Construction of Genetic Linkage Maps From a Hybrid Family of Large Yellow Croaker (*Larimichthys crocea*)

Xinxiu Yu[1,2]*, Rajesh Joshi[3], Hans Magnus Gjøen[1], Zhenming Lv[2] and Matthew Kent[1]

[1]Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, Ås, Norway, [2]National Engineering Research Centre of Marine Facilities Aquaculture, Marine Science and Technology College, Zhejiang Ocean University, Zhoushan, China, [3]GenoMar Genetics AS, Oslo, Norway

Consensus and sex-specific genetic linkage maps for large yellow croaker (*Larimichthys crocea*) were constructed using samples from an F_1 family produced by crossing a Daiqu female and a Mindong male. A total of 20,147 single nucleotide polymorphisms (SNPs) by restriction site associated DNA sequencing were assigned to 24 linkage groups (LGs). The total length of the consensus map was 1757.4 centimorgan (cM) with an average marker interval of 0.09 cM. The total length of female and male linkage map was 1533.1 cM and 1279.2 cM, respectively. The average female-to-male map length ratio was 1.2 ± 0.23. Collapsed markers in the genetic maps were re-ordered according to their relative positions in the *ASM435267v1* genome assembly to produce integrated genetic linkage maps with 9885 SNPs distributed across the 24 LGs. The recombination pattern of most LGs showed sigmoidal patterns of recombination, with higher recombination in the middle and suppressed recombination at both ends, which corresponds with the presence of sub-telocentric and acrocentric chromosomes in the species. The average recombination rate in the integrated female and male maps was respectively 3.55 cM/Mb and 3.05 cM/Mb. In most LGs, higher recombination rates were found in the integrated female map, compared to the male map, except in LG12, LG16, LG21, LG22, and LG24. Recombination rate profiles within each LG differed between the male and the female, with distinct regions indicating potential recombination hotspots. Separate quantitative trait loci (QTL) and association analyses for growth related traits in 6 months fish were performed, however, no significant QTL was detected. The study indicates that there may be genetic differences between the two strains, which may have implications for the application of DNA-information in the further breeding schemes.

Keywords: large yellow croaker, RAD sequencing, linkage map, collinearity, recombination rate

INTRODUCTION

Large yellow croaker (*Larimichthys crocea*) has become an important aquaculture species in southeast China, where Mindong and Daiqu are the two major strains farmed. Artificial breeding of the Mindong strain started in 1985, while the Daiqu strain has been bred since 1999 (Chen et al., 2018). In 2019, the total production of large yellow croaker exceeded 220,000 tons and accounted for more than 12% of the cultured marine fish production of China (Yu X. J. et al., 2020).

The production is supported by several breeding programs, of which the majority are based on classical basic selection methods, like phenotypic selection (Chen et al., 2018). However, modern breeding approaches, such as marker assisted selection (MAS), can further enhance the genetic gain for economically important traits. MAS can greatly increase the efficiency if a sufficiently large QTL is detected, typically through QTL linkage mapping and association studies (Zenger et al., 2019). An outstanding example of MAS applied in aquaculture was the discovery of a QTL imparting resistance to infectious pancreatic necrosis (IPN) in Atlantic salmon, accounting for about 80% of the total variation in this trait (Moen et al., 2009; Houston et al., 2012). Information from the QTL was used in selective breeding to generate IPN resistant fish, which now dominate production in Norway, leading to a remarkable reduction in IPN outbreaks (Norris, 2017). Subsequent studies have provided functional genomics data indicating that mutations in the epithelial cadherin gene (cdh1) affect virion internalization (Moen et al., 2015), demonstrating the power of genomic tools to help reveal the mechanistic basis for important traits. Although MAS can be useful for some traits where major QTLs have been identified, most traits of economic importance in aquaculture species (i.e., production traits) are assumed to be polygenic, and often have low-to-moderate heritabilities (Zenger et al., 2019). As a result, application of MAS to improve these complex traits may be inefficient. For such polygenic traits, *genomic selection* is a viable alternative, based on genomic breeding values predicted on a genome-wide scale, allowing even small QTLs to contribute (Meuwissen et al., 2001). For large yellow croaker, the estimates of heritability for body weight (0.31 ± 0.06), body length (0.33 ± 0.06) and body height (0.41 ± 0.07) in 6 months fish, and the genetic correlations between them ranged from 0.74 to 0.95 (Yu X. X. et al., 2020).

High-throughput sequencing has transformed genetics by making it relatively easy to generate genome-wide genetic marker datasets, which are a prerequisite for QTL identification in MAS. Significant progress was made through the discovery of cost-effective restriction-site associated DNA sequencing (RADseq) based strategies (Baird et al., 2008). RADseq can generate medium density SNP resources and has been successfully used in various fish species for genetic linkage maps, QTL analysis and population genetics (Davey and Blaxter, 2010), e.g., in Atlantic salmon (Houston et al., 2012; Gonen et al., 2014), channel catfish (Li et al., 2014) and Nile tilapia (Palaiokostas et al., 2013).

Several genetic linkage maps for large yellow croaker have been developed using different approaches (**Supplementary Table S1**). The first two genetic linkage maps made publicly available were constructed using amplified fragment length polymorphism (AFLP; Ning et al., 2007) and simple sequence repeats (SSR; Ye et al., 2014). However, next-generation sequencing technologies have made detection of large numbers of genome-wide SNP markers relatively easy, and Ao et al. (2015) constructed a SNP genetic linkage map with a total length of 5451.3 cM using RADseq, while Xiao et al. (2015) constructed a genetic map of 2,632 cM using RNA sequencing (RNAseq) of expressed genes. More recently, Kong et al. (2019) constructed a

double-digest restriction-site associated DNA (ddRAD) based genetic map using 5261 SNPs with a total length of 1885.67 cM. Despite using different approaches, these SNP linkage maps have one thing in common, as they were all developed using only the Mindong strain.

Daiqu strain of large yellow croaker has been successfully cultured since 1999 and the aquaculture production is on an industrial scale (Chen et al., 2018). Most consumers prefer lean large yellow croaker, and the body shape has become an important economic trait (Dong et al., 2019). The Daiqu strain has better performance for this trait, as the ratio of body length and body height is significantly higher than for the Mindong strain (Huang et al., 2006). The Daiqu strain also has later sexual maturation and better tolerance to lower temperatures than the Mindong strain (Liu and Mitcheson, 2008; Miao et al., 2014). The offspring from a crossing between Mindong and Daiqu displayed significant heterosis in body shape and growth of fish after 526 days (Li et al., 2010). Our study therefore sought to develop a genetic linkage map in a crossed (F_1) family arising from these strains.

The aim of this study was to construct consensus and sex-specific linkage maps based on a hybrid family from the Daiqu and Mindong strains using RADseq, to compare the linkage map to the latest physical map (*ASM435267v1*), and to perform a QTL analysis and association analysis for growth-related traits.

MATERIALS AND METHODS

Mapping Family
A female (F_3 of wild Daiqu strain, from an aquaculture farm in Xiangshan, Zhejiang province) and a male (approx. F_{10} of wild Mindong strain, from an aquaculture farm in Fuding, Fujian province) large yellow croaker were crossed to generate a fullsib family (Yu X. et al., 2017). One-hundred and twenty offspring were randomly selected at 6 months, and the following growth traits were recorded: body weight (BW), body length (BL) and body height (BH). Fin clips were preserved in 99% ethanol and sent to BGI Genomics Company (Shenzhen, China) for sequencing.

