He *et al. BMC Genomics* (2024) 25:931
https://doi.org/10.1186/s12864-024-10857-3

BMC Genomics

RESEARCH **Open Access**

Genome-wide association study reveals the underlying regulatory mechanisms of red blood traits in *Anadara granosa*

Xin He[1,2,3], Yushan Liao[2], Gaowei Yu[2], Shi Wang[3] and Yongbo Bao[1,2*]

Abstract

Background *Anadara granosa*, commonly known as the blood clam, exhibits the unusual characteristic of having red blood among invertebrates. There is significant individual variation in blood color intensity among blood clams; individuals with vibrant red blood are deemed healthier and exhibit stronger stress resistance. However, the molecular basis underlying these red blood traits (RBTs) remains poorly understood.

Results In this study, we performed genome-wide association studies (GWAS) in a population of 300 *A. granosa* individuals, focusing on RBTs as measured by hemoglobin concentration (HC), total hemocyte count (THC), and heme concentration (HEME). Our analysis identified 18 single nucleotide polymorphisms (SNPs) correlated with RBTs, subsequently selected 117 candidate genes within a 100 kb flanking region of these SNPs, potentially involved in the RBTs of *A. granosa*. Moreover, we discovered two haplotype blocks specifically associated with THC and HEME. Further analysis revealed eight genes (*Septin7*, *Hox5*, *Cbfa2t3*, *Avpr1b*, *Hhex*, *Eif2ak3*, *Glrk*, and *Rpl35a*) that significantly influence RBTs. Notably, a heterozygous A/T mutation in the 3'UTR of *Cbfa2t3* was found to promote blood cell proliferation. These genes suggest that the hematopoietic function plays a significant role in the variability of RBTs in *A. granosa*.

Conclusions Our findings reveal a conservation of the regulatory mechanisms of RBTs between blood clams and vertebrates. The results not only provide a scientific basis for selective breeding in blood clams, but also offer deeper insights into the evolutionary mechanisms of RBTs in invertebrates.

Keywords *Anadara granosa*, Red blood traits, GWAS, SNP, Hematopoietic function

Background

Anadara granosa, also known by its species name *Tegillarca granosa* and commonly referred to as the blood clam, is a marine bivalve species that is extensively farmed along the eastern coast of China and throughout Southeast Asia, representing a valuable marine bioeconomic resource [1, 2]. In the majority of invertebrates, hemolymph can appear slightly bluish due to the presence of hemocyanin as the respiratory protein, or it may be colorless in the absence of respiratory pigments [3–5]. Surprisingly, blood clams possess rare red-colored blood among mollusks, a characteristic attributed to the heme group within hemoglobin [4]. Moreover, hemoglobin is synthesized by erythrocytes, which constitute nearly 90% of the total hemocyte count in blood clams [6]. Therefore, the hemoglobin concentration (HC), total hemocyte count (THC), and heme concentration (HEME) serve as

---

*Correspondence:
Yongbo Bao
baoyongbo@zwu.edu.cn
[1] Ninghai Institute of Mariculture Breeding and Seed Industry, Zhejiang Wanli University, Ninghai 315604, China
[2] Key Laboratory of Aquatic Germplasm Resource of Zhejiang, College of Biological & Environmental Sciences, Zhejiang Wanli University, Ningbo 315100, China
[3] Key Laboratory of Marine Genetics and Breeding, Ocean University of China, Qingdao 266003, China

crucial parameters of the red blood traits (RBTs) in blood clams.

In vertebrates, hemoglobin is mainly responsible for oxygen transport [7], a role it also fulfills in the blood clam, *A. granosa* [4]. *A. granosa*, a species that buries itself in sediment, has hemoglobin that has evolved to adapt to extremely hypoxic environments [8]. In addition to low oxygen tolerance, hemoglobin of *A. granosa* also exhibits antimicrobial effects against *Vibrio parahaemolyticus* [9, 10] and likely functions as peroxidases, aiding in the defense mechanisms of bivalve mollusks [11]. Erythrocytes of *A. granosa* demonstrate lysosomal and oxidative capacities involved in immunological activities [6, 12]. Notably, the bioavailability of heme iron in blood clams significantly contributes to enhancing human immunity [13–15]. Our research indicates significant individual variation in blood color intensity among blood clams (Fig. 1a), with individuals displaying vibrant red blood deemed healthier and more adaptable to environments of high temperature or high salinity [16]. Therefore, investigating the RBTs of *A. granosa* holds substantial biological importance. Nevertheless, few studies have focused on the molecular mechanisms regulating these differences in RBTs among blood clams.

Previous studies showed that RBTs can be influenced by genetic factors [17]. Genes related to hemoglobin and their regulation pathways, as well as the regulation associated with heme synthesis, iron metabolism, and hematopoiesis, collectively affect the ultimate RBTs. Genome-wide association studies (GWAS) are a powerful tool for unraveling genetic variations related to complex quantitative traits and selecting the corresponding candidate genes, thereby offering genes and markers for selective breeding initiatives [18, 19].

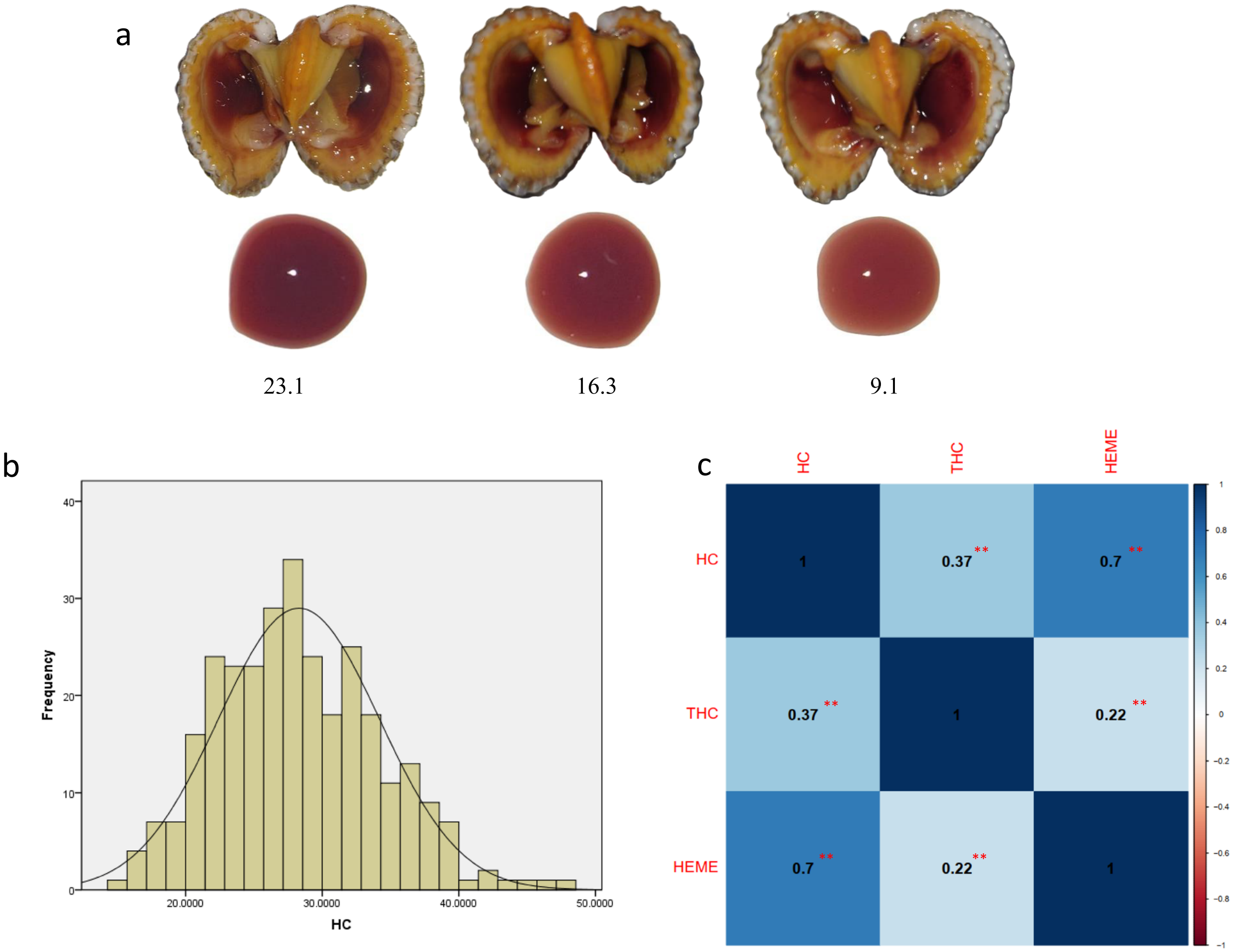

Fig. 1 **a** Variation in hemoglobin concentrations (HC) (g/L) observed in *Anadara granosa*. **b** Phenotypes of HC in 300 *A. granosa* individuals. For associated phenotypes, including total hemocyte count (THC) and heme concentration (HEME), refer to Fig. S2. **c** Phenotypic correlation analysis among red blood traits of *A. granosa*. The asterisk symbol '**' denotes a significant correlation at the 0.01 level (two-tailed)

