



Quantitative trait loci on LGs 9 and 14 affect the reproductive interaction between two *Oreochromis* species, *O. niloticus* and *O. aureus*

Andrey Shirak¹ · Tatyana Zak² · Lior Dor^{1,3} · Ayana Benet-Perlberg² · Joel I. Weller¹ · Micha Ron¹ · Eyal Seroussi¹

Received: 15 March 2018 / Revised: 11 July 2018 / Accepted: 17 July 2018 / Published online: 6 August 2018
© The Genetics Society 2018

Abstract

Effective farming of tilapia requires all-male culture, characterized by uniformity and high growth rate. Males of *O. aureus* (*Oa*) and females of *O. niloticus* (*On*) produce all-male offspring, but there is a behavioral reproductive barrier between the two species that prevents mass production. In crosses between *Oa* and *On* broodstocks, few hybrid females are attracted to the *Oa* male nests (denoted responders), and if they harbor the *On* alleles for the sex determination (SD) sites on linkage groups (LGs) 1, 3, and 23, all-male progeny are produced. Yet, without controlling for the alleles underlying SD, the parental stocks gradually lose their capability for all-male production. Hypothesizing that marker-assisted selection for female responders would allow production of sustainable broodstocks, we applied genotyping-by-sequencing to generate 4983 informative SNPs from 13 responding and 28 non-responding females from two full-sib families. Accounting for multiple comparisons in a genome-wide association study, seven SNPs met a false discovery rate of 0.061. Lowest nominal probabilities were on LGs 9 and 14, for which microsatellite DNA markers were designed within the candidate genes *PTGDSL* and *CASRL*, respectively. By increasing the sample size to 22 responders and 47 non-responders and by genotyping additional established microsatellites, we confirmed the association of these LGs with female responsiveness. The combined effects of microsatellites *GMI71* and *CARSL-LOC100690618* on LGs 9 and 14 explained 37% of the phenotypic variance of reproductive interaction ($p < 0.0001$). Based on these findings, we propose a strategy for mass production of all-male tilapia hybrids through selection for genomic loci affecting SD and female responsiveness.

Introduction

Sexual interactions between species, also termed reproductive interference, were first documented by behavioral biologists in 1929 and has been observed in a wide range of

animal taxa such as arthropods, insects, and vertebrates (Groning and Hochkirch 2008). Although cichlid fish species differ in traits relevant for mating preferences, courtship behavior, timing, and location, they can be readily hybridized in the laboratory to produce viable offspring (Stelkens et al. 2010). This makes cichlid hybrids an ideal model system for understanding the genetics of adaptation and speciation for these traits as well as reproductive interference (Henning and Meyer 2014). Nevertheless, there are no published reports of causative genes that affect behavioral sexual incompatibility between cichlids.

In tilapia, courtship behavior is known to involve chemical, visual, and auditory sensations (Keller-Costa et al. 2014; Longrie et al. 2013; Simoes et al. 2015). Of the multiple genomic loci that mediate these sensory pathways, genes that mediate olfactory stimuli conveyed by the urine may have a major impact (Keller-Costa et al. 2014; Simoes et al. 2015). Based on blood concentration profiles, expression pattern, level of glycosylation, and protein

Electronic supplementary material The online version of this article (<https://doi.org/10.1038/s41437-018-0131-9>) contains supplementary material, which is available to authorized users.

✉ Eyal Seroussi
seroussi@agri.huji.ac.il

¹ Institute of Animal Science, Agricultural Research Organization (ARO), Rishon LeTsiyon, Israel

² Dor Research Station, Division of Fishery and Aquaculture, Hof HaCarmel, Israel

³ Robert H. Smith Faculty of Agriculture, Food and Environment, Hebrew University of Jerusalem, Rehovot, Israel

structures, two families of major urinary protein (*MUP*) orthologs in tilapia, male-specific protein (*MSP*) and tributyltin-binding protein (*TBTBP*) gene families, were proposed to have a role as reproductive pheromones (Shirak et al. 2008; Shirak et al. 2012). Pheromones are mainly detected by their binding to neuronal G protein-coupled receptors, and little is known about the function of these receptors in cichlids (Keller-Costa et al. 2015). In cichlids, chemosensory receptors may belong to three gene families: trace amine-associated receptors (TAARs) and vomeronasal type 1 and 2 receptors (V1R and V2R). As fish lack a vomeronasal organ, this nomenclature follows their orthology to similar gene families in tetrapods (Keller-Costa et al. 2015). Metabolites such as F prostaglandins, C18, C19, and C21 sex steroid derivatives, and amino and bile acids have pheromonal function in phylogenetically distant fish species (Sorensen and Baker 2015). Yet, individuals of closely related species can distinguish between conspecific and allospecific partners. It was hypothesized that, while the set of pheromones is common among fish, their mixtures, behavioral context, timing, and location convey the species specificity (Stacey 2015). In goldfish and common carp, initial communal signals promoting female ovulation are not species specific. However, advanced stages of courtship result in interspecific barriers mostly by visual and auditory sensations (Kobayashi et al. 2002). Pheromone transporters such as lipocalins, which have a species-specific structure and expression, may also form interspecific barriers (Shirak et al. 2012).

Four tilapia species: Nile tilapia (*Oreochromis niloticus*, *On*), blue tilapia (*Oreochromis aureus*, *Oa*), Mozambique tilapia (*Oreochromis mossambicus*), and Wami tilapia (*Oreochromis urolepis hornorum*), and their hybrids are agriculturally important. Tilapia males grow faster than females and better perform when cultured in the absence of females, thus eliminating territorial and reproductive behavior. Most of the intensive production of tilapia relies on inducing mono-sex cultures of males through interspecific crossbreeding of *On* females and *Oa* males followed by post-larval masculinization using steroid hormones supplemented to the feed. This steroid treatment is essential to reduce residual production of females, which appears to result from contamination of both purebred stocks by hybrids (Beardmore et al. 2001; Rothbard and Pruginin 1975). The use of hormones has been criticized by consumer advocates and has even been banned in Europe. Moreover, steroid treatment still results in the appearance of 2–8% females (Desprez et al. 1995; Wohlfarth and Wedekind 1991).

In Israel, starting from 1960, crosses between *On* females and *Oa* males are practiced as the best strategy for achieving all-male population with uniform and high growth rate, cold tolerance, and most important, preventing uncontrolled

reproduction (Wohlfarth et al. 1994). Since the end of the 1970s, *On* stock in Bar-Ilan University originating from Ghana was the reference for a purebred stock (Galman et al. 1987; Timan et al. 2002). Artificial fertilization of females from this stock always produced all-male progeny with wild-type *Oa* males from the Dead Sea region (Trombka and Avtalion 1993). Although these two stocks were important models for immunologic and genetic studies, they were unsuitable for aquaculture, due to the reproductive barrier between them. Tilapia hybridization in Eastern Africa between local species was reported since the 1960s (Elder and Garrod 1961; Pruginin et al. 1975). This is probably due to contamination of *On* natural populations that were used to establish cultivated and laboratory broodstocks (Timan et al. 2002). Several studies have been focused on finding the optimal conditions for their crossbreeding in tanks. However, these studies resulted in fry numbers that were insufficient for commercial aquaculture (Mires 1980, 1995). Thus the problem of inefficient mass production of all-male tilapia by crossing purebred species remained unsolved.

In commercial ponds, males of *Oa* built normal nests along the pond perimeter but failed to attract purebred *On* females that remain in the center of the pond (Lovshin 1980; Mires 1980). In aquaria, the male cleans a place for spawning and tries to attract the *On* females, but the females try to avoid the males, frequently resulting in injuries and even in death, due to the male aggression (Mires 1995). In aquaria, *Oa* males may attract F2 hybrid females. However, when the courtship dance is too prolonged the male removes the non-responding females from the nest despite their ripeness. In natural conditions, a number of studies reported that, by experimental testing and through a specific scheme of crosses between *Oa* males and *On* females, it is possible to detect “responders”, which are hybrid females that interact with purebred *Oa* males bearing all or near all-male progeny (Lahav and Lahav 1994). However, over several generations, the parental stocks either lose their capability of interspecies communication or their ability to produce all male spawn. Hence, understanding the genetic regulation of responsiveness can be the key to sustainability.

