

XII International Symposium on Genetics of Aquaculture

SEX CHROMOSOME EVOLUTION AND MECHANISM FOR SEX DETERMINATION AND REVERSAL REVEALED BY WHOLE-GENOME SEQUENCING AND METHYLATION SEQUENCING IN HALF-SMOOTH TONGUE SOLE *Cynoglossus semilaevis*

Song-Lin Chen

**Yellow Sea Fisheries Research Institute, CAFS
266071 Qingdao, China
chensl@ysfri.ac.cn**

Why do we chose half-smooth tongue sole for genome sequencing and genetics studies?

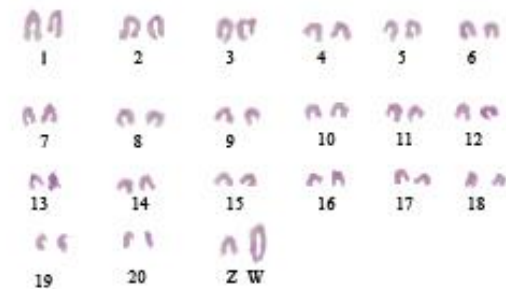
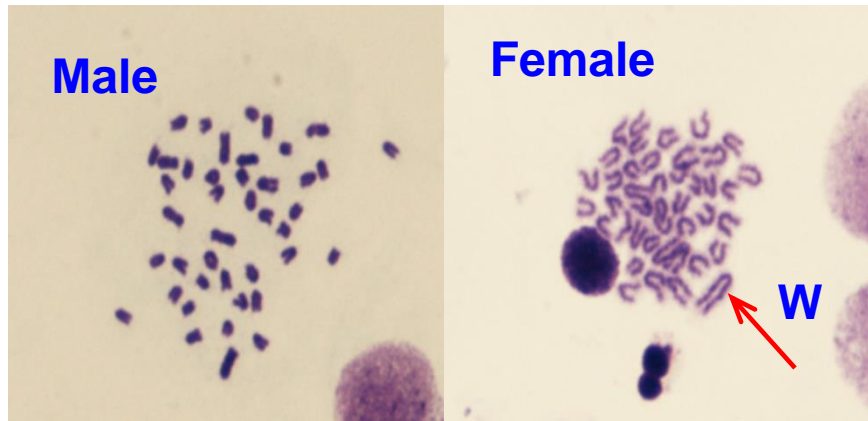
- Delicious sea food
- Economically important marine flatfish species in China
- Big difference in growth rate between female and male
- ZW sex determination with a big W chromosome



Female



Male



Not only an economically important marine fish, also a promising model fish for sex chromosome evolution and sex determination mechanism.

Contents

- ◆ **Whole genome assembly and analysis of the tongue sole**
- ◆ **Construction of high density genetic linkage map**
- ◆ **Structure and evolution of sex chromosomes**
- ◆ **Identification and characterization of sex determination gene Dmrt1**
- ◆ **Epigenetic inheritance of sex reversal**

Whole genome assembly and analysis

Tongue sole genome assembly

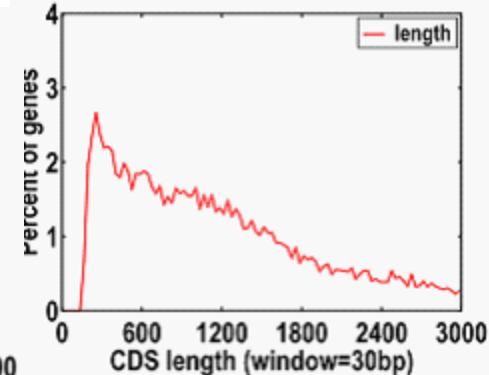
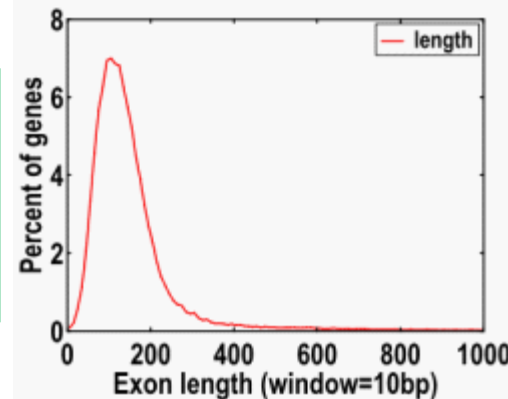
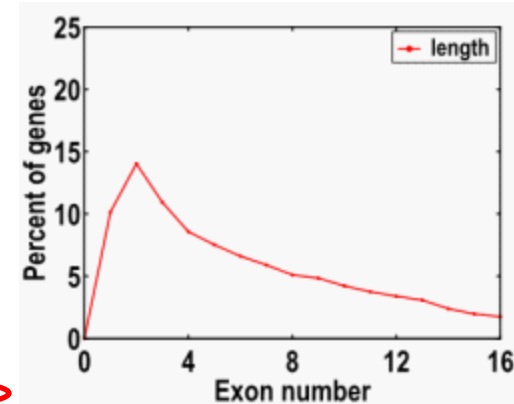
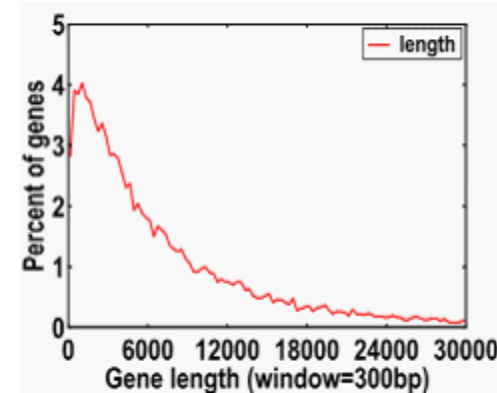
Type	Contig		scaffold	
	Size (bp)	Num.	Size (bp)	Num.
N90	4,319	20,363	228,546	568
N80	9,336	13,399	444,474	424
N70	14,345	9,503	585,702	332
N60	19,908	6,829	687,861	257
N50	26,123	4,844	867,445	194
N40	33,524	3,314	960,073	140
N30	42,762	2,117	1,078,682	93
N20	54,982	1,182	1,357,254	53
N10	75,253	467	1,703,876	21
Longest	194,815	1	4,694,140	1
Total	452,397,884	114,070	477,293,128	80,788

We successfully assembled the genome of tongue sole. The scaffold N50 is 867kb.

Gene prediction

Gene set		Number	Average gene length (bp)	Average CDS length (bp)	Average exons per gene	Average exon length (bp)	Average intron length (bp)
Denovo	glimmerhmm	43,309	8,502	1,001	6	158	1,408
	snap	40,096	12,898	1,139	8	134	1,571
	augustus	18,175	10,461	1,658	10	171	1,013
Homolog	Danio	16,319	8,519	1,541	10	161	812
	Gasterosteus	17,637	8,486	1,560	10	160	789
	Human	14,521	9,502	1,637	10	159	846
	Mouse	14,343	9,389	1,628	10	158	832
	Oryzias	17,530	8,043	1,514	9	161	775
	Takifugu	16,180	9,159	1,675	10	160	793
	Tetraodon	16,335	8,770	1,598	10	158	786
EST		9,489	2,287	502	4	141	859
Glean		21,516	7,950	1,535	9	171	804

We annotated 21516 protein coding genes with an average gene length of 7950bp.



