Candidate Gene Polymorphisms and their Association with Glycogen Content in the Pacific Oyster *Crassostrea gigas*

Zhicai She[1,2], Li Li[1]*, Haigang Qi[1], Kai Song[1], Huayong Que[1], Guofan Zhang[1]*

1 National & Local Joint Engineering Laboratory of Ecological Mariculture, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, Shandong, China, 2 University of Chinese Academy of Sciences, Beijing, China

* gfzhang@qdio.ac.cn (GFZ); lili@qdio.ac.cn (LL)

Abstract

Background

The Pacific oyster *Crassostrea gigas* is an important cultivated shellfish that is rich in nutrients. It contains high levels of glycogen, which is of high nutritional value. To investigate the genetic basis of this high glycogen content and its variation, we conducted a candidate gene association analysis using a wild population, and confirmed our results using an independent population, via targeted gene resequencing and mRNA expression analysis.

Results

We validated 295 SNPs in the 90 candidate genes surveyed for association with glycogen content, 86 of were ultimately genotyped in all 144 experimental individuals from Jiaonan (JN). In addition, 732 SNPs were genotyped via targeted gene resequencing. Two SNPs (Cg_SNP_TY202 and Cg_SNP_3021) in *Cg_GD1* (glycogen debranching enzyme) and one SNP (Cg_SNP_4) in *Cg_GP1* (glycogen phosphorylase) were identified as being associated with glycogen content. The glycogen content of individuals with genotypes TT and TC in Cg_SNP_TY202 was higher than that of individuals with genotype CC. The transcript abundance of both glycogen-associated genes was differentially expressed in high glycogen content and low glycogen content individuals.

Conclusions

This study identified three polymorphisms in two genes associated with oyster glycogen content, via candidate gene association analysis. The transcript abundance differences in *Cg_GD1* and *Cg_GP1* between low- and the high-glycogen content individuals suggests that it is possible that transcript regulation is mediated by variations of Cg_SNP_TY202, Cg_SNP_3021, and Cg_SNP_4. These findings will not only provide insights into the genetic basis of oyster quality, but also promote research into the molecular breeding of oysters.

Introduction

The Pacific oyster *Crassostrea gigas* is an economically important mollusk that is widely cultivated all over the world, though China is the largest producer. Global production expands continuously at a rate of four million tons per year (FAO, 2014, http://www.fao.org). Oysters are an important source of glycogen, as it can make up 20–40% of their dry weight. The glycogen in oysters is associated with their energy metabolism and summer mortality, and can affect both gonad development and gametogenesis [1–3].

There have been many studies of glycogen metabolism in humans, and its molecular mechanism has been well addressed [4–8]. As it is an important nutrient, glycogen has been extensively studied in domesticated animals such as pigs, lambs, and chickens [9–12]. In bivalves, glycogen-related research lags far behind that conducted on humans and domesticated animals. To date, the target traits for the genetic dissection and breeding of oysters have mainly been associated with rapid growth and increased stress resistance [13–17]. Studies investigating the genetic basis of glycogen content in mollusks are scarce.

Glycogen is a branched glucose polymer present in many organisms, from bacteria and archaea to humans [18]. Glycogen metabolism is very conservative, and the number of enzymes directly involved in this process is quite small [19]. Glycogen phosphorylase has been extracted and purified from the adductor muscles of oysters and scallops by Keiko Hata [20,21], while Bacca cloned the oyster glycogen synthase gene and carried out the real-time fluorescent quantitative PCR, finding out that the expression level of glycogen synthase was consistent with seasonal changes in glycogen content [22]; however, the molecular mechanism behind glycogen metabolism is still unclear.

Oyster glycogen content is a quantitative trait, changing seasonally. Its production varies widely among different individuals [22,23]. It cannot be detected visually or in living animal, and thus conventional breeding methods are not suited to efforts aiming to breed oysters for nutritional quality. Marker-assisted selection seems to be the best alternative, and requires the screening of the genomic variation related to the target traits. Association analysis, including genome wide association study (GWAS) and candidate gene association study (CGA), is utilized extensively in order to link genomic and phenotypic variations. The pathway-driven identification and selection of candidate genes has proven to be a successful strategy for association studies [24], especially those investigating simple metabolic pathways, such as the metabolism of glycogen.

Single nucleotide polymorphisms (SNPs), mainly include the polymorphisms in a DNA sequence caused by single nucleotide mutations. SNPs have been widely used in genetic research and especially in association analysis due to the advantages of higher abundance, genotyping efficiency, data quality, genome-wide coverage and analytical simplicity [25–29]. The oyster genome and transcriptome predict abundant SNPs and genes related to the oyster glycogen metabolism [30,31], providing the genetic resources for our study. Of the various methods of SNP genotyping, high resolution melting HRM is powerful and cost-effective, and does not have the drawbacks of other methods [32–35]. In addition, the improved two-step HRM using a small amplicon is a more economic and effective method for moderate-throughput SNP genotyping in organisms with high polymorphism [31]. The cost decrease of next generation sequencing makes genotyping by sequencing (GBS) a more accessible method for association analysis in many species [36–40]. Additionally, targeted resequencing is an ideal method for candidate gene association studies. Targeted resequencing, more cost-effective and rational for sequencing target regions [41], is usually used to validate genes and mutations screened by association analysis [42,43]. The regulation of gene expression is an important genetic component of

quantitative traits, so analysis of the mRNA expression of the associated genes is usually used to further validate results [24,44,45].