Three sex determination (SD) loci on linkage groups (LGs) 1, 3, and 23 have been detected in different *Oreochromis* species and their hybrids (Eshel et al. 2012; Shirak et al. 2006). In purebred tilapia species, SD is apparently mono-factorial (Hammerman and Avtalion 1979; Wohlfarth and Wedekind 1991). In *O. mossambicus*, an XX/XY SD system was detected on LG1 (Liu et al. 2013). In *Oa* and *O. urolepis hornorum*, the SD system is WZ/ZZ encoded by a locus on LG3 (Ezaz et al. 2004; Zhu et al. 2016). In *On* of the Ghana strain, the SD factor resides on LG23, resulting in an XX/XY system (Eshel et al. 2012; Shirak et al. 2006).

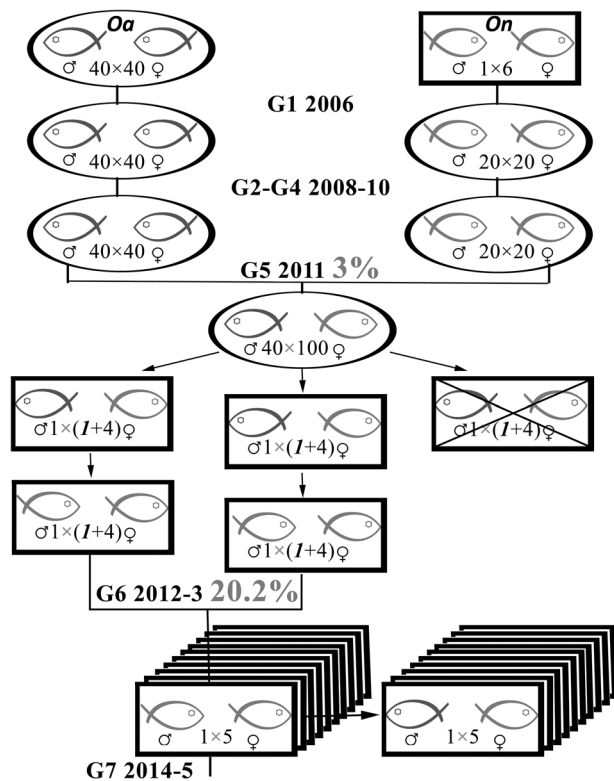


Fig. 1 Description of seven generations of natural matings and artificial insemination in experimental ponds and tanks and detection of female responders. Ponds are represented by ovals, and tanks by rectangles. The *Oa* species is denoted in blue and the *On* species in red. Purebred generations are denoted G1–G7. Lines between ovals (or rectangles) indicate parenthood and arrows denote physical relocation and re-mating of the same fish. Artificial inseminations or natural matings are indicated by black or green crosses, respectively. In generation G5, one pond and five tanks were used for mating progeny of G4 to produce the G6 generation and G6 hybrids. The numbers of known female responders in each tank are denoted by bold and italicized numerals. The unsuccessful tank mating is denoted by a rectangle covered by an “X”. A total of 46 tanks were used to mate 175 G6 female progeny to G6 males of both species: 23 tanks for *On* × *Oa* matings, and 23 for *On* × *On* matings, with 5 G6 females in each tank. The same 175 G6 *Oa* females were mated to both *Oa* and *On* males. The rate of detection of female responders is indicated in percentages (large red font) to the right of the year and generation annotation

In *On* broodstocks originating from Eastern Africa, several SD loci were observed, indicating multi-factorial SD (Cnaani et al. 2008; Lee et al. 2003; Palaiokostas et al. 2013) probably due to original stock contamination. To date, only the master key regulator, *AMH*, on LG23 has been identified (Eshel et al. 2014; Li et al. 2015). *AMH* is present in two and single copies on the Y and X chromosomes, respectively. On the Y chromosome, one copy is similar to that of the X chromosome and the other has a 5 bp insertion in exon 6 and 233 bp deletion in exon 7. A supernumerary copy of *AMH* is present in most of the *On* stocks that originated from Eastern and Western Africa (Eshel et al. 2014; Li et al. 2015).

Our preliminary study demonstrated that selecting hybrid females with *On* alleles, for the SD loci, restored the ability to produce all-male progeny when mated to *Oa* males (https://www.aquacultureinIsrael.com/he/component/k2/item/download/24_1c3cd4678c35e56910b6e448896d3c83). However, this resulted in a decrease of progeny yield of up to 90%. The objective of this study was to detect quantitative trait locus (QTL) for interspecific female responsiveness. Via marker-assisted selection, it should then be possible to increase progeny yield.

Materials and methods

Fish

On broodstock (*OnDOR2006*) of Dor Research Station was produced in 2006 by artificial egg fertilization of six *On* females (a strain originating from Ghana, Bar-Ilan University) by the milt of a single *On* male (a strain originating from Egypt). *Oa* Dor stock (*OaDOR2006*) originated from the Eynot Tzukim nature reserve - Ein Feshkha, Dead Sea region, Israel (Zak et al. 2014). All experimental protocols were approved by the Institutional Animal Care and Use Committee at Animal Experimentation Ethics Committee of the Agricultural Research Organization, Volcani Center, Rishon LeTsiyon, Israel.

Stock control by microsatellite markers

From the original 2006 broodstocks, current Dor broodstocks were established in years 2008 (G1, *OaDOR2008*; *OnDOR2008*) and 2011 (G5, *OaDOR2011*; *OnDOR2011*) (Fig. 1). Both stocks were genotyped for a panel of nine microsatellite markers (*BYL018*, *UNH168*, *TA13*, *UNH310*, *UNH362*, *GM472*, *UNH890*, *UNH954*, and *MIC176*; Table 1). Three markers *BYL018*, *UNH168*, and *MIC176* were positioned in the three SD loci on LG 1, 3, and 23, respectively (Shirak et al. 2006). Genotyping of markers for SD was performed in each generation (G2–G5) through the period 2008–2011 and individuals that carried alien alleles were removed. As all *On* microsatellite alleles were different in length from the *Oa* alleles, alleles observed in generations G2–G5 were considered alien for a broodstock if they were not observed in the G1 fish of this broodstock. Markers with the prefixes UNH and GM have been previously published and used for the construction of the tilapia second-generation linkage map (Lee et al. 2005); *BYL018* marker on LG1 was developed by Dr. Bo-Yung Lee in Professor T.D. Kocher's laboratory. *MIC176* on LG23 and *TA13* on LG3 were developed in our laboratory.

Table 1 Primer sequences used for microsatellite genotyping and variant sequencing

	Position		Forward 5'→3'	Reverse 5'→3'
<i>BYL018</i>	LG1	23,992,027	CAGGGCTTATTTGCACCACT	TGGGGAACCTTTCTCTGCATC
<i>UNH168</i>	LG3b	17,003,362	TAAGAAGGTTAGAAAAGAAAGTG	TATATAATAATTTCTAAACGGC
<i>TA13</i>	LG3b	16,887,105	ACCATCCAGCTGAACACACA	TCTTCCAGCGCAAATACAGA
<i>UNH310</i>	LG13	3,420,589	TATTGGAACACCCTCCCTCC	AGGTGTGCATGTACAAATGGTT
<i>UNH362</i>	LG17	24,602,691	GAACAGCTTTCAGACGGAGG	ACTGAGGCCAGGTGAAGAAA
<i>GM472</i>	LG10	5,419,889	CTAAATCTCCACGCAGTCC	TGTAATTGCTCCACAAATCTG
<i>UNH890</i>	LG12	22,385,222	AAAAGCATCCCTGCTTCTCA	TGGCTGTTGCAAGACACATT
<i>UNH954</i>	LG13	26,894,065	GGAAAACGTTTGGAGAGACG	AAACGGAGCTCCTGTCTGAA
<i>MIC176</i>	LG23	11,178,668	GCTGCATGTGGCCTGCGATGTA	GCACGGACCTTTTGGGAGGCTT
<i>GM171</i>	LG9	21,439,990	GAGTTAGGGCAGAGGTTATGT	GCCATCACCTCTTGCTT
<i>GM613</i>	LG9	22,156,320	TAGCACAGTGCCTGGGTGA	GTGTCACCTCACGCCTTCCTT
<i>PTGDS</i>	LG9	20,152,669	CCGAAAGTGCCTCTTACCAC	TGAGACCAGACGAGCACAGA
<i>GM103</i>	LG14	15,084,734	CCACCAATATAGAGTTTAGCC	CCCAGAGCTTTCATTACC
<i>GM237</i>	LG14	17,011,287	TTAATCCCATCTTAATCACAA	CTGAGCTAAGAGCAGGTG
<i>CASRL</i>	LG14	15,872,980	CACGCAATATAGTGTAGAAGAAAACAA	CCAAGTGGGCAATGAAGAGA
<i>CASRLex2</i>	LG14	15,862,198	GGAGCTTTTGCCCAATTACA	GGCTGGAGGATTCTTGACAC