RAD Sequencing and SNP Calling
Library preparation was performed by BGI according to Baird et al. (2008). In brief, individual genomic DNA samples were digested using the restriction enzyme *Pst I*, and the resulting fragments were ligated to a double-stranded Illumina sequencing primer containing a sample-specific barcode sequence. Libraries were then pooled and sheared by sonication, and fragments from 300–500 bp were separated by agarose gel electrophoresis and purified before ligating a Y-adapter to the sheared ends. Fragments including both barcode and Y-adapters were amplified with PCR to generate the final RAD libraries, which were then sequenced using a *Hiseq2000* platform to produce paired-end reads.

Raw reads were processed by BGI using the *Reseqtools* software package (https://github.com/BGI-shenzhen/Reseqtools) to remove adapter sequences and low-quality

reads, and to de-multiplex the pool. The retained reads were analysed and genotyped using *Stacks* (Catchen et al., 2013) and in-house analysis pipelines, and RAD-tags with too low (<2) or too high (>100) sequencing coverage were excluded.

SNP Filtering and Linkage Map Construction

SNPs missing in >10% of samples and minor allele frequency (MAF) < 0.05 were excluded using the PLINK software (Purcell et al., 2007). Markers were individually tested against the expected segregation ratio, based on parental genotypes, and those showing significant segregation distortion ($p < 0.05$, χ^2 test) were removed by PLINK.

The remaining SNPs were used to generate consensus and sex-specific maps using Lep-MAP2 software (Rastas et al., 2013). All SNP markers that passed filtering (n = 20,186) were used to produce the consensus map, while those markers polymorphic in the father (n = 11,684) or mother (n = 11,838) were used to construct their respective sex-specific linkage maps. LGs were developed using the *separate chromosomes* module, with a logarithm of odds (LOD) score ranging from 1 to 20. A LOD score of 9, which gave 24 LGs and the lowest number of single markers, was finally selected. The option *sizeLimit =100* was used to generate linkage groups of size ≥100 markers. The module *JoinSingles* could not assign any of the singular markers to any of the 24 LGs. Eventually, 20,147 SNPs were ordered using the *OrderMarkers* module, which assign the markers with paternal or maternal positions for the sex-specific maps. The option *sexAverage = 1* was applied during execution of *OrderMarkers* to get positions for the consensus map. To avoid the map distances being too long, especially when the number of markers per chromosome was much higher than the number of individuals, the parameter *minError = 0.15* was used. Finally, the *Kosambi* mapping function was used to calculate genetic distance between markers. The LG were numbered by the SNP size of each LG (i.e., the LG with the largest SNP number was labelled LG1). Illustrations of the consensus and sex-specific linkage maps were drawn using MapChart 2.32 (Voorrips, 2002).

QTL Analysis and Association Analysis

QTL analysis was initially performed using the QTL *IciMapping* software by the option *inclusive composite interval mapping with an additive effect* (ICIM-ADD) (Li et al., 2015; Meng et al., 2015). The LOD threshold for QTL significance of each trait was determined by a permutation test (1,000 replications) with a genome-wide significance level of 0.05. The permutation threshold method for QTL mapping estimates the null distribution of the genome-wide maximum LOD score by shuffling the phenotypes relative to the genotype data, breaking the association between the phenotype and the genotypes (Churchill and Doerge, 1994). The genome-wide LOD thresholds are calculated based on the 1-α quantiles of the genome-wide maximum LOD scores obtained from the permutations, where α is the significance level (α = 0.05 in our case).

As a complementary method for QTL mapping, a genome-wide association study (GWAS) was performed using SNPs subjected to a more stringent quality filtering than that was applied for linkage mapping to ensure a high QTL identification accuracy. Using PLINK, individuals displaying more than 5% missing genotypes were removed. Also, SNPs were removed in cases where missing genotypes >5% across samples and Hardy-Weinberg p value (Fishers exact test) < 10^{-9}. The final SNP set used for GWAS thus included 16,570 SNPs from 74 individuals.

The genome-wide association analysis was performed using a mixed linear model equation on BW, BL and BH by the *Genome-wide Complex Trait Analysis* (GCTA) program, with the *-mlma* function (Yang et al., 2011). The following model was used:

$$\mathbf{y} = \mathbf{a} + \mathbf{bx} + \mathbf{g} + \mathbf{e}$$

where **y** is the phenotypes (BW, BL, BH), a is the overall mean for each trait, **b** is the additive genetic effect of the candidate SNP to be tested for association, x is the incidence matrix for the candidate SNPs, **g** is the polygenic effect and **e** is the vector of random residual effects.

SNPs were considered genome wide significant when exceeding the Bonferroni threshold for multiple testing (α = 0.05) of 0.05/tg = 3.017502×10^{-6}, where tg = 16,570 (total number of genome-wide SNPs); and SNPs were graded as chromosome-wide significant when Bonferroni threshold for multiple testing (α = 0.05) surpassed 0.05/tc = 7.246377×10^{-5}, where tc = 690 (average number of SNPs per chromosome). The genome-wide significant threshold used in this study was $p \le 3.017502 \times 10^{-6}$ (–log10 (P) = 5.52), while chromosome-wide significant threshold was $p \le 7.246377 \times 10^{-5}$ (–log10 (P) = 4.14). SNPs were visualised along the linkage groups using the Manhattan function in the R package QQMAN (Turner, 2014).
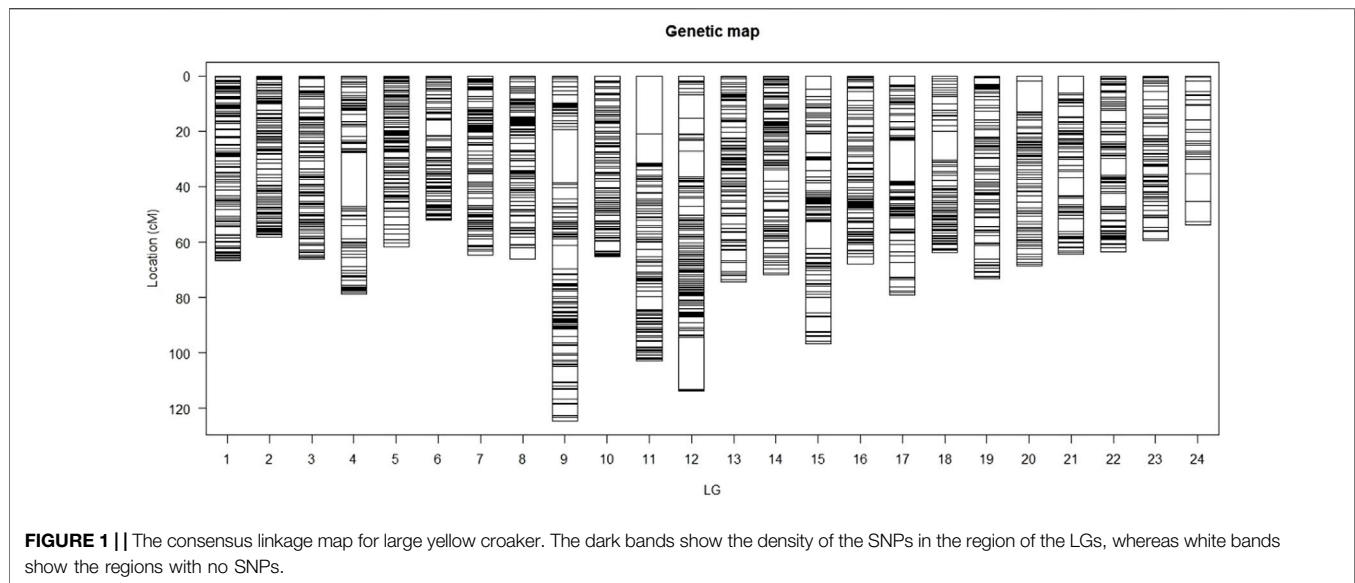
Collinearity Analysis: Genetic vs. Physical Map

To explore the level of agreement between our consensus genetic map and a recently published physical map, the large yellow croaker assembly, *ASM435267v1* (GenBank ID GCA_004352675.1), the RAD-tag sequences (82 bp) from the consensus linkage map were aligned to *ASM435267v1* using BLASTN (https://blast.ncbi.nlm.nih.gov/Blast.cgi) with the following parameters: *expect value e ≤ 1 × 10⁻¹⁵, identity ≥ 95%, matched length ≥ 81 bp, mismatches ≤ 1 and gap open = 0*. If a query sequence hit two or more loci in the physical assembly and the difference between the 1st and 2nd smallest e-values was greater than 10^3, the 1st smallest e-value was chosen to define the hit. Finally, 9885 SNPs from the consensus map hit the physical map.

