Evidence supports that sex determination (SD) in tilapia is controlled by major genetic factors that may interact with minor genetic as well as environmental factors, thus implying that SD should be analyzed as a quantitative trait. Quantitative trait loci (QTL) for SD in Oreochromis niloticus were previously detected on linkage groups (LG) 1 and 23. Twenty-one short single repeats (SSR) of >12 TGs and one single nucleotide polymorphism were identified using the unpublished tilapia genome sequence on LG23. All markers showed two segregating alleles in a mapping family that was obtained by a cross between O. niloticus male (XY) and sex-reversed female (ΔXY) yielding 29 females (XX) and 61 males (XY and YY). Interval mapping analysis mapped the QTL peak between SSR markers ARO172 and ARO177 with a maximum F value of 78.7 ($P < 7.6 \times 10^{-10}$). Twelve adjacent markers found in this region were homozygous in females and either homozygous for the alternative allele or heterozygous in males. This segment was defined as the sex region (SR). The SR encompasses 1.5 Mbp on a single tilapia scaffold (no. 101) harboring 51 annotated genes. Among 10 candidate genes for SD that were tested for gene expression, anti-Müllerian hormone (Amh), which is located in the center of the SR, showed the highest overexpression in male vs. female embryos at 3 to 7 days postfertilization.

Sex determination (SD) can be controlled by one or more genetic factors, environment or their interactions, involved SD factors located on sex chromosomes and/or on either autosomes (Bull 1981). The sex chromosomes are characterized by both morphologically undifferentiated and differentiated homologs, in simple and multiple systems with male or female heterogamy. Studies on organisms with differentiated sex chromosomes, male heterogametic (mammals and fly) and female heterogametic (birds and reptiles), have shown interesting similarities between the two systems (XY/XX and ZW/ZZ) of sex determination (Ellegren 2011).

Teleost species are an interesting model for SD research, with a variety of SD systems and capability of producing viable hybrids between closely related species having different SD systems (Mank et al. 2006). Teleost fish have diverged from land vertebrates more than 450 million years ago, after they diverged from birds and mammals, but before they diverged from each other (Bellott et al. 2010). Moreover, these fish species experienced whole-genome duplication (ancestral tetraploidy), followed by variable-rate reduction of ploidy that significantly complicates the identification of orthologs (Kasahara et al. 2007). Bellott et al. (2010) compared zebrafish, Tetraodon, pufferfish, and medaka genomes to mammalian X or avian Z chromosome and reported that most orthologs to Z and X genes occupy separate portions of each fish genome.

Different aspects of tilapia SD have been explored because tilapias are an important aquaculture commodity (Devlin and Nagahama 2002). Their commercial production relies on all-male monosex culture, which so far has proved difficult to maintain in large-scale production facilities (Cnaani and Levavi-Sivan 2009). A better understanding of the genetic basis of SD in tilapia is needed to overcome these difficulties.

The differences in SD mechanisms among closely related tilapia species and the influence of the environment (Baroiller et al. 2009) suggest that SD should be analyzed as a quantitative trait using a markers-based QTL approach. Various sex-linked markers have been identified in O. niloticus and O. aureus (Lee et al. 2003, 2004; Shirak et al. 2002, 2006; Eshel et al. 2010) and mapped to different LG. In purebred O. niloticus and O. niloticus × O. aureus hybrids, the QTL...
were detected on LG1, LG23 (Lee et al. 2003; Eshel et al. 2010) and on LG3 (Lee et al. 2005), respectively. The SD QTL on LG23 was mapped within a confidence interval (CI) of 16–21 cM (Eshel et al. 2010; Figure 1B), which harbors the genes Ansh and Dmrt2 that are involved in the vertebrate SD cascade (Shirak et al. 2006, Figure 1A). The first assembly version of the unpublished tilapia genome, consisting of 5900 scaffolds, was recently released (Accession no. PRJNA59571). Using this information enabled us to refine the confidence interval of SD QTL on LG23 and find positional candidate genes for the sex master-key regulators in the Swansea stock of *O. niloticus*.

The critical period for SD in tilapias is 0–18 days postfertilization (dpf). During this period, embryos are sensitive to androgens, estrogens, and precursors of steroids through immersion and dietary exposure (Devlin and Nagahama 2002). In a more recent study, Ijiri et al. (2008) demonstrated that differential expression of genes in XX and XY gonads of *O. niloticus* during the period of 9–10 dpf is critical for differentiation of primordial germ cells (PGC) into either ovary or testis. Rougeot et al. (2008) applied temperature treatment on all-female population embryos until hatching (3–4 dpf) and showed ~20% phenotypic sex change of females to males. Furthermore, Palti et al. (2002) demonstrated by using genetic markers that sex-specific mortality occurs shortly after hatching. On the basis of these findings, we hypothesized that master regulation genes initiating the SD cascade should be expressed before the detectable differences in the PGCs. Recent studies revealed that the major genes involved in the SD pathway are common to mammals and fish (Schartl 2004). Moreover, the upstream genes on the SD cascade may vary among organisms, but downstream components tend to be conserved (Charlesworth and Mank 2010). To study the onset of the SD cascade at early stages in embryonic development, we selected eight genes (*Lhx9*, *Amh*, *Foxl2*, *Cyp19a*, *Dmrt1*, *DAX1*, *Sox9a*, and *Sox9b*) with a known role in the SD pathway of other organisms (Birk et al. 2000; Shirak et al. 2006; Ijiri et al. 2008) and examined their expression from gastrula to late larval stage (2–9 dpf) in *O. niloticus* monosex progeny of XY males and XX females. Additionally, we analyzed two genes previously mapped to the SD region: *Sox14* (Cnaani et al. 2007) and ELAVL1 (A. Shirak, unpublished data) with ambiguous similarity to known genes on the SD pathway. Gene expression profiles represent the