Currently, GWAS is widely used to identify genes and SNP mutations associated with phenotypes related to RBTs in vertebrate species [17, 20, 21]. Genes associated with iron homeostasis, such as *TMPRSS6*, *HFE*, and *TFR2*, have been discovered through GWAS to be linked to traits including mean corpuscular hemoglobin content (MCH), the volume of red blood cells (MCV), and red blood cell count (RBC) [17, 22]. Notably, SNP mutations in *TMPRSS6* and *HFE* have also been independently validated for their association with hemoglobin levels in different populations [23]. GWAS has implicated both *HBS1L-MYB* and *BCL11A* in the regulation of fetal globin expression, and the locus of *HBS1L-MYB* has been involved in broader aspects of erythropoiesis. [17, 22, 24]. Furthermore, another GWAS conducted on a large population cohort from Sardinia revealed five variants at previously unidentified loci: *MPHOSPH9*, *PLTP-PCIF1*, *ZFPM1 (FOG1)*, *NFIX*, and *CCND3*. Besides this, among the signals at known loci, half of these variants also exhibited pleiotropic associations with various hemoglobin traits [25]. Based on research findings in vertebrates, we hypothesize that the variation in RBTs in *A. granosa* may also be associated with genetic variation at key gene loci. However, to date, no studies have been reported on the genetic variation in RBTs of *A. granosa*.

Hence, we conducted a GWAS on RBTs of *A. granosa*, identifying SNPs associated with HC, THC, and HEME as potential candidates. We screened nearby candidate genes related to these SNP markers, subsequently validating them through haplotype analysis and quantitative real-time polymerase chain reaction (qRT–PCR). The results of this study could enhance our understanding of the regulatory mechanisms underlying individual blood color variation in *A. granosa*.

Results

Phenotype statistics of red blood traits

RBTs, including HC, THC, and HEME, were measured in 300 individual *A. granosa* (Table 1). All phenotypic data displayed a normal distribution, making them suitable for GWAS analysis (Fig. 1b and S1). Phenotypic correlation analysis among HC, THC, and HEME was significant, and revealed a strong positive correlation between HC and HEME, with a coefficient of 0.7 (Fig. 1c). Genetic correlation results showed a strong correlation between HC and HEME (0.898 ± 0.095) (Fig. S2), consistent with the phenotypic correlation findings. However, the phenotypic correlations between THC and both HC and HEME were found to be weak, exhibiting coefficients ranging from 0.20 to 0.39 (Fig. 1c).

Table 1 Descriptive statistics for *A. granosa* red blood traits

	HC (g/L)	THC ($\times 10^6$ cell/ml)	HEME ($\times$6250 μM)
Num	300	300	300
Mean	28.29	112	0.16
SD	5.90	46.39	0.03
CV	0.21	0.42	0.20
Min	15.54	29	0.10
Max	47.65	260	0.26

Num Number, *SD* Standard deviation, *CV* Coefficient of variance, *HC* Hemoglobin concentration, *THC* Total hemocyte count, and *HEME* Heme concentration

Genome-wide association study (GWAS) of red blood traits

To uncover the genes and mechanisms associated with RBTs in *A. granosa*, a GWAS analysis was conducted. The sequencing of 300 individual *A. granosa*, along with kinship and PCA, was completed and detailed in a prior study [26]. A total of 3114 Gb of high-quality sequencing data was obtained, and 355,254 high-quality SNPs were filtered [26]. The results of the GWAS analysis, focusing on RBTs, were illustrated in the Manhattan and quantile–quantile (QQ) plots (Fig. 2).

We identified 5, 4, and 9 SNPs ($P < 10^{-5}$) associated with HC, THC, and HEME, respectively. Among these, one SNP each was located in the 3'UTR, 5'UTR, and downstream regions of genes. Additionally, two SNPs were found in gene exon regions, and three resided within gene intron regions. The remaining SNPs were located in intergenic regions (Table S1). The proportion of phenotypic variation explained (PVE) by these SNPs for the associated RBTs ranged from 12.68% to 17.55% (Table S1). Furthermore, a total of 117 candidate genes situated within 100 kb upstream and downstream of these SNPs were identified (Table S2).

Enrichment analysis

GO and KEGG enrichment analysis were conducted for candidate genes of HC, THC, and HEME. The candidate genes associated with HC are involved in intracellular processes, notably in the regulation of transcription, DNA-templated, RNA biosynthetic processes, and more (Fig. S3a). KEGG analysis indicated these genes were enriched in pathways like "Transcription factors", "CD molecules", "Starch and sucrose metabolism", "Aminoacyl-tRNA biosynthesis", and more (Fig. S3b). For THC, GO analysis revealed enrichment in molecular functions related to transmembrane signaling receptor activity, G protein-coupled receptor activity, signaling receptor activity, and molecular transducer activity (Fig. S3c). The candidate genes are

