Marine Biotechnology (2019) 21:655–670
https://doi.org/10.1007/s10126-019-09910-0

ORIGINAL ARTICLE

Genome-Wide Association Study of Growth and Body-Shape-Related Traits in Large Yellow Croaker (*Larimichthys crocea*) Using ddRAD Sequencing

Zhixiong Zhou[1,2] · Kunhuang Han[1,3] · Yidi Wu[2] · Huaqiang Bai[2] · Qiaozhen Ke[1,2] · Fei Pu[1,2] · Yilei Wang[1,3] · Peng Xu[1,2,4]

Received: 23 April 2019 / Accepted: 26 June 2019 / Published online: 22 July 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Large yellow croaker (*Larimichthys crocea*) is an economically important marine fish species of China. Due to overfishing and marine pollution, the wild stocks of this croaker have collapsed in the past decades. Meanwhile, the cultured croaker is facing the difficulties of reduced genetic diversity and low growth rate. To explore the molecular markers related to the growth traits of croaker and providing the related SNPs for the marker-assisted selection, we used double-digest restriction-site associated DNA (ddRAD) sequencing to dissect the genetic bases of growth traits in a cultured population and identify the SNPs that associated with important growth traits by GWAS. A total of 220 individuals were genotyped by ddRAD sequencing. After quality control, 27,227 SNPs were identified in 220 samples and used for GWAS analysis. We identified 13 genome-wide significant associated SNPs of growth traits on 8 chromosomes, and the beta *P* of these SNPs ranged from 0.01 to 0.86. Through the definition of candidate regions and gene annotation, candidate genes related to growth were identified, including important regulators such as *fgf18*, *fgf1*, *nr3c1*, *cyp8b1*, *fabp2*, *cyp2r1*, *ppara*, and *ccm2l.* We also identified SNPs and candidate genes that significantly associated with body shape, including *bmp7*, *col1a1*, *col11a2*, and *col18a1*, which are also economically important traits for large yellow croaker aquaculture. The results provided insights into the genetic basis of growth and body shape in large yellow croaker population and would provide reliable genetic markers for molecular marker-assisted selection in the future. Meanwhile, the result established a basis for our subsequent fine mapping and related gene study.

Keywords *Larimichthys crocea* · Growth-related traits · ddARD · GWAS

Zhixiong Zhou and Kunhuang Han contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10126-019-09910-0) contains supplementary material, which is available to authorized users.

✉ Yilei Wang
ylwang@jmu.edu.cn

✉ Peng Xu
xupeng77@xmu.edu.cn

[1] State Key Laboratory of Large Yellow Croaker Breeding, Ningde Fufa Fisheries Company Limited, Ningde 352130, China

[2] State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen 361102, China

[3] Key Laboratory of Healthy Mariculture for the East China Sea, Fisheries College, Jimei University, Xiamen 361021, China

[4] Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266071, China

Introduction

Growth-related traits are of great importance for many aquatic species as they exert direct influence on production. As the major target in most selective breeding programs, growth-related traits are known to be quantitative traits and controlled by multiple genes across the genome. The genetic basis of growth traits in aquatic species has been widely studied previously, and many important functional genes are revealed (Li et al. 2018; Yue 2014). Among which, genes whose encoded products can serve as carrier proteins or receptors on the somatotropic axis, such as growth hormone (*GH*), growth hormone–releasing hormone (*GHRH*), and insulin-like growth factors (*IGF*), play an import role in the regulation of metabolic and physiological processes of finfish growth (Renaville et al. 2002; De-Santis and Jerry 2007). Most of these genes have been located to be associated with superior growth traits in different fish species (Hu et al. 2013; Tsai et al. 2014; Feng

et al. 2015). In addition, genes encoding hormones in the IGF axis, such as glucocorticoids, gonadotropin, and thyroid hormone, can influence growth traits indirectly (Sternberg and Moav 1999; Wong et al. 2006). Other genes including those in the transforming growth factor superfamily, myogenic regulatory factors (MFR) who regulate the myogenesis, and myostatin (MSTN) which serve as a negative regulator of skeletal muscle in mice have also been identified to be associated with the growth of fish (Sanchez-Ramos et al. 2012; Mcpherron and Lee 1997). Besides, some genes are involved in the metabolic pathway of bone development, lipid, digestion, and so on (Gutierrez et al. 2015; Gonzalez-Pena et al. 2016). Although quite a few genes and their encoded proteins have been found to affect growth in fish, the polygenic nature of growth makes it difficult for us to fully understand the genetic basis. Therefore, more precise localization and identification for genes or single-nucleotide polymorphism (SNP) loci are needed to further explore the genetic mechanism of growth and the subsequent application in aquaculture.

Driven by the blooming of high-throughput genome sequencing technologies over the last decade, new sequencing technologies such as reduced representation sequencing, whole-genome resequencing, and SNP array have been applied in many fish to genotype growth and other important traits (Li et al. 2018; Mehinto et al. 2012; Liu et al. 2015; Fragomeni et al. 2014; Aleman 2017). Combined with the resources of whole genomic sequences and genetic linkage maps, these technologies facilitate the accurate identification and localization for key SNPs, genes, or genetic regions in genomic level (Wang et al. 2018a). For example, 22 quantitative trait loci (QTLs) for growth-related traits were identified in common carp with a 250 K SNP array and QTL mapping association analysis, and potential candidate genes including important regulators such as *kiss2*, *igf1*, *smtlB*, *npffr1*, and *cpe* were also identified (Peng et al. 2016).

Over the last decade, methods for genetic mapping of important traits have also been updated continuously, among which, GWAS and QTL mapping have been proved to be powerful genetic mapping tools in aquaculture (Imumorin et al. 2011). The combination of reduced-respiration sequencing and these genetic mapping methods have been widely used in aquaculture for their low cost and convenience (Li and Wang 2017). Most growth-related genes and genomic regions have been localized with these genetic tools in fishes including catfish, rainbow trout, turbot, tilapia, and so on (Imumorin et al. 2011; Li et al. 2018; Gonzalez-Pena et al. 2016; Wang et al. 2015; Robledo et al. 2016; Lin et al. 2016). As one of the most promising technologies, double-digest RADseq (ddRAD-Seq) reduces the difficulty of genotyping in complex genome by digesting the genomic DNA with two specific restriction enzymes (Peterson et al. 2012). The accuracy in identifying thousands of SNPs and genomic prediction makes it an attractive option for genetic mapping in species without a high-density SNP genotyping array (Robledo et al. 2018; Vallejo et al. 2016). A high-density genetic map of Asian seabass was constructed based on ddRAD genotyping and identified one genome-wide significant and five suggestive QTLs for growth traits (Wang et al. 2015). The polygenic nature of growth-related traits makes it difficult and expensive to mapping functional genes in genomic level; however, the combination of GWAS and ddRAD sequencing could be a cost-effective and convenient approach for accurate localization and identification of growth-related traits in fish.

Large yellow croaker (*Larimichthys crocea*), mainly distributed along the eastern and southern coast of China, is an economically important marine fish in East Asia (Mai et al. 2006; Wu et al. 2014). Although once as one of the four major fishery targets in China, the wild stock of *L. crocea* has nearly depleted due to increasing pressure of overfishing and habitat destruction (Ai et al. 2006; Wang et al. 2012). Moreover, the growing threat of global climate change and disease invasion, as well as the decreased genetic diversity caused by frequent inbreeding, has resulted in problems such as slower growth, smaller size, and poorer disease resistance in cultured *L. crocea* (Ye et al. 2014). Deterioration in germplasm has led to an urgent need for genetic improvement in the growth and resistance to stressful environment. However, understanding the genetic regulation mechanism of growth is the basis for genetic improvement. Nowadays, the increasingly improved sequencing technology and genetic mapping methods has promoted the complete whole genome sequencing and genetic linkage mapping of *L. crocea*, which provides us foundation for precise localization of growth traits in genomic level (Wu et al. 2014; Ao et al. 2015a, b). Here, we used a ddRAD approach to genotype 220 *L. crocea* individuals for GWAS of growth-related traits including body weight (BW), body length (BL), total length (TL), body depth (BD), body thickness (BT), and body-shape-related traits including BL/BD ratio, BL/BT ratio, and body mass index (BMI). Thirteen genome-wide significant SNPs and 40 candidate genes such as *fgf18*, *fgf1*, *nr3c1*, and *cyp8b1* were identified to be associated with the growth, and 8 significant SNPs and 25 candidate genes such as *bmp7*, *col1a1*, *col11a2*, and *col18a1* were identified to be associated with body shape of *L. crocea*. Our study provides new insight into the genetic basis of growth and will facilitate the future genetic mapping and marker-assisted selection in the breeding programs of *L. crocea*.

Materials and Methods

Experimental Fish and Sample Collection

The random mating population of *L. crocea* was generated using 100 female and 100 male parental fish at the National Hatchery Station of Large Yellow Croaker at

Ningde, Fujian Province, China, in March 2015. The fish were reared in a 40-m^2 (2.4 m in depth) cement pool and fed twice daily under standard feeding regime using compound feed (Tianma, Ningde). The water quality was maintained as follows: temperature = 26–27 °C, DO ≥ 5 mg/L, pH = 8–8.5. At the age of 13 months, a total of 314 individuals were randomly collected and the growth-related traits including BW, BL, TL, BD, and BT were measured. BL/BD, BL/BT, and BMI (BW/BL^2) were used to represent the body shape traits of large yellow croaker. The dorsal fin of each fish was then collected and stored in anhydrous ethanol for further analysis.