![Figure 1](https://example.com/figure1.png)

**Figure 1** Linkage and physical map of SR on LG23 and interval mapping of the SD QTL. (A) LG23 genetic map by genotyping of *O. niloticus*♂ × *O. aureus*♀ family (Shirak et al. 2006). (B) Sex determination QTL on LG23 for *O. niloticus* families based on nine SSR markers (Eshel et al. 2010). (C) Fine mapping of the QTL region with additional mapping family (sex reversed) with 33 markers. (D) Interval mapping of the QTL based on panel C. (E) LG23-related scaffolds. QTL, quantitative trait loci; SD, sex determination; SR, sex region; SSR, short single repeat.
primary level of integration between environmental factors and the genome, providing the basis for phenotypes, such as morphology and behavior. Therefore, we examined differences in candidate genes expression between genders during early embryonic development for initiation of the SD cascade in tilapia.

The objectives of this study were to refine the sex region on LG23 using both linkage and physical mapping and to identify candidate genes for SD in the region with differential expression at early embryonic development.

MATERIALS AND METHODS

Breeding of *O. niloticus* (Swansea stock) families used for this study was performed at the Agricultural Research Organization, Israel.

Mapping family

The inheritance of gender in a cross between *O. niloticus* male (XY) and a sex-reversed neofemale (DXY) that yielded 29 females (XX) and 61 males (XY and YY) was validated by segregation of the sex-linked marker UNH898 (Eshel et al. 2010).

**Monosex groups**

To obtain all-female (XX) and all-male (XY) progenies, eggs of a single *O. niloticus* female (XX) were divided into two groups, and each group was artificially fertilized with either milt of a sex-reversed male (DXY) or milt of a genetically modified male (YY). Sex was determined at age of three months by gonadal squash of at least 100 individuals per each full-sib group (Mair et al. 1997).

**Development of SSR markers:** We ran BLASTN search for the *Amh*, *Dmrt2*, and *Sox14* genes and for nine SSR markers that were previously mapped to LG23 against the unpublished tilapia genome (Accession no. PRJNA59571; http://cichlid.umd.edu/blast/blast.html). Hits were found in three scaffolds: no. 7, 101, and 29. We searched for tandem

| Scaffold | Marker/Gene | Forward Primer | Reverse Primer |
|-----------|-------------|----------------|----------------|
| 7         | ARO765      | CTTGAAACTCAGGGCTGTA | GCTTCACCAAAGGCAGCA |
| 7         | ARO772      | GCCTTTGTCACCTGAGGAG | AACCTCGCCTCTCTGGAA |
| 7         | ARO780      | TGTGGGCTTCTTGACAG | GAAACCCCTCTCTGGTG |
| 7         | ARO790      | TGAGAAAGAGGGCTGCTT | GCTGGGTCAGGGTTCGTA |
| 101       | ARO104      | AGAAGGACGCTCTGCTTGTC | TCTCCATCCACTCTGCA |
| 101       | ARO133      | GTGGGCTCAAGGCTTGCTT | GCCGCAATCACAGTCGA |
| 101       | ARO158      | GTGGGCAAAAACAGGCTT | TTGTTCCAGTAAAGGTG |
| 101       | ARO196      | GATTGTAAGCGCTGTAAGT | TCTGGTCTGTCTGTCG |
| 101       | ARO145      | CAATGTCGCTGCTGCTT | GCCTGTGACAGAAGC |
| 101       | ARO172      | AGGCTCTGATGCGTCTT | ACCCTTGAGTAAAGGCA |
| 101       | ARO177      | CCGCTCCCTGACTCTTGC | GTGCAAGGCAAATGAAAGC |
| 101       | ARO195      | CATGCTGATGAGGACGAG | TCAAGAGCTAAAGGAGT |
| 101       | ARO120      | AAGGGAAGTTGCTGCTGCT | GTGTCTCCCCACAGTTC |
| 101       | ARO121      | GTGGGACTGTTGTTAGT | GCTGGTAAAGCAAGCA |
| 101       | ARO114      | AGGAGAATTCAGGGATGCA | GCGCACTGGCTGTGTCAT |
| 101       | ARO118      | TGATCTTCCACAGGCAACA | GTGGTGGTCCAACAGC |
| 101       | ARO123      | TTAATCCGATGCCAGCTAAC | AAGCCTCAATTCATCTTGA |
| 101       | ARO124      | CGAGCTGCTTGGTCTGTA | CGAACCATGGAAATG |

SSRs, short single repeats.

| Table 1 Primers for SSRs and ELAV-like 1 gene used in this study and their locations on the unpublished tilapia genome scaffolds |
|-----------|----------------|----------------|----------------|
| Scaffold | Marker/Genes | Forward Primer | Reverse Primer |
|-----------|-------------|----------------|----------------|
| 7         | ARO765      | CTTGAAACTCAGGGCTGTA | GCTTCACCAAAGGCAGCA |
| 7         | ARO772      | GCCTTTGTCACCTGAGGAG | AACCTCGCCTCTCTGGAA |
| 101       | ARO104      | AGAAGGACGCTCTGCTTGTC | TCTCCATCCACTCTGCA |

