Genome Scan for Genomic Regions and Genes Associated with Growth Trait in Pacific White Shrimp *Litopeneaus vannamei*

Yang Yu [1,2] · Quanchao Wang [1] · Qian Zhang [1,3] · Zheng Luo [1,3] · Yue Wang [1,3] · Xiaojun Zhang [1,2] · Hao Huang [4] ·
Jianhai Xiang [1,2,5] · Fuhua Li [1,2,5]

Abstract

The Pacific white shrimp *Litopeneaus vannmei* (*L. vannmei*) is a predominant aquaculture shrimp species worldwide, and it is considered as the aquaculture species with the highest single output value. Advances in selective breeding have accelerated the development of *L. vannmei* aquaculture. Recently, the genome-wide association studies (GWAS) have been applied in aquaculture animals and markers associated with economic traits were identified. In this study, we focused on the growth trait of *L. vannamei* and performed GWAS to identify SNPs or genes associated with growth. Genomic regions in linkage group 7, 27, 33, and 38 were identified to be associated with body weight and body length of the shrimp. Further, candidate gene association analysis was performed in two independent populations and the result demonstrated that the SNPs in the genes *protein kinase C delta type* and *ras-related protein Rap-2a* were significantly associated with the growth trait of *L. vannamei*. This study showed that GWAS analysis is an efficient approach for screening trait-related markers or genes. The genomic regions and genes identified in this study are essential for further fine mapping of growth-related genes. The identified markers will provide useful information for marker-assisted selection in *L. vannamei*.

Keywords Genome-wide association studies · Growth trait · Candidate genes · *Litopenaeus vannamei*

Introduction

The Pacific white shrimp *Litopeneaus vannmei* (*L. vannmei*) is a worldwide cultured species with the total production of 4 million tons, and the total value of the output was 24 billion US dollars in 2016 (http://www.fao.org/fishery/statistics/global-aquaculture-production/query/en). It is

considered as an aquaculture species with the highest single output value. Advances in selective breeding have accelerate the development of *L. vannmei* aquaculture. For selective breeding, the growth trait is among the most important economic traits in shrimp aquaculture. Several breeding programs have been performed to select rapid growth broodstocks in *L. vannamei* since 1990s (Andriantahina et al. 2012; Argue et al. 2002; Castillo-Juarez et al. 2015; Moss et al. 2008; T et al. 2005). The heritability of the growth trait was reported to be 0.24–0.515 (Andriantahina et al. 2012; Castillo-Juarez et al. 2007; Li et al. 2015), which illustrated that the genetic factor played important role during the growing process. In order to improve the selective efficiency, marker-assisted selection (MAS) is a useful approach to accelerate the genetic progress. Besides, understanding the genetic architecture and identifying the QTLs for growth trait could facilitate the application of MAS in shrimp (Liu and Cordes 2004).

In the past decade, genetic mapping has been conducted to identify the growth-related QTLs or markers in *L. vannamei*. Compared with the other aquaculture species, such as Atlantic salmon (Gutierrez et al. 2012, 2015), tilapia (Liu et al. 2014)

Yang Yu and Quanchao Wang contributed equally to this manuscript.

✉ Fuhua Li
fhli@qdio.ac.cn

[1] Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China

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

[3] University of Chinese Academy of Sciences, Beijing 100049, China

[4] Hainan Grand Suntop Ocean Breeding Co., Ltd, Wenchang 571300, China

[5] Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao 266071, China

and catfish (Geng et al. 2016; Li et al. 2018), very limited information about the growth-associated markers or genes is known for the shrimp *L. vannamei*. Previous studies have identified some QTLs related to growth trait through linkage mapping method (Andriantahina et al. 2013; Yu et al. 2015). However, no genes related to growth trait were identified due to the low resolution of the linkage map or limited information of the whole genomic sequence of *L. vannamei*.

Association mapping is regarded as an efficient approach for dissecting the complex trait (Mackay 2001; Tan et al. 2018). The association mapping highly relies on the linkage disequilibrium (LD) between the markers and the genetic variations affecting the trait (Wen et al. 2014). Among the association mapping approaches, the genome-wide association (GWA) mapping is the most widely used method in plant and livestock (Hayes and Goddard 2010; Hirschhorn and Daly 2005; Huang et al. 2010). In the past few years, genome-wide association study (GWAS) analyses have also been conducted to identify growth-related QTLs in several aquaculture species (Li et al. 2018; Tsai et al. 2015; Gonzalez-Pena et al. 2016). For instance, markers significantly associated with growth were identified in Atlantic salmon (Gutierrez et al. 2015). In common carp, growth-related genes such as *KISS2*, *IGF1*, *SMTLB*, *NPFFR1*, and *CPE* were identified (Peng et al. 2016). In catfish, one single significant QTL located on the linkage group 5 was identified to be associated with growth traits (Li et al. 2018). Furthermore, nine genes such as *actn2*, *Adam12*, and *dock1* located in the QTL regions were also identified as candidate genes, which might function in the bone development of this species (Geng et al. 2017). Up to now, there is no application of GWAS in studying the growth traits of the penaeid shrimp *L. vannamei*. It is partially attribute to the lack of reference genome information and the high-throughput genotyping platform in shrimp. Recent advances in genomic study have facilitated the genotyping of genome wide SNP with a low cost in non-model species (Davey et al. 2011; Yu et al. 2015). So, the use of GWAS was therefore possible to improve the mapping efficiency and identify interested genes in *L. vannamei*.

In this study, we presented the results of GWAS for growth trait in *L. vannamei*. The body weight and body length were separately analyzed to identify growth-associated genomic regions. Among the QTL regions, two candidate growth-related genes were identified and analyzed.