Gene annotation

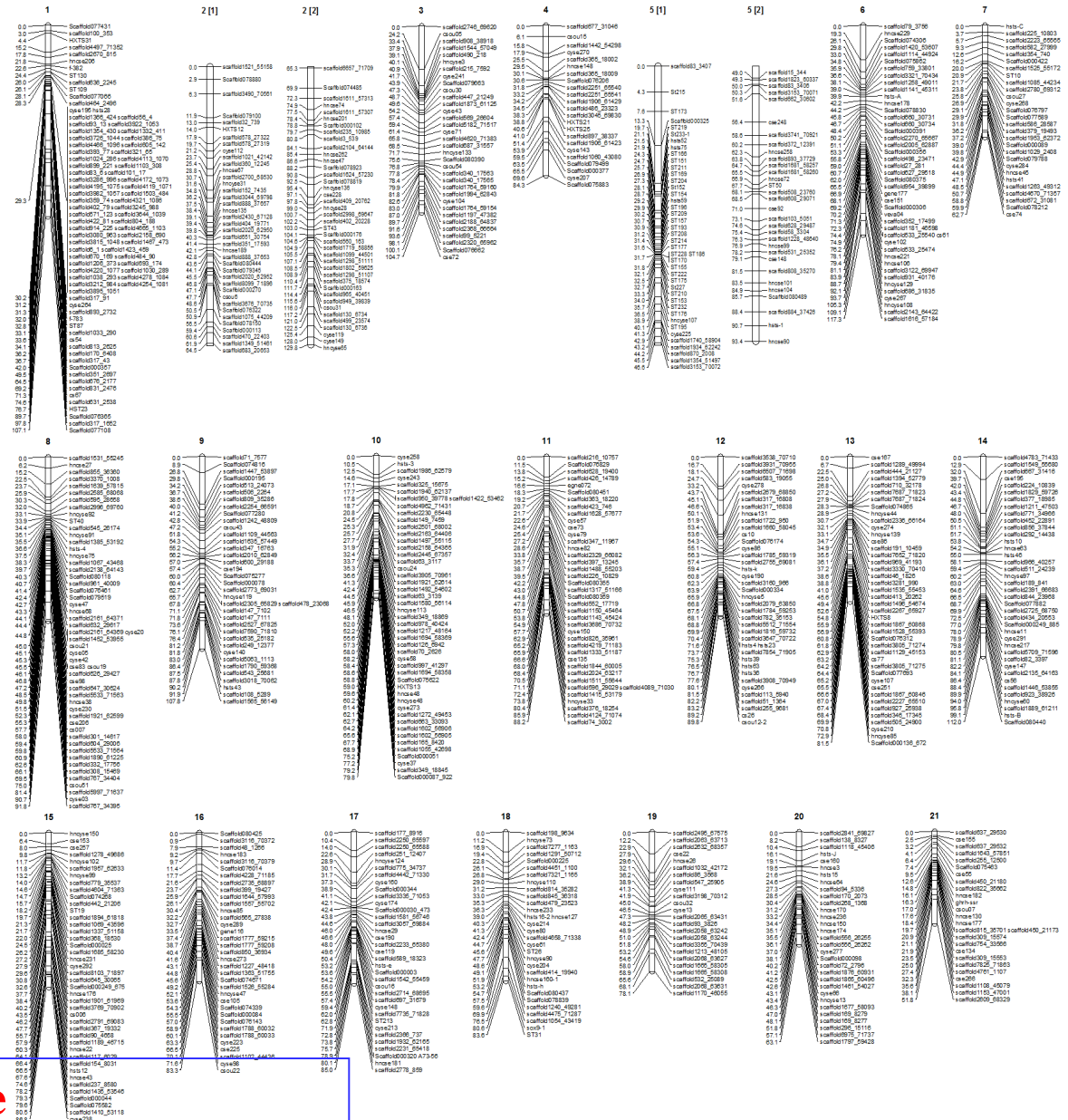
	Number	Percent(%)
Total	21,516	100.00
Annotated	21,309	99%
Swissprot	17,216	82.63
TrEMBL	17,895	85.89
InterPro	15,150	72.72
KEGG	13,237	63.54
GO	11,337	54.42
Unannotated	207	15.02

Among the 21,516 genes, 21,309 genes have known homologs, and only 207 genes have not homologs.

Construction of high density SSR and SNP genetic linkage map

Construction of SSR female genetic map

The female high density linkage map was consisted of 828 SSR markers, and was grouped in 21 linkage groups with an average resolution of 1.8 cM.



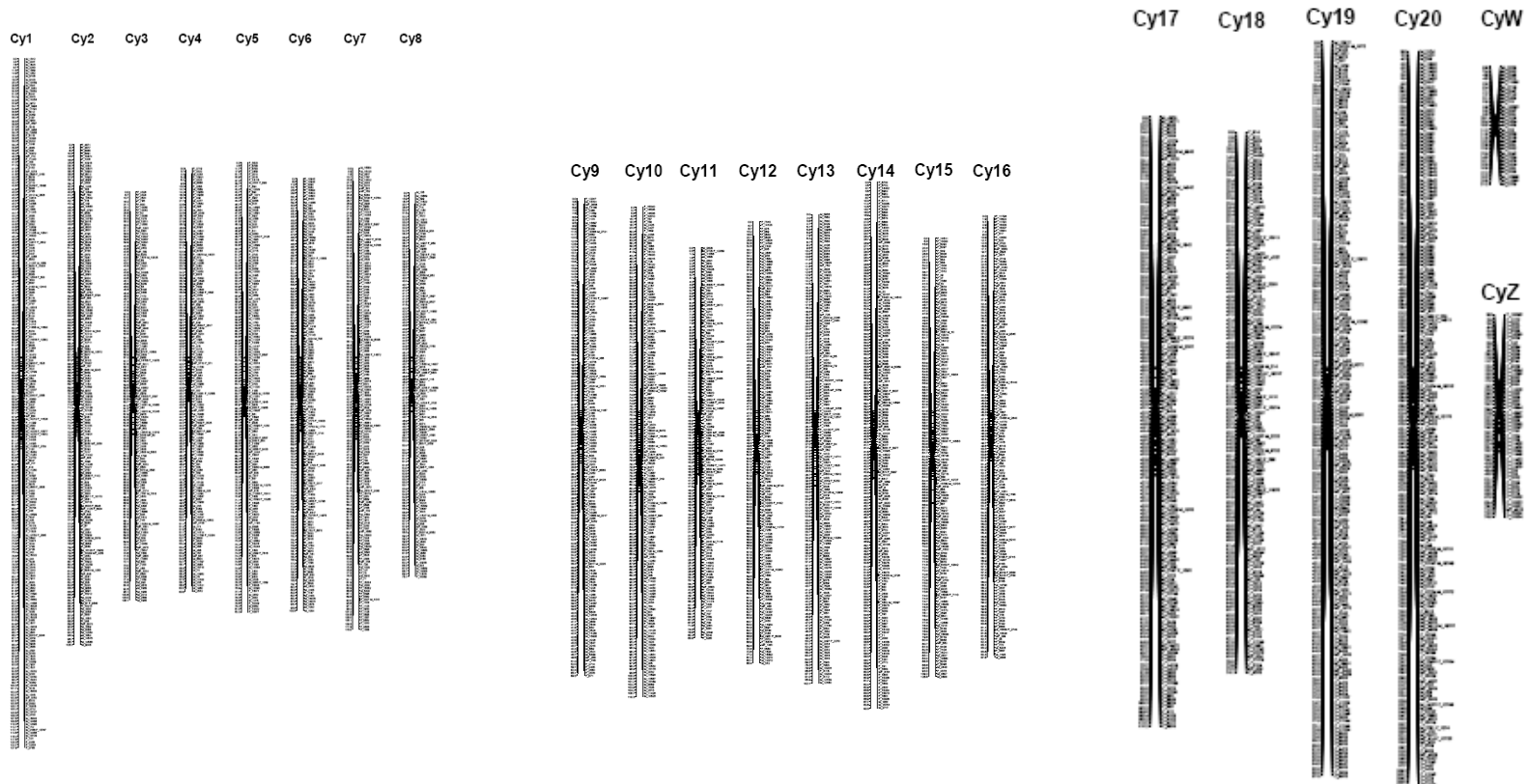
Number of SNP identified by RAD sequencing of individuals of mapping panel

Polymorphic in	No. of SNP	No. of anchored scaffolds	Length of anchored scaffolds (Mb)	Genome reference (scaffolds) coverage (%)
Female	5,237	805	318.3	66.7
Male	5,966	909	367.2	76.9
Both parents	939	487	246.3	51.6
Total	12142	1104	404.8	84.8

We have done RAD sequencing for a family of tongue sole (including female and male parents, and their 216 offspring individuals). 12,142 heterozygous SNP loci were obtained, of which 5,966 loci for male, 5,237 loci for female and 939 loci for both.