Genome DNA Extraction

Genome DNA was isolated from dorsal fin sample using standard protocols (Kong et al. 2019). Shredded tissue was dissociated after incubating at 56 °C for about 4 h with 500 μL cell lysis solution and 20 μL protease K (20 mg/mL), then reacted 30 min with 3 μL RNaseA in 36 °C. Subsequently, the protein and DNA were separated by equal volume Tris-phenol, and DNA was extracted again by equal mixed solution of chloroform/isopropanol (24:1). After 2-h precipitation at − 20 °C, genome DNA were collected by brief centrifugation, washed twice with 70% ethanol, air-dried, and resuspended in milli-Q water. Next, DNA concentrations were quantified using spectroscopy by Nanodrop2000 (Thermo Scientific), checked by 1.5% agarose gel electrophoresis stained with ethidium bromide for integrity, and then diluted to 100 ng/μL for genotyping. Finally, a total of 220 DNA samples met the quality requirement for ddRAD library construction.

ddARD Library Construction and Sequencing

Ten ddRAD libraries were constructed by multiplexing 220 individuals following the protocols described previously (Hodel et al. 2017; Zhao et al. 2018). Briefly, 1.5–2 μg of genome DNA from each fish was digested with *Eco*RI and *Msp*I (New England Biolabs, UK; NEB). The P1 adapter with forward amplification primer and a 5 bp barcode was added to the *Eco*RI overhang, and the P2 adapter with reverse amplification primer was added to the *Msp*I overhang, respectively. The DNA fragments of 300–400 bp were retrieved on E-Gel (Thermo Fisher Scientific, SH, China) and then amplified with 20 cycles of PCR with regular forward primer and indexes ligated reverse primers, followed by purifying with AMPure XP Beads (Beckman Coulter, Brea, MA, USA). The obtained ddRAD libraries were sequenced on an Illumina Hiseq2000 platform with 150 bp pair-end strategy at Novogene Corporation (Beijing, China).

SNP Identification and Quality Control

The Illumina sequence adaptors were first truncated from the raw data. Reads with more than 10% "N" and 50% low-quality base ($Q \leq 5$) were discarded. Finally, 1,725,758,354 clean reads with a data size of 258.86 Gb (Table S1) were obtained and filtered using STACKS v2.0 (Catchen et al. 2011, 2013), which was specifically developed to analyze short-read data generated through NGS (Davey et al. 2013). All reads were trimmed to 130 bp using *process_radtags* to avoid the slight sequencing errors after 130 bp. Using BWA v0.7.17 and Samtools v1.8 (Li and Durbin 2009; Li et al. 2009), the rad-tags were aligned to a newly developed *L. crocea* reference genome which was sequenced by PacBio Sequencing platform, assembled with the auxiliary of Hi-C technology and integrated with high-density genetic linkage map (unpublished data). We removed the individuals which genome coverage was lower than 1% or the average sequence depth was lower than 10×. The stacks of each sample and a catalog of loci were constructed; then, samples were matched against the catalog using *gstacks* with default parameters. The program *populations* was then used for genotyping, and only loci presented in at least 75% of the individuals were used for the subsequent analysis. Quality control was implemented by filtering low-confidence SNPs using PLINK v 1.07 (Purcell et al. 2007) with the following parameters: missingness per individual (-mind) > 0.8; minor allele frequency (-maf) < 0.05, individuals genotyping call rate (-geno) < 0.75. Finally, the nondimorphism SNPs were filtered by an in-house script, and the obtained 33,502 SNPs were further reduced to 13,417 by selecting representative SNPs with a distance more than 10 kb from the adjacent SNP.

GWAS Analysis

Basic statistic for all phenotypic data and the corresponding figures were performed and drawn with R. Data were checked for normality, and non-normal data were transformed prior to use in the subsequent analysis. A Pearson's correlation was employed to reflect the relationship between the growth-related traits. Using PLINK, the degree of linkage disequilibrium (LD, characterized by correlation coefficient r^2) among the obtained 13,417 SNPs were estimated in the 24 chromosomes, respectively. The 99th percentile of the r^2 distribution ($r^2 = 0.724$) was set as the background LD, and the 95th percentile ($r^2 = 0.330$) as the cutoff threshold for identifying GWAS candidate regions (see Fig. S2). SNPs with high degree of linkage disequilibrium were pruned by plink, with the parameters of "–indep-pairwise 50 5 0.724." Finally, 27,227 SNPs were used for GWAS analysis of growth-related traits. To display the population components and exclude potential outliers, population structure

analysis was performed before genetic analyses (Finlay et al. 2012; Geng et al. 2015). Multidimensional scaling (MSD) algorithm was implemented to estimate the potential genetic relatedness and thus display population structure of the 220 samples by MSD plot (Zhou et al. 2018).

Association analysis was performed with general linear model (GLM). The significant threshold for GWAS analysis were defined as 3.67×10^{-5} ($1/N$) on genome-wide level and $1/N_c$ on chromosome-wide level according to Bonferroni method, where N and N_c were the number of SNPs used in this study and on a particular chromosome, respectively (Zhong et al. 2017; Benjamini and Hochberg 1995; Gutierrez et al. 2015). The Manhattan plot of the -log10 (P value) and QQ-plot were drawn by R (Figs. 1 and 2).

Candidate Region and Gene Annotation

The linkage disequilibrium (LD) structure was calculated to evaluate whether SNPs in the associated linkage groups were independent (Arendt et al. 2015). To identify the candidate genome regions associated with growth traits, we evaluated the LD degree around each significant SNP that was identified by GWAS. Regions showing higher LD degree ($r^2 > 0.33$) with the significant SNPs (both genome-wide and chromosome-wide) were defined as candidate regions and then mapped onto the reference genome. If the LD degree with the significant SNPs was low, the up- and downstream 50-kb genome regions of the candidate SNPs would be screened for potential candidates. Finally, the obtained candidate regions were mapped to the contig by sequence similarity searches and annotated with an in-house script.

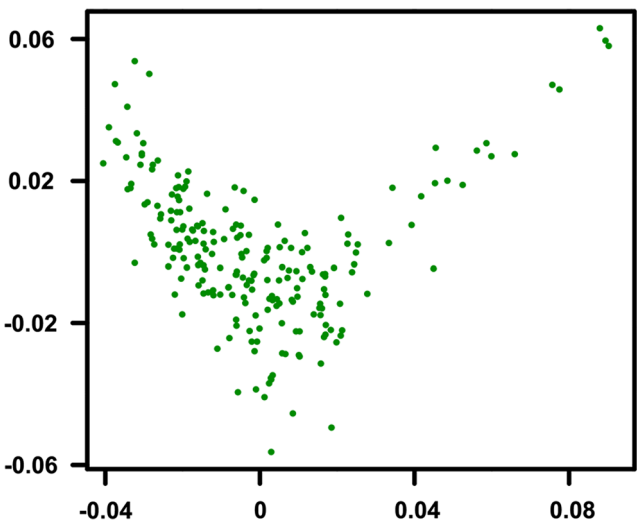

Fig. 1 Multidimensional scaling plot (MSD) of the studied 220 *L. crocea* samples

Results

Genotyping Result

In this study, genotyping was performed by ddRAD in 220 samples of *L. crocea*. After quality control, 1.73 billion reads were obtained for *stacks* assembly, and the average reads for each sample was 3.29 million (Table S2). An average of 3 million reads per sample were mapped to the reference genome with the mapping ratio and mean genome coverage ranged from 77.3 to 94.19% and from 0.49 to 4.45% for each sample, respectively (Table S2). Data of 25 samples were discarded for their low genome coverage (< 1%) or sequencing depth (< 10×). A final catalog consisting 1,215,798 loci was constructed by *gstacks*, of which 149,268 contained at least one SNP. Using *populations*, a total of 21,614 loci were genotyped with at least one SNP per locus in more than 75% of the samples. After filtering, 33,502 SNPs (from 220 samples) distributed on 24 chromosomes and unassembled scaffolds were obtained, with a total genotyping rate of 0.899859. The SNP numbers on each chromosome ranged from 1155 (Chr13) to 1675 (Chr8, see Fig. S1 and Table S3).

Statistics of Growth-Related Traits

As shown in Table 1 and Table S4, the mean BW, BL, TL, BD, and BT were 69.88 ± 32.01 g, 14.67 ± 2.15 cm, 17.34 ± 2.40 cm, 4.09 ± 0.68 cm, and 2.11 ± 0.46 cm, respectively. Shapiro–Wilk tests showed non-normal distributions in BW and BT, and thus, they were normally transformed as normal body weight (NBW) and normal body thickness (NBT) for further analysis (Templeton 2011). Correlation analysis showed strong and extremely significant correlations among the 5 growth traits ($P < 0.01$), with the Pearson correlation coefficient ranged from 0.8967 (BL and BT) to 0.9862 (BT and TL, see Table S5). The high phenotypic correlations among these traits suggest that there might be a coordinated regulatory and possibly same genes, loci, or genomic regions that control the studied growth traits in *L. crocea*.

GWAS of Growth-Related Traits

The associated genomic inflation value (λ) values of five growth-related traits were ranged from 1.27 to 1.37 (Table S6). Based on the genome-wide threshold of $P < 3.67 \times 10^{-5}$, a total of 11 SNPs on 8 different chromosomes were identified to be significant associated with the growth traits (Table 2). Among them, 5 SNPs located on Chr2, Chr7, Chr16, and Chr19 were associated with BW, and the most significant one was LC201285 on Chr16 ($P = 9.95 \times 10^{-6}$, Table 2 and Fig. 3a). Three SNPs located on Chr7 and Chr21 were identified to be associated with BL, and the most significant one was LC245517 on Chr7 ($P = 1.06 \times 10^{-5}$,