The relative positioning of RAD sequences in the genetic map and the physical map were graphically presented using *shinyCircos* (Yu Y. et al., 2017). The marker positions on genetic map were multiplied by 4×10^5 for better visualisation of the Circos plot.

Adjusting Genetic Maps Based on the Physical Map

The collinearity analysis highlighted 107 SNPs whose assignment to LGs disagreed with their physical assignment to chromosomes.

FIGURE 1 | The consensus linkage map for large yellow croaker. The dark bands show the density of the SNPs in the region of the LGs, whereas white bands show the regions with no SNPs.

Examples of this were seen in all LGs and, when detected, SNPs were reordered according to the physical map. The adjusted SNP order in each linkage group was used as an input of *evaluateOrder* option in the *OrderMarkers* module of Lep-MAP2, and the genetic distances were recalculated using the *Kosambi* mapping function. The integrated consensus and sex-specific linkage maps were drawn using *MapChart 2.32* (Voorrips, 2002).

Scatter plots were generated between the integrated linkage maps in cM distances and the physical map in Mb distances by using the *ggplot2* package in R. Recombination rates throughout the genome in the integrated female and male genetic maps were estimated using *MareyMap* online (Siberchicot et al., 2017) with a computed sliding window size of 3.37 Mb. The threshold markers number in a window was set to 8, the default value. The recombination rate changes throughout the genome in the integrated female and male maps were visualised by using the *ggplot2* package in R.

RESULTS

Sequencing and SNP Filtering
Approximately 1.9 billion reads were produced after sequencing two parents and 120 offspring, with each individual contributing roughly 15 ± 2.9 million reads. After reads filtering and RAD-tag SNP detection, approximately 370,000 variants were detected within each individual. The average heterozygosity rate was 32.5%. After filtering for segregation errors, MAF and missing genotypes, a final set of 20,186 SNP markers was used for linkage map construction.

Linkage Map Construction
The SNPs were assigned to 24 LGs, in accordance with the haploid chromosome number (Lou et al., 2015). The consensus map (**Figure 1**; **Table1**) covered 1757.4 cM, with individual linkage group lengths ranging from 51.9 cM (LG6)

to 124.6 cM (LG9). The number of markers per linkage group varied from 243 to 1,230, with an average genetic distance between markers of 0.09 cM and a standard deviation of 0.037 (**Table1**).

Separate male and female maps were constructed using segregating (heterozygous) markers from each parent. The total length of the male linkage map was 1,279 cM, and the total length of the female map was 1,533 cM (**Table1**, **Supplementary Figure S1**). In the female map, LG length ranged from 39.8 cM (LG22) to 86.7 cM (LG3), and the SNP number per LG varied from 114 to 783. In the male map, the length of each LG varied from 43.1 cM (LG2) to 83.5 cM (LG4), and the SNP number in each LG varied from 158 to 745. The average distance between markers for female and male is thus 0.11 and 0.13 cM, respectively. The female-to-male length ratio ranged from 0.7 (LG22) to 1.6 (LG3), with an average of 1.2 ± 0.23; most LGs in the female map were larger than those in the male map, with the exceptions of LG4, LG19, LG21 and LG24.

QTL Analysis and Association Analysis for Growth Traits
The growth traits, BW, BL and BH, recorded in 120 offspring at 6 months of age, are presented in **Table 2**. In the QTL analysis, the LOD threshold used was 8.81 for BW, 7.35 for BL, and 18.71 for BH. However, no QTL was above the LOD threshold for any of the growth traits, BW, BL or BH (**Supplementary Figure S3**). In the GWAS analysis, the estimated genomic heritabilities for the three traits were close to zero (**Table 3**). A total of 16,570 SNPs from 74 recorded individuals were used, however, no SNPs crossed the genome or chromosome-wide significant level (**Supplementary Figure S4**).

Collinearity Analysis
In total, 9885 SNPs from the consensus map hit the physical map (*ASM435267v1*), but there were 107 SNPs hitting non-

TABLE 1 | Key figures for the genetic linkage maps of large yellow croaker.

LG	Consensus			Female		Male	
	No. of markers	Size (cM)	Average distance (cM)	No. of markers	Size (cM)	No. of markers	Size (cM)
1	1,230	66.8	0.05	613	78.2	745	51.8
2	1,125	58.3	0.05	783	65.7	716	43.1
3	1,029	66.0	0.06	660	86.7	574	55.5
4	1,000	78.7	0.08	654	66.8	503	83.5
5	997	61.8	0.06	626	72.6	629	49.4
6	973	51.9	0.05	543	54.2	548	43.8
7	967	64.6	0.07	532	67.2	558	50.1
8	949	66.3	0.07	458	70.2	647	60.4
9	899	124.6	0.14	609	68.3	375	58.0
10	875	65.2	0.07	484	68.6	505	49.4
11	850	102.8	0.12	475	64.4	468	51.8
12	837	113.7	0.14	533	68.6	410	59.3
13	831	74.4	0.09	480	58.5	480	52.8
14	817	71.9	0.09	589	71.8	377	50.7
15	814	96.6	0.12	297	70.4	589	52.9
16	791	68.0	0.09	387	61.2	526	55.2
17	789	79.0	0.10	417	70.4	476	45.6
18	789	63.9	0.08	616	55.2	566	48.6
19	787	73.2	0.09	589	54.2	283	56.6
20	741	68.4	0.09	539	54.2	358	46.1
21	693	64.5	0.09	378	50.5	392	55.3
22	689	63.6	0.09	237	39.8	528	58.0
23	432	59.3	0.14	225	65.7	273	50.5
24	243	53.9	0.22	114	49.8	158	50.8
Total	20147	1757.4	0.09	11838	1533.1	11684	1,279.2

TABLE 2 | Mean ± SD, range and coefficient of variation (CV) of growth traits at 6 months.

	Body weight, g	Body length, cm	Body height, cm
Mean ± SD	45.2 ± 13.8	13.8 ± 1.6	3.6 ± 0.5
Range	20.8–89.3	9.5–17.7	2.8–7.6
CV (%)	31	11	15

corresponding chromosomes. A collinearity analysis, comparing the consensus linkage and physical maps, was performed (**Figure 2**). The average correlation coefficient between the genetic map and the physical map was 0.78 ± 0.16 (**Supplementary Table S2**). Each LG matches well with its corresponding chromosome of the physical map, with an average matching percentage of 98.92 ± 1.5%. There were 7 LGs that showed no mismatch between the genetic map and the physical map; LG7, LG9, LG10, LG17, LG18, LG20 and LG24 (**Supplementary Table S3**).

