A novel c.1759T>G variant in follicle-stimulating hormone-receptor gene is concordant with male determination in the flathead grey mullet (Mugil cephalus)

Arie Y. Curzon,1,2 Lior Dor,1,2 Andrey Shirak,1 Iris Meiri-Ashkenazi,3 Hana Rosenfeld,3 Micha Ron,1 and Eyal Seroussi 1,*

*Corresponding author: Agricultural Research Organization (ARO), Rishon LeTsiyon, P.O.B 15159 7528809, Israel. seroussi@agri.huji.ac.il

Abstract

Various master key regulators (MKRs) that control a binary switch of sex determination (SD) have been found in fish; these provide an excellent model for the study of vertebrate genetic SD. The SD region in flathead grey mullet has been previously mapped to a 1 Mbp region harboring 27 genes, of which one is follicle-stimulating hormone receptor (fshr). Although this gene is involved in gonad differentiation and function, it has not been considered as an MKR of SD. We systematically investigated polymorphism in mullet fshr using DNA shotgun sequences, and compared them between males and females. Capable of encoding nonconservative amino acid substitutions, c.1732G>A and c.1759T>G exhibited association with sex on a population level (N = 83; P ≤ 6.7 × 10⁻¹⁰). Hence, 1732A and 1759G represent a male-specific haplotype of the gene, designated as “fshry.” Additional flanking SNPs showed a weaker degree of association with sex, delimiting the SD critical region to 143 nucleotides on exon 14. Lack of homozygotes for fshry, and the resulting divergence from Hardy–Weinberg equilibrium (N = 170; P ≤ 3.9 × 10⁻⁵), were compatible with a male heterogametic model (XY/XX). Capable of replacing a phenylalanine with valine, c.1759T>G alters a conserved position across the sixth transmembrane domain of vertebrate FSHRs. Amino acid substitutions in this position in vertebrates are frequently associated with constant receptor activation and consequently with FSH/FSHR signaling alteration; thus, indicating a potential role of fshry as an MKR of SD.

Keywords: sex determination; XX/XY; bony fishes; Genetics of Sex; gonadotropin receptor

Introduction

Sex dimorphism in gonochoristic species is associated with the differentiation of gonads, but also includes other important behavioral, morphological, and physiological characteristics (Williams and Carroll 2009). Sex determination (SD) is variable and can involve genetic, environmental, and social factors. Studies of species with a genetic SD system, show that in most cases a single gene plays the role of a master key regulator (MKR), initiating the dichotomic division of sexes (males and females). Different MKRs were found in vertebrate and insect taxa including sry (Gubbay et al. 1990; Sinclair et al. 1990), sox3 (Takehana et al. 2014), amh (Eshel et al. 2014; Li et al. 2015), amhr2 (Kamiya et al. 2012), amrl1 (Raymond et al. 2000; Matsuda et al. 2002; Nanda et al. 2002), gsdf (Myosho et al. 2012), kds17b1 (Koyama et al. 2019), and sdY (Yano et al. 2012), and in many cases these are found to be associated with gene duplications (Eshel et al. 2014; Curzon et al. 2020). In Table 1, we surveyed the types of reported sequence variation of SD genes in vertebrates. Copy-number variations (CNVs) were the underlying polymorphism in eight instances whereas nonstructural polymorphisms were evident in five SD genes.

The existence of common MKRs for SD in taxonomically distant species suggests that the number of MKRs is limited to a set of factors, which belong to a conserved pathway (Marshall Graves and Peichl 2010; Graves 2013) that triggers the primary differentiation of the bipotential gonad up to sex hormone synthesis (Kim and Capel 2006). However, the surprising finding of an MKR that had not been previously associated with sex differentiation, the immune-related gene (sdY) in salmonids (Yano et al. 2012, 2013), indicates a more complex regulation of SD (Herpin and Schartl 2015). SD among different fish is considered to be extremely variable, even between families of the same genus for which different MKRs have been found, e.g. Orzysz (Matsuda and Sakaizumi 2016), Cichlidae (Ser et al. 2010), and zebrafish (Kossack and Draper 2019). Thus, fish are an excellent model for the elucidation of the SD pathway and the vertebrate sex differentiation cascade (Baroiller et al. 1999; Schartl 2004; Capel 2017).

The flathead grey mullet (Mugil cephalus) is the most widespread fish of the Mugilidae family and populates both subtropical and temperate coastal waters in the major oceans. Despite its abundance and the fact that it has adapted to a wide range of environments, it is not fully domesticated and fry-production for aquaculture is mainly dependent on collection from river estuaries. The flesh and roe of this mullet are highly prized commodities in Asian and Mediterranean markets (Aizen et al. 2005). The
ovaries, processed with salt, are sold as ‘Botarga Caviar’ or as ‘Karasumi’ in Asia, and are considered a delicacy in countries around the Mediterranean, Japan, and Taiwan (Hung and Shaw 2006; Katselis et al. 2016). Thus, on account of a faster growth rate and the high value of their products, production of all-female populations is economically advantageous.