SNP high density genetic linkage map

In total, the 12,142 SNPs were assigned to 22 linkage groups. The total linkage distance reached 1900cM with an average interval of 0.33 cM.

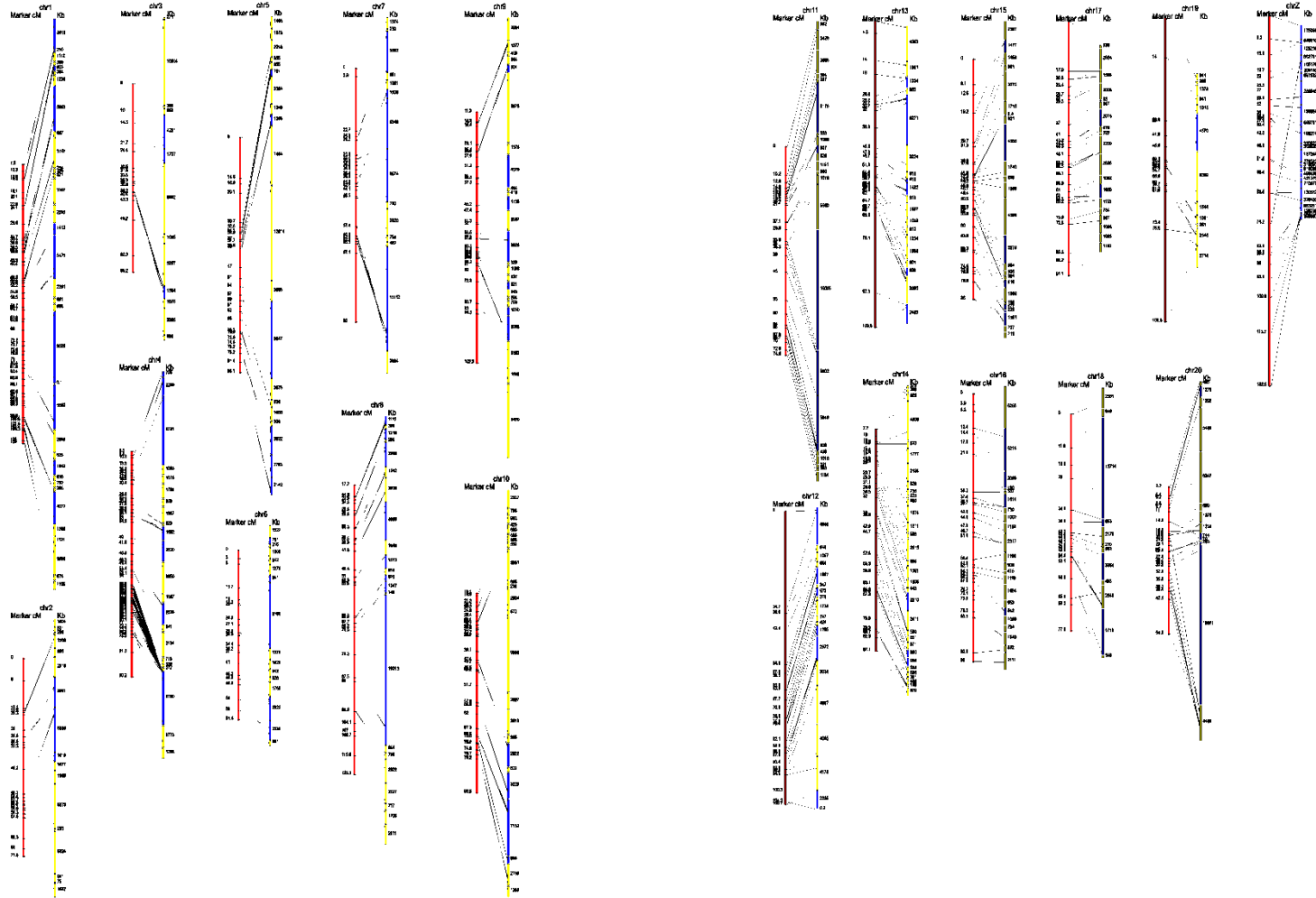


Creation of chromosome

Chr.	# SSR	# RAD-tag	Contig		Scaffold			# Genes
			#	Len.(bp)	#	Len.(bp)	Source	
1	81	1,184	2,410	32,791,084	53	34,529,112	Female	1,487
2	40	810	1,227	19,259,417	29	20,052,734	Female	911
3	29	484	1,189	15,467,848	25	16,253,993	Female	596
4	85	323	1,263	19,377,156	31	20,014,501	Female	846
5	43	89	1,147	18,609,661	29	19,279,693	Female	706
6	30	825	1,270	18,113,957	29	18,841,016	Female	978
7	54	54	993	13,185,383	15	13,814,722	Female	613
8	53	642	2,144	28,615,567	37	30,153,790	Female	1,395
9	50	454	1,314	18,790,677	31	19,618,599	Female	1,029
10	46	777	1,507	20,081,642	33	21,015,569	Female	1,037
11	42	949	1,428	19,676,390	34	20,528,432	Female	1,022
12	40	517	1,349	17,485,432	35	18,398,590	Female	745
13	43	865	1,518	20,959,882	34	21,922,143	Female	892
14	50	1,288	1,782	27,668,722	47	28,847,931	Female	1,228
15	46	703	1,478	19,132,837	32	20,094,621	Female	761
16	40	430	1,252	17,874,443	29	18,785,820	Female	809
17	38	246	1,333	15,583,495	25	16,472,647	Female	984
18	28	639	1,092	14,404,870	22	15,207,555	Female	783
19	33	553	1,108	17,115,378	24	17,747,288	Female	847
20	34	226	1,036	14,355,002	18	15,234,830	Female	881
Z	37	53	2,044	20,757,346	26	21,915,962	Male	930
W	34	NA	2,436	13,020,023	306	16,461,726	Female	320
Total	942	12,111	32,320	422,326,212	944	445,191,274	NA	19,800

944 scaffolds representing 445 Mb genomic sequences were anchored to 20 autosomes, one Z and one W chromosome.