Extension primer: [AAAAAA]CAAACACCTGAACGGACACAC^a

| Table 2 Primers for genes analyzed by qPCR |
|-----------|----------------|----------------|----------------|
| Gene      | Forward Primer | Reverse Primer |
| Amh       | GCACCCAGCTGAGTACAAC | GTGGGAAGGTCAAAATGCA |
| Cyp19a1a^a| GATAGGGCAGCCAGCAGAC | GTCCAGCTGAGTCAAGTGTA |
| Dax1      | CAGATCTGAGGATTGCC | GATGGATCAGCCCTGAG |
| Dmrt1^a   | CGGCCAGGTGCTGTGAG | CCAATCTGCTGCTGCA |
| ELAV-like 1 | CAGGCTCCCTGACTCTTCAG | GTGGTGTCAGCTGTTTGA |
| FoxL2     | CAGGCAAGGAGGAGGACGA | TGGCAAGGAGAGAGCATAG |
| Lhx9      | GATACCTACAGGCTTCCGTCGAG | TCAAGCTGATACAGGAGC |
| Sox14     | TGCTCAAGAGGACGGCTTACG | AAGACCCCTAGGAGAGTCC |
| Sox9a     | GCAAACTTGGGACGTTTCGA | TCGGCTGACAGCTCTTATG |
| Sox9b     | GAGAGAGTAGGAGTCA | TCAAGAGTGAGCTGAG |
| GAPDH     | GGCATGGCTGAGGCTCAT | CATTCTACCCAGGAGGCGT |

^a Ijiri et al. (2008).
At least 12 TG repeats in scaffold 7 (6,500,000–9,485,422 bp), in scaffold 29 (3,291,196–5,141,938 bp), and in the entire scaffold 101. We entered the sequence of 200 bp upstream and downstream of the TG repeats core to Primer3 software and developed 21 novel SSR markers. Polymorphism of these markers was tested in parents of the mapping family. To develop genetic markers in the vicinity of UNH216, we mapped in our family the marker UNH2190, which was derived from the Malawi cichlids hybrid *Metriaclima zebra* × *Labotrophodes fueleborni* and was mapped adjacent to UNH216 (Albertson et al. 2003).

**Development of the SNP marker:** On the basis of partial cDS sequence (GI: 93115149) of *O. mossambicus ELAVL1* (embryonic lethal, abnormal vision, Drosophila-like 1), we identified the SNP polymorphism A/G (Table 1) at nucleotide 391 in our mapping family.

**DNA extraction and genotyping of SSR and SNP markers:** DNA was isolated from fin samples by the “salting out” high-throughput procedure (Zilberman et al. 2006). The concentration of the DNA was quantified with NanoDrop spectrometer (NanoDrop Technologies, DE), and each DNA sample was diluted to a final concentration of ~10 ng/μl. PCR amplification was performed in a total volume of 10 μl with Super-Therm Taq DNA polymerase (JMR Holding, London), mixture of 2 mM dNTPs of each nucleotide, and primer concentration of 10 pmol/μl (Metabion GmbH, Germany). PCR conditions were 3 min at 94°C; 40 sec at 94°C, 40 sec at 61.5°C, 1 min at 72°C for 30 cycles and 10 min at 72°C. The mapping family was amplified for SSR markers, and genes with primers taken from NCBI database or designed based on scaffold sequence (detailed list in Table 1), where one primer in each pair was 5’ end-labeled by HEX, TET, or FAM fluorescent dyes (Operon Technologies, Alameda, CA). Size calling of PCR products was determined using ABI GeneMapper software version 4.0 (Applied Biosystems, Foster City, CA) after electrophoresis in a capillary gel on ABI-3130 apparatus. Sequencing and SNaPshot reaction for genotyping of SNP markers were also carried out on ABI-3130 according to the manufacturer instructions using the primers specified in Table 1. (For additional data, see supporting information, File S1.)

**Linkage and interval mapping:** The linkage map for LG23 using segregating markers in our mapping family was reconstructed by CRIMAP software (http://linkage.rockefeller.edu/soft/crimap/).

The interval mapping was based on a nonlinear regression using the method of Knott et al. (1996), with the program developed by Spelman et al. (1996). The test statistic and locus effects were evaluated at 1 cM intervals. The 95% confidence intervals (CI) for the QTL location and effect were determined by generation of 200 bootstrap samples.

**Identification of genes and annotation:** Annotation of genes positioned in the SR was performed by combining three bioinformatics resources: (1) EST contigs assembled via MIRA program with BLASTN (Lee et al. 2010); (2) comparative mapping to other fish genomes by BLASTX; and (3) BouillaBase annotation using Maker Gene pipeline (http://cichlid.umd.edu/cgi-in/gb2/gbrowse/Tilapia_broad_scaffolds_v1/?source=Tilapia_broad_scaffolds_v1) using Maker pipeline. Further analysis between tilapia and stickleback group VIII (15.4–16.7 Mbp) (http://www.ensembl.org/index.html) identified all 39 genes in the same order in both genomes. On the basis of this high level of orthology, we used stickleback orthologous region on UCSC genome browser (http://genome.ucsc.edu/) as anchor for similarity with Tetraodon (chr1:13.4–14.7 Mbp), medaka (chr4:7.4–8.8 Mbp), fugu (scaffold 25:0.4–1 Mbp), zebrafish (chr22:19.1–21.6 Mbp), and human (chr19p13.1:1.5–5 Mbp), and we detected 12 additional genes in tilapia.