The addition of sex steroids to larva feed can cause sex inversion in many fish species, and is often practiced in commercial production, giving rise to environmental and food safety concerns (Malilila et al. 2015; Chen et al. 2018). These worries could be allayed by technologies for mono-sex production through parent sex-reversal that do not require the use of steroids. Dor et al. (2016) analyzed the sex of M. cephalus progeny in two independent families, and based on genetic markers, showed that the alleles linked with maleness were transmitted from the fathers, indicating that mullets have a genomic XX/XY SD system. Thus, a female mono-sex population can be constructed by female-to-male sex reversal, followed by mating of the XX males with females. As opposed to sex reversing fish that go directly for human consumption, mono-sex production based on such mating presents an environmental friendly and safer alternative. Since sex reversal occurs at an early age prior to the appearance of any external signs of morphological sex differentiation, a genetic marker for SD is required to discriminate between sex-reversed and normal individuals of the flathead grey mullet. The most reliable genetic marker for sex would be the causal variation in the MKR of SD.

Materials and methods
Fish sampling and sexing
Adult flathead grey mullet specimens from two full-sib families (n = 6 and n = 7) and from a wild population (n = 70) were dissected and sexed based on external observation of gonad morphology (n = 83). The wild-sexed population consisted of fish that have been reared by commercial growers from fry caught in river estuaries around the Israeli Mediterranean coast in years 2016 (n = 46), 2017 (n = 7), and 2018 (n = 12), and of fish reared from fry caught in Ebro River, Spain (n = 5). Of these fish, the gonads of 27 individuals (17 males and 10 females) were also examined microscopically in order to improve confidence in the visual sexing. A caudal fin sample was taken from each of the adult specimens and preserved in ethanol (90–96%). Fin samples were also taken from two groups of 28 dph larvae (Group A, n = 7; Group B, n = 8, Dor et al. 2020), and from 87 nonsexed fingerlings of a wild population in year 2019 in river estuaries along the Israeli Mediterranean coast.

Sequence assembly of fshr in M. cephalus using whole-genome sequencing
Mugil cephalus fshr raw reads and previously built SOAPdenovo scaffolds (Dor et al. 2016, 2020) of a male and female were assembled separately. The Gap5 (Staden et al. 2000) wrapper and the
BWA program were used for alignment of reads and scaffolds against the masked gene sequence of *Dicentrarchus labrax* *fshr* (GenBank accession no. FQ310507). Masking was performed using CENSOR (Kohany et al. 2006) and Tilapia Repeat Masker version 2 (Shirak et al. 2010). Aligned sequences were then used for *de novo* assembly using Gap4 (Staden et al. 2000). Raw reads and scaffolds were used again for alignment against the newly assembled sequences using Gap5. A consensus sequence was obtained and used for further alignment. This process was repeated several times until the entire gene assembly was completed. Exon borders were deduced using mRNA sequences that had been previously built using the Trinity software (Dor et al. 2020).

### DNA extraction and sequencing procedure

A sample of the caudal fin (100–200 mg) was used for DNA extraction using a commercial kit (MasterPure DNA Purification, Madison). Polymerase chain reaction (PCR) was performed using relevant primers that were designed using Primer3 (Untergasser et al. 2012) (Table 2), and the Bio-X-ACT Long kit (Bioline Ltd., London, UK) according to the manufacturers’ instructions under the following conditions: 36 cycles for 30 s at 94°C, 30 s at 66°C, and 30 s at 72°C. Thereafter, the PCR products were excised from the gel and purified with a kit (Montage Gel Extraction, Millipore, Bedford, MA). Sanger sequencing was conducted from both directions of the amplified genomic region. For the purpose of allele-specific PCR, “X” specific and “Y” specific forward primers were designed following the method of Liu et al. (2012), including a mismatch in the third nucleotide position from the 3’ end. Annealing temperature was raised to 69°C.

### Characterization of sex-specific alleles for *fshr*

Analysis of genomic alleles was based on whole-genome sequencing (WGS) of a male and female deposited in the Short Read Archive (SRA) under project IDs PRJEB12265 and PRJEB34342, respectively. Based on sequence variation along the assembled *fshr* gene and read-pair information, sex-specific alleles were manually separated using Gap5’s viewer. The assembly of *fshr* (ENA accession LR860658) was also deposited in the BAM format, which provided further verification for each read assembly (over 3000 reads, ENA accession ERX4280005). In search for sequence variants, we preferentially targeted the gene regions that are likely to be functional as has been previously described (Dor et al. 2020). Briefly, most variants were excluded by observing their fitness to the XY system model in the male and female WGS data. Variants that fit the model were tested for association with sex by Sanger sequencing of sample of individuals from two families (families A1 and B, n = 17). Variants that passed this test were further examined on the population level by Sanger sequencing (n = 83, using primer pair #1, Table 2).

### Fshr gene expression

Expression of previously established RNA-seq libraries (Dor et al. 2020) included two groups of larvae A and B (n = 7 and 8), female and male brains (n = 3 and 2, respectively), female and male gonads (n = 2 and 1, respectively). Existence of *fshr* variants were examined in these libraries, using the BLASTN algorithm for alignment with 32 bp words: GCATGGCTCAAGGGATGCGGTC and GCATGGCTCAAGGGATGCGGTC. This represented the male-specific allele of *fshr* (*fshx*) and the nonspecific allele of *fshr* (*fshy*), respectively. All available *M. cephalus* SRA RNA-Seq submissions were screened including 92 runs (ERX353607-ERX3536169, a total of 1,038,774,806 reads) originating from Israel (Dor et al. 2020) and 4 runs from other sources (SRX153305 and SRX1817285-8, a total of 140,071,760 reads). No *fshr* reads were detected in the latter source, whereas 81 reads were found to contain the *fshx* probe word, most of them (85%) in adult male tests. The *fshr* expression values were calculated and normalized in a previous study using Trimmed mean of values (TMM) (Dor et al. 2020).