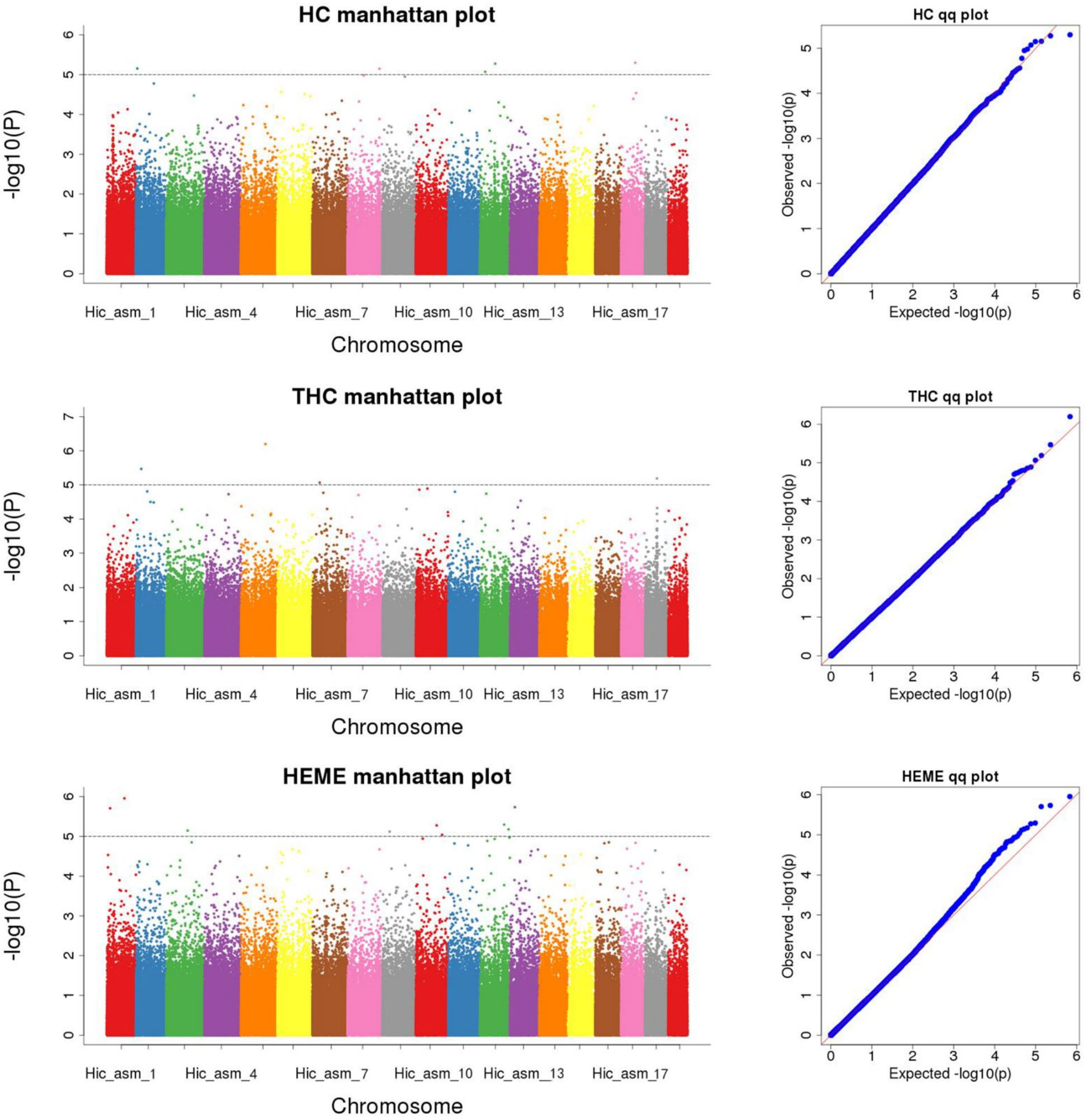

Fig. 2 Manhattan plots and QQ plots of genome-wide association studies for red blood traits in *A. granosa*. The Manhattan plot displays the − log10 (observed *P*-values) for the genome-wide SNPs (y-axis) mapped against their respective positions on each scaffold (x-axis), with the horizontal red line representing the genome-wide suggestive threshold (10^{-5}). In the QQ plot, the x-axis represents the expected − log10 transformed *P*-values, while the y-axis shows the observed − log10 transformed *P*-values

linked to the integral component of the membrane and are primarily involved in biological processes associated with cellular processes (Fig. S3c), engaging in pathways such as "Cytoskeleton proteins", "Acute myeloid leukemia", "Ubiquitin system", and more (Fig. S3d). HEME-associated candidate genes play roles in protein and cellular protein metabolic processes (Fig. S3e), with enriched pathways including "Glutamatergic synapse", "Ion channels", "Cutin, suberin and wax biosynthesis", "Protein families: metabolism", and more (Fig. S3f). These insights lay a foundation for identifying crucial genes associated with RBTs.

Haplotype analysis

Haplotype analysis of candidate genes associated with RBTs yielded two haplotype blocks. The gene Pec0223400, containing the SNP locus Hic_asm_18_15896362 ($P=6.46\times10^{-6}$ as determined by GWAS) associated with THC, includes a haplotype block consisting of four SNPs, all located in introns (Fig. 3a). The gene Pec0155450, containing the SNP locus Hic_asm_12_37973832 ($P=6.70\times10^{-6}$ as determined by GWAS) associated with HEME, forms a haplotype block comprising five SNPs, with one (12:37,970,826) in the 3'UTR and the others located in introns (Fig. 3b). Unfortunately, the SNPs within these haplotypes did not show a significant association with the traits in the GWAS.

Candidate gene validation

To deepen our understanding of the roles of candidate genes in RBTs, we conducted gene annotation and domain identification for 117 candidate genes. Based on gene annotation and primer design results, eight genes were identified for their potential relevance to RBTs and subsequently validated in an independent *A. granosa* population. Among these genes, *Septin7* is located 36 kb upstream of the SNP (Hic_asm_2_2013700), and *Hox5* is located 82 kb upstream of another SNP (Hic_asm_17_18653255); both are associated with HC. Our findings revealed significant upregulation of *Septin7* and *Hox5* in the high-HC (H-HC) group (Fig. 4a-b). *Cbfa2t3*, which harbors a SNP (Hic_asm_2_7083387) in its 3' UTR linked to THC, exhibited increased expression in the high-THC (H-THC) group (Fig. 4c). Conversely, *Avpr1b*, located 90 kb upstream of the SNP (Hic_asm_7_9279261), showed an inverse expression pattern (Fig. 4d). In studies targeting HEME, four genes were examined. *Hhex*, positioned 59 kb downstream from the SNP (Hic_asm_1_4147970), and *Rpl35a*, which harbors a SNP (Hic_asm_12_32057574) downstream, both exhibited significant upregulation in the high-HEME (H-HEME) group compared to the low-HEME (L-HEME) group. Conversely, *Eif2ak3* and *Glrk*, located 67 kb and 91 kb downstream of the SNPs Hic_asm_1_23141593 and Hic_asm_9_9773862 respectively, demonstrated increased expression levels in the L-HEME group (Fig. 4e-h).

SNP genotyping of *Cbfa2t3*

Among the eight validated candidate genes, *Cbfa2t3* was previously identified as being related to the hematopoiesis of *A. granosa* [27]. Notably, in the current study, a SNP (Hic_asm_2: 7,083,387; $P=3.40\times10^{-6}$) located in the