**RNA extraction and qPCR:** A pool of 20–30 embryos from each gender were placed in RNA later reagent (Qiagen) to stabilize the RNA and then stored at −20°C until RNA extraction. Total RNA was

### Table 3 Level of normalized relative expression ± SD and statistical significance of sex-specific differences for gene candidates for SD in embryos at 2 to 9 dpf

| Gene  | dfp  | All Female       | All Male       | Probability |
|-------|------|------------------|----------------|-------------|
| ELAVL1| 2    | 14.5 ± 2.2       | 3.2 ± 2.1      | **          |
|       | 5    | 42.1 ± 11.8      | 0.5 ± 11.8     | *           |
|       | 7    | 22.8 ± 5.8       | 1.2 ± 5.4      | *           |
|       | 9    | 2.7 ± 0.6        | 0.2 ± 0.6      | *           |
| Amh   | 3    | 0.02 ± 0.005     | 0.04 ± 0.005   | *           |
|       | 4    | 0.2 ± 0.2        | 3.5 ± 0.2      | ***         |
|       | 5    | 0.5 ± 0.9        | 5.2 ± 0.9      | **          |
|       | 6    | 0.4 ± 0.2        | 5.8 ± 0.2      | ***         |
|       | 7    | 0.3 ± 0.7        | 3.7 ± 0.7      | **          |
| Lhx9  | 2    | 3.1 ± 0.18       | 0.8 ± 0.2      | ***         |
| Sox9a | 1.6 ± 0.17       | 0.9 ± 0.2      | *           |
| Sox9b | 7    | 6.3 ± 0.6        | 4 ± 0.6        | *           |
| Fox2  | 0.9 ± 0.06       | 0.6 ± 0.06     | *           |
| Fox14 | 0.3 ± 0.07       | 0.55 ± 0.06    | *           |

Asterisks represent levels of significance for sex-specific expression differences: *P* ≤ 0.05, **P** ≤ 0.01, and ***P*** ≤ 0.001. dfp, days post fertilization.
was used as an internal standard. Final results were analyzed by Student
et al."
Expression data for all genes embedded within the SR on LG23 were
gene expression in SD-associated organs: Detection System, Applied Biosystems), was determined by comparison
amount (IA) according to User Bulletin #2 ABI (PRISM7700 Sequence
products. The relative amount of the target RNA, called the input
curve analysis was performed for each gene for speci-
differences for the other 4 genes (Table 2). The Amh gene is located between these markers. The
counting the newly developed markers narrowed down the SR to scaffold 101
between markers GM597 and ARO124 from 990,577 to 2,468,000 bp
(Figures 1E and 2). The Amh gene is located between these markers. The
classification of qPCR for each reaction
primers were designed so at
mRNA to an intron-exon boundary (Table 2). The
expression analysis of 10 SD-related genes and two genes mapped
detected for the remaining 7 genes as pre-
expression (Table 3). Figure 3A presents the continued elevation in
expression data, 4 out of the remaining 10 genes showed overexpression in at
expression by over 3-fold. Thirteen genes were found relevant to SD
in the brain; expression of 15 of these genes exceeded the median
expression values, whereas each bar along the X axis indicates a sam-
expression from
$P < 7\times 10^{-15}$ at 22 cm (Figure 1D). This region
cutaneous glands, was localized to scaffolds 7, 101, and 29. Physical mapping of the scaffolds
with the newly developed markers narrowed down the SR to scaffold 101
between markers GM597 and ARO124 from 990,577 to 2,468,000 bp
(Figures 1E and 2). The Amh gene is located between these markers. The
Gene expression at early developmental stages
Expression analysis of 10 SD-related genes and two genes mapped
within the SD QTL on LG23 was performed for embryos of known type (XX or XY) during 2–9 dpf. No significant differences between
genders were found for cyp19a, Dmrt1, or Dax1. Significant sex-
differential expression was detected for the remaining 7 genes as pre-
all 10 tested genes. Significant sex difference was found for Lhx9
expression ($P = 0.0002$) at 2 dpf, equivalent to the developmental
stage of segmentation (Fujimura and Okada 2007) but not later in
the embryonic development. Likewise, sex-differential expression was
detected for ELAVL1 ($P < 0.01$) at age of 2 dpf but was attenuated at
5, 7, and 9 dpf (P ≤ 0.05) (Figure 3B). Gender-specific expression
differences for the other 4 genes (Foxl2, Sox9a, Sox9b, and Sox14) were
detected at later developmental stages (6–9 pdf).
Characterization of genes positioned within the SR
Fifty-one genes were identified within the SR and are presented in
Table 4. Thirty-nine genes had expression data in a variety of 91 tissues
of mammals in the BioGPS database. We focused on three SD-related
organs that are relevant in the tilapia SD cascade: brain, testis, and
pituitary. Interestingly, 17 out of the 39 genes showed overexpression in
the brain; expression of 15 of these genes exceeded the median
expression by over 3-fold. Thirteen genes were found relevant to SD
following a literature survey. After removing 3 genes with no expression
data, 4 out of the remaining 10 genes showed overexpression in at
least one SD-associated organ (Notch2, Pias5, Zbtr7, or Celf5).
Comparative mapping
Comparative analysis of the genes positioned within the SR detected
high level of orthology between tilapia and six different species. Within
$< 1.3$ Mbp region of stickleback, Tetraodon, fugu, zebrafish,
medaka, and human, 40, 39, 29, 29, 29, and 21 orthologous genes,
Table 4 SD-related data for annotated genes in the SR on scaffold 101 between 990,577 and 2,468,000 bp