Genotyping of microsatellite markers

Fin samples or whole larvae were immersed in absolute alcohol and stored until extraction. DNA was extracted using the MasterPure™ DNA Purification Kit (Epicentre®; Biotechnologies). PCR was applied using Super-Therm Taq DNA polymerase (JMR Holding Inc., London, UK), dye-labeled (FAM, PET, VIC, or NED) forward primers (Table 1) and 20 ng of genomic DNA. The amplified products were separated on an ABI3130 DNA sequencer and automatically sized using the GeneMapper software v. 4.0 (Applied Biosystems) using GeneScan-500 LIZ size standard (Applied Biosystems) (Dor et al. 2016).

Reproduction in experimental ponds and tanks

From 2006 to 2011, five generations of purebred mating were produced by artificial insemination and rearing in ponds (Fig. 1). To reduce contamination from uncontrolled crosses in the facilities, samples of the *OaDOR2008* ($n = 50$), *OnDOR2008* ($n = 50$), *OnDOR2011* ($N = 100$ females), and *OaDOR2011* ($n = 40$ males) were genotyped; the unique alleles in each stock were identified and the numbers of their occurrences were determined (Table S1). Broodstocks (G1–G4) were routinely selected against alien alleles in the SD loci (*BYL018*, *UNH168*, and *MIC176*) by genotyping a sample of 80–90 fish for each gender. This resulted in culling 3–16% of the sampled fish that carried at least one alien allele.

To test for the rate of interspecific female responsiveness in the fifth generation (Fig. 1, G5), in 2011, 100 *OnDOR2011* females and 40 *OaDOR2011* males were

stocked at the age of 1.5 years in an experimental earth pond of $16 \times 25 \text{ m}^2$ and 1.2 m depth. The pond's water temperature varied between 26 and 29 °C. Prior to transfer, all individuals were tagged by individual PIT-tags and their fins were clipped for DNA extraction. Tilapia larvae and fingerlings were collected from the drained pond 40 days after stocking and were reared in tanks (28–29 °C) until they reached 2 g weight (about 2 months). Gender was determined by microscopic examination of the gonads using the squash technique (Guerrero and Shelton 1974). The first groups of larvae were observed near the pool shore 12 days after stocking. After a period of 40 days, we reduced to half the pond water level and collected ~6000 offspring. We observed two age groups: fingerlings born just after parent stocking, and larvae born at least 20 days later. A comparison of the genotypes between fingerlings ($n = 32$) and their putative parents revealed that they belonged to three pairs of parents (D51 × S4, D79 × S4, and D76 × S29), whereas larvae ($n = 32$) belonged to the same dams and a single sire (D51 × S4, D79 × S4, and D76 × S4). At the age of 4 months, sexing of a sample of hybrids from both groups ($n = 2400$) indicated that all were males. To determine the involvement of reproductive hierarchy in breeding performance of both genders, three experiments were performed following removal of: (1) reproducing males (38 males were retained); (2) reproducing females (97 females were retained); and (3) all reproducing individuals from the pond. After removing only the two reproducing males (S4 and S29) from the pool for 40 days, ~4,000 offspring belonging to two aging groups were collected and genotyped ($n = 32$, each). These fingerlings belonged to two families (D79 × S20 and D76 × S20); larvae were offspring

of three pairs of parents (D51 × S20, D79 × S20, and D76 × S20). At the age of 4 months, sexing of a sample of hybrids from both groups ($n = 1700$) indicated that all were males. When the three dams that responded to *Oa* males (D51, D76, and D79) were removed and the reproducing sires (S4 and S29) were re-introduced, or removed, no offspring were detected through a period of the next 40 days.

Intraspecific or interspecific reproductive activities were examined in tanks of 1 m³ at temperatures of 28–30 °C with five water exchanges a day. These tanks were used for three different purposes: first, an *On* female responder and four *On* non-responding females were introduced into each tank with an *Oa* male to re-examine the findings in the experimental pond; second, this male was replaced with an *On* male to confirm that the non-responding females can successfully reproduce with conspecific males; third, 5–6 daughters from the latter experiment (*On* × *On*) were introduced into each tank with an *Oa* male to prepare segregating families for genome-wide association study (GWAS). Every 7–10 days females were examined for the presence of eggs or larvae in their mouth cavity. Collected eggs were incubated in Zuger bottles (28–29 °C) to confirm their fertilization. To test sex ratio, fry of interspecific spawns were transferred into rearing tanks (28–29 °C) and gender was determined at the age of 4 months by microscopic examination of the squashed gonads after fry dissection.

Selecting a broodstock for female responders

To select for the sixth generation (Fig. 1, G6), progeny of the D51 and D76 dams from the intraspecific crosses were reared and sexed. A total of 175 daughters were distributed, 5 individuals to a tank, with an *On* male for a period of 3 weeks (Fig. 1, G6). Sixty (34%) of them reproduced. These females were removed and the rest were again tested with *On* males. After repeating this test for the third time, a total of 116 (66%) females successfully reproduced. These were assigned randomly to groups of 5–6 females and re-examined in 23 tanks each with an *Oa* male. Six females having several non-fertilized eggs in their mouth cavity and an injured female were removed from the experiment.

Genotyping-by-sequencing (GBS)

DNA samples (of 500 ng each) of 22 and 47 responding and non-responding females, respectively, from two full-sib families and their parents were submitted to the sequencing facility of University of Wisconsin Biotechnology Center. All samples were digested by *ApeKI* restriction enzyme. A set of barcoded adapters with an *ApeKI* overhang and a common Y-tail adaptor were designed according to Poland

et al. (2012). After the *ApeKI* digestion and the adaptors ligation, all samples were pooled and were PCR-amplified in a single tube, producing a library of the 73 samples, which was sequenced on a single lane of Illumina HiSeq2500 flow cell.

Analysis of single-nucleotide polymorphism (SNP) data

Using the barcode adapters, sequences of each individual were grouped into separate files. Sequence reads were mapped to the repeat-masked version of the *On* genome (GenBank accession No. ASM185804v2) using the Fast Alignment Search Tool suite (mrsFAST-ultra-3.2.00 (Hach et al. 2014)). The obtained BAM-formatted files were analyzed using the GenomeAnalysisTK module of Genome Analysis Toolkit (GATK), (McKenna et al. 2010) for variant discovery and genotyping, following the recommended pipeline of the Broad institute (<https://software.broadinstitute.org/gatk/best-practices/>), which includes realignment around indels, base recalibration, and variant discovery using the haplotype caller function.

