

Supplementary Information for

Genome editing reveals *dmrt1* as an essential male sex-determining gene in Chinese tongue sole (*Cynoglossus semilaevis*)

Zhongkai Cui^{1,2,3^}, Yun Liu^{4^}, Wenwen Wang^{1,2}, Qian Wang^{1,2}, Ning Zhang^{1,2}, Fan Lin^{1,2}, Na Wang^{1,2}, Changwei Shao^{1,2}, Zhongdian Dong^{1,2}, Yangzhen Li^{1,2}, Yingming Yang^{1,2}, Mengzhu Hu^{1,2}, Hailong Li^{1,2}, Fengtao Gao^{1,2}, Zhanfei Wei^{1,2}, Liang Meng^{1,2}, Yang Liu^{1,2}, Min Wei^{1,2}, Ying Zhu^{1,2}, Hua Guo^{1,2}, Christopher H.K. Cheng^{4*}, Manfred Schartl^{5*} & Songlin Chen^{1,2*}

1 Key Lab of Sustainable Development of Marine Fisheries, Ministry of Agriculture; Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, China,

2 Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266273, China,

3 College of Fisheries and Life, Shanghai Ocean University, Shanghai 201306, China,

4 School of Biomedical Sciences, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, China; School of Biomedical Sciences Core Laboratory, The Chinese University of Hong Kong Shenzhen Research Institute, Shenzhen 518057, China.

5 Physiologische Chemie, University of Würzburg, Biozentrum, and Comprehensive Cancer Center Mainfranken, University Clinic Würzburg, D-97074 Würzburg, Germany and Texas A&M Institute for Advanced Studies and Department of Biology, Texas A&M University, College Station, Texas 77843, USA

[^] The first two authors contributed equally to this paper.

* equally contributing senior authors and corresponding authors: S.C. (chensl@ysfri.ac.cn), M.S. (phchl@biozentrum.uni-wuerzburg.de), and C.H.K.C (chkcheng@cuhk.edu.hk)

Suppl 1

Microinjection was done in three steps: First, because *C. semilaevis* embryos are floating, more than half of the medium was aspirated from the embryo holding device. Leaving too much fluid makes the embryos falling in the trough. Second, the embryos were rotated to a position where the cell can be injected by the microinjection needle perpendicularly. Third, the microinjection needle was placed close to a glass slide. The tip of the needle gently touched the rough surface of the glass slide, until a small amount of *dmrt1*-TALENs mRNA flew out because of pressure from the injector device.

Suppl 2

The PCR reaction contained 2 μ l gDNA (~50ng/ μ l) as template, 5 μ l 10x reaction buffer (with MgCl²⁺), 4 μ l dNTPs, 0.5 μ l Takara Ex Taq DNA polymerase (Takara, Japan), 1.5 μ l of each primer (at 10 μ M) and 35.5 μ l ddH₂O to a total volume of 50 μ l. PCR conditions were one cycle at 95°C for 5 min, followed by 35 cycles at 95°C for 30 sec, 60°C for 30 sec, and 72°C for 30 sec, and a final elongation step at 72°C for 10 min. The PCR product was purified using Gel Extraction Kit (CW BIO, China). 5 μ l of purified PCR product added to 1.1 μ l 1x NEBuffer and 4.4 μ l ddH₂O were heated to 95°C for 5min and then cooled to room temperature. Then *T7E* I (NEB, USA) was added and incubated at 37°C for 30 min. The reaction

mixture was separated by agarose gel electrophoresis and detected by Gel Imaging System (Bio-Rad, USA). The purified PCR product was cloned into the PMD-18T vector (Takara, Japan) and transformed into Top10 competent cells. Individual bacterial colonies were randomly selected and screened by PCR in a total volume of 25 μ l as described above. Positive clones were sequenced and mutated sites compared to wild-type sequences. The mutation rate of each sample was calculated from the result of sequencing (mutated sequences/total sequences).