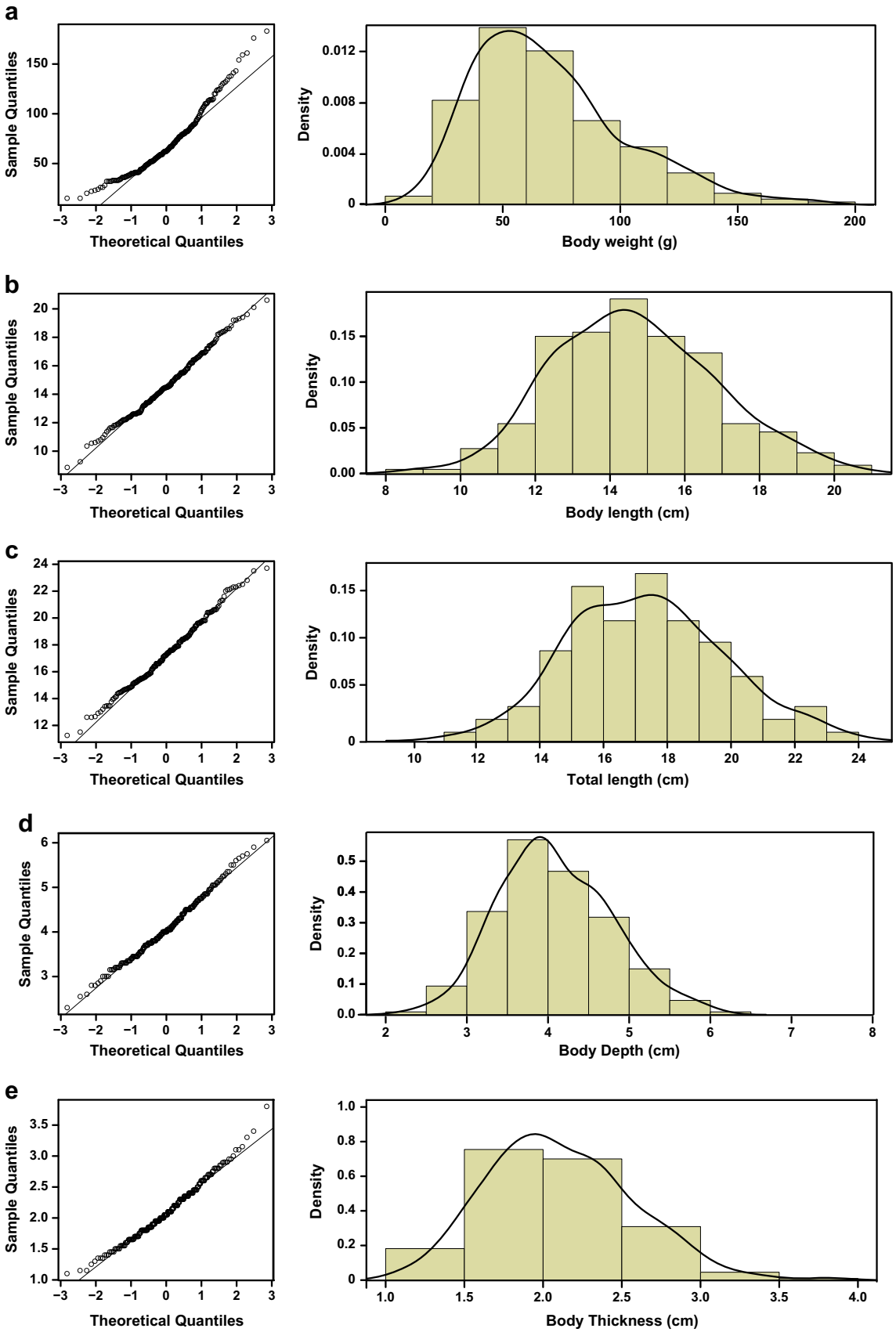

Fig 2 QQ-plots and the corresponding density distribution histograms of body weight (**a**), body length (**b**), total length (**c**), body depth (**d**), and body thickness (**e**). Solid lines in the histograms are the estimated density

Table 1 Statistics of growth traits in *L. crocea*

Traits	Min	Max	Mean	SD	W	Skewness	Kurtosis	*P*
Body weight (g)	15	183	69.88	32.01	0.93	0.65	0.94	8.69×10^{-8}
Body length (cm)	8.85	20.60	14.67	2.15	0.21	− 0.20	0.99	0.43
Total length (cm)	11.25	23.70	17.34	2.40	0.20	− 0.26	0.99	0.38
Body depth (cm)	2.30	6.05	4.09	0.68	0.28	− 0.11	0.99	0.27
Body thickness (cm)	1.10	3.80	2.11	0.46	0.45	0.19	0.99	0.03
Normal body weight (cm)	20.55	124.13	70.13	19.90	0.06	− 0.20	0.99	0.99
Normal body thickness (cm)	0.91	3.55	2.12	0.46	0.10	− 0.06	0.99	0.99

$N = 220$; W and P were the statistics of Shapiro–Wilk test in R; normal body weight and normal body thickness were the fitting value of BW and BT after normality transformation

Min minimum, *Max* maximum

Table 2 and Fig. 3b). One SNP located on Chr2 (LC164786, $P = 2.88 \times 10^{-5}$) was associated with TL (Table 2 and Fig. 3c). Three SNPs on Chr2, Chr5, and Chr7 were associated with BD, and the most significant one was LC217665 on Chr5 ($P = 5.16 \times 10^{-6}$, Table 2 and Fig. 3d). Finally, 6 SNPs on Chr7, Chr16, Chr21, Chr22, and Chr23 were associated with BT and the most significant one was LC201285 on Chr5 ($P = 9.96 \times 10^{-6}$, Table 2 and Fig. 3e).

As shown in Table 2, 5 SNPs were found to be significantly associated with not only one of the studied growth traits. More specifically, LC245517 was associated with all the studied traits except TL; LC245784 was associated with BW, BL, and BT; LC201285 was associated with both BW and BT; LC164786 was associated with BW and TL; and LC124204 was associated with BL and BT.

Genomic Regions Associated with Growth-Related Traits and Gene Annotation

As described above, the studied 5 growth-related traits showed strong and significant correlations, and 5 out of 11 genome-wide significant SNPs were identified to be significantly associated with 2–4 traits simultaneously. Therefore, candidate regions identified from these SNPs and jointly chromosome-wide significant SNPs were possibly associated with multiple growth-related traits. We combined these regions to expand the range of annotation and finally identified 34 candidate regions associated with multiple growth traits (Table S7). Besides, for growth-related trait identification of the 44 regions, 7 candidate regions on 5 chromosomes were BW-specific (CRW, candidate regions of BW), 3 regions on 2

Table 2 Genome-wide significant SNPs associated with growth-traits identified by GWAS analysis

	SNP	Chr	Position (bp)	P	Major_allele	Minor_allele	Beta *P*	SE (Beta *P*)
BW	LC201285	16	18,540,921	9.95×10^{-6}	T	C	0.83	0.39
	LC245517	7	2,113,540	1.05×10^{-5}	C	G	− 0.78	0.20
	LC245784	7	4,068,020	2.24×10^{-5}	A	T	− 0.80	0.22
	LC164786	2	12,825,691	2.84×10^{-5}	G	A	0.12	0.35
	LC256158	19	18,049,045	3.19×10^{-5}	C	T	0.86	0.26
BL	LC245517	7	2,113,540	1.06×10^{-5}	C	G	− 0.86	0.22
	LC245784	7	4,068,020	3.18×10^{-5}	A	T	0.87	0.24
	LC124204	21	14,041,748	3.55×10^{-5}	G	A	0.21	0.25
TL	LC164786	2	12,825,691	2.88×10^{-5}	G	A	1.39	0.43
BD	LC217665	5	11,392,643	5.16×10^{-6}	G	A	0.59	0.13
	LC165626	2	14,064,545	5.59×10^{-6}	T	A	0.21	0.06
	LC245517	7	2,113,540	9.68×10^{-6}	C	G	− 0.27	0.07
BT	LC201285	16	18,540,921	9.96×10^{-6}	T	C	0.20	0.09
	LC102497	22	5,806,271	1.42×10^{-5}	G	A	0.31	0.05
	LC177745	23	6,315,017	2.05×10^{-5}	T	A	− 0.16	0.09
	LC245784	7	4,068,020	2.41×10^{-5}	A	T	0.18	0.05
	LC124204	21	14,041,748	2.90×10^{-5}	G	A	0.01	0.05
	LC245517	7	2,113,540	2.94×10^{-5}	C	G	− 0.17	0.05