We herein sort out the gens related to glycogen metabolism and work out a direct and simplified pathway, from which we screen a set of candidate genes for association analysis applying HRM in a wild population. Then, we validate the results in an independent population using targeted gene resequencing. Our research aims to identify the genes and mutations associated with glycogen content, and provide genetic tools for the marker-assisted selection of oysters.

Materials and Methods

Ethics Statement

The oysters used in this study were marine-cultured animals, and all experiments were conducted according to the regulations established by the local and central government. No specific permissions were required to collect oysters or conduct the experiments described.

Experimental animals

Wild oyster larvae from Jiaonan (JN, N35°44′, E119°56′), Qingdao, China were attached and cultured locally. At the age of 16 months, 144 individuals were sampled randomly for association analysis. The adductor muscle and mantle of each individual were dissected individually and snap-frozen in liquid nitrogen prior to being stored at -80°C for later use. Meanwhile, for the independent analysis, 144 individuals from Shentanggou (STG, N36°22′, E120°42′, 110 kilometers from JN), Qingdao were attached, cultured, and sampled in the same manner as those from JN. Sixty-four of the 144 STG individuals, consisting of the 32 individuals with the highest glycogen content, and the 32 with the lowest, were selected for targeted gene resequencing. In addition, gill samples from STG oysters with extreme phenotypic rankings, divided into high glycogen content (n = 16) and low glycogen content (n = 16) groups, were used for mRNA expression assays for the associated genes. Prior to dissection, shell height, length, and width, total weight, and tissue weight were measured. Details of the samples used in the experiment are shown in S1 Table.

Glycogen content assay

The adductor muscles of oysters from both JN and STG were used for the glycogen content assay along with a kit for detecting liver and muscle glycogen content (Nanjing Jiancheng Bioengineering Institute), according to the procedure described below: First, 30–60 mg adductor muscle were removed and rinsed using physiological saline, then weighed with a 0.1 mg electronic balance after the moisture had been blotted up using filter paper. Then, the sample was added to a test tube with a volume of alkaline liquor three times the weight of the sample. Then the samples were incubated for 20 min at 100°C in a water bath, followed by cooling under running water. Then, a volume of distilled water 16-fold the sample was added to the above hydrolysate for subsequent usage. Subsequently, 2 mL of the color regent was added and the tubes were incubated in a 100°C water bath for 5 min. Finally, the OD value of each tube was measured using a spectrophotometer under light of 1 cm diameter. The blank control and standard were set for each group (S2 Table). The glycogen content was calculated according to the formula below:

Glycogen content (mg/g) = (OD of test group/OD of standard group) $\cdot$ 0.01 $\cdot$ 20 $\cdot$ 10/1.11

where 1.11 is the coefficient of glucose content detected using this method, converted to glycogen content.

Screening of candidate genes and SNPs genotyping

We summarized the glycogen metabolism pathway using established knowledge, and searched for the associated enzyme-encoding genes in the predicted oyster gene set [30]. All examined genes encoding important enzymes were assigned as candidate genes related with glycogen content, and each of the candidate genes was aligned using blastn and blastp from NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) for sequence confirmation. As important mutations associated with the target traits are more likely located in the key genes and their surrounding regulatory regions [38,42,43], we studied not only the candidate gene regions, but also their upstream and downstream regions. Specifically, SNP validation using HRM and association analysis were performed for all the candidate genes in JN to locate the associated genes and SNPs. The SNPs were further validated by HRM in STG population. In addition, some selected candidate genes in the STG population were resequenced via Illumina sequencing to acquire more trait-associated SNPs.

HRM analysis. An improved small-amplicon HRM analysis [31] was used for SNP validation and genotyping. The genomic DNA of each individual was extracted from the mantle using the E.Z.N.A SQ Tissue DNA Kit. For each SNP predicted by the genome and transcriptome sequences of the candidate genes [30,31], we designed specific primers using the Primer Premier 5.0 software. The amplified size was limited to 50–100 bp, and only one SNP was included. Primers producing a single clear band on the polyacrylamide gel were used for the HRM analysis with eight samples. After amplification, LC Green and an interior label were added to each sample, followed by denaturation at 94°C for 10 min. Samples were then placed in a Light-Scanner 96 device (Idaho Technology, Salt Lake City, UT, USA) for HRM analysis. Genotypes were identified using Light Scanner 96 software (Idaho Technology), according to their melting temperatures. All the validated SNPs were genotyped in the JN population, while the associated SNPs were genotyped in STG population using the HRM analysis method described above.

Targeted gene resequencing. The regions targeted for sequencing included the selected genes and their 3 kb upstream and downstream regions. In addition to genes associated with glycogen content at association analysis stage I (details in the materials and methods section "Association and other statistical analysis"), *Cg_GS* and *Cg_GP*, enzymes known to rate-limit the process of glycogen metabolism, were selected as target genes for resequencing.

Prior to resequencing, cloning and sequencing were performed to fill the gaps in the genome sequence and ensure that the PCR amplification was specific to the target region. The filled genome sequence then function as the reference genome for subsequent SNP calling. The targeted regions were amplified using several overlapping pairs of primers designed using primer premier 5.0 software. All amplicons were linked to T vectors, and then transformed into the competent *E. coli* cells. After one night of incubation, monoclones were picked for sequencing via the Sanger sequencing method. Finally, 42 pairs of primers were selected for use in amplification of the target genes.

For each of the 64 (S1 Table) samples from STG, all PCR products were detected via agarose gel electrophoresis, then equally mixed. Then, the mixture of each individual was fragmented randomly and barcoded, in order to identify individuals after sequencing. Finally, the 64 samples were divided into four groups, each containing 16 individuals. We constructed libraries for each group and sequenced them on Illumina HiSeq 2000 in BGI.