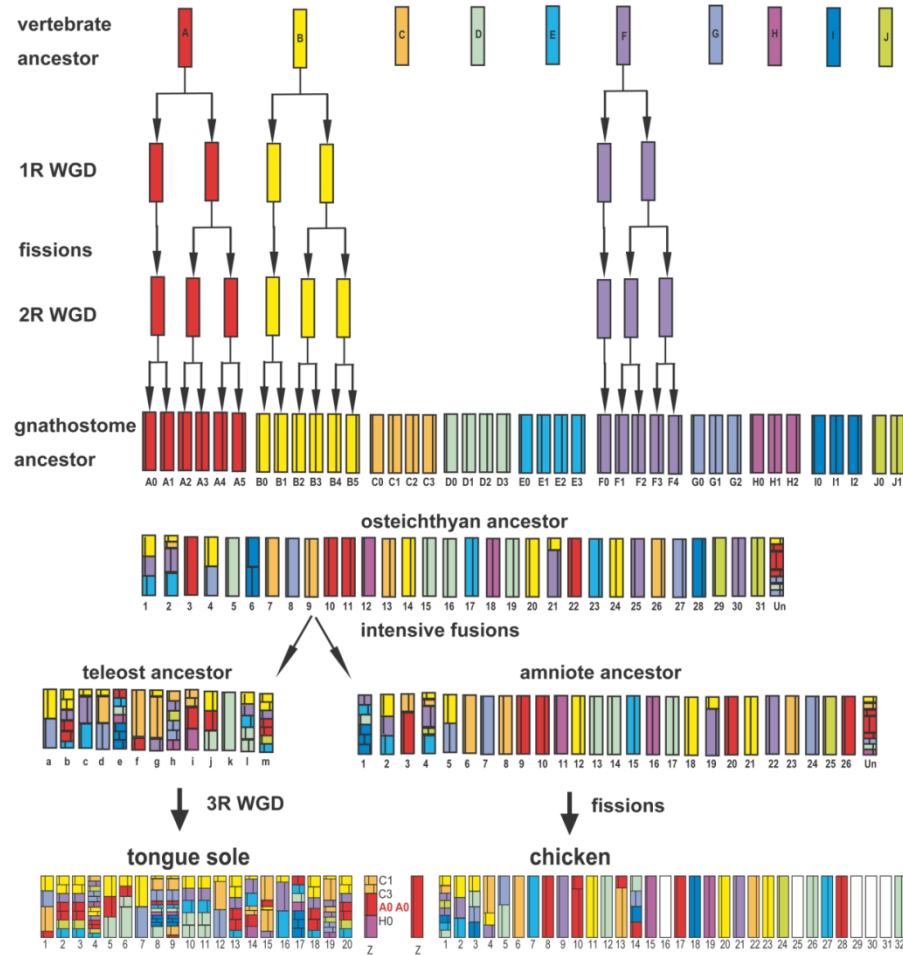
Construction of chromosome genome map



Left: Linkage group; Right: Chromosome

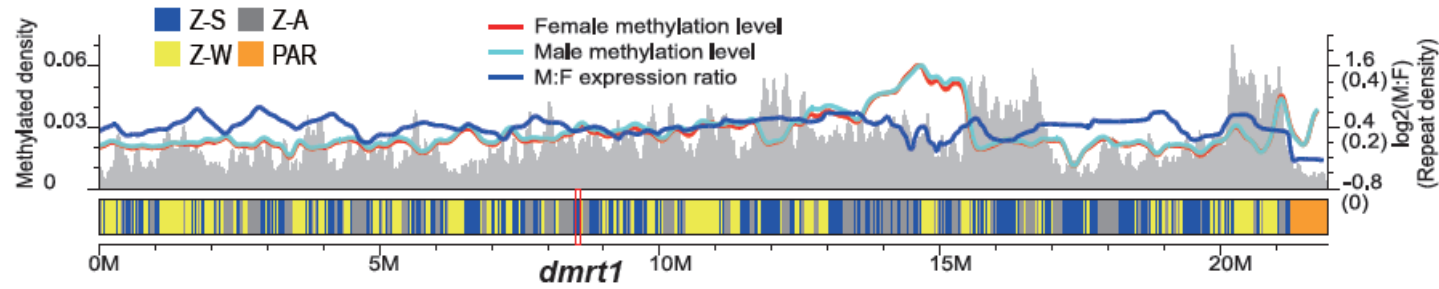
Structure and evolution of sex chromosomes

Origin and evolution of sex chromosomes



Using different vertebrate genomes from human, chicken, fugu, medaka and zebrafish, we induced ancestral karyotype, and found that the sex chromosomes of tongue sole were derived from the same ancestral vertebrate protochromosome as chicken Z and W chromosomes.

Structure of sex chromosomes



Type	Z (non-PAR)			W (non-PAR)		
	Functional genes	Pseudo genes	Total	Functional genes	Pseudo genes	Total
Z-W	286	11	297	272	67	339
Z-A	248	10	258	NA*	NA	NA
Z-S	370	12	382	NA	NA	NA
W-Z_random	NA	NA	NA	17	7	24
W-A	NA	NA	NA	26	4	30
W-S	NA	NA	NA	2	0	2
Total	904	33	937	317	78	395

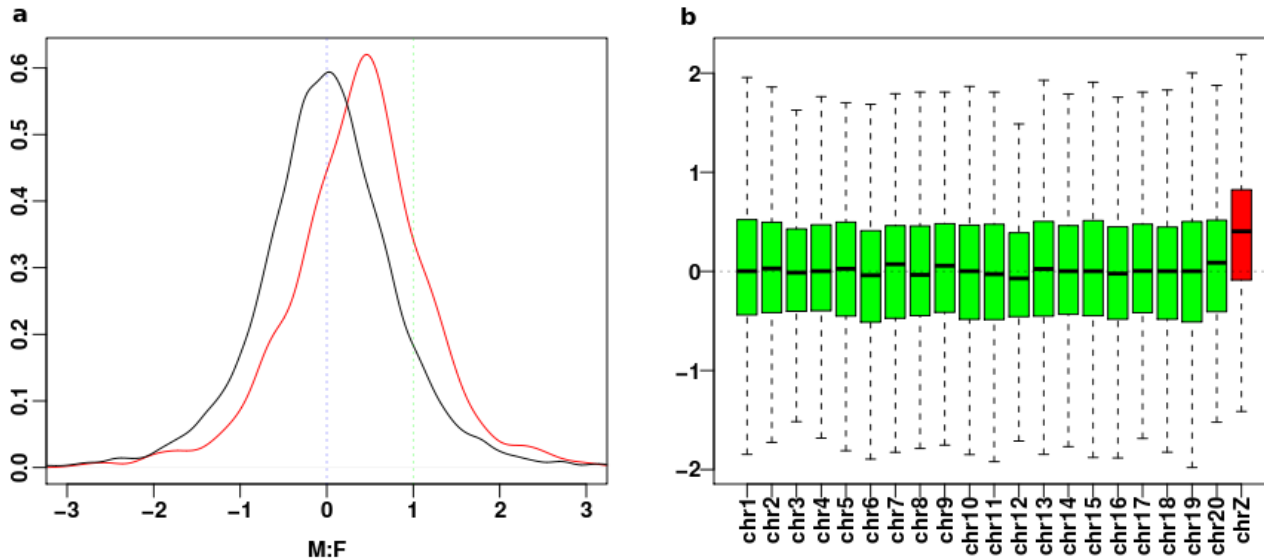
- Z chromosome contains 937 genes, of which 382 genes are Z-specific.
- W chromosome contains 395 genes, of which 339 genes are homologous to 297 Z genes, and two genes are W-specific.