Filtering and validation of SNP data

Initially, only SNPs that were informative in both families were considered. Hence, we further analyzed only two types of crosses (AB × AA; AB × AB). SNPs were deleted if Mendelian inheritance was not observed. This occurred in two cases: (i) obtaining BB progeny from a cross of AB × AA parents; (ii) not obtaining at least two homozygous progeny of each type (AA and BB) from an AB × AB parental cross. Of the 69 G6 progeny analyzed, 28 were deleted owing to low SNP call rate, leaving 13 responders and 28 non-responders. As an initial cutoff for the minimal read coverage, we set the number of 154 reads, which is required for representation of at least 5 reads for 75% of the 41 progeny. The 29,132 candidate SNPs that passed these criteria were further examined for balanced allele frequencies with minimal allele frequency (MAF = 0.15) in both families. Of the 23,255 SNPs with this MAF, 18,648 SNP candidates that showed no >75% heterozygotes were further analyzed, under the assumption that SNP with >75% heterozygotes probably represent sites with multiple genome copies. For assigning individual heterozygous genotypes, we required that the frequencies of reads of the two SNP alleles would have a ratio ranging between 1/5 and 4/5. SNPs with valid genotypes for <75% of the individuals were deleted, leaving 4983 valid SNPs. The obtained sequence reads were mapped onto the *On* genome and genotypes were called taking into account realignment around indels, base recalibration, and haplotype variation.

Deep sequencing of *Oa* female and male genomes

The *On* reference genome was used to find variations between the tilapia species that may underlay the differences in reproductive behavior. To allow this comparison, *Oa* single female and single male DNA were deep-sequenced using the Illumina (San Diego, CA, USA) HiSeq2000 platform according to the manufacturer's pair-end protocol (ENA accession no. PRJEB23203). Average fragment length was 580 bp, and 100-bp sequence reads were obtained from both ends. Each DNA sample was applied to one lane yielding ~32- and ~26-fold coverage for the male and female *Oa* genomes, respectively.

Analysis of candidate genes

Candidate textual search in NCBI Gene database was performed using the following text query: (Oreochromis niloticus[Organism]) AND LG9[Chromosome] (sex OR pheromone OR olfactory OR taste OR prostaglandin OR lipocalin). Expression of candidate genes was evaluated using NCBI's Sequence Read Archive (SRA), Nucleotide BLAST tool, and RNA-seq SRA data of *On* (GenBank accession No. PRJNA78915). For identifying variation between *Oa* and *On* within the coding sequence of candidate genes, the *On* reference gene sequences were used as templates for mapping DNA-seq reads obtained for *Oa* using the GAP5 software (Bonfield and Whitwham 2010). BWA options for this mapping were set to bam bwsw -t 8 -T 60 (Li and Durbin 2010). Further analysis of variation was performed with Sanger sequencing: DNA was amplified using PCR primers (Table 1) and the Bio-X-ACT™ Long Kit (Bioline Ltd., London, UK) according to the manufacturer's instructions under the following conditions: 30 cycles for 40 s at 92 °C, 60 s at 63 °C, and 60 s at 68 °C. The PCR products were separated on agarose gels, excised, and purified with the AccuPrep® Gel Purification Kit (Bio-Near Corp., Seoul, Korea). Chromatograms were obtained by ABI3730 sequencing using a BigDye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). Detection and characterization of indels was performed using ShiftDetector and the ABI tracefiles (Seroussi et al. 2002). Copy number proportions were estimated based on the ABI chromatograms peak-height ratios (Seroussi et al. 2013; Shirak et al. 2017).

Statistics

The 4983 SNP that passed the edits were tested for their effects on mating response of females from the two families that were selected as broodstock for female responsiveness. The effect of each SNP was determined by Student's *t* test comparing the 13 responders to the 28 non-responders and a

chi-squared test for random association of genotypes with respect to response status. For the *t* test for each SNP, homozygotes for the more frequent allele were scored as 0, heterozygotes as 1, and homozygotes for the less frequent allele as 2. Owing to the large number of comparisons, experiment-wise significance was determined first by application of the Bonferroni correction for multiple tests and also by controlling the false discovery rate (FDR) at 15% (Benjamini and Hochberg 1995). That is, the highest probability was determined with an expectation that no >15% of the selected SNPs would obtain this level of nominal significance by chance. SNPs that met this criterion were also analyzed by Fisher's exact test, which unlike the chi-squared test can be applied even if the expected cell frequencies are <5.

The effects of 9 microsatellites on response status of 69 G6 progeny were tested by the General Linear Model (GLM) procedure of SAS (Statistical Analysis System 1999). In the first stage, the effect of genotypes of each of the microsatellite markers on breeding response was tested with responders coded as "one" and non-responders coded as "zero." Since progeny were from two families, and the parental genotypes were different in each family, the effect of progeny genotype was nested within family, and the model also included a family effect. All effects were considered categorical effects. In the second stage, all two-way combinations of markers with significant effects were tested, and coefficients of determination were computed. The two-way combinations were also analyzed by the GLMIX procedure of SAS that assumes a categorical dependent variable. However, this procedure does not estimate the coefficient of determination.

Results

Association between infiltration of foreign alleles and interspecific reproductive interaction

To investigate the relationship between cross-contamination and the ability to produce all-male progeny in *On* × *Oa* crosses, we analyzed our broodstocks for the presence of foreign genetic alleles for the three SD loci of tilapia. These broodstocks were maintained by artificial insemination generating a new generation each year in most years (G1–G6, Fig. 1). To ensure purebred quality, the parents of the *Oa* stock (*OaDOR2006*) were isolated from a nature reserve in the wild. The initial *On* broodstock (G1, *OnDOR2006*) originated of laboratory broodstocks and its females produced all-male progeny when artificially inseminated with *OaDOR2006* sperm (data not shown). In the 2008 purebred stocks, each of the *Oa* and *On* stocks had specific range of allele lengths. Thus alleles of these

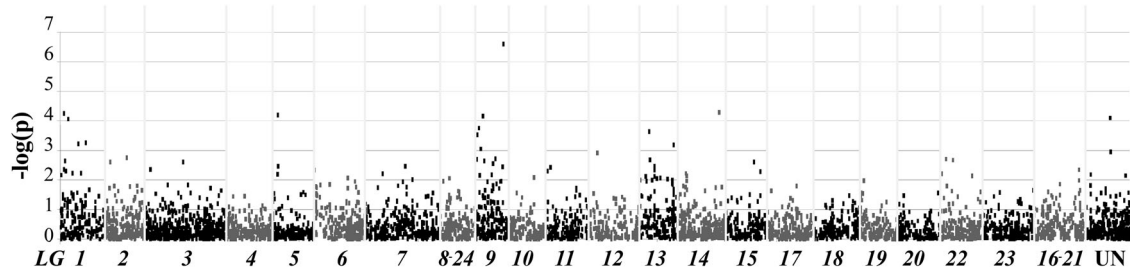


Fig. 2 Whole-genome association study of 4983 SNPs for females' reproductive interaction

Table 2 Top ten ranking SNPs associated with female responsiveness controlling for false discovery rate of 15%

SNP Position	Nominal <i>P</i> value	False discovery rate
LG9^a	22,897,740	2.49E-07
LG14	34,153,894	5.03E-05
LG1	3,099,249	5.46E-05
LG5	3,952,415	6.25E-05
LG9	5,647,143	6.78E-05
UN	156,511	7.87E-05
LG1	6,707,487	8.51E-05
LG9	2,523,212	1.74E-04
LG13	7,819,252	2.28E-04
LG9	1,032,491	2.88E-04

^aSNPs that also satisfy the criterion of 6% FDR are in bold

markers can be used to estimate the level of cross-contaminations in current broodstocks. In the period between 2009 and 2010, broodstocks were selected against alien alleles in the SD loci; and thus in the 2011 stocks, no contaminating alleles were observed for the three microsatellites residing in the SD regions. However, in 5 out of 6 other markers placed on autosomal regions that are scattered in the genome, we detected *Oa* alleles in the *OnDOR2011* stock with low frequency of 4–5.5% (Table S1, red font on bluish background). Hence, the number of alleles in *OaDOR2008* and *OaDOR2011* remained stable, whereas *OnDOR2011* had more alleles than the original *OnDOR2008* stock, suggesting infiltration of autosomal alleles that may mediate interspecific reproductive interaction with *Oa*.

