Aquaculture 566 (2023) 739194

Contents lists available at ScienceDirect

Aquaculture

journal homepage: www.elsevier.com/locate/aquaculture

First genome-wide association study and genomic prediction for growth traits in spotted sea bass (*Lateolabrax maculatus*) using whole-genome resequencing

Chong Zhang, Haishen Wen, Yonghang Zhang, Kaiqiang Zhang, Xin Qi, Yun Li *

Key Laboratory of Mariculture, Ministry of Education (KLMME), Ocean University of China, Qingdao 266003, China

ARTICLE INFO

Keywords:
Lateolabrax maculatus
Growth traits
GWAS
Genomic selection

ABSTRACT

Spotted sea bass (*Lateolabrax maculatus*), widely distributed along the Chinese coasts, is an economically important aquaculture fish species. Recently, degeneration of genetic characteristics such as the decline in the growth rate severely hampers the development of its industry, and genetic improvement for this species is urgently required. In this study, the first genome-wide association study (GWAS) for growth traits (body weight, body height, total length and body length) were conducted and the potential performance of genomic selection (GS) were evaluated by genomic prediction of breeding values. Based on >4 million single-nucleotide polymorphisms (SNPs) genotyped by whole-genome resequencing for 514 individuals from Dongying (DY, 301 individuals) and Tangshan populations (TS, 213 individuals), GWAS detected a total of 66 growth-related SNPs located in multiple chromosomes but no major QTL, suggesting that growth traits were controlled by a polygenic genetic architecture. Candidate growth associated genes were identified to be involved in cytoskeleton reorganization, neuromodulation, angiogenesis and cell adhesion, and vascular endothelial growth factor (VEGF) and estrogen signaling pathways were considered to play important roles for growth. Predictive accuracies of the genomic estimated breeding value (GEBV) were compared among rrBLUP, BayesB, BayesC and BL models, and rrBLUP was determined as the optimal model for growth traits. Furthermore, the predictive performance based on different selection strategies of SNPs were compared, indicating using GWAS-informative SNPs was more efficient than random selected markers. These results highlighted the potential of GWAS to improve predictive accuracies of GS and reduce genotyping cost substantially. Our study laid the basis for further elucidate genetic mechanisms and demonstrated the application potential of GS approach for growth traits in spotted sea bass, which will facilitate future breeding of fast growth strains.

1. Introduction

Aquaculture is the fastest growing food sector in the world, providing a source of high-quality protein for human consumption (Ke et al., 2022; Tang et al., 2018). According to the Food and Agriculture Organization of the United Nations, global fish production had reached about 179 million tons in 2018, of which aquaculture accounted for 46% of the total production (FAO, 2020). Growth is one of the most important economic traits in aquaculture industry as it exerts direct influence on production (Li et al., 2018; Zhou et al., 2019). Therefore, many studies have focused on the genetic basis of growth-related traits and identified several genes and regulatory factors that could be applied to genetic improvement of growth in fish species. For examples, genes on somatotropic axis, such as growth hormone (*GH*), insulin-like growth factor (*IGF*), and their associated carrier proteins and receptors, have been demonstrated to play vital roles in the regulation of metabolic and physiological processes of fish growth (Caldarone et al., 2016; Chandhini et al., 2021; De-Santis and Jerry, 2007; Eivers et al., 2005; Hu et al., 2013). In addition, myogenic regulatory factors (*MRFs*), myostatin (*MSTN*), myomaker (*MYMK*) and myomixer (*MYMX*) were proved as essential myogenic genes for muscle development and growth (De-Santis and Jerry, 2007; Landemaine et al., 2019; Perello-Amoros et al., 2021; Sánchez-Ramos et al., 2012; Sun et al., 2012), as well as bone morphogenetic proteins (*bmps*) have been demonstrated as important regulators for skeletal development and bone formation in several aquaculture fish species (Geng et al., 2017; Wu et al., 2016). Despite

* Corresponding author.
E-mail addresses: wenhaishen@ouc.edu.cn (H. Wen), zkq@ouc.edu.cn (K. Zhang), qx@ouc.edu.cn (X. Qi), yunli0116@ouc.edu.cn (Y. Li).

https://doi.org/10.1016/j.aquaculture.2022.739194
Received 16 October 2022; Received in revised form 20 December 2022; Accepted 21 December 2022
Available online 23 December 2022

this, as the growth traits are known to be quantitative traits that controlled by multiple genes, it is difficult to elucidate the underlying genetic mechanisms of growth through studies at few gene or pathway level. Therefore, studies concentrated on precise localization of casual loci and identification for more candidate genes are needed to further revealing the genetic mechanism of fish growth.

Based on the rapid development of high-throughput sequencing technologies, several efficient tools have been discovered to understand the polygenic genetic architecture of growth traits (Li et al., 2018; Tsai et al., 2015b; Ye et al., 2014; Zhou et al., 2019). Among which, quantitative trait loci (QTL) mapping and genome-wide association studies (GWAS) became powerful methods for exploration of growth-related loci and genes. QTL mapping through linkage relationship relied on the family samples, which limits the number of recombinational events and leads to the identified QTLs localized in large chromosomal regions (Gutierrez et al., 2015; Korte and Farlow, 2013; Tsai et al., 2015a; Xu et al., 2019; Zhu et al., 2019). In contrast, GWAS utilizes tens of thousands SNP markers and natural populations that mimic a large number of historical recombination, which provides a relatively higher resolution and accuracy for identifying candidate genes (Scherer and Christensen, 2016; Tsai et al., 2015b; Xu et al., 2019; Yu et al., 2016). To date, GWAS for growth traits has successfully conducted for several aquaculture fish species such as the Atlantic salmon (*Salmo salar*) (Gutierrez et al., 2015; Tsai et al., 2015b), common carp (*Cyprinus carpio*) (Palaiokostas et al., 2018), hybrid catfish (*Ictalurus* spp.) (Geng et al., 2017; Li et al., 2018), rainbow trout (*Oncorhynchus mykiss*) (Gonzalez-Pena et al., 2016) and large yellow croaker (*Larimichthys crocea*) (Dong et al., 2016b; Zhou et al., 2019). Many novel candidate genes and genomic regions spread over the entire genome have been identified through these studies, facilitating the elucidation of molecular mechanisms regulating growth in fish.

Based on genetic markers covering the whole genome, genomic selection (GS) is a precise way to evaluate breeding potential of candidate populations and greatly facilitate genetic breeding progress for target traits (Barría et al., 2021; Liu et al., 2019). GS considers all SNPs with micro-effect sizes contributing the phenotypic variation and better predict the genomic estimated breeding values (GEBV) of genotyped candidate individuals, which tends to be more practical in the selective breeding of traits with a polygenic genetic architecture or low heritability (Meuwissen et al., 2001; Zhao et al., 2021). Compared with traditional breeding approach, GS possesses many advantages such as higher predictive accuracy, increased genetic gains, reduced breeding cycle and no requirement for complex pedigree recording (Shan et al., 2021; Zhao et al., 2021). Therefore, GS has been widely applied to economic species (Cuyabano et al., 2019; Crossa et al., 2017; Georges et al., 2019) and has made great progress in the selective breeding. With the significant reduction of sequencing cost, GS studies focusing on growth traits have been conducted in several aquaculture fish species such as Atlantic salmon (Tsai et al., 2015b), large yellow croaker (Dong et al., 2016a; Dong et al., 2016b), rock bream (*Oplegnathus fasciatus*) (Gong et al., 2021), yellow drum (*Nibea albiflora*) (Liu et al., 2019), common carp (Palaiokostas et al., 2018), and rainbow trout (García-Ballesteros et al., 2022). Genomic predictive accuracy is a key indicator determining the success of GS program and it depends on several factors including selection models, SNP sets of varying densities, selection strategies of SNPs, as well as population structure and genetic relatedness. These suggested that when applying GS in genetic breeding program, it's required to consider various influencing factors and compare predictive accuracies of different GS models and selection strategies, ultimately choose the most suitable GS approach for target trait (Dong et al., 2016a; Khatkar, 2017; Liu et al., 2019).

Spotted sea bass (*Lateolabrax maculatus*), as a euryhaline and eurythermic fish species, is widely distributed along the Chinese coasts (Wang et al., 2016). Due to its high nutritional value and pleasant taste (Liu et al., 2020; Sun et al., 2021), spotted sea bass became an economically important aquaculture fish species in China, with the annual production exceeding 190,000 tons in recent years (China Fishery Statistical Yearbook, 2021(Li et al., 2021). However, due to the lack of scientific genetic breeding, industry of spotted sea bass is threatened by degeneration of genetic characteristics such as decline in the growth rate and decreased disease-resistant ability. Additionally, its long-term generation interval (3–4 years) made the development of efficient molecular breeding program urgently required (Liu et al., 2020; Wang et al., 2017). Therefore, in order to accelerate the progress in genetic improvement of growth performance, in this study, a total of 514 individuals derived from two populations of spotted sea bass were genotyped using whole-genome resequencing methods, and their phenotypes of growth traits including body weight (BW), body height (BH), total length (TL) and body length (BL), were recorded. Our study aims to (1) identify SNPs and candidate genes associated with growth traits for spotted sea bass by GWAS method, (2) compare the accuracies of genomic prediction (GP) among different GS models and SNP selection strategies for growth traits, thereafter to evaluate the potential of GS for genetic improvement. Our study provides valuable resource for further elucidating the genetic mechanisms and preliminarily explores the optimal GS approach for growth traits in spotted sea bass, which will facilitate future selection breeding of its fast growth strains.

2. Materials and methods

2.1. Ethics statement

This study was conducted in accordance with approved guidelines of the respective Animal Research and Ethics Committees of Ocean University of China (Permit Number: 20141201). The present study did not include endangered or protected species.

2.2. Fish sample and phenotype measurement

Samples of spotted sea bass were collected from two local fish farms located in Dongying (DY) and Tangshan (TS), China, respectively. In details, a total of 301 one-year-old fish from DY population, which were natural population collected from the Yellow Sea and the Bohai Sea, as well as 213 five-year-old fish from TS population that were reared as brook stock and originally derived from multiple sources, were randomly selected as experimental materials respectively. Four growth traits including body weight (BW), body height (BH), total length (TL) and body length (BL) were measured, of which, BW was measured using electronic scale with a precision of 0.01 g; BH, TL, and BL were measured using tpsDig v2.0 software with precision of 0.01 cm. Pearson's correlation was employed to reflect the relationships among these four growth traits. The pectoral fins of fish individuals were sampled and stored in anhydrous ethanol for DNA extractions.

2.3. Sequencing, genotyping, and quality control

Genomic DNA of each sample was extracted with TIANamp Genomic DNA Kit (TIANGEN, Beijing, China). The concentration and quality of genomic DNA were determined by a Biodropsis BD-1000 nucleic acid analyzer (OSTC, Beijing) and electrophoresis in 1% agarose gel. Subsequently, high-quality DNA was sequenced with whole-genome resequencing technique implemented via BGISEQ-500 platform to generate paired-end 150 bp reads (BGI Genomics Co., Ltd., CHINA) following the protocols described previously (Fang et al., 2018). Thereafter, the raw reads were obtained and filtered using SOAPnuke v2.0 (Chen et al., 2018) to remove reads that contained the adapters, reads that possessed N bases (ambiguous bases) >10% of the total read length and reads with low Phred score ($Q \leq 12$) bases >50% of the total read length. After filtering, the clean reads were aligned to the reference genome of *L. maculatus* (PRJNA407434) using BWA v0.7.17 with default settings (Li and Durbin, 2009). Variant calling was conducted with GATK v4.1.8 software (McKenna et al., 2010) with the following filtering conditions:

quality by depth/variant confidence ≥2.0, fisher strand ≤60.0, RMS mapping quality ≥40.0, MQRankSum ≥ −12.5, ReadPosRankSum ≥ −8.0 and StrandOddsRatio (SOR) >3.0. Then we executed variant joint with GATK software for the following three genotype data sets (1) 301 individuals from DY population, (2) 213 individuals from TS population and (3) all 514 individuals, namely as ALL population. SNP data were further filtered using Plink v1.9 (Purcell et al., 2007) with the following parameters: (1) SNP sites with minor allele frequency (MAF) below 5% were removed; (2) individuals with variant missing rate >2% (mind >0.02) and SNP sites with genotyping call rate below 5% (geno <0.05) were discarded; and (3) SNP sites with *P* value of the Hardy Weinberg equilibrium chi-square test <0.05 (hwe < 0.05) were removed. Missing genotypes were imputed using software Beagle v5.2 (Browning and Browning, 2016). The SnpEff v5.0 program (Cingolani et al., 2012) was used to annotate SNPs based on the information of the reference genome of spotted sea bass. Finally, within each population, totally 4,660,345 SNPs were identified for DY population, and 4,288,765 SNPs were characterized for TS population, respectively. Additionally, 3,754,961 SNPs were shared among all 514 individuals from ALL population. Those SNPs were retained and were used for subsequent analyses.