Structure of sex chromosomes

Chr.	Genes per megabase	Total Tes	DNA Tes	LINE	LTR	SINE	Unclassified Tes	Average gene size (bp)
chr1	43	5.11%	2.41%	0.69%	0.01%	0.32%	1.68%	8,930
chr2	45	4.39%	2.15%	0.45%	0.01%	0.23%	1.55%	9,101
chr3	37	4.97%	2.15%	0.83%	0.06%	0.17%	1.76%	10,650
chr4	44	4.82%	2.17%	0.51%	0.02%	0.20%	1.92%	9,351
chr5	37	3.35%	1.55%	0.33%	0.01%	0.21%	1.25%	10,264
chr6	52	4.10%	1.73%	0.31%	0.01%	0.18%	1.87%	8,039
chr7	47	4.31%	1.90%	0.62%	0.04%	0.12%	1.63%	8,556
chr8	49	5.05%	2.42%	0.68%	0.05%	0.22%	1.68%	9,428
chr9	52	4.65%	2.16%	0.55%	0.02%	0.24%	1.68%	7,833
chr10	49	4.72%	2.34%	0.57%	0.03%	0.24%	1.54%	8,500
chr11	51	3.42%	1.42%	0.40%	0.01%	0.22%	1.37%	8,612
chr12	40	3.46%	1.70%	0.29%	0.01%	0.14%	1.32%	10,820
chr13	43	4.30%	2.15%	0.52%	0.03%	0.20%	1.40%	9,080
chr14	43	4.49%	1.93%	0.64%	0.02%	0.28%	1.62%	9,550
chr15	39	5.08%	2.04%	0.58%	0.08%	0.20%	2.18%	8,594
chr16	43	4.99%	2.42%	0.53%	0.02%	0.19%	1.83%	9,839
chr17	60	4.56%	2.23%	0.50%	0.01%	0.17%	1.65%	7,061
chr18	51	3.55%	1.62%	0.27%	0.01%	0.13%	1.52%	7,630
chr19	48	3.05%	1.32%	0.33%	0.03%	0.18%	1.19%	8,782
chr20	58	2.32%	0.91%	0.21%	0.01%	0.08%	1.11%	7,733
Autosomes	46	4.33%	1.99%	0.51%	0.02%	0.21%	1.60%	8,876
chrZ	42	13.13%	4.74%	3.95%	0.23%	0.43%	3.78%	9,857
chrW	19	29.94%	8.74%	9.39%	1.09%	0.46%	10.26%	12,156

- The gene density on CseZ is 42 genes/Mb, slightly lower than the average value of autosomes (46 genes/Mb).
- The CseW has a gene density of 19 genes/Mb, much lower than that of the autosomes.

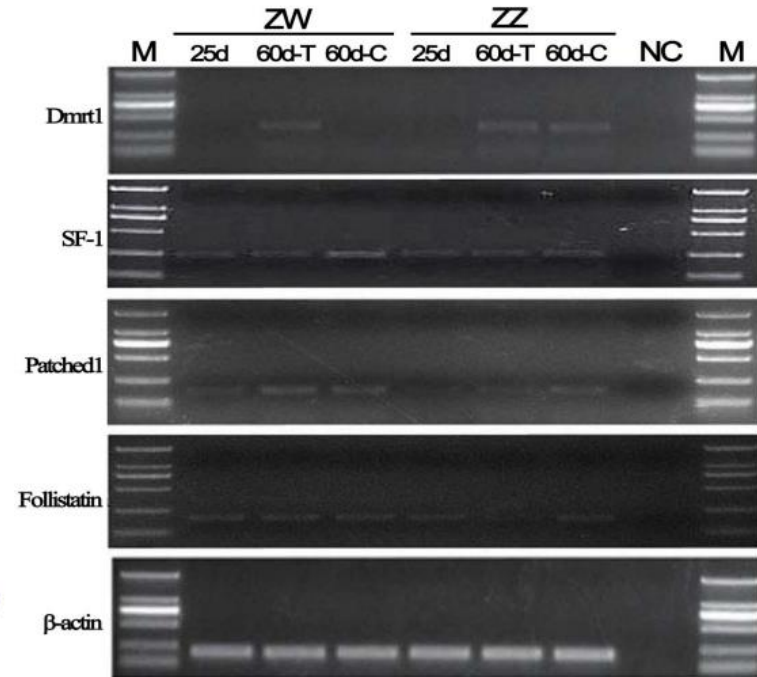
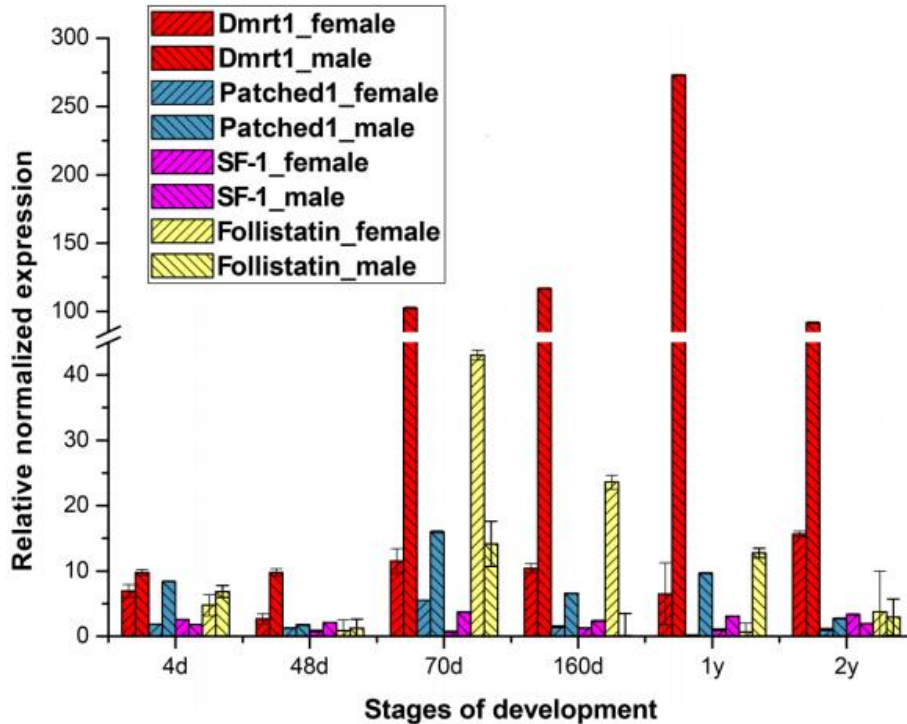
Dosage compensation of the Z chromosome in tongue sole



For those 763 Z-linked genes, male expression in whole-body (without gonad) transcriptomes was on average 1.32 times higher than female expression, indicating incomplete gene dosage compensation.

Identification and characterization of sex determination gene Dmrt1

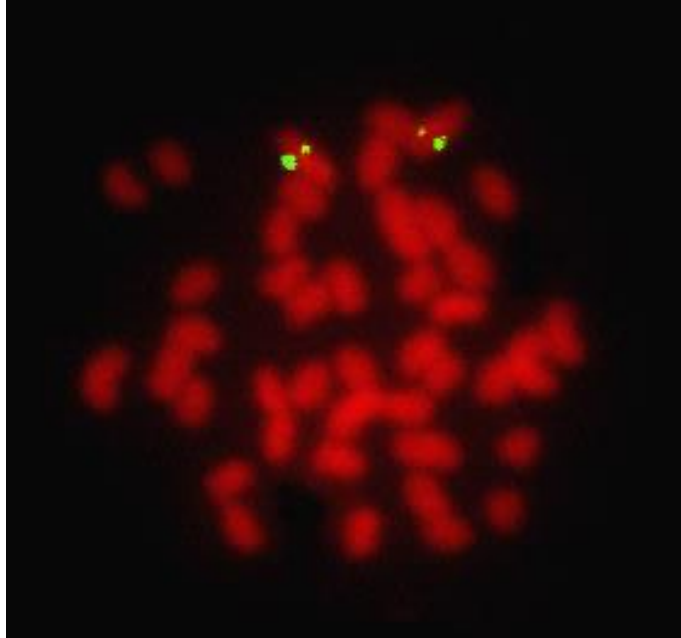
Characterization of sex related genes on Z chromosome