In 2012, *OnDOR2011* females ($n = 100$) and *OaDOR2011* males ($n = 40$) were crossed in natural conditions in an experimental pool (G5, ×, green font, Fig. 1). Genetic analysis of the ~6000 offspring revealed that they belonged to three dams, which were denoted responders. These females were placed in tanks with other four non-responders and tested for interspecific and intraspecific interactions with males. The tests confirmed that only these dams were capable of interspecific reproduction, whereas all other females were healthy and capable of intraspecific

reproduction, thus indicating detection frequency of responders of 3% (Fig. 1). This rate is close to the rate of autosomal *Oa* alleles that infiltrated the *OnDOR2011*. Therefore, we devised an experiment that would examine the hypothesis that the responsiveness trait is genetically inherited and related to the infiltrating *Oa* alleles. Progeny of two of the dams from the intraspecific crosses (G6, Fig. 1) were reared and further tested for interspecific and intraspecific interactions with males. The detection frequency of responders in G6 was estimated to be ~20% (Fig. 1). The responding status of each remaining females was corroborated by re-examining the responders and non-responders in separated groups. In all cases, the original status indicated in the first test was confirmed. Thus 22 responders (20.2%) were detected, and 87 were determined to be non-responders. The frequency of responders in each family was similar: 12 of 57 (21.1%) and 10 of 52 (19.2%). Thus, by application of a single generation of genetic selection, the rate of responders that produce nearly all-male offspring increased from 3.0% (3/100) to 20.2% (22/109). The progeny ($n = 1080$) of the 22 responders were reared for 4 months and sexed. Unlike in the ponds for which all progeny were males, in the tanks 0.74% of the progeny were females. The elevation in the rate of detection of female responders supported the hypothesis that the responsiveness trait is genetically inherited and suggested that the G6 population may be used to genetically map the genomic regions that underlie female responsiveness.

GWAS for female responsiveness

Following the indication that the percentage of female responders increases (~7×) in a broodstock raised from female responders, we concluded that this trait is likely to have significant heritability and that a genome-wide scan of genetic variants in responders and non-responders females may point to the major genomic regions involved. This study can be facilitated with the use of SNPs genotyped by sequencing. GWAS *t* test results based on the 41 females with valid genotypes and determination of responder status are given in Fig. 2. The number of individuals with valid genotypes for each SNP varied from 31 to 41 with a mean

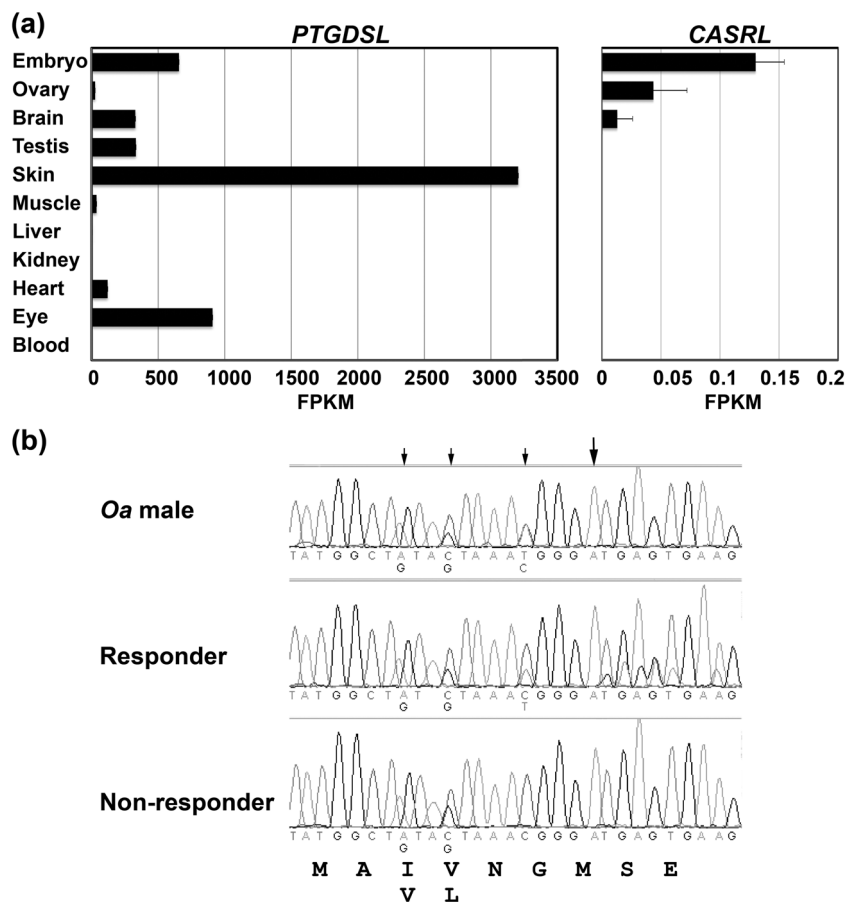


Fig. 3 Expression and sequence variation of candidate genes. **a** Transcript abundance histogram showing the expression of *PTGDSL* and *CASRL* in 11 tissues of Nile tilapia. Data of RNA sequencing of *Oreochromis niloticus* from a project in which each tissue sample was sequenced in three runs (GenBank accession No. PRJNA78915) was meta-analyzed and transcript abundance was estimated in fragments per kilobase of transcript per million (FPKM). The standard error was approximated from three runs (error bars). **b** Typical sequencing chromatograms of the second exon of *CASRL*. Genomic DNA

extracted from *Oreochromis aureus* males ($n = 4$) and responsive and non-responsive females ($n = 10$, each) was sequenced using Sanger sequencing of a PCR amplicon (Table 1) from the forward primer. Three arrows point to three polymorphic positions that encode I130V, V131L amino acid variations, and a synonymous nucleotide substitution. A larger fourth arrow indicates the first position in the codon for the methionine (ATG), which is missing in the gene variant with the shorter exon 2

of 36.5. The mean frequency of heterozygotes was only 0.35, as compared to the expected value of 0.5. This bias may be partially due to the requirement that the frequencies of reads of the two SNP alleles would have a ratio ranging between 1/5 and 4/5 for calling heterozygotes. The most significant association with females' reproductive interaction was at position 22.9 Mb on LG9. The association with this SNP was the only one that met experiment-wise significance after Bonferroni correction (experiment-wise $p = 0.001$). Six additional SNPs on LGs 1, 5, 14, and an unknown location of scaffold 320 met an FDR of 6% (Benjamini and Hochberg 1995), (Table 2). It is expected that at least five of these markers represent real effects. Chi-squared and Fisher's exact test probabilities for the SNPs with the lowest probabilities were higher (data not shown). It should be also noted that the observed distribution of SNPs with the greatest effects on LGs 9 and 14 was not

concentrated in a single peak, as these were positioned on both ends of these LGs (Fig. 2).

Selection of candidate genes on LGs 9 and 14

Seeking corroboration for the existence of major QTLs for female responsiveness, we screened for candidate genes within LGs 9 and 14, for which the most significant SNPs were located. A gene was considered a strong candidate if its orthologs were shown to mediate species-specific reproductive interaction; its expression pattern fits this role; and if there was significant functional variation between the coding sequences of the *On* and *Oa* gene, for which we determined the sequence as part of our whole-genome *Oa* sequencing project. Using key words (sex, pheromone, olfactory, taste, prostaglandin, lipocalin), a search in the *On* annotated genome build (NCBI Gene

database) revealed 12 and 74 relevant genes on LGs 9 and 14, respectively.