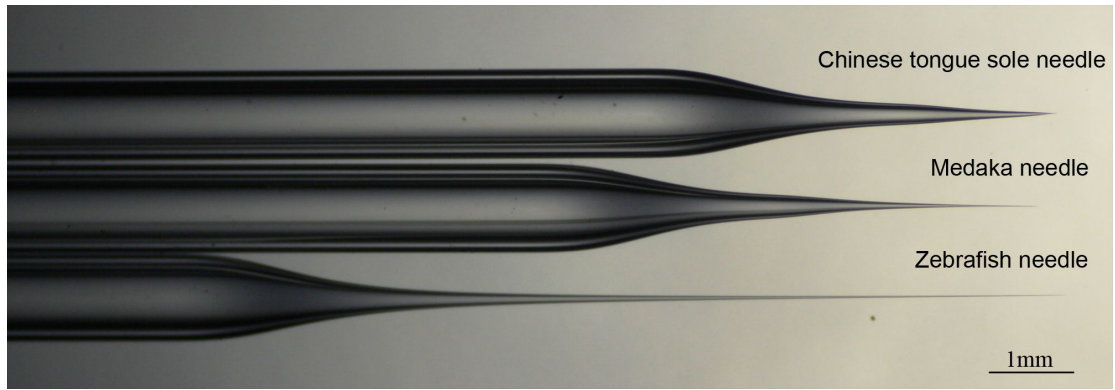


Figure S1 Comparison of *C. semilaevis*, medaka and zebrafish needles. The tip of a needle for embryo microinjection of *C. semilaevis* is shorter and thick pointed, to allow penetration of the tough chorion for injecting into the egg.

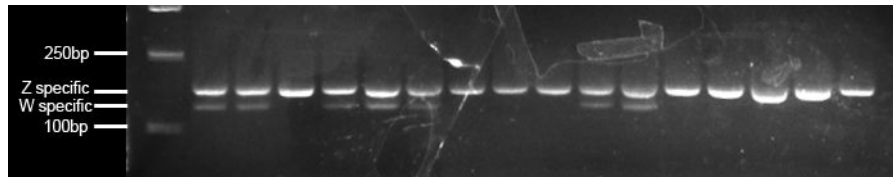


Figure S2 Determination of the genetic sex of the *dmrt1* mutated fish (partial). Different sized products for the Z (169 bp) and W (134 bp) chromosomes obtained by SSR PCR.

1# 1R2R
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGTCT CGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA A-----GGGCCA CAAACGCTACT -24 1/5

2# 2R3R
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG T-----GTCTCCGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTC-----GTTGAA GGGC CACAAACGCTACT -8 2/5
 CCC-----TCCAGT CA CAGTA GC-----GTT-----GTTGAA GGGC CACAAACGCTACT (-38, +19) 1/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTCC GTTGA AGGA AGGGC TTGAA GGGC CACAAACGCTACT (-1, +5) 1/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTGG TAA-----GGCTCCGTTGAA GGGC CACAAACGCTACT (-1, +5) 1/5

4# 2RSR
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGTCT CGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCT-----CCGTTGA AGGGC CACAAACGCTACT -8 2/5

5# 3NR
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGTCT CGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTC-----CTCCGTTGAA GGGC CACAAACGCTACT -4 1/5

7# 4RSR
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGTCT CGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTCG-----TTGA AGGGC CACAAACGCTACT -8 3/5
 -----TTGA AGGGC CACAAACGCTACT -42 1/5

9# 2O3O
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGTCT CGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTCT-----CCGTTGAA GGGC CACAAACGCTACT (-6, +2) 1/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTCC-----GTTGA AGGGC CACAAACGCTACT -8 1/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTC-----GTTGA AGGGC CACAAACGCTACT (-10, +2) 2/5

10# 2O4O
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGTCT CGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTCTCTGCTCGTT GAA GGGC CACAAACGCTACT (-2, +2) 1/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCT-----TCGTTGAA GGGC CACAAACGCTACT (-9, +1) 1/5
 CCCAAGTGC TCCGCTGCA G-----CCGTGGA AGGGC CACAAACGCTACT (-24, +5) 1/5

12# 3O4O
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TG-----TCTCCGTTGA AGGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTCG TG GGGTCTCC GTTGA AGGGC CACAAACGCTACT +3 1/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTC-----CGTTGA AGGGC CACAAACGCTACT -7 2/5

16# 2RSO
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGTCT CGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTC-----CTCCGTTGAA GGGC CACAAACGCTACT (-7, +1) 1/5

18# 3RSO
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGTCT CGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCT-----CCGTTGAA GGGC CACAAACGCTACT -8 1/5

19# 1Y2Y
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGTCT CGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTC-----TCCGTTGAA GGGC CACAAACGCTACT -7 1/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTC-----TCCGTTGAA GGGC CACAAACGCTACT -8 3/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTC-----TTGTTGAA GGGC CACAAACGCTACT (-9, +1) 1/5