Materials and Methods

Samples and Phenotype Data

The samples used for GWAS analysis were derived from a full-sib family with a total of 205 progenies. It is the same as

the mapping family described previously (Yu et al. 2015). Briefly, the full-sib family was created in the breeding center of Guangxi Institute of Fisheries (Nanning, Guangxi Province, China) and all the shrimp were cultured under the same condition. The growth-related phenotypes including body weight and body length were measured for each individual after 3-months cultivation. In order to validate the identified growth-related markers, another two independent populations which was derived from a commercial breeding program developed by Hainan Grand Suntop Marine breeding company (Wenchang, Hainan province, China) were used. The shrimp family was constructed in 2015 and the individuals from the same family were cultured in one tank for seeding and culturing. After the offspring reached around 3 cm in length, a total of 13 families were labeled with visible implant elastomer (VIE) tags and pooled together as the first validation population, in which each family contained 50 offspring. The same strategy was applied to construct another family pool, in which the families were different from the previous one. After almost 3 month's cultivation, the phenotype of body weight and body length was measured for the two populations respectively, and the sex of each individual was recorded. A total of 143 individuals and 200 individuals were randomly collected and measured for the two populations separately.

DNA Extraction, SNP Genotyping, and GWAS Analysis

Genomic DNA of the mapping family and the two validation populations were extracted using TIANGEN Plant DNA extraction kit (TIANGEN, Beijing, China). The concentration of extracted DNA was determined using a NanoDrop 1000 Spectrophotometer (NanoDrop, Wilmington, DE, USA). DNA integrity was evaluated by gel electrophoresis. The whole genome SNP data used for genome-wide association analysis of the family have been described previously (Yu et al. 2015). Briefly, the individuals of the family were genotyped by SLAF-seq approach (Sun et al. 2013). As the GWAS only use the bi-allele marker, the tri-allele and four-allele markers were filtered. The markers were further filtered with minor allele frequency (> 0.05) and call rate for SNPs (> 90%). The location of each SNPs along the chromosome was obtained from the linkage map published previously (Yu et al. 2015).

In order to determine the SNPs associated with body weight and body length, GWAS was performed using the GenABEL package in the R environment (Aulchenko et al. 2007b). A genomic kinship matrix was calculated from the SNP data using the *gkin* function. Heritability was estimated using a linear mixed model implemented on polygenic function. Genome-wide rapid association using the mixed model and regression-Gamma (GRAMMAR-Gamma) method was used to identify the associated SNPs (Aulchenko et al. 2007a).

The -log10 (*P* value) of each SNP across the genome was calculated to illustrate the GWAS results. The threshold *P value* for genome-wide significance was calculated using Bonferroni correction based on the estimated number of independent markers. Considering the marker number and the number of genotyping populations, the suggestive threshold for the association significance was arbitrarily set as 0.01.

Annotation of Significant Markers

The flanking sequence of the significant SNPs was blasted to the assembled *L. vannamei* genome (Zhang et al. 2019) using BLASTN. The upstream 40 kb and downstream 40 kb sequence around the top BLAST hit region were examined for candidate genes according to their locations and functions. The genes within these regions were predicted using FGENESH (Solovyev et al. 2006) and annotated by BLAST analysis against the non-redundant protein database in National Center for Biotechnology Information (NCBI) database.

Candidate Gene Association Analysis

Based on the annotation of the significant SNP markers, three genes including *E3 ubiquitin-protein ligase HECW1*, *protein kinase c*, and *ras-related protein Rap-2a* near the Marker33701, Marker1286, and Marker1121 were selected for candidate gene association analysis. Primers were designed to amplify the gene regions in the genomic DNAs. A pool of DNAs from the two validation populations were used as template for PCR amplification. The PCR products were sequenced using Sanger sequencing by ABI Prism 3730xl and SNPs in the gene regions of the three genes were detected. Those regions containing SNPs were further amplified and sequenced in the individual DNA sample to obtain the genotype information for each shrimp. As the body weight and body length shown a significant correlation, only the body weight were analyzed and the association test were performed using R/SNPassoc package.

Expression Analysis of the Candidate Genes

In order to analyze the function of the candidate growth-related genes, the expression patterns of these genes were analyzed. Healthy shrimps with an average weight of 11 g

were randomly selected. Hemocytes were collected from the shrimp with anticoagulant and isolated by centrifugation at 800×*g* for 10 min under 4 °C. Different tissues including epidermis, gill, hemocyte, hepatopancreas, heart, intestine, muscle, nerve, lymphoid organ, and stomach were collected and preserved in liquid nitrogen for total RNA extraction. Total RNA was extracted from various tissues using RNAiso Plus reagent (TaKaRa, Japan). A total of 1 μg RNA for each was used for cDNA synthesis in a total volume of 25 μl using PrimeScript RT Reagent Kit (TaKaRa, Japan), the temperature was set at 37 °C for 15 min and then 85 °C for 5 s.

Quantitative real-time PCR (qPCR) was performed to analyze the expression pattern of candidate genes. The housekeeping gene 18S rRNA was employed as a reference gene. The reaction contained 5 μl SuperReal PreMix Plus (SYBR Green) (Tiangen, China), 0.3 mM primers, 1 μl cDNA template, and then adding RNase free water to 10 μl. The qPCR program was set as the following: 95 °C for 20 min; 40 cycles of 95 °C for 15 s, annealing temperature for 15 s and 72 °C for 20 s. In order to confirm whether the PCR product was specific, a melt cycle was added. The data was analyzed by $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen 2001).

Result

Phenotypic Data of the Growth Trait

The phenotypic data of the two validation populations shown a relative high level of variation compared with the mapping family (Table 1). The average body weight and body length of the mapping family was 20.32 ± 3.72 g and 117.13 ± 7.36 mm, respectively. The average body weight and body length of the validation population 1 was 8.94 ± 2.80 g and 90.35 ± 9.69 mm, and those for the validation population 2 were 5.56 ± 2.16 g and 76.99 ± 9.95 mm separately.

