Heredity (2020) 125:317–327
https://doi.org/10.1038/s41437-020-0340-x

the geneticssociety

ARTICLE

A duplication of the *Anti-Müllerian hormone* gene is associated with genetic sex determination of different *Oreochromis niloticus* strains

A. Y. Curzon[1,2] · A. Shirak[2] · L. Dor[1,2] · T. Zak[3] · A. Perelberg[3] · E. Seroussi[2] · M. Ron[2]

Received: 29 December 2019 / Accepted: 25 June 2020 / Published online: 9 July 2020

Abstract

Sex determination (SD) mechanisms are ancient and conserved, yet much diversity is exhibited in primary sex-determining signals that trigger male or female development. In *O. niloticus*, SD is associated with a male-specific locus on linkage group (LG) 23 which harbors the Y-linked *Anti-Müllerian hormone (amh)* gene, and a truncated duplication, denoted *amhΔy*. We have evaluated the possible role of identified indels and SNPs in the *amh* gene on SD, based on conservation in different *O. niloticus* strains. A fluorescent assay for the detection of a 5 bp insertion in *amhΔy* exon VI, efficiently discriminated between XX, XY, and YY genotypes. Concordance rate between *amhΔy* and sex varied in six *Oreochromis* strains, from 100% (Ghana) through 90% (Swansea) to 85% (Thai-Chitralada). The association of *amhΔy* with sex was found to be conserved in all tested *O. niloticus* strains, and thus supports its key role in SD. However, the previously identified missense SNP (C/T) in *amh* exon II was found only in the Swansea strain, thus excluding its candidacy for the causal variation of SD across all strains. Effects of markers on LGs 1, 3, and 23 (*amhΔy*) fully explained sex distribution in one Thai-Chitralada family ($R^2 = 1.0$), whereas in another family only the major effect of LG23 (*amhΔy*) was significant ($R^2 = 0.37$). Thus, *amhΔy* on LG23 is associated with genetic SD, either as a single causal gene in different *O. niloticus* strains, or in combination with segregating genes on LGs 1 and 3 in the Thai-Chitralada hybrid strain.

Introduction

Various breeding programs (GIFT, GET-EXCEL, FaST, and GST) have utilized hybridization between different *Oreochromis* species and strains to improve growth rate. Followed by selection, these interspecific crosses have demonstrated a wide range of genetic variation in growth rate (Bolivar 1998; Tayamen 2004; Zimmermann and Natividad 2004). The Thai-Chitralada *O. niloticus* strain is derived from a strain of 50 fish from Japan whose lineage originated in Egypt (Mair et al. 2004); and which was introduced to the Royal Chitralada Palace in Thailand in 1965. Chitralada has been considered a potential leading strain in Asian countries, as it has desirable traits for aquaculture (Macaranas et al. 1997). Based on mitochondrial and genomic analyses, it has been demonstrated that *O. aureus*, *O. niloticus* (Egypt), and *O. mossambicus* contributed to the formation of Thai-Chitralada (Curzon et al. 2019). Effective farming of *O. niloticus* requires all-male culture, characterized by uniformity and high growth rate (Shirak et al. 2018). This can be achieved by deciphering the genetic regulation of sex determination (SD), and by developing the manipulation thereof, for the production of an all-male population (Chen et al. 2018).

A gene whose action is sufficient to initiate an entire developmental program with a binary switch that determines whether an individual will develop as a male or a female is designated a "master key regulator" (MKR; Harrison 2007). Different MKR genes such as *sox3*, *dmrt1*, *irf9*, *Anti-Müllerian hormone (amh)*, *amhrII*, and *gsdf* are involved in SD of vertebrate and insect taxa (Volff et al. 2003, 2007; Mei and Gui 2015; Li and Gui 2018). Many

Supplementary information The online version of this article (https://doi.org/10.1038/s41437-020-0340-x) contains supplementary material, which is available to authorized users.

✉ M. Ron
micha.ron@mail.huji.ac.il

[1] Robert H. Smith Faculty of Agriculture, Food and Environment, Hebrew University of Jerusalem, 76100 Rehovot, Israel

[2] Institute of Animal Science, Agricultural Research Organization, 50250 Bet Dagan, Israel

[3] Dor Research Station, Fisheries and Aquaculture Department, Ministry of Agriculture and Rural Development, Bet Dagan, Israel

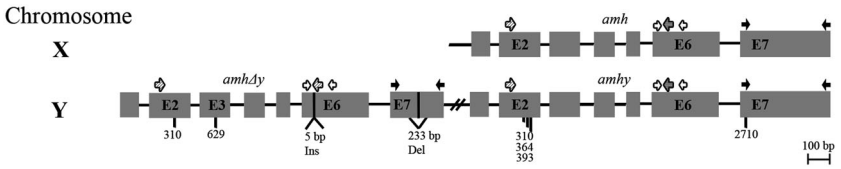


Fig. 1 Schematic representation of the structure of the X and Y copies of the *amh* gene in *O. niloticus*. Adapted from Eshel et al. (2014) and Li et al. (2015). Black vertical lines and their nucleotide position relative to the beginning of the ORF, on exon 1 of each gene copy, represent alteration of the DNA sequence (deletion, insertion, and SNP). Relative positions of pairs of primers for PCR are marked by identical arrows in opposing directions. The PCR based on forward primer on E2 has a different reverse primer on E6 for *amhΔy* than that for *amh* and *amhy*. 100 bp size ladder is indicated.

MKR genes of SD in animals demonstrate copy number variation (CNV) between sexes (Volff and Schartl 2001). The sex-determining region *Y* (*sry*) is an MKR that initiates testis development in most eutherian mammals, and may have evolved from *sox3* through duplication and regulatory mutation in the proto-Y chromosome allele (Gubbay et al. 1990; Sinclair et al. 1990). A sex-specific copy of *dmrt1* has been proposed as the prime candidate for initiating GSD in several species (Nanda et al. 2002; Matsuda et al. 2002; Yoshimoto et al. 2008; Smith et al. 2009; Chen et al. 2014; Cui et al. 2017). In the case of rainbow trout (*Oncorhynchus mykiss*), it has been suggested that *sdY*, a male-specific MKR universal in salmonids, evolved through a duplication event (Yano et al. 2013). *amh* is a member of the TGF-β super family, which has been suggested as an MKR for GSD in *O. niloticus* (Shirak et al. 2006; Eshel et al. 2012), as well as in other distant fish species (Hattori et al. 2012; Yamamoto et al. 2014; Rondeau et al. 2016; Bej et al. 2017). In *O. niloticus*, *amhΔy* is located near the regular gene of *amh*, and differs from the sequence of *amh* by a 233 bp deletion in exon 7 (Eshel et al. 2014) (Fig. 1). A more recent study shows an insertion of 5 bp in exon 6 of *amhΔy*, but not in *amh*; furthermore, the Y-linked *amh* (*amhy*) differs from the X-linked *amh* by a deletion in the promoter region, and by a missense mutation in exon 2 (Li et al. 2015).