Integration of Physical and Genetic Maps

SNP position information based on the *ASM435267v1* genome assembly was used to produce the physically informed consensus, female and male linkage maps (**Supplementary Figure S2**). A summary of the integrated maps is shown in **Supplementary Table S3**. A comparison of map positions between the integrated genetic and physical maps for different LGs is shown in **Figure 3**, in which most LGs exhibited sigmoidal patterns of recombination, with greater recombination rates toward the middle and low

recombination rates toward the ends of the chromosomes. Large gaps or jumps can be seen in some of the plots, viewed from the *x*-axis or from the *y*-axis (**Supplementary Table S4**). Viewed from the *x*-axis, representing the physical position, large gaps were observed on LG24 (2.12 Mb) and LG1 (1.9 Mb), whereas viewed from the *y*-axis, representing the genetic position, large gaps were observed e.g., in LG4 (36.36 cM) of the integrated male map and in LG21 (50.37 cM) of the integrated female map. The markedly large jump downward in LG23 of the integrated consensus map was due to the fragmented linkage group LG23.1 assigned to LG23 in this case. Recombination rates of the three integrated maps are shown in **Supplementary Table S3**, and the recombination rate variation comparison of integrated female and male maps is visualised in **Figure 4**. The average recombination rate in the female was 3.55 cM/Mb whereas it in the male was 3.05 cM/Mb. The pattern of the recombination rates was different between male and female in some LGs, as there was a higher recombination rate for the male than for the female in the beginning of some LGs (e.g., LG09), whereas in other LGs the pattern was just opposite (e.g., LG20).

DISCUSSION

Linkage Map Construction and Collinearity Analysis

The total genetic length of the consensus linkage map in our study was 1757.4 cM. The genetic map length (1885.67 cM) using the Mindong strain only, found by Kong et al. (2019), is slightly larger than that of our study. However, the linkage map length, also

TABLE 3 | Estimates of variance components and heritability with standard errors (in parenthesis) using the genomic relationship matrix in GWAS by GCTA.

Traits	σ_g^2	σ_e^2	σ_P^2	Genomic h²
BW	0.00022 (48.66)	194.5570 (42.94)	194.5572 (35.96)	0.000001 (0.25)
BL	0.000003 (0.69)	2.925069 (0.70)	2.925072 (0.52)	0.000001 (0.24)
BH	0 (0.08)	0.377716 (0.09)	0.37716 (0.07)	0.000001 (0.22)

σ_g^2, Genetic variance; σ_P^2, Phenotypic variance; σ_e^2, Residual variance; h², Heritability.
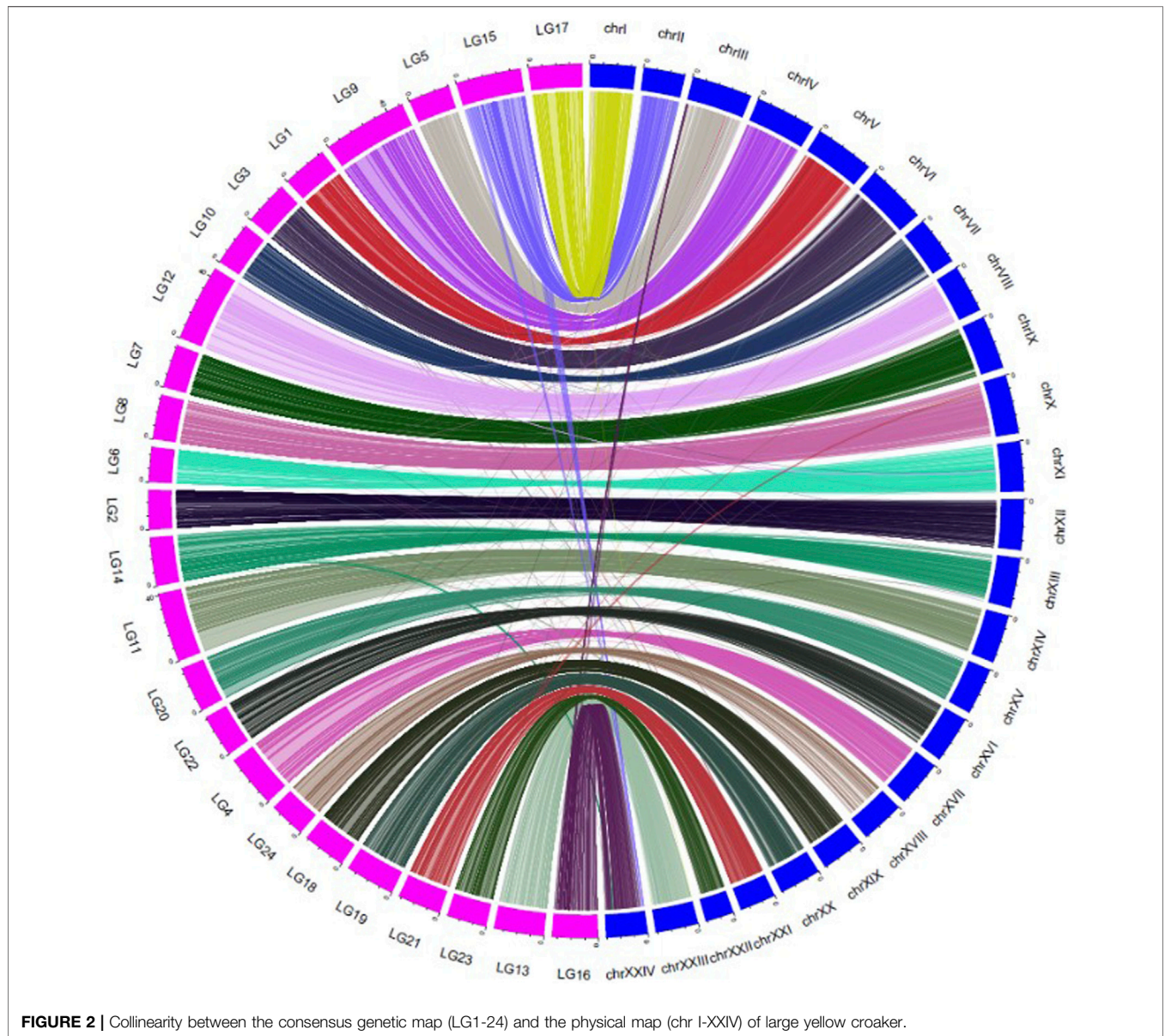
FIGURE 2 | Collinearity between the consensus genetic map (LG1-24) and the physical map (chr I-XXIV) of large yellow croaker.

using Mindong strain, found by Ao et al. (2015) was 5451.3 cM, is much larger than in our study. The differences in total genetic length could be caused by the mapping family used in our study, which was a cross between Mindong strain and Daiqu strain. Suppressed recombination rates have also been reported in rainbow-Yellowstone cutthroat trout (*Oncorhynchus clarkii bouvieri*) hybrids, there explained by chromosome