### Statistics

The JMP® statistical package (Pro 13, SAS Institute, Cary, NC) was used for conducting statistical tests. Fisher’s exact Chi-squared test was applied for association study of individual genetic markers and sex. Pearson Chi-squared test was applied for testing deviation from Hardy–Weinberg equilibrium.

### Data availability

DNA- and RNA-seq reads are available at the Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) under accession numbers PRJEB12265 and PRJEB34342. The *M. cephalus* assembled and annotated transcriptome is available at http://cowry.agri.huji.ac.il/. Supplementary material includes Figure S1 Comparison of predicted Fshr of *M. cephalus* with other vertebrate FSHRs (complete alignment) and the annotated mullet *fshr* sequences (File S2). Supplementary material is available at figshare: DOI: https://doi.org/10.25387/g3.13352078.

### Results

#### Assembly of *fshr* in *M. cephalus*

Our established genomic libraries allowed the assembly of the whole *fshr* gene in two full-sibs of *M. cephalus*, representing both sexes. Separation of sequencing fragments based on polymorphism in these paired-read sequences revealed four different

### Table 2 PCR primer pairs

| Pair | Sequence (5’ to 3’) | Product size (bp) | Target of amplification* |
|------|---------------------|-------------------|-------------------------|
| 1    | GGATGTGGAGGATCTGGTGT| 221               | SD region (Exon 14)     |
| 2    | GAGACGGCCAAGAAGGAGAT| 491/481           | Upstream of SD region (Exon 14) |
| 3    | GCCGGCCAGGCGACGCTCA| 618/608           | fshy                    |
| 4    | GCCGGCCAGGCGACGCTCA| 618/608           | fshx                    |
| 5    | CGGCTGAGCATCTGATGTG| 230               | Non-synonymous SNP (Exon 1) |

* *fshr* represents the putative non-specific allele of *fshr*, while *fshx* represents the male-specific allele.
fsr alleles, with no evidence of a fifth allele that might indicate a CNV. Annotation based on the RNA-seq data (Dor et al. 2020) was used to characterize the fshr gene structure with its 14 exons, its conserved exon–intron boundaries, and its putative protein sequence, all of which were found similar to those of known orthologous FSHRs (Table 3, Supplementary Figure S1, ENA accession No. LR860658, Additional File 2). Using this predicted Fshr protein and the BLASTp program, we searched GenBank for the reference sequences of mammals, amphibians, birds, and reptiles. This search indicated that the assembled gene was fshr, i.e. 50% identity and 67% similarity to human FSHR, NP_000136; 51 and 66% similarity to clawed frog Fshr, XP_017949625; 50% identity and 66% similarity to duck FSHR, XP_005012152; and 51% identity and 65% similarity to python Fshr, XP_007491488. However, nomenclature for many similar proteins in other fish differed. 

Analysis of polymorphism in fshr

According to the predicted transcript, two nonsynonymous male-specific SNPs were found in exon 14 at 1732 and 1759 nucleotide positions of the fshr assembled sequence (Table 4). Using PCR primers (pair 1, Table 2), these SNPs were investigated in two full-sib families (n = 6 and n = 7), and were found to be fully associated with sex. Another two nonsynonymous SNPs in exons 1 and 14 (Table 4, nucleotide positions at 131 and 2047 bp, respectively) were found in fshr’s assembly; however, these did not match the expected XY/XX SD model, since in the two established genomic libraries of both sexes used, both SNPs were homozygous for each variant. Using PCR primers (pairs 2 and 5, Table 2), we amplified and sequenced the relevant regions in exons 1 and 14 in a small panel of WT males (n = 4) and females (n = 4), precluding association with sex of these other SNPs. Moreover, they did not conform to any simple genetic sex chromosome model (XY/XX or WZ/ZZ), as homozygous individuals were found for each variant. Apart from these instances, Sanger sequencing revealed another two nonsynonymous SNPs at nucleotide position 149 and 1781 bp, in exons 1 and 14, respectively (Table 4), however, these were not associated with sex.

Two nonsynonymous substitutions in fshr are male-specific

The two male-specific SNPs, c.1732G>A and c.1759T>G representing fshr, found in the full-sib families which are capable of encoding nonconservative amino acid substitutions (Table 4; Figure 1), were further investigated for association with sex in WT samples. All tested samples (n = 83) demonstrated that these two SNPs were a stable haplotype which was significantly associated with sex, with only 4.8% of mismatches (P < 6.7 × 10^-10, Table 5). Samples showing discordance between fshr haplotype and sex were all phenotyped for sex without microscope inspection. In addition, neither Sanger sequencing of these samples (n = 83), nor the nonsexed fingerlings (n = 87), were homozygous for the fshr haplotype, even though, based on Hardy–Weinberg equilibrium, their expected frequency was 6.6% (Table 6). Most importantly, on the population level, additional SNPs flanking the fshr haplotype, including other nonsynonymous substitutions (Table 4, positions 131–1689 and 141–1875), did not show significant association with sex, and thus the observed SD associated region is delimited between positions 1732 and 1875, to 143 nucleotides on exon 14 (Table 4). It should be noted that we have...