2.4. Linkage disequilibrium (LD) and population structure analysis

The LD coefficient (r^2) between two SNPs was calculated using the PopLDdecay v3.41 package (Zhang et al., 2019) with the parameters of "-MaxDist 300 kb", and the genome-wide pattern of LD decay with distance was plotted by ggplot2 v3.3.6 package. Principal component analysis (PCA) was used to investigate the genetic structure for all individuals using Plink v1.9 (Purcell et al., 2007), and the PCA plot was visualized based on PC1–2 using ggplot2 package. For a better understanding of the population structure of all individuals, the Admixture v1.3.0 software (Alexander et al., 2009) was further employed to calculate population structure using the maximum likelihood model for DY, TS and ALL populations, respectively. The K-value (the putative number of genetic group) was set to 1–8, 1–10 and 1–10 respectively, and the K-value corresponding to the minimum cross-validation (CV) value was considered as the optimal population stratification number. In addition, genetic relatedness between individuals were evaluated by GCTA v1.93.2 software (Yang et al., 2011), and the corresponding heatmap was visualized using hist function in R.

2.5. Genome-wide association study (GWAS)

Based on the SNP genotype and growth traits phenotypic data, the Mixed Linear Model (MLM) of GEMMA v0.98.1 (Zhou and Stephens, 2012) was used to executed univariate GWAS to detect SNPs associated with the four growth traits in DY and TS populations, respectively. The calculation model is as follow:

$$y = W\alpha + X\beta + Z\mu + \varepsilon.$$

Where y is the phenotype data of growth traits; W is a covariate matrix of fixed effects, a is vector of corresponding coefficients including the first three principal components; X is the matrix for the fixed effects, β is the allele substitution effect of each SNP; Z is the genomic kinship matrix based on SNPs, μ is the additive genetic effect; and ε is the vector of residual errors. In addition, we carried out meta-analyses by METAL software (Willer et al., 2010) to integrate the results of two univariate GWAS analyses (DY and TS populations), and calculated the overall weighted *P*-values for each marker in ALL population. Meta-analysis is powerful to achieve the same effect as increasing the sample size by combining the results of multiple GWAS projects, and detect more variation sites with low frequency. The results of univariate GWAS analyses and Meta-analyses were visualized by the "CMplot v4.1.0" package (Yin et al., 2021) in R platform. LD pruning was conducted using Plink v1.9 (Purcell et al., 2007) with the parameter of "–indep-pairwise 1000kb 1 0.2" and generating independent SNPs. Based on Bonferroni correction with the estimated number of independent SNPs (N), the genome-wide significance association threshold was set as 0.05/N and the suggestive association threshold was set as 1/N. Gene was considered as candidate gene of growth traits when significant SNP is located in its gene region, or the nearest genes located at the upstream and downstream of the significant SNP. To further elucidate the regulation mechanism of growth traits, Gene Ontology (GO) enrichment analysis of all candidate genes was evaluated in DAVID (https://david.ncifcrf.gov/) online website. Finally, significantly enriched GO terms ($p < 0.05$) were visualized with ggplot2 v3.3.6 package.

2.6. Genomic prediction (GP)

To assess the feasibility of genomic selection (GS) for improving the growth traits in spotted sea bass, we conducted GP to evaluate predictive accuracies of TL trait in DY and TS populations. Considering the multicollinearity of GS model might be caused by high correlation among the adjacent SNPs, the tagging SNP, representing SNP in a haplotype region of the genome, were selected by Plink v1.9 (Purcell et al., 2007) and were used for subsequent GP analyses. Four models, including ridge regression best linear unbiased prediction (rrBLUP) (Meuwissen et al., 2001), BayesB (Cheng et al., 2015; Meuwissen et al., 2001), BayesC (Sun et al., 2011) and Bayesian LASSO (BL) (Li et al., 2018) were applied in our GS prediction. In short, the general form of these models was:

$$y = Xb + Zg + e.$$

Where y is the vector of phenotypic values; b is the vector of fixed effect including the first three principal components; g is the vector of additive genetic values (SNPs effect); e is the vector of residual effect; X and Z are incidence matrices relating the fixed effect and additive genetic values. Of which, rrBLUP model was performed using R package "rrBLUP" (Endelman, 2011), and the other three Bayes models were operated using R package "BGLR v1.1.0" (Pérez, and de los Campos, G., 2014).

Predictive accuracies of different GS models were compared through randomly selecting different numbers (100, 200, 400, 800, 1600, 3200, 6400, 12,800, 25,600, 51,200, 102,400, 204,800, 409,600 and all) of SNPs using ten-fold cross-validation with ten replicates (Spindel et al., 2015). In the procedure of ten-fold cross validation, the training set including 90% individuals was used to build GS model and calculate markers effect, then the GS model was used to calculate the genomic breeding values (GEBVs) of the validation set comprising the remaining 10% individuals. Finally, the predictive accuracy was calculated as the mean correlation coefficient between the predicted GEBVs and the actual phenotypic values of the validation set. Different numbers of SNP as stated above were randomly selected with five replicates. The predictive accuracy results for each GS model were averaged over five replicates to reduce random effects, and the model possessing the highest accuracy was selected as the optimal GS model for subsequent analyses.

Since the cost of high throughput sequencing is still high for aquaculture species compared with their market values, it's economically significant to obtain considerable predictive accuracy by using as few markers as possible. To explore the impact of reduced SNP density on predictive accuracy of GS models, we built 12 different SNP sets with the number of 10, 50, 100, 500, 1000, 5000, 10,000, 50,000, 100,000, 200,000, 500,000 and all SNPs, and compared the predictive accuracies using the optimal GS model of three SNP selection strategies under the following: (1) selecting the SNPs with highest ranked *P*-value based on the result of GWAS implemented only in the training population (GWAS), (2) selecting the SNPs with highest ranked P-value based on the result of GWAS implemented in whole population (GWAS1) and (3) SNPs were randomly selected with five replicates (Random).

3. Results

3.1. Statistics of the growth traits

The basic phenotypic statistics of four growth traits including BW, BH, TL and BL were described in Table 1 and Fig. S1, and the detailed phenotypes of growth traits for two populations were provided in Table S1. As shown in Table 1, the mean values of BW, BH, TL and BL were 73.62 ± 37.99 g, 4.37 ± 0.66 cm, 21.32 ± 2.94 cm and 18.51 ± 2.69 cm for DY population, and 2.85 ± 0.68 kg, 13.56 ± 1.14 cm, 65.64 ± 4.81 cm and 58.64 ± 4.53 cm for TS population, respectively. Among them, the coefficient of variation of BW were higher than the other traits (Table 1). The distribution histograms displayed that the four growth traits were close to normal distribution (Fig. S1). Significant high correlations were identified among the four growth traits ($P < 0.01$), with the Pearson correlation coefficient ranged from 0.8516 to 0.9899 for DY population, and from 0.8299 to 0.9949 for TS population (Fig. S2).

3.2. Genotyping results and marker distribution

A total of 3,754,961 SNPs were shared for ALL population, and the transition to transversion ratio was 1.545 (2,279,513/1,475,448). Among those SNPs, 1,634,217 (29.62%) were located in the intergenic region, 148,534 (2.69%) were in the exon region, 1,868,195 (33.86%) were in the intron region, and 1,723,112 (31.23%) were in the 1 kb region upstream or downstream of the gene coding sequences (Table S2). Among the SNPs belonging to the coding regions, the ratio of synonymous SNPs to nonsynonymous SNPs was 2.32:1 (103,221/44,532). The total physical distance covered by these SNPs was 597.61 Mb, and these SNPs are densely and evenly distributed in the genome with the average density of SNP/154 bp. Among the 24 chromosomes, chromosome 20 harbored the highest SNP marker density of SNP/137 bp, while chromosome 4 possessed the lowest SNP marker density of SNP/172 bp (Fig. 1A).

3.3. Population structure analysis

LD analysis results revealed that the squared correlation coefficient (r2) of two loci decreased rapidly as the distance between each pair of SNPs increased, and r2 decayed to 0.1 when the distance was around 150 bp (Fig. 1B). The result of principal component analysis (PCA) indicated that the genetic group of DY individuals were relatively simple, whereas TS individuals derived from multiple sources have more genetic groups. In addition, all DY individuals and partial TS individuals shared overlapped cluster, indicating that they might belong to the same genetic group (Fig. 1C). Moreover, population structure analysis indicated that the CV value reached the minimum when K = 1, 7 and 7 for DY, TS and ALL populations, respectively (Fig. S3), which were selected as the most suitable genetic group numbers for the three populations. As shown in Fig. 1D, the result of population structure analysis further confirmed that all DY individuals and part of TS individuals were clustered to the same genetic group. In addition, the analysis of genetic relatedness indicated no genetic relatedness was detected among DY individuals (Fig. S4A), and weak genetic relatedness existed among TS individuals (Fig. S4B).

Table 1
Statistics of phenotypic data for four growth traits in DY and TS populations.

Populations	Traits	Mean (SD)	Min	Max	CV (%)
DY	BW (g)	73.62 (37.99)	25.64	278.20	51.60%
($N = 301$)	BH (cm)	4.37 (0.66)	2.79	6.65	15.10%
	TL (cm)	21.32 (2.94)	15.99	32.05	13.79%
	BL (cm)	18.51 (2.69)	12.54	28.63	14.54%
TS	BW (kg)	2.85 (0.68)	1.350	5.35	23.86%
($N = 213$)	BH (cm)	13.56 (1.14)	10.15	17.16	8.41%
	TL (cm)	65.64 (4.81)	48.01	81.87	7.33%
	BL (cm)	58.64 (4.53)	42.98	72.36	7.73%

BW: Body weight; BH: Body height; TL: Total length; BL: Body length.
SD: standard deviation; Min: minimum; Max: maximum; CV: coefficient of variation.

3.4. GWAS analysis results

Due to the population stratification identified across the DY and TS populations, we firstly performed univariate GWAS analyses to identify significant SNPs associated with the four growth traits for DY and TS populations separately. Thereafter, meta-analyses which may exert great power to detect common and new candidate SNPs were applied for ALL population. LD pruning by Plink were performed, which were used to set a stringent threshold that guaranteed confidence of SNPs and avoided false positives, producing 2,236,541, 1,181,885 and 1,885,069 independent SNPs for DY, TS and ALL populations, respectively.

For BW trait, univariate GWAS analyses identified 3 genome-wide significant SNPs and 15 suggestive SNPs for DY population, and the phenotypic variance explained (PVE) by the 3 significant SNPs and 15 suggestive SNPs ranged from 9.45% to 10.36%, and from 8.24% to 10.36%, respectively. For TS population, 1 genome-wide significant SNP with the PVE of 13.53%, and 3 suggestive SNPs with the PVE from 10.85% to 11.17% were identified. Additionally, meta-analysis identified 1 novel significant SNP and 9 novel suggestive SNPs for ALL population (Table 2). For BH trait, univariate GWAS analyses identified 9 suggestive SNPs with the PVE from 8.18% to 9.57% for DY population, and 10 suggestive SNPs with the PVE from 11.16% to 12.63% were identified in TS population. Moreover, meta-analyses identified 2 new suggestive SNPs for ALL population (Table S3). The significant SNPs associated with TL and BL traits were basically consistent attributed to the extremely high correlations (0.99 for both DY and TS populations) between the two traits, thus we selected SNPs associated with TL trait as the representative of both traits. Univariate GWAS analyses of TL trait identified 1 significant SNP with the PVE of 11.06% and 12 suggestive SNPs with the PVE from 8.16% to 9.40% for DY population. For TS population, 2 significant SNPs with the PVE from 13.74% to 14.88% and 7 suggestive SNPs with the PVE from 10.88% to 11.42% were detected (Table S4). In addition, 2 new suggestive SNPs were detected for ALL population using meta-analyses.