Three SD loci on linkage groups (LGs) 1, 3, and 23 have been detected by several independent groups for different *Oreochromis* species and their hybrids (Shirak et al. 2002; Lee et al. 2003, 2004; Cnaani et al. 2008; Eshel et al. 2011; Lühmann et al. 2012; Liu et al. 2013; Palaiokostas et al. 2013; Sun et al. 2014; Lin et al. 2016). Other LGs that are associated with SD have been reported, e.g., LGs 2 and 6 (Shirak et al. 2002), LG12 (Shirak et al. 2008), and LG14 (Gammerdinger et al. 2019). It has also been demonstrated that the effect on GSD might be mimicked by embryonic or adult sex-specific mortality (Shirak et al. 2002, 2008). Environmental temperature may also influence sex, as high incubation temperature (over 32 °C) during weeks 1–3 post hatching causes masculinization of *O. niloticus* (Baroiller et al. 2009). GSD in *Oreochromis* species was proposed to be monofactorial (Wohlfarth and Wedekind 1991). The XX/XY SD system in *O. mossambicus*, the WZ/ZZ system in *O. aureus* and *O. urolepis hornorum*, and the XX/XY system in *O. niloticus* were mapped to LGs 1, 3, and 23, respectively (Lee et al. 2003, 2004; Ezaz et al. 2004; Liu et al. 2013; Eshel et al. 2014; Zhu et al. 2016). Consequently, *amh* was mapped as the causative gene for SD on LG23 (Eshel et al. 2014; Li et al. 2015).

The objectives of this study were: (1) to evaluate the possible role of identified indels and SNPs in the *amh* gene on GSD, based on association with sex of different *O. niloticus* strains; (2) to develop an efficient assay for different *O. niloticus* strains capable of discriminating between XX, XY, and YY genotypes, in order to develop and maintain an all-male population; and (3) to determine additional genomic regions affecting GSD in the Thai-Chitralada hybrid strain.

Materials and methods

Genetic strains

A collection of six strains was assembled, comprising the Chitralada hybrid, and five *O. niloticus* strains that included a local strain of Ghana at Dor Research Station (Shirak et al. 2018), plus strains from the University of Swansea, and from Canada, Egypt, and Ghana (Shirak et al. 2009). The complete data on genetic strains, full-sib progeny, and number of samples from both sexes are provided in Table 1. The strains and full-sib progeny were bred and maintained in Dor research station (Israel). All *O. niloticus* strains were previously analyzed for *COI* sequences that corresponded to the respective *O. niloticus* sequences in the BOLD database (Shirak et al. 2009).

Fish rearing and sexing

Three of the strains described above, Chitralada, Swansea, and Ghana, are currently bred in Dor station. A family for each of these strains was produced by artificial insemination at the Dor research station. Two additional families of Thai-Chitralada, consisting of 130 and 50 full-sib progeny, were

reared at Dor Research Station (Tables 5 and 6). Fertilized eggs were incubated in Zuger bottles (28–29 °C). Fry of each strain was transferred into separate rearing tanks (28–29 °C), and sex was determined at the age of 4 months by microscopic examination of squashed gonads after fry dissection (Guerrero and Shelton 1974).

Fin samples, DNA extraction, and purification of PCR products

Fin samples were preserved in alcohol. DNA was extracted from 100 to 200 mg of caudal fin using MasterPure DNA Purification Kit (Madison, USA), and PCR was performed using Super-Therm Taq DNA polymerase (JMR Holding Inc, London, UK) at 60° for annealing (Table 2). Reactions were performed in a 10-µl volume containing: 1 × PCR buffer with 2 mM MgCl, 1 U DNA polymerase, 100 µM dNTP, 10 pM primer, and 10–30 ng of genomic DNA. The band of relevant size was excised from the gel, purified with DNA Montage Gel Extraction Kit (Millipore, Bedford, MA, USA), and used as template for sequencing.

Table 1 Genetic strains and samples used for the different experiments[a].

Strain	Origin of samples	No. of samples (M:F)[b]	Experiment[c]	Table of results
Canada	Local strain (LS)	4:4	1	3
Egypt	LS	4:3	1	3
Swansea	LS	5:5	1	3
	Full-sib progeny (FSP)	99:20	2	4
Ghana	LS	4:4	1	3
Ghana_dor	LS	45:45	1	3
	FSP	31:39	2	4
Chitralada	LS	50:150	1	3
	FSP	31:58	2	4
	FSP	33 (XY and YY):12	3	NA
	FSP	110:20	4	5
	FSP	10:40	4	6

[a]Maintained in Dor experimental station.

[b]Males:females.

[c]1—concordance of *amh* exon 2 polymorphism with sex; 2—concordance of *amh* polymorphism with sex; 3—assay for sex-specific marker; 4—effects of LGs 1, 3, and 23 on sex.

Genotyping of microsatellite markers on LGs 1 and 3

PCR for microsatellite markers used dye-labeled forward and unlabeled reverse primers for *BYL18* and *UNH168* markers on LGs 1 and 3, respectively (Shirak et al. 2018). The amplified products were separated on an ABI3130 DNA sequencer, and sized using the GeneMapper software v. 4.0 (Applied Biosystems) with GeneScan-500 LIZ size standard (Applied Biosystems) (Dor et al. 2014).

amh sequencing and detection of polymorphism

Polymorphism in exon 2 of *amh* was tested by sequencing four males and females from each strain (except for Egypt, where only three females were present). A forward primer in the beginning of exon 2 and two reverse primers in exon 6 (Table 2) were designed to investigate polymorphism in exon 2 of *amh, amhΔy*, and *amhy* (Eshel et al. 2014; Li et al. 2015). Orientation and position of primers are presented in Fig. 1 and Table 2, respectively. As intronic regions were not conserved, sequencing was analyzed only for exon 2 for the different strains. Presence or absence of 5 bp insertion in exon 6 (Ins6) was analyzed for all available strain samples by using PCR primer pair Ins6F/R (Table 2 and Fig. 2). Presence or absence of 233 bp deletion in exon 7 (ILL2) was analyzed following Eshel et al. (2014).