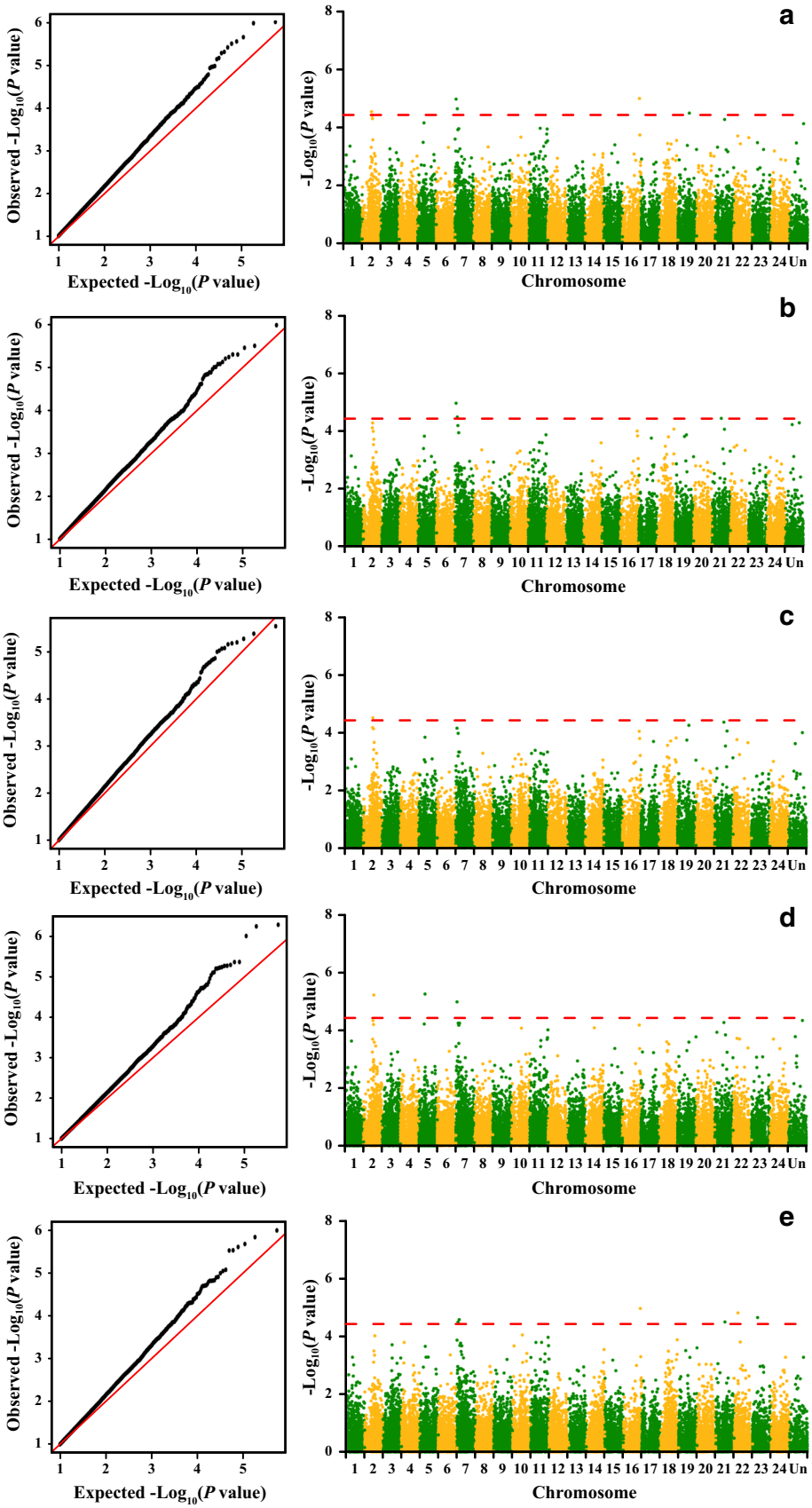

Fig. 3 QQ-plots and Manhattan plots for body weight (**a**), body length (**b**), total length (**c**), body depth (**d**), and body thickness (**e**) in *L. crocea*. Red dotted lines represent the threshold −log10 (*P* value) for genome-wide significance

chromosomes and contig 110 were TL-specific (CRL, candidate regions of BL), 8 regions on 8 chromosomes were BD-specific (CRD, candidate regions of BD), and 26 regions on 15 chromosomes were BT-specific (CRT, candidate regions of BT). The details are shown in Table S7.

After mapping these candidate regions to the reference genome, 13 genes including *fgf18*, *ganab*, *acsl5*, *lrp2*, *aggf1*, *slc6a2*, *mapk9*, *arpc4*, *megf6*, *hs3st3b1*, *tbata*, *pkm*, and *cyp2r1* were annotated to be associated with multiple growth-related traits in *L. crocea* (Table 3). Gene annotation

Table 3 Summary of growth-related candidate genes identified from GWAS in *L. crocea*

Traits	CR	Chr	Gene	Annotation
Growth	CRG2	2	*acsl5*	Long-chain-fatty-acid--CoA ligase 5
	CRG3	3	*hs3st3b1*	Heparan sulfate glucosamine 3-O-sulfotransferase 3B1
	CRG5	2	*lrp2*	Low-density lipoprotein receptor-related protein 2
	CRG7	2	*tbata*	Protein TBATA
	CRG10	5	*aggf1*	Angiogenic factor with G patch and FHA domains 1
	CRG11	5	*arpc4*	Actin-related protein 2/3 complex subunit 4
	CRG13	7	*ganab*	Neutral alpha-glucosidase AB
	CRG15	7	*fgf18*	Fibroblast growth factor 18
	CRG16	7	*mapk9*	Mitogen-activated protein kinase 9
	CRG22	11	*slc6a2*	Sodium-dependent noradrenaline transporter
		11	*pkm*	Pyruvate kinase PKM
	CRG24	11	*cyp2r1*	Vitamin D 25-hydroxylase
	CRG28	18	*megf6*	Multiple epidermal growth factor-like domains protein 6
BW	CRW2	8	*ccm2l*	Cerebral cavernous malformations 2 protein-like
	CRW4	11	*cdh15*	Cadherin-15
	CRW7	20	*csgp4*	Chondroitin sulfate proteoglycan 4
TL	CRL2	10	*bach1*	Transcription regulator protein BACH1
	CRL3	ctg110*	*dynlrb1*	Dynein light chain roadblock-type 1
BD	CRD2	3	*mthfd1*	C-1-tetrahydrofolate synthase, cytoplasmic
	CRD6	17	*pitpnm2*	Membrane-associated phosphatidylinositol transfer protein 2
	CRD6	17	*mzt2*	Mitotic-spindle organizing protein 2
	CRD7	19	*jarid2*	Protein Jumonji
	CRD8	20	*ppara*	Peroxisome proliferator-activated receptor alpha
	CRD8	20	*cdkn1b*	Cyclin-dependent kinase inhibitor 1B
BT	CRT2	3	*slc4a11*	Sodium bicarbonate transporter-like protein 11
	CRT5	5	*paqr7*	Membrane progestin receptor alpha
	CRT7	7	*myot*	Myotilin
	CRT8	9	*map2k4*	Dual specificity mitogen-activated protein kinase kinase 4
	CRT9	10	*ccni*	Cyclin-I
		10	*med37c*	Heat shock 70 kDa protein 4
		10	*fgf1*	Fibroblast growth factor 1
		10	*nr3c1*	Glucocorticoid receptor
	CRT14	11	*glce*	D-glucuronyl C5-epimerase
	CRT16	13	*dmd*	Dystrophin
	CRT17	14	*scap*	Sterol regulatory element-binding protein cleavage-activating protein
		14	*cspg5*	Chondroitin sulfate proteoglycan 5
	CRT18	18	*st3gal4*	CMP-N-acetylneuraminate-beta-galactosamide-alpha-2,3-sialyltransferase 4
	CRT20	20	*man2c1*	Alpha-mannosidase 2C1
	CRT26	24	*prkcsh*	Glucosidase 2 subunit beta
		24	*fabp2*	Fatty acid-binding protein, intestinal

CR candidate region

*Contig that is unassembled on chromosomes

showed that these genes are involved in the fibroblast growth factor signaling pathway, glucose metabolism, fatty acid metabolism, vasculogenesis, hypothalamus-hypophysis-adrenal axis, cell proliferation, actin development, epidermal growth factor pathway, vitamin D metabolic, and so on. In addition, another 27 genes were annotated to be trait-specific (Table 3). Among which, genes including *ccm2l*, *cdh15*, and *cspg4* are BW-specific, and they play roles in the Heart of Glass-Cerebral Cavernous Malformation (Heg-CCM) pathway, muscle differentiation, and cell proliferation. Two genes (*bach1* and *dynlrb1*) which involved in the cell mitotic and cytoskeleton formation are TL-specific. For BD, 6 specific genes (*mthfd1*, *pitpnm1*, *ppara*, *mzt2*, *jarid2*, and *cdkn1b*) which involved in the folic acid metabolic, lipid metabolic, cellular fission, and embryonic development are annotated. A total of 16 genes were annotated to be BT-specific, and they play a part in pathway including fibroblast growth factor signaling pathway, glucocorticoid signaling pathway, cholesterol metabolic, bile acid biosynthetic, muscle development, fatty acid metabolic, and oocyte maturation.

GWAS of Body-Shape-Related Traits and Gene Annotation

The associated genomic inflation value (λ) values of three body-shape-related traits ranged from 1 to 1.31 (Table S6), and the phenotypes of body-shape-related traits are listed in Table S8. Based on the genome-wide threshold of $P < 3.67 \times 10^{-5}$, no SNP were identified to be significant associated with the BL/BD (Fig. 5a). LC12941 on Chr15 ($P = 9.96 \times 10^{-6}$, Table S9 and Fig. 5b) was the most significant SNP of BL/BT. Seven SNPs located on Chr5, Chr11, Chr15, Chr16, and Chr22 were identified to be associated with BMI, and the most significant one was LC104526 ($P = 1.07 \times 10^{-7}$, Table S9 and Fig. 5c).

Of the 34 identified genome regions that associated to body-shape-related traits, 14 candidate regions on 12 chromosomes were BL/BD-specific (CRLD, candidate regions of BL/BD), 7 regions on 4 chromosomes were BL/BT-specific (CRLT, candidate regions of BL/BT), 13 regions on 8 chromosomes were BMI-specific (CRM, candidate regions of BMI). The details are shown in Table S10. After mapping these candidate regions to the reference genome, 11 genes including *glcci1*, *col1a1*, *col11a2*, *col18a1*, and *lipg* were annotated to be associated with BL/BD trait in *L. crocea* (Table 4). For BL/BT, 5 specific genes (*slc34a2*, *slc35g1*, *rbp4a*, *glmn*, and *grm8*) which involved in the ion transport and visual perception are annotated (Table 4). A total of 9 genes were annotated to be candidate genes for BMI, and they play a part in pathway including glucocorticoid signaling pathway, bone development, lipid catabolic process, and muscle development (Table 4).

Discussion

Genetic markers can be used to directly select promising parents, which avoids the deficiency in aquatic fish that well-documented pedigree information is always lacking. However, the accuracy of marker-assisted selection depends largely on the existing genomic resources and the numbers of genotyped samples. Generally, the more sample the more cost, which limits the widespread application of marker-assisted selection in aquaculture. In this study, we used a ddRAD approach to genotype SNP loci associated with five important growth traits in *L. crocea*. Based on two RE digestion and pooling sequencing, genotyping was enormously simplified and a total of 27,227 SNPs were identified in only 220 samples. Using these SNPs, the subsequent GWAS analysis identified 11 genome-wide significant SNPs associated with growth traits and 8 genome-wide significant SNPs associated with body shape traits.