GWAS Analysis of the Body Length in the Mapping Family

A total of 3960 SNPs were obtained after filtration and quality control. These SNPs distributed along 44 linkage groups with one SNP per 1.04 cM in average. The heritability estimated by the SNPs information for body length was 0.25. After Bonferroni correction, the genome-wide significance threshold

Table 1 The phenotype information of three populations for growth trait

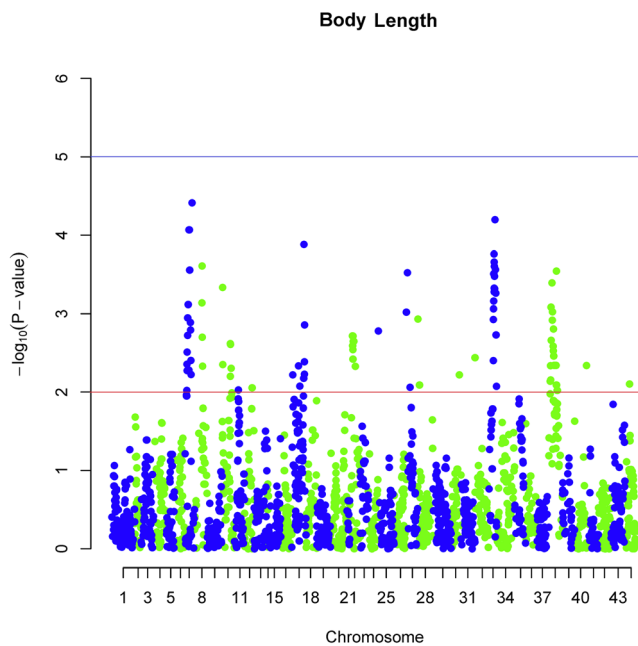|	Mapping family		Validation population 1		Validation population 2	
	Mean ± SD	CV(%)	Mean ± SD	CV(%)	Mean ± SD	CV(%)
Body weight (g)	20.32 ± 3.72	18.29	8.94 ± 2.80	31.22	5.56 ± 2.16	38.76
Body length (mm)	117.13 ± 7.36	6.29	90.35 ± 9.69	10.74	76.99 ± 9.95	12.92

Body Length

Fig. 1 Genome-wide association analysis result for body length of *L. vannamei*. The red line indicated the suggestive threshold for significant markers associated with body length. The blue line indicated the genome-wide threshold for significant markers

was 1.6×10^{-5}. Consequently, GWAS analysis identified a total of 52 suggestively associated markers ($P < 0.01$) (Fig. 1). Among these markers, the most significant marker was Marker19909 located in 52.69 cM along LG 7, and the next are Marker4108 and Marker1309 located in 38.62 cM of LG 33 (Table 2). Interestingly, among the top 10 significant markers, eight markers located in LG7, and the other two markers located in LG33. These data gave us a clue that the linkage groupLG7 and LG33 might contain QTLs for body length.

GWAS Analysis on Body Weight in the Mapping Family

Similar to the body length trait, only suggestively associated markers were identified. A total of 51 suggestively associated markers were screened ($P < 0.01$) (Fig. 2). The most significant markers were Marker4108 and Marker1309 located in 38.62 cM of LG33 (Table 3). The heritability estimated by SNPs information for body weight was 0.42. Similar to the body length, the LG33 contained eight markers among the top 20 significant markers, which illustrate the importance of these regions

Table 2 The markers associated with body length of *L. vannamei* at $P < 0.01$

| Marker	LG	Location (cM)	P value	Marker	LG	Location (cM)	P value
Marker19909	7	52.69	7.63E-04	Marker4670	38	38.59	3.82E-03
Marker4108	33	38.62	1.07E-03	Marker53616	10	2.32	4.19E-03
Marker1309	33	38.62	1.07E-03	Marker60667	33	32.89	4.25E-03
Marker4295	7	34.12	1.32E-03	Marker13920	33	35.82	4.30E-03
Marker37724	7	34.12	1.32E-03	Marker21557	33	32.33	4.57E-03
Marker20665	7	34.12	1.32E-03	Marker1077	33	44.68	4.71E-03
Marker3072	7	34.12	1.32E-03	Marker45687	33	28.52	5.51E-03
Marker2719	7	34.12	1.32E-03	Marker7900	8	87.95	5.72E-03
Marker16098	7	34.12	1.32E-03	Marker91840	7	30.81	5.91E-03
Marker9183	7	34.12	1.32E-03	Marker3095	7	31.00	5.91E-03
Marker15416	17	76.64	1.77E-03	Marker15621	38	25.43	6.24E-03
Marker8720	33	31.98	2.13E-03	Marker4976	33	27.46	6.44E-03
Marker1079	33	32.49	2.51E-03	Marker30069	38	42.93	6.84E-03
Marker5211	8	88.34	2.71E-03	Marker33701	27	4.36	6.88E-03
Marker31892	33	33.57	2.73E-03	Marker7660	27	4.40	6.88E-03
Marker6298	33	33.57	2.73E-03	Marker41944	7	27.34	7.75E-03
Marker1121	33	40.13	2.92E-03	Marker5243	7	28.50	7.75E-03
Marker15848	7	35.15	2.95E-03	Marker10112	7	28.50	7.75E-03
Marker40314	7	35.15	2.95E-03	Marker15906	28	1.89	7.93E-03
Marker7680	7	35.15	2.95E-03	Marker2715	28	1.89	7.93E-03
Marker15779	38	56.60	3.01E-03	Marker13166	33	27.07	8.00E-03
Marker19992	27	16.29	3.11E-03	Marker13684	38	42.38	8.11E-03
Marker16804	33	28.57	3.18E-03	Marker1208	7	36.72	8.48E-03
Marker12804	33	28.74	3.18E-03	Marker50862	17	84.89	8.94E-03
Marker16675	33	34.08	3.34E-03	Marker508	38	52.91	9.71E-03
Marker3267	33	34.08	3.34E-03	Marker25487	7	38.23	9.90E-03
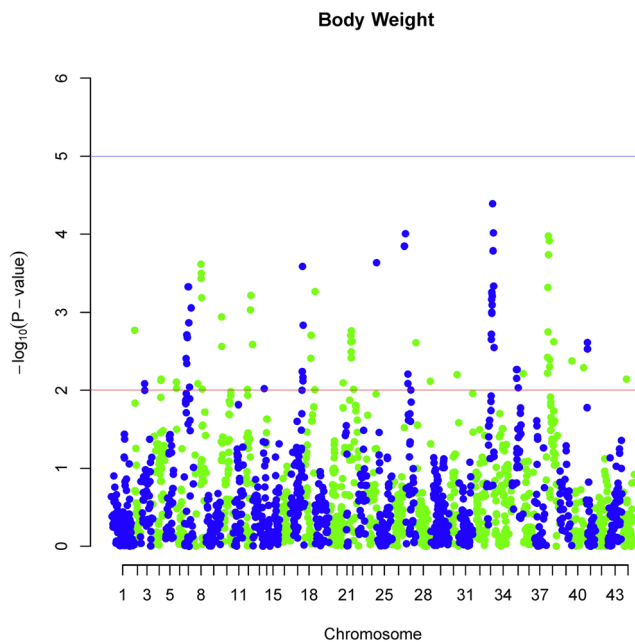
Body Weight