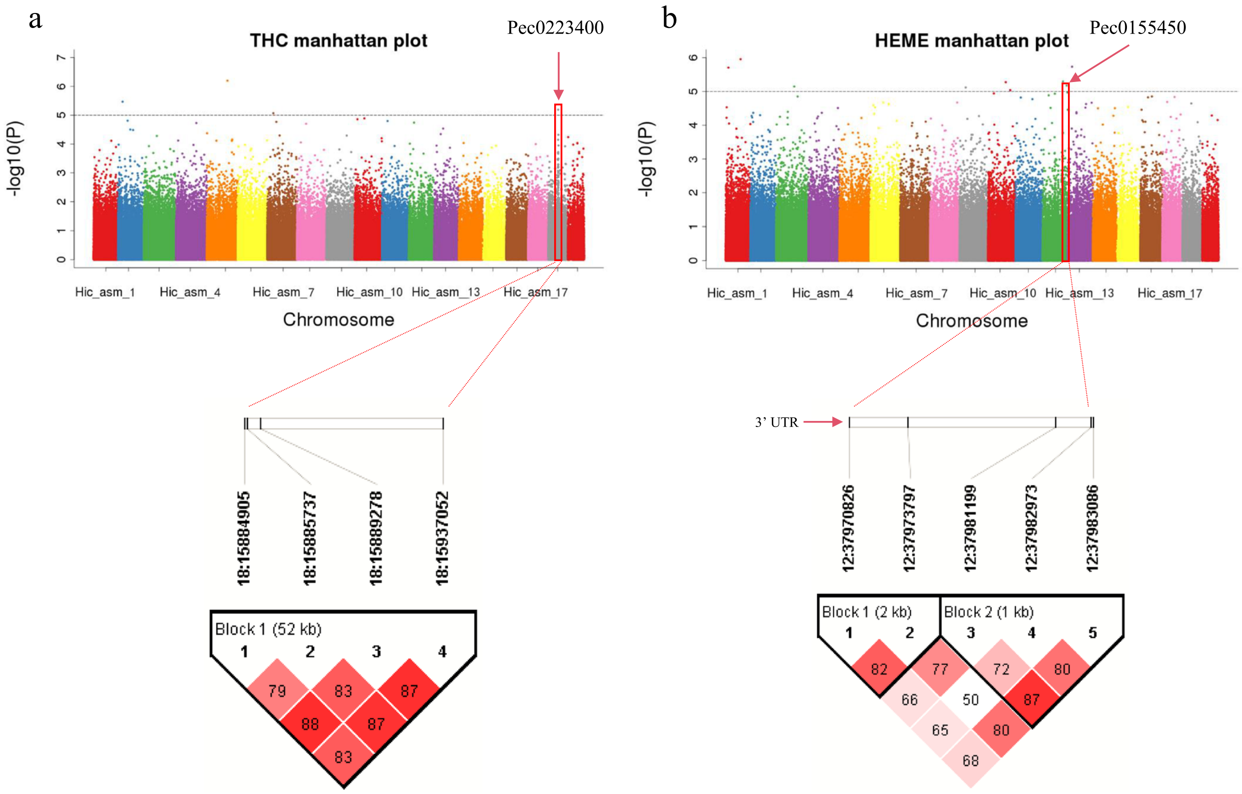

Fig. 3 Haplotype analysis of (**a**) gene Pec0223400 and **b** gene Pec0155450

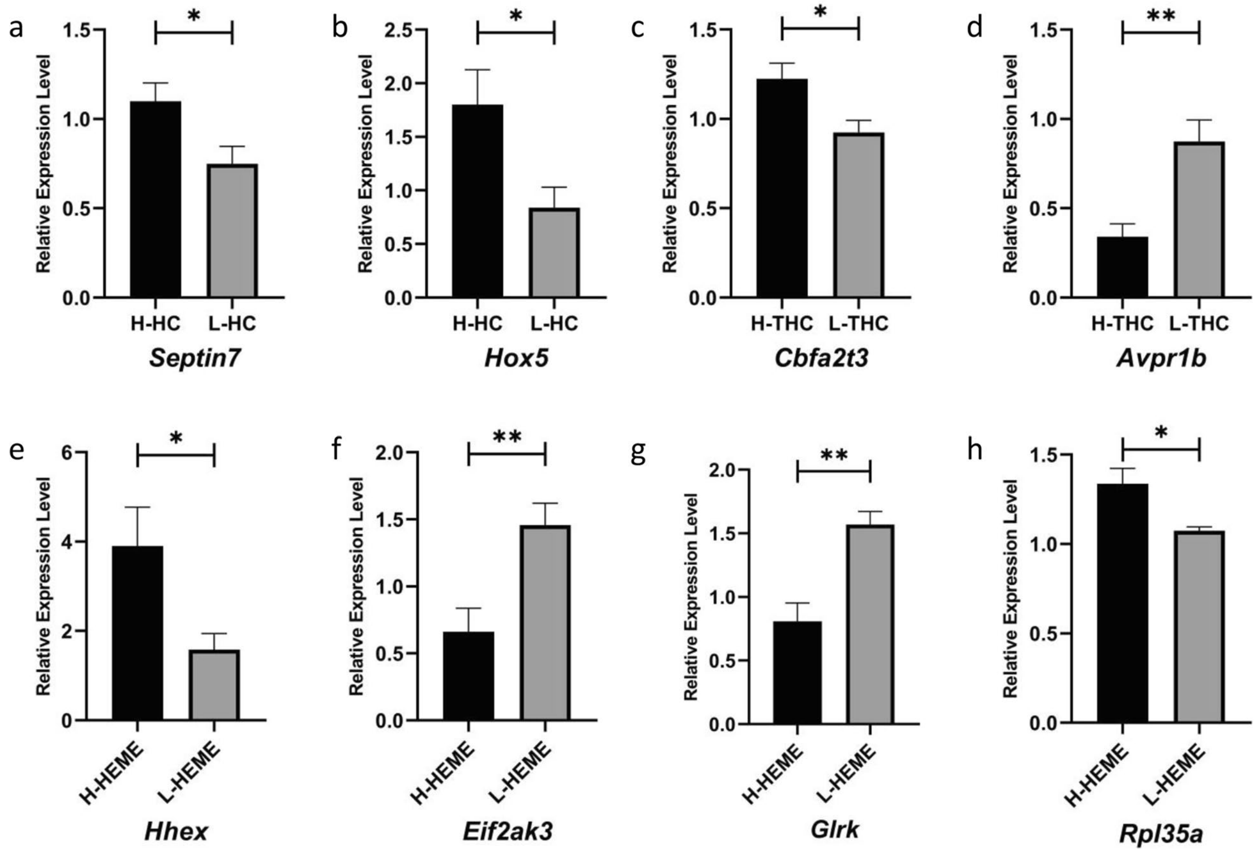

Fig. 4 Independent population validation for candidate genes **a** *Septin7*, **b** *Hox5*, **c** *Cbfa2t3*, **d** *Avpr1b*, **e** *Hhex*, **f** *Eif2ak3*, **g** *Glrk*, and **h** *Rpl35a* in *A. granosa*

3’UTR of *Cbfa2t3* (Pec0133180) was significantly associated with THC (Table S1). To further investigate the function of SNPs, an additional 37 *A. granosa* individuals were subjected to genotyping. This analysis revealed that, within the sampled population, 12 individuals possessed the heterozygous A/T genotype, while 25 were homozygous for the A/A genotype. Notably, blood cell counts in individuals with the heterozygous A/T genotype were significantly higher than those in individuals with the homozygous A/A genotype (Fig. 5). This suggests that the heterozygous A/T mutation may play a role in promoting blood cell proliferation.

Discussion

A. granosa exhibits the unusual characteristic of red blood, a feature uncommon among invertebrates. Studies on RBTs in vertebrates have been abundant; however, the molecular regulatory mechanisms of RBTs in blood clams have yet to be elucidated. The identified SNPs and candidate genes could provide a theoretical framework for exploring the molecular regulatory mechanisms of RBTs.

Currently, six hemoglobin genes have been identified, including *HbI*, *HbIIA*, *HbIIB*, *HbIII*, *HbIII_Like*, and *Hb_Like* [4]. Among these hemoglobins’ subunits, HbI, HbIIA, and HbIIB can bind heme, but the other subunits cannot [8]. The correlation analysis of RBTs in this study also showed that the relationship between HC and HEME was not completely proportional, likely due to the existence of unique hemoglobin genes in *A. granosa* that do not bind to heme. Additionally, although the expression level of myoglobin in *A. granosa* is lower than that of hemoglobin [4], it could influence heme concentration as well.

However, a weak correlation was observed between THC and both HC and HEME, contrary to the findings of previous studies [16]. This discrepancy may be associated with the mean corpuscular hemoglobin concentration (MCHC), which is defined as the ratio of Hb to THC [17]. The erythrocytes in different *A. granosa* individuals exhibit varying abilities to express hemoglobin and heme, leading to a weak correlation between these parameters. In vertebrates, including fish and humans, MCHC is considered a fundamental hematological parameter [17,

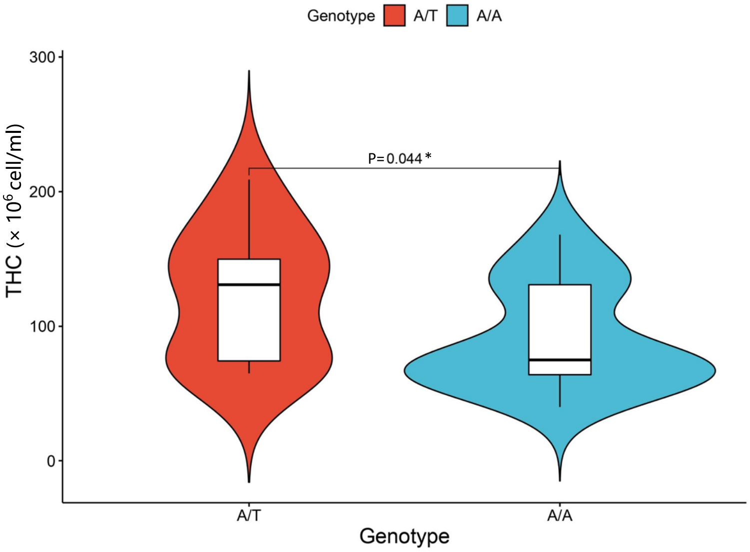

Fig. 5 THC differences among various genotypes of *Cbfa2t3* in *A. granosa*

28, 29]. This insight guides the direction of our future research endeavors, suggesting that MCHC should be considered as one of the key parameters in measuring RBTs.

Subsequently, we screened SNPs and genes related to RBTs, followed by an analysis of candidate gene haplotypes. Finally, two haplotype blocks were identified. One block, located in gene Pec0223400, was annotated as MFS-type transporter *Slc18b1*, containing both the MFS-1 domain and a membrane-spanning domain. The SLC18B1 protein is responsible for vesicular storage and release of polyamines, serving as a vesicular polyamine transporter (VPAT). It may also functionally regulate polyamine levels [30], facilitating the vesicular storage of spermine (spm) and spermidine (spd) in astrocytes, affecting glutamatergic neuronal transmission and memory formation [31]. Furthermore, spd and spm serve as potent secretagogues for histamine release from mast cells, originating from hematopoietic stem cells [32]. However, little is known about hematopoietic stem cell generation in mollusks. Our results indicated that this gene is associated with THC, suggesting that *Slc18b1* may also function as VPAT in *A. granosa*, yet its cellular effects remain unclear. Unfortunately, another gene, Pec0155450, related to HEME, was not annotated, likely due to the incomplete assembly of the chromosome or because it may represent an unknown gene. In summary, these two haplotype blocks hold potential for future applications in enhancing the RBTs performance of *A. granosa* through genetic improvement.