On LG9, we flagged a gene annotated as lipocalin (*LOC100712094*, *PTGDSL*, GenBank accession No. XP_003443553), which was highly orthologous (identity 30%, similarity 49%) to the human prostaglandin-H2 D-isomerase (*PTGDS*, GenBank accession No. NP_000945). As prostaglandins comprise the postovulatory female sex pheromone in goldfish (Sorensen et al. 1988) and their synthesis also play a key role in ovulation, with species-specific differences in the regulation or timing (Lister and Van der Kraak 2009), we further analyzed this gene expression in *On* using RNA-seq data that was deposited in Short Read Archive (GenBank accession No. PRJNA78915). This analysis indicated that *PTGDSL* is mostly expressed in the skin, which is compatible with mediating exogenous interactions (Fig. 3a). Comparison of the coding region between the putative *On* and *Oa* prostaglandin synthases revealed that the latter had a non-conservative amino acid substitution encoded by the fourth exon (M86T).

On LG14, we observed three clusters of chemosensory receptors each including a large number of candidate genes. Of special interest was the telomeric cluster stretching over a million bases (15.1–16.1 Mb on LG14) with over 90 genes and pseudo-genes encoding extracellular calcium-sensing receptor-like (CaSRLs) protein domains, one of which, *LOC100690618* (*CASRL*, GenBank accession No. XP_005464210), was specifically transcribed in the ovary (Fig. 3a). The fact that CaSR genes are expressed in ovarian surface epithelial cells, and function in oocyte proliferation, maturation, and follicle survival, supports the possible involvement of *CASRL* in the calcium signal that activate oocytes maturation. The signal for egg activation is species dependent (Ellinger 2016). Analysis of variation in *CASRL* coding sequences between *On* and *Oa* was complicated by copy number variation and indicated an amino acid deletion (Δ M134) and non-conservative amino acid substitutions encoded by the second exon (Fig. 3b). Variant Δ M134 was associated with female responders for which the peak ratio in sequencing chromatogram also suggested the presence of additional copy (e.g., ratio 1:3 in the synonymous substitution, Fig. 3b).

To further evaluate association of the selected candidate genes with female responsiveness, we designed microsatellite markers within these genes. These markers were positioned at the 5' intergenic region and in the untranslated region of the last exon of *PTGDSL* and *CASRL*, respectively (Table 1).

Confirmation of effects on LGs 9 and 14 using microsatellite markers

To confirm the SNP associations obtained in our GWAS analysis, we tested microsatellite markers on the two most

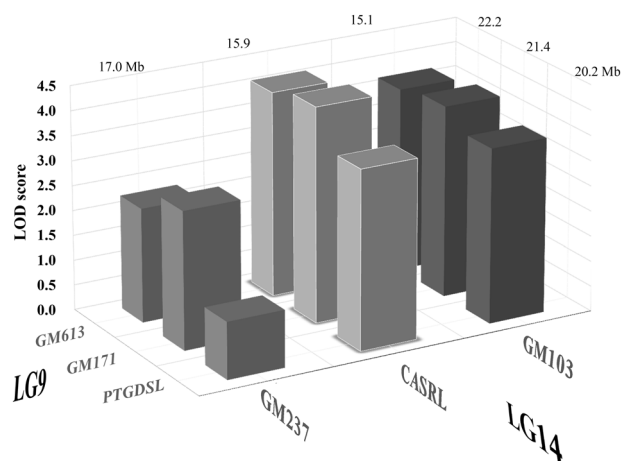


Fig. 4 Combined effects of microsatellite markers originating from LGs 9 and 14 on females' reproductive interaction. LOD scores were derived from the F -values for the complete models that include the effects of family and marker genotypes for the two markers nested within family

significant LGs. For each of LGs 9 and 14, we applied two established markers (*GM171*, *GM613* and *GM103*, *GM237*, respectively (Lee et al. 2005)) and one marker developed within an attractive gene candidate (*PTGDSL-LOC100712094* and *CASRL-LOC100690618*, respectively). All markers were significantly associated with female responsiveness ($p < 0.0001$) in the individual marker GLM analyses. Significance was also determined for the nine two-way combinations of the six markers, including one marker from each LG. The limit of detection (LOD) scores ranged from 1.9 to 6.1 (Fig. 4). LOD scores and coefficients of determination were highest for the combination of the *GM171* and *CASRL* microsatellites, supporting the hypothesis that the *CASRL* gene cluster affects this trait. LOD scores were lower for combinations with *PTGDSL*, suggesting that the related gene is a less likely candidate. The combined effects of *GM171* and *CASRL* from LGs 9 and 14, respectively, resulted in a coefficient of determination of 0.37 for female responsiveness. Although this includes the "family effect", as noted previously there was virtually no difference between response rate in the two families, and this effect was not significant in any of the models tested.

Discussion

Our study aimed at better understanding of the genetics of adaptation and speciation for traits related to reproductive incompatibility between two closely related tilapia species and their hybrids. Identification of the master genes for SD and the genes for the behavioral reproductive barrier between *Oa* and *On* will have important implications for the

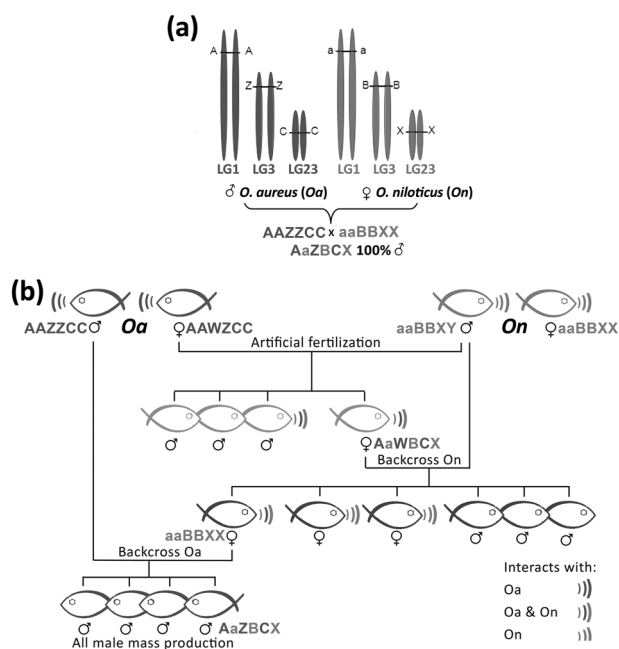


Fig. 5 A proposed strategy for mass production of all-male tilapia hybrids. **a** The allele patterns that are assumed to determine all-male progeny by an *Oa* male \times *On* female cross are represented using a delineation of the tilapia sex determination (SD) systems. These are derived from: *O. mossambicus*, which has a XX/XY system on LG1 (alleles A, A) (Hammerman and Avtalion 1979); *O. aureus* and *O. hornorum* having WZ/ZZ system on LG3 (alleles Z, B) (Ezaz et al. 2004; Zhu et al. 2016); and *O. niloticus*, which has a XX/XY system on LG23 (alleles C, X) (Shirak et al. 2006). **b** A putative crossing scheme for mass production of all-male tilapia. The allele patterns that determine SD and types of sexual interaction are represented using a delineation of two crossing steps for acquiring *On* female with the original SD systems (aaBBXX) and restored capability of sexual interaction. The final crossing step between this female (aaBBXX) and *Oa* male (AAZZCC) is expected to derive all-male progeny. Lettering of SD alleles follows that of **a**. Types of sexual interaction of individual females are symbolized by arches colored according to their sexual-mediated interaction with either *Oa* (blue) or *On* (red)

production of monosex culture required for effective aquaculture (Beardmore et al. 2001). In tilapia, it has been previously demonstrated that LGs 1, 3, and 23 play a major role in SD and that *Oa* male has a super male genotype (LG3, ZZ, Fig. 5a) when crossed with the *On* female. In both broodstocks and in wild crosses, contamination between these species may occur (Bakhoum et al. 2009; Rognon and Guyomard 2003) and therefore careful monitoring of the master SD genes alleles involved is advised when using this hybridization scheme for maintaining all-male culture. However, currently only the master gene *AMH* on LG23 was identified; and genomic selection for the relevant SD alleles is based on microsatellite analysis (Eshel et al. 2014; Shirak et al. 2006).