3.5. Candidate genes identification

For BW trait, angiopoietin like 4 (*angptl4*), valyl-tRNA synthetase 1 (*vars*) and receptor activity modifying protein 1 (*ramp1*) exceeding the significant threshold, and perforin 1 (*prf1*), integrator complex subunit 8 (*ints8*), vascular endothelial growth factor receptor 3 (*vegfr3*), solute carrier family 39 member 10 (*slc39a10*), vav guanine nucleotide exchange factor 2 (*vav2*), NCK associated protein 1 (*nckap1*) and 8 other genes passed the suggestive threshold, were recognized as putative candidate genes associated for DY population (Fig. 2A). Neuritin 1 (*nrn1*), phospholipase C beta 3 (*plcb3*), solute carrier family 8 member A3 (*slc8a3*) and other 3 genes as putative candidate genes were identified for TS population (Fig. 2B). In addition, meta-analysis identified ribosomal protein L38 (*rpl38*), sidekick cell adhesion molecule 2 (*sdk2*), immunoglobin superfamily member 21 (*igsf21*), phosphatase and tensin homolog (*pten*) and other 11 novel genes, as BW associated candidate genes (Fig. 2C). For BH trait, nuclear receptor coactivator 3 (*ncoa3*), laminin subunit alpha 4 (*lama4*) and other 8 genes were detected as putative candidate genes for DY population (Fig. S5A), and neuritin 1 (*nrn1*), cadherin 18 (*cdh18*), AVL9 cell migration associated (*AVL9*), C2 calcium dependent domain containing 4A (*c2cd4a*) and other 8 genes were deduced as candidate genes for TS population (Fig. S5B). Moreover, Golgi associated, gamma adaptin ear containing, ARF binding protein 3 (*gga3*), melanin-concentrating hormone receptor 1 (*mchr1*), cyclin dependent kinase 13 (*cdk13*) and RALY RNA binding protein like gene (*ralyl*) were identified as putative BH associated genes using meta-

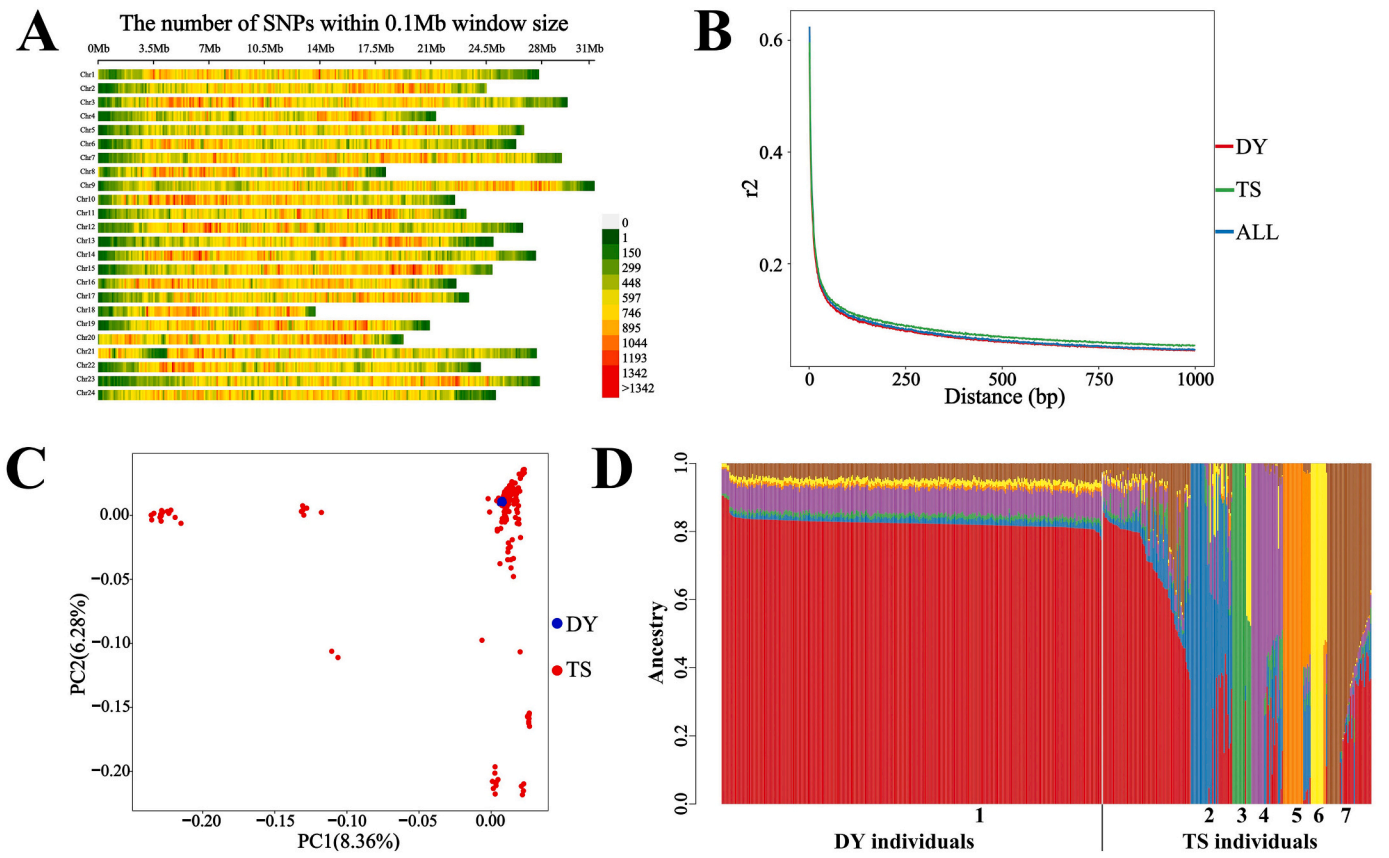

Fig. 1. (A) SNPs distribution pattern in different chromosomes of spotted sea bass. Different colors represent the corresponding number of SNPs within 0.1 Mb distance according to the legend. (B) LD decay plot of SNPs for DY, TS and ALL populations. (C) PCA plot of all individuals based on PC1 and PC2. (D) The population structure of ALL population when K = 7. The numbers represent the corresponding populations.

analyses (Fig. S5C). For TL trait, microtubule-actin cross-linking factor 1 (*macf1*) exceeding the significant threshold, and growth differentiation factor 15 (*gdf15*), mitochondrial ribosomal protein L10 (*mrpl10*) and other 9 candidate genes passed the suggestive threshold, were considered as putative candidate genes for DY population (Fig. S6A). Neuritin 1 (*nrn1*), ring finger protein 152 (*rnf152*), RAS responsive element binding protein 1 (*rreb1*) and other 8 genes were identified as putative candidate genes for TS population (Fig. S6B). In addition, meta-analyses detected two novel genes, solute carrier organic anion transporter family member 5A1 (*slco5a1*) and *N*-acetyltransferase 8 (*nat8*) were TL trait associated (Fig. S6C).

Enrichment analysis of GO indicated that these candidate genes serve essential functions in multiple growth-related procedures, such as angiogenesis, multicellular organism development, central nervous system development, ossification, homophilic cell adhesion, positive regulation of actin filament polymerization, and cell migration (Fig. 3A). In addition, two signaling pathways including vascular endothelial growth factor (VEGF) signaling pathway (Fig. 3B) and estrogen signaling pathway (Fig. 3C) were detected in our study and considered to potentially play critical roles in mammal growth (Coultas et al., 2005; Ferrara et al., 2003; Roman-Blas et al., 2009). The detailed potential functional mechanism of candidate genes and signaling pathways is described in "Discussion."

3.6. Genomic prediction (GP) for TL trait

After selected by Plink, 507,714 and 518,798 tagging SNPs were retained for DY and TS populations, separately. Using a tenfold cross-validation analysis, we firstly evaluated the predictive performance of four GS models (rrBLUP, BayesB, BayesC and BL) through randomly selecting different numbers of SNPs. Results showed that with marker density increasing, the predictive accuracies of three Bayes models always maintained a low level for DY population, and improved obviously for TS population. Meanwhile the predictive accuracy of rrBLUP model was generally higher than Bayes models at different marker densities for two populations (Table S5). For example, when the number of SNPs were 12,800 and 409,600, rrBLUP model had significantly higher predictive performance than Bayes models for both populations (19%–92% higher than other three models) (Fig. 4, Table S5). Therefore, rrBLUP model was considered as the optimal model for both populations and it was used to compare the predictive performance of different selection strategies.

The predictive accuracies of TL trait were compared among different SNP selection strategies using rrBLUP model (Fig. 5). For both DY and TS populations, the general predictive performance by using the same SNP numbers for different selection strategies was: GWAS1 > GWAS > Random. The predictive accuracies had reached >0.9 for both populations in strategy of GWAS1 when the number of SNPs was only 500, which were significantly higher than other selection strategies. We consider this result is highly biased and effect is overestimated, which was explained and discussed in details in "Discussion". In strategies of GWAS, the predictive accuracies increased gradually with the increasing of SNP markers, which reached the predictive plateau (0.37 and 0.42 for DY and TS population), and remained constant or decreased slightly when marker density continued to increase until employing all SNPs (Fig. 5). The general pattern of predictability of Random strategy was similar with GWAS, but the overall predictive accuracy was significantly lower than using GWAS strategy. In addition, we also found that the predictive accuracies of GWAS strategy could reach the prediction plateau at a lower marker density compared to the Random strategy. For

Table 2
Summary of identified SNPs and candidate genes associated with BW trait.

Population	SNP ID	Allele	Location	MAF	P-value	PVE	Candidate gene
DY	2–8,902,605	T/C	downstream	0.073	9.55E-09	0.1036	*angptl4*
	3–6,566,767	A/G	Extron	0.066	3.81E-07	0.0824	*prf1*
	5–9,225,217	A/G	Intergenic	0.125	1.23E-08	0.1027	*vars*
	11–18,404,621	T/C	Intronic	0.075	2.00E-07	0.0861	*ints8*
	12–17,266,377	G/A	Intronic	0.110	6.74E-08	0.0929	*vegfr3*
	13–4,929,312	T/A	Intergenic	0.058	1.70E-07	0.1011	*slc39a10*
	13–5,640,281	C/T	Intronic	0.071	1.55E-08	0.0945	*ramp1*
	13–21,236,985	A/G	Intergenic	0.050	4.56E-08	0.0871	*ahr*
	15–22,695,502	C/G	Intergenic	0.055	1.36E-07	0.0884	*sgcd/mrpl22*
	16–16,852,732	G/C	Intronic	0.053	1.36E-07	0.0881	*vav2*
	19–11,149,861	A/G	Intronic	0.060	2.35E-07	0.0952	*atrnl1*
	19–15,466,132	A/G	Intronic	0.070	4.11E-08	0.0909	*rpl38*
	19–15,466,142	T/A	Intronic	0.066	1.58E-07		
	19–15,466,147	T/A	Intronic	0.076	8.39E-08		
	22–9,077,242	G/A	Intergenic	0.063	4.46E-07	0.0920	*dph6/cdin1*
	22–9,077,745	G/A	Intergenic	0.063	7.33E-08		
	22–17,934,651	A/C	Intronic	0.071	4.01E-07	0.1036	*lrfn5*
	24–4,022,157	C/T	Intronic	0.070	4.19E-07	0.1036	*nckap1*
TS	4–19,888,299	T/C	Intronic	0.293	2.92E-08	0.1353	*nrn1*
	18–12,316,059	T/C	Intergenic	0.054	5.90E-07	0.1108	*plcb3/slc8a3*
	23–7,281,563	A/C	Intergenic	0.176	7.92E-07	0.1085	*chst14/cracd*
	23–14,632,325	G/A	Intronic	0.099	4.88E-07	0.1117	*npr2*
Meta	7–9,281,619	C/T	Intergenic	0.234	2.29E-07		*igsf21/arhgef10l*
	7–9,281,623	C/A	Intergenic	0.234	2.29E-07		
	7–17,918,414	A/G	Extron	0.220	3.37E-07		*spen*
	12–16,441,342	A/G	Intergenic	0.054	3.78E-07		*NA*
	13–5,640,281	C/T	Intronic	0.080	3.47E-07		*ramp1*
	19–15,794,281	T/C	Intergenic	0.095	7.62E-09		*rpl38/sdk2*
	22–12,321,843	C/G	Intronic	0.053	4.84E-07		*eml1*
	22–12,739,568	A/G	Intronic	0.075	1.69E-07		*ephb2*
	22–22,043,912	A/G	Intronic	0.054	2.03E-07		*pten*
	23–18,037,472	G/C	Intergenic	0.057	3.15E-08		*uhrf1/kcnn1*
	24–3,467,373	T/C	Intergenic	0.210	6.60E-08		*cdh11/mcm8*

Allele: minor/major allele; MAF: minor allele frequency; PVE: phenotypic variance explained.