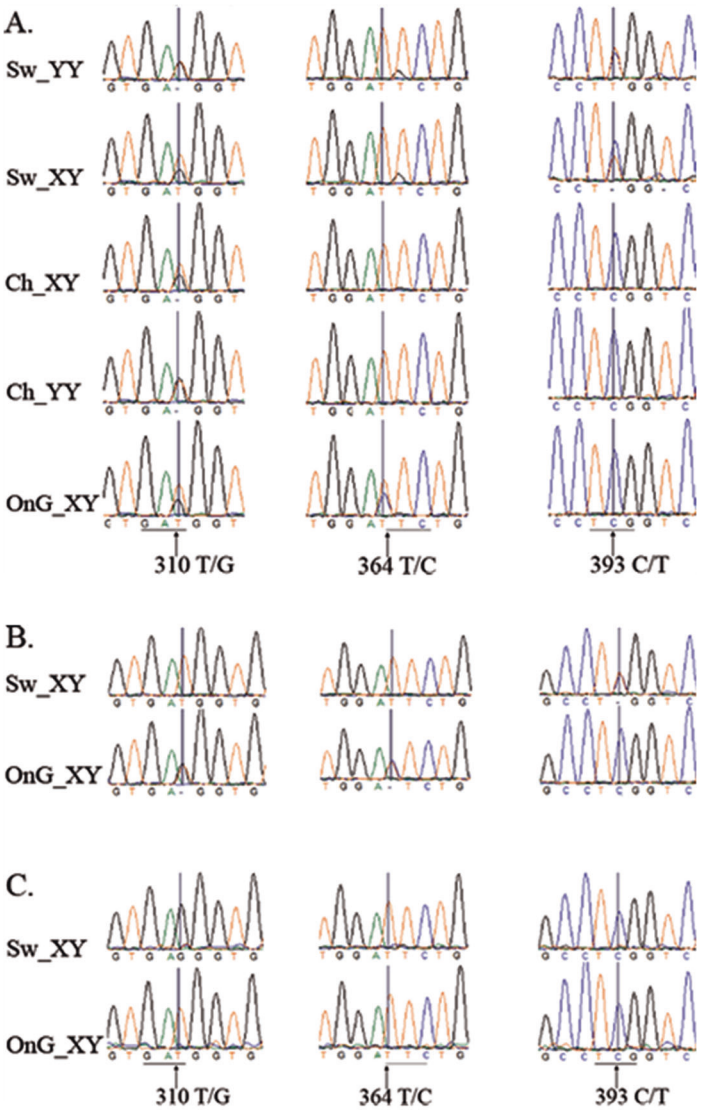
Statistics

Association of marker genotypes and sex in *O. niloticus* strains was tested using a chi-square test. The JMP® statistical package (Pro 13, SAS Institute, Cary, NC) was used

Table 2 Primers for detection and analysis of *amh* copies.

Primer	Exon	Position[a]	Target	Sequence
ILL2F	7	−25	233 bp deletion (*amhΔy*)	TGTGTTTTCTTTCTGCGTCCGCCA (Eshel et al. 2014)
ILL2R	7	393		AGCAGCTCTAGCGGCATCCACA (Eshel et al. 2014)
Ins6F	6	20	5 bp insertion (*amhΔy*)	CCGGCATCTCTGCAAACC
Ins6R	6	305		TGTCATTTGCACCAAAGTCTG
E2F	2	21	Exon 2	AAGACCCCATCATCACCATC (Eshel et al. 2014)
E6Ri	6	51	*amhΔy*	GGAAGCGTTTCATCGACATT
E6Rni	6	46	*amh* and *amhy*	AGGAAGCGTTTCATCTCACAA

[a]The number of nucleotide within an exon that corresponds to the start of the primer is indicated. Number of nucleotide following a minus sign refers to the position from the end of the preceding intron. Relative positions of primers are presented in Fig. 1.



D.

O. niloticus strain	Exon 2 SNP position (in bp) for *amhΔy/amhy/amh*		
	310 T/G	364 T/C	393 C/T
Swansea	G/T/T	T/T/T	C/T/C
Ghana_dor	T/G/T	T/C/T	C/C/C
Chitralada	G/T/T	T/T/T	C/C/C
	64 aa D/E	82 aa F/L	92 aa S/L

Fig. 2 Missense mutations in exon 2 of *amh* copies of different *O. niloticus* strains. Sequencing results of *amh* for XY and YY genotypes of different strains (Sw Swansea, Ch Chitralada, OnG Ghana) based on ratio of peak height of 5 bp insertion on *amhΔy* exon 6 (312 bp) designated as "Y," and its absence (307 bp) designated as "X." The PCR for sequencing included E2 forward primer and either both reverse primers on E6 for all *amh* versions (**a**) or copy specific reverse primer on E6 for *amh* and *amhy* lacking the 5 bp insertion (**b**) or copy specific reverse primer for *amhΔy* harboring the 5 bp insertion (**c**). A summary of deduced mutations is given based on the structure of the *amh* gene (Fig. 1) (**d**). Each single nucleotide change is denoted by its position (in bp) relative to the beginning of the ORF, on exon 1, and the prevalent and novel nucleotides before and after the dividor symbol (/), respectively. Novel nucleotides for the different gene copies and strains are in gray. Each single nucleotide change capable of encoding an amino acid is denoted by its position (in aa) and by the single-letter amino acid codes of the prevalent and novel amino acids before and after the asterisk, respectively.

for logistic regression of sex on markers on LGs 1 and 3, and on *amh* gene on LG23. Separate analysis was used for each one of the two families of Chitralada consisting of 130 and 50 full-sib progeny, respectively. Fisher's exact chi-squared test was applied for association study of individual genetic markers and sex.

Results

Exon 2 sequencing of *amh, amhΔy*, and *amhy*

Allele-specific amplification for different copies of the *amh* gene identified three SNPs in the different *Oreochromis* strains. The schematic representation of the gene following our previous findings (Eshel et al. 2014) and that of Li et al. (2015), and the identified polymorphism, is presented in Fig. 1. The sequencing results for exon 2 in different *Oreochromis* strains, and the nucleotide changes capable of encoding amino acid substitutions, are presented in Fig. 2. Three non-synonymous SNPs in positions 310, 364, and 393 nucleotides of exon 2, relative to the beginning of the open reading frame (ORF) in exon 1, were detected in the Ghana and Swansea strains of *O. niloticus* and in the hybrid of Chitralada. None of the polymorphisms was conserved in all strains. Individuals of the *O. niloticus* strain of Ghana at Dor station showed the same profile of SNPs as the original Ghana strain. Polymorphism was not detected in any of the individuals of Egypt and Canada strains. Fluorescent peak ratios in SNP positions of XY individuals were 2:1, as compared to 1:1 for YY individuals (Fig. 2). This reflects the number of different *amh* copies underlying the chromatograms of Sanger sequencing (Seroussi et al. 2013).