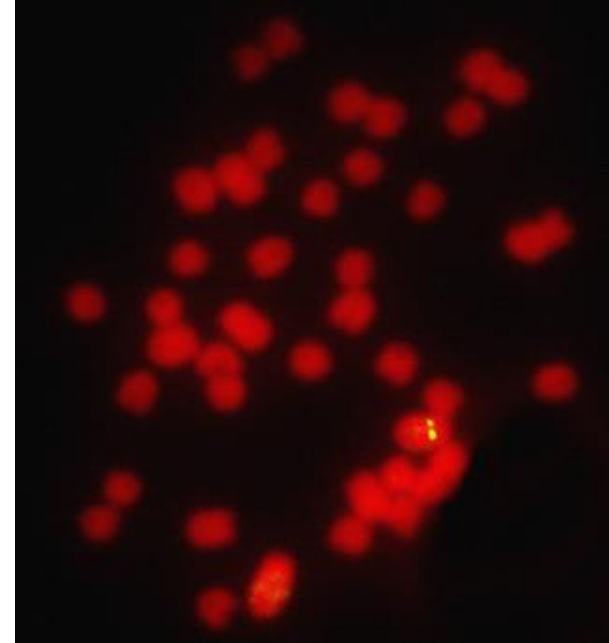
We examined the expression pattern of four sex related genes including *dmrt1*, *patched1*, *sf1* and *follistatin* during sex determination stages, and found only *dmrt1* gene expressed specifically in the gonads of ZZ male and ZW pseudomale, making it an outstanding candidate for a master sex-determining gene.

dmr1 FISH hybridization

Male (ZZ)



Female (ZW)

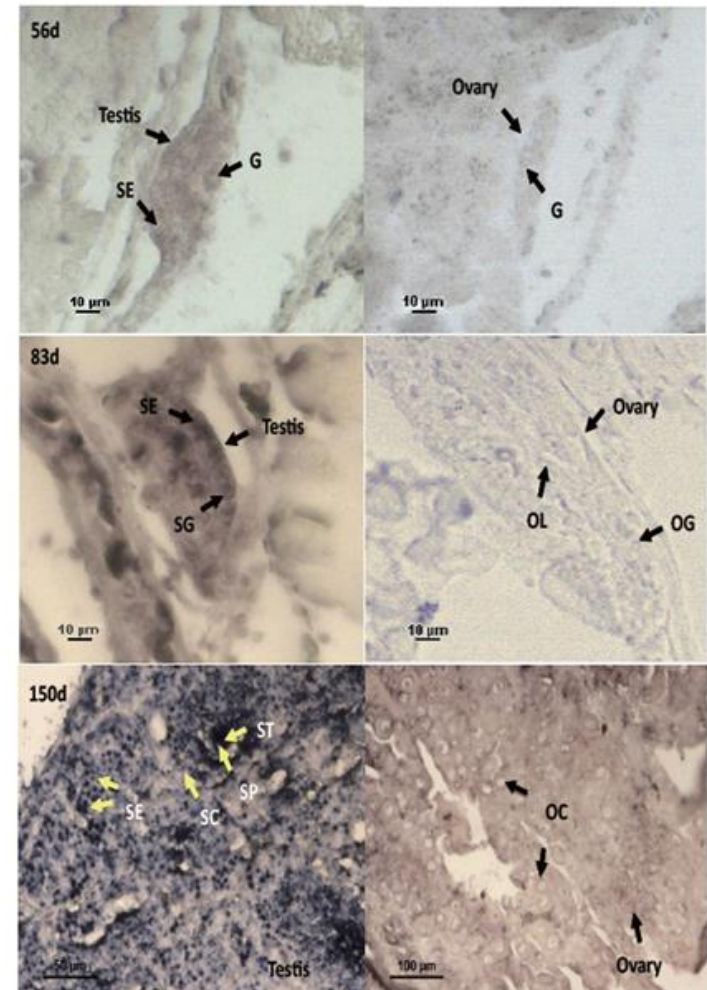
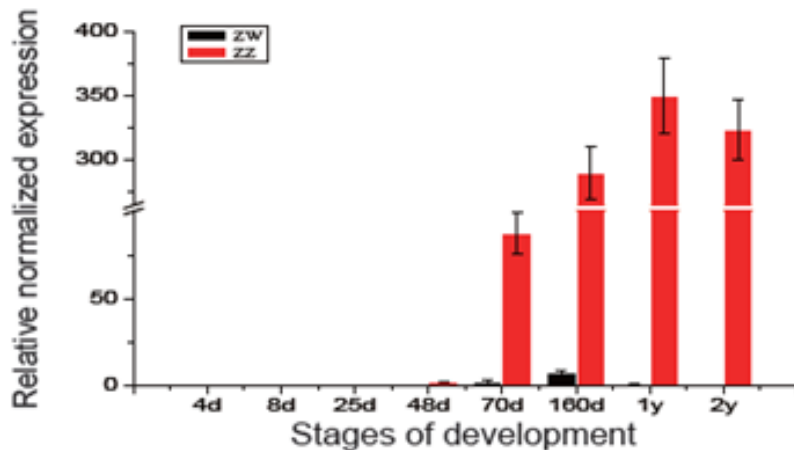


- ◆ Chromosomes FISH demonstrated that Dmrt1 gene hybridizes only with two Z chromosomes in ZZ male, and one Z chromosome in ZW female. But no hybridization signal was observed on W chromosome.

Chen et al., 2014 Nat Genet. 46:253-260.

Gonad *in situ* hybridization using dmrt1 as probe

- *In situ* hybridization showed that dmrt1 gene expressed specifically in testes at the sex-determination stage and persists at high levels during testis development. No expression was observed in ovaries.
- RT-PCR demonstrated that Dmrt1 gene expressed in larvae at 48 dah, and increased significantly from 70D till 2 Y with a peak at 1Y. But almost no expression was detected in ovaries at any stages in females.

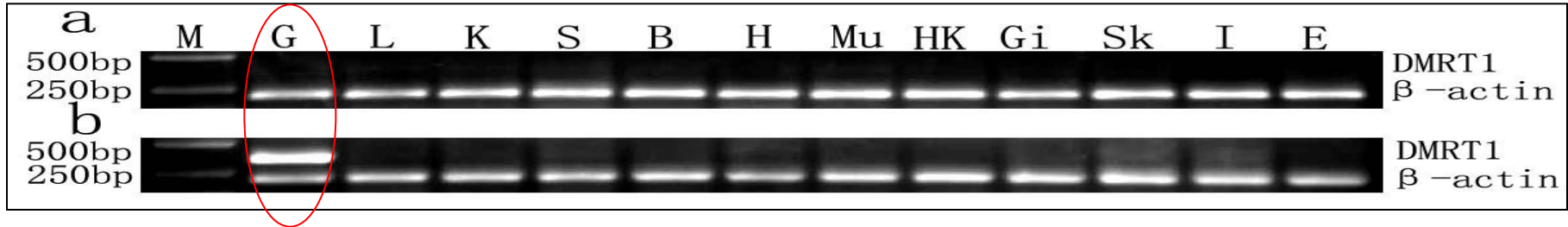


Testes

Ovaries

Chen et al., 2014 Nat Genet. 46:253-260.