rearrangements (Ostberg et al., 2013). And in a hybrid cross of Human Pathogenic Fungus, *Cryptococcus neoformans*, the linkage map length (197 cM) was much shorter than those (1,356.3 cM) observed in a single strain (Sun and Xu, 2007).
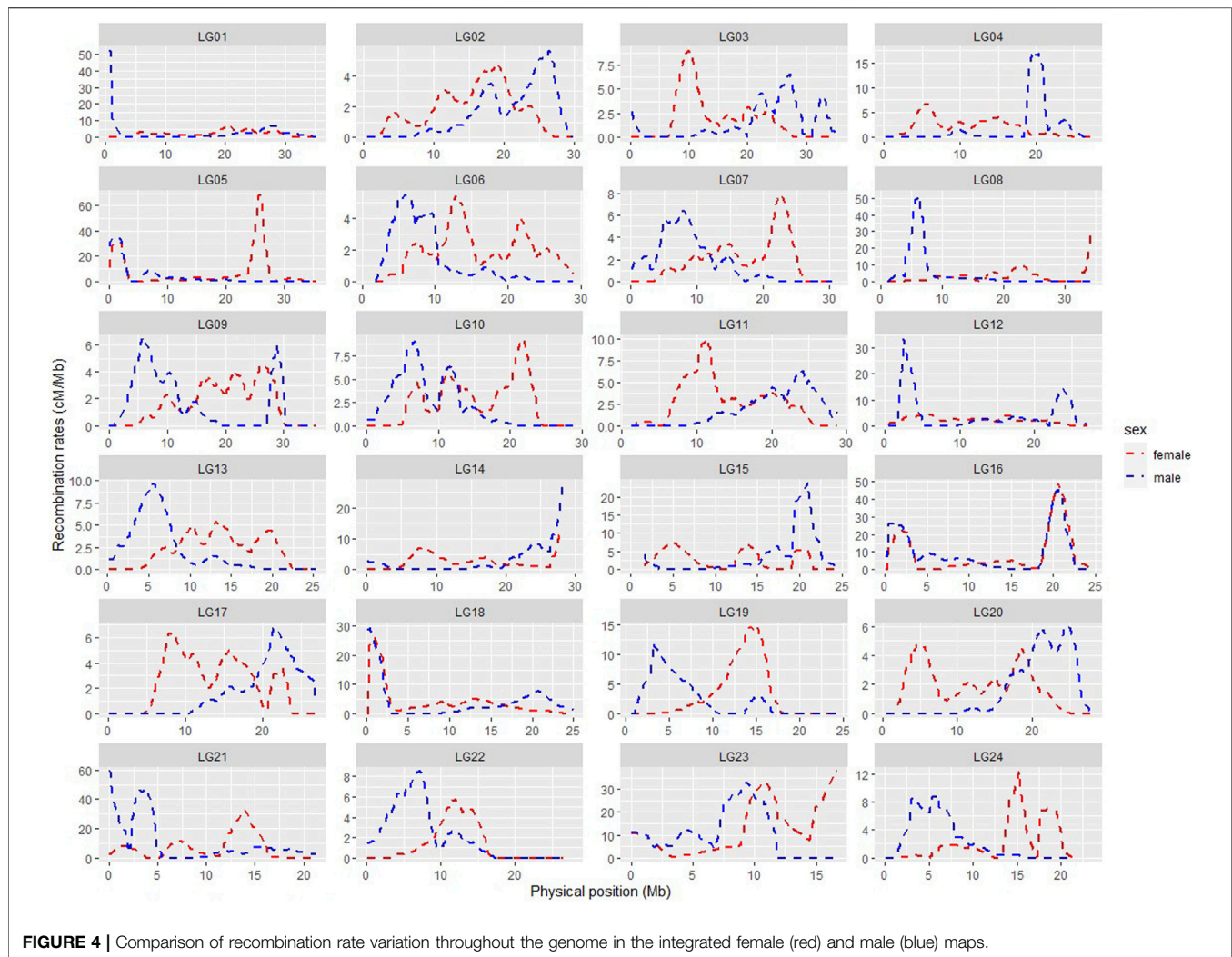
The average female-to-male map length ratio was 1.2 ± 0.23 in our study, indicating more recombination events happening in females, which is consistent with an earlier study in large yellow

FIGURE 3 | Scatter plots showing SNP linkage map positions (cM; *y*-axis) versus physical positions (Mb; *x*-axis) for integrated female (red), male (blue) and consensus (black) genetic maps. (Fragmented linkage group LG23.1 was assigned to LG 23.)
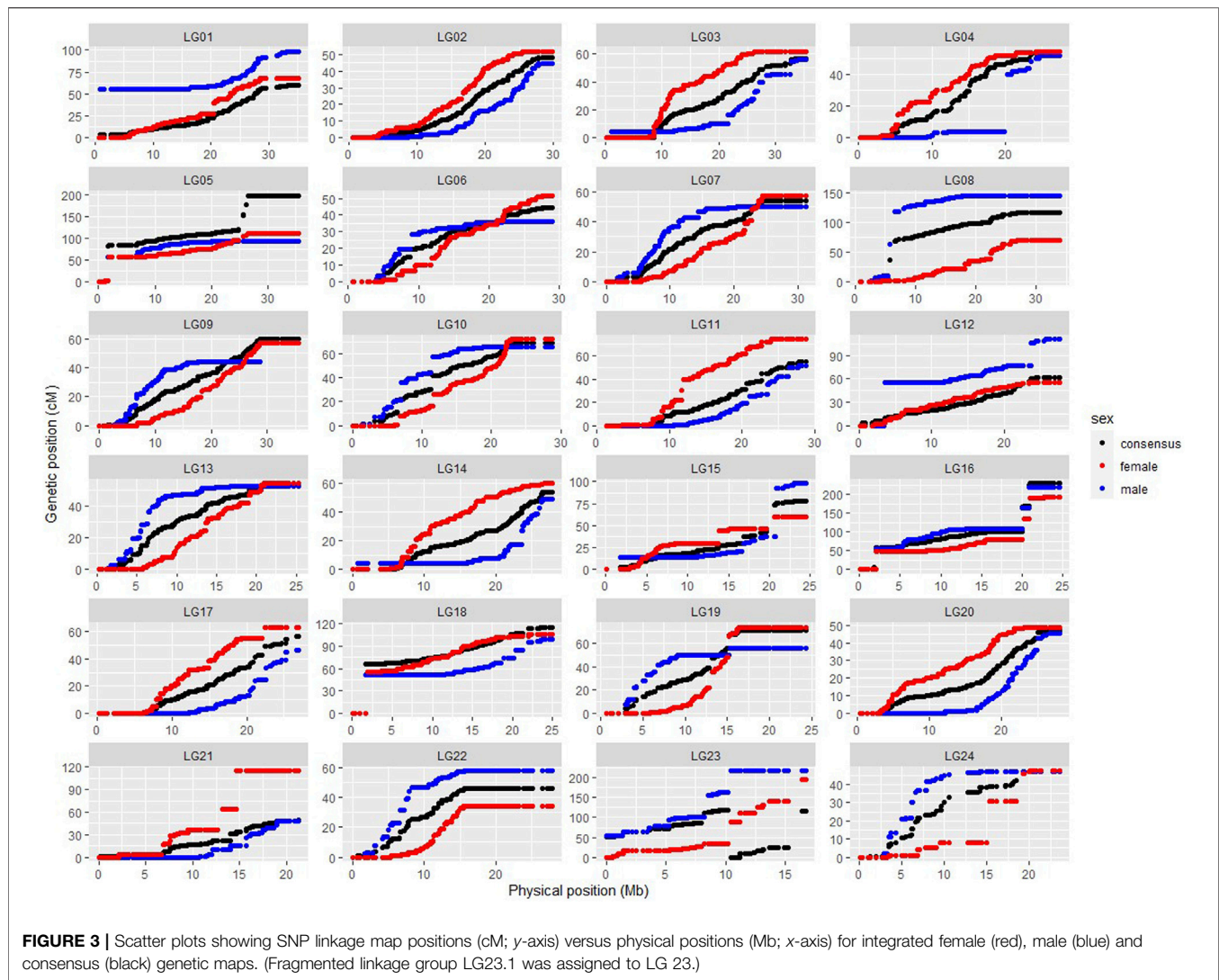
croaker by Ning et al. (2007). The phenomenon of heterochiasmy, i.e., sex differences in recombination rates between the two sexes, has been found in many fish species. Higher recombination rate in female fish, as in our case, was also reported in Atlantic salmon (1.38) (Lien et al., 2011), gilthead sea bream (1.61) (Tsigenopoulos et al., 2014), Nile tilapia (1.2) (Joshi et al., 2018) and Gasterosteus sticklebacks (1.64) (Sardell et al., 2018), where in all cases it seems that the heterogametic sex has lower recombination rates. In our study, most LGs in the female map were larger than those in the male map, whereas the male map was larger in LG4, LG19, LG21 and LG24. Similar cases were also found in other fish, such as the gilthead seabream and Nile tilapia (Tsigenopoulos et al., 2014; Joshi et al., 2018). The molecular mechanisms for the sex differences in recombination rates are still not well understood. The differences may be caused by sexually antagonistic selection, meiotic drive in females, selection during the haploid phase of the life cycle, selection against aneuploidy, or mechanistic constraints; however, no single hypothesis can adequately explain the evolution of heterochiasmy in all cases (Sardell and Kirkpatrick, 2020).