| Gene or Symbol | Accession or Ensembl No. | Scaffold 101 (bp) | SD-related Publications | Gene Expression |
|----------------|--------------------------|-------------------|-------------------------|-----------------|
|               |                          | Start             | End                      | Brain | Testis | Pituitary |
| PLIN3         | NP_001167399.1           | 1187253           | 1192169                 |       |        |           |
| ZFAND6        | XP_002199446.1           | 1194645           | 1198048                 |       |        |           |
| FAM108C1      | XP_001342996.2           | 1205385           | 1210274                 |       |        |           |
| RGL1          | NP_991200.1              | 1216521           | 1224747                 |       |        |           |
| GLT25D2       | ENSGACT00000016402       | 1245440           | 1265427                 |       |        |           |
| Novel protein | ENSGACT00000016408       | 1273512           | 1277395                 | ND    |        |           |
| C1orf21       | ENSGACT00000016410       | 1273570           | 1287700                 |       |        |           |
| EDEM3         | XP_688275.4              | 1305007           | 1311942                 |       |        |           |
| NPL           | NP_001133311.1           | 1314238           | 1318468                 |       |        |           |
| SEC22B        | ACM09163.1               | 1328257           | 1332317                 |       |        |           |
| NOTCH2a       | NP_001108566.1           | 1348777           | 1380740                 | Zhu et al. 2007 |       |           |
| SLC35A3       | ACN10890.1               | 1384970           | 1391950                 |       |        |           |
| FAM7BB        | CAQ14615.1               | 1390770           | 1404535                 | ND    |        |           |
| C19orf60      | NP_001158740.1           | 1406814           | 1410851                 |       |        |           |
| CRLF1         | NP_001002650.1           | 1412045           | 1421020                 |       |        |           |
| TMEM59L       | ENSGACT00000016558       | 1422768           | 1431927                 |       |        |           |
| Novel protein | ENSGACT00000016563       | 1431927           | 1433134                 | ND    |        |           |
| SSBP4         | NP_001184034.1           | 1502998           | 1511875                 |       |        |           |
| FKB8          | NP_001133417.1           | 1569993           | 1579098                 |       |        |           |
| ELLb          | NP_956001.1              | 1603737           | 1613759                 | Zhou et al. 2009 |       |           |
| DOT1          | CAP09616.1               | 1667805           | 1679590                 |       |        |           |
| Amh           | AB558513.1               | 1688658           | 1695299                 |       |        |           |
| Novel protein | ENSGACT00000016737       | 1691664           | 1695317                 |       |        |           |
| OAZ1          | NP_001134904.2           | 1696653           | 1701176                 |       |        |           |
| dkey-3k20.4   | ENSGACT00000016747       | 1705254           | 1704891                 | Dai et al. 2009 |       |           |
| ORGp          | NP_001093824.1           | 1709490           | 1716883                 |       |        |           |
| LINGO3        | ENSGACT00000016753       | 1769300           | 1771737                 |       |        |           |
| Novel protein | ENSGACT00000016755       | 1783055           | 1787782                 | ND    |        |           |
| AT8P33        | XP_003201012.1           | 1808332           | 1824572                 |       |        |           |
| ONECUT3       | ENSGACT00000016776       | 1832236           | 1848746                 |       |        |           |
| PIAS4a        | AAI57528.1               | 1876824           | 1885646                 |       |        |           |
| MAP2K2a       | XP_002761634.1           | 1892758           | 1902715                 |       |        |           |
| ZBTB7a        | CAK04316.1               | 1917084           | 1923570                 |       |        |           |
| TCF3a         | NP_001187227             | 1948472           | 1961119                 |       |        |           |
| OCR10a        | ACQ58208.1               | 1968852           | 1973040                 |       |        |           |
| MBBD3b        | CAK10918.1               | 1973187           | 1979287                 |       |        |           |
| MIYOSb        | CAK10917.1               | 1980557           | 1993470                 |       |        |           |
| UNCI3A        | NP_001038363.1           | 1997106           | 2036658                 |       |        |           |
| HMG20a        | NP_001018387.1           | 2039148           | 2044881                 |       |        |           |
| EE2F2         | NP_956572.2              | 2047515           | 2063389                 |       |        |           |
| SNORD37       | ENSGACT000000029482      | 2050851           | 2051406                 |       |        |           |
| Rxfp3b        | NP_001077348.1           | 2080930           | 2081856                 | Wilson et al. 2009 |       |           |
| CREB3         | NP_001018509.1           | 2088906           | 2099759                 |       |        |           |
| CELFS5        | NP_001124260.1           | 2263338           | 2289589                 |       |        |           |
| RIGMA         | NP_001133684.1           | 2299255           | 2306831                 |       |        |           |
| H5D11B1Lb     | NP_001098261.1           | 2342628           | 2365654                 |       |        |           |
| QIL1          | AC169344.1               | 2365927           | 2368604                 |       |        |           |
| SPIN1         | XP_001339043.4           | 2370662           | 2372447                 |       |        |           |
| CFD           | AC169308.1               | 2392555           | 2404506                 |       |        |           |
| BTBD8a        | EAW73104.1               | 2406880           | 2407219                 |       |        |           |
| CYL1P1        | ACN58730.1               | 2459356           | 2488052                 |       |        |           |

1 indicates 2-3x from median expression in BioGPS; √√ indicates >3x from median expression in BioGPS; ND, no data; SD, sex determination; SR, sex region.