In this study, eight genes were selected for validation in an independent *A. granosa* population. Among them, *Septin7*, a filament-forming cytoskeletal GTPase crucial for actin cytoskeleton organization [33], interacts with Borg4 to regulate the polar distribution of *Cdc42*, *Borg4*, and *Septin7* in hematopoietic stem cells (HSCs) [34]. *Hox5*, part of the homeobox transcription factor family, plays key roles in embryonic axis development, tissue differentiation, and growth regulation [35, 36]. Vertebrate studies highlight the significance of *Hox5* in hematopoiesis. For instance, *HOXA5* plays a key role in balancing myeloid and erythroid differentiation [37], and has been suggested to influence hematopoietic lineage determination by promoting differentiation within myelopoietic lineages [38, 39]. *HOXB5* is identified as a functional marker for long-term HSCs [40], while *HOXC5* is associated with immature acute myelogenous leukemia [41, 42]. Our findings demonstrate that *Septin7* and *Hox5* both involved in the terms of genetic information processing (Table S3), were higher expressed in the H-HC group, suggesting their potential role in promoting erythrocyte proliferation and in regulating RBTs through the occurrence of HSCs in *A. granosa*.

Cbfa2t3 is a member of the myeloid translocation gene family and acts as a significant transcriptional corepressor in hematopoiesis [27, 43]. It regulates the proliferation and differentiation of erythroid progenitors by repressing the expression of *TAL1* target genes [44]. Our previous studies have identified *Cbfa2t3* in *A. granosa* as a critical gene in hematopoiesis, as evidenced by WGCNA and RNAi analyses [27]. In this study, a SNP (A/T) in the 3'UTR of *Cbfa2t3* was identified related to THC via GWAS (Table S1). Additionally, our findings show that *Cbfa2t3* expression was higher in the H-THC group, suggesting its role in hemocyte proliferation. In this study, genotype analysis showed that individuals with the A/T heterozygous genotype had significantly higher THC levels than those with the A/A homozygous genotype, indicating that A/T heterozygous individuals may experience enhanced blood cell proliferation. This suggests the SNP might be a potential regulatory site for *Cbfa2t3* in the proliferation of *A. granosa* blood cells, but the specific regulation mechanism requires further analysis. In summary, *Cbfa2t3* potentially regulates blood cell proliferation via SNP sites, impacting the RBTs of *A. granosa.*

Avpr1b encodes the receptor for arginine vasopressin (AVP), whose activity is mediated by G proteins, which activate a phosphatidylinositol-calcium second messenger system. *Avpr1b* plays a crucial role in the regulation of erythropoiesis in mammals by initiating rapid blood cell replenishment, accelerating both the proliferation and differentiation of bone marrow erythroid precursors during anemia, and the release of RBCs from the bone marrow [45]. This regulation of blood cell proliferation is consistent with our results. Elevated expression of *Avpr1b* in the L-THC group of *A. granosa* indicates that *Avpr1b* may play a role in negatively regulating hemocyte proliferation.

Four genes (*Hhex*, *Eif2ak3*, *Glrk*, and *Rpl35a*) related to heme concentration were validated in an independent population. *Hhex* encodes a homeodomain transcription factor that is widely expressed across hematopoietic stem and progenitor cell populations. It plays a role in maintaining long-term HSCs and in lineage allocation from multipotent progenitors, especially under conditions of stress hematopoiesis [46]. High expression of this gene in the H-HEME group suggests that its function in *A. granosa* is likely similar to that in vertebrates. The regulation of the translation initiation factor 2 (eIF2), critical to heme and hemoglobin synthesis, involves triggering a heme-regulated inhibitor that leads to eIF2 phosphorylation, resulting in decreased eIF2 availability and ultimately inhibiting protein synthesis [47, 48]. *Eif2ak3* encodes one of the eIF2α kinases, a metabolic-stress sensing protein kinase that phosphorylates the alpha subunit of eukaryotic translation initiation factor 2 in response to a variety of stress conditions [49]. Our findings suggest that *Eif2ak3* also plays a role in heme regulation. *Glrk* encodes a glutamate receptor that functions as a ligand-gated ion channel in the central nervous system and plays a crucial role in excitatory synaptic transmission. Research has shown that glutamate receptors are functionally linked to heme oxygenase in cerebral microvascular endothelium [50]. Our results demonstrated that the expression of *Glrk* was significantly higher in the L-HEME group than in H-HEME group, suggesting a potential regulatory role for *Glrk* in heme homeostasis. *Rpl35a* encodes the large ribosomal subunit protein eL33, an essential component of the ribonucleoprotein complex that facilitates protein synthesis within cells [51]. This protein is also essential for the proliferation and viability of hematopoietic cells [52]. In *A. granosa*, elevated *Rpl35a* expression in the H-HEME group suggests it may influence heme concentration by regulating hematopoietic cell proliferation.

In addition to previously identified genes associated with RBTs, other candidate genes may also play a role in the regulation of RBTs. For example, a gene annotated as a Toll-like receptor (*Tollo*) linked to HC (Table S2) suggests a potential relationship between RBTs and mollusc innate immunity [53]. Regarding the HEME, a candidate gene identified as Metalloproteinase inhibitor 2 (*Timp2*) encodes complexes that irreversibly inactivate metalloproteinases by binding to their catalytic zinc cofactor [54], highlighting its potential significance in the regulatory mechanisms of heme. Although *Timp2* has been associated with heme binding in myoglobin [55], its relationship with heme in hemoglobin remains unexplored. Dynein regulatory complex protein 9 (*Iqcg*), located 100 bp downstream of *Rpl35a*, interacts with calmodulin (CaM) and functions as a regulator upstream of CaM-dependent kinase IV. In the human chromosome 3, the genes *IQCG*, *RPL35A*, *PCYT1A*, and *LRCH3* span a 2-Mb genomic region, which is syntenic with the genomic locus of *Iqcg* in zebrafish on chromosome 18 [56]. The reduction in numbers of hematopoietic stem cells and multilineage-differentiated cells in iqcg-deficient embryos suggests that *Iqcg* and *Rpl35a* likely play a role in the proliferation of hematopoietic cells and heme regulation in *A. granosa* [56].

Previous studies on the origin of Hb have demonstrated that Hb evolved convergently in blood clams and vertebrates [4]. Furthermore, homologous genes involved in vertebrate hematopoiesis, such as *CBFA2T3*, *TAL1*, and *FLI1*, have been identified in *A. granosa* as factors that enhance RBTs through the promotion of hemocyte proliferation [27, 57]. Consequently, we speculate that genes related to RBTs in *A. granosa* function in a manner similar to their homologous gene in vertebrates. In this study, the majority of genes identified through GWAS, including *Slc18b1*, *Septin7*, *Hox5*, *Avpr1b*, *Hhex*, and *Rpl35a*, have been previously reported in vertebrates with red blood and are predominantly associated with hematopoiesis or hemocyte proliferation. This suggests that hematopoietic function in *A. granosa* plays a significant role in RBT variability, hinting that the regulatory mechanisms of RBTs in blood clams and vertebrates might exhibit convergent evolution. However, limitations due to incomplete gene sequences and suboptimal primer designs have impeded a more detailed analysis of these genes. Future studies will focus on the functional verification of these genes.