20# 1Y3Y
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG T-----GTCTCCGTTGA AGGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTCC-----GTTGA AGGGC CACAAACGCTACT (-8, +1) 1/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTC-----TTGA AGGGC CACAAACGCTACT -12 1/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTGG TTTGGCC TCCGTTGA AGGGC CACAAACGCTACT +4 1/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTCG GGTGAA CCC GTTGA AGGGC CACAAACGCTACT +7 1/5

21# 1Y4Y
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGTCT CGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTCG T-----TGAAGGGC CACAAACGCTACT -8 1/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTC-----CCGTTGA AGGGC CACAAACGCTACT -8 1/5

22# 1Y5Y
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGTCT CGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTC-----CCGTTGA AGGGC CACAAACGCTACT -8 1/5

23# 2Y3Y
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGT-----CTCCGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTCG TGTCTCGTTGAA GGGC CACAAACGCTACT +2 1/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTTCG T-----TGAAGGGC CACAAACGCTACT -8 1/5

24# 2Y4Y
 CCCAAGTGC TCCGCTGCA GGA ACCA GGCTTCG TGTCT CGTT TAA GGGC CACAAACGCTACT WT
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTC-----TCGTTGAA GGGC CACAAACGCTACT (-9, +1) 2/5
 CCCAAGTGC TCCGCTGCA GGA ACCA CGGCTC-----CCGTTGAA GGGC CACAAACGCTACT -8 2/5

Figure S3 Wild-type sequence and *dmrt1* mutations of 16 injected individuals (partial). Mutation frequency ranging from 20% to 100%.

A

```

CCCAAGTGCTCCCGCTGCAGGAACCAAGGCTTCGTGTCTCCGTTGAAGGGCCACAAACGCTACT WT
CCCAAGTGCTCCCGCTGCAGGAACCACGGCTTCG-----CTCCGTTGAAGGGCCACAAACGCTACT -3 1/12
CCCAAGTGCTCCCGCTGCAGGAACCACGGCT-----CCGTTGAAGGGCCACAAACGCTACT -8 7/12
-----TCCGTTGAAGGGCCACAAACGCTACT -97 2/12

```

B

```

CCCAAGTGCTCCCGCTGCAGGAACCAAGGCTTCGTGTCTCCGTTGAAGGGCCACAAACGCTACT WT
CCCAAGTGCTCCCGCTGCAGGAACCACGGCT-----TCGTTGAAGGGCCACAAACGCTACT (-9, +1) 2/5
CCCAAGTGCTCCCGCTGCAGGAACCACGGCT-----CCGTTGAAGGGCCACAAACGCTACT -8 2/5

```

Figure S4 Wild-type sequence and *dmrt1* mutations in the ovary (A) and fins (B) of the same fish. The frequency of *dmrt1* mutations was about 80% in this individual.

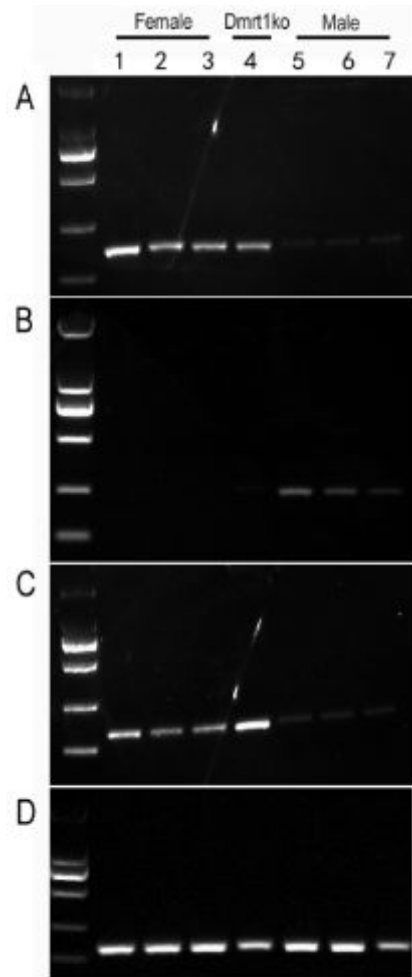


Figure S5 Expression of *dmrt1*, *foxl2* and *cyp19a1a* in *dmrt1*-deficient gonads of one year old fish determined by semiquantitative RT-PCR. *β-actin* was used as control and for calibration of RNA amounts. **A, mRNA expression of *cyp19a1a* in *dmrt1*-deficient testis is comparable to wt females, while only amplification barely above background is recorded in wild-type males. **B**, RT-PCR for *dmrt1* with RNA from *dmrt1*-deficient testis revealed considerably less amplification compared with wild-type males. **C**, mRNA expression of *foxl2* in *dmrt1*-deficient testis appears to be even higher in comparison with wild-type females. **D**, mRNA expression of *β-actin* in *dmrt1*-deficient tongue sole, wild-type females and wild-type males.**