2 Gene related to SD in transcriptional processes.

3 Gene associated with gonadal development/function.

respectively, have been found in the same order. GO term enrichment analysis of these genes with DAVID software (Huang et al. 2009) yielded 4 genes, Notch2, ELL, Amh, and TCF, which are involved in biological processes of cell differentiation, cellular, and anatomical structure development based on zebrafish background of 839 genes.

**DISCUSSION**

Different SD systems with remarkable variation have been observed in teleosts (Vollff and Scharl 2002). Evidence supports that sex determination (SD) in tilapia is controlled by major genetic factors that may interact with minor genetic as well as environmental factors, thus implying that SD should be analyzed as a quantitative trait.
QTL for SD in *Oreochromis niloticus* were previously detected on LG1 and LG23 (Lee et al. 2003; Cnaani et al. 2004; Eshel et al. 2010). In the present study, interval mapping analysis using 33 markers on LG23 detected the QTL peak between two adjacent genetic markers: ARO172 and ARO177. However, the confidence interval was still rather large between 5 cm (Eshel et al. 2010; 156 individuals) and 30 cm in the current study. Thus, mapping QTL to confidence interval < 5 cm is not a viable option using genetic markers and segregating families of moderate size (Ron and Weller 2007). However, using physical mapping based on the unpublished tilapia genome sequence, all 26 markers in the QTL were physically mapped to three scaffolds on LG23. Furthermore, recombination in two females were used to identify the boundaries of the SR between markers GM597 and ARO124 on a single scaffold (no. 101; Figure 2). This explains the lack of power of the interval mapping that is based on bootstrap analysis of a family of 90 individuals of which only two are informative. Figure 2 demonstrates the distinct contrast of genotypes for markers between genders in the specific sex region. The absence of recombination along a region of 12 genetic markers may reflect the moderate size of the family, but it also conforms to the theory that the evolution of sex chromosomes involves suppressed recombination between homologous chromosomes to maintain sex-related gene blocks (Bergero and Charlesworth 2009). The SR encompasses 1.5 Mbp harboring 51 annotated genes. Our assumption that the SR harbors sex-related or male-determining genes is strengthened by the conservation of this region in other teleost fish. Out of 51 genes that were positioned within the SR, 40 and 39 orthologous genes have been found within <1.3 Mbp region of stickleback and Tetraodon, respectively. Information from the literature indicates the putative role of 13 out of the 51 genes in SD: 10 genes in transcriptional processes related to SD and 3 in gonadal development and function (Table 4).

We examined expression of genes in the SD pathways at early developmental stages of tilapia. Previous studies on SD-related gene expression in tilapia focused on brain, PGS, and gonads (Jiri et al. 2008; Poonlaphdecha et al. 2011). The results from our study on expression data of 10 candidate genes indicate that the onset of the SD cascade begins at 2 dpf at the gastrulation stage, based on overexpression of *Lhx9* and *ELAVL1* in females. *Lhx9* was found to be essential for mouse gonad formation (Birk et al. 2000). *ELAVL1* is a member of CELF proteins implicated in cell-specific and developmentally regulated alternative splicing (Ladd et al. 2001). Additional SD-related genes were *Sox9*, which is necessary and sufficient to cause testicular differentiation in mammals (Vidal et al. 2001). Likewise, *Fox2* plays a role in ovarian sex differentiation and has been suggested to function as a repressor of the male pathway during ovarian development (Ottolenghi et al. 2005). Significant differences in expression between genders for *Sox9a, Sox9b, Fox2*, and *Sox14* genes were detected in later stages of embryonic development and may indicate their downstream role in the SD cascade. We detected higher expression of *SOX9* in females than in males at 7 dpf, in contrast to the results of Jiri et al. (2008) of higher expression in male gonads at 37–70 days posthatching. *Fox2* was also highly expressed in females at 8 dpf as was previously reported (Jiri et al. 2008). Among 10 candidate genes, *Amh*, which is located in the center of the SR, showed the highest expression in male vs. female embryos. Our observation was supported by Poonlaphdecha et al. (2011) who reported on dimorphic expression of *Amh* between genders in adult gonads and brains as well as in embryo heads at 10 and 15 dpf. GO term enrichment analysis detected 4 genes, including *Amh*, that are involved in biological processes of cell differentiation, cellular development, and anatomical structure development. Genes playing a role in SD initiation with dimorphic expression between genders may be considered as candidate genes and should be further investigated.

To test the role of *Amh* and other candidate genes in SD of tilapia, targeted strategies could be considered, such as (i) mutant detection in candidate genes, as performed in zebrafish (Demarest et al. 2011); (ii) gene silencing using siRNA technology, as applied in the giant freshwater prawn (Ventura et al. 2009); and (iii) transgenesis using the *To12* system, which was demonstrated for Nile tilapia (Fujimura and Kocher 2011). Large-scale experiments might involve (i) genomic mutagenesis together with sex reversal, phenotypic mutant screening, and sequence analysis, as was applied in a medaka SD study (Otake et al. 2006); and (ii) a whole-transcriptome scan for gene expression at early embryonic development to identify the key regulators of SD. A complete computational approach was pursued to design a 44K features microarray (O. Eshel, unpublished data) based on the unpublished tilapia genome sequence annotation and EST libraries (Lee et al. 2010) for construction of the full tilapia gene list.

