Mar Biotechnol (2015) 17:364–374
DOI 10.1007/s10126-015-9625-0

ORIGINAL ARTICLE

Transcriptome Analysis of Shell Color-Related Genes in the Clam *Meretrix meretrix*

Xin Yue · Qing Nie · Guoqiang Xiao · Baozhong Liu

Received: 26 October 2014 / Accepted: 19 January 2015 / Published online: 14 February 2015

Abstract Color polymorphism has received much attention due to its strong implications for speciation and adaptation. In contrast to body color, little is currently known about the molecular mechanism of shell color formation. This study represents the first analysis of the relationship between whole-scale gene expression and shell color variations in the marine bivalve mollusks via comparative transcriptome analyses. Three clam *Meretrix meretrix* strains with different and monotonous shell color patterns, which were developed by our 10-year artificial selection, combined with clams with nearly white shell color were used in the analyses. The results supported the idea that there was a relationship between gene expression and shell pigmentation in the clam *M. meretrix*, and complex signal transduction were involved. It was proposed that Notch signaling pathway played a crucial role in shell pigmentation in a gene-dosage dependent pattern and also potentially involved in the shell color patterning. Calcium signaling process may equally be implicated in shell color formation via activation of Notch pathway. Other differentially expressed genes (e.g., *Myl*, *Mitf*) potentially implicated in shell color pigmentation were also noticed. This study provides information on the expression profiles of clams with different shell color morphs and sheds light on color formation mechanism of shell.

Keywords Shell color · *Meretrix meretrix* · Comparative transcriptome · Differentially expressed genes

Introduction

Color polymorphism is an interesting phenomenon which could be found commonly throughout the animal kingdom, with functions in camouflage (Barbato et al. 2007), thermoregulation (Heath 1975), mating selection (Houde and Endler 1990), social interactions (Rodgers et al. 2010), desiccation resistance (Parkash et al. 2009), salinity adaptation (Sokolova and Berger 2000), immunity (Scheil et al. 2013), and so on. As one of the most diverse phenotypical traits in nature and one of the driving forces of speciation, coloration, and color pattern formation attract many biologists involved in the study (Bagnara and Matsumoto 2007). And visible coloration polymorphisms provide a tractable system within which to examine the molecular basis of adaptation and evolution because of their often-simple patterns of inheritance and the general ease with which morph/allele frequencies can be estimated (Protas and Patel 2008).

In mollusks, shell color variations are known to exist in many species (Adamkewicz and Castagna 1988; Cain 1988; Kobayashi et al. 2004; Liu et al. 2009; Mitton 1977). These amazing shell color patterns aroused the questions on the mechanism underlying this intricate phenotype, e.g., what is the molecular basis of coloration patterns? How many genes are involved in orchestrating differences in shell coloration between individuals within a population? As one of the marine bivalve molluscan species, clam *Meretrix meretrix* exhibits highly variable shell color morphs based on the distribution of overlying pigments, which is a suitable species for investigating shell color variation. The methods used to study coloration include genetic crosses, gene expression analyses, complementation studies, and linkage

Electronic supplementary material The online version of this article (doi:10.1007/s10126-015-9625-0) contains supplementary material, which is available to authorized users.

X. Yue · Q. Nie · B. Liu (✉)
Key Laboratory of Experimental Marine Biology,
Institute of Oceanology, Chinese Academy of Sciences,
7 Nanhai Road, Qingdao 266071, China
e-mail: bzliu@qdio.ac.cn

G. Xiao
Zhejiang Mariculture Research Institute, Wenzhou 325005, China

mapping (Sivka et al. 2013; Protas and Patel 2008). To the bivalve mollusks, the existing studies for the shell coloration were few and mainly concerned the genetic basis analysis through selected crosses (Adamkewicz and Castagna 1988; Innes 1977; Winkler et al. 2001). And the genes and pathways underlying shell color formation of the marine bivalve are poorly understood. Considering shell color variation is likely to stem from differences in gene expression, gene expression analysis is a priority for the nonmodel species with rare genetic resources. Next-generation sequencing technologies permit rapid profiling the genes globally and functionally expressed, which has been applied to the coloration studies in the mammal, fish, and insect (Croucher et al. 2013; Fan et al. 2013; Henning et al. 2013; Nie et al. 2014; Wang et al. 2014).