Development of a marker for sex based on 5 bp insertion in *amhΔy* exon 6

Following the observation of 5 bp insertion on exon 6 (Li et al. 2015), we developed a fluorescent assay, using primers Ins6F/R, for the detection of the insertion in *amhΔy*, representing exclusively the Y chromosome; and for its absence in *amh* and *amhy* copies, representing both the X and Y chromosomes (Fig. 3). The resulting genotypes 307/307, 307/312, and 312/312 were in complete concordance with XX, XY, and YY genotypes, respectively (Fig. 3a). Full-sib progeny of Thai-Chitralada XY × XY cross was analyzed for calibration of the 5 bp insertion marker using the fluorescent assay (Fig. 3b). Twelve XX females exhibited a single product of 307 bp. The realized ratios between the 307 and 312 bp peak height alleles for 22 XY males varied from 1.87 to 2.03, with a mean of 1.947 ± 0.042. The respective ratios for 11 YY males varied from 0.959 to 1.002, with a mean of 0.980 ± 0.015. Thus, the realized

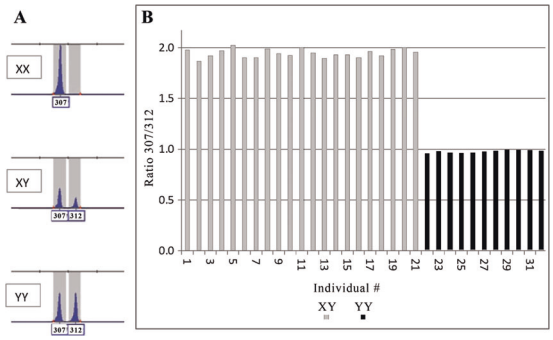


Fig. 3 Development of a marker for sex identification based on a 5 bp insertion in exon 6 of *amhΔy*. Analysis of a 5 bp insertion on *amhΔy* exon 6 allowed detection of fragments with lengths of 312 and 307 bp that are associated with sex. The insertion (312 bp) is designated as "Y," and its absence (307 bp) as "X." To the left, alleles' patterns of three individuals with XX, XY, and YY are shown (**a**). To the right, based on peak height ratios, determination of XY and YY genotypes of 32 Chitralada individuals is exemplified (**b**). *X*-axis represents the individuals ordered by their genotypes (21 XY males followed by 10 YY males). *Y*-axis represents peak height ratios between these allele fragments that were calculated based on their fluorescence values.

peak height ratios in XX, XY, and YY individuals were 1:0, 2:1, and 1:1 based on the fluorescent fragment analysis, this being compatible with the expected dosage of X and Y chromosomes.

Associations of 5 bp insertion in *amhΔy* exon 6 with sex

The concordance rate between the 5 bp insertion in *amhΔy* and sex varied in samples of six *Oreochromis* strains from 85 to 100% (Table 3). Significant associations between the marker and sex were found for Chitralada and Ghana strains ($p < 0.001$): 170 out of 200 Chitralada individuals (85%) showed the expected sex, whereas complete concordance was obtained for the Ghana strain ($n = 90$). Other *O. niloticus* strains had small samples of individuals that were under the minimum threshold required for statistical analysis by chi-square test. Nevertheless, the overall association of this marker and sex for 323 individuals over all strains was highly significant ($p < 0.0001$) (Table 3). Full-sib progeny of the current strains of Chitralada, Swansea, and Ghana showed significant associations with sex, with concordance rates of 91%, 92%, and 100%, respectively (Table 4).

Analysis of the 233 bp deletion in *amhΔy* exon 7

Genotyping exon 7 of *amh* revealed 209/434 and 434/434 genotypes. By sequencing of the different alleles, we found several SNP differences between species, while maintaining the size and location of the 233 bp deletion in exon 7.

Table 3 Associations of 5 bp insertion in exon 6 of *amhΔy* and sex distribution in six *O. niloticus* strains.

O. niloticus strain	Males	Females	*amhΔy* genotype[a]	Concordance with sex (%)[b]	Significance[c]
Canada	3	0	XY	87	NA
	1	4	XX		
Chitralada	33	13	XY	85	*
	17	137	XX		
Egypt	3	0	XY	86	NA
	1	3	XX		
Ghana	4	0	XY	100	NA
	0	4	XX		
Ghana_dor	45	0	XY	100	*
	0	45	XX		
Swansea	4	0	XY	90	NA
	1	5	XX		
Total	92	13	XY	90	*
	20	198	XX		

NA not analyzed due to insufficient samples.

[a]Occurrence of 5 bp insertion on *amhΔy* exon 6 (312 bp) designated as "Y," while its absence (307 bp) designated as "X."

[b]Number of concordant individuals for *amhΔy* and sex (left to right diagonal) divided by total number of tested individuals.

[c]Chi-squared test for divergence of sex distribution from equality between *amh* genotypes.

*$p < 0.0001$.

Table 4 Associations of 5 bp insertion in exon 6 of *amhΔy* and sex distribution in three families of different *O. niloticus* strains[a].

O. niloticus strain	*amhΔy* genotype[b]		Progeny			
	Sire	Dam	Females	Males	*amhΔy* genotype[b]	Concordance with sex (%)[c]
Chitralada	XY	XX	53	3	XX	91
			5	28	XY	
Ghana_dor	XY	XX	39	0	XX	100
			0	31	XY	
Swansea	XY	XY	20	11	XX	92
			0	99	XY/YY	

[a]Significant chi-squared tests for divergence of sex distribution from equality between *amh* genotypes ($p < 0.0001$).

[b]Occurence of 5 bp insertion in *amhΔy* on exon 6 (312 bp) designated as "Y," while its absence (307 bp) designated as "X."

[c]Number of concordant individuals for *amhΔy* and sex (left to right diagonal) divided by total number of tested individuals.

However, the peak height ratio of the 209–434 alleles was not stable, with significant bias toward amplification of the smaller 209 bp allele. This may be expected in view of the twofold difference in allele fragment length. It should be noted, however, that the observed size difference between these alleles was only 225 bp, in comparison to the expected 233 bp. Sequencing of the different alleles validated the expected size difference, and the size calling of the two alleles may have been affected by its altered migration on the acrylamide gel. Nevertheless, complete correspondence was obtained between the 233 bp deletion on exon 7 and the 5 bp insertion on exon 6, and between both markers and sex in Chitralada full-sib progeny, as well as in all other strains.

Effects of markers in LGs 1, 3, and 23 (*amhΔy*) on sex in Chitralada

Genotypes of full-sib progeny of Chitralada XY × XY mating (Family 1) were analyzed for the effects of *BYL018* (LG1) and *UNH168* (LG3) markers, and of *amhΔy* (LG23) on sex (Table 5). Dam and sire genotypes of Family 1, for markers on LGs 1, 3 and 23, formed 24 possible genotypic

Table 5 Genotype effects of *BYL018* (LG1), *UNH168* (LG3) markers, and *amhΔy* (LG23) on sex distribution (female:male ratio) of Chitralada full-sibs progeny of Family 1.