---

**Table 3** Genomic organization of the M. cephalus fshr gene, based on the male fshry variant

| Intron | Exon | Intron |
|--------|------|--------|
| No.    | Size | Size   |
| ATGATG | 1    | 188    | G TGCTG TGTgaagctgc 3,119 |
| tgcacacgAGGAACACT | 2 72 | GAGACTgaagctttata 120 |
| tctattctacGACACT | 3 75 | TTACAGTTgaagctgttt 526 |
| cgccttctacAAGCAT | 4 75 | TGAGAATgaagctgaa 107 |
| tcttcctcagAACAT | 5 75 | GTTATATgaagctttct 885 |
| tctctctcagGACACT | 6 75 | TCACTGAgaagctgttc 319 |
| tctctctcagTCACCT | 7 81 | GAAGAAGTgaagctgct 1,506 |
| tctctctcagAGTGGCT | 8 69 | AAGACAGTgaagctct 101 |
| ttcctctcagATTTCT | 9 75 | AGTGCTgaagctgca 380 |
| tttctctcagGTAGCAT | 10 186 | AAAACAGTgaagctgct 115 |
| tttctctcagGATCAG | 11 218 | TGTGCTgaagctggtt 140 |
| tttctctcagGCAGCA | 12 182 | TTCTCAGTgaagctttg 97 |
| tttctctcagGTGCTT | 13 192 | AGCAAAAGTgaagctgta 1,046 |
| tttctctcagGTGAGC | 14 529 | CGTGGCTGTAG  |

---

**Table 4** Polymorphism in M. cephalus predicted fshr gene

| Exon 1 | Exon 14 |
|--------|---------|
| Position | 131 | 149 | 1689 | 1732 | 1759 | 1781 | 1875 | 1917 | 2031 | 2047 | 2122 |
| Nucleotide | T/C | T/C | C/A | G/A | T/G | C/A | C/T | C/T | G/A | G/A | C/A |
| AA | 1/P | M/T | T/T | V/M | F/V | A/D | I/I | R/R | S/S | A/T | 15/5 |
| Males | 6/2 | 4/4 | 60/26 | 44/42 | 44/42 | 78/8 | 2/2 | 18/6 | 18/6 | 3/21 | 10/14 |
| Females | 7/1 | 3/5 | 80/0 | 77/3 | 77/3 | 78/2 | 4/0 | 9/11 | 6/14 | 5/15 | 15/5 |

---

* Intron and exon sizes are given in base pairs, and their sequences are written in lowercase and uppercase letters, respectively. The first and last two bases of the introns are presented in bold type (gt and ag for donor and acceptor splice sites, respectively). The initiation and stop codons are shown in bold type (ATG and TAG) starting from the initiation codon. The putative transcript sizes of the fshr gene are 10,553 and 2,118 bp, respectively.

* Two nonsynonymous SNPs comprising a haplotype that fits an XY model of SD are shown in bold underlined font. The “Y” fshry haplotype is presented on the right side of the divider.

* Nucleotide position in the coding DNA sequence (in bp). Nucleotide and amino acid (AA) variation is denoted in both sides of divider.

* Number of counts of alternative alleles in each position for samples of males and females, respectively.
ing the primer designed for the "A" variant. As opposed to female formed for each individual. Only males gave a product when us-

were evident within familial structures only. However, the associ-

positions 1732 and 1759 of the predicted cds) are marked by an arrow.

Conservation of FSHR protein among vertebrate species

The protein sequence of the fshry variant was used as a template for BLASTP search against GenBank (nonredundant protein sequences). While the V/M substitution (nt 1732 on exon 14) was found in other organisms, the F/V substitution (nt 1759 on exon 14) is novel and causes an alteration in amino acid, which is conserved across vertebrate species (Figures 1 and 2; Supplementary Figure S1).

Analysis of ploidy of fshr by allele specific PCR

As indicated in Introduction, scenarios in which the SD gene is in a diploidic state, in which each individual carries two alleles for the gene (paternal and maternal), two males and two females were analyzed by allele-specific PCR. Two forward primers were designed to amplify both variants of the c.1732G>A alteration (primer pair #3 and #4, Table 2) and two PCR reactions were performed for each individual. Only males gave a product when using the primer designed for the "A" variant. As opposed to female analyzed additional sex-linked polymorphism in candidate genes besides fshr (Dor et al. 2020). In these genes, associations with sex were evident within familial structures only. However, the association with genetic factor affecting sex on the population level was displayed only for exon 14 polymorphism of fshr.

Analysis of full-sib groups for sex-biased expression

In a previous study, two full-sib groups of fry (28 dph) of a mullet family (n = 15), were assigned to different sex groups, designated as A and B, with differential expression of genes that are located on the critical SD chromosomal region (Dor et al. 2020). We geno-
typed fshr for these groups and revealed that they represent fshr haplotypes "XX" and "XY," respectively (Table 6). Interestingly, fshry was not found in any expression library, whereas expression of fshrx varied between larvae, adult brains and gonads, with a tendency toward higher values in male tissues (Figure 3). Since the promoter and gene structure of fshrx and fshry were similar, we expected that in males both genes would be similarly expressed. However, as detailed in "Method" section, BLASTN search using a 32bp word spanning the critical variation sequence revealed only 0, 3, and 68 fshr reads in male larvae, brain, and testis, respectively, but no fshry reads. This highly signif-
ificant result (P < 4 × 10⁻²⁵) and the differentially expressed pattern of fshrx (Figure 3) are compatible with the observation of Dor et al. (2020) that genes in the LG9 SD region are sex-biased differentially expressed and thus may be subjected to partial chromosome silencing.