In the present study, we showed that contamination with 4–5.5% *Oa* alleles for both the SD and autosomal loci accumulated during two generations of 2009 and 2010 in

our *On* stock. According to the reports of the Israeli Division of Fishery and Aquaculture in 2011, 100 females of this stock produced with 40 *Oa* males approximately 60,000 fry in an experimental pond over a period of 40 days, but 27–31% females were detected in the hybrid broodstock. As a control, *On* females harboring a single *Oa* allele for any of the SD loci were crossed with *Oa* males, yielding up to 50% of females in their fry (data not shown). After culling of hybrid females with alien *Oa* alleles for the three SD loci (Table S1) and mating with *Oa* males, all-maleness was restored in progeny, although the extent of fry production was reduced by 90% (4000–6000). Thus culling of *On* females containing *Oa* alleles prior to mating with *Oa* males is essential to maintain all-male progeny.

Interestingly, LG1 was involved in SD of these hybrids' offspring, in spite of the fact that only LGs 3 and 23 are associated with SD in the parental *Oa* and *On* stocks, respectively. This may be in line with the hypothesis that some SD loci that do not segregate in intraspecific crosses become polymorphic and influential only in interspecific crosses (Hammerman and Avtalion 1979). Other explanations suggest that LG1 contains the original SD locus of *On*, as it is active in Eastern Africa *On* sub-species (Palaiokostas et al. 2013) or that it is an ancient cross-contamination of *On* stock by *O. mossambicus* or by other tilapia species with original SD locus on LG1 (Elder and Garrod 1961; Pruginin et al. 1975). It has also been shown that interaction between genetic factors and environmental temperatures may mediate LG1's effect on SD in *Oa* female \times *On* male crosses (Baroiller et al. 2009).

Under our experimental conditions, interspecific reproductive activity was apparent in only 3 out of 100 *On* females (3%). This coincides with a similar rate of contamination of alien *Oa* alleles for autosomal loci in *On* stock (Table S1; 4–5%). Moreover, we detected that 20% daughters of these active females also responded to all-specific males, producing all-male progeny or nearly all-male progeny. Enrichment of stock by responsive females from 3% to 20% through one generation supported our initial hypothesis for genetic control of the reproductive barrier between *Oa* male and *On* female. This prompted us to conduct GWAS aimed at the initial characterization of the genetic loci underlying the observed female responsiveness.

Using the GBS method, we analyzed 4983 SNPs for GWAS and found 7 SNPs with an FDR < 0.06 that were distributed in 5 regions: LGs 1, 5, 9, 14, and an unmapped scaffold 320, with the most significant effects on LGs 9 and 14. Neither *t* test nor Chi-squared or Fisher's exact tests are completely valid. *t* Test assumes a continuous trait distribution, Chi-squared test is generally considered valid only if the expected value is >5 individuals for each cell, and Fisher's exact test has been found to be too conservative (Crans and Shuster 2008), which is critical for the

very low probabilities considered. The effect of the number of rare alleles in each SNP (0, 1, 2) on the response status was analyzed. Estimating the number of rare allele as a regression effect, as is done by *t* test, is more powerful than Chi-squared or Fisher's exact tests, which only test for deviation from statistical independence. Logistic regression could not be applied owing to the small sample size and the presence of empty cells.

The combined effects of microsatellites on LGs 9 and 14 explain 37% of the variance between female responders and non-responders. Unlike GWAS, which included genotypes for part of the sample that passed the filtering criteria (36.5 individuals, on the average), the microsatellite analysis included 69 individuals. The results from this follow-up experiment, which included additional individuals and tested for new markers, support the validity of the most significant findings from GWAS of QTLs on LGs 9 and 14 affecting female responsiveness.

SNPs with the greatest effects on LGs 9 and 14 were positioned on both ends of their LGs (Fig. 2). We suspect that this distribution may be an artifact due to assembly problems in the current *On* genome version, which also makes use of linkage data. Erroneous end construction in linkage maps is often due to existence of one obligate crossing-over on short chromosomal arms (Lukaszewski and Curtis 1993) and discrepancies between the current *On* genome build and the *On* radiation-hybrid map have been indeed observed in the lower parts of LGs 9 and 14 (Guyon et al. 2012). Thus repeating GWAS with larger families and with a better genome map and coverage is essential to further narrow the critical regions for female responsiveness, for which the current confidence intervals apparently encompass the entire LGs.

Once the critical regions are obtained, marker-assisted selection (MAS) can be applied to identify females with sexual interaction. In Fig. 5, we present the putative genetic systems for SD and for reproductive interaction. A strategy of three crossing steps is proposed to control the SD system and restore the sexual interaction, thus enabling all-male progeny production. In the first crossing of *Oa* female and *On* male, both females and males are obtained in the progeny (Fig. 5b). Based on genetic test for reproductive interaction, a female responder offspring will be selected and backcrossed to *On* male. In the final cross, based on genetic tests for the original genotype for SD (aaBBXX) and for sexual interaction, the desired female will be selected from the progeny and crossed with *Oa* male (AAZZCC) to derive all-male progeny.

In conclusion, our study has shown that female responsiveness in tilapia has high heritability. We have taken the initial steps toward identifying the major genomic loci responsible for sexual incompatibility among tilapia

species. The identification of the genomic loci underlying both SD and female responsiveness will enable mass production of all-male tilapia through *Oa* male and *On* female hybridization, backcrossing, and MAS (Fig. 5b).

Data archiving

Sequencing data from this article have been deposited in ENA <https://www.ebi.ac.uk/ena/data/view/PRJEB23203>.

Acknowledgements We are grateful to Dr. Joshua R. Hyman, Director, UWBC DNA Sequencing Facility University of Wisconsin - Madison and to Mr. Eran Lahav for their advice in planning of the experiments. This project was supported by grant 362–0430 from the Chief Scientist, Ministry of Agriculture and Rural Development, Israel.