**ACKNOWLEDGMENTS**

We thank Berta Levavi-Sivan for providing us the ΔXX males developed in her laboratory. We thank the Broad Institute Genome Sequencing Platform and Genome Sequencing and Analysis Program, Federica Di Palma, and Kerstin Lindblad-Toh for making the unpublished genome sequence data for *Oreochromis niloticus* available. This work is a contribution from the Institute of Animal Science, Agricultural Research Organization, Bet Dagan, Israel, No. 592/11. This research was supported by Grant IS-3995-07 from the United States–Israel Binational Agricultural Research and Development (BARD) Fund and by Grant No. 801/11 from the Israeli Science Foundation.

**LITERATURE CITED**

Albertson, R. C., J. T. Streelman, and T. D. Kocher, 2003 Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. Proc. Natl. Acad. Sci. USA 100: 5252–5257.

Baroiller, J. F., H. D’Cotta, and E. Saillant, 2009 Environmental effects on fish sex determination and differentiation. Sex Dev. 3: 118–135.

Bellott, D. W., H. Skeatsky, T. Pynitikova, E. R. Mardis, T. Graves, et al. 2010 Convergent evolution of chicken Z and human X chromosomes by expansion and gene acquisition. Nature 466: 612–616.

Bergero, R., and D. Charlesworth, 2009 The evolution of restricted recombination in sex chromosomes. Trends Ecol. Evol. 24: 94–102.

Birch, O. S., D. E. Casiano, C. A. Wassif, T. Cogliati, L. P. Zhao et al., 2000 The LIM homeobox gene *Lhx9* is essential for mouse gonad formation. Nature 403: 909–913.

Bull, J. J., 1981 *Evolution of Sex Determining Mechanisms*. Benjamin-Cummings, California.

Charlesworth, D., and J. E. Mank, 2010 The birds and the bees and the flowers and the trees: lessons from genetic mapping of sex determination in plants and animals. Genetics 186: 9–31.

Cnaani, A., and B. Levavi-Sivan, 2009 Sexual development in fish, practical applications for aquaculture. Sex Dev. 3: 164–175.

Cnaani, A., N. Zilberman, S. Tinman, G. Hulata, and M. Ron, 2004 Genome-scan analysis for quantitative trait loci in an F-2 tilapia hybrid. Mol. Genet. Genomics 272: 162–172.

Cnaani, A., B. Y. Lee, C. Ozouf-Costaz, C. Bonillo, F. Baroiller et al., 2007 Mapping of *Sox2* and *Sox14* in tilapia (*Oreochromis* spp.). Sex Dev. 1: 207–210.

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

Couderc, J. L., D. Godt, S. Zollman, J. Chen, M. Li et al., 2002 The birc a brac locus consists of two paralogous genes encoding RTB/POZ domain proteins and acts as a homeotic and morphogenetic regulator of imaginal development in Drosophila. Development 129: 2419–2433.
Ottolenghi, C., S. Omari, J. E. Garcia-Ortiz, M. Uda, L. Crisponi et al., 2005 Fine-mapping of a locus on linkage group 23 for sex determination in Nile tilapia (Oreochromis niloticus). Anim. Genet. 42: 222–224.

Ellegren, H., 2011 Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. Nat. Rev. Genet. 12: 157–166.

Lee, B. Y., G. Hulata, and T. D. Kocher, 2004 Two unlinked loci controlling line of tilapia (Oreochromis niloticus). Anim. Genet. 42: 222–224.

Fujimura, K., and T. D. Kocher, 2011 Tol2-mediated transgenesis in tilapia (Oreochromis niloticus). Aquaculture 319: 342–346.

Lee, B. Y., M. A. Conte, H. D’Cotta, E. Pepey, et al., 2010 Development of the embryo, larva and early juvenile of Nile tilapia Oreochromis niloticus (Pisces: Cichlidae). Developmental staging system. Dev. Growth Differ. 49: 301–324.

Iijiri, S., H. Kaneko, T. Kobayashi, D. S. Wang, F. Sakai et al., 2008 Sex dimorphic expression of genes in gonads during early differentiation of a teleost fish, the Nile tilapia Oreochromis niloticus. Biol. Reprod. 78: 333–341.

Lee, B. Y., C. H. Wang, M. L. Wu, F. M. Yang, et al., 2005 PIASy inhibits LRH-1-dependent CYP11A1 expression by competing for SRC-1 binding. Biochem. J. 419: 201–209.

Huang, D. W., B. T. Sherman, and R. A. Lempicki, 2009 Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4: 44–57.

Knott, S. A., J. M. Elsen, and C. S. Haley, 1996 Methods for multiple-marker mapping of quantitative trait loci in half-sib populations. Theor. Appl. Genet. 93: 71–80.

Ladd, A. N., N. Charlet-B, and T. A. Cooper, 2001 The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. Mol. Cell. Biol. 21: 1285–1296.

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

Lee, B. Y., G. Hulata, and T. D. Kocher, 2004 Two unlinked loci controlling the sex of blue tilapia (Oreochromis aureus). Heredity 92: 543–549.

