Characterization of the sex determining region of channel catfish (*Ictalurus punctatus*) and development of a sex-genotyping test.

Ming Wen[1,2,*], Qiaowei Pan[2,3], Wes Larson[4], Camille Eché [5], Yann Guiguen[2*]

Affiliations

[1] State Key Laboratory of Developmental Biology of Freshwater Fish, College of Life Science, Hunan Normal University, Changsha, China

[2] INRAE, LPGP, 35000 Rennes, France.

[3] Department of Ecology and Evolution, University of Lausanne, Switzerland.

[4] National Oceanographic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, Auke Bay Laboratories, 17109 Point Lena Loop Road, Juneau, AK 99801, USA.

[5] INRAE, GeT-PlaGe, Castanet-Tolosan, France.

[*] Corresponding authors: Ming Wen (ming.wen@hunnu.edu.cn) and Yann Guiguen (yann.guiguen@inrae.fr)

Abstract

Channel catfish is an important species for aquaculture that exhibits a sexually dimorphic growth in favor of males. Genetic sexing and development of sex markers is crucial for the early identification of sex and of particular genotypes (YY males) for the production of all-males population in channel catfish aquaculture. In this study, we sequenced genomic DNA from pools of males and pools of females to better characterize the sex determining region (SDR) of channel catfish and to develop sex-specific markers for genetic sexing. Perfroming comparative analyses on males and females pooled genomic reads, we identified a large SDR (~ 8.3 Mb) in the middle of channel catfish linkage group 4 (LG04). This non-recombining SDR contains a high-density of male-specific (Y chromosome) fixed single nucleotide polymorphisms (SNPs) along with ~185kb male-specific insertions or deletions. This SDR contains 95 annotated protein-encoding genes, including the recently reported putative channel catfish master sex determining (MSD) gene, breast cancer anti-estrogen resistance protein 1 (*bcar1*), located at one edge of the SDR. No sex-specific SNPs and/or indels were found in the coding sequence of *bcar1*, but one male-specific SNP were identified in its first intron. Based on this genomic information, we developed a PCR-based sex-specific genetic test. Genotyping results confirmed strong linkage between phenotypic sexes and the identified SDR in channel catfish. Our results confirms, using a Pool-Seq approach, that channel catfish is male heterogametic (XX-XY) with a large SDR on the LG04 sex chromosome. Furthermore, our genotyping primers can be used to identify XX, XY and YY fish that will facilitate future research on sex determination and aquaculture applications in channel catfish.

Key words: Channel catfish, Pool-sequencing, Sex markers, Sex chromosome, Sex determining region

Introduction

Channel catfish, *Ictalurus punctatus*, belongs to the Ictaluridae family within the order Siluriformes (catfishes), and is known as a highly adaptive species with a high tolerance to low water quality and resistance to various infectious pathogens [1]. Its native geographic distribution encompasses south-east of Canada, the eastern part of the United States, and north Mexico. It is now widely farmed worldwide and is an important aquaculture species especially in North America and in Asia where it has been introduced. Channel catfish exhibits a sexual growth dimorphism with male growing faster than female [2]. Therefore, the establishment of all-male mono-sex populations of channel catfish are sought after for aquaculture production [2]. This made research on sex determination (SD) mechanisms in channel catfish especially interesting. Channel catfish SD system has been described as being male heterogametic (XX/XY) with morphologically undifferentiated sex chromosomes [3]. This study identified the breast cancer anti-estrogen resistance protein 1 gene (*bcar1*) as the potential channel catfish master sex determining (MSD) gene encoding a male-specific isoform transcript [4].

In the teleost fish, SD mechanisms are highly diverse with either genetic or environmental (temperature, pH, social interactions) triggers [5, 6]. In channel catfish, like in some other aquaculture fish species such as goldfish and Nile tilapia, previous studies also reported that SD could be under genetic determinants (genetic sex determination, GSD) with an influence of temperature [7-11], a phenomenon known as thermal effects on GSD (GSD+TE) [12]. However, in channel catfish the effective temperature threshold for triggering male-to-female sex-reversal is 34℃ [7], questioning the significance of this GSD+TE effect in wildtype populations of channel catfish. GSD in teleosts is characterized by a high turnover of SD systems, sex chromosomes and MSD genes. Classical monofactorial male (XX/XY) or female (ZZ/ZW) heterogametic SD systems have been described [13] along with multiples sex chromosomes [14, 15]. This high SD turnover [16] is even observed among closely-related species [17] or among different populations of the same species [18, 19]. This plasticity in SD of teleost makes them interesting models to study the evolution of SD in vertebrates.

Growth is one of the most valuable economic traits in fish and benefiting from a better growth rate of either males or females is key for modern aquaculture practices [20]. In most teleost fish, sex chromosomes are cytogenetically and morphologically undifferentiated [21], but it is feasible to develop molecular approaches to identify sex markers in fish for aquaculture applications. These molecular approaches firstly used different molecular marker techniques such as amplified fragment length polymorphism length polymorphism (AFLP) [22, 23], random amplified polymorphism DNA marker technology (random amplified polymorphic DNA, RAPD) [24, 25], microsatellite marker technology simple sequence repeats (SSR) [26, 27]. With the development of high throughput sequencing technologies,

sex-linked markers were successfully characterized using Restriction-site Associated DNA (RAD) sequencing [28, 29], Quantitative Trait Loci (QTL) mapping [30, 31], and whole-genome sequencing [32, 33]. In the present study, we implemented a male versus females pooled whole genome sequencing approach (Pool-sequencing, Pool-Seq) [34] to identify sex-linked genetic variations in channel catfish, leveraging the available genome assemblies of an XX female [1] and a YY male [4]. With this strategy, we identified a large SDR on channel catfish linkage group 4 (LG04), confirming previous results demonstrating the LG04 was the sex chromosome in that species [4]. We also developed sex-specific markers for channel catfish that will be important for SD research and the development of genetically all-males population for channel catfish aquaculture.

Materials and methods

Fish sampling and genomic DNA extraction

Fish used in this experiment were collected from a wild population in the Wisconsin River (Grant county, Wisconsin state, USA). The phenotypic sex of the 49 males and 35 females sampled was identified based on sex-specific differences in the vent area of each fish. Fin clips were stored in 95% ethanol until genomic DNA (gDNA) extraction.

For routine sex genotyping, fin clips were cut into pieces, and lysed by 20 mg Proteinase K within 5% Chelex at 55℃ for 2 hours, followed by Proteinase K denatured at 99℃ for 2 min. After brief centrifugation, supernatant containing genomic DNA was transferred to a new PCR tube. Finally, DNA was diluted with equal volume of water and stored at -20℃ before genotyping.

For Pool-Seq, Fin clips' gDNA from 30 phenotypic females and 30 males were extracted individually using the NucleoSpin Kits for Tissue (Macherey-Nagel, Duren, Germany) following the manufacturer's instructions. Concentrations of gDNA were measured using a Qubit3 flurometer and the Qubit dsDNA HS Assay Kit (Invitrogen, Carlsbad, CA) and gDNA from each individual was normalized to reach equimolar gDNA concentration within each sex. All normalized gDNA were then pooled according to sex to produce equimolar gDNA pools of males and females.

Whole genome pool-sequencing (Pool-seq)

Pool-seq libraries (male and female pools) were built constructed using the Truseq nano kit (Illumina, ref. FC-121-4001) following the manufacturer's instructions. Briefly, each pool was sonicated using a Bioruptor (Diagenode). Firstly, 200 ng of gDNA from the male and female pools was briefly sonicated with Bioruptor sonication device (Diagenode, Liege, Belgium). The sonicated DNA was size selected on magnetic beads filtering for a 550 bp insert size and adenylated on their 3′ ends. Adenylated DNA was ligated to Illumina's specific adapters and, after purification on magnetic beads, was amplified in an 8 cycles PCR. Libraries were purified using magnetic beads, checked on a Fragment Analyzer (Agilent) using the HS NGS

Fragment kit (DNF-474-33) and quantified by qPCR using the KAPA Library quantification kit (Roche, ref. KK4824). Finally, 2 x150 paired reads for male and female pool were generated.