			LG23[a]		
	LG1	LG3	XX	XY	YY
Parents					
Sire	**245**/247	151/151		+	
Dam	247/249	151/**159**		+	
Progeny					
Genotype[b]	**245**/247 & **245**/249	151/151	0:11	0:38	0:15
		151/**159**	4:00		
	247/247 & 247/249	151/151 & 151/**159**	16:0	0:34	0:12
Logistic analysis[c] (R square)	$p < 10E^{-22}$ (1)				

[a]5 bp insertion on *amhΔy* exon 6 (312 bp) designated as "Y," and its absence (307 bp) designated as "X."

[b]Interacting marker alleles on LGs 1 and 3 that are associated with sex are in bold.

[c]Significance value of logistic analysis for the effects of the three loci on sex with R squared value within brackets.

Table 6 Genotype effects of *BYL018* (LG1), *UNH168* (LG3) markers, and *amhΔy* (LG23) on sex distribution (female:male ratio) of Chitralada full-sibs progeny of Family 2.

			LG23[a]	
	LG1	LG3	XX	XY
Parents				
Sire	**245**/247	145/169		+
Dam	247/283	169/169	+	
Progeny				
Genotype[b]	**245**/283 & **245**/247	145/169 & 169/169	17:00	5:00
	247/247 & 247/283		10:00	8:10
Fisher's exact test[c]	$p < 0.0056$		$p < 0.0001$	
(R square)	(0.1)		(0.37)	

[a]5 bp insertion on amhΔy exon 6 (312 bp) designated as "Y," and its absence (307 bp).

[b]Interacting marker allele on LG1 that is associated with sex is in bold.

[c]Significance value of Fisher's exact test ($p < 0.01$) is presented with R squared value within brackets.

combinations ($4 \times 2 \times 3$). Genotype effects of the main combinations that fully explained sex are presented in Table 5 (logistic probability of 10^{-22}). All 99 individuals bearing the Y chromosome (*amhΔy*) on LG23 are males. In the absence of the Y chromosome on LG23 (*amhΔy*), alleles of 245 bp on LG1 and 159 bp on LG3 were associated with females. An XX genotype on LG23 without a 159 bp allele on LG3, and with a 245 bp allele on LG1, resulted in males, whereas all other allele combinations were females. The genotypes for the three loci and sex of Chitralada XY × XX full-sib progeny of Family 2 are presented in Table 6 and Supplementary File 1. Markers on LGs 1 and 23 showed significant effects by Fisher's exact test, with R squared values of 0.10 and 0.37, respectively. However, genotypes of both markers were highly correlated, despite segregation in different LGs ($p < 0.02$), thus indicating an equivocal effect on LG1.

Discussion

SD mechanisms are ancient and conserved, yet, there is diversity in primary sex-determining signals that trigger male or female development (Bachtrog et al. 2014). *O. niloticus* demonstrates different regulators of GSD in closely related strains, accordingly, comparative genomic analysis may reveal the different regulators that affect GSD in the different strains.

Conservation of *amhΔy* and its potential role in GSD

A novel male-specific duplication of *amh*, denoted *amhΔy*, lacking the TGF-β domain was identified and mapped to the QTL region on LG23 for SD, thus indicating its potential role in GSD of *O. niloticus* (Eshel et al. 2014, Li et al. 2015). Li et al. (2015) found that *amhΔy* is located upstream to another copy of *amh* on the Y chromosome (*amhy*). The schematic representation of the structure of the X and Y versions of the *amh* gene in *O. niloticus* is presented in Fig. 1. Knockout of *amhy* by CRISPR/Cas9 resulted in male to female sex reversal in XY fish, while *amhy*, but not *amh*, overexpression resulted in female to male sex reversal (Li et al. 2015). A missense SNP (C/T) in the ORF of *amh* (exon 2) has been suggested as regulating male SD (Li et al. 2015). However, the results of our study showed that, although in all strains sex was associated with polymorphism in *amh*, this SNP in exon 2 was found only in a single strain (Swansea). In a recent study all GIFT individuals were found to have the base C in this position, indicating absence of the suggested SNP (Taslima et al. 2020). Thus, the candidacy of this SNP for the causal variation of SD is excluded. Expression analysis previously showed that both *amhΔy* and *amhy* genes are transcribed, but *amhΔy* is probably not translated (Li et al. 2015). However, interestingly, some nonfunctional genomic pseudogenes which are in-tandem duplicates of the intact gene may act as regulators of gene expression (Han et al. 2011). As *amhΔy* is not translated to a protein, it may act by regulating *amh* gene expression. The *amh* gene was recently discovered as an MKR in additional distant teleost species, e.g., lingcod, *Ophiodon elongates* (Rondeau et al. 2016) and silverside, *Hypoatherina tsurugae* (Bej et al. 2017). Similar to *O. niloticus*, both males and females share one copy of *amh*, however, a second duplicate copy appears solely in male individuals, suggesting that convergent evolution of an

additional copy of *amh* gene may have occurred in multiple teleost lineages, thus adopting an MKR role. Sequence analysis of different strains of *O. niloticus* for *amhΔy* clearly shows the high association between the occurrence of an additional copy of *amh* and maleness in strains originating from Eastern and Western Africa; and thus supports a putative conserved function of *amh*. Other polymorphisms between *amhΔy* and *amhy*, such as the large deletion in *amhy* promoter, have been suggested as being involved in GSD (Li et al. 2015). Thus, further study is needed to test this association and to explore the causative region in *amh* for SD.

Development of an accurate assay for genotyping of *amhΔy*

We developed a sensitive fluorescent assay for detecting the occurrence of a 5 bp insertion on exon 6 of *amhΔy*, and its absence in *amh* and *amhy* copies, yielding complete discrimination between XX, XY, and YY genotypes. Complete concordance was evident between the 233 bp deletion on exon 7 and the 5 bp insertion on exon 6, with increased accuracy for the latter. The realized fluorescent peak height ratios in XX, XY, and YY individuals were compatible with the expected dosage of X and Y chromosomes. Similar assays demonstrated identification of CNV (Seroussi et al. 2013). Thus, we used the 5 bp insertion on exon 6 of *amhΔy* as the marker of choice for GSD. Although the assay provides high but not complete prediction of GSD, it can be used for selection of individuals with desired *amh* genotypes for mating toward the production of all-male populations in different *O. niloticus* strains.