Adapter sequences and contaminated sequences were removed, and reads in which low-quality ($Q \leq 5$) bases accounted for more than half of the read were filtered out. Then, we detected SNPs for each sample using SOAPsnp (http://soap.genomics.org.cn/soapsnp.html), after aligning the reads to the reference sequence using SOAP2 (http://soap.genomics.org.cn/

soapaligner.html) software. The SNP calling was made following the filtering criteria below: base quality higher than 20, sequencing depth higher than five and lower than 200, unique read alignment, missing rate lower than 10%, and a minor allele frequency higher than 0.05.

Association and other statistical analysis

SPSS version 16.0 was used to conduct the descriptive statistical analysis of the glycogen content. We also performed a correlation analysis to determine whether glycogen content was correlated with shell height, shell length, shell width, total weight, or tissue weight. Glycogen content among different genotypes was compared via ANOVA, using SPSS version 16.0.

We performed the first-stage candidate gene association analysis using the genotyped SNPs from JN to screen for associated loci and genes. Significantly associated loci were further validated in the STG population via HRM analysis. In addition, we carried out the second-stage association analysis with SNPs obtained via resequencing the associated genes from the STG population. We performed this analysis using GAPIT version 2.08 software via a mixed linear model approach implemented in EMMA, taking into account population structure and relative kinship [46–48]. The statistical model was:

$$Y = X\beta + Zu + e$$

where Y is the phenotype glycogen content, β is an unknown vector containing fixed effects such as the genetic marker, population structure (Q), and the intercept u, an unknown vector of random additive genetic effects. X and Z are the known design matrices, and e is the unobserved vector of the random residual.

We used PHASE version 2.1.1 to predict the SNP haplotype of the associated genes in the 64 resequenced samples from STG, then analyzed the association between haplotype and glycogen content with SPSS version 16.0, using one-way ANOVA.

The mRNA expression analysis of associated genes

The transcript abundances of the associated genes was detected using real-time quantitative PCR (Q-PCR). Total RNA was isolated from the gills of each individual in both the high glycogen content and the low glycogen content groups, using TRIzol reagent (Invitrogen). The integrity and concentration was detected by agarose gel electrophoresis and Nanodrop 2000 spectrophotometer (Thermo), respectively. First-strand cDNA was synthesized using Prime-Script RT reagent Kit with gDNA Eraser (TaKaRa), following the manufacturer's instructions. Quantitative Real-Time PCR was performed using an ABI 7500 FAST Real-time PCR System. The elongation factor (EF) gene was used as a normalizer gene. The amplification was carried out in a reaction volume of 20 μL, containing 10 μL of 2X SYBR Green PCR Master Mix (ABI), 0.4 μL each of the forward and reverse primers (10 μM), 1 μg of template cDNA, and ddH$_2$O for a final volume of 20 μL. The cycling program was as follows: 50°C for 2 min, 95°C for 2 m, then 40 cycles of 95°C for 3 s and 60°C for 30 s. The comparative CT method was used to determine fold-changes in gene expression, calculated as $2^{-\Delta\Delta CT}$.

Results and Discussion

Phenotypic statistics

The glycogen content of the experimental oysters ranged from 2.1 to 10.4 mg/g, with a mean ± std. value of 5.3 ± 1.7. The degree of variation was more than five-fold. Glycogen content was normally distributed (P > 0.05) (S1 Fig). The correlation analysis showed that glycogen content was not correlated with shell height, shell length, shell width, total weight, or tissue

Table 1. Statistics for the polymorphisms of the glycogen content candidate genes.

Enzymes	Encoding genes	Designed Primers	Validated SNPs	Genotyped SNPs
Cg_GS	1	22	9	5
Cg_PG	4	37	12	0
Cg_HKH	3	23	14	1
Cg_GP	2	42	14	1
Cg_GD	2	65	33	21
Cg_GT	74	400	216	50
Cg_GPK	3	76	26	6
Cg_PP	1	12	4	2
Total	90	677	295	86

Validated SNPs were validated by HRM using 8 samples, genotyped SNPs were genotyped by HRM using all the 144 samples. *Cg_GS*, glycogen synthase. *Cg_PG*, phosphoglucomutase. *Cg_HKH*, HK hexokinase. *Cg_GP*, glycogen phosphorylase. *Cg_GD*, glycogen debranching enzyme. *Cg_GT*, glycosyl transferase. *Cg_GPK*, glycogen phosphorylase kinase. *Cg_PP*, protein phosphatase.

doi:10.1371/journal.pone.0124401.t001

weight. This indicates that glycogen content is a quantitative trait suitable for association analysis.

Candidate gene screening of the glycogen metabolism pathway

The glycogen metabolism pathway is summarized in S2 Fig. Ninety genes encoding eight important enzymes were ultimately selected as candidate genes (S3 Table). They were classified into three groups: 1) glycogen anabolism enzyme, including glycogen synthase (*Cg_GS*, one encoding gene), phosphoglucomutase (*Cg_PG*, four encoding genes), and HK hexokinase (*Cg_HKH*, three encoding genes); 2) glycogen catabolism, including glycogen phosphorylase (*Cg_GP*, two encoding genes), glycogen debranching enzyme (*Cg_GD*, two encoding genes), and glycosyltransferase (*Cg_GT*, 74 encoding genes); 3) glycogen metabolism regulators, including glycogen phosphorylase kinase (*Cg_GPK*, three encoding genes) and protein phosphatase (*Cg_PP*, one encoding gene).