Identification of sex determining region (SDR)

The published two genome assemblies, one XX female [GCA_001660625.2] and one YY male [GCA_004006655.3] were used as genome references for our whole genome analysis. Illumina reads, from both male and female pools, were respectively mapped onto the two individual assemblies using BWA mem version 0.7.17 [35] with default setting. Then, for each assembly, the aligned sequences were sorted, grouped, and merged using the default settings of Picard tool version 2.18.2 (https://broadinstitute.github.io/picard/). PCR duplicated were removed using the Picard tool. In addition, reads with mapping quality less than 20 and that not mapped uniquely were removed using samtools version 1.9 [36]. Samtools mpileup was then used to generate a pileup file using the two sex BAM files by setting per-base alignment quality disabled (-B). Subsequently, with this mpileup file a sync file was generated using popoolation mpileup2sync version 1.201 [37] by setting minimum quality to 20. The nucleotide composition of each sex for each position in the reference was included in the sync file. SNPs and window coverages of the male and female pooled datasets were calculated using PSASS (https://github.com/SexGenomicsToolkit/PSASS). To identify sex-biased SNP enriched region, a 100 kb sliding window with an output every 10 kb was applied using PSASS. The parameters for running PSASS were as follows: minimum depth to consider a site was set to 10 (--min-depth 10), range of frequency for a sex-linked SNP in the heterogametic sex was set to 0.5±0.2 (--freq-het 0.5, --range-het 0.1), frequency of a sex-linked SNP in the homogametic sex set to more than 0.95 (--freq-hom 1, --range-hom 0.02). The resulting analyses were plotted the PSASS supporting R package (http://github.com/RomainFeron/PSASS-vis).

Sex-specific primers design, SNP verification, and genotyping

The channel catfish SDR identified with the Pool-Seq approach described above was visually inspected with the Integrative Genomics Viewer (IGV) [38] in order to find short insertions or deletions characterized by a half read coverage in one sex and no or very few reads (less than 2X) in the other sex. Primers were designed flanking these potential sex-specific fragments using Primer3 version 4.1.0 (http://primer3.ut.ee). The same strategy was applied to design two primer sets amplifying fragments containing sex-specific SNPs. All primers are listed in Table S1 and PCRs were performed with 0.25 units of JumpStart Taq DNA Polymerase (Sigma Aldrich), 1 µl of 10× PCR Buffer, 100 µM dNTP mixture, 1 µl of 50 ng/µl gDNA, 0.1 µM of each primer in a total volume of 25 µl. The conditions of PCR were: denaturation at 94℃ for 30s, primer annealing at 56℃ for 30s and polymerase extension at 72℃ for 30s for 35 cycles. Finally, PCR products were visualized on 1.5% agarose gels.

Results

Whole genome pool-sequencing data of channel catfish

Whole genome sequencing of male and female pools yielded respectively 266,697,484 and 231,722,384 paired-end clean reads in total (Table 1). With the help of genome reference of YY channel fish, mapping rates of paired end reads from male pool and female pool were 99.55% and 99.50% respectively, and the average depth of male pool and female pool were 34 and 30 respectively. Besides, the nucleotide statistics of sequencing data showed that the GC contents were 40.0% and 40.0%, the averages of Q20 were 96.7% and 96.8%, and the averages of Q30 were 91.9 and 92.0 in male pool and female pool respectively (Table 1). These results show that the sequencing datasets of male and female channel catfish pools are of high and equivalent quality.

Identification of sex chromosome and sex-differentiation region in channel catfish

To characterize the sex-chromosome and SDR of channel catfish, we first searched for sex-biased signals, including sex coverage differences and sex-biased SNPs using the published channel catfish genome as reference. The whole genome analysis of the sex-specific SNP distribution using the male assembly as a reference showed that a high density of male-biased SNPs was present on LG04 (Y chromosome) with no obvious sex-biased region (Figure 1, Figure 2B and 2C). A similar SDR region was also identified on LG04 (X chromosome) using the female assembly as a reference (Figure S1, Figure S2B). Furthermore, population genetics analyses showed significant genetic differentiation with high F_{ST} value, between males and females, were observed in the position of SDR (Figure 2A, Figure S2A). These results confirmed previous published results describing LG04 as the sex-chromosome in channel catfish [1], but differences in the distribution of male-biased SNPs were observed in our Pool-Seq analyses depending on the reference assembly used. With the male genome assembly as a reference, male-biased SNPs were more tightly clustered (Figure 2B, Figure S2B) in an 8.3 Mb SDR compared to a 9.8 Mb SDR with the female assembly. Comparison between X and Y chromosomes showed a good alignment in the pseudo-autosomal regions around the SDR (Figure S3), whereas, the X and Y alignments were fragmented in the SDR (Figure S3), giving rise to the discrepancy in the SNP distribution pattern observed in our data. Furthermore, these small structural changes between the X and Y chromosome could potential be involved in the suppression of the recombination between sex chromosomes.

Development of sex-specific chromosome markers and SNPs verification in channel catfish

Using Integrative Genomics Viewer (IGV), we manually checked the mapping data throughout the Y chromosome and identified around 185 kb of potential Y-specific insertions based on coverage difference between male and female pool (Supplemental excel file 3). Then

we designed three PCR primer sets in conserved male and female (Y and X) regions around some of these small male-specific (Y-specific) insertions (Figure S4). These three sex-linked primer sets produce PCR products of different sizes in male and female (Figure 3A) with a single amplicon in females (X-amplicon) and two amplicons in males (X- and Y- amplicons) (Figure 3B). PCR of 12 additional phenotypic males and 26 phenotypic females (new individuals not included on the male and female pools) demonstrated a complete association (p-value = 7.08e-09) of the phenotypic sex with the sex genotype (XX or XY) inferred from the PCR tests for all three sex-specific primer sets used (Figure S5). In addition, we also designed conserved X and Y primers around male-specific SNPs and four SNPs with specific heterozygosity in males were verified by Sanger sequencing of the resulting PCR products sequencing (Figure 3C).

Search for a potential MSD genes in the channel catfish SD

The MSD gene controlling genetic SD in channel catfish likely resides in the SDR on Y chromosome [39]. We extracted all protein-encoding genes (N = 95) from the 8.3 Mb SDR on the Y chromosome that we characterized using Pool-Seq (Supplemental excel file 1). In the Y SDR, we observed 10,190 male SNPs (average of 1.23 SNPs per kb), and most of them (9,953, 97.7 %) were located on non-coding sequences and fewer (237, 2.3 %) on coding sequence (Table S2). Among the 95 protein-encoding genes within the channel catfish SDR, we did not found any obvious candidate MSD gene apart from *bcar1*, that has been recently described as the potential channel catfish MSD gene (BIB). This *bcar1* gene was localized on the edge of the Y chromosome non-recombinant region in a region with one male-biased SNP (see Figure S6). No male-biased SNPs were detected in the coding region of the *bcar1* gene but in the first intron (Supplemental excel file 2).

Discussion

Channel catfish, belongs to the Ictaluridae family, is a native species from North America, and was introduced worldwide for recreational fishing and aquaculture. The understanding of its SD mechanism would not only benefit to research on the evolution of SD in vertebrates, but will have also important impact on aquaculture practice. In teleost, GSD systems include both male heterogametic (XX/XY) and female heterogametic (ZZ/ZW) like for example, goldfish [40], rainbow trout [41], and medaka [42] that have a male heterogametic SD system, or Chinese half-smooth tongue sole that has a female SD heterogametic system [43]. Channel catfish is known as a species with a male SD heterogametic system [7]. In this study, using a novel pool-sequencing approach, we confirmed previous results [4] showing that this species is male heterogametic with a large non-recombining SDR on LG04. Pool-sequencing is increasingly applied for searching population genomic polymorphisms [44], and has been also used recently to characterize SD in teleost fish with a strict GSD system [33, 40, 45, 46]. In channel catfish, environmental factors including temperature, like in goldfish [8] and Nile

tilapia [10], have been shown to interfere with the GSD system, thus preventing the use of a pooled strategy that can be biased by sex-reversed individuals [7]. This influence of temperature on channel catfish sex-differentiation has been shown to produce male-to-female sex reversals [38] that would result in XY females in species with male heterogametic SD system. These sexes reversed XY females, depending on their proportion in a pool of phenotypic females, could prevent the characterization of a sex-biased signal using a pooled strategy like the one we used. However, such sex-reversal are only induced at a high temperature (34°c) for channel catfish that is unlikely to be observed in the wild-type population that we sampled within its native range of distribution (Wisconsin, USA). The results of our whole genome Pool-Seq analysis and the individual genotyping of additional individuals confirmed the absence of such male-to-female sex reversions in our wildtype population. This suggests that channel catfish SD is mainly driven by a GSD system at least in its native habitats. It would be however interesting to explore further different aquaculture populations to investigate if captive breeding conditions could produce male-to-female sex reversals.

Growth in channel catfish has been described as sexually dimorphic with males growing faster than females [2]. Therefore, the establishment of all-male population of channel catfish would increase production in commercial aquaculture operations. All-male or all-female fish populations can be generated through gynogenesis or the production of special sex-reversed breeders using hormonal sex control [47]. The optimal strategies to produce monosex population are variable depending on the desired sex phenotype and the GSD system of the species. Genetic tools such as sex markers could simplify and shorten these procedures by quickly identifying special phenotype/genotype association [41, 43, 48] like for instance potential XY females or even YY males that would be sought after for the production of all-male population in channel catfish aquaculture. The characterization of sex-specific variations for the development of a genetic sexing method like the one we provide here in channel catfish are then of great interest for their application for better sex-control in aquaculture. As our genotyping approach in channel catfish allows the identification of both X- and Y-alleles, different genotypes (sex XX, XY, and YY)can be easily characterized. Therefore, it is possible to generate YY fish by combining temperature or steroid-induced sex reversals to first produce XY females that can be crossed with a normal XY male in order to produce YY males that are needed in order to generate genetically all-male populations. Both XY females and YY males can be easily detected with our sex genotyping tests, allowing a much quicker genotyping than the progeny testing protocols that were used without such genetic sexing tools.