In this study, we employed the powerful tool of high-throughput RNA sequencing and digital gene expression (DGE) analysis in identifying the genes potentially involved in shell color determination of the clam *M. meretrix*, by comparing the gene expression profiles in the mantle of the clams with different shell color morphs because the pigment is deposited by cells of the mantle. In the wild, the coloration patterns of *M. meretrix* consist of discrete (e.g., blotching), alternating (e.g., zigzag), radial (e.g., V-shaped), and so on. The shell color morph of one randomly picked clam individual may exhibit elaborate shell color-combined different patterns. In order to eliminate the influence of mixed shell color patterns, we spent 10 years developing three strains of *M. meretrix* with different and monotonous shell color morphs by artificial selection, and the expression analyses were mainly based on these strains. This study represents the first analysis of the relationship between whole-scale gene expression and shell color variations in the marine bivalve mollusks and would provide primary information for the future study about the molecular basis of the shell color pattern formation.

Materials and Methods

Clams and Sample Collection

Three clam strains separately exhibiting steady and hereditary shell color morph of black blotching (strain SB), purple zigzag (strain SP), or whole maroon (strain WM) were developed by 10-year successive generations selection. Briefly, clams with specific shell color morph (SB, SP, or WM) were collected from the progeny of the wild population. The morph was steady after more than four generations intra-morph artificial breeding during which the unrepresentative individuals were picked out. In this research, clams separately with four kinds of different and monotonous shell color morphs (*M. meretrix*) (illustrated in Fig. 1) were applied. Three kinds were collected from the three clam strains (strain SB, SP, and WM), respectively. And the other kind, which was nearly white shell color morph (W), was found and collected in the population who were cultured in the same estuarine area (Lingkun County, Wenzhou, Zhejiang Province). All collected clams (adult, 2-year-old) were acclimated in the water (18–20 °C, 25‰ salinity) under continuous aeration and fed with *Isochrysis galbana* daily for 2 weeks. Then for each of the four kinds, mantle from five clams were dissected and immediately ground into a powder mixture in liquid nitrogen and then reserved at −80 °C before processing for RNA extraction.

RNA Extraction and Quality Control

Total RNA was extracted separately from the five mantle powder mixtures of WM, SB, SP, or W using the SV Total RNA Isolation System (Promega, USA) according to the manufacturer's instructions. RNA degradation and contamination was monitored on 1 % agarose gels. RNA purity was checked using the NanoPhotometer spectrophotometer (Implen, GER). RNA concentration was measured using Qubit RNA Assay Kit in Qubit 2.0 Fluorometer (Life Technologies, CA, USA). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA) and expressed as RNA Integrity Number (RIN). A total amount of 3 μg qualified RNA (RIN>8) was used. RNA of four shell color morphs (WM, SB, SP, and W) was applied to library preparation for DGE sequencing, respectively, and a pooled RNA including SP, SB, WM, and W were applied to library preparation for de novo transcriptome sequencing.

Library Preparation, Sequencing, and Quality Control

Library was generated using Illumina TruSeq™ RNA Sample Preparation Kit (Illumina, USA) following manufacturer's recommendations. Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature in Illumina proprietary fragmentation buffer. First-strand cDNA was synthesized using random oligonucleotides and SuperScript II. Second-strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities and enzymes were removed. After adenylation of 3′ ends of DNA fragments, Illumina PE adapter oligonucleotides were ligated to prepare for hybridization. The library fragments were purified with AMPure XP system (Beckman Coulter, USA) to select cDNA fragments of preferentially 200 bp in length. DNA fragments with ligated adaptor molecules on both ends were selectively enriched using Illumina PCR Primer Cocktail in a 10-cycle PCR reaction. Products were purified using AMPure XP system (Agencourt, USA) and quantified using the Agilent high sensitivity DNA assay on the Agilent Bioanalyzer 2100 system.

Fig. 1 Clams (*M. meretrix*) separately with four kinds of shell color morphs. *SP*, clams with purple zigzag shell color; *SB*, clams with black blotching shell color; *WM*, clams with whole maroon shell color; *W*, clams with nearly whole white shell color

Clustering was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina, USA) according to the manufacturer's instructions. After cluster generation, the libraries were sequenced on an Illumina Hiseq 2000 platform. And 100-bp single-end reads were generated for libraries prepared for DGE, 100-bp paired-end reads were generated for libraries prepared for de novo transcriptome. The left-end reads of de novo transcriptome generated were pooled into one big left.fastq file named Mantle_1, while the right-end reads were pooled into one big right.fastq file named Mantle_2.