Sigmodal patterns of recombination in large yellow croaker, with greater recombination rates toward the middle and lower recombination rates toward the ends, have also been seen in other species, like Nile tilapia (Joshi et al., 2018), whereas, salmon, channel catfish, etc (Li et al., 2014; Tsai et al., 2015) have shown opposite patterns, with higher recombination at the end of the LGs. The segments with little or no recombination may suggest possible location of centromeres. The karyotypes of large yellow croaker were earlier categorised into 10 pairs of sub-telocentric and 14 pairs of acrocentric chromosomes (Xu et al., 2017), implying that the centromeres are located at the end of the chromosomes, matching the low recombination rates seen towards the end of these LGs in our study. Recombination rate profiles within each LG also differed between males and female, with distinct regions containing potential recombination hotspots.

FIGURE 4 | Comparison of recombination rate variation throughout the genome in the integrated female (red) and male (blue) maps.

Physical gaps, as viewed from the *x*-axis in **Figure 3**, indicate lack of SNPs in these regions. One reason for these gaps could be massive repeat sites, unrecognisable by the *Pst I* enzyme during RADseq. Identifying additional markers with a different enzyme should thus help to fill these gaps. Another related reason could be the random and consequently partly uneven distribution of detected markers across the genomes, which is a disadvantage of RAD based technologies. Thus, RADseq usually generates medium density SNP linkage maps, leading to a low genome coverage (Robledo et al., 2017). Furthermore, the strain used in the *ASM435267v1* genome assembly, called DH2-L1, is a double haploid obtained by artificial gynogenesis from the Mindong strain only (Cai et al., 2010), whereas the population used in our genetic map is a cross between Daiqu strain and Mindong strain. The strain difference could thus be another reason for the physical gaps, as chromosomal rearrangements, including deletions, duplications, inversions, and translocations, could be different among strains.

Large jumps in some of the LGs were also viewed from the *y*-axis in **Figure 3**. These regions, with significantly elevated recombination rates, may be due to recombination hotspots,

insufficient SNP coverage caused by the randomness of RAD sequencing explained above, and/or low level of polymorphism in the F$_1$ family. A similar problem of large intervals was also presented in the genetic linkage map of the small yellow croaker (*Larimichthys polyactis*), also from only one fullsib family (Liu et al., 2020). Thus, use of multiple fullsib or halfsib families should be preferred, as done for instance with the high-density linkage map developed in Nile tilapia using 41 fullsib families (Joshi et al., 2018).

QTL Analysis and GWAS for Growth Traits

QTL analysis and GWAS are two types of strategies to detect potential causal genes for quantitative traits. QTL analysis detect associations between marker intervals and phenotypes, while GWAS identifies associations between single DNA markers and phenotypes, and thus the two methods complement each other (Sonah et al., 2015).

For fish less than 10 months of age, the gonads are hard to assess only by naked eye observation and there is hardly any gender difference to use for sex determination (Wang and Cai, 2018). Thus, no gender information was available for the fish at

6 months in the present study. Growth differences have already been observed between the Mindong and Daiqu strains and some phenotype segregation may be expected in the F_1, but no significant QTLs or SNPs were detected in our QTL or GWAS analysis. This was probably due to the complex genetic nature of the three growth traits which generally have been found to be controlled by many genes, each with minor effects. Also, the power of QTL analysis and GWAS will often not be sufficient with only one test family, due to the categorical nature of QTLs, for which a significant variant may or may not be present in any given family. For instance, the highly significant QTL variant that induced high resistance to IPN virus in the study of Moen et al. (2009), was only present in ca 5 % of the breeding nucleus. The QTL plot of BH is close to the threshold by 1,000 permutations, while the Manhattan plot of BH is far from the suggestive threshold by Bonferroni correction, which has been reported to be overly conservative in some cases (Kaler and Purcell, 2019). Also, no SNPs were identified to be significantly associated with the BL/BD ratio or the BL/BH ratio in large yellow croaker (Dong et al., 2019; Zhou et al., 2019). This may be due to low power in all these studies, but the results correspond well with the assumed polygenic nature of these traits. However, Xiao et al. (2015) identified several potential QTLs for growth traits (total weight, total length and total height) by composite interval mapping using 72 individuals from one fullsib family. But LOD score significance thresholds were not given in the plots in this study.

Using one F_1 fullsib family, as in the present study, Kong et al. (2019) identified seven significant QTLs linked to white spot disease resistance. The probability of identifying the QTLs in disease resistance traits could be higher than in growth traits, as it is often found that they are controlled by some major QTLs (Fraslin et al., 2020). However, these studies, using one F_1 family, only provide preliminary results of QTL mapping, and studies involving a more representative sample of the breeding population are required to conduct a marker-assisted selection scheme. One fullsib family is thus not ideal for identifying candidate genes, and a larger sample size and more families should be used to improve the power and to reveal potential associations (Korte and Farlow, 2013).

CONCLUSION

A consensus genetic linkage map for large yellow croaker was constructed with 20,147 SNPs from RAD sequencing, based on an F_1 family from Mindong strain and Daiqu strain. The total length of the consensus map was 1757.4 cM with an average marker interval of 0.09 cM. The female-to-male linkage map length ratio was 1.2. The map was adjusted based on the physical map, and integrated consensus and sex-specific linkage maps were generated. The recombination pattern mostly showed sigmoidal pattern of recombination. In most LGs, higher recombination rates were found in the integrated female map, compared to the integrated male map. No significant QTLs for growth related traits in fish at 6 months were found, probably due to the low detection power in only one family and the polygenic and complex nature of growth traits that are controlled by many genes with minor efforts. The present study indicates that there may be genetic differences between the two strains Daiqu and Mindong, which may have implications for breeding programs using DNA-information in a future selection scheme.

DATA AVAILABILITY STATEMENT

The sequenced data used for SNP detection has been deposited in Sequence Read Archive in NCBI (PRJNA786283). The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://figshare.com/, doi: 10.6084/m9. figshare. 16779160.

ETHICS STATEMENT

The animal study was reviewed and approved by the Animal Care and Use Committee of Zhejiang Ocean University.