Mammals and most birds have cytologically differentiated sex chromosomes. In mammals the Y chromosome is highly degenerated and contains large repetitive sequences, small

pseudo-autosomal regions and very few protein-encoding genes compared to the X chromosome. In teleost fishes, most species do not have morphologically differentiated sex chromosomes that can be distinguished on simple metaphase chromosomes karyotypes [4]. However, this does not mean that these morphologically undifferentiated sex chromosomes could not have a large differentiated SDR at the molecular level. For instance relatively large SDR were reported in species with morphologically undifferentiated sex chromosomes, for instance in goldfish (SDR ~11.7 Mb) [40], blackchin tilapia, *Sarotherodon melanotheron* (SDR ~17.9 Mb) [46], and Nile tilapia *Oreochromis niloticus* (~10.7 Mb) [49]. In contrast, small SDRs were identified through whole genome analysis in yellow perch, *Perca flavescens* (SDR ~0.1 Mb) [33] and in common seadragon, *Phyllopteryx taeniolatus* (SDR ~47 kb) [50]. In the Takifugu, *Takifugu rubripes* [51], the SDR is restricted to the minimum size with only a missense SNP on the Y chromosome associated with SD. In channel catfish, the Y chromosome (38.5Mb) is slightly longer than X chromosome (34.5Mb), and we identified a relative large SDR (8.3 Mb) on LG4 as previously described using the whole genome sequencing of a YY male [4]. This region is highly enriched with male-specific SNPs, male-specific insertions, and many chromosomal rearrangements between X and Y which can suppress recombination between X and Y chromosome [52].

A high diversity of MSD genes have been reported in teleosts[16], and most of these genes belong to a few gene families known as "usual suspects" MSD genes [53] like the Doublesex/Mab3 (DM) domain family [42, 43, 54], transforming growth factor beta (TGF-β) family [32, 33, 45, 50, 51, 55-58], steroid enzymes [59], and SRY-related HMG-Box (Sox) family [60]. However, exceptions to this "usual suspect" rule have been found such as the *sdY* in rainbow trout [41], and *paics* in blue tilapia, *Oreochromis aureus* [61]. In channel catfish, *bcar1*, a member of the Crk-associated substrate (CAS) family of scaffold proteins, was strongly suggested as a MSD gene due to its location within the SDR, the expression of a male-specific isoform differentiating male gonads, and the male-to-female sex reversal of *bcar1* knockout males [4]. In our study, we found that one sex-biased SNP in the *bcar1* locus are localized in the first intron, and that *bcar1* is surprisingly located at the edge of the channel catfish SDR in a region with relatively few male-specific SNPs. There were no sex-biased SNPs observed on genomic sequence of bcar1 corresponding to previous results [4]. Most of reported SD genes were originated from allelic diversification or gene duplication [53]. However, bcar1 as a SD gene was not derived from both of them, but to evolve as gonad specific transcript for initializing early gonad to differentiate to either testis or ovary [4].

Conclusions

Our results confirmed that channel catfish has a male heterogametic (XX/XY) system, and that SD in that species is strongly regulated by genetic factors. Using a complimentary

approach than those previously used before [40, 45], a large non-recombining region (~8.3 Mb) was characterized on LG04 of channel catfish, confirming the size of this SDR region and that LG04 is the sex chromosome of this species. The already described channel catfish *bcar1* MSD gene was also found in this large non-recombining LG04 region but surprisingly in a region with a relatively low density of male-specific SNPs, on the edge of this large SDR. There are no genomic differences observed at *bcar1* coding sequences. In addition, we used some small male- and Y-specific insertions in this SDR region to develop simple PCR-based sex genotyping tools that enable the identification of XX, XY and YY genotypes, that can be now used for the production of all-male populations in channel catfish aquaculture.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Data availability:

Male and female channel catfish pool-sequencing reads have been deposited in the Sequence Read Archive (SRA), under BioProject PRJNA821372.

Acknowledgements

Yann Guiguen conceived the project, designed the experiments, acquired funding, analyzed genomic data, and wrote the manuscript. Ming Wen, conducted investigation, analyzed genomic data, and wrote the manuscript. Qiaowei Pan conducted investigation and wrote the manuscript. Wes Larson collected samples and modified this manuscript. Camille Eché analyzed genomic data.

Authors' contributions:

This project was supported by funds from the “Agence Nationale de la Recherche” (PhyloSex project, grant No. ANR-13-ISV7-0005). The GeT-PlaGe sequencing facility was supported by France Génomique National infrastructure, funded as part of “Investissement d’avenir” program managed by Agence Nationale pour la Recherche (grant No. ANR-10-INBS-09). We are grateful to the Genotoul bioinformatics platform Toulouse Midi-Pyrenees (Bioinfo Genotoul) for providing computing and/or storage resources.

Reference

1. Liu Z, Liu S, Yao J, Bao L, Zhang J, Li Y, Jiang C, Sun L, Wang R, Zhang Y *et al*: **The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts**. *Nat Commun* 2016, **7**:11757.
2. Simco BA, Goudie CA, Klar GT, Parker NC, Davis KBJTotAFS: **Influence of sex on growth of channel catfish**. 1989, **118**(4):427-434.
3. Tiersch TR, Simco BA, Davis KB, Chandler RW, Wachtel SS, Carmichael GJ: **Stability of genome size among stocks of the channel catfish**. *Aquaculture* 1990, **87**(1):15-22.
4. Bao L, Tian C, Liu S, Zhang Y, Elaswad A, Yuan Z, Khalil K, Sun F, Yang Y, Zhou T *et al*: **The Y chromosome sequence of the channel catfish suggests novel sex determination mechanisms in teleost fish**. *BMC Biology* 2019, **17**(1):6.
5. Baroiller J-F, D'cotta H: **Environment and sex determination in farmed fish**. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* 2001, **130**(4):399-409.
6. Kobayashi Y, Nagahama Y, Nakamura M: **Diversity and plasticity of sex determination and differentiation in fishes**. *Sexual Development* 2013, **7**(1-3):115-125.
7. Patiño R, Davis KB, Schoore JE, Uguz C, Strüssmann CA, Parker NC, Simco BA, Goudie CA: **Sex differentiation of channel catfish gonads: normal development and effects of temperature**. *Journal of Experimental Zoology* 1996, **276**(3):209-218.
8. Goto-Kazeto R, Abe Y, Masai K, Yamaha E, Adachi S, Yamauchi K: **Temperature-dependent sex differentiation in goldfish: Establishing the temperature-sensitive period and effect of constant and fluctuating water temperatures**. *Aquaculture* 2006, **254**(1-4):617-624.
9. Wen M, Feron R, Pan Q, Guguin J, Jouanno E, Herpin A, Klopp C, Cabau C, Zahm M, Parrinello HJBg: **Sex chromosome and sex locus characterization in goldfish, Carassius auratus (Linnaeus, 1758)**. 2020, **21**(1):1-12.
10. Wessels S, Hörstgen-Schwark G: **Temperature dependent sex ratios in selected lines and crosses with a YY-male in Nile tilapia (Oreochromis niloticus)**. *Aquaculture* 2011, **318**(1-2):79-84.
11. Hattori R, Gould R, Fujioka T, Saito T, Kurita J, Strüssmann C, Yokota M, Watanabe S: **Temperature-dependent sex determination in Hd-rR medaka Oryzias latipes: gender sensitivity, thermal threshold, critical period, and DMRT1 expression profile**. *Sexual Development* 2007, **1**(2):138-146.
12. Ospina-Alvarez N, Piferrer F: **Temperature-dependent sex determination in fish revisited: prevalence, a single sex ratio response pattern, and possible effects of climate change**. *PloS one* 2008, **3**(7):e2837.
13. Devlin RH, Nagahama Y: **Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences**. *Aquaculture* 2002, **208**(3-4):191-364.
14. Sember A, Nguyen P, Perez MF, Altmanová M, Ráb P, Cioffi MdB: **Multiple sex chromosomes in teleost fishes from a cytogenetic perspective: state of the art and future challenges**. *Philosophical Transactions of the Royal Society B* 2021, **376**(1833):20200098.
15. Moore EC, Roberts RB: **Polygenic sex determination**. *Current Biology* 2013, **23**(12):R510-R512.
16. Pan Q, Anderson J, Bertho S, Herpin A, Wilson C, Postlethwait JH, Schartl M, Guiguen Y: **Vertebrate sex-determining genes play musical chairs**. *C R Biol* 2016, **339**(7-8):258-262.
17. Matsuda M, Sakaizumi M: **Evolution of the sex-determining gene in the teleostean genus Oryzias**. *General and comparative endocrinology* 2016, **239**:80-88.
18. Dong Z, Li X, Yao Z, Wang C, Guo Y, Wang Q, Shao C, Wang Z: **Oryzias curvinotus in Sanya Does Not Contain the Male Sex-Determining Gene dmy**.