Genotyping by HRM and resequencing

High resolution melting (HRM) analysis. We designed a total of 677 pairs of primers in total for the 90 candidate genes (Table 1). Two hundred and ninety-five SNPs were validated preliminarily (S4 Table) for eight individuals from JN, and 86 were finally genotyped for all 144 JN experimental individuals (S1 Table). The overall ratio of SNP validation for eight oysters was 43.6% (295/677), which was consistent with the results of the oyster transcriptome SNP validation, 43.9% (1301/2962) [31]. The SNPs genotyped across all 144 JN individuals only accounted for 29.2% (86/295) of the total validated SNPs. This failure of the HRM analysis was mainly due to the presence of unpredicted SNPs in the PCR amplicon [31]. Only small amplified regions with a single SNP were well-genotyped via the HRM method, as unexpected SNPs in the PCR amplicon led to the failure of the HRM analysis. The presence of the unexpected SNPs in the oyster was derived from the high polymorphism and limited number of individuals used for SNP prediction in the non-genic region.

Targeted gene resequencing. To validate the candidate SNP markers from the stage I association analysis and to get more glycogen content related SNPs, we sequenced five selected target genes (S5 Table) in the independent population of STG. The five target genes included the associated *Cg_GD1* at the stage I and four extra key genes selected basing on the prior

Table 2. The distribution of SNPs for each resequenced gene.

Gene ID	Length(kb)	Total number	CDS	3'UTR	5'UTR
CGI_10016530	23.8	155	13	10	0
CGI_10023027	8.9	54	14	5	0
CGI_10005175	14.9	32	13	0	0
CGI_10005875	21	235	28	13	3
CGI_10017418	22	256	32	8	27
Total	90.6	732	100	36	30

doi:10.1371/journal.pone.0124401.t002

knowledge. We designed 85 pairs of primers in total to amplify the target gens as well as the 3 kb of both upstream and downstream of them. After PCR amplification, cloning, and sequencing, 42 pairs of primers were retained while 43 pairs discarded for producing no amplification products, or non-targeted amplicons. The details of the remaining 42 pairs of primers are shown in S6 Table. Sixty-four oysters consisting of 32 high-glycogen and 32 low-glycogen content oysters from the original 144 wild individuals (S1 Table) were sequenced. The descriptive statistics of the 32 low- and the 32 high-glycogen content individuals are shown in S7 Table. The mean glycogen content of the low group was 3.73 mg/g, with a range of 4.69 mg/g—2.19 mg/g. In the high glycogen-content group, the mean glycogen content was 8.83 mg/g, while the maximum and the minimum values were 10.2 mg/g and 7.84 mg/g, respectively.

The Q20 values of all filtered 1.26G reads were greater than 97.69%, indicating that the sequencing quality was very high. By aligning the reads to the 90.6 kb reference sequence, we identified 732 high quality SNPs. Table 2 shows the distribution of the SNPs on each candidate gene. The total number on each of the sequenced genes ranged from 32 in *Cg_GD2* to 256 in *Cg_GS*. On average, one SNP calling was made per 124 bp, which was sufficient for association analysis. The read coverage rate of each individual mostly ranged from 60–70% and the mapping rate from 40–50%, which was much lower than that achieved via the method using selector probes for targeted region enrichment [41]. This might be a result of the existence of highly duplicated sequences in the oyster genome [30], and the fact that the sequenced amplicons were not the targeted ones.

Candidate gene association analysis

Through the first stage of candidate gene association analysis with the 86 SNPs genotyped by HRM, we got one SNP (Cg_SNP_TY202, Y (C/T)) associated with glycogen content (P<0.01). It was located in the third exon of *Cg_GD1*. It was a synonymous mutation, and did not change the amino acid composition. The location and gene information of the SNP is shown in Table 3 and Fig 1. In the JN population, individuals with the TT and TC genotypes had much higher glycogen contents than those with the CC genotype (P<0.05). Individuals with the CC

Table 3. Information for all the SNPs associated with glycogen content.

SNPs	Scaffold	Position	Allels	P.value	Enzyme	Area	Genotyping method
Cg_SNP_TY202	41994	8325	C/T	0.0056	*Cg_GD1*	exon	HRM
Cg_SNP_3021	41994	3510	A/G	0.0058	*Cg_GD1*	intron	sequencing
Cg_SNP_4	1083	167235	C/T	0.0064	*Cg_GP1*	up stream	sequencing

Area was the relative location of the SNP to the gene. *Cg_GD1*, glycogen debranching enzyme encoding gene 1. *Cg_GP1*, glycogen phosphorylase encoding gene 1.
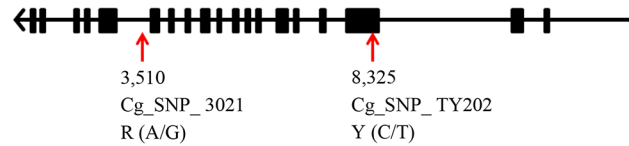
doi:10.1371/journal.pone.0124401.t003

Fig 1. Location of the two associated SNPs in *Cg_GD1*.

doi:10.1371/journal.pone.0124401.g001

genotype were observed at a much lower frequency in the STG population (0.007) than in the JN population (0.11). We only found one individual with the CC genotype among the 144 individuals from the STG population (S1 Table). In addition, the average glycogen content of STG individuals was significantly higher at 6.17mg/g (P<0.01) than in JN individuals (5.35 mg/g) (S8 Table). This might support the finding that TT and TC individuals from the JN population had much higher glycogen contents than those with the CC genotype.