DY population, 50 K GWAS-informative SNPs could achieve the prediction plateau, while almost all SNPs were required to obtain the highest predictive accuracy using Random strategy (Fig. 5A). For TS population, only 1 K GWAS-informative SNPs were needed to achieve the prediction plateau, however, 10 K SNPs were required for Random strategy (Fig. 5B).

4. Discussion

In this study, for growth traits of spotted sea bass, we identified 66 growth-related SNPs using the MLM model, of which, only two SNPs were located in the exon region and caused synonymous mutations, while most significant SNPs were located in intron and intergenic regions. These results suggested that most genetic variations may affect growth traits by regulating gene expression rather than altering gene products directly. In addition, the significant SNPs associated with growth traits were widely distributed in multiple chromosomes, and the phenotypic variance explained by these SNPs varied from 8.16% to 14.88%, which were slightly higher than previous studies in other fishes, such as catfish (2.67–6.72%) (Li et al., 2018), Atlantic salmon (up to 12%) (Tsai et al., 2015b), and rainbow trout (0.1–0.18%) (Gonzalez-Pena et al., 2016). These results confirmed that growth traits of fish species are complex polygenic genetic architecture and they are regulated by many micro-effect genes rather than several major QTLs (Li et al., 2018; Liu et al., 2019; Tsai et al., 2015b). However, significant SNPs detected in two populations were quite different, and they were also inconsistent with SNPs detected by QTL mapping for growth traits of spotted sea bass by using 2b-RAD method (Liu et al., 2020). In addition to methodological differences between GWAS and QTL mapping, population structure, genetic relatedness, genotyping strategy, marker density and some other factors may affect the final outcome (Rosenberg et al., 2010; Salisbury et al., 2003; Wu et al., 2019). Compared with our previous study of QTL mapping using 6883 SNPs generated from 333 F1 individuals in a full-sib family (Liu et al., 2020), in the current study, much higher density of SNP markers (>4 million) and complex population structure could detect more genetic variants associated with growth traits. Furthermore, genetic relatedness and population structure added to the model as a covariate could improve the accuracy and universality of the GWAS results (San et al., 2021). Our study indicated that population structure might be mainly responsible for the final outcome of GWAS, thus a wider range of population sources are needed to verify and further explore the polygenic regulation mechanisms underlying the growth traits of spotted sea bass.

In DY population, the most significant SNP associated with TL/BL traits were located in intron regions of *macf1* gene. Macf1 is an important cytoskeletal protein which could cross-link microfilaments and microtubules through the N-terminal actin-binding domain and C-terminal microtubule-binding domain (Hu et al., 2016), playing vital roles in several cellular processes including cell migration, cell proliferation and cell differentiation (Hu et al., 2015; Ka and Kim, 2016; Ka et al., 2014). In addition, the GO enrichment analysis also detected some cytoskeleton reorganization related categories such as microtubule (*slc8a3, macf1, dlg1, kif2a, eml1*), positive regulation of actin filament polymerization (*nckap1, dlg1, cracd*) and cytoskeletal protein binding (*macf1, dlg1, nebl*), indicating that cytoskeleton reorganization may play important roles in growth regulation of spotted sea bass. Neuromodulation has be identified as a major factor affecting fish growth (Liu et al., 2020; Su et al., 2018), and body weight is determined by a balance between food intake and energy expenditure, which is regulated by multiple neural circuits (Rui, 2013). In our study, many neuromodulation-related procedures including learning or memory (*slc8a3, ptprz1, pten, ephb2*), central nervous system development (*nckap1, ptprz1, pten, msi2*), synapse (*slc8a3, nrn1, ptprz1, pten, kcnn1, rpl38, sdk2*), neuron projection (*dlg1, pten, kcnn1, ephb2, mchr1*) and

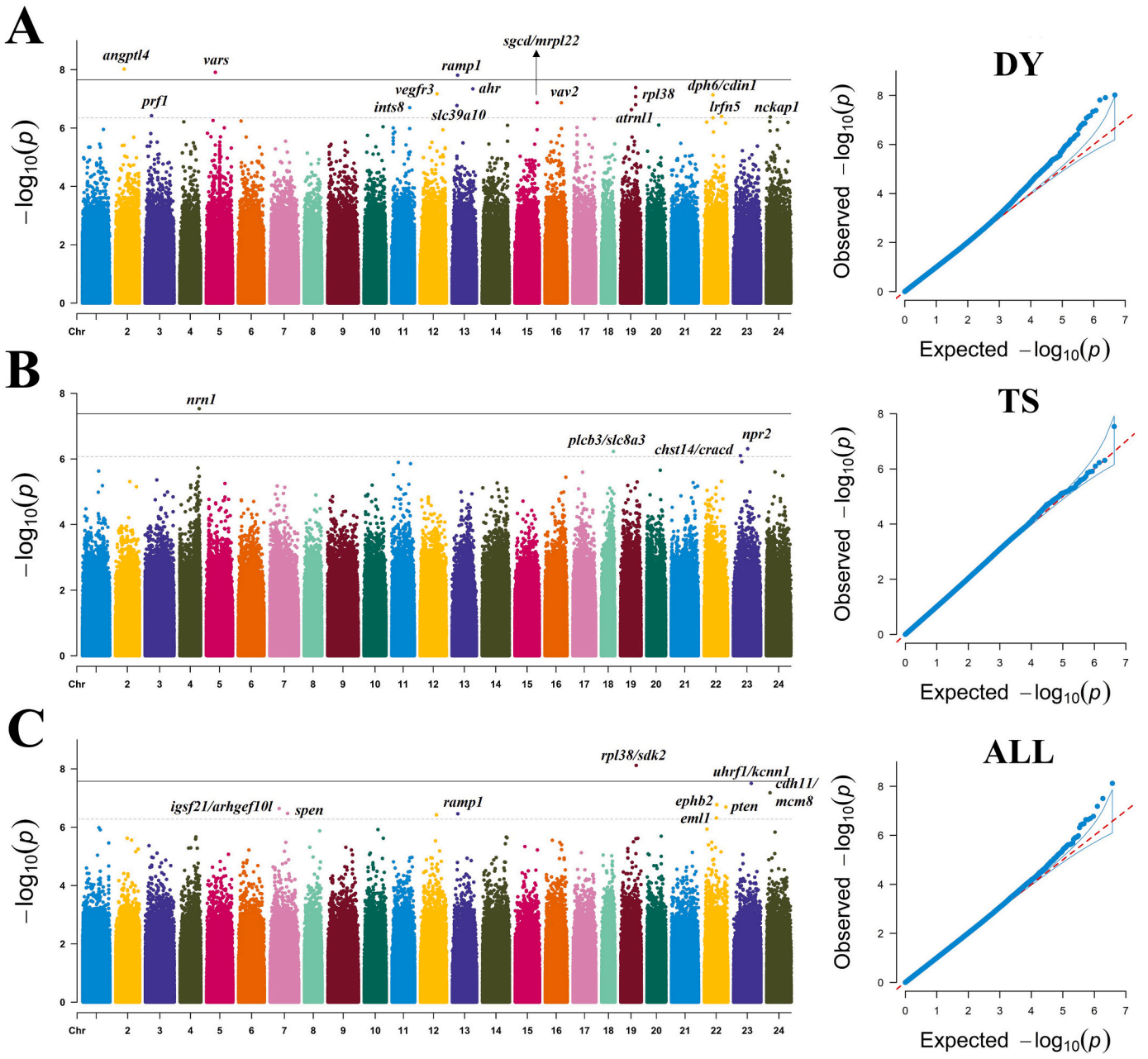

Fig. 2. Manhattan plots (left) and Quantile-Quantile (Q-Q) plots (right) of BW trait in GWAS analysis for (A) DY, (B) TS and (C) ALL populations. In Manhattan plots, the solid lines indicate the threshold *P* value for genome-wide significance association and the dashed lines indicate the threshold P value for suggestive association.

neuromuscular junction (*slc8a3, dlg1, lama4*) were enriched. In addition, the lowest *P*-values SNP were associated with BW/TL/BL/ traits for TS population and located in *nrn1* gene. Nrn1 is a neurotrophic factor closely related to the plasticity of the nervous system, which can promote neurite outgrowth and synaptic maturation, regulate the formation of synaptic circuits, and prevent the degeneration or apoptosis of nerve cells. These functions have positive significance for the regeneration and repair of nerve injury (Di Giovanni et al., 2005; Putz et al., 2005). Moreover, cell adhesion plays important roles in cell-cell communication and the development and maintenance of tissues (Khalili and Ahmad, 2015), and some related functional categories including adherens junction (*igsf2, dlg1, cdh11, cdh18*), cadherin binding (*macf1, plcb3, dlg1, cdh11, cdh18*) and homophilic cell adhesion via plasma membrane adhesion molecules (*igsf21, cdh11, sdk2, cdh18*) were detected in our study. Among them, cadherins (*cdhs*) are a superfamily of calcium-dependent transcellular membrane proteins playing an important role in mediating cell–cell adhesion (Bruner and Derksen, 2018; Mège and Ishiyama, 2017). Moreover, *col4a3*, *col4a4* and *pcdh10* genes related to cell adhesion has been also detected in previous QTL mapping for growth traits of spotted sea bass (Liu et al., 2020), indicating its significant role in teleost growth. However, the detailed mechanisms of cell adhesion regulating teleost growth are still undefined and need further investigations.

Notably, angiogenesis related genes (*vegfr3, pten, rora, angptl4, ephb2, ramp1, vav2*) and pathway were significantly enriched and studies have demonstrated that muscle regeneration is closely related to the formation of new blood vessels in mammals (Latroche et al., 2015a; Latroche et al., 2015b). In addition, VEGF signaling pathway which was known to play the essential role in regulating angiogenesis (Grünewald et al., 2010; Karaman et al., 2018) was also detected in our study and its potential regulation mechanism for growth was described in Fig. 3B. In detail, VEGF could activate downstream signaling pathways including