Conclusions

In summary, our study successfully identified 18 SNPs and 117 candidate genes associated with RBTs in *A. granosa* through GWAS, uncovering two significant haplotype blocks linked to THC and HEME, respectively. Among these, eight genes (*Septin7*, *Hox5*, *Cbfa2t3*, *Avpr1b*, *Hhex*, *Eif2ak3*, *Glrk*, and *Rpl35a*), validated within an independent *A. granosa* population, have been implicated in the regulation of RBTs. Notably, a SNP located in the 3'UTR of *Cbfa2t3* was found to potentially promote blood cell proliferation. Our findings indicate that the hematopoietic function in *A. granosa* plays a pivotal role in the variability of RBTs. The results of this study enable a detailed analysis of the correlation between gene variation and genetic mechanisms related to RBTs in *A. granosa*, potentially offering deeper insights into the evolutionary mechanisms of RBT. The significant SNPs and candidate genes identified herein provide a wealth of genetic resources and lay a solid foundation for future functional research and the molecular breeding of *A. granosa*. Our findings suggest a conservation of the regulatory mechanisms of RBT between blood clams and vertebrates, aligning with the evolutionary conservation of hemoglobin. This shared regulatory framework illuminates the fundamental principles of RBT regulation across the vast evolutionary divide between invertebrates and vertebrates and provides a scientific basis for selective breeding in blood clams.

Methods

Sample collection and phenotypic measurement

The population of *A. granosa* utilized for GWAS was constructed in a preceding investigation [26]. A total of 300 two-year-old individuals, collected from Ninghai, Zhejiang, were used for resequencing. The HC of each *A. granosa* individual was measured using the hemoglobin assay kit (Real-Tech Biological Technology, Beijing, China). A standard curve was initially established using cyanogenic methemoglobin at five concentrations (0, 25, 50, 75, and 100 g/L). Absorbance was measured at 540 nm in triplicate using a UV–Vis spectrophotometer (Cary 3500, Agilent, USA). Then, 1 mL of HC determination reagent was employed to calibrate the spectrophotometer. Afterwards, 5 µL of hemolymph from each individual was mixed with 1 mL of this reagent. After reacting for 1 min at room temperature, the HC was determined by measuring the absorbance at 540 nm. THC (cell/mL) was calculated by a Neubauer hemocytometer at 100× magnification. 10 µL of hemolymph from each blood clam was added to 1 mL PBS, and the resulting mixture was placed on a blood cell counting plate to estimate the THC using a microscope. The HEME measurement was conducted using the Heme Assay Kit (Sigma-Aldrich, USA), with product information provided below. The total heme concentration of a sample can be determined by the following equation: (OD of sample − OD of blank) × (OD of calibrator − OD of blank)$^{-1}$ × 62.5 × (Dilution Factor) µM.

Phenotypic statistics

The analysis of these RBTs was performed using IBM SPSS Statistics 20. Phenotypic correlation analysis, utilizing the Pearson correlation coefficient, was conducted with the "Corrplot" package in R v3.6.3 (R Core Team). The strength of the Pearson correlation coefficient was interpreted according to a commonly accepted definition: 'very weak' for values between 0.00 and 0.19, 'weak' for values between 0.20 and 0.39, 'moderate' for values between 0.40 and 0.59, 'strong' for values between 0.60 and 0.79, and 'very strong' for values between 0.80 and 1.0 [58]. Results of the phenotypic analysis are presented in Table 1. Genetic correlation analysis was performed using GCTA software (http://cnsgenomics.com/software/gcta/).

Genome-wide association study (GWAS)

The reference genome of *A. granosa* (NCBI accession number: JABXWC000000000) was applied in this study. The methods and data pertaining to sampling, genome sequencing, and SNP calling were executed in accordance with our previously published work [26]. A total of 3114 Gb of high-quality sequencing data was obtained, with high sequencing quality ($Q30 \geq 89.27\%$) and a normal GC distribution [26]. The average mapping rate achieved was 91.93%, with an average sequencing depth of around 13×. SNP calling was executed using SAMtools, applying filtering criteria of dp4 coverage depth, $MISS < 0.3$, and $MAF > 0.01$. The identified SNPs were then annotated using ANNOVAR, yielding a total of 355,254 high-quality SNPs [26]. GWAS correlation analysis, kinship, and principal component analysis (PCA) were conducted using GEMMA software (http://www.xzlab.org/software.html) using the compressed mixed linear model (MLM), as described in our previous work [26]. The PCA results indicated that most individuals belong to a single population [26]. The Manhattan and QQ plots were performed by plot function of R (version 3.6.3). The raw sequence data were deposited in the NCBI Sequence Read Archive under accession number PRJNA988240.

To mitigate false positive associations, we exclusively chose SNPs in the GWAS with a minor allele frequency

(MAF) greater than 0.01 and a missing rate less than 0.3 within the population. Considering that the Bonferroni test threshold (0.05/N) is too strict, we established the SNP GWAS threshold at a P-value of less than 10^{-5}, and the r^2 value exceeded 0.02 beyond a 3 kb range within the populations [26]. To broaden the search for potential red blood candidate genes, we considered a 100 kb range upstream or downstream from significant SNPs within a scaffold, consistent with previous studies [59–61]. The identified candidate SNPs and genes are listed in Tables S1 and S2.

GO and KEGG enrichment and haplotype analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were conducted and visualized using TBtools (version 1.108) [62]. In the *A. granosa* genome, 24,398 protein-coding genes were identified [4]. The proportion of genes annotated in GO was 47.8% (11,667 out of 24,398), and in KEGG, it was 39.6% (9,668 out of 24,398). The enrichment analyses for GO and KEGG were conducted using TBtools (version 1.108), with the enrichment backend provided by TBtools. The Benjamini/Hochberg method (BH method) was applied for P-value correction [63]. The enrichment results are presented in Tables S3 and S4, with the top 10 terms displayed for categories containing more than 10 terms.

Haplotype analysis

We utilized Bcftools (version 1.9) to pinpoint SNP sites within candidate gene regions [64], while Haploview (version 4.2) was employed for haplotype analysis and visualization of haploblocks [65]. For haplotype analysis, the Hardy–Weinberg P-value cutoff for the haplotype blocks was set at 0.001, with a minimum minor allele frequency of 0.01, a minimum genotype call rate of 75%, and a maximum of one Mendelian inheritance error. A threshold of 0.7 was used to divide the haplotype blocks.

Independent population validation

The HC, THC, and HEME of another independent *A. granosa* population from Ninghai, Zhejiang, China, were measured. From this population, individuals exhibiting the top 5% in terms of HC, THC, and HEME were categorized into a high (H) group, whereas those in the bottom 5% were classified into a low (L) group (Fig. S4). After synthesizing the physiological states of the individuals, 16 *A. granosa* specimens for each trait were respectively selected for the quantitative verification of candidate genes.

Gills of *A. granosa* were cut for RNA extraction using TRIzol. RNA quality was assessed using 1.0% agarose gel electrophoresis, and RNA concentration was quantified with UV spectrophotometry using a Nanodrop 2000 spectrophotometer. Primers for the genes of interest were designed with Primer Premier 5 (version 5.0) [66]. Reverse transcription used HiScript III RT SuperMix for qPCR (Vazyme, Nanjing, China), following the manufacturer’s protocol, with 1 μg of total RNA. Synthesized cDNA was diluted 20-fold for real-time PCR analysis. ChamQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing, China) was used for qPCR. The 18S ribosomal RNA gene (18S) was the internal reference, and the $2^{-\Delta\Delta Ct}$ method was used to determine relative gene expression levels, following prior studies [16]. Figures were generated using GraphPad Prism 8. Detailed information about the candidate genes and their primers is available in Table S5.

SNP genotyping analysis

Thirty-seven healthy *A. granosa* individuals from a common population were selected for SNP genotyping, using the blood cell counting method described in Sect. 2.1 to ascertain the number of blood cells. To perform Sanger sequencing on sequences upstream and downstream of the target SNP (Hic_asm_2:7,083,387) in *A. granosa*, specific primers Cbfa2t1-F and Cbfa2t1-R were used. The primer sequences are: Cbfa2t3-F: 5’-ATGTGGACAAGTTGGTCTTTGATAC-3’ and Cbfa2t3-R: 5’-GTCCAACTAATTCTGTGGCATCTAC-3’.

Abbreviations

RBT	Red blood trait
HC	Hemoglobin concentration
THC	Total hemocyte count
HEME	Heme concentration
GWAS	Genome-wide association studies
SNP	Single nucleotide polymorphism

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10857-3.

Additional file 1: Supplementary Figures.

Additional file 2: Supplementary Tables.

Acknowledgements

We express our sincere gratitude to all members of the laboratory for their invaluable technical advice and enriching discussions.