Compliance with ethical standards

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

References

- Bakhoun SA, Sayed-Ahmed MA, Ragheb EA (2009) Genetic evidence for natural hybridization between Nile tilapia (*Oreochromis niloticus*; Linnaeus, 1757) and blue tilapia (*Oreochromis aureus*; Steindachner, 1864) in Lake Edku, Egypt. *Glob Vet* 3:91–97
- Baroiller JF, D'Cotta H, Bezault E, Wessels S, Hoerstgen-Schwark G (2009) Tilapia sex determination: where temperature and genetics meet. *Comp Biochem Phys A* 153:30–38
- Beardmore JA, Mair GC, Lewis RI (2001) Monosex male production in finfish as exemplified by tilapia: applications, problems, and prospects. *Aquaculture* 197:283–301
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* 57:289–300
- Bonfield JK, Whitwham A (2010) Gap5--editing the billion fragment sequence assembly. *Bioinformatics* 26:1699–1703
- Cnaani A, Lee BY, Zilberman N, Ozouf-Costaz C, Hulata G, Ron M et al. (2008) Genetics of sex determination in tilapiine species. *Sex Dev* 2:43–54
- Crans GG, Shuster JJ (2008) How conservative is Fisher's exact test? A quantitative evaluation of the two-sample comparative binomial trial. *Stat Med* 27:3598–3611
- Desprez D, Melard C, Philippart JC (1995) Production of a high percentage of male offspring with 17-alpha-ethynylestradiol sex-reversed *Oreochromis-aureus*. 2. Comparative reproductive-biology of females and F2 pseudofemales and large-scale production of male progeny. *Aquaculture* 130:35–41
- Dor L, Shirak A, Rosenfeld H, Ashkenazi IM, Band MR, Korol A et al. (2016) Identification of the sex-determining region in flat-head grey mullet (*Mugil cephalus*). *Anim Genet* 47:698–707
- Elder HY, Garrod DJ (1961) A natural hybrid of *Tilapia nigra* and *Tilapia leucostictus* from Lake Naivasha Kenya colony. *Nature* 191:722–728
- Ellinger I (2016) The calcium-sensing receptor and the reproductive system *Front Physiol* 7:371
- 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*). G3 (Bethesda) 2:35–42
- 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
- Galman OR, Moreau J, Hulata G, Avtalion RR (1987). The second international symposium on tilapia in aquaculture. In: Pullin RSV, Bhukaswan T, Tonguthai K and Maclean JL (eds). ICLARM conference proceedings, vol 15, The Department of Fisheries, Bangkok, Thailand, The International Center for Living Aquatic Resources Management, Manila, Philippines; Bangkok, Thailand, pp 177–181.
- Groning J, Hochkirch A (2008) Reproductive interference between animal species. Q Rev Biol 83:257–282
- Guerrero RD, Shelton WL (1974) An aceto-carmin squash technique for sexing juvenile fishes. Prog Fish Cult 36:56
- Guyon R, Rakotomanga M, Azzouzi N, Coutanceau JP, Bonillo C, D'Cotta H et al (2012) A high-resolution map of the Nile tilapia genome: a resource for studying cichlids and other percomorphs. BMC Genomics 13:222
- Hach F, Sarrafi I, Hormozdiari F, Alkan C, Eichler EE, Sahinalp SC (2014) mrsFAST-Ultra: a compact, SNP-aware mapper for high performance sequencing applications. Nucleic Acids Res 42: W494–W500
- Hammerman IS, Avtalion RR (1979) Sex determination in *Sarotherodon* (tilapia). 2. Sex-ratio as a tool for the determination of genotype - model of autosomal and gonosomal Influence. Theor Appl Genet 55:177–187
- Henning F, Meyer A (2014) The evolutionary genomics of cichlid fishes: explosive speciation and adaptation in the postgenomic era. Annu Rev Genomics Hum Genet 15:417–441
- Keller-Costa T, Canario AVM, Hubbard PC (2015) Chemical communication in cichlids: a mini-review. Gen Comp Endocrinol 221:64–74
- Keller-Costa T, Hubbard PC, Paetz C, Nakamura Y, da Silva JP, Rato A et al. (2014) Identity of a tilapia pheromone released by dominant males that primes females for reproduction. Curr Biol 24:2130–2135
- Kobayashi M, Sorensen PW, Stacey NE (2002) Hormonal and pheromonal control of spawning behavior in the goldfish. Fish Physiol Biochem 26:71–84
- Lahav M, Lahav E (1994) The development of all-male tilapia hybrids in Nir-David. Isr J Aquacult-Bamid 42:58–61
- Lee BY, Lee WJ, Streebman JT, Carleton KL, Howe AE, Hulata G et al. (2005) A second-generation genetic linkage map of tilapia (*Oreochromis* spp.). Genetics 170:237–244
- Lee BY, 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 H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595
- Li MH, Sun YL, Zhao JE, Shi HJ, Zeng S, Ye K et al (2015) A tandem duplicate of anti-mullerian hormone with a missense SNP on the Y chromosome is essential for male sex determination in Nile tilapia, *Oreochromis niloticus*. PLoS Genet 11:e1005678
- Lister AL, Van der Kraak GJ (2009) Regulation of prostaglandin synthesis in ovaries of sexually-mature zebrafish (*Danio rerio*). Mol Reprod Dev 76:1064–1075
- 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
- Longrie N, Poncin P, Denoel M, Gennotte V, Delcourt J, Parmentier E (2013) Behaviours associated with acoustic communication in Nile tilapia (*Oreochromis niloticus*). PLoS ONE 8:e61467
- Lovshin LL. (1980). The international conference on the biology and culture of tilapias. In: Pullin RSV, Lowe-McConnell RH (eds) ICLARM conference proceedings, vol 7, The International Center for Living Aquatic Resources Management, Manila, Philippines; Bellagio, Italy, pp 279–308.
- Lukaszewski AJ, Curtis CA (1993) Physical distribution of recombination in b-genome chromosomes of tetraploid wheat. Theor Appl Genet 86:121–127
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A et al (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303
- Mires D (1980). The international conference on the biology and culture of tilapias. In: Pullin RSV, Lowe-McConnell RH (eds) ICLARM conference proceedings, vol 7, The International Center for Living Aquatic Resources Management, Manila, Philippines; Bellagio, Italy, pp 317–329.
- Mires D (1995) The tilapias. In: Nashh CE, Novotny AJ (eds) Production of aquatic animals: fishes. Elsevier Science, Amsterdam, p 133–152
- 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 PLoS ONE 8:e68389
- Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach PLoS ONE 7: e32253
- Pruginin Y, Rothbard S, Wohlfarth G, Halevy A, Moav R (1975) All male broods of *Tilapia nilotica* X *Tilapia 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
- Rothbard S, Pruginin Y (1975) Induced spawning and artificial incubation of tilapia. Aquaculture 5:315–321
- 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
- Seroussi E, Ron M, Kedra D (2002) ShiftDetector: detection of shift mutations. Bioinformatics 18:1137–1138
- Shirak A, Golik M, Lee BY, 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, Reicher S, Uliel S, Mehlman T, Shainskaya A, Ron M et al (2012) Gender and sexual behavior modulate the composition of serum lipocalins in Nile tilapia (*Oreochromis niloticus*). Gene 504:22–30
- Shirak A, Seroussi E, Cnaani A, Howe AE, Domokhovskiy R, Zilberman N et al. (2006) *Amh* and *Dmrt2* genes map to tilapia (*Oreochromis* spp.) linkage group 23 within quantitative trait locus regions for sex determination. Genetics 174:1573–1581
- Shirak A, Seroussi U, Gootwine E, Seroussi E (2017) Sequence motifs capable of forming DNA stem-loop structures act as a replication diode. FEBS Open Bio 7:944–952
- Simoes JM, Barata EN, Harris RM, O'Connell LA, Hofmann HA, Oliveira RF (2015) Social odors conveying dominance and reproductive information induce rapid physiological and neuro-molecular changes in a cichlid fish BMC Genomics 16:114
- Sorensen PW, Baker C (2015) Species-specific pheromones and their roles in shoaling, migration, and reproduction: a critical review

- and synthesis. In: Sorensen PW, Wisenden BD (eds) Fish pheromones and related cues. Wiley-Blackwell, Iowa, pp 11–32
- Sorensen PW, Hara TJ, Stacey NE, Goetz FW (1988) F-prostaglandins function as potent olfactory stimulants that comprise the post-ovulatory female sex-pheromone in goldfish. *Biol Reprod* 39:1039–1050
- Stacey N (2015) Hormonally-derived pheromones in teleost fishes. In: Sorensen PW, Wisenden BD (eds) Fish pheromones and related cues. Wiley-Blackwell, Iowa, pp 33–88
- Statistical Analysis System (1999) SAS/STAT user's guide, version 8. SAS Institute, Cary, N.C., p 3884
- Stelkens RB, Young KA, Seehausen O (2010) The accumulation of reproductive incompatibilities in African cichlid fish. *Evolution* 64:617–632
- Timan B, Shirak A, Segal R, Machnes Z, Avtalion RR (2002) Epitopes of tilapia red blood cells. I. species-specific antibodies for the control of tilapia breeding stocks. *Isr J Aquacult-Bamid* 54:5–11
- Trombka D, Avtalion R (1993) Sex determination in tilapia - a review. *Isr J Aquacult-Bamid* 45:26–37
- Wohlfarth GW, Rothbard S, Karplus I, Harpaz S, Halevy A (1994) The association between initial weight and growth in tilapias during communal testing. I. Pond testing. *Isr J Aquacult-Bamid* 46:83–88
- Wohlfarth GW, Wedekind H (1991) The heredity of sex determination in tilapias. *Aquaculture* 92:143–156
- Zak T, Deshev R, Benet-Perlberg A, Naor A, Magen I, Shapira Y et al (2014) Genetic improvement of Israeli blue (Jordan) tilapia, *Oreochromis aureus* (Steindachner), through selective breeding for harvest weight. *Aquac Res* 45:546–557
- Zhu HP, Liu ZG, Lu MX, Gao FY, Ke XL, Ma DM et al (2016) Screening and identification of a microsatellite marker associated with sex in Wami tilapia, *Oreochromis urolepis hornorum*. *J Genet* 95:283–289