Fig. 2 Genome-wide association analysis for the body weight of *L. vannamei*. The red line indicated the suggestive threshold for significant markers associated with body weight. The blue line indicated the genome-wide threshold for significant markers

for growth trait. The LG38, LG27, and LG8 were the other three linkage groups containing significant markers.

Annotation on the Associated Genomic Regions

Based on the genome-wide association results, we analyzed the genomic regions associated with body length and body weight. The top 10 significant markers were selected for each trait, separately. As there were two common markers for the two traits, a total of 18 growth-related markers were annotated. Among them, 11 markers were next to genes and seven were annotated as known genes (Table 4). For the body length, the genomic regions along the top 10 markers contained four genes, of which one gene was annotated as Beta-2 adrenergic receptor and the function of the three other genes were unknown. For the body weight, the associated genomic regions contained eight genes, in which six genes were annotated by non-redundant protein database. Some growth trait-related genes such as *protein kinase C delta (PKC-delta) ras-related protein Rap-2a (Rap-2a)* and *E3 ubiquitin-protein ligase HECW2* happen to be located in these regions.

Table 3 The markers associated with body weight of *L. vannamei* at *P* < 0.01

Marker	LG	Location (cM)	*P* value	Marker	LG	Location (cM)	*P* value
Marker4108	33	38.62	7.77E-04	Marker16098	7	34.12	4.18E-03
Marker1309	33	38.62	7.77E-04	Marker9183	7	34.12	4.18E-03
Marker1077	33	44.68	1.41E-03	Marker32584	38	19.53	4.25E-03
Marker19992	27	16.29	1.43E-03	Marker3452	18	87.45	4.61E-03
Marker15621	38	25.43	1.49E-03	Marker1079	33	32.49	4.67E-03
Marker4670	38	38.59	1.66E-03	Marker31892	33	33.57	4.81E-03
Marker33701	27	4.36	1.84E-03	Marker6298	33	33.57	4.81E-03
Marker7660	27	4.40	1.84E-03	Marker17576	12	61.69	5.00E-03
Marker1121	33	40.13	2.02E-03	Marker1340	12	62.28	5.00E-03
Marker1286	38	27.85	2.18E-03	Marker13920	33	35.82	5.07E-03
Marker24697	25	0.57	2.56E-03	Marker16888	8	94.25	5.23E-03
Marker7900	8	87.95	2.64E-03	Marker8720	33	31.98	5.43E-03
Marker15416	17	76.64	2.76E-03	Marker16675	33	34.08	6.05E-03
Marker11961	8	91.26	3.17E-03	Marker3267	33	34.08	6.05E-03
Marker5211	8	88.34	3.53E-03	Marker19909	7	52.69	6.43E-03
Marker529	33	45.69	4.12E-03	Marker19981	12	56.33	6.68E-03
Marker17240	33	45.69	4.12E-03	Marker60667	33	32.89	7.02E-03
Marker26923	33	45.69	4.12E-03	Marker21557	33	32.33	7.18E-03
Marker21487	33	45.69	4.12E-03	Marker53616	10	2.32	7.73E-03
Marker4295	7	34.12	4.18E-03	Marker15848	7	35.15	8.72E-03
Marker37724	7	34.12	4.18E-03	Marker40314	7	35.15	8.72E-03
Marker20665	7	34.12	4.18E-03	Marker7680	7	35.15	8.72E-03
Marker3072	7	34.12	4.18E-03	Marker50862	17	84.89	9.14E-03
Marker2719	7	34.12	4.18E-03				

Table 4 Annotation of top associated markers with body length or body weight in *L. vannamei*