Authors’ contributions

XH designed the research, analyzed the data, and wrote the manuscript. YL conducted validation experiments. GY collected experimental materials. SW and YB critically revised the manuscript. All authors have read and approved the final version of the manuscript.

Funding

This work was financially supported by the National Science Foundation of China (32273123), the Key Natural Science Foundation of Ningbo (2023J042), the Zhejiang Major Program of Science and Technology (2021C02069-7), and the Ningbo Public Benefit Research Key Project (2021S014).

Availability of data and materials

The data supporting this article are included within the article itself and in its Additional files. The sequencing data files are available in the NCBI Sequence Read Archive (BioProject: PRJNA988240) and can be accessed using accession numbers SRR25343508-SRR25343807 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA988240).

Declarations

Ethics approval and consent to participate

Ethical approval for this study was obtained from the Experimental Animal Ethics Committee of Zhejiang Wanli University. All the experimental procedures were approved by the Experimental Animal Ethics Committee of Zhejiang Wanli University.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 16 May 2024 Accepted: 1 October 2024

References

1. Astorga MP. Genetic considerations for mollusk production in aquaculture: current state of knowledge. Front Genet. 2014;5:435.
2. Syahira N, Nithiyaa N, Nooraini I, Aileen Tan SH. Preliminary study on the growth development of blood cockle (Tegillarca granosa) by using different substrates in the hatchery system. J Survey Fisheries Sci. 2021;7(2):71–8. https://doi.org/10.17762/sfs.v7i2.121.
3. Terwilliger NB. Functional adaptations of oxygen-transport proteins. J Exp Biol. 1998;201(8):1085–98.
4. Bao Y, Zeng Q, Wang J, Zhang Z, Zhang Y, Wang S, Wong NK, Yuan W, Huang Y, Zhang W, et al. Genomic insights into the origin and evolution of molluscan red-bloodedness in the blood clam Tegillarca granosa. Mol Biol Evol. 2021;38(6):2351–65. https://doi.org/10.1093/molbev/msab030.
5. Bao Y, Wang Q, Lin Z. Hemoglobin of the bloody clam Tegillarca granosa (Tg-HbI) is involved in the immune response against bacterial infection. Fish Shellfish Immunol. 2011;31(4):517–23.
6. Kim JH, Lee HM, Cho YG, Shin JS, You JW, Choi KS, Hong HK. Flow cytometric characterization of the hemocytes of blood cockles Anadara broughtonii (Schrenck, 1867), Anadara kagoshimensis (Lischke, 1869), and Tegillarca granosa (Linnaeus, 1758) as a biomarker for coastal environmental monitoring. Mar Pollut Bull. 2020;160: 111654.
7. Glenn A, Armstrong CE. Physiology of red and white blood cells. Anaesthes Intens Care Med. 2019;20(3):170–4.
8. Yuan W, Liu H, Zhang W, Lin Z, Li C, Bao Y. Identificaiton and characterization of a novel hemoglobin gene (Tgr-HbIII) from blood clam Tegillarca granosa. Gene. 2023;862: 147256.
9. Yang S, Dong Y, Aweya JJ, Xie T, Zeng B, Zhang Y, Liu GM. Antimicrobial activity and acting mechanism of Tegillarca granosa hemoglobin-derived peptide (TGH1) against Vibrio parahaemolyticus. Microb Pathog. 2020;147: 104302.
10. Bao Y, Li P, Dong Y, Xiang R, Gu L, Yao H, Wang Q, Lin Z. Polymorphism of the multiple hemoglobins in blood clam Tegillarca granosa and its association with disease resistance to Vibrio parahaemolyticus. Fish Shellfish Immunol. 2013;34(5):1320–4.
11. Wang S, Yu X, Lin Z, Zhang S, Xue L, Xue Q, Bao Y. Hemoglobins likely function as peroxidase in blood clam tegillarca granosa hemocytes. J Immunol Res. 2017;2017:7125084.
12. de la Ballina NR, Maresca F, Cao A, Villalba A. Bivalve Haemocyte Subpopulations: a review. Front Immunol. 2022;13: 826255.
13. Taniguchi CN, Dobbs J, Dunn MA. Heme iron, non-heme iron, and mineral content of blood clams (Anadara spp.) compared to Manila clams (V. philippinarum), Pacific oysters (C. gigas), and beef liver (B. taurus). J Food Compos Anal. 2017;57:49–55.
14. Li G, Zhan J, Xu A, Tan B, Sun N, Wang C, Jia R, Li C, Zhang J, Yang W. Determination of the iron bioavailability, conformation, and rheology of iron-binding proteins from Tegillarca granosa. J Food Biochem. 2021;45(1): e13517.
15. Sun B, Zhang P, Wei H, Jia R, Huang T, Li C, Yang W. Effect of hemoglobin extracted from Tegillarca granosa on iron deficiency anemia in mice. Food Res Int. 2022;162:112031.
16. Yang Z, He X, Jin H, Su D, Lin Z, Liu H, Bao Y. Hemocyte proliferation is associated with blood color shade variation in the blood clam Tegillarca granosa. Aquaculture. 2023;571:739447.
17. Ganesh SK, Zakai NA, van Rooij FJ, Soranzo N, Smith AV, Nalls MA, Chen MH, Kottgen A, Glazer NL, Dehghan A, et al. Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. Nat Genet. 2009;41(11):1191–8.
18. Korte A, Farlow A. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods. 2013;9(1): 29.
19. Ning X, Li X, Wang J, Zhang X, Kong L, Meng D, Wang H, Li Y, Zhang L, Wang S, et al. Genome-wide association study reveals E2F3 as the candidate gene for scallop growth. Aquaculture. 2019;511:734216.
20. Andrews NC. Genes determining blood cell traits. Nat Genet. 2009;41(11):1161–2.
21. Chambers JC, Zhang W, Li Y, Sehmi J, Wass MN, Zabaneh D, Hoggart C, Bayele H, McCarthy MI, Peltonen L, et al. Genome-wide association study identifies variants in TMPRSS6 associated with hemoglobin levels. Nat Genet. 2009;41(11):1170–2.
22. Soranzo N, Spector TD, Mangino M, Kuhnel B, Rendon A, Teumer A, Willenborg C, Wright B, Chen L, Li M, et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. Nat Genet. 2009;41(11):1182–90.
23. Sorensen E, Rigas AS, Didriksen M, Burgdorf KS, Thorner LW, Pedersen OB, Hjalgrim H, Petersen MS, Erikstrup C, Ullum H. Genetic factors influencing hemoglobin levels in 15,567 blood donors: results from the Danish Blood Donor Study. Transfusion. 2019;59(1):226–31.
24. Uda M, Galanello R, Sanna S, Lettre G, Sankaran VG, Chen W, Usala G, Busonero F, Maschio A, Albai G. Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of β-thalassemia. Proc Natl Acad Sci. 2008;105(5):1620–5.
25. Danjou F, Zoledziewska M, Sidore C, Steri M, Busonero F, Maschio A, Mulas A, Perseu L, Barella S, Porcu E, et al. Genome-wide association analyses based on whole-genome sequencing in Sardinia provide insights into regulation of hemoglobin levels. Nat Genet. 2015;47(11):1264–71.
26. He X, Liao Y, Yang Z, Liu H, Wang S, Bao Y. A genome-wide association study to identify the genes associated with growth-related traits in Tegillarca granosa. Aquaculture. 2024;578: 740127.
27. Jin H, Liu H, Wang J, Zhang W, Bao Y. CBFA2T3 affects red blood phenotype in the blood clam Tegillarca granosa by regulating hemocyte proliferation. Aquaculture. 2024;581:740462.
28. Arnaudov A, Arnaudova D. Erythrocytes and hemoglobin of fish: potential indicators of ecological biomonitoring. In: Animal Models and Experimental Research in Medicine. IntechOpen; 2022. https://doi.org/10.5772/intechopen.107053.
29. Witeska M, Kondera E, Bojarski B. Hematological and hematopoietic analysis in fish toxicology—A review. Animals. 2023;13(16): 2625.
30. Fredriksson R, Sreedharan S, Nordenankar K, Alsio J, Lindberg FA, Hutchinson A, Eriksson A, Roshanbin S, Ciuculete DM, Klockars A, et al. The polyamine transporter Slc18b1(VPAT) is important for both short and long time memory and for regulation of polyamine content in the brain. PLoS Genet. 2019;15(12):e1008455.
31. Hiasa M, Miyaji T, Haruna Y, Takeuchi T, Harada Y, Moriyama S, Yamamoto A, Omote H, Moriyama Y. Identification of a mammalian vesicular polyamine transporter. Sci Rep. 2014;4:6836.
32. Moriyama Y, Hatano R, Moriyama S, Uehara S. Vesicular polyamine transporter as a novel player in amine-mediated chemical transmission. Biochim Biophys Acta Biomembr. 2020;1862(12): 183208.
33. Kremer BE, Adang LA, Macara IG. Septins regulate actin organization and cell-cycle arrest through nuclear accumulation of NCK mediated by SOCS7. Cell. 2007;130(5):837–50.
34. Kandi R, Senger K, Grigoryan A, Soller K, Sakk V, Schuster T, Eiwen K, Menon MB, Gaestel M, Zheng Y, et al. Cdc42-Borg4-Septin7 axis regulates HSC polarity and function. EMBO Rep. 2021;22(12):e52931.