Growth-Related Candidate Genes

As a crucial economic trait of *L. crocea*, extensive investigation about growth has been conducted in several methods. For example, with extreme phenotypic sampling, a GWAS for n-3 highly unsaturated fatty acids (HUFA) and eviscerated weight (EW) traits in *L. crocea* has been conducted, which suggests that genes including *fabp*, *dgat*, *atp8b1*, *faf2*, and *cers2*, as well as *igf2*, *broa*, *cyp1a1*, *crtp1*, and *hox* genes are potential candidate genes of n-3 HUFA and EW, respectively (Wan et al. 2018). In the present study, 11 genome-wide significant SNPs on 8 chromosomes were identified to be significantly associated with growth-related traits of *L. crocea*, and the subsequent analysis identified 78 candidate regions and 40 candidate genes, of which 34 candidate regions and 13 candidate genes were associated with multiple growth-related traits (Table 2 and Table 3). The beta *P* of significant SNPs ranged from 0.12 to 0.86 in BW, 0.21 to 0.87 in BL, 0.39 in TL, 0.21–0.59 in BD, and 0.01–0.31 in BD, and it betoken the reliable efforts in the future genetic improvement of growth traits (Table 2). The identified candidate genes serve essential functions in multiple biological processes closely related to growth and various metabolism procedures, such as glucose and fatty acid metabolism, vasculogenesis, and cell proliferation. We found fibroblast growth factor 18 (*fgf18*) in the candidate region CRG15 on chromosome 7 (Fig. 4a). *Fgf18* is a member of FGF family which is an essential component in the regulatory of cell growth, differentiation, skeletal growth, survival, and numerous developmental processes (Muenke and Schell 1995; Su et al. 2014). Previous study demonstrated that *fgf18* is an important regulator in the developmental processes of both chondrogenesis and osteogenesis. In mouse osteoblasts, *fgf18* promotes osteogenesis by upregulated expression of Bmp2 as well as upregulation or maintenance of Fgfr1,

Table 4 Summary of body-shape-related candidate genes identified from GWAS in *L. crocea*

Traits	CR	Chr	GENE	Annotation
BL/BD	CRLD3	6	*glcci1*	Glucocorticoid-induced transcript 1 protein
	CRLD4	8	*klf15*	Krueppel-like factor 15
	CRLD5	10	*ppm1a*	Protein phosphatase 1A
	CRLD5	10	*col1a1*	Collagen alpha-1(I) chain
	CRLD8	17	*col11a2*	Collagen alpha-2(XI) chain
	CRLD10	18	*b3glct*	Beta-1,3-glucosyltransferase
	CRLD11	19	*col18a1*	Collagen alpha-1(XVIII) chain
	CRLD12	20	*tgm2*	Protein-glutamine gamma-glutamyltransferase 2
	CRLD13	22	*lipg*	Endothelial lipase
	CRLD13	22	*cspg4*	Chondroitin sulfate proteoglycan 4
	CRLD14	22	*dbnl*	Drebrin-like protein
BL_BT	CRLT2	1	*slc34a2*	Sodium-dependent phosphate transport protein 2B
	CRLT4	2	*slc35g1*	Solute carrier family 35 member G1
	CRLT4	2	*rbp4a*	Retinol-binding protein 4-A
	CRLT6	15	*glmn*	Glomulin
	CRLT7	23	*grm8*	Metabotropic glutamate receptor 8
BMI	CRM3	3	*sipa1l1*	Signal-induced proliferation-associated 1-like protein 1
	CRM3	3	*pcdh15*	Protocadherin-15
	CRM5	5	*cdh26*	Cadherin-like protein 26
	CRM5	5	*bmp7*	Bone morphogenetic protein 7
	CRM5	5	*dtnbp1*	Dysbindin
	CRM5	5	*arpc4*	Actin-related protein 2/3 complex subunit 4
	CRM7	6	*lpcat1*	Lysophosphatidylcholine acyltransferase 1
	CRM9	15	*hs2st1*	Heparan sulfate 2-O-sulfotransferase 1
	CRM10	16	*liph*	Lipase member H

Fgfr2, and Fgfr3 expression (Nagayama et al. 2013). Another candidate gene in FGF family, *fgf1*, was also identified to be associated with BT in our study (Fig. 4b). This suggests that FGF family may play an important role in the skeletal

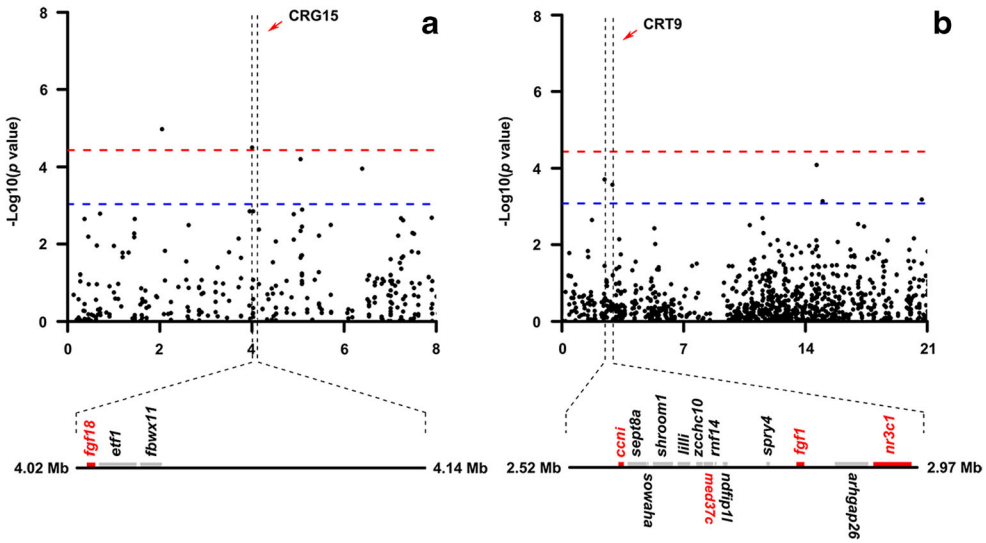

Fig. 4 Regional genome scan for two genomic regions significantly associated with growth. **a** Regional amplification of the candidate region CRG15 identified from the significantly growth-related SNP LC245517 on Chr7. Manhattan plot was based on GWAS result of body length. **b** Regional amplification of the candidate region CRT19 identified from the SNPs significantly associated with body thickness on Chr10. Manhattan plot was based on GWAS result of body thickness. Red and blue dotted line represent the threshold −log10 (*P* value) for genome-wide significance and −log10 (*P* value) for chromosome-wide significance, respectively

development of *L. crocea*, and subsequently affect multiple growth traits. Except for genes in FGF family, other important gene such as *cyp2r1* was also identified in this group. Generally, *cyp2r1* controls the production of 25-hydroxylase which affects the activation of vitamin D, thereby influences the formation of bones and teeth. Evidences have been provided that *cyp2r1* can also influence the susceptibility to type 1 diabetes in human (Ramos-Lopez et al. 2007). Therefore, we inferred that *cyp2r1* may play an important role in the growth of *L. crocea* by regulating the metabolic process of vitamin D. Combined with the genotyping and analytical methods we used, several novel candidate genes including *acsl5*, *ganab*, *aggf1*, and *slc6a2* were identified to be associated with multiple growth-related traits.

Besides, four genes (*ccni*, *med37c*, *fgf1*, and *nr3c1*) located within 0.45 Mb of chromosome 10 were identified to be associated with BT (Fig. 4b). Glucocorticoid receptor (*nr3c1*) is essential for the maintenance of energy supply and growth control (Mueller et al. 2012). It regulates various metabolic reactions (including insulin-stimulated glucose uptake and promoting proteolysis) in skeletal muscle and exerts effects on growth and fatness of both human and pig (Kuo et al. 2012) (Fig. 5). In mice, *nr3c1* can act as a coactivator for Stat5-dependent transcription upon growth hormone (GH) stimulation and thereby affect body growth (Tronche et al. 2004). Therefore, *nr3c1* may indirectly influence the BT of *L. crocea* by participating in the GH/IGF-I axis (Fig. 6a). Other gene such as *fabp2* on chromosome 24 was also identified to be associated with BT. The encoded protein of *fabp2* is intestinal fatty acid-binding protein (I-FABP) which is known to take part in the intracellular transport of long-chain fatty acids (Glatz et al. 1998). I-FABP can also influence insulin resistance and dietary fat in humans (Pratley et al. 2000). The mapping result we found suggests that *fabp2* influence the BT of *L. crocea* by affecting the fatty acid metabolism in intestine. Two genes on chromosome 20 were associated with BD, of which, *ppara* (peroxisome proliferator-activated receptor alpha) is a ligand-activated transcription