We used the 732 high quality SNPs obtained via sequencing for the second stage of our association analysis. Two of the five targeted genes were located on two different linkage groups (unpublished data), and the other three were not mapped. We assumed that the different targeted genes are located on different chromosomes (S9 Table), for convenience of analysis and display. A quantile-quantile plot (Fig 2) and a Manhattan plot of the p-value (Fig 3) were produced. The association analysis results showed that two SNPs (P<0.01) were associated with the target trait glycogen content. One was Cg_SNP_3021 (R (A/G)), located on the 15th intron of *Cg_GD1*. Cg_SNP_3021 was 4815 bp from Cg_SNP_TY202, which further validated that *Cg_GD1* was associated with glycogen content. Cg_SNP_TY202 failed to be located through genotyping by sequencing. According to the results of the HRM analysis, 54 of the 64 (S1 Table) sequenced oysters were homozygous at the locus Cg_SNP_TY202, and only 11 samples had allele C. We analyzed the sequencing reads of the 11 individuals with allele C at Cg_SNP_TY202, and found that the depth was either too high or too low for SNP calling. The other SNP (Cg_SNP_4, Y (C/T)) associated with glycogen content was located upstream of *Cg_GP1*, 2773 bp from 5'UTR. Table 3 and Fig 1 provide detailed information for all associated SNPs.
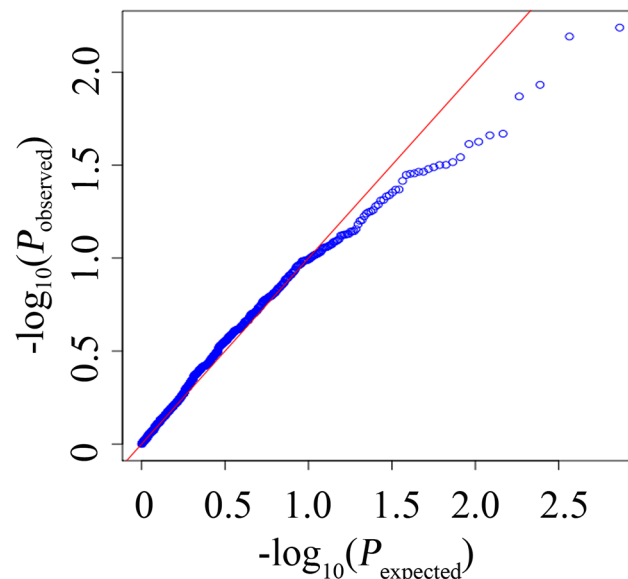

Fig 2. Quantile-quantile plot for glycogen content.

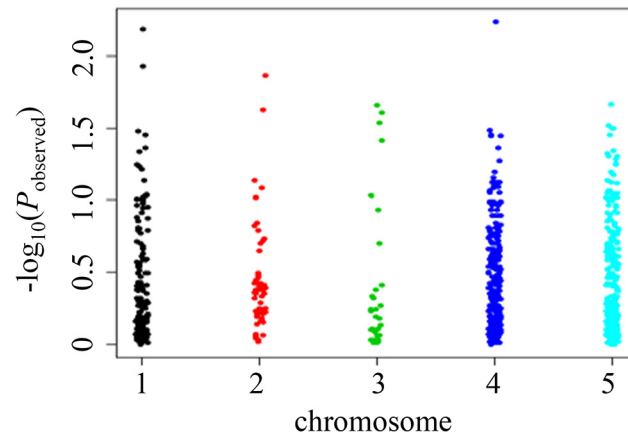doi:10.1371/journal.pone.0124401.g002

Fig 3. A Manhattan plot showing the associations of all SNPs.

doi:10.1371/journal.pone.0124401.g003

We then performed haplotype analysis. For *Cg_GD1*, three haplotypes were established by PHASE using Cg_SNP_3021 and Cg_SNP_TY202 (S10 Table). The haplotypes were AT, AC, and GT, and the frequencies were 0.90, 0.08, and 0.02 respectively. Because of the uneven distribution of the allele frequency and the lack of haplotype GC individuals, there was no clear pattern of glycogen content across different haplotypes.

The mRNA expression of associated genes

The glycogen content of the high- and low-glycogen content groups was 9.39 ± 0.48 (mg/g) and 3.15 ± 0.62 (mg/g), respectively. For *Cg_GP1*, the transcript abundance in the high glycogen content group was significantly down-regulated, compared that of the low glycogen content group (P<0.05). For *Cg_GD1*, the expression difference between the two groups was not significant (P>0.05), but the transcript abundance in the high glycogen content group was almost 50% lower than that of the low group (Fig 4). The lack of significant differences in *Cg_GD1* expression might due to the presence of multiple mutations in the genes determining
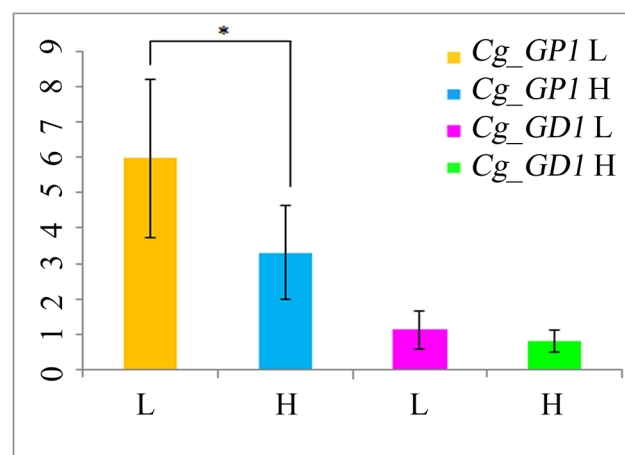