Animals 2021, **11**(5):1327.
19. Pan Q, Feron R, Jouanno E, Darras H, Herpin A, Koop B, Rondeau E, Goetz FW, Larson WA, Bernatchez L: **The rise and fall of the ancient northern pike master sex-determining gene**. *Elife* 2021, **10**:e62858.
20. Mei J, Gui JF: **Genetic basis and biotechnological manipulation of sexual dimorphism and sex determination in fish**. *Sci China Life Sci* 2015, **58**(2):124-136.
21. Sandra G-E, Norma M-M: **Sexual determination and differentiation in teleost fish**. *Reviews in Fish Biology and Fisheries* 2009, **20**(1):101-121.
22. Felip A, Young WP, Wheeler PA, Thorgaard GH: **An AFLP-based approach for the identification of sex-linked markers in rainbow trout (Oncorhynchus mykiss)**. *Aquaculture* 2005, **247**(1-4):35-43.
23. Lee BY, Coutanceau JP, Ozouf-Costaz C, D'Cotta H, Baroiller JF, Kocher TD: **Genetic and physical mapping of sex-linked AFLP markers in Nile tilapia (Oreochromis niloticus)**. *Mar Biotechnol (NY)* 2011, **13**(3):557-562.
24. Vale L, Dieguez R, Sanchez L, Martinez P, Vinas A: **A sex-associated sequence identified by RAPD screening in gynogenetic individuals of turbot (Scophthalmus maximus)**. *Mol Biol Rep* 2014, **41**(3):1501-1509.
25. Kovács B, Egedi S, Bártfai R, Orbán LJG: **Male-specific DNA markers from African catfish (Clarias gariepinus)**. 2000, **110**(3):267-276.
26. Shikano T, Herczeg G, Merilä JJBRN: **Molecular sexing and population genetic inference using a sex-linked microsatellite marker in the nine-spined stickleback (Pungitius pungitius)**. 2011, **4**(1):1-6.
27. Fuji K, Yoshida K, Hattori K, Ozaki A, Araki K, Okauchi M, Kubota S, Okamoto N, Sakamoto T: **Identification of the sex-linked locus in yellowtail, Seriola quinqueradiata**. *Aquaculture* 2010, **308**:S51-S55.
28. Kafkas S, Khodaeiaminjan M, Guney M, Kafkas E: **Identification of sex-linked SNP markers using RAD sequencing suggests ZW/ZZ sex determination in Pistacia vera L**. *BMC Genomics* 2015, **16**:98.
29. Gamble T: **Using RAD‐seq to recognize sex‐specific markers and sex chromosome systems**. *Molecular ecology* 2016, **25**(10):2114-2116.
30. Dong C, Jiang P, Zhang J, Li X, Li S, Bai J, Fan J, Xu P: **High-Density Linkage Map and Mapping for Sex and Growth-Related Traits of Largemouth Bass (Micropterus salmoides)**. *Front Genet* 2019, **10**:960.
31. Zhou Y-L, Wang Z-W, Guo X-F, Wu J-J, Lu W-J, Zhou L, Gui J-F: **Construction of a high-density genetic linkage map and fine mapping of QTLs for growth and sex-related traits in red-tail catfish (Hemibagrus wyckioides)**. *Aquaculture* 2021, **531**.
32. Wen M, Pan Q, Jouanno E, Montfort J, Zahm M, Cabau C, Klopp C, Iampietro C, Roques C, Bouchez O *et al*: **An ancient truncated duplication of the anti-Mullerian hormone receptor type 2 gene is a potential conserved master sex determinant in the Pangasiidae catfish family**. 2022:2022.2001.2014.475871.
33. Feron R, Zahm M, Cabau C, Klopp C, Roques C, Bouchez O, Eche C, Valiere S, Donnadieu C, Haffray P *et al*: **Characterization of a Y-specific duplication/insertion of the anti-Mullerian hormone type II receptor gene based on a chromosome-scale genome assembly of yellow perch, Perca flavescens**. *Mol Ecol Resour* 2020, **20**(2):531-543.
34. Futschik A, Schlötterer C: **The Next Generation of Molecular Markers From Massively Parallel Sequencing of Pooled DNA Samples**. *Genetics* 2010, **186**(1):207-218.
35. Li H, Durbin R: **Fast and accurate short read alignment with Burrows–Wheeler transform**. *Bioinformatics* 2009, **25**(14):1754-1760.
36. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM: **Twelve years of SAMtools and BCFtools**. *Gigascience* 2021, **10**(2):giab008.
37. Kofler R, Pandey RV, Schlötterer C: **PoPoolation2: identifying differentiation**

between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* 2011, **27**(24):3435-3436.
38. Thorvaldsdottir H, Robinson JT, Mesirov JP: **Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration**. *Brief Bioinform* 2013, **14**(2):178-192.
39. Zhang X, Wagner S, Holleley CE, Deakin JE, Matsubara K, Deveson IW, O'Meally D, Patel HR, Ezaz T, Li Z *et al*: **Sex-specific splicing of Z- and W-borne nr5a1 alleles suggests sex determination is controlled by chromosome conformation**. *Proc Natl Acad Sci U S A* 2022, **119**(4).
40. Wen M, Feron R, Pan Q, Guguin J, Jouanno E, Herpin A, Klopp C, Cabau C, Zahm M, Parrinello H *et al*: **Sex chromosome and sex locus characterization in goldfish, Carassius auratus (Linnaeus, 1758)**. *BMC Genomics* 2020, **21**(1):552.
41. Yano A, Guyomard R, Nicol B, Jouanno E, Quillet E, Klopp C, Cabau C, Bouchez O, Fostier A, Guiguen Y: **An immune-related gene evolved into the master sex-determining gene in rainbow trout, Oncorhynchus mykiss**. *Curr Biol* 2012, **22**(15):1423-1428.
42. Matsuda M, Nagahama Y, Shinomiya A, Sato T, Matsuda C, Kobayashi T, Morrey CE, Shibata N, Asakawa S, Shimizu N: **DMY is a Y-specific DM-domain gene required for male development in the medaka fish**. *Nature* 2002, **417**(6888):559.
43. Cui Z, Liu Y, Wang W, Wang Q, Zhang N, Lin F, Wang N, Shao C, Dong Z, Li Y: **Genome editing reveals dmrt1 as an essential male sex-determining gene in Chinese tongue sole (Cynoglossus semilaevis)**. *Scientific reports* 2017, **7**:42213.
44. Schlötterer C, Tobler R, Kofler R, Nolte V: **Sequencing pools of individuals—mining genome-wide polymorphism data without big funding**. *Nature Reviews Genetics* 2014, **15**(11):749-763.
45. Pan Q, Feron R, Yano A, Guyomard R, Jouanno E, Vigouroux E, Wen M, Busnel J-M, Bobe J, Concordet J-P: **Identification of the master sex determining gene in Northern pike (Esox lucius) reveals restricted sex chromosome differentiation**. *PLoS genetics* 2019, **15**(8):e1008013.
46. Gammerdinger WJ, Conte MA, Baroiller J-F, D'cotta H, Kocher TD: **Comparative analysis of a sex chromosome from the blackchin tilapia, Sarotherodon melanotheron**. *BMC genomics* 2016, **17**(1):808.
47. Hendry CI, Martin-Robichaud DJ, Benfey TJ: **Hormonal sex reversal of Atlantic halibut (Hippoglossus hippoglossus L.)**. *Aquaculture* 2003, **219**(1-4):769-781.
48. Dan C, Mei J, Wang D, Gui JF: **Genetic differentiation and efficient sex-specific marker development of a pair of Y- and X-linked markers in yellow catfish**. *Int J Biol Sci* 2013, **9**(10):1043-1049.
49. Gammerdinger WJ, Conte MA, Acquah EA, Roberts RB, Kocher TD: **Structure and decay of a proto-Y region in Tilapia, Oreochromis niloticus**. *Bmc Genomics* 2014, **15**(1):975.
50. Qu M, Liu Y, Zhang Y, Wan S, Ravi V, Qin G, Jiang H, Wang X, Zhang H, Zhang B *et al*: **Seadragon genome analysis provides insights into its phenotype and sex determination locus**. 2021, **7**(34):eabg5196.
51. Kamiya T, Kai W, Tasumi S, Oka A, Matsunaga T, Mizuno N, Fujita M, Suetake H, Suzuki S, Hosoya S *et al*: **A trans-species missense SNP in Amhr2 is associated with sex determination in the tiger pufferfish, Takifugu rubripes (fugu)**. *PLoS Genet* 2012, **8**(7):e1002798.
52. Ponnikas S, Sigeman H, Abbott JK, Hansson B: **Why Do Sex Chromosomes Stop Recombining?** *Trends Genet* 2018, **34**(7):492-503.
53. Herpin A, Schartl M: **Plasticity of gene‐regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers, and usurpators**. *EMBO reports* 2015, **16**(10):1260-1274.
54. Nanda I, Kondo M, Hornung U, Asakawa S, Winkler C, Shimizu A, Shan Z, Haaf T, Shimizu N, Shima A *et al*: **A duplicated copy of DMRT1 in the sex-determining region of the Y chromosome of the medaka, Oryzias latipes**. *Proc Natl Acad Sci U*