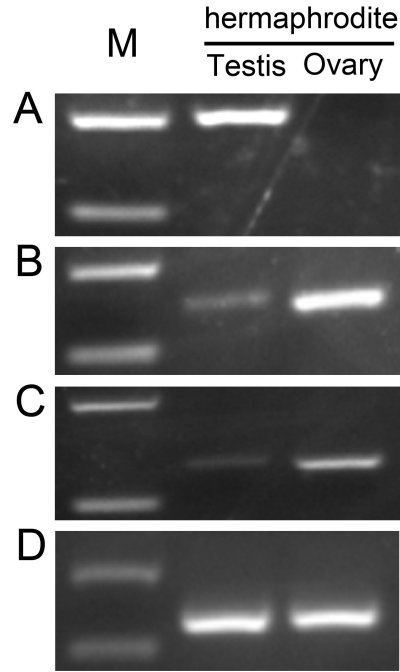


Figure S7 Expression of *dmrt1*, *foxl2* and *cyp19a1a* in gonads of the hermaphroditic fish. β -actin was used as control and for calibration of RNA amounts. **A**, mRNA expression of *dmrt1* in intersexual testis is significantly higher than that in intersexual ovary. **B**, RT-PCR for *cyp19a1a* with RNA from intersexual testis revealed less amplification compared with intersexual ovary. **C**, mRNA expression of *foxl2* in intersexual ovary appears to be higher in comparison with intersexual testis. **D**, mRNA expression of β -actin in gonads of intersexual tongue sole.

Table S1. Primers used in this study.

Primer	Sequence (5'-3')	Purpose
<i>dmrt1</i> -TALEN-F	CGGGCAAAGGGAGAAGG	<i>dmrt1</i> amplification
<i>dmrt1</i> -TALEN-R	AAAAACATCTCCTGAGGGCTAA	
cs-sex-F	CCTAAATGATGGATGTAGATTCTGTC	determination of genetic sex
cs-sex-R	GATCCAGAGAAAATAAACCCAGG	
<i>dmrt1</i> -RT-F	CCGGACGGCTTCGTGTC	real-time PCR
<i>dmrt1</i> -RT-R	CTCCACAGGGAGCAGGCAGT	
<i>foxl2</i> -RT-F	GCACCCAATCCGTTACAGC	real-time PCR
<i>foxl2</i> -RT-R	GCCATTCGTCACCCTCT	
<i>cyp19a1a</i> -RT-F	CTCTGTTCCCTCAGGTTTCTCTC	real-time PCR
<i>cyp19a1a</i> -RT-R	GATGTGACCCAGTGTGTGTTG	
<i>Sox9a</i> -RT-F	CAGGCAGGTAATGTTGGGGT	Real-time PCR
<i>Sox9a</i> -RT-R	AAGGAGCCGTAGGTGATGTG	
<i>Amh</i> -RT-F	ATGTGGCTGTAGAAGAAGAC	Real-time PCR
<i>Amh</i> -RT-R	CTGTTAGCAGGATGTATCG	

Table S2. Growth traits of *dmrt1* mutant and wild-type *C. semilaevis*.

type	2014 (1 year old)			2015 (8 months old)		
	Weight (g)	Body length (cm)	Body width (cm)	Weight (g)	Body length (cm)	Body width (cm)
WTF 1	335.7	40.3	11.3	134.1	27.4	8.4
WTF 2	259.7	32.9	9.0	104.5	25.5	7.5
WTF 3	295.1	34.5	9.7	105.2	25.8	7.5
DDM 1	250.5	34.0	9.5	120.0	26.7	7.7
DDM 2	202.4	32.6	9.0	116.2	26.5	7.7
DDM 3	270.7	34.4	10.5	102.1	26.0	7.5
WTM 1	113.2	25.4	7.4	53.2	21.0	6.0
WTM 2	100.6	25.6	6.7	61.8	22.3	6.4
WTM 3	117.0	27.3	7.0	52.6	21.0	6.0

WTF wild-type female *C. semilaevis*, DDM *dmrt1* mutant male *C. semilaevis*, WTM wild-type male *C. semilaevis*,