Fig 4. The mRNA expression levels of the high glycogen content and the low glycogen content groups. * P < 0.05, the difference between the high glycogen content and the low glycogen content group of *Cg_GD1* was significant. *Cg_GD1*, glycogen debranching enzyme encoding gene 1. *Cg_GP1*, glycogen phosphorylase encoding gene 1.

doi:10.1371/journal.pone.0124401.g004

glycogen content. In addition, the identified loci could only explain a part of phenotypic varia-tion, and thus did not govern the consistency of *Cg_GD1* expression within groups. This could lead to differential expression within groups and a lack of significant differences between groups. Whether significant or not, the expression differences between the two genes implied that it is possible that polymorphisms associated with the Cg_SNP_TY202, Cg_SNP_3021, and Cg_SNP_4 might regulate the expression of *Cg_GP1* and *Cg_GD1* [38]. These data support the results of the association analysis, which indicated that *Cg_GP1* and *Cg_GD1* are associated with glycogen content at the transcript level.

The two enzymes both participated in the catabolism of glycogen, suggesting that it might be catabolism rather than anabolism that decides an individual's glycogen content. It has been reported that triploid oysters had higher glycogen contents and lower physiological activity lev-els than diploid oysters; the mortality rate of triploids was about half that of diploids, and trip-loids produced higher quality meat than diploids [1,2]. It was deduced that oysters with higher glycogen contents, which did not exhibit active glycogen catabolism, performed better with re-spect to environmental stress resistance and nutritional status. The identified SNPs located on *Cg_GP1* and *Cg_GD1*, which participate glycogen catabolism, are potentially useful for the mo-lecular breeding of the oysters.

Conclusions

From the glycogen metabolism pathway of oysters, we screened 90 encoding genes of eight en-zymes. Two hundred and ninety-five SNPs were validated via HRM analysis; 732 SNPs were called by resequencing the candidate genes and the 3 kb upstream and downstream regions of the genes. The candidate glycogen content genes and SNPs are valuable genetic resources for further studies on oyster. Cg_SNP_TY202 and Cg_SNP_3021 in *Cg_GD1*, and Cg_SNP_4 in *Cg_GP1* are associated with the glycogen content. The transcript abundance differences be-tween the low and the high glycogen content individuals suggest that it is possible that the tran-script regulation of *Cg_GD1* and *Cg_GP1* is mediated via the variations associated with associated SNPs. These findings will not only provide insights into the genetic basis of oyster quality traits, but also promote the research on the molecular breeding of oysters.

Supporting Information

S1 Fig. Histogram of glycogen content.
(TIF)

S2 Fig. The glycogen metabolism pathway. Yellow indicates enzymes that we selected as can-didate genes. White indicates the substrates of those enzymes. UDPG is uridine diphosphate glucose. Glycogen G (n+1) unbranched contains Alpha1, 4-glycosidic bond only. Glycogen G (n+1) contains both Alpha1, 4-glycosidic bond and Alpha1, 6-glycosidic bond.
(TIF)

S1 Table. Statistics of samples used in the experiment. 64 included 32 with the highest glyco-gen content and 32 with the lowest from the original 144 STG individuals. 32 included 16 with the highest glycogen content and 16 with the lowest from the original 144 STG individuals.
(XLSX)

S2 Table. Operation table for glycogen content assay. The 0.01 mg/mL standard solution was diluted from the 1 mg/mL standard reserve liquid. The color reagent was made from color powder and concentrated sulfuric acid.
(XLSX)

S3 Table. List of candidate genes for the glycogen metabolism pathway.
(XLSX)

S4 Table. Information for the 295 validated SNPs. Gene was the gene ID in which SNP located. CDS, coding sequence. INT, intron. INTER, intergenic region.
(XLSX)

S5 Table. Information for genes selected for targeted gene resequencing.
(XLSX)

S6 Table. Details for the remaining 42 pairs of primers for targeted gene resequencing.
(XLSX)

S7 Table. Descriptive statistics for the low and high glycogen content groups.
(XLSX)

S8 Table. The descriptive statistics of different genotypes for Cg_SNP_TY202 in the two populations. The unit is mg/g.
(XLSX)

S9 Table. The putative chromosome number for each gene.
(XLSX)

S10 Table. The descriptive statistics of different haplotypes.
(XLSX)

Acknowledgments

Author Contributions

Conceived and designed the experiments: LL GFZ ZCS. Performed the experiments: ZCS. Analyzed the data: ZCS KS HGQ. Contributed reagents/materials/analysis tools: LL HYQ GFZ. Wrote the paper: ZCS LL.

References

1. Maguire GB, Gardner NC, Nell JA, Kent GN, Kent AS (1995) Studies on triploid oysters in Australia. 2. Growth, condition index gonad area, and glycogen content of triploid and diploid Pacific oysters, *Crassostrea gigas*, from oyster leases in Tasmania, Australia. Aquaculture 137: 357–357.

2. Akashige S, Fushimi T (1992) Growth, survival, and glycogen-content of triploid Pacific oyster *Crassostrea gigas* in the waters of Hiroshima, Japan. Nippon Suisan Gakkaishi 58: 1063–1071.

3. Samain JF (2011) Review and perspectives of physiological mechanisms underlying genetically-based resistance of the Pacific oyster *Crassostrea gigas* to summer mortality. Aquatic Living Resources 24: 227–236.