Concordance between *amhΔy* polymorphism and sex

The concordance rate between the 5 bp insertion on exon 6 of *amhΔy* and sex varied in samples from six *O. niloticus* strains and three strains that are currently maintained in Dor station (Table 4). Interestingly, Ghana strains showed full concordance, but all strains that originated from Egypt showed incomplete concordance. Unbalanced sex ratios in progeny of wild and domesticated *O. niloticus*, have been attributed to environmental effects such as temperature, or to polygenic SD (Bezault et al. 2007; Baroiller et al. 2009; Palaiokostas et al. 2015; Baroiller and D'Cotta 2016, 2018; Sissao et al. 2019). Additional studies suggested that specific variants of *amh* are sensitive to thermal effects (Wessels et al. 2014, 2017). Thus, it is possible that *O. niloticus* from an Egyptian origin carries variants of *amh* that are sensitive to temperature. In Israeli growing conditions, temperatures above 29 °C seem to prevent spawning, and thus may not be responsible for sex reversal (unpublished data). Furthermore, although *Oreochromis* strains have been observed even at extreme temperatures (Bezault et al. 2007; Nyingi et al. 2009), they were found to be genetically distant and morphologically different from those analyzed in the current study (Ndiwa et al. 2016).

Multiple loci affecting GSD in the Thai-Chitralada hybrid strain

In our study, we examined the effects of markers on LGs 1, 3, and 23 (*amhΔy*) in Thai-Chitralada families. Interestingly, these markers fully explained sex distribution in a Thai-Chitralada family (R square of unity). Nevertheless, only the effects of markers on LGs 1 and 23 were significant in a second Thai-Chitralada family, with partial R-squared values of 0.10 and 0.37, respectively, indicating the existence of additional unidentified loci that regulate SD. However, the significant correlation between the genotypes of both markers, although segregating in different LGs, undermines the validity of the statistical analysis, thus indicating that, in this family, the effect of LG1 is probably associated through its correlation to *amhΔy* on LG23. Previously, such co-segregation of markers has been attributed to sex-specific mortality (Shirak et al. 2002; Cnaani et al. 2008), or nonrandom segregation of chromosomes during meiosis, resulting in pseudo-linkage (Woram et al. 2004). The two alleles of 245 and 159 bp of genetic markers on LGs 1 and 3, respectively, seem to be linked to the genetic contributors of the effects on sex of Family 1 (Table 5), whereas absence of allele of 159 bp for *UNH168* on LG3 in Family 2 may explain the loss of effect of this locus, or the distant location of the marker from the causative polymorphism (Table 6 and Supplementary File 1).

We have recently shown that *O. mossambicus, O. aureus*, and *O. niloticus* contributed to the formation of the Thai-Chitralada hybrid (Curzon et al. 2019). Thus, involvement of LGs 1, 3, and 23 in GSD of Thai-Chitralada could be explained by its progenitor species harboring these segregating loci (Lee et al. 2004; Liu et al. 2013; Shirak et al. 2018). Recent studies support the major role of *amh* in GSD of farmed *O. niloticus* (Cáceres et al. 2019; Taslima et al. 2020), although incomplete linkage between *amhΔy* and the deletion in the *amhy* promoter was also reported, suggesting that multiple segregating loci may be found in wild *O. niloticus* strains (Sissao et al. 2019). Nevertheless, the purity of these wild strains may be suspect in view of mitochondrial DNA heterogeneity (Rognon and Guyomard 2003), indicating a contamination of West African strains, possibly through the escape of *Oreochromis* hybrids from farms (Elder and Garrod 1961; Pruginin et al. 1975). Currently, breeding strains can be tested for their origin using the *COI* gene, and other assays developed specifically for *Oreochromis* species (Syaifudin et al. 2019; Lind et al. 2019; Curzon et al. 2019).

Conclusions

A sensitive fluorescent assay for the detection of a 5 bp insertion in *amhΔy*, and its absence in *amh* and *amhy* copies, efficiently discriminated between XX, XY, and YY genotypes. The association of *amhΔy* with sex was conserved in all tested *O. niloticus* strains, and thus supports its key role in GSD. However, the previously identified missense SNP (C/T) in *amh* exon II was found only in the Swansea strain, thus excluding its candidacy for the causal variation of SD across all strains. Thus, *amhΔy* on LG23 is associated with GSD, either as a single causal gene in different *Oreochromis* strains, or in combination with segregating genes on LGs 1 and 3 in the Thai-Chitralada hybrid strain.

Acknowledgements This work was supported by the Fish Breeders Organization, Israel (Grant Number 07-2018). Permission for animal experiments was granted by Health Bureau certification (#12844) and by the Institutional Animal Care and Use Committee (#739). We would like to thank Dr. J.I. Weller for his statistical advice.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Bachtrog D, Mank JE, Peichel CL, Kirkpatrick M, Otto SP, Ashman T-L et al. (2014) Sex determination: why so many ways of doing it? PLoS Biol 12:e1001899

Baroiller JF, D'Cotta H (2016) The reversible sex of gonochoristic fish: insights and consequences. Sex Dev 10:242–266

Baroiller JF, D'Cotta H (2018). Sex control in Tilapias. In: Sex control in aquaculture, John Wiley & Sons, Ltd, Chichester, p 189–234

Baroiller JF, D'Cotta H, Saillant E (2009) Environmental effects on fish sex determination and differentiation. Sex Dev 3:118–135

Bej DK, Miyoshi K, Hattori RS, Strüssmann CA, Yamamoto Y (2017) A duplicated, truncated amh gene is involved in male sex determination in an old world silverside. Genes Genomes Genet 7:2489–2495

Bezault E, Clota F, Derivaz M, Chevassus B, Baroiller JF (2007) Sex determination and temperature-induced sex differentiation in three natural populations of Nile tilapia (*Oreochromis niloticus*) adapted to extreme temperature conditions. Aquaculture 272:S3–S16

Bolivar R (1998) Estimation of response within-family selection for growth in Nile tilapia (*O. niloticus*). (PhD Dissertation). Dalhousie University, Halifax, Canada

Cáceres G, López ME, Cadiz MI, Yoshida GM, Jedlicki A, Palma-Véjares R et al. (2019) Fine mapping using whole-genome sequencing confirms anti-Müllerian hormone as a major gene for sex determination in farmed Nile tilapia (*Oreochromis niloticus L.*). Genes Genomes Genet 9:3213–3223

Chen J, Fan Z, Tan D, Jiang D, Wang D (2018) A review of genetic advances related to sex control and manipulation in tilapia. J World Aquac Soc 49:277–291

Chen S, Zhang G, Shao C, Huang Q, Liu G, Zhang P 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

Cnaani A, Lee B-Y, Zilberman N, Ozouf-Costaz C, Hulata G, Ron M et al. (2008) Genetics of sex determination in tilapiine species. Sex Dev 2:43–54