Marker	LG	Site	Flanking gene annotation
Marker19909	7	52.69	Unknown[a]
Marker4108	33	38.62	Unknown
Marker1309	33	38.62	NA[b]
Marker4295	7	34.12	Unknown
Marker37724	7	34.12	NA
Marker20665	7	34.12	NA
Marker3072	7	34.12	NA
Marker2719	7	34.12	NA
Marker16098	7	34.12	NA
Marker9183	7	34.12	Beta-2 adrenergic receptor
Marker1077	33	44.68	NA
Marker19992	27	16.29	Unknown
Marker15621	38	25.43	Hypothetical protein X975_01911
Marker4670	38	38.59	Protein FAM186A isoform X1
Marker33701	27	4.36	E3 ubiquitin-protein ligase HECW2
Marker7660	27	4.40	Glutaminase kidney isoform, mitochondrial
Marker1121	33	40.13	Ras-related protein Rap-2a
Marker1286	38	27.85	Protein kinase C delta type

Unknown[a] means that the function of annotated gene is unknown. NA[b] means that there was no predicted gene in this region

Association Analyses of Candidate Genes

Based on the *P* value and the annotated gene function of the identified markers, we selected three genes including *PKC-delta*, *Rap-2a*, and *E3 ubiquitin-protein ligase HECW2* for further validation and candidate genes association analyses. From the sequencing of the genomic DNA of the three genes, a total of five, six, and three SNPs were identified in the amplified regions of *PKC-delta*, *Rap-2a*, and *E3 ubiquitin-protein ligase HECW2*, respectively. After genotyping these SNPs in the two validation populations, one SNP located in the 3′UTR region of *PKC-delta* was significantly associated

with the body weight under both the dominant and codominant model ($P < 0.05$). One SNP in the intron of *Rap-2a* was also validated to be significantly associated with body weight (Table 5). Interestingly, only two genotypes were observed for the marker in *Rap-2a*, one type of homozygote was lost in both populations.

Gene Expression Pattern of Candidate Genes

The expression pattern of the *PKC-delta* and *Rap-2a* in different tissues was analyzed by qRT-PCR using the primers listed in Table 6. The *PKC-delta* was highly expressed in muscle, heart, and stomach. The expression was low in the other tissues (Fig. 3a). The *Rap-2a* was highly expressed in lymphoid organ, besides, the expression level in the digestive organs including hepatopancreas, intestine, and stomach was also high (Fig. 3b).

Discussion

Growth traits including body weight and body length are the most concerned economic traits in the selective breeding of *L. vannamei* (Benzie 2009; Caballero-Zamora et al. 2015). In this study, we identified several genomic regions associated with body weight and body length by genome-wide association analysis. Based on the assembled genome sequence, candidate growth-related genes were identified.

Genome-wide association study has become the major approach for mapping genes associated with the interested traits. It has been widely used to identify the causative genes of disease in humans and explore the SNPs related to economical traits in crops and livestock (Bouwman et al. 2018; Hirschhorn and Daly 2005; Huang and Han 2014; Zhang et al. 2012). In aquaculture species, several GWAS have been performed in catfish, tilapia, common carp, Atlantic salmon, rainbow trout, and Pacific oyster (Correa et al. 2015; Geng et al. 2016; Gutierrez et al. 2015; Li et al. 2017, 2018;

Table 5 Association analysis result of SNPs in the genes *PKC* and *Rap-2a*

Genes	SNP information	Validation population 1				Validation population 2			
		Genotype	Body weight (g)	*P* value in codominant model	*P* value in dominant model	Genotype	Body weight (g)	*P* value in codominant model	*P* value in dominant model
PKC-delta	CGCTGACTGGGAAAATACTC TTACC[A/G]CTTTTTTT CCCCCCGACGTCGATGA	A/A A/G G/G	8.22 ± 2.65 10.10 ± 3.28 9.31 ± 2.52	0.053	0.018*	A/A A/G G/G	5.31 ± 2.16 5.63 ± 2.18 6.15 ± 1.98	0.029*	0.011*
Rap-2a	TCGTTGGCACCCGCGGCGGC GCTTGCG[G/A]GGGCGGCC GCCAGGGGGCCAGTGGTGC	G/A A/A	9.61 ± 2.67 8.61 ± 2.86	0.039*	NA	G/A A/A	5.86 ± 2.00 5.28 ± 2.25	0.018*	NA

*P value < 0.05

Primer ID		Primer sequence (5′–3′)	Fragment length (bp)	Ta (°C)
18S	Forward	TATACGCTAGTGGAGCTGGAA	136	55
	Reverse	GGGGAGGTAGTGACGAAAAAT		
PKC	Forward	GTGCTGAGCCTCGGAACCA	132	56
	Reverse	GCCGCAGTGTTGTATGTGGA		
Rap-2a	Forward	GCCGTGCGTGCTTGAGAT	135	55
	Reverse	TTGATGTCCTGGAAGGTCTGG		

Ta, the optimal annealing temperature

Sodeland et al. 2013; Song et al. 2018; Wang et al. 2017; Zheng et al. 2016; Zhou et al. 2017). Genomic regions or candidate genes associated with disease resistance, fillet quality, sexual maturation, shell color, and growth rate were

Fig. 3 Relative expression pattern of *protein kinase C delta* (**a**) and *ras-related protein Rap-2a* (**b**). Epi, epidermis; Gi, gill; Hc, hemocyte; Hp, hepatopancreas; Ht, heart; In, intestine; Ms, muscle; Ner, nerve; Oka, lymphoid organ; St, stomach

identified. In our study, by combing GWAS and candidate gene association approach, several genomic regions along LG7, LG27, LG33, and LG38 associated with the growth traits were identified in *L. vannamei*. Two genes including *PKC-delta* and *Rap-2a* were identified as candidate growth-related genes.