S A 2002, **99**(18):11778-11783.
55. Reichwald K, Petzold A, Koch P, Downie Bryan R, Hartmann N, Pietsch S, Baumgart M, Chalopin D, Felder M, Bens M *et al*: **Insights into Sex Chromosome Evolution and Aging from the Genome of a Short-Lived Fish**. *Cell* 2015, **163**(6):1527-1538.
56. Rondeau EB, Messmer AM, Sanderson DS, Jantzen SG, von Schalburg KR, Minkley DR, Leong JS, Macdonald GM, Davidsen AE, Parker WA: **Genomics of sablefish (Anoplopoma fimbria): expressed genes, mitochondrial phylogeny, linkage map and identification of a putative sex gene**. *BMC genomics* 2013, **14**(1):452.
57. Myosho T, Otake H, Masuyama H, Matsuda M, Kuroki Y, Fujiyama A, Naruse K, Hamaguchi S, Sakaizumi M: **Tracing the emergence of a novel sex-determining gene in medaka, Oryzias luzonensis**. *Genetics* 2012, **191**(1):163-170.
58. Hattori RS, Murai Y, Oura M, Masuda S, Majhi SK, Sakamoto T, Fernandino JI, Somoza GM, Yokota M, Strüssmann CA: **A Y-linked anti-Müllerian hormone duplication takes over a critical role in sex determination**. *Proceedings of the National Academy of Sciences* 2012, **109**(8):2955-2959.
59. Sandra G-E, Norma M-M, Fisheries: **Sexual determination and differentiation in teleost fish**. *Reviews in Fish Biology* 2010, **20**(1):101-121.
60. Takehana Y, Matsuda M, Myosho T, Suster ML, Kawakami K, Shin IT, Kohara Y, Kuroki Y, Toyoda A, Fujiyama A *et al*: **Co-option of Sox3 as the male-determining factor on the Y chromosome in the fish Oryzias dancena**. *Nat Commun* 2014, **5**:4157.
61. Tao W, Xu L, Zhao L, Zhu Z, Wu X, Min Q, Wang D, Zhou Q: **High-quality chromosome-level genomes of two tilapia species reveal their evolution of repeat sequences and sex chromosomes**. 2021, **21**(2):543-560.

Tables

Table 1. Summary of the male and female channel catfish pool-sequencing datasets.

Sample	Clean reads	Clean reads (Gb)	Mapped ratio (%)	Q20 (%)	Q30 (%)	GC (%)	Depth (x)
Male pool	266,697,484	40.3	98.5	96.7%	91.9%	40	34
Female pool	231,722,384	35.0	98.3	96.8%	92.0%	40	30

Table S1. Sequences of the primers used for sex genotyping in channel catfish

names	**Primers Sequence (5' - 3')**	**PCR product (bp)**	**Genome location** HNU_Msal_1.0
Marker 1	Forward: CTGGAGACGAGTTCAGCACGAT Reverse: TCCAAATCGTAAGACAATAGCG	893 bp in males 596 bp in females	LG4: 17,847,495-17,848,945
Marker 2	Forward: CTTCTGCCCAATGTCTTAACCA Reverse: CTGGCACTACCATCAGATCCATAC	922 bp in males 755 bp in females	LG4: 24,650,445-24,651,895
Marker 3	Forward: AAAATGGCACCTGTAAACTGAA Reverse: GCTAACTGCCCTAACAACCTCA	944 bp in males 611 bp in females	LG4: 22242899-22244349
SNP 1	Forward: GCAGAGGTTAGCAGGCAGAGTC Reverse: AGAGTGTCACGCACAGATTTCG	330 bp in males and in females	LG4: 24,736,909-24,737,239
SNP 2	Forward: GCCTTATTTGGATGCTGCCTGTG Reverse: ACGTGATGCCTGTCGTCTTTGC	403 bp in males and in females	LG4: 24,737,586-24,737,989

Three primer pairs (marker 1 to 3) flanking a male- and Y-specific insertion and two primer pairs for surrounding a male-specific SNP (SNP 1 and 2) were designed based on our Pool-Seq analyses of the channel catfish SDR. Linkage groups (LGs) names and nucleotide position of each amplicons are given in the genome location column.

Table S2. Summary of sex-specific SNPs distribution in SDR in YY channel catfish

	Male SNPs	Female SNPs
Category		

	Male coding SNPs	Male non-coding SNPs	Female coding SNPs	Female non-coding SNPs
SNPs number	239	9,952	0	18
Total SNPs	10,191		18	

Figure legends

Figure 1. Pool-sequencing results of male and female identify a male heterogametic system with a large sex determination region on LG04 using YY channel fish as reference. A 100kb sliding window with an output point every 10kb was used to calculate overall SNPs and depth. All the 29 linkage groups (LGs) were labelled with their LG number with all unplaced scaffolds combined together. (**A**) The distribution of relative male depth. (**B**) The distribution of relative female depth. (**C**) The distribution of male-SNPs showing a region with enriched male-specific SNPs on LG4. (**D**) The distribution of female-SNPs.

Figure 2. Characterization of sex chromosome and sex-locus in channel catfish. Overall F_{ST} (**A**), sex-specific SNPs (**B**) and male / female relative read depth (**C**) were calculated using a 100 kb sliding window with an output point every 10 kb. A significant sex-biased signal with a high F_{ST} was observed on LG04 within a 8.3 Mb sex determining region (SR, highlighted with a black box on (**B**)) characterized by high density of male-specific SNPs (**B**). The *bcar1* gene on the edge of the SDR is shown by an orange arrow. Female- and male-specific SNPs are respectively indicated by red and blue color. (**C**) Relative read depth ratio between male (blue) and female (red).

Figure 3. Sex specific fragments and sex-specific polymorphic nucleotides. (A) Male specific fragments were observed within/around SDR with no/few reads mapped in female pool and half average depth in male pool. Sex-specific primers were designed by spanning the male-specific fragments, generating different size PCR products between male and female. (**B**) PCR validation of three sex-specific markers presenting in agarose gel with 6 males (♂, left side) and 6 females (♀, right side). The additional analysis on 12 males and 26 females was shown in Figure S5. (**C**) Random male-specific SNPs from SDR were manually checked using IGV, and verified by Sanger sequencing.

Supplementary figures

Figure S1: Pool-sequencing results of male and female identify a male heterogametic system with a large sex determination region on LG04 using XX channel fish as reference. A 100kb sliding window with an output point every 10kb was used to calculate overall SNPs and depth. All the 29 linkage groups (LGs) were labelled with their LG number with all unplaced scaffolds combined together. (**A**) The distribution of relative male depth. (**B**) The distribution of relative female depth. (**C**) The distribution of male-SNPs showing a region with enriched male-specific SNPs on LG4. (**D**) The distribution of female-SNPs.

Figure S2: Significant genetic differences identified on LG4 of female assembly. Overall F_{ST} (A), sex-specific SNPs (B) and male / female relative read depth were calculated using a 100kb sliding window with an output point every 10kb. A significant sex-biased signal with high F_{ST} was observed on LG04 with a Mb sex determining region (SR, highlighted with a

black box on (**B**)) characterized by high density of male-specific SNPs (**B**). The *bcar1* gene on the SDR, but not at the SNP enriched region is shown by an orange arrow. Female- and male-specific SNPs are respectively indicated by red and blue color. (**C**) Relative read depth ratio male (blue) and female (red).

Figure S3: Dot plot alignment of X and Y sequences of channel catfish. Alignment of X and Y showed a clear synteny relationship outside the sex determination region (SDR, boxed) but many rearrangements with this SDR.

Figure S4: IGV plots of the male-specific (Y-specific) insertions. (**A-D**) Read depth differences with few/no reads mapped in females and a half mean coverage in males, characterize potential male-specific insertions (highlighted with an orange box) within the sex determination region of channel catfish.

Figure S5. Sex genotyping with primer sets flanking male-specific (Y-specific) insertions. Male and female symbols indicated phenotypic sex. M= DNA ladder marker

Figure S6: IGV plot of the sex-biased SNP on the first intron of *bcar1*.