35. Mallo M. Reassessing the role of Hox genes during vertebrate development and evolution. Trends Genet. 2018;34(3):209–17.
36. Rezsohazy R, Saurin AJ, Maurel-Zaffran C, Graba Y. Cellular and molecular insights into Hox protein action. Development. 2015;142(7):1212–27.
37. Bhatlekar S, Fields JZ, Boman BM. Role of HOX genes in stem cell differentiation and cancer. Stem Cells Int. 2018;2018:3569493.
38. Steens J, Klein D. HOX genes in stem cells: Maintaining cellular identity and regulation of differentiation. Front Cell Dev Biol. 2022;10:1002909.
39. Argiropoulos B, Humphries R. Hox genes in hematopoiesis and leukemogenesis. Oncogene. 2007;26(47):6766–76.
40. Chen JY, Miyanishi M, Wang SK, Yamazaki S, Sinha R, Kao KS, Seita J, Sahoo D, Nakauchi H, Weissman IL. Hoxb5 marks long-term haematopoietic stem cells and reveals a homogenous perivascular niche. Nature. 2016;530(7589):223–7.
41. Van Oostveen J, Bijl J, Raaphorst F, Walboomers J, Meijer C. The role of homeobox genes in normal hematopoiesis and hematological malignancies. Leukemia. 1999;13(11):1675–90.
42. Bijl J, Van Oostveen J, Walboomers J, Brink A, Vos W, Ossenkoppele G, Meijer C. Differentiation and cell-type-restricted expression of HOXC4, HOXC5 and HOXC6 in myeloid leukemias and normal myeloid cells. Leukemia. 1998;12(11):1724–32.
43. Ajore R, Kumar P, Dhanda RS, Gullberg U, Olsson I. The leukemia associated nuclear corepressor ETO homologue genes MTG16 and MTGR1 are regulated differently in hematopoietic cells. BMC Mol Biol. 2012;13:1–16.
44. Goardon N, Lambert JA, Rodriguez P, Nissaire P, Herblot S, Thibault P, Dumenil D, Strouboulis J, Romeo PH, Hoang T. ETO2 coordinates cellular proliferation and differentiation during erythropoiesis. EMBO J. 2006;25(2):357–66.
45. Mayer B, Németh K, Krepuska M, Myneni VD, Maric D, Tisdale JF, Hsieh MM, Uchida N, Lee H-J, Nemeth MJ. Vasopressin stimulates the proliferation and differentiation of red blood cell precursors and improves recovery from anemia. Sci Trans Med. 2017;9(418):eaao1632.
46. Goodings C, Smith E, Mathias E, Elliott N, Cleveland SM, Tripathi RM, Layer JH, Chen X, Guo Y, Shyr Y, et al. Hhex is Required at Multiple Stages of Adult Hematopoietic Stem and Progenitor Cell Differentiation. Stem Cells. 2015;33(8):2628–41.
47. Paolini NA, Moore KS, di Summa FM, Fokkema I. t Hoen PAC, von Lindern M: Ribosome profiling uncovers selective mRNA translation associated with eIF2 phosphorylation in erythroid progenitors. PLoS ONE. 2018;13(4):e0193790.
48. Chen J-J, Zhang S. Heme-regulated eIF2α kinase in erythropoiesis and hemoglobinopathies. Blood J Am Soc Hematol. 2019;134(20):1697–707.
49. Harding HP, Zhang Y, Ron D. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. Nature. 1999;397(6716):271–4.
50. Parfenova H, Fedinec A, Leffler CW. Ionotropic glutamate receptors in cerebral microvascular endothelium are functionally linked to heme oxygenase. J Cereb Blood Flow Metab. 2003;23(2):190–7.
51. Li H, Huo Y, He X, Yao L, Zhang H, Cui Y, Xiao H, Xie W, Zhang D, Wang Y. A male germ-cell-specific ribosome controls male fertility. Nature. 2022;612(7941):725–31.
52. Farrar JE, Nater M, Caywood E, McDevitt MA, Kowalski J, Takemoto CM, Talbot CC Jr, Meltzer P, Esposito D, Beggs AH, et al. Abnormalities of the large ribosomal subunit protein, Rpl35a diamond-blackfan anemia. Blood. 2008;112(5):1582–92.
53. Liu S, Shi W, Guo C, Zhao X, Han Y, Peng C, Chai X, Liu G. Ocean acidification weakens the immune response of blood clam through hampering the NF-kappa β and toll-like receptor pathways. Fish Shellfish Immunol. 2016;54:322–7.
54. Stetler-Stevenson WG, Krutzsch HC, Liotta LA. Tissue inhibitor of metalloproteinase (TIMP-2): a new member of the metalloproteinase inhibitor family. J Biol Chem. 1989;264(29):17374–8.
55. Kai HS-T, Butler GS, Morrison CJ, King AE, Pelman GR, Overall CM. Utilization of a novel recombinant myoglobin fusion protein expression system to characterize the Tissue Inhibitor of Metalloproteinase (TIMP)-4 and TIMP-2 C-terminal domain and tails by Mutagenesis: the importance of acidic residues in binding the MMP-2 hemopexin C domain. J Biol Chem. 2002;277(50):48696–707.
56. Chen LT, Liang WX, Chen S, Li RK, Tan JL, Xu PF, Luo LF, Wang L, Yu SH, Meng G, et al. Functional and molecular features of the calmodulin-interacting protein IQCG required for haematopoiesis in zebrafish. Nat Commun. 2014;5:3811.
57. Jin H, Liu H, Wang J, Zhang W, Bao Y. TAL1-mediated regulation of hemocyte proliferation influences red blood phenotype in the blood clam Tegillarca granosa. Aquaculture. 2024;586:740801.
58. Joshi A, Suragimath G, Zope SA, Ashwinirani S, Varma SA. Comparison of gingival biotype between different genders based on measurement of dentopapillary complex. J Clin Diagn Res: JCDR. 2017;11(9):ZC40.
59. Meng J, Song K, Li C, Liu S, Shi R, Li B, Wang T, Li A, Que H, Li L, et al. Genome-wide association analysis of nutrient traits in the oyster Crassostrea gigas: genetic effect and interaction network. BMC Genomics. 2019;20(1):625.
60. Shi R, Li C, Qi H, Liu S, Wang W, Li L, Zhang G. Construction of a high-resolution genetic map of Crassostrea gigas: QTL mapping and GWAS applications revealed candidate genes controlling nutritional traits. Aquaculture. 2020;527:735427.
61. He X, Li C, Qi H, Meng J, Wang W, Wu F, Li L, Zhang G. A genome-wide association study to identify the genes associated with shell growth and shape-related traits in Crassostrea gigas. Aquaculture. 2021;543:736926.
62. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13(8):1194–202.
63. Thissen D, Steinberg L, Kuang D. Quick and easy implementation of the Benjamini-Hochberg procedure for controlling the false positive rate in multiple comparisons. J Educ Behav Statist. 2002;27(1):77–83.
64. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA. Davies RM: twelve years of SAMtools and BCFtools. Gigascience. 2021;10(2):giab008.
65. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005;21(2):263–5.
66. Lalitha S. Primer premier 5. Biotech Software Int Rep. 2000;1(6):270–2.

Publisher's Note