Discussion

Multiple genetic SD factors are considered to be upstream of gonadal hormones, and the SD cascade is believed to be initiated by a single upstream MKR (Nakamura 2010). Many distant vertebrate organisms share common MKRs, suggesting that only a limited number of genes can initiate the fate of the bipotential gonad (Marshall Graves and Peichel 2010; Graves 2013). For example, dmrt1 determines sex in Medaka (Nanda et al. 2002), Spotted scat (Mustapha et al. 2018), African clawed frog (Yoshimoto et al. 2008), Chicken (Smith et al. 2009), and Half-smooth tongue sole (Cui et al. 2017). Anti Mullerian hormone is a MKR of Nile tilapia (Eshel et al. 2014), Patagonian pejerrey

**Table 6** Contingency table comparing observed counts and frequencies of fshr haplotypes with those expected from Hardy–Weinberg equilibrium in the genotyped population (N = 170)².

| fshr haplotype | Actual count (frequency) | Expected count (frequency) |
|----------------|--------------------------|----------------------------|
| "XX"          | 83 (0.48)                | 94 (0.55)                  |
| "XY"          | 87 (0.52)                | 65 (0.38)                  |
| "YY"          | 0 (0)                    | 12 (0.07)                  |

² Pearson Chi-square test for divergence of fshr haplotypes from Hardy–Weinberg equilibrium (P ≤ 3.9 × 10⁻³³).

sequences, the sequenced products of males (608-618 bp), for both variants “A/G,” showed complete homozygosity over multiple SNPs (including positions 1759-2112, Table 4), suggesting no evidence for CNV in males.
The network properties and their degree of suggest that SD may act as a network, through genes with antagonistic (Tilmann and Capel 2002). Furthermore, new models it difficult to identify MKRs and their position in the sex case-
during early development (the early developmental period), ren-
However, the initial low level of MKR’s extra-gonadal expression
sion genes (cassettes) among different species (Capel 2017 ).
(Figure S1). The highly conserved phenylalanine (F) residue (nt 1759 in exon 14, Table 4), which was
indicated above the amino acid alignment. Identical and similar amino acid residues in at least two of four sequences are indicated by a black and grey background, respectively. White boxes indicate non-conservative amino acid changes between the proteins.

(Hattori et al. 2012), lingcod (Rondeau et al. 2016), and silverside (Bej et al. 2019). Moreover, a conserved cascade for initial activa-
tion of the gonad is implied by conservation of sex-biased expres-
sion genes (cassettes) among different species (Capel 2017). However, the initial low level of MKR’s extra-gonadal expression during early development (the early developmental period), renders it difficult to identify MKRs and their position in the sex cascade (Tilmann and Capel 2002). Furthermore, new models suggest that SD may act as a network, through genes with antagonistic effects on the manifestation of sex, without a conserved order (Capel 2017). The network properties and their degree of conservation require further study (Herpin and Schartl 2015). Fish are an excellent model for SD studies, as different, frequ-
cently closely related species exhibit varying MKRs (Herpin and Schartl 2015). Moreover, MKRs in fish include genes that are pre-
sumed to be involved in SD by controlling different types of mechanisms at different levels, including differentiation of supporting cells, proliferation of germ cells, and regulation of estrogen and androgen production, allowing a broader understanding of the different mechanisms (Capel 2017). The elucidation of whole MKR sets in different species is important for a wider understanding of the SD cascade (Herpin and Schartl 2015). Here, we have for the first time detected a putative MKR of M. cephalus,

| Human | Mouse | Mugil Female | Mugil Male | Seriola | Lavrax | Tilapia | Trout | Fugu |
|-------|-------|--------------|-----------|---------|--------|---------|-------|------|
| Cytoplasmic | Transmembrane helix #6 | Extracellular |
| IGYIVKYPVNLQLYSSST- | - | - |
| LSIYLMPPSSYPASHALT | - | - |
| LSIYLMPPSSYPASHALT | - | - |
| ALGYVRKPSHATGTT | - | - |
| 601 | 610 | 620 | 630 | 640 | 650 |

**Figure 2** Comparison of predicted Fshr of M. cephalus with other vertebrate FSHRs (GenBank: Human NP_000136; Mouse NP_038551; Tilapia NP_001266517; Seriola XP_022599280; Lavrax AAV48628; Trout NP_001117799; Fugu XP_029691459). Amino acid numbering follows the complete gene alignment presented in supplementary file (Figure S1). The highly conserved phenylalanine (F) residue (nt 1759 in exon 14, Table 4), which was

**Figure 3** Expression data for fshr in terms of normalized Trimmed Mean of values (TMM) for 28 dph larvae groups: A (“XX”) and B (“XY”), for adult brains and for gonads of both sexes. Standard error bar is shown at the top of columns with biological replicates.

| Observed distribution | Expected distribution | p^b |
|-----------------------|-----------------------|-----|
| 1 | 0 | 2.97 \times 10^{-7} |
| 11.3 | 21.4 | 10.3 |
a Hardy–Weinberg equilibrium: p^2, 2pq, q^2.