Expression pattern of dmrt1



The expression pattern of DMRT1 in female (a) and male (b) tissues
G: gonad, L: liver, K, kidney, S: spleen, B: brain, H: heart, Mu: muscle, HK: head kidney, Gi: gill, Sk: skin, I: intestine, E: eye



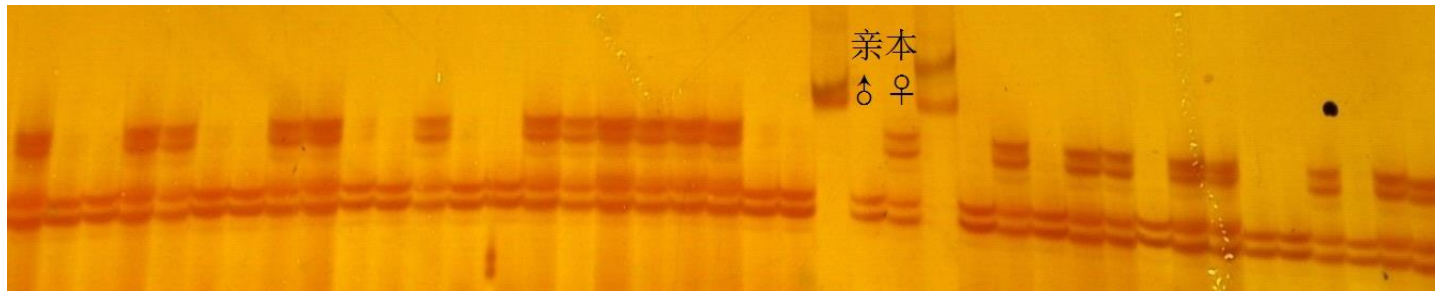
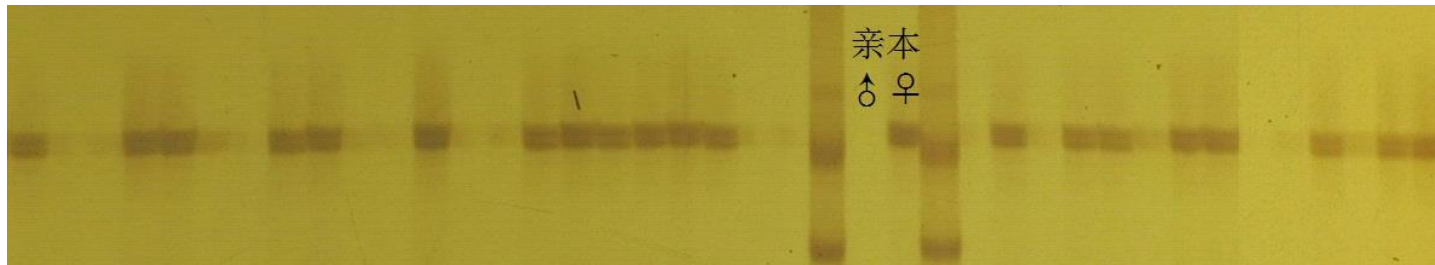
1, 2. ovary of normal female; 3 and 4. testes of normal male; 5-8. testes of pseudomales

- **Dmrt1 was only expressed in male gonad, but not in other tissues.**
- **When genetic females were reversed to phenotypic male, Dmrt1 expression was significantly induced in the gonad of pseudomale, demonstrating that Dmrt1 plays an important role in testis development and sex reversal in tongue sole.**

Epigenetic inheritance of sex reversal

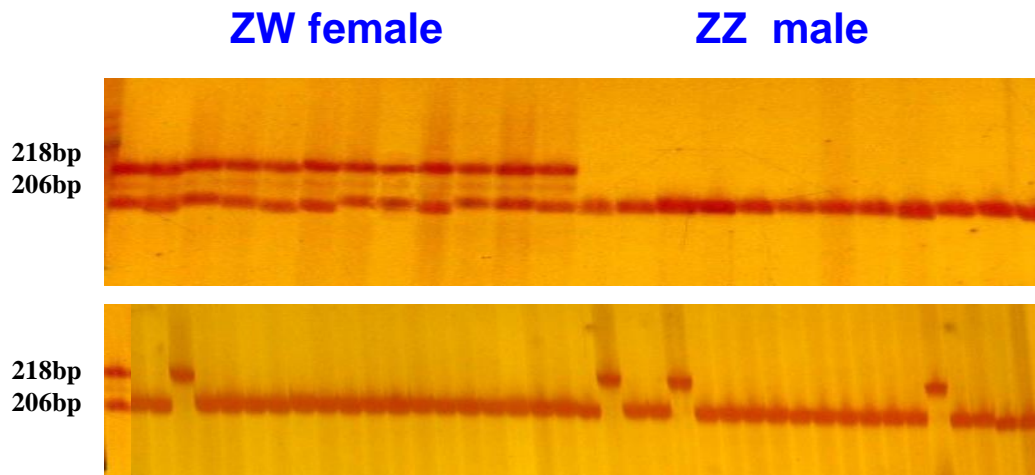
Isolation of sex-linked SSR markers

Sex-linked SSR markers were isolated by comparing the female and male genome sequences. There are two kinds of the markers: 1) single band in female, but no band in male; 2) two bands in female, and only one band in male.



Identification of WW super female

We developed the technique for identify the ZZ male , ZW female, ZW neo-male and WW superfemale.

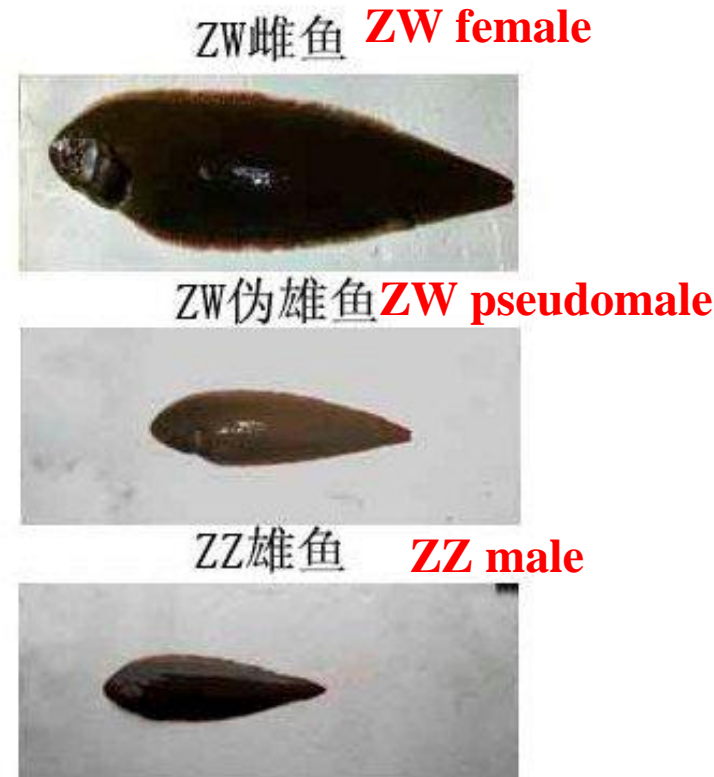


Confirmation of sex-linked SSR markers

218 bp : W chromosome specific allele

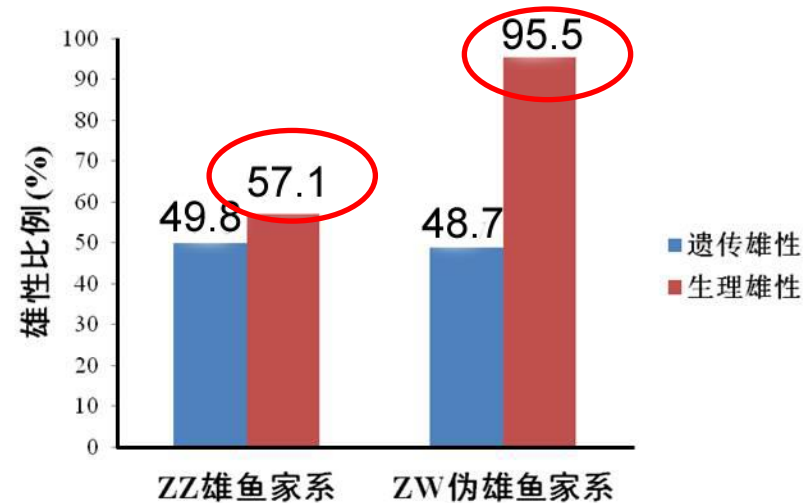
206 bp : Z chromosome specific allele

Among the 39 mitogynogenetic embryos, four embryos are WW superfemale.