Usually the Bonferroni correction is generally applied to avoid false positive (type I error) rates. However, it is known that the Bonferroni-corrected P value is too strict in a GWAS analysis (Hong and Park 2012; Spencer et al. 2009). In our study, the genome-wide significance threshold was around 1.6×10^{-5} after Bonferroni correction. Considering the small sample size and large number of SNPs, the corrected significant threshold is too strict to identify candidate markers. Similar results were reported in other aquaculture species (Gutierrez et al. 2015). The growth trait is a polygenic trait controlled by many genes with small effect (Gutierrez et al. 2015; Meng et al. 2017). The body weight and body length might be controlled by many genes with moderate to low effects in shrimp. So, the p value of the most detected markers was not significant for the adjusted threshold. However, the top associated markers were clustered in genomic regions along the LG7, LG33, and LG38, which illustrated effective genes controlling growth might be located in these regions.

In order to identify traits-related markers, the candidate markers should be validated in other populations (Yue 2014). With the help of genome reference, the annotations of the significant markers were obtained. By combing the GWAS result and the annotation information, it can give evidence for identifying the candidate genes. In this study, three candidate genes were selected for association analyses and two genes were confirmed. The *PKC-delta* is a metabolism-related gene and it has been proved to play an important role in apoptosis (Zhao et al. 2012) and it also presented a regulation function during the molting process (Chen et al. 2017). The growth of shrimp is discontinuous and accompanied by frequent molting, so the *PKC-delta* may influence the growth by regulating the molting (Shyamal et al. 2018). *Rap-2a* is a member of the Ras-related protein family. It is part of several signaling
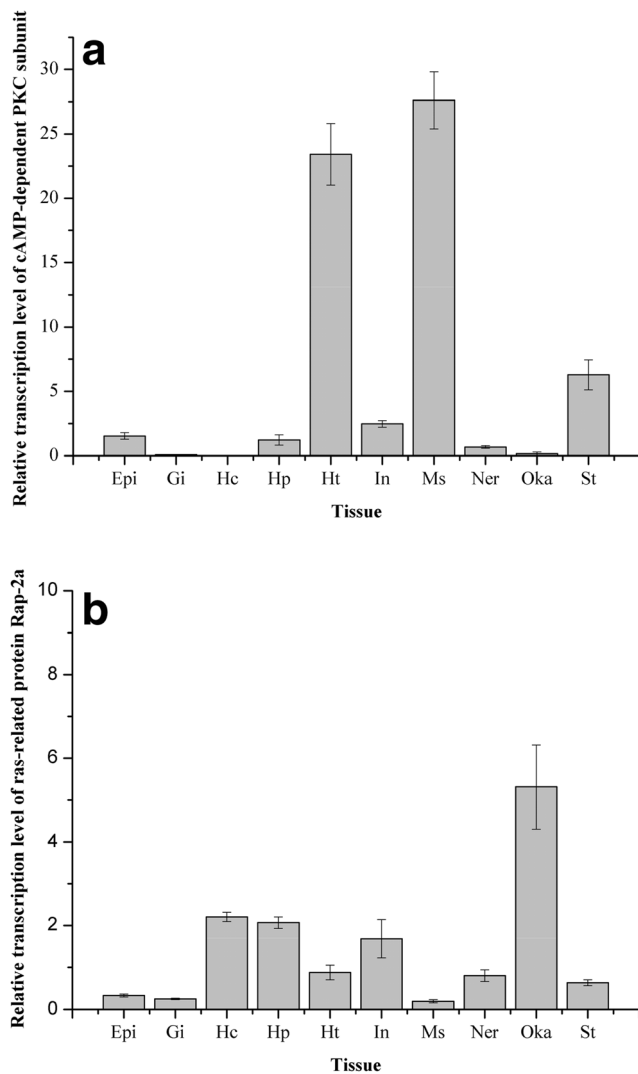
cascades and may regulate cytoskeletal rearrangements, cell migration, and cell spreading (Taira et al. 2004). Interestingly, the Ras gene was identified to be associated with head size in catfish, and it located in the QTL hub of the catfish genome which played an important role in growth regulation (Geng et al. 2016). The expression pattern of *PKC-delta* shown that it was high expressed in the muscle tissue, as the muscle was the major portion of shrimp body, it may play crucial rule in growth. The *Rap-2a* shown highly expression in the intestine tissue. The intestine was the major tissue for energy absorption and was related to development and growing. The expression analysis results gave us more evidence for the two genes as growth-related genes.

The marker-assisted selection (MAS) has been widely used in crop and animal breeding (Goddard and Hayes 2009). The advantage of MAS was to improve the selection accuracy and speed up the genetic gain through early selection (Yue 2014). The QTLs and SNPs associated with growth trait identified in this study will be useful for the further MAS application. Furthermore, the gene-assisted selection (GAS) was more accurate than MAS (Dekkers 2004). In this study, the two identified growth-related genes will also accelerate the GAS study in *L. vannamei*. However, as the restriction of mapping population and the marker density, the causative SNPs or mutations influencing the growth trait have not been detected. So, further fine mapping and function analysis of these genes are necessary.

Conclusion

In this study, SNPs and genomic regions associated with growth trait were identified by genome wide-association mapping approach. Based on the assembled genome reference, two growth-related genes including *PKC* and *Rap-2a* were identified. It illustrated the power of mapping QTLs or genes using GWA mapping in *L. vannamei*. The associated markers or genes will be useful to marker-assisted selection and the identified genes will expand our knowledge about the mechanisms of growth regulation in *L. vannamei*.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no competing interests.