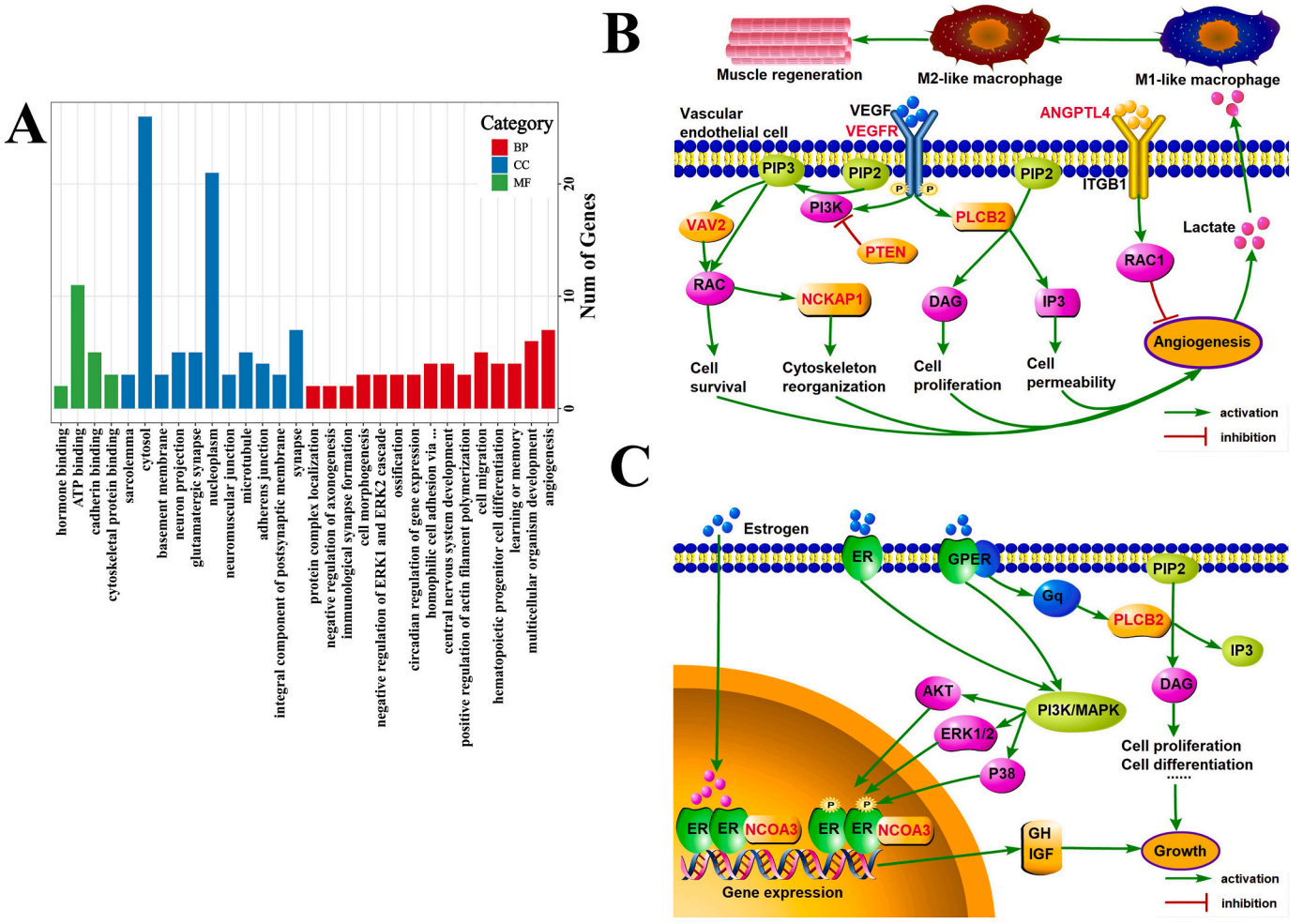

Fig. 3. (A) Gene ontology (GO) enrichment analysis of candidate genes associated with growth traits. (B) Schematic diagram of putative growth regulation mechanism by vascular endothelial growth factor (VEGF) signaling pathway. (C) Schematic diagram of putative growth regulation mechanism by estrogen signaling pathway. These growth traits associated candidate genes identified in this study were shown in red font. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

phosphatidylinositol 4,5-bisphosphate (PIP2) signaling pathway and rac family small GTPase (RAC) signaling pathway by binding to VEGF receptor (VEGFR), thereby triggering multiple biological processes like cytoskeleton reorganization, cell survival, cell proliferation and cell permeability, which ultimately promote angiogenesis (Grünewald et al., 2010; Koch and Claesson-Welsh, 2012; Meadows et al., 2001). Furthermore, angiopoietin like 4 (ANGPTL4) detected in our GWAS results could also affect angiogenesis via integrin-mediated RAC1 signaling pathway (Fernández-Hernando and Suárez, 2020; Huang et al., 2011). Upon angiogenesis, activated vascular endothelial cell could secrete lactate that instructs M1-like macrophages polarization toward a pro-regenerative M2-like macrophages, promoting muscle regeneration by stimulating the proliferation and differentiation of myogenic progenitor cells (MPCs) (Almada and Wagers, 2016; Zhang et al., 2020). Moreover, M2-like macrophages could also upregulate the expression of VEGF, thereby creating a positive feedback loop that further promotes angiogenesis (Zhang et al., 2020). These results indicate that angiogenesis may play an important role in muscle proliferation and differentiation of spotted sea bass, thereafter influence its growth performance, however, studies of angiogenesis for teleost growth are still limited and it's detailed functional mechanisms need to be further verified.

Moreover, estrogen signaling pathway, known to regulate fish growth mainly by affect the expression of growth hormone (GH) and insulin-like growth factor (IGF) (Chandhini et al., 2021; Holloway and Leatherland, 1997; Trudeau et al., 1992), was identified in our study and its potential regulation mechanism for growth was described in Fig. 3C. Estrogen primarily exerts growth-promoting effects by activating estrogen receptor (ER) and G protein-coupled estrogen receptor (GPER) on target tissues, mainly via two activation modes: (1) estrogen could bind to ER localized at nucleus to form ER dimers, (2) estrogen could bind to ER or GPER localized at or near cell membrane to activate the phosphatidylinositol-3/Akt (PI3K/Akt) and/or protein kinase C/mitogen activated protein kinase (PKC/MAPK) signal transduction pathways, thereby activate ER localized at nucleus by kinase-mediated phosphorylation and form ER dimers (Revankar et al., 2005; Roman-Blas et al., 2009). Activated ER dimers act as transcription factors which could specifically bind to estrogen response elements (EREs) in the promoters of target genes (GH, IGF) to regulate their expression, and the transcriptional activator functions could be enhanced by nuclear receptor coactivator 3 (NCOA3) that was also detected by our GWAS analysis (Klinge, 2001; Li et al., 2022; Roman-Blas et al., 2009; Xu et al., 2009). In addition, GPER could activate phospholipase c beta 2 (PLCB2) and then trigger PIP2 signaling pathway, thereby affecting cell proliferation, differentiation and other cell responses (Mooibroek and Wang, 1988; Prossnitz and Barton, 2014; Qiu et al., 2006). These complex biological responses may together contribute to the function of estrogen signaling pathway to regulate growth performance in spotted sea bass.

In this study, four GS models were performed and their predictive accuracies of GEBV were compared using ten replicates of ten-fold cross-

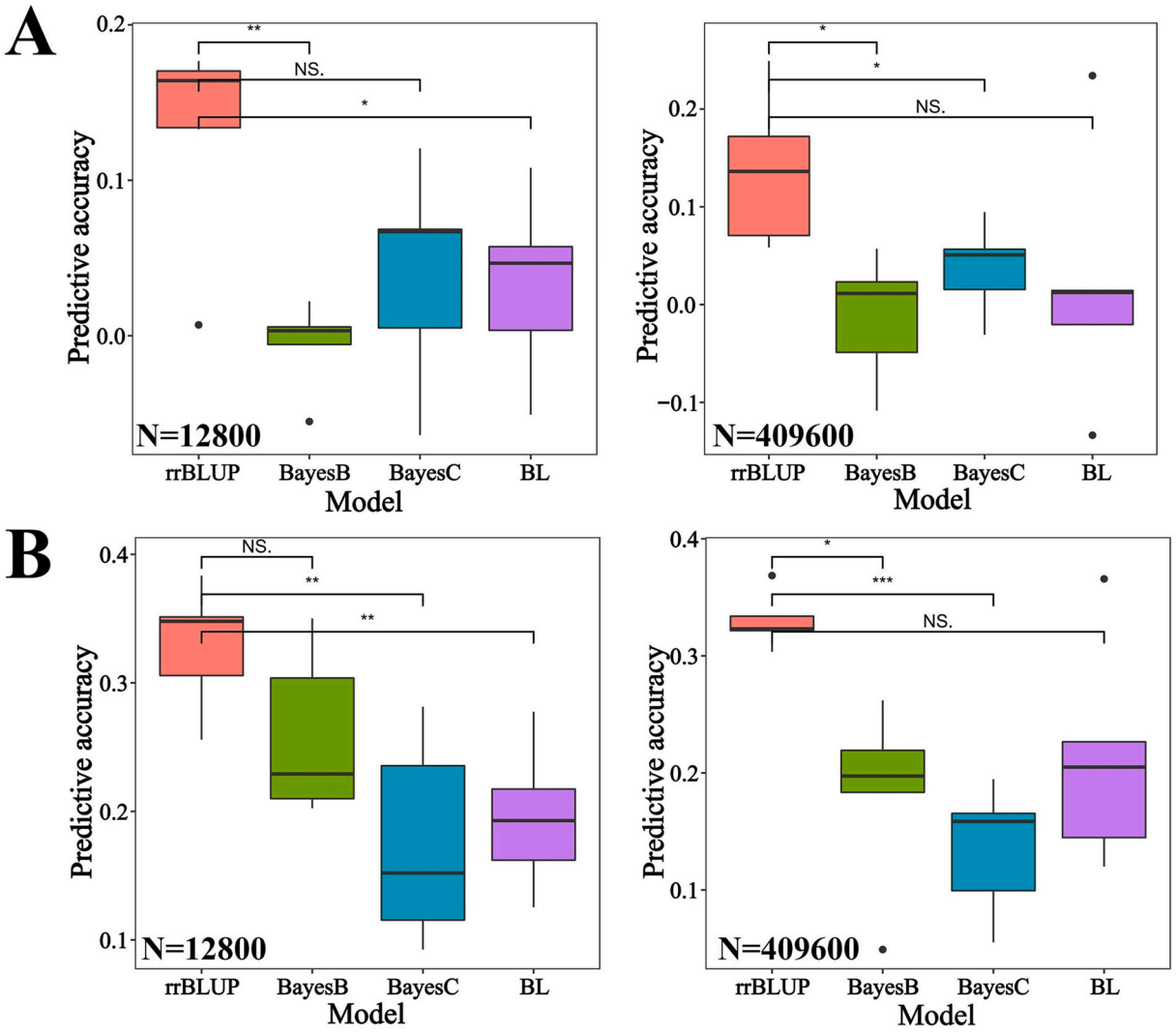

Fig. 4. Comparison of predictive accuracies using different models for TL trait in (A) DY and (B) TS populations. N is the number of randomly selected SNPs with five replicates, and *t*-test was used for significance analysis.

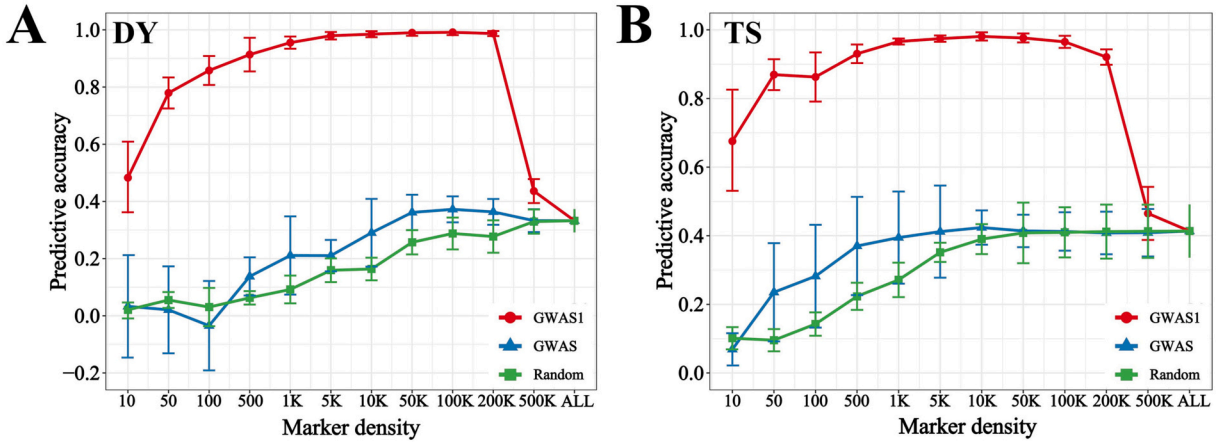

Fig. 5. Comparison of predictive accuracies using different selection strategies for TL trait in (A) DY and (B) TS populations. The rrBLUP model was used for the analysis.

validation. The predictive accuracies of rrBLUP model were significantly better than the other three Bayes models for TL trait of two populations, indicating rrBLUP model was more suitable for growth traits of spotted sea bass. Generally, the Bayes models have better performance for traits regulated by major quantitative trait loci due to the prior hypothesis of SNP effects (Gianola, 2013; Wang et al., 2014), while rrBLUP models are

more suitable for polygenic regulated traits because it assumes that every SNP following the same normal distribution has an effect (Endelman, 2011). In addition, the predictive accuracies of rrBLUP model for TS population seemed to be higher than DY population, which might due to the low genetic relationship of DY individuals and the discrepancy of population structure and size between two populations. The result indicated that population structure, sample size and genetic relatedness are important factors affecting the predictive accuracies of GS, which has been reported in many studies (Liu et al., 2019; Morgante et al., 2018; Yu et al., 2023). Hence, when applying genomic selection for genetic breeding, the above factors should be considered carefully in advance and then choose the most suitable model for a specific trait (Liu et al., 2019; Wang et al., 2018).