AUTHOR CONTRIBUTIONS

ZL, XY: Conceptualization. HG, MK, and RJ: Supervision. RJ, XY, and MK: Data curation and analysis. XY: Samples preparation, result visualization and writing the original draft. RJ, MK, and HG: Reviewing and Editing.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.792666/full#supplementary-material

REFERENCES

Ao, J., Li, J., You, X., Mu, Y., Ding, Y., Mao, K., et al. (2015). Construction of the High-Density Genetic Linkage Map and Chromosome Map of Large Yellow Croaker (*Larimichthys Crocea*). *Ijms* 16 (11), 26237–26248. doi:10.3390/ijms161125951

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., et al. (2008). Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS One* 3 (10), e3376. doi:10.1371/journal.pone.0003376

Cai, M., Wu, Q., Liu, X., Yao, C., Chen, Q., and Wang, Z. (2010). Artificial Induction of Mito-Gynogenetic Diploids in Large Yellow Croaker (*Pseudosciaena Crocea*) by Hydrostatic Pressure. *Chin. J. Ocean. Limnol.* 28, 713–719. doi:10.1007/s00343-010-908510.1007/s00343-010-9085-3

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., and Cresko, W. A. (2013). Stacks: an Analysis Tool Set for Population Genomics. *Mol. Ecol.* 22 (11), 3124–3140. doi:10.1111/mec.12354

Chen, S., Su, Y., and Hong, W. (2018). "Aquaculture of the Large Yellow Croaker," in *Aquaculture in China* (Hoboken, New Jersey: Wiley), 297–308. doi:10.1002/9781119120759.ch3_10

Churchill, G. A., and Doerge, R. W. (1994). Empirical Threshold Values for Quantitative Trait Mapping. *Genetics* 138, 963–971. doi:10.1093/genetics/138.3.963

Davey, J. W., and Blaxter, M. L. (2010). RADSeq: Next-Generation Population Genetics. *Brief. Funct. Genomics* 9, 416–423. doi:10.1093/bfgp/elq031

Dong, L., Han, Z., Fang, M., Xiao, S., and Wang, Z. (2019). Genome-wide Association Study Identifies Loci for Body Shape in the Large Yellow Croaker (*Larimichthys Crocea*). *Aquacult. Fish.* 4, 3–8. doi:10.1016/j.aaf.2018.05.001

Fraslin, C., Quillet, E., Rochat, T., Dechamp, N., Bernardet, J.-F., Collet, B., et al. (2020). Combining Multiple Approaches and Models to Dissect the Genetic Architecture of Resistance to Infections in Fish. *Front. Genet.* 11, 677. doi:10.3389/fgene.2020.00677

Gonen, S., Lowe, N. R., Cezard, T., Gharbi, K., Bishop, S. C., and Houston, R. D. (2014). Linkage Maps of the Atlantic salmon (*Salmo salar*) Genome Derived from RAD Sequencing. *BMC Genomics* 15, 166. doi:10.1186/1471-2164-15-166

Houston, R. D., Davey, J. W., Bishop, S. C., Lowe, N. R., Mota-Velasco, J. C., Hamilton, A., et al. (2012). Characterisation of QTL-Linked and Genome-wide Restriction Site-Associated DNA (RAD) Markers in Farmed Atlantic salmon. *BMC Genomics* 13, 244. doi:10.1186/1471-2164-13-244

Huang, L. M., Xie, Y. J., and Su, Y. Q. (2006). Studies on Genetic Diversities of Daiqu Stock and Min-Yue Stock. *Pseudosciaena Crocea. J. Xiamen Univ. (Natural Science)* 45 (6), 836–840.

Joshi, R., Árnyasi, M., Lien, S., Gjøen, H. M., Alvarez, A. T., and Kent, M. (2018). Development and Validation of 58K SNP-Array and High-Density Linkage Map in Nile tilapia (*O. niloticus*). *Front. Genet.* 9, 472. doi:10.3389/fgene.2018.00472

Kaler, A. S., and Purcell, L. C. (2019). Estimation of a Significance Threshold for Genome-wide Association Studies. *BMC Genomics* 20, 618. doi:10.1186/s12864-019-5992-7

Kong, S., Ke, Q., Chen, L., Zhou, Z., Pu, F., Zhao, J., et al. (2019). Constructing a High-Density Genetic Linkage Map for Large Yellow Croaker (*Larimichthys Crocea*) and Mapping Resistance Trait against Ciliate Parasite *Cryptocaryon Irritans*. *Mar. Biotechnol.* 21, 262–275. doi:10.1007/s10126-019-09878-x

Korte, A., and Farlow, A. (2013). The Advantages and Limitations of Trait Analysis with GWAS: a Review. *Plant Methods* 9, 29. doi:10.1186/1746-4811-9-29

Li, M.-y., Hu, Y.-z., Miao, L., Chen, J., Shi, Y.-h., Xue, L.-y., et al. (2010). Studies on the Growth Characteristics and Heterosis of Genealogies of *Pseudosciaena Crocea. J. Fish. China* 34 (6), 679–684. doi:10.3724/sp.j.1231.2010.06750

Li, S., Wang, J., and Zhang, L. (2015). Inclusive Composite Interval Mapping of QTL by Environment Interactions in Biparental Populations. *PLoS One* 10 (7), e0132414. doi:10.1371/journal.pone.0132414

Li, Y., Liu, S., Qin, Z., Waldbieser, G., Wang, R., Sun, L., et al. (2014). Construction of a High-Density, High-Resolution Genetic Map and its Integration with BAC-Based Physical Map in Channel Catfish. *DNA Res.* 22, 39–52. doi:10.1093/dnares/dsu038

Lien, S., Gidskehaug, L., Moen, T., Hayes, B. J., Berg, P. R., Davidson, W. S., et al. (2011). A Dense SNP-Based Linkage Map for Atlantic salmon (*Salmo salar*) Reveals Extended Chromosome Homeologies and Striking Differences in Sex-specific Recombination Patterns. *BMC Genomics* 12, 615. doi:10.1186/1471-2164-12-615

Liu, F., Zhan, W., Xie, Q., Chen, H., Lou, B., and Xu, W. (2020). A First Genetic Linage Map Construction and QTL Mapping for Growth Traits in *Larimichthys Polyactis*. *Sci. Rep.* 10, 11621. doi:10.1038/s41598-020-68592-0

Liu, M., and de Mitcheson, Y. S. (2008). Profile of a Fishery Collapse: Why Mariculture Failed to Save the Large Yellow Croakerfile of a Fishery Collapse Why Mariculture Failed to Save the Large Yellow Croaker. *Fish and Fihseries* 9, 219–242. doi:10.1111/j.1467-2979.2008.00278.x

Lou, J. F., Lei, S. Y., Zhu, J. Q., and Wu, X. F. (2015). Chromosome Karyotype of *Pseudosciaena Crocea* Daiqu Population. *Chin. J. Zoolog.* 50 (1), 148–152. doi:10.13859/j.cjz.201501019

Meng, L., Li, H., Zhang, L., and Wang, J. (2015). QTL IciMapping: Integrated Software for Genetic Linkage Map Construction and Quantitative Trait Locus Mapping in Biparental Populations. *Crop J.* 3, 269–283. doi:10.1016/j.cj.2015.01.001

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-wide Dense Marker Maps. *Genetics* 157, 1819–1829. doi:10.1093/genetics/157.4.1819

Miao, L., Li, M. Y., Chen, J., and Zhang, H. (2014). Breeding of Fast Growth and Low Temperature Tolerance of New Variety Donghai No.1 Large Yellow Croaker (*Pseudosciaena Crocea*). *J. Agric. Biotechnol.* 22 (10), 1314–1320. doi:10.3969/j.issn.1674-7968.2014.10.015

Moen, T., Baranski, M., Sonesson, A. K., and Kjøglum, S. (2009). Confirmation and fine-mapping of a Major QTL for Resistance to Infectious Pancreatic Necrosis in Atlantic salmon (*Salmo salar*): Population-Level Associations between Markers and Trait. *BMC Genomics* 10, 368. doi:10.1186/1471-2164-10-368

Moen, T., Torgersen, J., Santi, N., Davidson, W. S., Baranski, M., Ødegård, J., et al. (2015). Epithelial Cadherin Determines Resistance to Infectious Pancreatic Necrosis Virus in Atlantic salmon. *Genetics* 200, 1313–1326. doi:10.1534/genetics.115.175406