Cui Z, Liu Y, Wang W, Wang Q, Zhang N, Lin F et al. (2017) Genome editing reveals *dmrt1* as an essential male sex-determining gene in Chinese tongue sole (*Cynoglossus semilaevis*). Sci Rep 7:1–10

Curzon AY, Shirak A, Dor L, Zak T, Benet-Perelberg A, Seroussi E et al. (2019) Hybrid origin of the Thai-chitralada tilapia strain using DNA barcoding and microsatellite analysis. Isr J Aquac 71:1–7

Dor L, Shirak A, Gorshkov S, Band MR, Korol A, Ronin Y et al. (2014) Construction of a microsatellites-based linkage map for the White grouper (*Epinephelus aeneus*). Genes Genomes Genet 4:1455–1464

Elder HY, Garrod DJ (1961) A natural hybrid of Tilapia nigra and Tilapia leucosticta from Lake Naivasha, Kenya Colony. Nature 191:722–724

Eshel O, Shirak A, Dor L, Band M, Zak T, Markovich-Gordon M et al. (2014) Identification of male-specific amh duplication, sexually differentially expressed genes and microRNAs at early embryonic development of Nile tilapia (*Oreochromis niloticus*). BMC Genomics 15:774

Eshel O, Shirak A, Weller JI, Hulata G, Ron M (2012) Linkage and physical mapping of sex region on LG23 of Nile Tilapia (*Oreochromis niloticus*). Genes Genomes Genet 2:35–42

Eshel O, Shirak A, Weller JI, Slossman T, Hulata G, Cnaani A et al. (2011) Fine-mapping of a locus on linkage group 23 for sex determination in Nile tilapia (*Oreochromis niloticus*). Anim Genet 42:222–224

Ezaz MT, Harvey SC, Boonphakdee C, Teale AJ, McAndrew BJ, Penman DJ (2004) Isolation and physical mapping of sex-linked AFLP markers in Nile tilapia (*Oreochromis niloticus* L.). Mar Biotechnol 6:435–445

Gammerdinger WJ, Conte MA, Sandkam BA, Penman DJ, Kocher TD (2019) Characterization of sex chromosomes in three deeply diverged species of Pseudocrenilabrinae (Teleostei:Cichlidae). Hydrobiologia 832:397–408

Gubbay J, Collignon J, Koopman P, Capel B, Economou A, Münsterberg A et al. (1990) A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature 346:245–250

Guerrero RD, Shelton WL (1974) An aceto-carmine squash method for sexing juvenile fishes. Progress Fish Culturist 36:56

Han YJ, Ma SF, Yourek G, Park Y-D, Garcia JGN (2011) A transcribed pseudogene of *MYLK* promotes cell proliferation. FASEB J 25:2305–2312

Harrison DA (2007) Sex determination: controlling the master. Curr Biol 17:R328–R330.

Hattori RS, Murai Y, Oura M, Masuda S, Majhi SK, Sakamoto T et al. (2012) A Y-linked anti-Mullerian hormone duplication takes over a critical role in sex determination. Proc Natl Acad Sci 109:2955–2959

Lee B-Y, Hulata G, Kocher TD (2004) Two unlinked loci controlling the sex of Blue tilapia (*Oreochromis aureus*). Heredity 92:543–549

Lee B-Y, Penman DJ, Kocher TD (2003) Identification of a sex-determining region in Nile tilapia (*Oreochromis niloticus*) using bulked segregant analysis. Anim Genet 34:379–383

Li M, Sun Y, Zhao J, Shi H, Zeng S, Ye K et al. (2015) A tandem duplicate of anti-Müllerian Hormone with a missense snp on the y chromosome is essential for male sex determination in Nile tilapia, *Oreochromis niloticus*. PLoS Genet 11:1–23.

Li XY, Gui JF (2018) Diverse and variable sex determination mechanisms in vertebrates. Sci China Life Sci 61:1503–1514

Lin G, Chua E, Orban L, Yue GH (2016) Mapping QTL for sex and growth traits in salt-tolerant tilapia (*Oreochromis* spp. X *O. mossambicus*) (JL Rummer, ed). PLoS ONE 11:e0166723

Lind CE, Agyakwah SK, Attipoe FY, Nugent C, Crooijmans RPMA, Toguyeni A (2019) Genetic diversity of Nile tilapia (*Oreochromis niloticus*) throughout West Africa. Sci Rep 9:16767

Liu F, Sun F, Li J, Xia JH, Lin G, Tu RJ 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

Lühmann LM, Knorr C, Hörstgen-Schwark G, Wessels S (2012) First evidence for family-specific QTL for temperature-dependent sex reversal in Nile tilapia *(Oreochromis niloticus).* Sex Dev 6:247–256

Macaranas JM, Mather PB, Lal SN, Vereivalu T, Lagibalavu M, Capra MF (1997) Genotype and environment: a comparative evaluation of four tilapia stocks in Fiji. Aquaculture 150:11–24

Mair GC, Lakapunrat S, Jere WL, Bart A (2004) Comparisons of reproductive parameters among improved strains of Nile tilapia. In: Bolivar R, Mair GC, Fitzsimmons K (eds) New dimensions of farmed tilapia, proceedings from the sixth international symposium on tilapia in Aquaculture, Manila, the Philippines, p 142–156

Matsuda M, Nagahama Y, Shinomiya A, Sato T, Matsuda C, Kobayashi T et al. (2002) *DMY* is a Y-specific DM-domain gene required for male development in the medaka fish. Nature 417:559–563

Mei J, Gui JF (2015) Genetic basis and biotechnological manipulation of sexual dimorphism and sex determination in fish. Sci China Life Sci 58:124–136

Nanda I, Kondo M, Hornung U, Asakawa S, Winkler C, Shimizu A et al. (2002) A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. Proc Natl Acad Sci 99:11778–11783

Ndiwa TC, Nyingi DW, Claude J, Agnèse J-F (2016) Morphological variations of wild populations of Nile tilapia (*Oreochromis niloticus*) living in extreme environmental conditions in the Kenyan Rift-Valley. Environ Biol Fishes 99:473–485

Nyingi D, De Vos L, Aman R, Agnèse J-F (2009) Genetic characterization of an unknown and endangered native population of the Nile tilapia *Oreochromis niloticus* (Linnaeus, 1758) (Cichlidae; Teleostei) in the Loboi Swamp (Kenya). Aquaculture 297:57–63

Palaiokostas C, Bekaert M, Taggart JB, Gharbi K, McAndrew BJ, Chatain B et al. (2015) A new SNP-based vision of the genetics of sex determination in European sea bass (*Dicentrarchus labrax*). Genet Sel Evol 47:68