In aquaculture industry, cost-efficiency is important for implementation of GS as the relative low unit price of aquatic product compared to livestock (Liu et al., 2019; Shan et al., 2021; You et al., 2020). The use of lower marker density to achieve relatively high predictive accuracy, which could reduce the genotyping and computation cost dramatically, is a viable cost-efficiency strategy (Gong et al., 2021; Ke et al., 2022). Thus, we compared the predictive accuracies among different selection strategies of SNPs (GWAS1, GWAS and Random) for TL trait to explore the minimum required marker number. Our result showed that the strategy of GWAS1 displayed the best performance, which could achieve a very high predictive accuracy (>0.9) using only 500 SNPs. This conclusion is generally consistent with previous GS study of orange-spotted grouper (Shan et al., 2022; Shan et al., 2021). However, the predictive accuracies based on the GWAS1 strategy were significantly overestimated, which may be due to the disclosure of the phenotypes of validation set in the association analysis. This information has made a certain contribution to calculating the significant SNPs associated with target traits, leading to the bias for ranking of SNPs, and ultimately overestimated the predictive accuracy. Despite this, we considered that this strategy is still an effective way to optimize GS model using cross-validation in the reference population, which could obtain the minimum required marker number and predict GEBVs for selection candidates accurately. The approach has been proven to be an effective method to improve disease resistance to *Cryptocaryon irritans* in large yellow croaker (Zhao et al., 2021).

For GWAS and Random strategy, our results showed that when the SNP density was low, their predictive accuracies maintained at a low level (<0.1), but improved significantly when the number of SNP increased, indicating that a sufficient marker density is essential for obtaining high predictive accuracies of GS models (Spindel et al., 2015). Furthermore, the predictive accuracies of GWAS strategy were significantly higher than that of Random strategy when marker density achieved >500 and 50 SNPs for DY and TS populations, respectively, which suggested that using GWAS-informative SNPs had prominent advantage in genomic prediction. Moreover, we also found that Random strategy required more SNPs than GWAS strategy to achieve the plateau of predictive accuracies The similar results have also been reported for growth traits in pacific oyster (Gutierrez et al., 2018), rock bream (Gong et al., 2021), Atlantic salmon (Tsai et al., 2015b) and Nile Tilapia (Yoshida et al., 2019), suggesting that GWAS strategy is a viable cost-efficiency approach to achieve high predictive accuracies.

Moreover, our results showed that the performance of genomic prediction of TS population was generally better than DY population and DY population also required more SNPs to achieve maximum accuracy despite under GWAS strategy, which further indicated that population structure, sample size and genetic relatedness could affect the final outcome of GP. There is a major challenge for GP study with limited population size, as in our case, which is insufficient to estimate effects of all SNPs accurately and lead to bias to predict GEBV of validation individuals. It has been demonstrated that increasing population size could improve predictive accuracy effectively in many GS studies (Wang et al., 2022; Yu et al., 2023; Zhu et al., 2021). Therefore, more independent populations are needed to increase the population size to obtain better GS performance of growth traits in future breeding program of spotted sea bass.

5. Conclusion

In this study, we demonstrated that the combination of GWAS and GS would be a high-efficiency and cost-effective approach for elucidating genetic mechanism and predicting GEBV of growth traits, which has great potential for genetic improvement in spotted sea bass. A total of 66 significant SNPs were detected for growth traits, and associated candidate genes were involved in cytoskeleton reorganization, neuromodulation, angiogenesis and cell adhesion. Furthermore, VEGF and estrogen signaling pathway that significantly enriched in our results were also considered to play an important role for growth. Predictive performance of different GS models and marker selection strategies were compared, which indicated that rrBLUP was the optimal GS model for growth traits and using GWAS-informative SNPs was more efficient than random selected markers. Our study provided the important basis for elucidating the genetic mechanisms of growth regulation in spotted sea bass, and proved the feasibility and effectiveness of GS approach for genetic improvement of growth traits, which will facilitate future selection breeding of fast growth strains in this species.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aquaculture.2022.739194.

CRediT authorship contribution statement

Chong Zhang: Conceptualization, Methodology, Software, Writing – original draft. **Haishen Wen:** Conceptualization, Funding acquisition, Resources. **Yonghang Zhang:** Visualization, Software. **Kaiqiang Zhang:** Methodology, Resources. **Xin Qi:** Conceptualization, Resources. **Yun Li:** Conceptualization, Funding acquisition, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declared that the research did not involve any potential conflict of interest.

Data availability

Data will be made available on request.

Acknowledgments

This research was funded by National Key R&D Program of China [grant number: 2018YFD0900101]; and China Agriculture Research System (CARS for Marine Fish Culture Industry) [grant number: CARS-47].

References

Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19 (9), 1655–1664. https://doi.org/10.1101/gr.094052.109.

Almada, A.E., Wagers, A.J., 2016. Molecular circuitry of stem cell fate in skeletal muscle regeneration, ageing and disease. Nat. Rev. Mol. Cell Biol. 17 (5), 267–279. https://doi.org/10.1038/nrm.2016.7.

Barría, A., Benzie, J.A.H., Houston, R.D., De Koning, D.J., de Verdal, H., 2021. Genomic selection and genome-wide association study for feed-efficiency traits in a farmed Nile Tilapia (*Oreochromis niloticus*) population. Front. Genet. 12, 737906 https://doi.org/10.3389/fgene.2021.737906.

Browning, B.L., Browning, S.R., 2016. Genotype imputation with millions of reference samples. Am. J. Hum. Genet. 98 (1), 116–126. https://doi.org/10.1016/j.ajhg.2015.11.020.

Bruner, H.C., Derksen, P.W.B., 2018. Loss of E-cadherin-dependent cell-cell adhesion and the development and progression of cancer. Cold Spring Harb. Perspect. Biol. 10 (3), a029330 https://doi.org/10.1101/cshperspect.a029330.

Caldarone, E., MacLean, S., Beckman, B., 2016. Evaluation of nucleic acids and plasma IGF1 levels for estimating short-term responses of postsmolt Atlantic salmon (*Salmo*