References

Andriantahina F, Liu XL, Huang H (2013) Genetic map construction and quantitative trait locus (QTL) detection of growth-related traits in *litopenaeus vannamei* for selective breeding applications. PLoS ONE 8(9):e75206

Andriantahina F, Liu XL, Huang H, Xiang JH (2012) Response to selection, heritability and genetic correlations between body weight and body size in Pacific white shrimp, *Litopenaeus vannamei*. Chin J Oceanol Limnol 30:506–506

Argue BJ, Arce SM, Lotz JM, Moss SM (2002) Selective breeding of Pacific white shrimp (*Litopenaeus vannamei*) for growth and resistance to Taura syndrome virus. Aquaculture 204:447–460

Aulchenko YS, de Koning DJ, Haley C (2007a) Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. Genetics 177:577–585

Aulchenko YS, Ripke S, Isaacs A, Van Duijn CM (2007b) GenABEL: an R library for genome-wide association analysis. Bioinformatics 23:1294–1296

Benzie JAH (2009) Use and exchange of genetic resources of penaeid shrimps for food and aquaculture. Rev Aquac 1:232–250

Bouwman AC, Daetwyler HD, Chamberlain AJ, Ponce CH, Sargolzaei M, Schenkel FS, Sahana G, Govignon-Gion A, Boitard S, Dolezal M, Pausch H, Brøndum RF, Bowman PJ, Thomsen B, Guldbrandtsen B, Lund MS, Servin B, Garrick DJ, Reecy J, Vilkki J, Bagnato A, Wang M, Hoff JL, Schnabel RD, Taylor JF, Vinkhuyzen AAE, Panitz F, Bendixen C, Holm LE, Gredler B, Hozé C, Boussaha M, Sanchez MP, Rocha D, Capitan A, Tribout T, Barbat A, Croiseau P, Drögemüller C, Jagannathan V, Vander Jagt C, Crowley JJ, Bieber A, Purfield DC, Berry DP, Emmerling R, Götz KU, Frischknecht M, Russ I, Sölkner J, van Tassell CP, Fries R, Stothard P, Veerkamp RF, Boichard D, Goddard ME, Hayes BJ (2018) Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. Nat Genet 50:362–367

Caballero-Zamora A, Cienfuegos-Rivas EG, Montaldo HH, Campos-Montes GR, Martinez-Ortega A, Castillo-Juarez H (2015) Genetic parameters for spawning and growth traits in the Pacific white shrimp (Penaeus (*Litopenaeus*) *vannamei*). Aquac Res 46:833–839

Castillo-Juarez H, Campos-Montes GR, Caballero-Zamora A, Montaldo HH (2015) Genetic improvement of Pacific white shrimp [Penaeus (*Litopenaeus*) *vannamei*]: perspectives for genomic selection. Front Genet 6:93

Castillo-Juarez H, Casares JCQ, Campos-Montes G, Villela CC, Ortega AM, Montaldo HH (2007) Heritability for body weight at harvest size in the Pacific white shrimp, Penaeus (*Litopenaeus*) *vannamei*, from a multi-environment experiment using univariate and multivariate animal models. Aquaculture 273:42–49

Chen CH, Pan J, Di YQ, Liu W, Hou L, Wang JX, Zhao XF (2017) Protein kinase C delta phosphorylates ecdysone receptor B1 to promote gene expression and apoptosis under 20-hydroxyecdysone regulation. Proc Natl Acad Sci U S A 114:E7121–E7130

Correa K, Lhorente JP, López ME, Bassini L, Naswa S, Deeb N, di Genova A, Maass A, Davidson WS, Yáñez JM (2015) Genome-wide association analysis reveals loci associated with resistance against Piscirickettsia salmonis in two Atlantic salmon (*Salmo salar L.*) chromosomes. BMC Genomics 16:854

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet 12:499–510

Dekkers JC (2004) Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. J Anim Sci 82:313–328

Geng X, Liu S, Yao J, Bao L, Zhang J, Li C, Wang R, Sha J, Zeng P, Zhi D, Liu Z (2016) A genome-wide association study identifies

multiple regions associated with head size in catfish. G3 (Bethesda) 6:3389–3398

Geng X, Liu SK, Yuan ZH, Jiang YL, Zhi DG, Liu ZJ (2017) A genome-wide association study reveals that genes with functions for bone development are associated with body conformation in catfish. Mar Biotechnol 19:570–578

Goddard ME, Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat Rev Genet 10:381–391

Gonzalez-Pena D et al (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 et al (2012) Genetic mapping of quantitative trait loci (QTL) for body-weight in Atlantic salmon (Salmo salar) using a 6.5 K SNP array. Aquaculture 358:61–70

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(3):e0119730

Hayes B, Goddard M (2010) Genome-wide association and genomic selection in animal breeding. Genome 53:876–883

Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95–108

Hong EP, Park JW (2012) Sample size and statistical power calculation in genetic association studies. Genomics Inform 10:117–122

Huang XH, Han B (2014) Natural variations and genome-wide association studies in crop. Plants. Annu Rev Plant Biol 65:531–551

Huang XH, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z, Li M, Fan D, Guo Y, Wang A, Wang L, Deng L, Li W, Lu Y, Weng Q, Liu K, Huang T, Zhou T, Jing Y, Li W, Lin Z, Buckler ES, Qian Q, Zhang QF, Li J, Han B (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genet 42:961–976

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. Methods 25:402–408

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 293:587–599

Li WJ, Luan S, Luo K, Sui J, Xu XD, Tan J, Kong J (2015) Genetic parameters and genotype by environment interaction for cold tolerance, body weight and survival of the Pacific white shrimp Penaeus vannamei at different temperatures. Aquaculture 441:8–15

Li Y, Geng X, Bao L, Elaswad A, Huggins KW, Dunham R, Liu Z (2017) A deletion in the Hermansky-Pudlak syndrome 4 (Hps4) gene appears to be responsible for albinism in channel catfish. Mol Gen Genomics 292:663–670

Liu F et al (2014) A genome scan revealed significant associations of growth traits with a major QTL and GHR2 in tilapia. Sci Rep 4:7256

Liu ZJ, Cordes JF (2004) DNA marker technologies and their applications in aquaculture genetics. Aquaculture 238:1–37