Lee, B. Y., W. J. Lee, T. T. Streelman, K. L. Carleton, A. E. Howe et al., 2005 A second-generation genetic linkage map of tilapia (Oreochromis niloticus). Genetics 170: 237–244.

Lee, B. Y., A. E. Howe, M. A. Conte, H. D’Cotta, E. Pepey et al., 2010 An EST resource for tilapia based on 17 normalized libraries and assembly of 116,899 sequence tags. BMC Genomics 11: 278.

Mair, G. C., J. S. Abucay, D. O. F. Skibinski, T. A. Abella, and J. A. Beardmore, 1997 Genetic manipulation of sex ratio for the large-scale production of all-male tilapia, Oreochromis niloticus. Can. J. Fish. Aquat. Sci. 54: 396–404.

Mank, J. E., D. E. L. Promislow, and J. C. Avise, 2006 Evolution of alternative sex-determining mechanisms in teleost fishes. Biol. J. Linn. Soc. Lond. 87: 83–93.

Otake, H., A. Shinomiya, M. Matsuda, S. Hamaguchi, and M. Sakaizumi, 2006 Wild-derived XY sex-reversal mutants in the medaka, Oryzias latipes. Genetics 173: 2083–2090.

Ottolenghi, C., S. Omari, J. E. Garcia-Ortiz, M. Uda, L. Crisponi et al., 2005 Foxl2 is required for ovary differentiation. Hum. Mol. Genet. 14: 2053–2062.

Ozaki, Y., M. Higuchi, C. Miura, S. Yamaguchi, Y. Tozawa et al., 2006 Roles of 11 beta-hydroxysteroid dehydrogenase in fish spermatogenesis. Endocrinology 147: 5139–5146.

Palti, Y., A. Shirak, A. Cnaani, G. Hulata, R. R. Avtalion et al., 2002 Detection of genes with deleterious alleles in an inbred line of tilapia (Oreochromis aureus). Aquaculture 206: 151–164.

Poonlapdecha, S., E. Pepey, S. H. Huang, M. Canonne, et al., 2011 Elevated amh gene expression in the brain of male tilapia (Oreochromis niloticus) during testis differentiation. Sex Dev. 5: 33–47.

Ron, M., and J. I. Weller, 2007 From QTL to QTN identification in livestock—winning by points rather than knock-out: a review. Anim. Genet. 38: 429–439.

Rougeot, C., P. Prignon, C. V. N. Kengne, and C. Melard, 2008 Effect of high temperature during embryogenesis on the sex differentiation process in the Nile tilapia, Oreochromis niloticus. Aquaculture 276: 205–208.

Schartl, M., 2004 Sex chromosome evolution in non-mammalian vertebrates. Curr. Opin. Genet. Dev. 14: 634–641.

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

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

Spelman, R. J., W. Coppieters, L. Karim, J. A. M. vanArendonk, and H. Bovenhuis, 1996 Quantitative trait loci analysis for five milk production traits on chromosome six in the Dutch Holstein-Friesian population. Genetics 144: 1799–1807.

Su, A. L., T. Wiltshire, S. Batalov, H. Lapp, K. A. Ching et al., 2004 A gene atlas of the mouse and human protein-encoding transcriptional. Proc. Natl. Acad. Sci. USA 101: 6062–6067.

Sumoy, L., L. Carim, M. Escarceller, M. Nadal, M. Gratacos et al., 2000 HMG20A and HMG20B map to human chromosomes 15q24 and 19p13.3 and constitute a distinct class of HMG-box genes with ubiquitous expression. Cytogenet. Cell Genet. 88: 62–67.

Ventura, T., R. Manor, E. D. Atlalo, S. Weil, S. Raviv et al., 2009 Temporal silencing of an androgenic gland-specific insulin-like gene affecting phenotypical gender differences and spermatogenesis. Endocrinology 150: 1278–1286.

Vidal, V. P. I., M. C. Chaboissier, D. G. de Rooij, and A. Schedl, 2001 Sox9 induces testis development in XX transgenic mice. Nat. Genet. 28: 216–217.

Volf, J. N., and M. Schartl, 2002 Sex determination and sex chromosome evolution in the medaka, Oryzias latipes, and the platfish, Xiphophorus maculatus. Cytogenet. Genome Res. 99: 170–177.

Wilson, B. C., D. Burnett, R. Rappaport, L. J. Parry, and E. K. Fletcher, 2009 Relaxin-3 and RXFP3 expression, and steroidogenic actions in the ovary of teleost fish. Comp. Biochem. Physiol. A Mol. Integr. Physiol. 153: 69–74.

Wu, C., C. Orozco, J. Boyer, M. Leglise, J. Goodale et al., 2009 BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biol. 10: R130.

Zilberman, N., S. Reikhav, G. Hulata, and M. Ron, 2006 High-throughput genomic DNA extraction protocol from tilapia’s fin tissue. Aquaculture 255: 597–599.

Zhou, J. G., X. Feng, B. Ban, J. X. Liu, Z. Wang et al., 2009 Elongation factor ELL (eleven-nineteen lysine-rich leukemia) acts as a transcription factor for direct thrombospondin-1 regulation. J. Biol. Chem. 284: 19142–19152.

Zhu, X. Y., A. S. Gleiberman, and M. G. Rosenfeld, 2007 Molecular physiology of pituitary development: signaling and transcriptional networks. Physiol. Rev. 87: 933–963.