salar) to food availability. Fish. Bull. 114, 288–301. https://doi.org/10.7755/FB.114.3.3.
Chandhini, S., Trumboo, B., Jose, S., Varghese, T., Rajesh, M., Kumar, V.J.R., 2021. Insulin-like growth factor signalling and its significance as a biomarker in fish and shellfish research. Fish Physiol. Biochem. 47, 1011–1031. https://doi.org/10.1007/s10695-021-00961-6.
Chen, Y., Chen, Y., Shi, C., Huang, Z., Zhang, Y., Li, S., Li, Y., Ye, J., Yu, C., Li, Z., Zhang, X., Wang, J., Yang, H., Fang, L., Chen, Q., 2018. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. GigaScience. 7, 1–6. https://doi.org/10.1093/gigascience/gix120.
Cheng, H., Qu, L., Garrick, D.J., Fernando, R.L., 2015. A fast and efficient Gibbs sampler for BayesB in whole-genome analyses. Genetics, Selection, Evolution : GSE. 14 (47), 80. https://doi.org/10.1186/s12711-015-0157-x.
Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly. 6 (2), 80–92. https://doi.org/10.4161/fly.19695.
Coultas, L., Chawengsaksophak, K., Rossant, Fau -, Rossant, J., 2005. Endothelial cells and VEGF in vascular development. Nature. 438 (7070), 937–945. https://doi.org/10.1038/nature04479.
Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de Los Campos, G., Burgueño, J., González-Camacho, J.M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., Varshney, R.K., 2017. Genomic selection in plant breeding: methods, models, and perspectives. Trends Plant Sci. 22 (11), 961–975. https://doi.org/10.1016/j.tplants.2017.08.011.
Cuyabano, Castro Dias, Wackel, H., Shin, D., Gondro, C., 2019. A study of genomic prediction across generations of two Korean pig populations. Animals. 9 (9), E672. https://doi.org/10.3390/ani9090672.
De-Santis, C., Jerry, D.R., 2007. Candidate growth genes in finfish — where should we be looking? Aquaculture. 272 (1–4), 22–38. https://doi.org/10.1016/j.aquaculture.2007.08.036.
Di Giovanni, S., Faden, A.I., Yakovlev, A., Duke-Cohan, J.S., Finn, T., Thouin, M., Knoblach, S., De Biase, A., Bregman, B.S., Hoffman, E.P., 2005. Neuronal plasticity after spinal cord injury: identification of a gene cluster driving neurite outgrowth. Federat. Am. Soc. Exp. Biol. 19 (1), 153–154. https://doi.org/10.1096/fj.04-2694fje.
Dong, L., Xiao, S., Wang, Q., Wang, Z., 2016a. Comparative analysis of the GBLUP, emBayesB, and GWAS algorithms to predict genetic values in large yellow croaker (*Larimichthys crocea*). BMC Genomics 17, 460. https://doi.org/10.1186/s12864-016-2756-5.
Dong, L., Xiao, S., Chen, J., Wan, L., Wang, Z., 2016b. Genomic selection using extreme phenotypes and pre-selection of SNPs in large yellow croaker (*Larimichthys crocea*). Mar. Biotechnol. 18 (5), 575–583. https://doi.org/10.1007/s10126-016-9718-4.
Eivers, E., McCarthy, K., Glynn, C., Nolan, C., Byrnes, L., 2005. Insulin-like growth factor (IGF) signalling is required for early dorso-anterior development of the zebrafish embryo. Int. J. Develop. Biol. 48 (10), 1131–1140. https://doi.org/10.1387/ijdb.041913ee.
Endelman, J., 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. The Plant Genome. 4 (3), 250–255. https://doi.org/10.3835/plantgenome2011.08.0024.
Fang, C., Zhong, H., Lin, Y., Chen, B., Han, M., Ren, H., Lu, H., Luber, J.M., Xia, M., Li, W., Stein, S., Xu, X., Zhang, W., Drmanac, R., Wang, J., Yang, H., Hammarström, L., Kostic, A.D., Kristiansen, K., Li, J., 2018. Assessment of the cPAS-based BGISEQ-500 platform for metagenomic sequencing. GigaScience. 7 (3), 1–8. https://doi.org/10.1093/gigascience/gix133.
FAO, 2020. The State of World Fisheries and Aquaculture 2020. United Nations Food and Agricultural Organization.
Fernández-Hernando, C., Suárez, Y., 2020. ANGPTL4: a multifunctional protein involved in metabolism and vascular homeostasis. Curr. Opin. Hematol. 27 (3), 206–213. https://doi.org/10.1097/moh.0000000000000580.
Ferrara, N., Fau, Gerber Hp, LeCouter, J., LeCouter, J., 2003. The biology of VEGF and its receptors. Nat. Med. 9 (6), 669–676. https://doi.org/10.1038/nm0603-669.
García-Ballesteros, S., Fernández, J., Kause, A., Villanueva, B., 2022. Predicted genetic gain for carcass yield in rainbow trout from indirect and genomic selection. Aquaculture. 554, 738119 https://doi.org/10.1016/j.aquaculture.2022.738119.
Geng, X., Liu, S., Yuan, Z., Jiang, Y., Zhi, D., Liu, Z., 2017. A genome-wide association study reveals that genes with functions for bone development are associated with body conformation in catfish. Mar. Biotechnol. 19 (6), 570–578. https://doi.org/10.1007/s10126-017-9775-3.
Georges, M., Charlier, C., Hayes, B., 2019. Harnessing genomic information for livestock improvement. Nat. Rev. Genet. 20 (3), 135–156. https://doi.org/10.1038/s41576-018-0082-2.
Gianola, D., 2013. Priors in whole-genome regression: the bayesian alphabet returns. Genetics. 194 (3), 573–596. https://doi.org/10.1534/genetics.113.151753.
Gong, J.A.-O., Zhao, J., Ke, Q., Li, B., Zhou, Z.A.-O., Wang, J., Zhou, T., Zheng, W., Xu, P. A.-O., 2021. First genomic prediction and genome-wide association for complex growth-related traits in Rock Bream (*Oplegnathus fasciatus*). Evol. Appl. 15 (4), 523–536. https://doi.org/10.1111/eva.13218.
Gonzalez-Pena, D., Gao, G., Baranski, M., Moen, T., Cleveland, B.M., Kenney, P.B., Vallejo, R.L., Palti, Y., Leeds, T.D., 2016. Genome-wide association study for identifying loci that affect fillet yield, carcass, and body weight traits in rainbow trout (*Oncorhynchus mykiss*). Front. Genet. 7, 203. https://doi.org/10.3389/fgene.2016.00203.
Grünewald, F.S., Fau, Prota Ae, Giese, A., Fau, Giese A., Ballmer-Hofer, K., Ballmer-Hofer, K., 2010. Structure-function analysis of VEGF receptor activation and the role of coreceptors in angiogenic signaling. Biochim. Biophys. Acta 1804 (3), 567–580. https://doi.org/10.1016/j.bbapap.2009.09.002.
Gutierrez, A.P., Yáñez, J., Fukui, S., Swift, B., Davidson, W., 2015. Genome-wide association study (GWAS) for growth rate and age at sexual maturation in Atlantic Salmon (*Salmo salar*). PLoS One 10 (3), e0119730. https://doi.org/10.1371/journal.pone.0119730.
Gutierrez, A.P., Matika, O., Bean, T.P., Houston, R.D., 2018. Genomic selection for growth traits in Pacific oyster (*Crassostrea gigas*): potential of low-density marker panels for breeding value prediction. Front. Genet. 9, 391. https://doi.org/10.3389/fgene.2018.00391.
Holloway, A.C., Leatherland, J.F., 1997. Effect of gonadal steroid hormones on plasma growth hormone concentrations in sexually immature rainbow trout *Oncorhynchus mykiss*. Gen. Comp. Endocrinol. 105 (2), 246–254. https://doi.org/10.1006/gcen.1996.6826.
Hu, X., Li, C., Shi, L., 2013. A novel 79-bp insertion/deletion polymorphism in 3′-flanking region of IGF-I gene is associated with growth-related traits in common carp (*Cyprinus carpio* L.). Aquac. Res. 44, 1632–1638. https://doi.org/10.1111/are.12091.
Hu, L., Su, P., Li, R., Yan, K., Chen, Z., Shang, P., Qian, A., 2015. Knockdown of microtubule actin crosslinking factor 1 inhibits cell proliferation in MC3T3-E1 osteoblastic cells. BMB Rep. 48 (10), 583–588. https://doi.org/10.5483/bmbrep.2015.48.10.098.
Hu, L., Su, P., Li, R., Yin, C., Zhang, Y., Shang, P., Yang, T., Qian, A., 2016. Isoforms, structures, and functions of versatile spectraplakin MACF1. BMB Rep. 49 (1), 37–44. https://doi.org/10.5483/BMBRep.2016.49.1.185.
Huang, R.-L., Teo, Z., Chong, H.C., Zhu, P., Tan, M.J., Tan, C.K., Lam, C.R.I., Sng, M.K., Leong, D.T.W., Tan, S.M., Kersten, S., Ding, J.L., Li, H.Y., Tan, N.S., 2011. ANGPTL4 modulates vascular junction integrity by integrin signaling and disruption of intercellular VE-cadherin and claudin-5 clusters. Blood. 118 (14), 3990–4002. https://doi.org/10.1182/blood-2011-01-328716.
Ka, M., Kim, W.Y., 2016. Microtubule-actin crosslinking factor 1 is required for dendritic arborization and axon outgrowth in the developing brain. Mol. Neurobiol. 53 (9), 6018–6032. https://doi.org/10.1007/s12035-015-9508-4.
Ka, M., Jung, E.M., Mueller, U., Kim, W.Y., 2014. MACF1 regulates the migration of pyramidal neurons via microtubule dynamics and GSK-3 signaling. Dev. Biol. 395 (1), 4–18. https://doi.org/10.1016/j.ydbio.2014.09.009.
Karaman, S.A.-O., Leppänen, V.M., Alitalo, K.A.-O., 2018. Vascular endothelial growth factor signaling in development and disease. Development. 145 (14), dev151019. https://doi.org/10.1242/dev.151019.
Ke, Q., Wang, J., Bai, Y., Zhao, J., Gong, J., Deng, Y., Qu, A., Suo, N., Chen, J., Zhou, T., Xu, P., 2022. GWAS and genomic prediction revealed potential for genetic improvement of large yellow croaker adapting to high plant protein diet. Aquaculture. 553, 738090 https://doi.org/10.1016/j.aquaculture.2022.738090.
Khalili, A.A., Ahmad, M.R., 2015. A review of cell adhesion studies for biomedical and biological applications. Int. J. Mol. Sci. 16 (8), 18149–18184. https://doi.org/10.3390/ijms160818149.
Khatkar, M.S., 2017. Genomic selection in aquaculture breeding programs. Bioinformat. Aquacult. 380-391 https://doi.org/10.1002/9781118782392.ch21.
Klinge, C.M., 2001. Estrogen receptor interaction with estrogen response elements. Nucleic Acids Res. 29 (14), 2905–2919. https://doi.org/10.1093/nar/29.14.2905.
Koch, S., Claesson-Welsh, L., 2012. Signal transduction by vascular endothelial growth factor receptors. Cold Spring Harbor Perspect. Med. 2 (7), a006502 https://doi.org/10.1101/cshperspect.a006502.
Korte, A., Farlow, A., 2013. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9, 29. https://doi.org/10.1186/1746-4811-9-29.
Landemaine, A., Ramirez-Martinez, A., Monestier, O., Sabin, N., Rescan, P.Y., Olson, E. A.-O., Gabillard, J.A.-O.X., 2019. Trout myomaker contains 14 minisatellites and two sequence extensions but retains fusogenic function. J. Biol. Chem. 294 (16), 6364–6374. https://doi.org/10.1074/jbc.ra118.006047.
Latroche, C., Gitiaux, C., Chrétien, F., Desguerre, I., Mounier, R., Chazaud, B., 2015a. Skeletal muscle microvasculature: a highly dynamic lifeline. Physiology. 30 (6), 417–427. https://doi.org/10.1152/physiol.00026.2015.
Latroche, C., Matot, B., Martins-Bach, A., Briand, D., Chazaud, B., Wary, C., Carlier, P.G., Chrétien, F., Jouvion, G., 2015b. Structural and functional alterations of skeletal muscle microvasculature in dystrophin-deficient mdx mice. Am. J. Pathol. 185 (9), 2482–2494. https://doi.org/10.1016/j.ajpath.2015.05.009.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics. 25 (14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
Li, N., Zhou, T., Geng, X., Jin, Y., Wang, X., Liu, S., Xu, X., Gao, D., Li, Q., Liu, Z., 2018. Identification of novel genes significantly affecting growth in catfish through GWAS analysis. Mol. Gen. Genomics. 293 (3), 587–599. https://doi.org/10.1007/s00438-017-1406-1.
Li, B., Tian, Y., Wen, H., Qi, X., Wang, L., Zhang, J., Li, J., Dong, X., Zhang, K., Li, Y., 2021. Systematic identification and expression analysis of the sox gene family in spotted sea bass (*Lateolabrax maculatus*). Comparat. Biochem. Physiol. Part D, Genom. Proteom. 38, 100817 https://doi.org/10.1016/j.cbd.2021.100817.
Li, Y., Liang, J., Dang, H., Zhang, R., Chen, P., Shao, Y., 2022. NCOA3 is a critical oncogene in thyroid cancer via the modulation of major signaling pathways. Endocrine. 75 (1), 149–158. https://doi.org/10.1007/s12020-021-02819-6.
Liu, G., Dong, L., Gu, L., Han, Z., Zhang, W., Fang, M., Wang, Z., 2019. Evaluation of genomic selection for seven economic traits in yellow drum (*Nibea albiflora*). Mar. Biotechnol. 21 (6), 806–812. https://doi.org/10.1007/s10126-019-09925-7.
Liu, Y., Wang, H., Wen, H., Shi, Y., Zhang, M., Qi, X., Zhang, K., Gong, Q., Li, J., He, F., Hu, Y., Li, Y., 2020. First high-density linkage map and QTL fine mapping for