Analysis of ratio of female to male

- ◆ We developed more than 100 tongue sole families using ZZ males or ZW pseudomales as paternal parents.
- ◆ Compared to 57.1% phenotypic male in ZZ normal family, there was a huge male bias, about 95% phenotypic male in pseudomale family, indicating the offsprings of pseudomales are easier to reverse to pseudomales.
- ◆ Surprisingly, no W spermatozoa was detected in pseudomales, leading to 50% genetic male in pseudomale's offsprings.



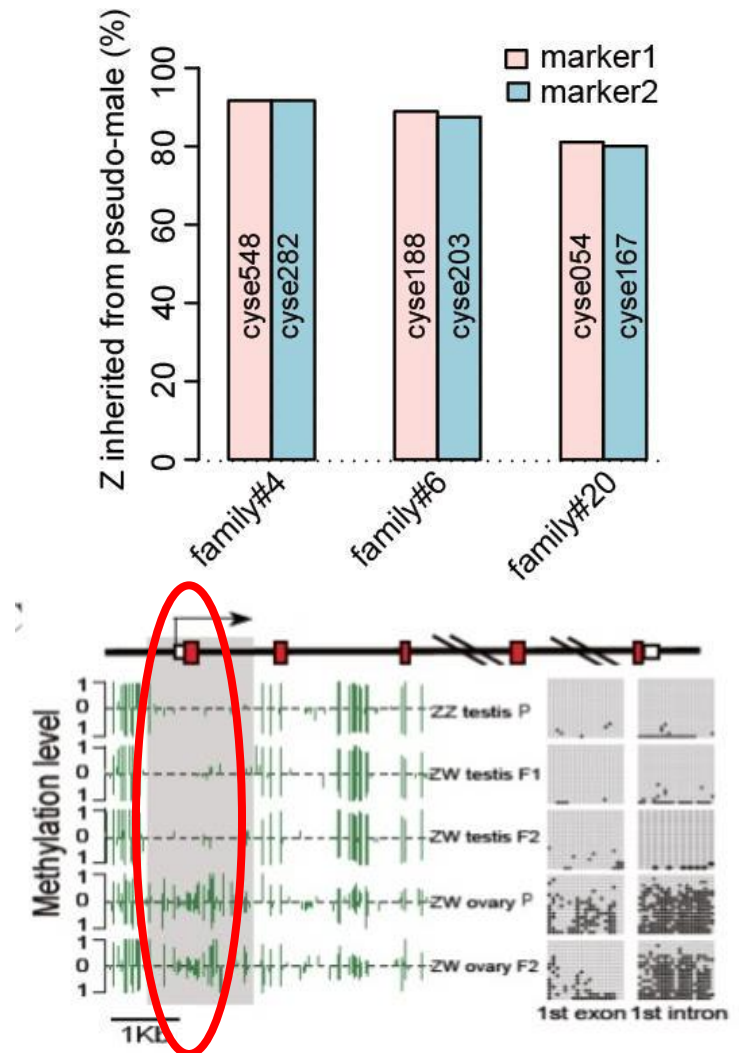
ZZ male family ZW pseudomale family
Blue bar: Genetic male; Red bar: phenotypic male



↑
Sperm of pseudomale

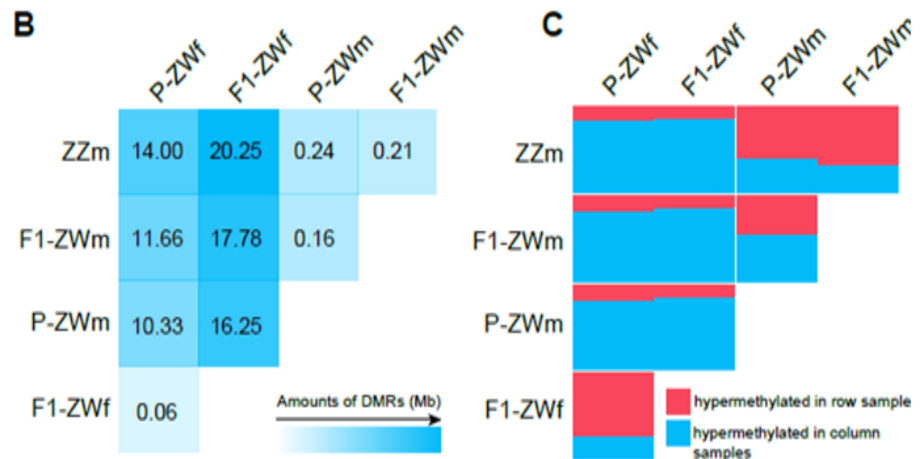
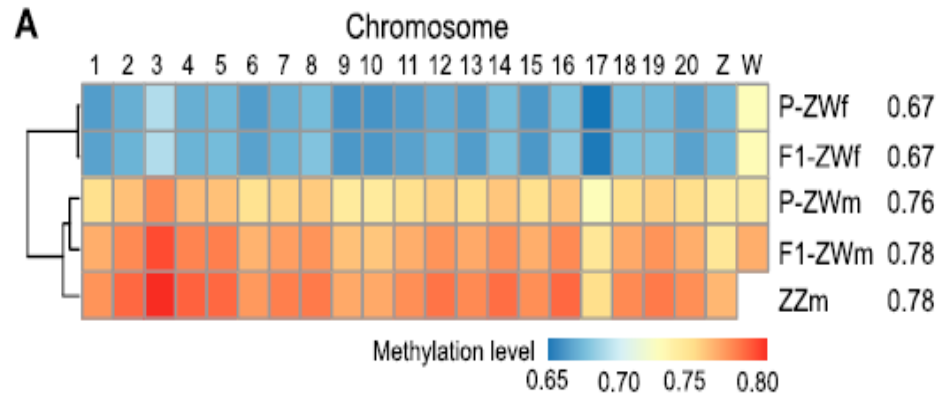
Epigenetic inheritance of sex reversal

- ◆ Microsatellite markers that distinguish paternal and maternal Z chromosomes revealed that 95% second-generation pseudomales have inherited the Z chromosome from their sex-reversed fathers.
- ◆ Dmrt1 in the offsprings of pseudomales retained the paternal methylation pattern, implying that transgenerational inheritance of DNA methylation.



Epigenetic inheritance of sex reversal

- Whole-genome methylation sequencing demonstrated that the methylation level in the testis from pseudomale and their progeny (pseudomale) is similar to testes of normal male, but clearly different from the ovary.
- Up to 86% of the DMRs between P-ZWm and ovaries are maintained in F1-ZWm, indicating that the overall change in methylation status of the genome after sex reversal had been inherited by the next generation.



In summary

- Whole genome assembly and analysis of tongue sole were completed.
- High resolution genetic linkage maps (SSR and SNP) were constructed.
- Convergent evolution of sex chromosomes was observed.
- Dmrt1 was found to be Z-chromosome linked, male specially expressed, essential for testes development in tongue sole.
- Epigenetic regulation was found to play multiple crucial roles in sexual reversal in tongue sole.

Acknowledgements

In collaboration with Jun Wang in BGI at Shenzhen
and M. Schartl in Germany

Yellow Sea Fish. Institute:

Changwei Shao

Qisheng Tang

Wentao Song

Mingshu Xie

Zhenxia Sha

Na Wang

BGI at Shenzhen:

Jun Wang

Guojie Zhang

Quanfei Huang

Na An

Qiye Li

Wuerzburg University

in Germany:

M. Schartl

University of Singapore:

Y. Hong

Thank you