Palaiokostas C, Bekaert M, Khan MGQ, Taggart JB, Gharbi K, McAndrew BJ et al. (2013) Mapping and validation of the major sex-determining region in Nile tilapia (*Oreochromis niloticus* L.) using RAD sequencing (L Orban, ed). PLoS ONE 8:e68389

Pruginin Y, Rothbard S, Wohlfarth G, Halevy A, Moav R, Hulata G (1975) All-male broods of Tilapia nilotica × T. aurea hybrids. Aquaculture 6:11–21

Rognon X, Guyomard R (2003) Large extent of mitochondrial DNA transfer from *Oreochromis aureus* to *O. niloticus* in West Africa. Mol Ecol 12:435–445

Rondeau EB, Laurie CV, Johnson SC, Koop BF (2016) A PCR assay detects a male-specific duplicated copy of anti-Müllerian hormone (*amh*) in the Lingcod (*Ophiodon elongatus*). BMC Res Notes 9:10–13

Seroussi E, Klompus S, Silanikove M, Krifucks O, Shapiro F, Gertler A et al. (2013) Nonbactericidal secreted phospholipase A2s are potential anti-inflammatory factors in the mammary gland. Immunogenetics 65:861–871

Shirak A, Cohen-Zinder M, Barroso RM, Seroussi E, Ron M, Hulata G (2009) DNA barcoding of Israeli indigenous and introduced cichlids. Isr J Aquac 61:83–88

Shirak A, Golik M, Lee B-Y, Howe AE, Kocher TD, Hulata G et al. (2008) Copy number variation of lipocalin family genes for male-specific proteins in tilapia and its association with gender. Heredity 101:405–415

Shirak A, Seroussi E, Cnaani A, Howe AE, Domokhovsky R, Zilberman N et al. (2006) *Amh* and *Dmrta2* genes map to tilapia (*Oreochromis* spp.) linkage group 23 within quantitative trait locus regions for sex determination. Genetics 174:1573–1581

Shirak A, Palti Y, Cnaani A, Korol A, Hulata G, Ron M et al. (2002) Association between loci with deleterious alleles and distorted sex ratios in an inbred line of tilapia (*Oreochromis aureus*). J Hered 93:270–276

Shirak A, Zak T, Dor L, Benet-Perlberg A, Weller JI, Ron M et al. (2018) Quantitative trait loci on LGs 9 and 14 affect the reproductive interaction between two *Oreochromis* species, *O. niloticus* and *O. aureus*. Heredity 122:341–353

Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths BL, Smith MJ et al. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature 346:240–244

Sissao R, D'Cotta H, Baroiller JF, Toguyeni A (2019) Mismatches between the genetic and phenotypic sex in the wild Kou population of Nile tilapia *Oreochromis niloticus*. PeerJ 7:e7709

Smith CA, Roeszler KN, Ohnesorg T, Cummins DM, Farlie PG, Doran TJ et al. (2009) The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. Nature 461:267–271

Sun Y-L, Jiang D-N, Zeng S, Hu C-J, Ye K, Yang C et al. (2014) Screening and characterization of sex-linked DNA markers and marker-assisted selection in the Nile tilapia (*Oreochromis niloticus*). Aquaculture 433:19–27

Syaifudin M, Bekaert M, Taggart JB, Bartie KL, Wehner S, Palaiokostas C et al. (2019) Species-specific marker discovery in tilapia. Sci Rep 9:13001

Taslima K, Wehner S, Taggart JB, de Verdal H, Benzie JAH, Bekaert M et al. (2020) Sex determination in the GIFT strain of tilapia is controlled by a locus in linkage group 23. BMC Genet 21:49

Tayamen MM (2004) Nationwide dissemination of GET-EXCEL tilapia in the Philippines. In: Bolivar R, Mair GC, Fitzsimmons K (eds) New dimensions of farmed tilapia, proceedings from the sixth international symposium on tilapia in Aquaculture, Manila, the Philippines, p 75–88

Volff J-N, Nanda I, Schmid M, Schartl M (2007) Governing sex determination in fish: regulatory putsches and ephemeral dictators. Sex Dev 1:85–99

Volff J-N, Schartl M (2001) Variability of genetic sex determination in poeciliid fishes. Genetica 111:101–110

Volff J-N, Zarkower D, Bardwell VJ, Schartl M (2003) Evolutionary dynamics of the DM domain gene family in Metazoans. J Mol Evol 57:S241–S249.

Wessels S, Krause I, Floren C, Schütz E, Beck J, Knorr C (2017) ddRADseq reveals determinants for temperature-dependent sex reversal in Nile tilapia on LG23. BMC Genomics 18:531

Wessels S, Sharifi RA, Luehmann LM, Rueangsri S, Krause I, Pach S et al. (2014) Allelic variant in the anti-Müllerian hormone gene leads to autosomal and temperature-dependent sex reversal in a selected Nile Tilapia line (L Orbán, ed). PLoS ONE 9:e104795

Wohlfarth GW, Wedekind H (1991) The heredity of sex determination in tilapias. Aquaculture 92:143–156

Woram RA, McGowan C, Stout JA, Gharbi K, Ferguson MM, Hoyheim B et al. (2004) A genetic linkage map for Arctic char (*Salvelinus alpinus*): evidence for higher recombination rates and

segregation distortion in hybrid versus pure strain mapping parents. Genome 47:304–315

Yamamoto Y, Zhang Y, Sarida M, Hattori RS, Strüssmann CA (2014) Coexistence of genotypic and temperature-dependent sex determination in pejerrey *Odontesthes bonariensis*. PLoS ONE 9:1–8

Yano A, Nicol B, Jouanno E, Quillet E, Fostier A, Guyomard R et al. (2013) The sexually dimorphic on the Y-chromosome gene (*sdY*) is a conserved male-specific Y-chromosome sequence in many salmonids. Evol Appl 6:486–496

Yoshimoto S, Okada E, Umemoto H, Tamura K, Uno Y, Nishida-Umehara C et al. (2008) A W-linked DM-domain gene, *DM-W*, participates in primary ovary development in *Xenopus laevis*. Proc Natl Acad Sci 105:2469–2474

Zhu H, Liu Z, Lu M, Gaao F, Ke X, Ma D et al. (2016) Screening and identification of a microsatellite marker associated with sex in Wami tilapia, *Oreochromis urolepis hornorum*. J Genet 95:283–289

Zimmermann S, Natividad JM (2004) Comparative pond performance evaluation of genomar supreme tilapia GST 1 and GST3 Groups. In: Bolivar R, Mair GC, Fitzsimmons K (eds) New dimensions of farmed tilapia, proceedings from the sixth international symposium on tilapia in Aquaculture, Manila, the Philippines, p 89