Raw data were obtained by base calling (all raw tag data have been deposited in Short Read Archive (SRA) of the National Center for Biotechnology Information (NCBI)) and then were filtered through in-house perl scripts, by which clean data were obtained by removing reads containing adapter, reads containing poly-N, and low-quality reads from raw data. At the same time, Q20, Q30, GC-content, and sequence duplication level of the clean data were calculated. All the downstream analyses were based on clean data with high quality.

De novo Transcriptome Assembly and Gene Function Annotation

Transcriptome assembly was accomplished using Trinity (Grabherr et al. 2011), by which transcripts and unigenes (the longest transcript of a set of transcripts that appear to stem from the same transcription locus) were obtained. Gene function was annotated based on the following databases with a cutoff *E* value of 1.0×10^{-5}: Nr (NCBI non-redundant protein sequences); Nt (NCBI non-redundant nucleotide sequences); Pfam (Protein family); COG (Clusters of Orthologous Groups of proteins); Swiss-Prot (A manually annotated and reviewed protein sequence database); KO (KEGG Orthology database); GO (Gene Ontology).

Sequence Mapping and Gene Expression Quantification

The assembled de novo transcriptome was used as the reference database, and gene expression levels were estimated for each sample. Briefly, clean data of DGE were mapped back to the reference transcriptome by Bowtie v0.12.9, and read count for each gene was obtained from the mapping results by RSEM (Li and Dewey 2011). And then RPKM (Reads Per Kilobase of exon model per Million mapped reads) of each gene, which is currently the most commonly used method for estimating gene expression levels, was calculated based on the length of the gene and reads count mapped to this gene (Mortazavi et al. 2008).

Differential Expression Analysis

For each sequenced library, the read counts were adjusted by edgeR program package through one scaling normalized factor. Differential expression analysis of two samples was performed using the DEGSeq R package (1.12.0). *P* values were adjusted using the Benjamini-Hochberg procedure (Mortazavi et al. 2008). Corrected *P* value of 0.005 and log2 (fold_change) of 1 were set as the threshold for significantly differential expression. Volcano plots were applied to intuitively show the differentially expressed genes. Hierarchical cluster analysis of differentially expressed genes union was performed to assess the transcriptional pattern variations among SP, SB, WM, and W using Cluster 3.0 (de Hoon et al. 2004). Venn charts were drawn using VennDiagram R package to exhibit shared or specific differentially expressed genes between different pairwise comparisons.

Gene ontology (GO) enrichment analysis of differentially expressed genes was implemented by the GOseq R package, in which gene length bias was corrected. GO terms with corrected *P* value less than 0.05 were considered significantly enriched by differential expressed genes (Young et al. 2010).

Quantitative Real-Time Validation

To validate our transcriptome sequencing results, six differentially expressed genes, mainly the ones in the enriched GO terms, were selected for quantitative real-time PCR (qRT-PCR) analysis. RNA was extracted separately from 30 samples (10 samples from each of WM, SB, and SP) and then cDNA was synthesized from RNA, which was applied as the template. Primers for qRT-PCR were designed using Premier Primer 5 and were listed in Supplemental Table S1. Each pair

of primers could generate a single fragment which was proved to be identical to the sequence of corresponding gene by sequencing. *β-Actin* was employed as a reference standard to normalize the expression levels between samples (Mori et al. 2008; Wang et al. 2009). The amplification was carried out on Mastercycler ep realplex 4S (Eppendorf, GER) machine using QuantiFast SYBR Green PCR Kit (Qiagen, GER), with the following profile: 95 °C for 5 min, then 40 cycles of 95 °C for 10 s, 60 °C for 30 s. The measurements were run in four replicates. Relative gene expression data was analyzed using the $2^{-\Delta\Delta CT}$ method, where CT is the threshold cycle (Livak and Schmittgen 2001). Data were examined for homogeneity of variances (*F* text), then analyzed by *t* test using software SPSS13.0. Difference was considered statistically significant if $P<0.05$.