4. Bollig G (2013) McArdle's disease (glycogen storage disease type V) and anesthesia-a case report and review of the literature. Pediatric Anesthesia 23: 817–823. doi: 10.1111/pan.12164 PMID: 23565573

5. Akram M, Asif HM, Akhtar N, Shah PA, Uzair M, Shaheen G, et al. (2011) Glycogen metabolism and glycogen storage diseases: a review. Journal of Medicinal Plants Research 5: 4980–4983.

6. Bezborodkina NN, Chestnova AY, Okovity SV, Kudryavtsev BN (2014) Activity of glycogen synthase and glycogen phosphorylase in normal and cirrhotic rat liver during glycogen synthesis from glucose or fructose. Experimental and Toxicologic Pathology 66: 147–154. doi: 10.1016/j.etp.2013.12.001 PMID: 24373751

7. Kao WHL, Coresh J, Shuldiner AR, Boerwinkle E, Brancati FL (1999) Associations of type 2 diabetes mellitus with gene variants of glucagon receptor (GCGR), sulfonylurea receptor (SUR), peroxisome proliferator activated receptor-gamma 2 (PPAR gamma 2), and glycogen-targeting subunit of type I protein phosphatase (PPP1R3) in African Americans: the atherosclerosis risk in communities (ARIC) study. American Journal of Epidemiology 149: S3–S3.

8. Dmitrzak-Weglarz M, Szczepankiewicz A, Slopien A, Rybakowski E, Hauser J, Rajewski A, et al. (2005) An association study of the glycogen synthase kinase 3-beta (GSK3-beta) gene in anorexia nervosa. European Neuropsychopharmacology 15: S600–S600.

9. Díaz MT, Vieira C, Pérez C, Lauzurica S, González de Chávarri E, Sánchez M, et al. (2014) Effect of lairage time (0 h, 3 h, 6 h or 12 h) on glycogen content and meat quality parameters in suckling lambs. Meat Science 96: 653–660. doi: 10.1016/j.meatsci.2013.10.013 PMID: 24200553

10. Chaosap C, Parr T, Wiseman J (2011) Effect of compensatory growth on forms of glycogen, postmortem proteolysis, and meat quality in pigs. Journal of Animal Science 89: 2231–2242. doi: 10.2527/jas.2010-2953 PMID: 21317347

11. Berri C, Debut M, Santé-Lhoutellier V, Arnould C, Boutten B, Sellier N, et al. (2005) Variations in chicken breast meat quality: implications of struggle and muscle glycogen content at death. British Poultry Science 46: 572–579. PMID: 16359110

12. Ciobanu D, Bastiaansen J, Malek M, Helm J, Woollard J, Plastow G, et al. (2001) Evidence for new alleles in the protein kinase adenosine monophosphate-activated gamma (3)-subunit gene associated with low glycogen content in pig skeletal muscle and improved meat quality. Genetics 159: 1151–1162. PMID: 11729159

13. Zhao XL, Yu H, Kong LF, Li Q (2012) Transcriptomic responses to salinity stress in the Pacific oyster Crassostrea gigas. PLoS ONE 7: e46244. doi: 10.1371/journal.pone.0046244 PMID: 23029449

14. Meng J, Zhu QH, Zhang LL, Li CY, Li L, She ZC, et al. (2013) Genome and transcriptome analyses provide insight into the euryhaline adaptation mechanism of Crassostrea gigas. PLoS ONE 8: e58563. doi: 10.1371/journal.pone.0058563 PMID: 23554902

15. Du XD, Li L, Zhang SD, Meng F, Zhang GF (2014) SNP identification by transcriptome sequencing and candidate gene-based association analysis for heat tolerance in the bay scallop Argopecten irradians. PLoS ONE 9: e104960. doi: 10.1371/journal.pone.0104960 PMID: 25121601

16. Frank-Lawale A, Allen SK, Degremont L (2014) Breeding and domestication of eastern oyster (Crassostrea virginica) lines for culture in the Mid-Atlantic, USA: line development and mass selection for disease resistance. Journal of Shellfish Research 33: 153–165.

17. Gionet C, Herbinger C, Despres M (2012) Next steps towards the development of the American oyster (Crassostrea virginica) breeding program in New Brunswick. Aquaculture Canada 2012—Frontiers: Bridging Technology and Economic Growth: 29–31.

18. Roach PJ, Depaoli-Roach AA, Hurley TD, Tagliabracci VS (2012) Glycogen and its metabolism: some new developments and old themes. Biochem J 441: 763–787. doi: 10.1042/BJ20111416 PMID: 22248338

19. Roach PJ (2002) Glycogen and its metabolism. Curr Mol Med 2: 101–120. PMID: 11949930

20. Hata K, Hata M, Matsuda K (1993) Purification and properties of glycogen-phosphorylase from the adductor muscle of the oyster, Crassostrea gigas. Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology 105: 481–486.

21. Hata K, Yokoyama I, Suda M, Hata M, Matsuda K (1987) Purification and properties of glycogen-phosphorylase from the adductor muscle of the scallop, Patinopecten yessoensis. Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology 87: 747–753.

22. Bacca H, Huvet A, Fabioux C, Daniel JY, Delaporte M, Pouvreau S, et al. (2005) Molecular cloning and seasonal expression of oyster glycogen phosphorylase and glycogen synthase genes. Comparative Biochemistry and Physiology Part B, Biochemistry & Molecular Biology 140: 635–646.

23. Aagesen AM, Hase CC (2014) Seasonal effects of heat shock on bacterial populations, including artificial vibrio parahaemolyticus exposure, in the Pacific oyster, Crassostrea gigas. Food Microbiology 38: 93–103. doi: 10.1016/j.fm.2013.08.008 PMID: 24290632

24. Yan JB, Warburton M, Crouch J (2011) Association mapping for enhancing maize (Zea mays L.) genetic improvement. Crop Science 51: 433.