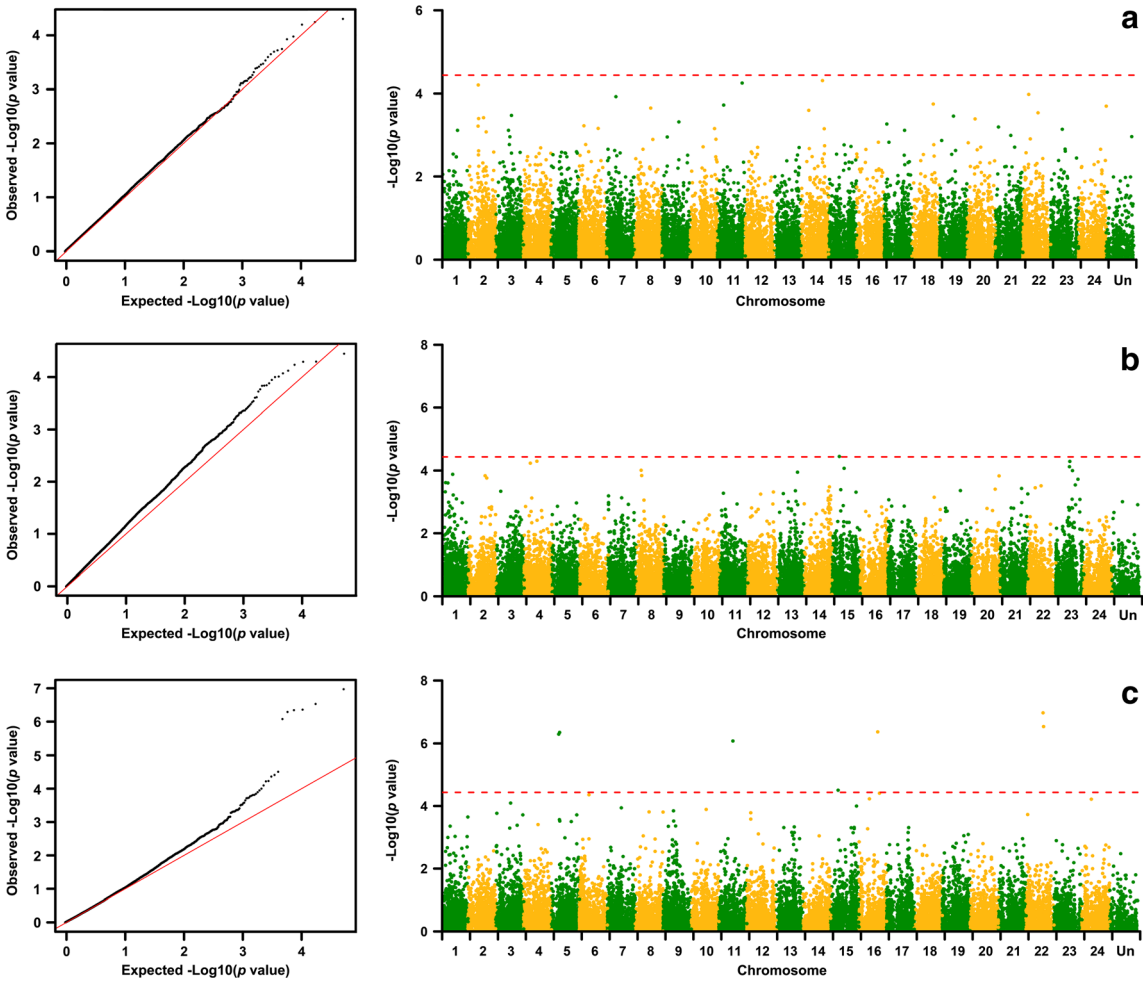

Fig. 5 **a** QQ-plots and Manhattan plots for BL/BD ratio (**a**), BL/BT ratio (**b**), and BMI (**c**) in *L. crocea*. Red dotted lines represent the threshold −log10 (*P* value) for genome-wide significance

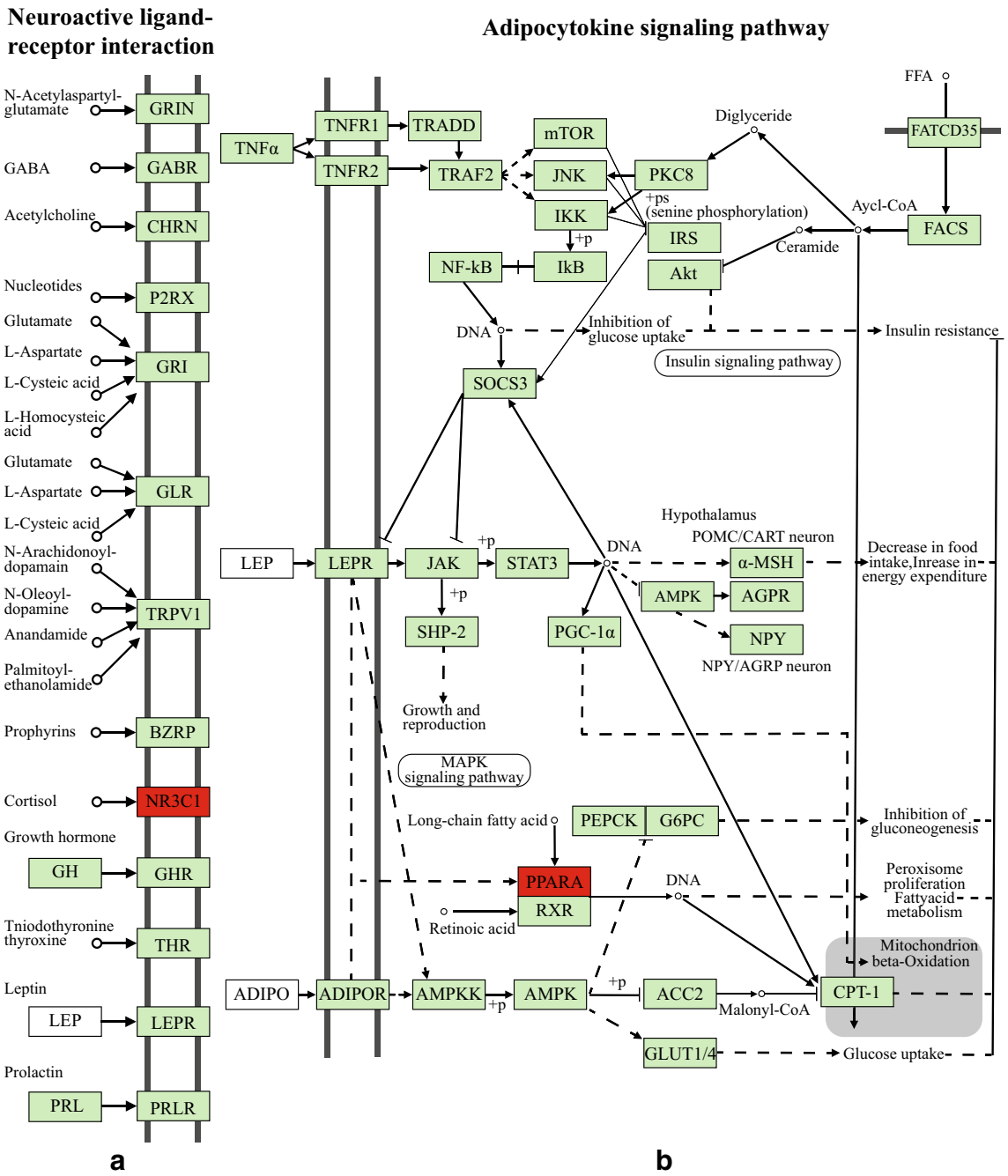

Fig. 6 **a** The Neuroactive ligand-receptor interaction pathway in *L. crocea*. Gene *n3c1* was marked by red block. **b** The Adipocytokine pathway in *L. crocea*. Gene *ppara* was marked by red block

factor in adipocytokine signaling pathway (Fig. 6b), and a key regulator of lipid metabolism in liver and regulates the peroxisomal beta-oxidation pathway of fatty acids, whose expression can protect male mice from high fat (Abdelmegeed et al. 2011). In human, several GWAS showed that *ppara* were associated with the concentrations of growth factors (Ahola-Olli et al. 2017) and lipid levels (Willer et al. 2013). By eGWAS in pig, *ppara* were identified as key regulators of lipid metabolism and overexpressed in males (Puig-Oliveras et al. 2016). Dietary fatty acids could regulate the gene expression of *srebf1* and *ppara* in Japanese seabass, and these two genes influence the long-chain polyunsaturated fatty acids (LC-PUFAs) through regulating *fads2* transcription in rainbow trout, Japanese seabass, and large yellow croaker (Dong et al. 2015; Dong et al. 2017). Thus, we speculated that *ppara* and *fabp2* could influence the fat metabolism and relate to abdominal fat accumulation in *L. crocea*. Overall, the comparison between BD-specific and BT-specific genes revealed that two common biological processes including lipid metabolism and some reproduction process (such as folic acid

metabolic and oocyte maturation) may be critical in determining the growth of *L. crocea*.

In addition to the genes and biological processes discussed above, cardiovascular development is also important to growth, especially in the early developmental stages of fish. During the cardiovascular development of zebrafish and mouse, Heart of Glass-Cerebral Cavernous Malformation (Heg-CCM) pathway is necessary (Rosen et al. 2013). Wherein, cerebral cavernous malformations 2 protein-like (*ccm2l*) is a critical component in Heg-CCM pathway and it regulates vessel stability and growth during cardiovascular development (Zheng et al. 2012). In the present study, *ccm2l* was identified to be associated with BW, suggesting the important role of cardiovascular development in the growth and production of *L. crocea*. However, phenotypic growth-related traits of the fish in this study were collected only at the age of 13 months when sexual maturity had yet been reached. Previous study showed that growth, as typical polygenic traits, might be controlled by different major genes at different life stages of fish.

Body-Shape-Related Candidate Genes

Due to differences on culture conditions (such as stocking density, nutrition, and environment), there is a significant difference between the body shape of cultured and wild *L. crocea* groups. Generally, the cultured croakers show shorter and thicker body shapes, whereas the wild ones are more slender. As a result of market preference in China, *L. crocea* individuals with slender body shape usually have higher market value. A GWAS for BL/BD of *L. crocea* identified 4 QTLs and 10 candidate genes, including *fgf13*, *fgfr3*, *tubgcp3*, *erbb4*, *phactr1*, *cyp26b1*, *wnt3a*, *myh10*, *mgp*, and *arrb2*, and all these genes were related to bone development (Dong et al. 2019). In the present study, we identified 25 candidate genes to be associated with body-shape-related traits of *L. crocea* (Table S10 and Table 4). The beta *P* of significant SNPs was 0.07 in BL/BT and less than 0.01 in BMI (Table S9). In view of the slight efforts in these significant SNPs, we need to further identify body-shape-related loci with bigger sample size. Gene annotation indicated that these genes involve in multiple biological processes, such as glucose and lipid metabolism, bone development, and muscle development. For example, a gene on chromosome 5—bone morphogenetic protein 7 (*bmp7*)—was identified from the candidate region CRM5. It belongs to the bone morphogenetic protein gene family, which induced cartilage and bone formation, and played an important role in bone homeostasis and calcium regulation (Hogan 1996). Many members of this gene family has been clearly identified to regulate bone development and body shape in fish (Peng et al. 2016; Chen et al. 2018). In chicken (*Gallus gallus*), *bmp7* was identified to be associated with growth and carcass traits (Wang et al. 2018b). We speculate that the BMP gene family may be one of the important regulatory factors of body shape in *L. crocea*. Moreover, in the GWAS of BL/BD, we identified three genes that encode type I collagen, including *col1a1*, *col11a2*, and *col18a1*. Type I collagen was an important component of bone, muscle, scales, and skin in teleost fish and involved in the growth and development of many important tissues (Kimura 2002). Type I collagen may take part in bone and muscle development and indirectly affect the body morphological character of *L. crocea*. In addition to the genes discussed above, we also identified some candidate genes related to glucose and lipid metabolism, as *glcci1*, *b3cglct*, *liph*, and *lipg*, which indicated that body shape was closely related to lipid metabolism like growth traits. Further exploration on the functions of these candidate genes is necessary to better understand their roles on body shape determination, and facilitates genetic breeding of market favored body shape of *L. crocea*.