growth-related traits of spotted sea bass (*Lateolabrax maculatus*). Mar. Biotechnol. 22 (4), 526–538. https://doi.org/10.1007/s10126-020-09973-4.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20 (9), 1297–1303. https://doi.org/10.1101/gr.107524.110.
Meadows, K.N., Bryant, P., Pumiglia, K., 2001. Vascular endothelial growth factor induction of the angiogenic phenotype requires Ras activation. J. Biol. Chem. 276 (52), 49289–49298. https://doi.org/10.1074/jbc.m108069200.
Mège, R.M., Ishiyama, N., 2017. Integration of cadherin adhesion and cytoskeleton at adherens junctions. Cold Spring Harb. Perspect. Biol. 9 (5), a028738 https://doi.org/10.1101/cshperspect.a028738.
Meuwissen, T.H., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics. 157 (4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819.
Mooibroek, M.J., Wang, J.H., 1988. Integration of signal-transduction processes. Biochem. Cell Biol. 66 (6), 557–566. https://doi.org/10.1139/o88-066.
Morgante, F., Huang, W., Maltecca, C., Mackay, T.F.C., 2018. Effect of genetic architecture on the prediction accuracy of quantitative traits in samples of unrelated individuals. Heredity. 120 (6), 500–514. https://doi.org/10.1038/s41437-017-0043-0.
Palaiokostas, C., Kocour, M., Prchal, M., Houston, R.D., 2018. Accuracy of genomic evaluations of juvenile growth rate in common carp (*Cyprinus carpio*) using genotyping by sequencing. Front. Genet. 9, 82. https://doi.org/10.3389/fgene.2018.00082.
Perello-Amoros, M., Rallière, C., Gutiérrez, J., Gabillard, J.C., 2021. Myomixer is expressed during embryonic and post-larval hyperplasia, muscle regeneration and differentiation of myoblats in rainbow trout (*Oncorhynchus mykiss*). Gene. 790, 145688 https://doi.org/10.1016/j.gene.2021.145688.
Pérez, P., de los Campos, G., 2014. Genome-wide regression and prediction with the BGLR statistical package. Genetics. 198 (2), 483–495. https://doi.org/10.1534/genetics.114.164442.
Prossnitz, E.R., Barton, M., 2014. Estrogen biology: new insights into GPER function and clinical opportunities. Mol. Cell. Endocrinol. 389 (1–2), 71–83. https://doi.org/10.1016/j.mce.2014.02.002.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81 (3), 559–575. https://doi.org/10.1086/519795.
Putz, U., Harwell, C., Nedivi, E., 2005. Soluble CPG15 expressed during early development rescues cortical progenitors from apoptosis. Nat. Neurosci. 8 (3), 322–331. https://doi.org/10.1038/nn1407.
Qiu, J., Bosch, M.A., Tobias, S.C., Krust, A., Graham, S.M., Murphy, S.J., Korach, K.S., Chambon, P., Scanlan, T.S., Rønnekleiv, O.K., Kelly, M.J., 2006. A G-protein-coupled estrogen receptor is involved in hypothalamic control of energy homeostasis. J. Neurosci. 26 (21), 5649–5655. https://doi.org/10.1523/jneurosci.0327-06.2006.
Revankar, C.M., Fau, Cimino Df, Sklar, L.A., La Fau, Sklar, Arterburn, J.B., Fau, Arterburn Jb, Prossnitz, E.R., Prossnitz, E.R., 2005. A transmembrane intracellular estrogen receptor mediates rapid cell signaling. Science. 307 (5715), 1625–1630. https://doi.org/10.1126/science.1106943.
Roman-Blas, J.A., Castañeda, S., Largo, Fau -, Fau, R. Largo R., Herrero-Beaumont, G., Herrero-Beaumont, G., 2009. Osteoarthritis associated with estrogen deficiency. Arthrit. Res. Ther. 11 (5), 241. https://doi.org/10.1186/ar2791.
Rosenberg, N.A., Huang, L., Jewett, Fau -, Fau, E.M. Jewett Em, Szpiech, Z.A., Fau, Szpiech Za, Jankovic, I., Fau, Jankovic I., Boehnke, M., Boehnke, M., 2010. Genome-wide association studies in diverse populations. Nat. Rev. Genet. 11 (5), 356–366. https://doi.org/10.1038/nrg2760.
Rui, L., 2013. Brain regulation of energy balance and body weight. Rev. Endocr. Metab. Disord. 14 (4), 387–407. https://doi.org/10.1007/s11154-013-9261-9.
Salisbury, B.A., Pungliya, M., Choi, Fau -, Fau, J.Y. Choi Jy, Jiang, R., Fau, Jiang R., Sun, X.J., Fau, Sun Xj, Stephens, J.C., Stephens, J.C., 2003. SNP and haplotype variation in the human genome. Mutat. Res. 526 (1–2), 53–61. https://doi.org/10.1016/s0027-5107(03)00014-9.
San, L.-Z., Liu, B.-S., Liu, B., Zhu, K.-C., Guo, L., Guo, H.-Y., Zhang, N., Jiang, S.-G., Zhang, D.-C., 2021. Genome-wide association study reveals multiple novel SNPs and putative candidate genes associated with low oxygen tolerance in golden pompano *Trachinotus ovatus* (Linnaeus 1758). Aquaculture. 544, 737098 https://doi.org/10.1016/j.aquaculture.2021.737098.
Sánchez-Ramos, I., Cross, I., Mácha, J., Martínez-Rodríguez, G., Krylov, V., Rebordinos, L., 2012. Assessment of tools for marker-assisted selection in a marine commercial species: significant association between MSTN-1 gene polymorphism and growth traits. TheScientificWorldJournal. 2012, 369802 https://doi.org/10.1100/2012/369802.
Scherer, A., Christensen, G., 2016. Concepts and relevance of genome-wide association studies. Sci. Prog. 99 (Pt 1), 59–67. https://doi.org/10.3184/003685016X14558068452913.
Shan, X., Xu, T., Ma, Z., Zhang, X., Ruan, Z., Chen, J., Shi, Q., You, X., 2021. Genome-wide association improves genomic selection for ammonia tolerance in the orange-spotted grouper (*Epinephelus coioides*). Aquaculture. 533, 736214 https://doi.org/10.1016/j.aquaculture.2020.736214.
Shan, X., Zhang, X., Ruan, Z., Chen, J., Shi, Q., Xu, J., You, X., 2022. Genomic selection of orange-spotted grouper (*Epinephelus coioides*) based on multiplex PCR enrichment capture sequencing. Aquacult. Fish. https://doi.org/10.1016/j.aaf.2022.08.006. In Press.
Spindel, J., Begum, H., Akdemir, D., Virk, P., Collard, B., Redoña, E., Atlin, G., Jannink, J.L., McCouch, S.R., 2015. Genomic selection and association mapping in rice (*Oryza sativa*): effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. PLoS Genet. 11 (6), e1005350 https://doi.org/10.1371/journal.pgen.1004982.
Su, S., Li, H., Du, F., Zhang, C., Li, X., Jing, X., Liu, L., Li, Z., Yang, X., Xu, P., Yuan, X., Zhu, J., Bouzoualegh, R., 2018. Combined QTL and genome scan analyses with the help of 2b-RAD identify growth-associated genetic markers in a new fast-growing carp strain. Front. Genet. 9, 592. https://doi.org/10.3389/fgene.2018.00592.
Sun, X., Habier, D., Fernando, R.L., Garrick, D.J., Dekkers, J.C., 2011. Genomic breeding value prediction and QTL mapping of QTLMAS2010 data using Bayesian methods. BMC Proc. 5 (Suppl. 3), S13. https://doi.org/10.1186/1753-6561-5-s3-s13.
Sun, Y., Fau, Yu X., Tong, J., Tong, J., 2012. Polymorphisms in Myostatin gene and associations with growth traits in the common carp (*Cyprinus carpio* L.). Int. J. Mol. Sci. 13 (11), 14956–14961. https://doi.org/10.3390/ijms131114956.
Sun, Y., Wen, H., Tian, Y., Mao, X., Li, X., Li, J., Hu, Y., Liu, Y., Li, J., Li, Y., 2021. HSP90 and HSP70 families in *Lateolabrax maculatus*: genome-wide identification, molecular characterization, and expression profiles in response to various environmental stressors. Front. Physiol. 12, 784803 https://doi.org/10.3389/fphys.2021.784803.
Tang, Q., Han, D., Shan, X., Zhang, W., Mao, Y., 2018. Species composition in Chinese aquaculture with reference to trophic level of cultured species. In: Aquaculture in China: Success Stories and Modern Trends, 70-91. https://doi.org/10.1002/9781119120759.ch1_5.
Trudeau, V.L., Fau, Somoza Gm, Nahorniak, C.S., Fau, Nahorniak Cs, Peter, R.E., Peter, R.E., 1992. Interactions of estradiol with gonadotropin-releasing hormone and thyrotropin-releasing hormone in the control of growth hormone secretion in the goldfish. Neuroendocrinology. 56 (4), 483–490.
Tsai, H.Y., Hamilton, A., Guy, D.R., Tinch, A.E., Bishop, S.C., Houston, R.D., 2015a. The genetic architecture of growth and fillet traits in farmed Atlantic salmon (*Salmo salar*). BMC Genet. 16, 51. https://doi.org/10.1186/s12863-015-0215-y.
Tsai, H.Y., Hamilton, A., Tinch, A.E., Guy, D.R., Gharbi, K., Stear, M.J., Matika, O., Bishop, S.C., Houston, R.D., 2015b. Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array. BMC Genomics 16, 969. https://doi.org/10.1186/s12864-015-2117-9.
Wang, C., Ding, X., Liu, J., Yin, Z., Zhang, Q., 2014. Bayesian methods for genomic breeding value estimation. Hereditas. 36 (2), 111–118. https://doi.org/10.3724/sp.j.1005.2014.00111.
Wang, J., Xue, D.X., Zhang, B.D., Li, Y.L., Liu, B.J., Liu, J.X., 2016. Genome-wide SNP discovery, genotyping and their preliminary applications for population genetic inference in spotted sea bass (*Lateolabrax maculatus*). PLoS One 11 (6), e0157809. https://doi.org/10.1371/journal.pone.0157809.
Wang, W., Ma, C., Chen, W., Zhang, H., Kang, W., Ni, Y., Ma, L., 2017. Population genetic diversity of Chinese sea bass (*Lateolabrax maculatus*) from southeast coastal regions of China based on mitochondrial COI gene sequences. Biochem. Syst. Ecol. 71, 114–120. https://doi.org/10.1016/j.bse.2017.01.002.
Wang, J., Zhou, Z., Zhang, Z., Li, H., Liu, D., Zhang, Q., Bradbury, P.J., Buckler, E.S., Zhang, Z., 2018. Expanding the BLUP alphabet for genomic prediction adaptable to the genetic architectures of complex traits. Heredity. 121 (6), 648–662. https://doi.org/10.1038/s41437-018-0075-0.
Wang, Y., Ren, Q., Zhao, L., Li, M., Kong, X., Xu, Y., Hu, X., Hu, J., Bao, Z., 2022. Estimating genetic parameters of muscle imaging trait with 2b-RAD SNP markers in Zhikong scallop (*Chlamys farreri*). Aquaculture. 549, 737715 https://doi.org/10.1016/j.aquaculture.2021.737715.
Willer, C.J., Li, Y., Abecasis, G.R., 2010. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 26, 2190–2191. https://doi.org/10.1093/bioinformatics/btq340.
Wu, M., Chen, G., Li, Y.P., 2016. TGF-β and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease. Bone Res. 4, 16009. https://doi.org/10.1038/boneres.2016.9.
Wu, L., Yang, Y., Li, B., Huang, W., Wang, X., Liu, X., Meng, Z., Xia, J., 2019. First genome-wide association analysis for growth traits in the largest coral reef-dwelling bony fishes, the Giant grouper (*Epinephelus lanceolatus*). Mar. Biotechnol. 21 (5), 707–717. https://doi.org/10.1007/s10126-019-09916-8.
Xu, J., Fau, Wu Rc, O'Malley, B.W., O'Malley, B.W., 2009. Normal and cancer-related functions of the p160 steroid receptor co-activator (SRC) family. Nat. Rev. Cancer 9 (9), 615–630. https://doi.org/10.1038/nrc2695.
Xu, T., Zhang, X., Ruan, Z., Yu, H., Chen, J., Jiang, S., Bian, C., Wu, B., Shi, Q., You, X., 2019. Genome resequencing of the orange-spotted grouper (*Epinephelus coioides*) for a genome-wide association study on ammonia tolerance. Aquaculture. 512, 734332 https://doi.org/10.1016/j.aquaculture.2019.734332.
Yang, J., Lee, S.H., Goddard, M.E., Visscher, P.M., 2011. GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 88 (1), 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011.
Ye, H., Liu, Y., Liu, X., Wang, X., Wang, Z., 2014. Genetic mapping and QTL analysis of growth traits in the large yellow croaker *Larimichthys crocea*. Mar. Biotechnol. 16 (6), 729–738. https://doi.org/10.1007/s10126-014-9590-z.
Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., Yuan, X., Zhu, M., Zhao, S., Li, X., Liu, X., 2021. rMVP: A memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. Genom. Proteom. Bioinformat. 19 (4), 619–628. https://doi.org/10.1016/j.gpb.2020.10.007.
Yoshida, G.M., Lhorente, J.P., Correa, K., Soto, J., Salas, D., Yáñez, J.M., 2019. Genome-wide association study and cost-efficient genomic predictions for growth and fillet yield in Nile Tilapia (*Oreochromis niloticus*). *G3: genes*. Genom. Genet. 9, 2597–2607. https://doi.org/10.1534/g3.119.400116.

You, X., Shan, X., Shi, Q., 2020. Research advances in the genomics and applications for molecular breeding of aquaculture animals. Aquaculture. 526, 735357 https://doi.org/10.1016/j.aquaculture.2020.735357.

Yu, H., You, X., Li, J., Liu, H., Meng, Z., Xiao, L., Zhang, H., Lin, H.R., Zhang, Y., Shi, Q., 2016. Genome-wide mapping of growth-related quantitative trait loci in Orange-spotted grouper (*Epinephelus coioides*) using double digest restriction-site associated DNA sequencing (ddRADseq). Int. J. Mol. Sci. 17 (4), 501. https://doi.org/10.3390/ijms17040501.

Yu, H., Sui, M., Yang, Z., Cui, C., Hou, X., Liu, Z., Wang, X., Dong, X., Zhao, A., Wang, Y., Huang, X., Hu, J., Bao, Z., 2023. Deciphering the genetic basis and prediction genomic estimated breeding values of heat tolerance in Zhikong scallop *Chlamys farreri*. Aquaculture. 565, 739090 https://doi.org/10.1016/j.aquaculture.2022.739090.

Zhang, C., Dong, S.S., Xu, J.Y., He, W.M., Yang, T.L., 2019. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. Bioinformatics. 35 (10), 1786–1788. https://doi.org/10.1093/bioinformatics/bty875.

Zhang, J., Muri, J., Fitzgerald, G., Gorski, T., Gianni-Barrera, R., Masschelein, E., D'Hulst, G., Gilardoni, P., Turiel, G., Fan, Z., Wang, T., Planque, M., Carmeliet, P., Pellerin, L., Wolfrum, C., Fendt, S.M., Banfi, A., Stockmann, C., Soro-Arnáiz, I., Kopf, M., De Bock, K., 2020. Endothelial lactate controls muscle regeneration from ischemia by inducing M2-like macrophage polarization. Cell Metab. 31 (6), 1136–1153.e1137. https://doi.org/10.1016/j.cmet.2020.05.004.

Zhao, J., Bai, H., Ke, Q., Li, B., Zhou, Z., Wang, H., Chen, B., Pu, F., Zhou, T., Xu, P., 2021. Genomic selection for parasitic ciliate *Cryptocaryon irritans* resistance in large yellow croaker. Aquaculture. 531, 735786 https://doi.org/10.1016/j.aquaculture.2020.735786.

Zhou, X., Stephens, M., 2012. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. 44 (7), 821–824. https://doi.org/10.1038/ng.2310.

Zhou, Z., Han, K., Wu, Y., Bai, H., Ke, Q., Pu, F., Wang, Y., Xu, P., 2019. Genome-wide association study of growth and body-shape-related traits in large yellow croaker (*Larimichthys crocea*) using ddRAD sequencing. Mar. Biotechnol. 21 (5), 655–670. https://doi.org/10.1007/s10126-019-09910-0.

Zhu, C., Liu, H., Pan, Z., Chang, G., Wang, H., Wu, N., Ding, H., Yu, X., 2019. Construction of a high-density genetic linkage map and QTL mapping for growth traits in *Pseudobagrus ussuriensis*. Aquaculture. 511, 734213 https://doi.org/10.1016/j.aquaculture.2019.734213.

Zhu, X., Ni, P., Xing, Q., Wang, Y., Huang, X., Hu, X., Hu, J., Wu, X.-L., Bao, Z., 2021. Genomic prediction of growth traits in scallops using convolutional neural networks. Aquaculture. 545, 737171 https://doi.org/10.1016/j.aquaculture.2021.737171.