Mackay TFC (2001) The genetic architecture of quantitative traits. Annu Rev Genet 35:303–339

Meng Q, Wang K, Liu X, Zhou H, Xu L, Wang Z, Fang M (2017) Identification of growth trait related genes in a Yorkshire purebred pig population by genome-wide association studies. Asian Australas J Anim Sci 30:462–469

Moss DR, Arce SM, Otoshi CA, Moss SM (2008) Inbreeding effects on hatchery and growout performance of Pacific white shrimp, Penaeus (Litopenaeus) vannamei. J World Aquacult Soc 39:467–476

Peng WZ, Xu J, Zhang Y, Feng J, Dong C, Jiang L, Feng J, Chen B, Gong Y, 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

Shyamal S, Das S, Guruacharya A, Mykles DL, Durica DS (2018) Transcriptomic analysis of crustacean molting gland (Y-organ) regulation via the mTOR signaling pathway. Sci Rep 8(1):7307

Sodeland M, Gaarder M, Moen T, Thomassen M, Kjoglum S, Kent M, Lien S (2013) Genome-wide association testing reveals quantitative trait loci for fillet texture and fat content in Atlantic salmon. Aquaculture 408:169–174

Song JL, Li Q, Yu Y, Wan S, Han LC, Du SJ (2018) Mapping genetic loci for quantitative traits of golden shell color, mineral element contents, and growth-related traits in Pacific oyster (Crassostrea gigas). Mar Biotechnol 20:666–675

Solovyev V, Kosarev P, Seledsov I, Vorobyev D (2006) Automatic annotation of eukaryotic genes, pseudogenes and promoters. Genome Biol 7(Suppl 1):P10.1–10.12

Spencer CC, Su Z, Donnelly P, Marchini J (2009) Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. PLoS Genet 5:e1000477

Sun XW, Liu D, Zhang X, Li W, Liu H, Hong W, Jiang C, Guan N, Ma C, Zeng H, Xu C, Song J, Huang L, Wang C, Shi J, Wang R, Zheng X, Lu C, Wang X, Zheng H (2013) SLAF-seq: an efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. PLoS One 8(3):e58700

Gitterle T, Rye M, Salte R, Cock J, Johansen H, Lozano C, Suárez JA, Gjerde B (2005) Genetic (co)variation in harvest body weight and survival in Penaeus (Litopenaeus) vannamei under standard commercial conditions. Aquaculture 243:83–92

Tan SX, Zhou T, Wang W, Jin Y, Wang X, Geng X, Luo J, Yuan Z, Yang Y, Shi H, Gao D, Dunham R, Liu Z (2018) GWAS analysis using interspecific backcross progenies reveals superior blue catfish alleles responsible for strong resistance against enteric septicemia of catfish. Mol Gen Genomics 293:1107–1120

Taira K, Umikawa M, Takei K, Myagmar BE, Shinzato M, Machida N, Uezato H, Nonaka S, Kariya KI (2004) The Traf2- and Nck-interacting kinase as a putative effector of Rap2 to regulate actin cytoskeleton. J Biol Chem 279:49488–49496

Tsai HY, Hamilton A, Tinch AE, Guy DR, Gharbi K, Stear MJ, Matika O, Bishop SC, Houston RD (2015) Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array. BMC Genomics 16:969

Wang X, Liu S, Jiang C, Geng X, Zhou T, Li N, Bao L, Li Y, Yao J, Yang Y, Zhong X, Jin Y, Dunham R, Liu Z (2017) Multiple across-strain and within-strain QTLs suggest highly complex genetic architecture for hypoxia tolerance in channel catfish. Mol Gen Genomics 292:63–76

Wen ZX, Tan R, Yuan J, Bales C, du W, Zhang S, Chilvers MI, Schmidt C, Song Q, Cregan PB, Wang D (2014) Genome-wide association mapping of quantitative resistance to sudden death syndrome in soybean. BMC Genomics 15:809

Yu Y, Zhang X, Yuan J, Li F, Chen X, Zhao Y, Huang L, Zheng H, Xiang J (2015) Genome survey and high-density genetic map construction provide genomic and genetic resources for the Pacific white shrimp Litopenaeus vannamei. Sci Rep 5:15612

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

Zhang H, Wang ZP, Wang SZ, Li H (2012) Progress of genome wide association study in domestic animals. J Anim Sci Biotechnol 3:26

Zhang XJ, Yuan J, Sun Y, Li S, Gao Y, Yu Y, Liu C, Wang Q, Lv X, Zhang X, Ma KY, Wang X, Lin W, Wang L, Zhu X, Zhang C, Zhang J, Jin S, Yu K, Kong J, Xu P, Chen J, Zhang H, Sorgeloos P, Sagi A, Alcivar-Warren A, Liu Z, Wang L, Ruan J, Chu KH, Liu B, Li F, Xiang J (2019) Penaeid shrimp genome provides insights into benthic adaptation and frequent molting. Nat Commun 10:356

Zhao M, Xia L, Chen GQ (2012) Protein kinase cdelta in apoptosis: a brief overview. Arch Immunol Ther Exp 60:361–372

Springer

Zheng XH, Kuang YY, Lv WH, Cao DC, Sun ZP, Sun XW (2016) Genome-wide association study for muscle fat content and abdominal fat traits in common carp (*Cyprinus carpio*). PLoS One 11(12): e0169127

Zhou T, Liu S, Geng X, Jin Y, Jiang C, Bao L, Yao J, Zhang Y, Zhang J, Sun L, Wang X, Li N, Tan S, Liu Z (2017) GWAS analysis of QTL for enteric septicemia of catfish and their involved genes suggest evolutionary conservation of a molecular mechanism of disease resistance. Mol Gen Genomics 292:231–242