Supplementary information

Supplemental excel file 1: Information of annotated genes in the channel catfish sex determination region of LG4.

Supplemental excel file 2: SNP distribution among genes from sex differentiated region in the channel catfish.

Supplemental excel file 3: Genome location of Potential male specific sequences.

Figures:

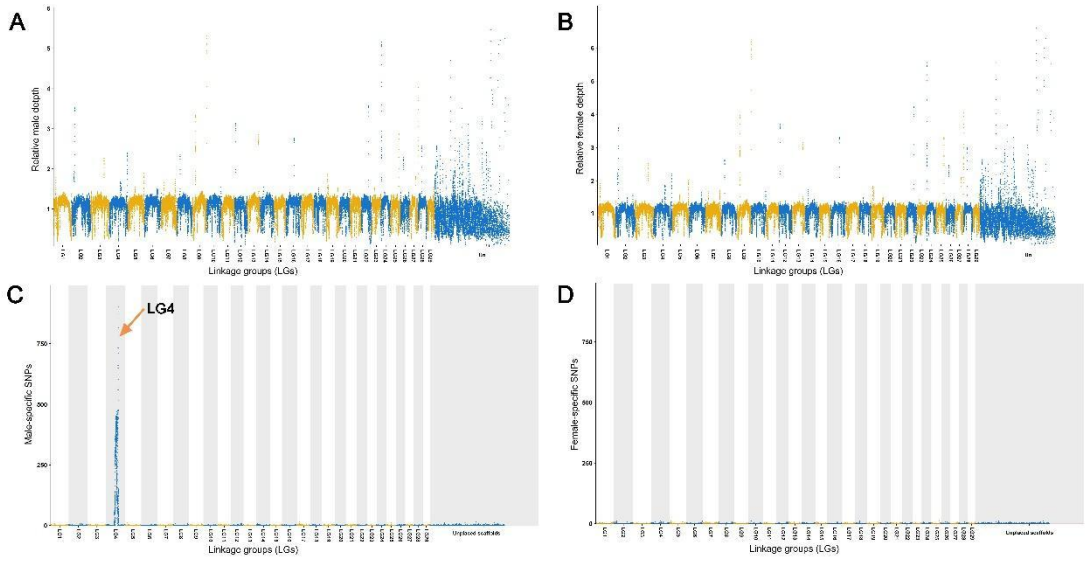

Figure 1

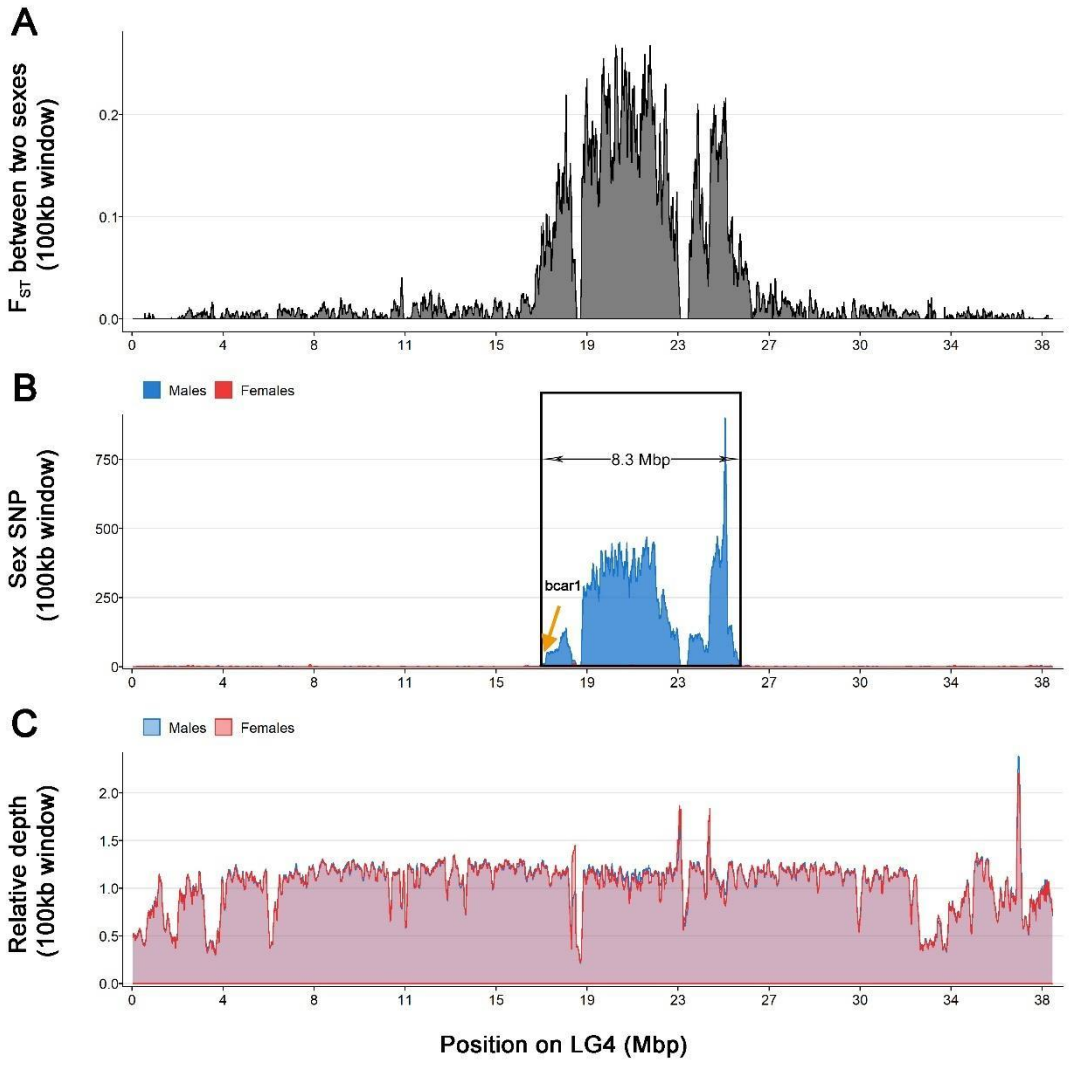