Ning, Y., Liu, X., Wang, Z. Y., Guo, W., Li, Y., and Xie, F. (2007). A Genetic Map of Large Yellow Croaker Pseudosciaena Crocea. *Aquaculture* 264, 16–26. doi:10.1016/j.aquaculture.2006.12.042

Norris, A. (2017). Application of Genomics in salmon Aquaculture Breeding Programs by Ashie Norris. *Mar. Genomics* 36, 13–15. doi:10.1016/j.margen.2017.11.013

Ostberg, C. O., Hauser, L., Pritchard, V. L., Garza, J. C., and Naish, K. A. (2013). Chromosome Rearrangements, Recombination Suppression, and Limited Segregation Distortion in Hybrids between Yellowstone Cutthroat trout (*Oncorhynchus Clarkii Bouvieri*) and Rainbow trout (*O. mykiss*). *BMC Genomics* 14, 570. doi:10.1186/1471-2164-14-570

Palaiokostas, C., Bekaert, M., Khan, M. G. Q., Taggart, J. B., Gharbi, K., McAndrew, B. J., et al. (2013). Mapping and Validation of the Major Sex-Determining Region in Nile tilapia (*Oreochromis niloticus* L.) Using RAD Sequencing. *PLoS One* 8 (9), e68389–1371. doi:10.1371/journal.pone.0068389

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* 81, 559–575. doi:10.1086/519795

Rastas, P., Paulin, L., Hanski, I., Lehtonen, R., and Auvinen, P. (2013). Lep-MAP: Fast and Accurate Linkage Map Construction for Large SNP Datasets. *Bioinformatics* 29, 3128–3134. doi:10.1093/bioinformatics/btt563

Robledo, D., Palaiokostas, C., Bargelloni, L., Martínez, P., and Houston, R. (2017). Applications of Genotyping by Sequencing in Aquaculture Breeding and Genetics. *Rev. Aquacult* 10, 670–682. doi:10.1111/raq.12193

Sardell, J. M., Cheng, C., Dagilis, A. J., Ishikawa, A., Kitano, J., Peichel, C. L., et al. (2018). Sex Differences in Recombination in Sticklebacks. *G3 Genes|Genomes|Genetics* 8 (6), 1971–1983. doi:10.1534/g3.118.200166

Sardell, J. M., and Kirkpatrick, M. (2020). Sex Differences in the Recombination Landscape. *The Am. Naturalist* 195 (2), 361–379. doi:10.1086/704943

Siberchicot, A., Bessy, A., Guéguen, L., and Marais, G. A. (2017). MareyMap Online: a User-Friendly Web Application and Database Service for Estimating Recombination Rates Using Physical and Genetic Maps. *Genome Biol. Evol.* 9 (10), 2506–2509. doi:10.1093/gbe/evx178

Sonah, H., O'Donoughue, L., Cober, E., Rajcan, I., and Belzile, F. (2015). Identification of Loci Governing Eight Agronomic Traits Using a GBS-GWAS Approach and Validation by QTL Mapping in Soya Bean. *Plant Biotechnol. J.* 13 (2), 211–221. doi:10.1111/pbi.12249

Sun, S., and Xu, J. (2007). Genetic Analyses of a Hybrid Cross between Serotypes A and D Strains of the Human Pathogenic Fungus Cryptococcus Neoformans. *Genetics* 177 (3), 1475–1486. doi:10.1534/genetics.107.078923

Tsai, H.-Y., Hamilton, A., Tinch, A. E., Guy, D. R., Gharbi, K., Stear, M. J., et al. (2015). Genome Wide Association and Genomic Prediction for Growth Traits in Juvenile Farmed Atlantic salmon Using a High Density SNP Array. *BMC Genomics* 16, 969. doi:10.1186/s12864-015-2117-9

Tsigenopoulos, C. S., Louro, B., Chatziplis, D., Lagnel, J., Vogiatzi, E., Loukovitis, D., et al. (2014). Second Generation Genetic Linkage Map for the Gilthead Sea Bream *Sparus Aurata* L. *Mar. Genomics* 18, 77–82. doi:10.1016/j.margen.2014.09.008

Turner, S. D. (2014). QQMAN: An R Package for Visualizing GWAS Results Using Q-Q and manhattan Plots. *bioRxiv*, 005165. doi:10.1101/005165

Voorrips, R. E. (2002). MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *J. Hered.* 93, 77–78. doi:10.1093/jhered/93.1.77

Wang, Z.-Y., and Cai, M.-Y. (2018). Artificial Gynogenesis and Sex Control in Large Yellow Croaker. *Sex Control Aquac. Artif. gynogenesis sex Control large yellow croaker* 11, 751–773. doi:10.1002/9781119127291.ch39

Xiao, S., Wang, P., Zhang, Y., Fang, L., Liu, Y., Li, J.-T., et al. (2015). Gene Map of Large Yellow Croaker (*Larimichthys Crocea*) Provides Insights into Teleost Genome Evolution and Conserved Regions Associated with Growth. *Sci. Rep.* 5, 18661. doi:10.1038/srep18661

Xu, D., Molina, W. F., Yano, C. F., Zhang, Y., de Oliveira, E. A., Lou, B., et al. (2017). Comparative Cytogenetics in Three Sciaenid Species (Teleostei, Perciformes): Evidence of Interspecific Chromosomal Diversification. *Mol. Cytogenet.* 10, 37. doi:10.1186/s13039-017-0338-0

Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: a Tool for Genome-wide Complex Trait Analysis. *Am. J. Hum. Genet.* 88 (1), 76–82. doi:10.1016/j.ajhg.2010.11.011

Ye, H., Liu, Y., Liu, X., Wang, X., and Wang, Z. (2014). Genetic Mapping and QTL Analysis of Growth Traits in the Large Yellow Croaker Larimichthys Crocea. *Mar. Biotechnol.* 16, 729–738. doi:10.1007/s10126-014-9590-z

Yu, X. J., Xu, L. J., and Wu, F. X. (2020a). *China Fishery Statistical Yearbook*. Beijing, China: China Agriculture Press, 22–24.

Yu, X., Wu, C., and Gjøen, H. M. (2017a). Artificial Fertilization and Generating Families for a Selective Breeding Programme of Large Yellow Croaker (*Larimichthys Crocea*). *Int. Aquat. Res.* 9, 161–167. doi:10.1007/s40071-017-0164-3

Yu, X. X., Ådnøy, T., Lv, Z. M., Wu, C., and Gjøen, H. M. (2020b). Phenotypic and Genetic Parameter Estimation for Growth Traits in Juvenile Large Yellow Croaker (*Larimichthys Crocea*). *Fish. Aqua J.* 11, 274. doi:10.35248/2150-3508.20.11.274

Yu, Y., Ouyang, Y., and Yao, W. (2017b). shinyCircos: an R/Shiny Application for Interactive Creation of Circos Plot. *Bioinformatics* 34, 1229–1231. doi:10.1093/bioinformatics/btx763

Zenger, K. R., Khatkar, M. S., Jones, D. B., Khalilisamani, N., Jerry, D. R., and Raadsma, H. W. (2019). Genomic Selection in Aquaculture: Application, Limitations and Opportunities with Special Reference to marine Shrimp and Pearl Oysters. *Front. Genet.* 9, 693. doi:10.3389/fgene.2018.006910.3389/fgene.2018.00693

Zhou, Z., Han, K., Wu, Y., Bai, H., Ke, Q., Pu, F., et al. (2019). Genome-wide Association Study of Growth and Body-Shape-Related Traits in Large Yellow Croaker (*Larimichthys Crocea*) Using ddRAD Sequencing. *Mar. Biotechnol.* 21, 655–670. doi:10.1007/s10126-019-09910-0