**Table 7** Goodness of fit of inheritance models with the observed distribution of c.1759T>G genotypes in male population (n = 43)

| Inheritance models | Observed distribution | Expected distribution | p^b |
|--------------------|-----------------------|-----------------------|-----|
| Autosomal locus | 14.9 | 28.1 | 2.83 \times 10^{-4} |
| Lethal GG homozygote | 0 | 43 | 0.99 |
| Male determination by G^0 | 0 | 43 | 0.99 |

| p | q | G |
|---|---|---|
| \frac{1}{2} | \frac{1}{2} | \frac{1}{2} |
| \frac{1}{2} | \frac{1}{2} | \frac{1}{2} |

| p | q | G |
|---|---|---|
| \frac{1}{2} | \frac{1}{2} | \frac{1}{2} |
| \frac{1}{2} | \frac{1}{2} | \frac{1}{2} |

| p | q | G |
|---|---|---|
| \frac{1}{2} | \frac{1}{2} | \frac{1}{2} |
| \frac{1}{2} | \frac{1}{2} | \frac{1}{2} |

| p | q | G |
|---|---|---|
| \frac{1}{2} | \frac{1}{2} | \frac{1}{2} |
| \frac{1}{2} | \frac{1}{2} | \frac{1}{2} |

| p | q | G |
|---|---|---|
| \frac{1}{2} | \frac{1}{2} | \frac{1}{2} |
| \frac{1}{2} | \frac{1}{2} | \frac{1}{2} |

**Table 7** Goodness of fit of inheritance models with the observed distribution of c.1759T>G genotypes in male population (n = 43)
the causative nature of the “Y” variant. The latter is a more likely explanation in view of the presumed XX/XY SD system, and its strong association with sex. Both inheritance models testing a segregating autosomal locus or lethal homozygosity of the GG genotypes were excluded (Table 7).

Many SD genes in fish, birds, reptiles, and crocodiles, belong to a gene pathway that drives the balance of sex steroids (Capel 2017). The assumed finding of fshr as an MKR of SD in M. cephalus, together with the knowledge of its involvement in gonad differentiation from other studies, support a general putative role of SD genes through 17\(^\beta\)-estradiol (E2) and 11-ketotestosterone (11-KT) equilibrium in fish (Baroiller et al. 1999). The manner in which fshr governs SD could be attributed to its recorded role in SD in amphibians and fish. In mammals, FSH signaling in granulosa cells is known to stimulate the cytochrome P450 aromatase (CYP19A1) expression, affecting estrogen biosynthesis and consequently the conversion of androgens to estrogens (Whitlock 1986; Steinkampf et al. 1987; Fitzpatrick and Richards 1991; Manuel Silva and Price 2000). Different studies in fish and amphibians have shown the involvement of fshr in functions that are related to SD, presumably by controlling cyp19a1 and estrogen biosynthesis (Yamaguchi et al. 2007; Suda et al. 2011; Nakamura 2013; Murozumi et al. 2014). However, findings in Honeycomb grouper (Epinephelus merra), (Kobayashi et al. 2010) and Japanese eel (Anguilla japonica) (Ohta et al. 2007), suggest its role is in upregulating androgen (11KT) biosynthesis, perhaps through the control of P450 11b-hydroxylase (P45011b). Indeed, disruption of fshr causes masculinization in medaka and zebrafish whereas in the latter loss of fshr lead sexual reversal to fertile males (Murozumi et al. 2014, Zhang et al., 2015). Thus, FSH/FSHR’s signaling role in sex hormone biosynthesis may possibly influence male determination in M. cephalus. The SD associated haplotype on fshr is capable of encoding two adjacent protein domains: transmembrane \(\alpha\)-helices six (TM6), and intracellular loop 3 (IL3). This loop is assumed to be a G protein interaction domain (Inglese et al. 1993; Kosugi et al. 1993), which is conserved in G protein-coupled receptors (GPCRs), such as FSHR, TSHR, and LHR (Kotlar et al. 1997, Figure 4). Many studies have demonstrated that this region, including the two domains, is of crucial importance in maintaining the G protein receptors in an inactive conformation (Tao 2008), furthermore, in humans, mutations in this domain are known to cause alteration of signaling mostly by activation (Kotlar et al. 1997, Figure 4). The first nucleic acid variation (V/M) is in IL3. However, this substitution is common as it was found in the WT fshr of fugu and trout (Figure 2). In contrast, the second nucleic acid substitution (F/V i.e. phenylalanine/valine) is in a highly conserved position in TM6, and based on comparative analysis, valine is a novel amino acid in this position in vertebrate FSHRs, which is likely to induce FSH/FSHR signaling alteration (Figure 4). A mutation in the same conserved amino acid, was found in thyrotropin receptor gene, and causes cAMP basal activation without induction (Porcellini et al. 1994), which may lead to fshry control of sex via upregulation of androgen biosynthesis. Alternatively, the male fshry may be inactive, thus inducing masculinity by lowering the signal which is needed for female gonad differentiation. Indeed, we were not able to detect the expression of fshry in 28 dph larvae and adult brains and testes. However,
our predicted \textit{fshr} displayed no pseudogene characteristics, as its entire reading frame was intact, which implies that it is functional (Supplementary Figure S1). Thus, it may be postulated that \textit{fshr} is expressed in a narrow time window, not yet resolved. A more complex mechanism may be postulated based on a hypothesis that LG9 has adopted some primitive characteristics of a sex chromosome, which may involve gene silencing in parts of the chromosome being packaged into transcriptionally inactive structures (heterochromatin). Indeed, of the 107 sex-biased genes that were differentially expressed by \textit{M. cephalus} larvae, 23 displayed sex-specific inactivation and 22 of these genes were positioned to LG9 (Dor et al. 2020). Moreover, the few genotypes (4 out of 83) that did not fit with the proposed mono factorial genetic model for male determination might be attributed to environmental sex reversal (Baroiller et al. 1999). However, sex phenotyping is prone to errors and thus may be responsible for the occurrence of such observations.