Overall, we performed a ddRAD sequencing on 220 large yellow croakers and obtained 33,502 high-quality SNPs. Using these SNPs, we identified 18 significant SNPs and multiple candidate genes on 14 chromosomes. Several novel genes were also identified to be associated with the growth of *L. crocea*. Although the functions of these candidate genes have been annotated and discussed, future studies are still needed to verify and further explore the polygenic regulation mechanisms underlying the growth of fish. The current study suggests that ddRAD sequencing is a satisfactory and cost-effective tool to obtain high-throughput genotyping data in aquatic fish. The identified SNP markers can be useful tools in marker-assistance parental fish selection. Finally, our study provides new insight into the genetic basis of growth and will facilitate the future genetic mapping and marker-assisted selection in the breeding programs of *L. crocea*.

Authors' Contribution PX and YLW conceived and supervised the study. ZZ construct the ddRAD library, analyzed the data, and drafted the manuscript. HB helped on data analysis. FP, KH, and QK helped with croaker breeding and sample collection. PX and YDW revised the manuscript. All authors have read and approved the manuscript.

Funding information We acknowledge financial support from the State Key Laboratory of Large Yellow Croaker Breeding (Fujian Fuding Seagull Fishing Food Co., Ltd) (Nos. LYC2017ZY01, LYC2017RS05, and LYC2016RS02), the Fundamental Research Funds for the Central Universities (Nos. 20720180123 and 20720160110), the Science and Technology Platform Construction of Fujian Province (No. 2018N2005), the Natural Science Foundation of Fujian Province, China (Nos. 2015J06019 and 2017J06022), the Local Science and Technology Development Project Guide by The Central Government (No. 2017L3019), the National Key Research and Development Program of China (No. 2018YFC1406305), China Agriculture Research System (No. CARS-47), and the Major Special Projects of Fujian Province (No. 2016NZ0001).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval This study was approved by the Animal Care and Use committee of State Key Laboratory of Large Yellow Croaker Breeding. All the methods used in this study were carried out in accordance with approved guidelines.

References

Abdelmegeed MA, Yoo SH, Henderson LE, Gonzalez FJ, Woodcroft KJ, Song BJ (2011) PPARalpha expression protects male mice from high fat-induced nonalcoholic fatty liver. J Nutr 141:603–610

Ahola-Olli AV, Wurtz P, Havulinna AS, Aalto K, Pitkanen N, Lehtimaki T, Kahonen M, Lyytikainen LP, Raitoharju E, Seppala I, Sarin AP, Ripatti S, Palotie A, Perola M, Viikari JS, Jalkanen S, Maksimow M, Salomaa V, Salmi M, Kettunen J, Raitakari OT (2017) Genome-wide association study identifies 27 loci influencing concentrations of circulating cytokines and growth factors. Am J Hum Genet 100: 40–50

Ai QH, Mai KS, Tan BP, Xu W, Zhang WB, Ma HM, Liufu ZG (2006) Effects of dietary vitamin C on survival, growth, and immunity of large yellow croaker, Pseudosciaena crocea. Aquaculture 261:327–336

Aleman F (2017) The necessity of diploid genome sequencing to unravel the genetic component of complex phenotypes. Front Genet 8:148

Ao JQ, Li J, You XX, Mu YN, Ding Y, Mao KQ, Bian C, Mu PF, Shi Q, Chen XH (2015a) Construction of the high-density genetic linkage map and chromosome map of large yellow croaker (Larimichthys crocea). Int J Mol Sci 16:26237–26248

Ao JQ, Mu YN, Xiang LX, Fan DD, Feng MJ, Zhang SC, Shi Q, Zhu LY, Li T, Ding Y, Nie L, Li QH, Dong WR, Jiang L, Sun B, Zhang XH, Li MY, Zhang HQ, Xie SB, Zhu YB, Jiang XT, Wang XH, Mu PF, Chen W, Yue Z, Wang Z, Wang J, Shao JZ, Chen XH (2015b) Genome sequencing of the perciform fish Larimichthys crocea provides insights into molecular and genetic mechanisms of stress adaptation. Plos Genet 11:e1005118

Arendt ML, Melin M, Tonomura N, Koltookian M, Courtay-Cahen C, Flindall N, Bass J, Boerkamp K, Megquir K, Youell L, Murphy S, Mccarthy C, London C, Rutteman GR, Starkey M, Lindblad-Toh K (2015) Genome-wide association study of golden retrievers identifies germ-line risk factors predisposing to mast cell tumours. Plos Genet 11:e1005647

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. J R Stat Soc B Methodol 57:289–300

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. G3 (Bethesda) 1:171–182

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. Mol Ecol 22: 3124–3140

Chen L, Peng WZ, Kong SN, Pu F, Chen BH, Zhou ZX, Feng JX, Li XJ, Xu P (2018) Genetic mapping of head size related traits in common carp (Cyprinus carpio). Front Genet 9:448

Davey JW, Cezard T, Fuentes-Utrilla P, Eland C, Gharbi K, Blaxter ML (2013) Special features of RAD sequencing data: implications for genotyping. Mol Ecol 22:3151–3164

De-Santis C, Jerry DR (2007) Candidate growth genes in finfish - where should we be looking? Aquaculture 272:22–38

Dong X, Xu H, Mai K, Xu W, Zhang Y, Ai Q (2015) Cloning and characterization of SREBP-1 and PPAR-α in Japanese seabass Lateolabrax japonicus, and their gene expressions in response to different dietary fatty acid profiles. Comp Biochem Physiol B Biochem Mol Biol 180:48–56

Dong X, Tan P, Cai Z, Xu H, Li J, Ren W, Xu H, Zuo R, Zhou J, Mai K (2017) Regulation of FADS2 transcription by SREBP-1 and PPAR-α influences LC-PUFA biosynthesis in fish. Sci Rep 7:40024

Dong L, Han Z, Fang M, Xiao S, Wang Z (2019) Genome-wide association study identifies loci for body shape in the large yellow croaker (*Larimichthys crocea*). Aquac Fish 4:3–8

Feng X, Yu XM, Pang MX, Liu HY, Tong JO (2015) Molecular characterization and expression of three preprosomatostatin genes and their association with growth in common carp (Cyprinus carpio). Comp Biochem Physiol B 182:37–46

Finlay EK, Berry DP, Wickham B, Gormley EP, Bradley DG (2012) A genome wide association scan of bovine tuberculosis susceptibility in Holstein-Friesian dairy cattle. PLoS One 7:e30545

Fragomeni BD, Misztal I, Lourenco DL, Aguilar I, Okimoto R, Muir WM (2014) Changes in variance explained by top SNP windows over generations for three traits in broiler chicken. Front Genet 5:332

Geng X, Sha J, Liu S, Bao L, Zhang J, Wang R, Yao J, Li C, Feng J, Sun F, Sun L, Jiang C, Zhang Y, Chen A, Dunham R, Zhi D, Liu Z (2015) A genome-wide association study in catfish reveals the presence of functional hubs of related genes within QTLs for columnaris disease resistance. BMC Genomics 16:196

Glatz JF, Van Breda E, Van Der Vusse GJ (1998) Intracellular transport of fatty acids in muscle. Role of cytoplasmic fatty acid-binding protein. Adv Exp Med Biol 441:207–218

Gonzalez-Pena D, Gao GT, Baranski M, Moen T, Cleveland BM, Kenney PB, Vallejo RL, Palti Y, Leeds TD (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

Gutierrez AP, Yanez JM, Fukui S, Swift B, Davidson WS (2015) Genome-wide association study (GWAS) for growth rate and age at sexual maturation in Atlantic salmon (Salmo salar). Plos One 10: e0119730

Hodel RGJ, Chen SC, Payton AC, Mcdaniel SF, Soltis P, Soltis DE (2017) Adding loci improves phylogeographic resolution in red mangroves despite increased missing data: comparing microsatellites and RAD-Seq and investigating loci filtering. Sci Rep 7:17598

Hogan BL (1996) Bone morphogenetic proteins in development. Curr Opin Genet Dev 6:432–438

Hu XS, Li CT, Shi LY (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

Imumorin IG, Kim EH, Lee YM, De Koning DJ, Van Arendonk JA, De Donato M, Taylor JF, Kim JJ (2011) Genome scan for parent-of-origin QTL effects on bovine growth and carcass traits. Front Genet 2:44

Kimura S (2002) Structure and origin of fish type I collagen. Nippon Suisan Gakk 68:637–645

Kong S, Ke Q, Chen L, Zhou Z, Pu F, Zhao J, Bai H, Peng W, Xu P (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

Kuo T, Lew MJ, Mayba O, Harris CA, Speed TP, Wang JC (2012) Genome-wide analysis of glucocorticoid receptor-binding sites in myotubes identifies gene networks modulating insulin signaling. Proc Natl Acad Sci U S A 109:11160–11165

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760

Li YH, Wang HP (2017) Advances of genotyping-by-sequencing in fisheries and aquaculture. Rev Fish Biol Fish 27:535–559

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Proc GPD (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079

Li N, Zhou T, Geng X, Jin YL, Wang XZ, Liu SK, Xu XY, Gao DY, Li Q, Liu ZJ (2018) Identification of novel genes significantly affecting growth in catfish through GWAS analysis. Mol Gen Genomics 293: 587–599

Lin G, Chua E, Orban L, Yue GH (2016) Mapping QTL for sex and growth traits in salt-tolerant tilapia (Oreochromis spp. X O. mossambicus). Plos One 11:e0166723