Figure 2

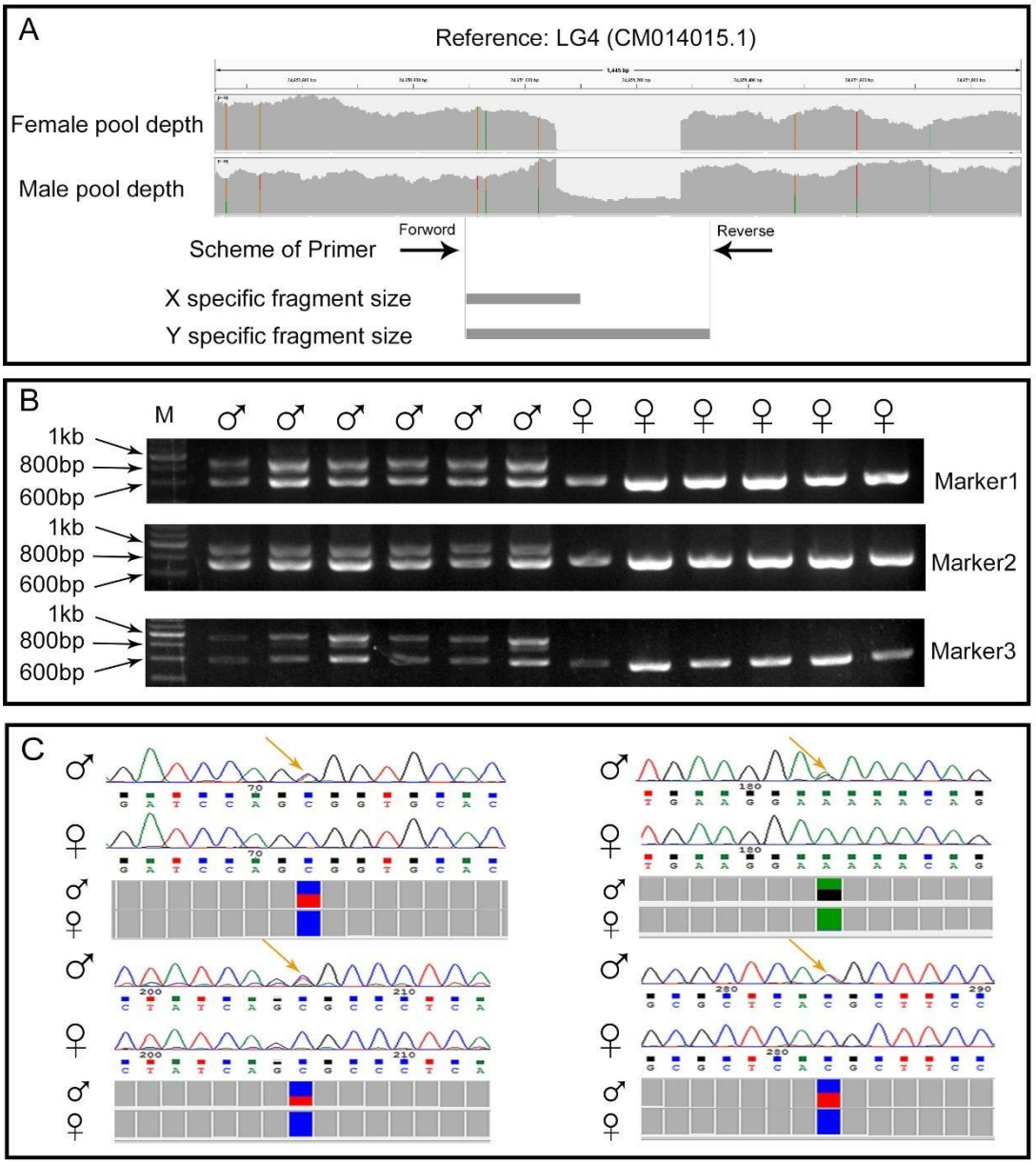

Figure 3

Supplementary figures

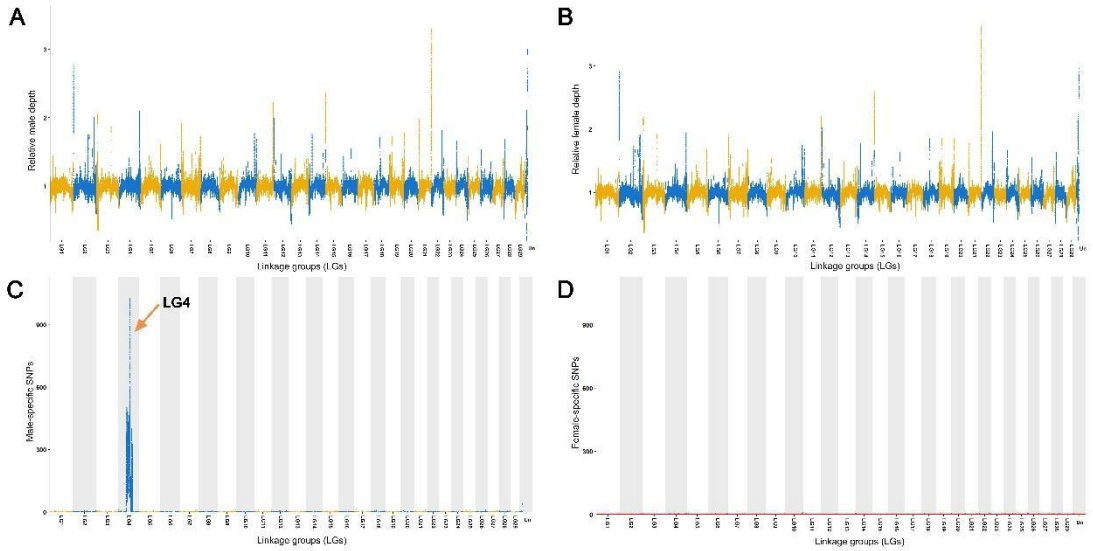

Figure S1

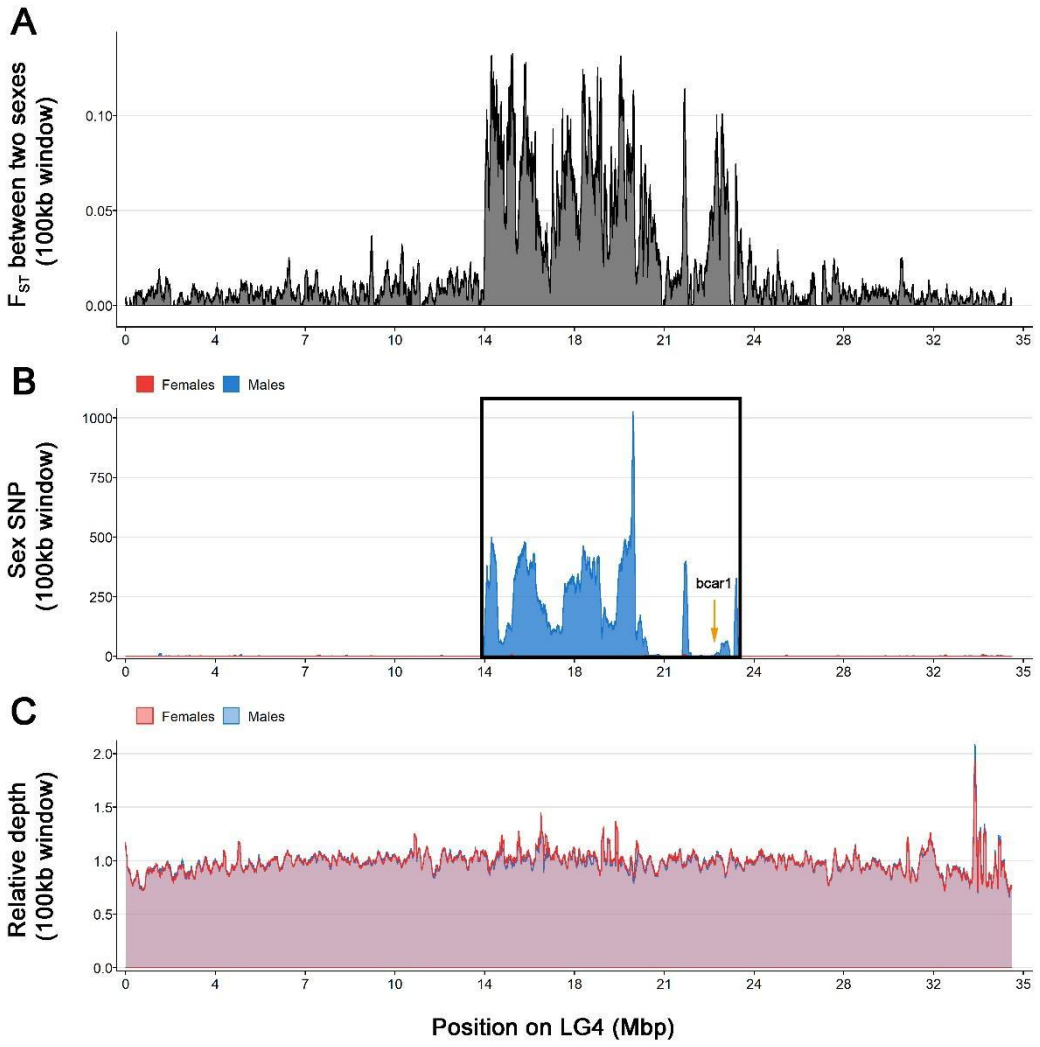

Figure S2

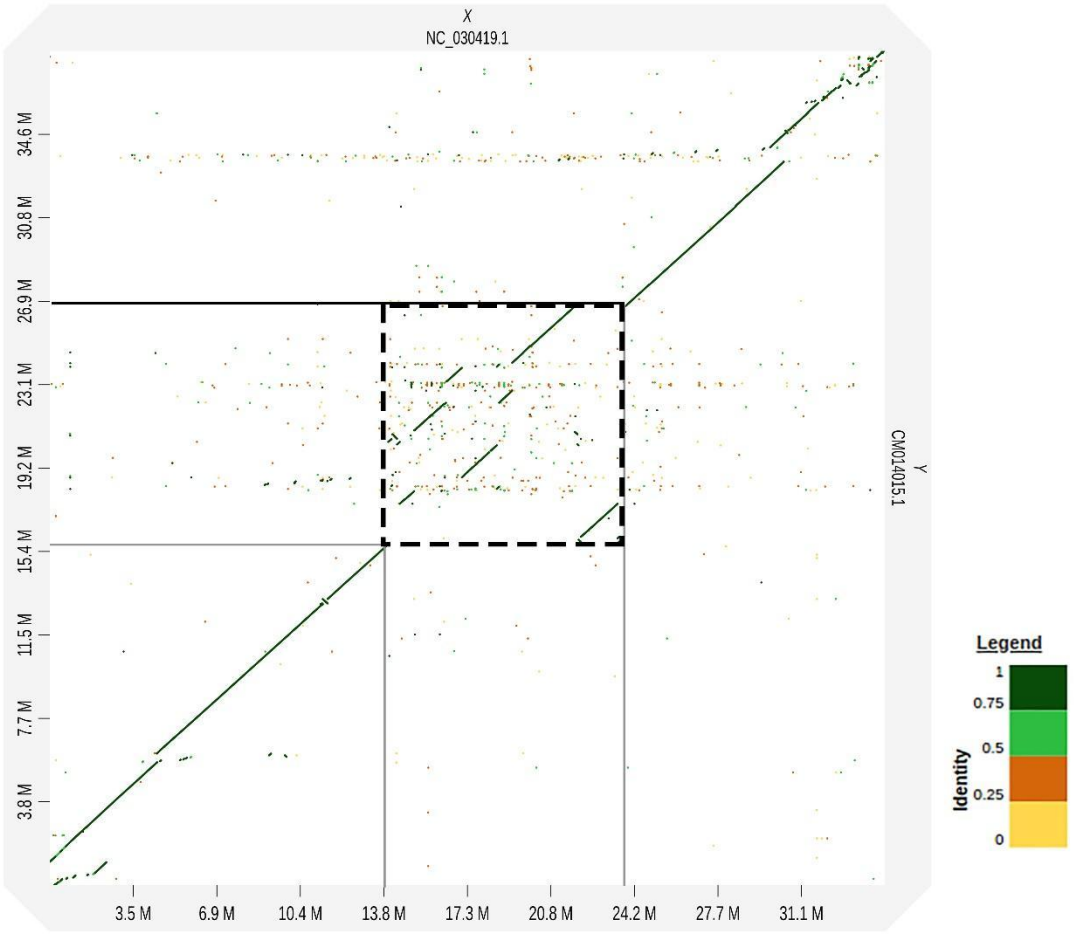
X
NC_030419.1
34.6 M
30.8 M
26.9 M
23.1 M
19.2 M
15.4 M
11.5 M
7.7 M
3.8 M
3.5 M
6.9 M
10.4 M
13.8 M
17.3 M
20.8 M
24.2 M
27.7 M
31.1 M
Y
CM014015.1
Legend
Identity
1
0.75
0.5
0.25
0