Previous studies have shown that gene duplications are frequently associated with SD (Eshel et al. 2014; Curzon et al. 2020). Our data from DNA-seq assembly, together with the results from specific PCR and sequencing of the regions representing the “X” and “Y” chromosomes, provide no evidence for duplications or null alleles in \textit{M. cephalus} SD mechanism. In \textit{Oreochromis niloticus}, an \textit{Anti Mullerian hormone} (amh) male-specific duplication, \textit{amhy}, was suggested as the SD key regulator (Eshel et al. 2014, Curzon et al. 2020). Li et al. (2015) showed that an \textit{amhy} knockout resulted in male to female sex reversal in \textit{XY} tilapia, while its overexpression resulted in female to male sex reversal in \textit{XY} tilapia, while its overexpression might be attributed to environmental sex reversal (Baroiller et al. 1999). However, sex phenotyping is prone to errors and thus may be responsible for the occurrence of such observations.

**Acknowledgments**

A. S., H. R., I. M.A., M.R., and E.S. have, together with Kidum—R&D Applications and Technology Transfer, filed an Israeli patent application based on the research described within this article (#275926; dated: 08/07/2020; titled: Nucleic acid sequences for development of methods of using such sequence).

**Funding**

This work was supported by the Nizan project, the Ministry of Agriculture and Rural Development, Israel. Grant: 30-04-0010.

**Conflicts of interest:** None declared.

**Literature cited**

Aizen J, Meiri I, Tzchori I, Levavi-Sivan B, Rosenfeld H. 2005. Enhancing spawning in the grey mullet (\textit{Mugil cephalus}) by removal of dopaminergic inhibition. Gen Comp Endocrinol. 142: 212–221. doi:10.1016/j.ygeneco.2005.01.002

Altshuler D, Daly MJ, Lander ES. 2008. Genetic mapping in human disease. Science. 322: 881–888. doi:10.1126/science.1156409

Bao L, Tian C, Liu S, Zhang Y, Elaswad A, et al. 2019. The \textit{Y} chromosome sequence of the channel catfish suggests novel sex determination mechanisms in teleost fish. BMC Biol. 17:16. doi:10.1186/s12915-019-0627-7

Baroiller JF, Guiguen Y, Fostier A. 1999. Endocrine and environmental aspects of sex differentiation in fish. Cell Mol Life Sci. 55: 910–931. doi:10.1007/s000180050344

Bej DK, Miyoshi K, Hattori RS, Strüssmann CA, Yamamoto Y. 2019. A duplicated, truncated amh gene is involved in male sex determination in an old world silverside. G3 (Bethesda). 7:2489–2495. doi:10.1534/g3.117.042697

Capel B. 2017. Vertebrate sex determination: Evolutionary plasticity of a fundamental switch. Nat Rev Genet. 18:675–689. doi:10.1038/nrrg.2017.60

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

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 Aquacult Soc. 49:277–291. doi:10.1111/jwas.12479

Cui Z, Liu Y, Wang W, Wang Q, Zhang N, et al. 2017. Genome editing reveals \textit{dmt1} as an essential male sex-determining gene in Chinese tongue sole (\textit{Cynoglossus semilaevis}). Sci Rep. 7:1–10. doi:10.1038/srep42213

Curzon AY, Shirak A, Dor L, Zak T, Perelberg A, et al. 2020. A conserved truncated male-specific copy of \textit{amh} is associated with sex determination of Nile tilapia (\textit{Oreochromis niloticus}). Heredity. 124:317–327. doi:10.1038/s41437-020-0340-x

Dor L, Shirak A, Rosenfeld H, Ashkenazi IM, Band MR, et al. 2016. Identification of the sex-determining region in flathead grey mullet (\textit{Mugil cephalus}). Anim Genet. 47:698–707. doi:10.1111/age.12486

Dor L, Shirak A, Curzon AY, Rosenfeld H, Ashkenazi IM, et al. 2020. Preferential mapping of sex-biased differentially-expressed genes of larvae to the sex-determining region of flathead grey mullet (\textit{Mugil cephalus}). Front Genet. 11:839. doi:10.3389/fgene.2020.00839