25. Brookes AJ (1999) The essence of SNPs. Gene 234: 177–186. PMID: 10395891

26. Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM, et al. (2011) A map of local adaptation in Arabidopsis thaliana. Science 334: 86–89. doi: 10.1126/science.1209271 PMID: 21980109

27. Sanchez CC, Smith TP, Wiedmann RT, Vallejo RL, Salem M, Yao J, et al. (2009) Single nucleotide polymorphism discovery in rainbow trout by deep sequencing of a reduced representation library. BMC Genomics 10: 559. doi: 10.1186/1471-2164-10-559 PMID: 19939274

28. Park YJ, Lee JK, Kim NS (2009) Simple sequence repeat polymorphisms (SSRPs) for evaluation of molecular diversity and germplasm classification of minor crops. Molecules 14: 4546–4569. doi: 10.3390/molecules14114546 PMID: 19924085

29. Morin PA, Luikart G, Wayne RK, SNP workshop group (2004) SNPs in ecology, evolution and conservation. Trends in Ecology & Evolution 19: 208–216.

30. Zhang GF, Fang XD, Guo XM, Li L, Luo RB, Xu F, et al. (2012) The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490: 49–54. doi: 10.1038/nature11413 PMID: 22992520

31. Wang JF, Qi HG, Li L, Que HY, Wang D, Zhang GF, et al. (2014) Discovery and validation of genic single nucleotide polymorphisms in the Pacific oyster *Crassostrea gigas*. Molecular Ecology Resources 11:123–135.

32. Seeb LW, Antonovich A, Banks AA, Beacham TD, Bellinger AR, Blankenship SM, et al. (2007) Development of a standardized DNA database for Chinook salmon. Fisheries 32: 540–552.

33. Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS (2005) A genome-wide scalable SNP genotyping assay using microarray technology. Nature Genetics 37: 549–554. PMID: 15838508

34. Fan JB, Oliphant A, Shen R, Kermani BG, Garcia F, Gunderson KL, et al. (2003) Highly parallel SNP genotyping. Cold Spring Harbor Symposia on Quantitative Biology 68: 69–78. PMID: 15338605

35. Yan JB, Yang XH, Shah T, Sánchez-Villeda H, Li JS, Warburton M, et al. (2009) High-throughput SNP genotyping with the GoldenGate assay in maize. Molecular Breeding 25: 441–451.

36. Rubin CJ, Zody MC, Eriksson J, Meadows JR, Sherwood E, Webster MT, et al. (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. Nature 464: 587–591. doi: 10.1038/nature08832 PMID: 20220755

37. Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV (2010) Population resequencing reveals local adaptation of Arabidopsis lyrata to serpentine soils. Nature Genetics 42: 260–263. doi: 10.1038/ng.515 PMID: 20101244

38. Li H, Peng ZY, Yang XH, Wang WD, Fu JJ, Wang JH, et al. (2013) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. Nature Genetics 45: 43–50. doi: 10.1038/ng.2484 PMID: 23242369

39. Salem M, Vallejo RL, Leeds TD, Palti Y, Liu S, Sabbagh A, et al. (2012) RNA-Seq identifies SNP markers for growth traits in rainbow trout. PLoS ONE 7: e36264. doi: 10.1371/journal.pone.0036264 PMID: 22574143

40. Zhou XM, Wang BS, Pan Q, Zhang JB, Kumar S, Sun XQ, et al. (2014) Whole-genome sequencing of the snub-nosed monkey provides insights into folivory and evolutionary history. Nature Genetics 46:1303–1310. doi: 10.1038/ng.3137 PMID: 25362486

41. Johansson H, Isaksson M, Sörqvist EF, Roos F, Stenberg J, Sjöblom T, et al. (2011) Targeted resequencing of candidate genes using selector probes. Nucleic Acids Research 39: e8. doi: 10.1093/nar/gkq1005 PMID: 21059679

42. Wang ZP, Qu LJ, Yao JF, Yang XL, Li GQ, Zhang YY, et al. (2013) An EAV-HP insertion in 5' flanking region of *SLCO1B3* causes blue eggshell in the chicken. PLoS Genetics 9: e1003183. doi: 10.1371/journal.pgen.1003183 PMID: 23359636

43. Yan JB, Kandianis CB, Harjes CE, Bai L, Kim EH, Yang XH, et al. (2010) Rare genetic variation at *Zea mays crtRB1* increases beta-carotene in maize grain. Nature Genetics 42: 322–327. doi: 10.1038/ng.551 PMID: 20305664

44. Sun YF, Zhao GP, Liu RR, Zheng MQ, Hu YD, Wu D, et al. (2013) The identification of 14 new genes for meat quality traits in chicken using a genome-wide association study. BMC Genomics 14.

45. Larsen PF, Schulte PM, Nielsen EE (2011) Gene expression analysis for the identification of selection and local adaptation in fishes. Journal of Fish Biology 78: 1–22. doi: 10.1111/j.1095-8649.2010.02834.x PMID: 21235543

46. Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, et al. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature Genetics 38: 203–208. PMID: 16380716

47. Zhang ZW, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, et al. (2010) Mixed linear model approach adapted for genome-wide association studies. Nature Genetics 42: 355–U118. doi: 10.1038/ng.546 PMID: 20208535

48. Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, et al. (2008) Efficient control of population structure in model organism association mapping. Genetics 178: 1709–1723. doi: 10.1534/genetics.107.080101 PMID: 18385116