Liu SX, Vallejo RL, Palti Y, Gao GT, Marancik DP, Hernandez AG, Wiens GD (2015) Identification of single nucleotide polymorphism markers associated with bacterial cold water disease resistance and spleen size in rainbow trout. Front Genet 6:298

Mai KS, Zhang CX, Ai QH, Duan QY, Xu W, Zhang L, Liufu ZG, Tan BP (2006) Dietary phosphorus requirement of large yellow croaker, Pseudosciaena crocea R. Aquaculture 251:346–353

Mcpherron AC, Lee SJ (1997) Double muscling in cattle due to mutations in the myostatin gene. Proc Natl Acad Sci USA 94:12457–12461

Mehinto AC, Martyniuk CJ, Spade DJ, Denslow ND (2012) Applications for next-generation sequencing in fish ecotoxicogenomics. Front Genet 3:62

Mueller KM, Themanns M, Friedbichler K, Kornfeld JW, Esterbauer H, Tuckermann JP, Moriggl R (2012) Hepatic growth hormone and glucocorticoid receptor signaling in body growth, steatosis and metabolic liver cancer development. Mol Cell Endocrinol 361:1–11

Muenke M, Schell U (1995) Fibroblast-growth-factor receptor mutations in human skeletal disorders. Trends Genet 11:308–313

Nagayama T, Okuhara S, Ota MS, Tachikawa N, Kasugai S, Iseki S (2013) FGF18 accelerates osteoblast differentiation by upregulating Bmp2 expression. Congenit Anom 53:83–88

Peng WZ, Xu J, Zhang Y, Feng JX, Dong CJ, Jiang LK, Feng JY, Chen BH, Gong YW, Chen L, Xu P (2016) An ultra-high density linkage map and QTL mapping for sex and growth-related traits of common carp (Cyprinus carpio). Sci Rep 6:26693

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One 7: e37135

Pratley RE, Baier L, Pan DA, Salbe AD, Storlien L, Ravussin E, Bogardus C (2000) Effects of an Ala54Thr polymorphism in the intestinal fatty acid-binding protein on responses to dietary fat in humans. J Lipid Res 41:2002–2008

Puig-Oliveras A, Revilla M, Castello A, Fernandez AI, Folch JM, Ballester M (2016) Expression-based GWAS identifies variants, gene interactions and key regulators affecting intramuscular fatty acid content and composition in porcine meat. Sci Rep 6:31803

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MaR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575

Ramos-Lopez E, Bruck P, Jansen T, Herwig J, Badenhoop K (2007) CYP2R1 (vitamin D 25-hydroxylase) gene is associated with susceptibility to type 1 diabetes and vitamin D levels in Germans. Diabetes Metab Res 23:631–636

Renaville R, Hammadi M, Portetelle D (2002) Role of the somatotropic axis in the mammalian metabolism. Domest Anim Endocrinol 23: 351–360

Robledo D, Fernandez C, Hermida M, Sciara A, Alvarez-Dios JA, Cabaleiro S, Caamano R, Martinez P, Bouza C (2016) Integrative transcriptome, genome and quantitative trait loci resources identify single nucleotide polymorphisms in candidate genes for growth traits in turbot. Int J Mol Sci 17:243

Robledo D, Palaiokostas C, Bargelloni L, Martinez P, Houston R (2018) Applications of genotyping by sequencing in aquaculture breeding and genetics. Rev Aquac 10:670–682

Rosen JN, Sogah VM, Ye LY, Mably JD (2013) ccm2-like is required for cardiovascular development as a novel component of the Heg-CCM pathway. Dev Biol 376:74–85

Sanchez-Ramos I, Cross I, Macha J, Martinez-Rodriguez 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. ScientificWorldJournal 2012:369802

Sternberg H, Moav B (1999) Regulation of the growth hormone gene by fish thyroid retinoid receptors. Fish Physiol Biochem 20:331–339

Su N, Jin M, Chen L (2014) Role of FGF/FGFR signaling in skeletal development and homeostasis: learning from mouse models. Bone Res 2:14003

Templeton GF (2011) A two-step approach for transforming continuous variables to normal: implications and recommendations for IS research. Commun Assoc Inf Syst 28:41–58

Tronche F, Opherk C, Moriggl R, Kellendonk C, Reimann A, Schwake L, Reichardt HM, Stangl K, Gau D, Hoeflich A, Beug H, Schmid W, Schutz G (2004) Glucocorticoid receptor function in hepatocytes is essential to promote postnatal body growth. Genes Dev 18:492–497

Tsai HY, Hamilton A, Guy DR, Houston RD (2014) Single nucleotide polymorphisms in the insulin-like growth factor 1 (IGF1) gene are associated with growth-related traits in farmed Atlantic salmon. Anim Genet 45:709–715

Vallejo RL, Leeds TD, Fragomeni BO, Gao G, Hernandez AG, Misztal I, Welch TJ, Wiens GD, Palti Y (2016) Evaluation of genome-enabled selection for bacterial cold water disease resistance using progeny performance data in rainbow trout: insights on genotyping methods and genomic prediction models. Front Genet 7:96

Wan L, Dong L, Xiao S, Han Z, Wang X, Wang Z (2018) Genomewide association study for economic traits in the large yellow croaker with different numbers of extreme phenotypes. J Genet 97:887–895

Wang L, Shi XF, Su YQ, Meng ZN, Lin HR (2012) Loss of genetic diversity in the cultured stocks of the large yellow croaker, Larimichthys crocea, revealed by microsatellites. Int J Mol Sci 13: 5584–5597

Wang L, Wan ZY, Bai B, Huang SQ, Chua E, Lee M, Pang HY, Wen YF, Liu P, Liu F, Sun F, Lin G, Ye BQ, Yue GH (2015) Construction of a high-density linkage map and fine mapping of QTL for growth in Asian seabass. Sci Rep 5:16358

Wang XH, Fu BD, Yu XM, Qu CY, Zhang Q, Tong JO (2018a) Fine mapping of growth-related quantitative trait loci in Yellow River carp (Cyprinus carpio haematoperus). Aquaculture 484:277–285

Wang Y, Guo F, Qu H, Luo C, Wang J, Shu D (2018b) Associations between variants of bone morphogenetic protein 7 gene and growth traits in chickens. Br Poultry Sci 59:264–269

Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, Beckmann JS, Bragg-Gresham JL, Chang HY, Demirkan A, Den Hertog HM, Do R, Donnelly LA, Ehret GB, Esko T, Feitosa MF, Ferreira T, Fischer K, Fontanillas P, Fraser RM, Freitag DF, Gurdasani D, Heikkila K, Hypponen E, Isaacs A, Jackson AU, Johansson A, Johnson T, Kaakinen M, Kettunen J, Kleber ME, Li X, Luan J, Lyytikainen LP, Magnusson PKE, Mangino M, Mihailov E, Montasser ME, Muller-Nurasyid M, Nolte IM, O'connell JR, Palmer CD, Perola M, Petersen AK, Sanna S, Saxena R, Service SK, Shah S, Shungin D, Sidore C, Song C, Strawbridge RJ, Surakka I, Tanaka T, Teslovich TM, Thorleifsson G, Van Den Herik EG, Voight BF, Volcik KA, Waite LL, Wong A, Wu Y, Zhang W, Absher D, Asiki G, Barroso I, Been LF, Bolton JL, Bonnycastle LL, Brambilla P, Burnett MS, Cesana G, Dimitriou M, Doney ASF, Doring A, Elliott P, Epstein SE, Ingi Eyjolfsson G, Gigante B, Goodarzi MO, Grallert H, Gravito ML, Groves CJ, Hallmans G, Hartikainen AL, Hayward C, Hernandez D, Hicks AA, Holm H, Hung YJ, Illig T, Jones MR, Kaleebu P, Kastelein JJP, Khaw KT, Kim E et al (2013)

Discovery and refinement of loci associated with lipid levels. Nat Genet 45:1274–1283

Wong AOL, Zhou H, Jiang YH, Ko WKW (2006) Feedback regulation of growth hormone synthesis and secretion in fish and the emerging concept of intrapituitary feedback loop. Comp Biochem Physiol A Mol Integr Physiol 144:284–305

Wu CW, Zhang D, Kan MY, Lv ZM, Zhu AY, Su YQ, Zhou DZ, Zhang JS, Zhang Z, Xu MY, Jiang LH, Guo BY, Wang T, Chi CF, Mao Y, Zhou JJ, Yu XX, Wang HL, Weng XL, Jin JG, Ye JY, He L, Liu Y (2014) The draft genome of the large yellow croaker reveals well-developed innate immunity. Nat Commun 5:5227

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 16:729–738

Yue GH (2014) Recent advances of genome mapping and marker-assisted selection in aquaculture. Fish Fish 15:376–396

Zhao YF, Peng WZ, Guo HY, Chen BH, Zhou ZX, Xu J, Zhang DC, Xu P (2018) Population genomics reveals genetic divergence and adaptive differentiation of Chinese sea bass (Lateolabrax maculatus). Mar Biotechnol 20:45–59

Zheng XJ, Xu C, Smith AO, Stratman AN, Zou ZY, Kleaveland B, Yuan LJ, Didiku C, Sen A, Liu X, Skuli N, Zaslavsky A, Chen M, Cheng L, Davis GE, Kahn ML (2012) Dynamic regulation of the cerebral cavernous malformation pathway controls vascular stability and growth. Dev Cell 23:342–355

Zhong X, Wang X, Zhou T, Jin Y, Tan S, Jiang C, Geng X, Li N, Shi H, Zeng Q, Yang Y, Yuan Z, Bao L, Liu S, Tian C, Peatman E, Li Q, Liu Z (2017) Genome-wide association study reveals multiple novel QTL associated with low oxygen tolerance in hybrid catfish. Mar Biotechnol 19:379–390

Zhou ZX, Chen L, Dong CJ, Peng WZ, Kong SN, Sun JS, Pu F, Chen BH, Feng JX, Xu P (2018) Genome scale association study of abnormal scale pattern in Yellow River carp identified previously known causative gene in European mirror carp. Mar Biotechnol 20:573–583

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.