex markers by cDNA-AFLP in *Takifugu rubripes*. *Aquaculture* 257: 30–36.
58. Chen S-L, Deng S-P, Ma H-Y, Tian Y-S, Xu J-Y, et al. (2008) Molecular marker-assisted sex control in half-smooth tongue sole (*Cynoglossus semilaevis*). *Aquaculture* 283: 7–12.
59. Brunelli JP, Thorgaard GH (2004) A new Y-chromosome-specific marker for Pacific salmon. *Transactions of the American Fisheries Society* 133: 1247–1253.
60. Palaiokostas C, Bekaert M, Davie A, Cowan ME, Oral M, et al. (2013) Mapping the sex determination locus in the Atlantic halibut (*Hippoglossus hippoglossus*) using RAD sequencing. *BMC genomics* 14: 566. doi: [10.1186/1471-2164-14-566](https://doi.org/10.1186/1471-2164-14-566) PMID: [23957753](https://pubmed.ncbi.nlm.nih.gov/23957753/)
61. Rondeau EB, Messmer AM, Sanderson DS, Jantzen SG, von Schalburg KR, et al. (2013) Genomics of sablefish (*Anoplopoma fimbria*): expressed genes, mitochondrial phylogeny, linkage map and identification of a putative sex gene. *BMC genomics* 14: 452. doi: [10.1186/1471-2164-14-452](https://doi.org/10.1186/1471-2164-14-452) PMID: [23829495](https://pubmed.ncbi.nlm.nih.gov/23829495/)
62. Knapp R, Marsh-Matthews E, Vo L, Rosencrans S (2011) Stress hormone masculinizes female morphology and behaviour. *Biology Letters* 7: 150–152. doi: [10.1098/rsbl.2010.0514](https://doi.org/10.1098/rsbl.2010.0514) PMID: [20659923](https://pubmed.ncbi.nlm.nih.gov/20659923/)
63. Senior AM, Nakagawa S (2013) A comparative analysis of chemically induced sex reversal in teleosts: challenging conventional suppositions. *Fish and Fisheries* 14: 60–76.
64. Senior AM, Lim JN, Nakagawa S (2012) The fitness consequences of environmental sex reversal in fish: a quantitative review. *Biological Reviews* 87: 900–911. doi: [10.1111/j.1469-185X.2012.00230.x](https://doi.org/10.1111/j.1469-185X.2012.00230.x) PMID: [22540898](https://pubmed.ncbi.nlm.nih.gov/22540898/)
65. Williamson KS, May B (2005) Inheritance studies implicate a genetic mechanism for apparent sex reversal in Chinook salmon. *Transactions of the American Fisheries Society* 134: 1253–1261.
66. Nanao K, Takada G, Takahashi E, Seki N, Komatsu Y, et al. (1994) Structure and chromosomal localization of the Aminomethyltransferase Gene (AMT). *Genomics* 19: 27–30. PMID: [8188235](https://pubmed.ncbi.nlm.nih.gov/8188235/)
67. Campos HH, Hubbs C (1971) Cytomorphology of six species of gambusiine fishes. *Copeia*: 566–569.
68. Rab P (1984) Chromosome study of four poeciliid fishes from Cuba. *Folia zoologica* 33: 229–234.
69. Lydeard C, Wooten MC, Meyer A (1995) Molecules, morphology, and area cladograms: a cladistic and biogeographic analysis of *Gambusia* (Teleostei: Poeciliidae). *Systematic Biology* 44: 221–236.
70. Lydeard C, Wooten MC, Meyer A (1995) Cytochrome b sequence variation and a molecular phylogeny of the live-bearing fish genus *Gambusia* (Cyprinodontiformes: Poeciliidae). *Canadian Journal of Zoology* 73: 213–227.
71. Bachtrog D, Kirkpatrick M, Mank JE, McDaniel SF, Pires JC, et al. (2011) Are all sex chromosomes created equal? *Trends in Genetics* 27: 350–357. doi: [10.1016/j.tig.2011.05.005](https://doi.org/10.1016/j.tig.2011.05.005) PMID: [21962970](https://pubmed.ncbi.nlm.nih.gov/21962970/)

72. Chen S, Zhang G, Shao C, Huang Q, Liu G, et al. (2014) Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. *Nat Genet* 46: 253–260. doi: [10.1038/ng.2890](https://doi.org/10.1038/ng.2890) PMID: [24487278](https://pubmed.ncbi.nlm.nih.gov/24487278/)
73. Graves JAM (2014) The epigenetic sole of sex and dosage compensation. *Nat Genet* 46: 215–217. doi: [10.1038/ng.2903](https://doi.org/10.1038/ng.2903) PMID: [24569234](https://pubmed.ncbi.nlm.nih.gov/24569234/)