Figure S3

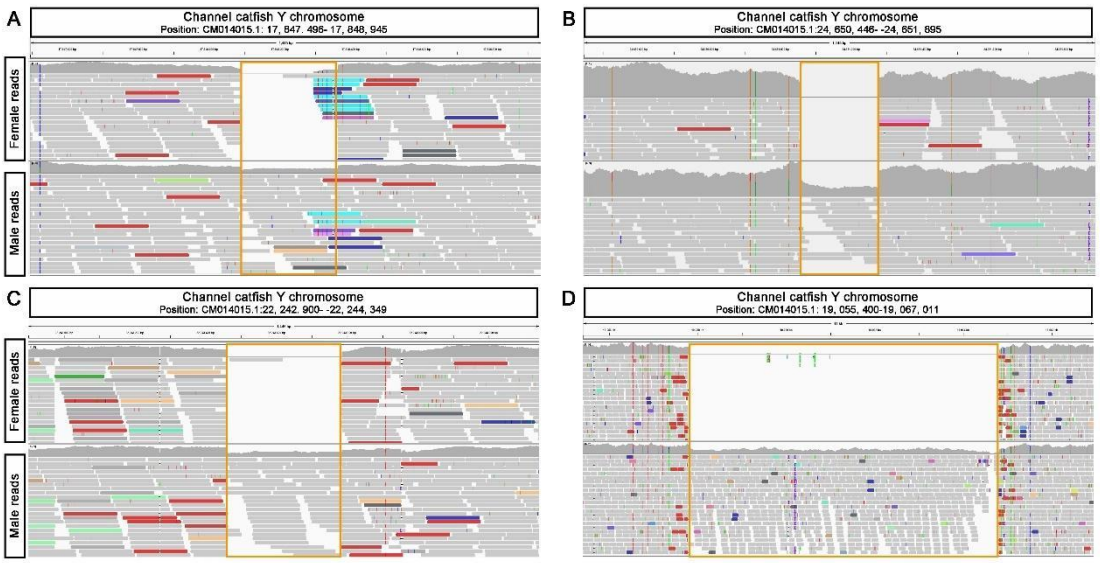

Figure S4

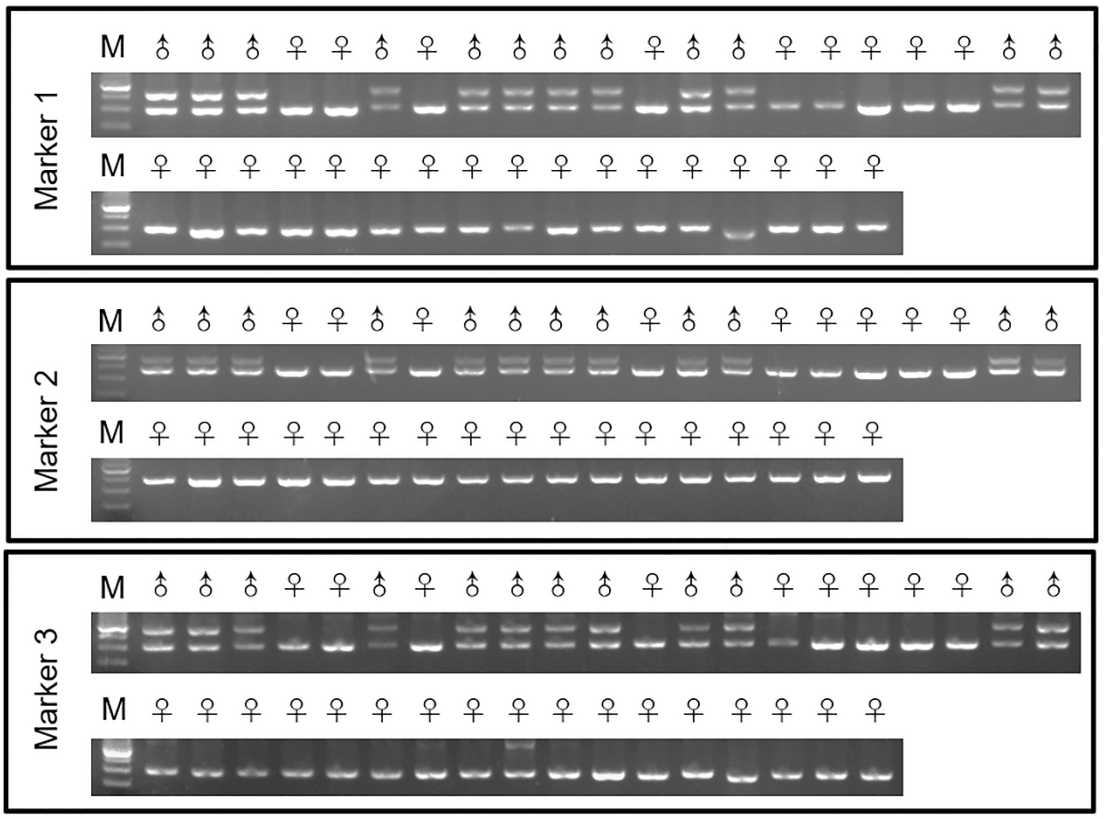

Figure S5

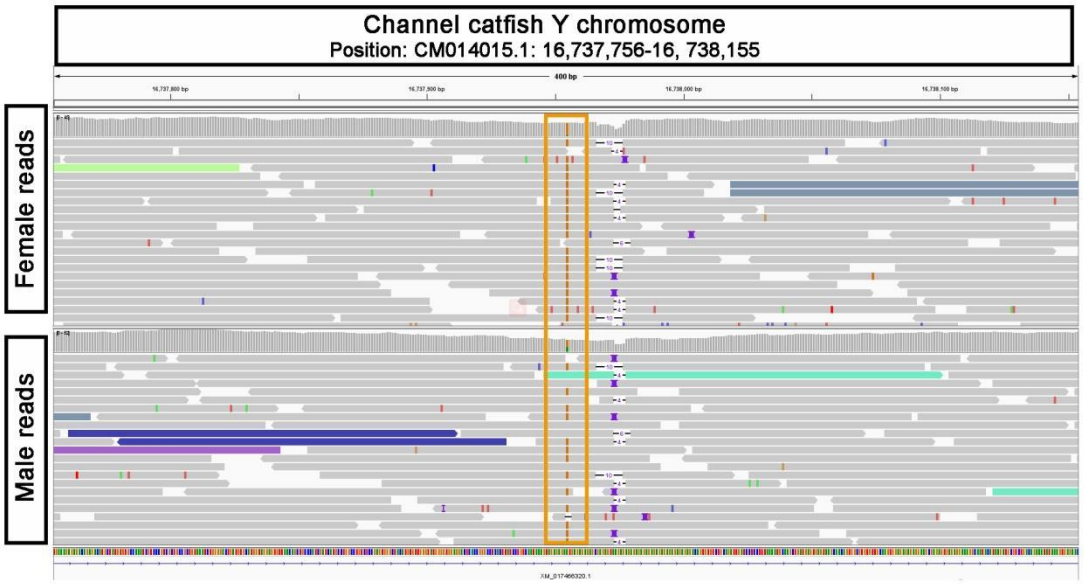
Channel catfish Y chromosome
Position: CM014015.1: 16,737,756-16, 738,155
400 bp
Female reads
Male reads

Figure S6