Eshel A, Shirak A, Dor L, Band M, Zak T, et al. 2014. Identification of male-specific \textit{amh} duplication, sexually differentially expressed genes and microRNAs at early embryonic development of Nile tilapia (\textit{Oreochromis niloticus}). BMC Genomics. 15:774. doi:10.1186/1471-2164-15-774

Feron R, Zahm M, Cabau C, Klopp C, Roques C, et al. 2020. Characterization of a \textit{Y}-specific duplication/insertion of the anti-Mullerian hormone type II receptor gene based on a chromosome-scale genome assembly of yellow perch, \textit{Perca fluviatilis}. Mol Ecol Resour. 20:531–543. doi:10.1111/1755-0998.13133

Fitzpatrick SL, Richards JS. 1991. Regulation of cytochrome P450 aromatase messenger ribonucleic acid and activity by steroids and gonadotropins in rat granulosa cells. Endocrinology. 129: 1452–1462. doi:10.1210/endo-129-3-1452

Graves JM. 2013. How to evolve new vertebrate sex determining genes. Dev Dyn. 242:354–359. doi:10.1002/dvdy.23887

Gubbay J, Collignon J, Koopman P, Capel B, Economou A, et al. 1990. A gene-mapping to the sex-determining region of the mouse \textit{y}-chromosome is a member of a novel family of embryonically expressed genes. Nature. 346:245–250. doi:10.1038/346245a0

Hattori RS, Murai Y, Oura M, Masuda S, Majhi SK, et al. 2012. A \textit{Y}-linked anti-Mullerian hormone duplication takes over a critical role in sex determination. Proc Natl Acad Sci USA. 109:2955–2959. doi:10.1073/pnas.1018392109

Herpin A, Schartl M. 2015. Plasticity of gene regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers, and usurpators. EMBO Rep. 16:1260–1274. doi:10.1525/embr.201540667
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: doi:10.1186/s13104-016-2030-6

Schartl M. 2004. Sex chromosome evolution in non-mammalian vertebrates. Curr Opin Genet Dev. 14:634–641. doi:10.1016/J.GDE.2004.09.005

Ser JR, Roberts RB, Kocher TD. 2010. Multiple interacting loci control sex determination in lake malawi cichlid fish. Evolution (N. Y). 64:486–501. doi:10.1111/j.1558-5646.2009.00871.x

Shirak A, Grabherr M, Palma FD, Lindblad-Toh K, Hulata G, Ser JR, Roberts RB, Kocher TD. 2010. Identification of repetitive elements in the genome of Oreochromis niloticus: Tilapia repeat masker. Mar Biotechnol. 12:121–125. doi:10.1007/s10126-009-9236-8

Shirak A, Seroussi E, Cnaani A, Howe AE, Domokhovsky R, et al. 2006. Amh and Dmrt1 genes map to tilapia (Oreochromis spp.) linkage group 23 within quantitative trait locus regions for sex determination. Genetics. 174:1573–1581. doi:10.1534/genetics.106.059030

Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths BL, et al. 2006. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature. 346:240–244. doi:10.1038/346240a0

Smith CA, Roeszler KN, Ohnesorg T, Cummins DM, Farlie PG, et al. 2004. Sex chromosome evolution in non-mammalian vertebrates. Curr Opin Genet Dev. 14:634–641. doi:10.1016/J.GDE.2004.09.005

Suda M, Kodama M, Oshima Y, Yamamoto K, Nakamura Y, et al. 2011. Up-regulation of FSHR expression during gonadal sex determination in the frog Rana rugosa. Gen Comp Endocrinol. 172:465–471. doi:10.1016/j.ygcen.2011.04.011

Takehana Y, Matsuda M, Myoshio T, Suster ML, Kawakami K, et al. 2014. Co-option of Sox3 as the male-determining factor on the y chromosome in the fish Oryzias dancena. Nat Commun. 5:1–10. doi:10.1038/ncomms5157

Tao Y-X. 2008. Constitutive activation of G protein-coupled receptors and diseases: insights into mechanisms of activation and therapeutics. Pharmacol Ther. 120:129–148. doi:10.1016/J.PHARM THERA.2008.07.005

Tilmann C, Capel B. 2002. Cellular and molecular pathways regulating mammalian sex determination. Recent Prog Horm Res. 57:1–18. doi:10.1012/tp.57.1.1

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, et al. 2012. Primer3—new capabilities and interfaces. Nucleic Acids Research. 40:e115–e115.

Ventura T, Sagi A. 2012. The insulin-like androgenic gland hormone in crustaceans: From a single gene silencing to a wide array of sexual manipulation-based biotechnologies. Biotechnol Adv. 30:1543–1550. doi:10.1016/j.biotechadv.2012.04.008

Whitlock JP. 1986. The regulation of cytochrome P-450 gene expression. Annu Rev Pharmacol Toxicol. 26:333–369. doi:10.1146/annurev.pa.26.040186.002001

Williams TM, Carroll SB. 2009. Genetic and molecular insights into the development and evolution of sexual dimorphism. Nat Rev Genet. 10:797–804. doi:10.1038/nrg2687

Yamaguchi T, Yamaguchi S, Hirai T, Kitano T. 2007. Follicle-stimulating hormone receptor (fshr) gene by TALEN leads to failed follicle activation in females followed by sexual reversal to males. Endocrinology. 156:3747–3762. doi:10.1210/en.2015-1039

Communicating editor: R. Houston