Results

De novo Transcriptome Sequencing, Assembly, and Gene Function Annotation

A total of 109,108,128 clean reads filtered from 113,282,886 raw reads (SRR1533149) (Table 1, Mantle_1 and Mantle_2) were generated through Illumina sequencing and assembled into 177,867 transcripts and 85,141 unigenes with a mean length of 1108 and 826 bp, respectively. These unigenes were annotated, among which 21,600 genes were annotated in NR, 1694 genes were annotated in NT, 7832 genes were annotated in KO, 16,007 genes were annotated in Swiss-Prot, 20,928 genes were annotated in PFAM, 20,928 genes were annotated in GO, and 11,102 genes were annotated in KOG. The functions of predicted unigenes were classified in GO, KOG, and KEGG. According to the GO annotation, mantle unigenes of the clam *M. meretrix* participated in 11 molecular functions (MF) (Fig. 2a). And these unigenes were involved in 26 COG classifications (Fig. 2b) and 31 KEGG pathways (Fig. 2c).

Table 1 Information and quality of RNA-seq

Sample	Clean bases	Q20 (%)	Q30 (%)	GC (%)
Mantle_1	5.46G	97.47	91.59	40.38
Mantle_2	5.46G	96.51	90.27	40.41
WM	1.21G	98.02	93.48	40.98
SB	1.32G	98.15	93.83	39.63
SP	1.48G	97.77	93.00	41.15
W	1.42G	98.13	93.69	40.11

Mantle_1: the left-end sequencing of the mantle transcriptome; Mantle_2: the right-end sequencing of the mantle transcriptome; the total clean reads of mantle transcriptome was the sum of clean reads of Mantle_1 and Mantle_2, the same to the total raw reads

DGE Library Sequencing, Mapping, and Gene Expression Quantification

Four DGE libraries of the clam *M. meretrix* with different shell color morph were sequenced, and the raw data were deposited in SRA of the NCBI with accession numbers of SRR1533133 (WM), SRR1533145 (SB), SRR1533146 (SP), and SRR153347 (W). The total clean reads in each DGE library ranged from 1.21 to 1.48 G (Table 1, WM, SB, SP, W), and the percentage of the clean reads which could be mapped to the unigenes from the reference de novo transcriptome was 87.24 (SP), 86.66 (SB), 87.31 (WM), and 88.71 % (W), respectively. Read counts and RPKM of each mapped clean read was calculated, which was considered as the expression level of the gene. Density distribution of expression level in each DGE library based on $\log_{10}$ (RPKM) was exhibited in Fig. 3a, which showed that the distribution patterns in WM, SB, and SP were similar and were obviously different from that in W.

Gene Expression Variations Among Clams with Different Shell Color Morphs

The expression level of genes based on the read counts are compared among four DGE libraries in pairs and the differentially expressed genes (DEGs) were detected (corrected *P* value <0.005 & $|\log_2$ (fold_change)$|>1$). The global expression profiles of DEGs union in each DGE library were estimated by hierarchical clustering (Fig. 4), which exhibited that the global expression pattern of DEGs in W was distinguishable from that in other three samples (SP, SB, and WM), and WM clustered more closely with SB rather than with SP. Consistent results were observed in volcano plots (Fig. 3b), which exhibited that there were more DEGs (red dots) between W and any of other three morphs, compared to those between SP, SB, and WM. The detailed numbers of DEGs unique or shared among multiple pairwise comparisons was summarized by Venn diagram (Fig. 5). There were 78, 54, 62, and 790 shared DEGs for the library of SP (Fig. 5a), SB (Fig. 5b), WM (Fig. 5c), and W (Fig. 5d), respectively, when comparing to other libraries. The shared DGEs were listed separately in Supplemental Table S2–S5 and were considered to be specifically expressed in SP, SB, WM, and W, respectively. To the DGE library of W, there were 1351, 1578, and 1408 DEGs detected in the comparison with SP, SB, and WM, respectively, while only 316, 228 and 254 DEGs were detected in the comparisons of WM vs SP, WM vs SB, and SB vs SP, respectively. The results of Venn diagram indicated that there was relatively less DEGs between WM, SB, and SP than those between W and any of other three strains (WM, SB, and SP), which was consistent with the results of clustering and volcano plots.

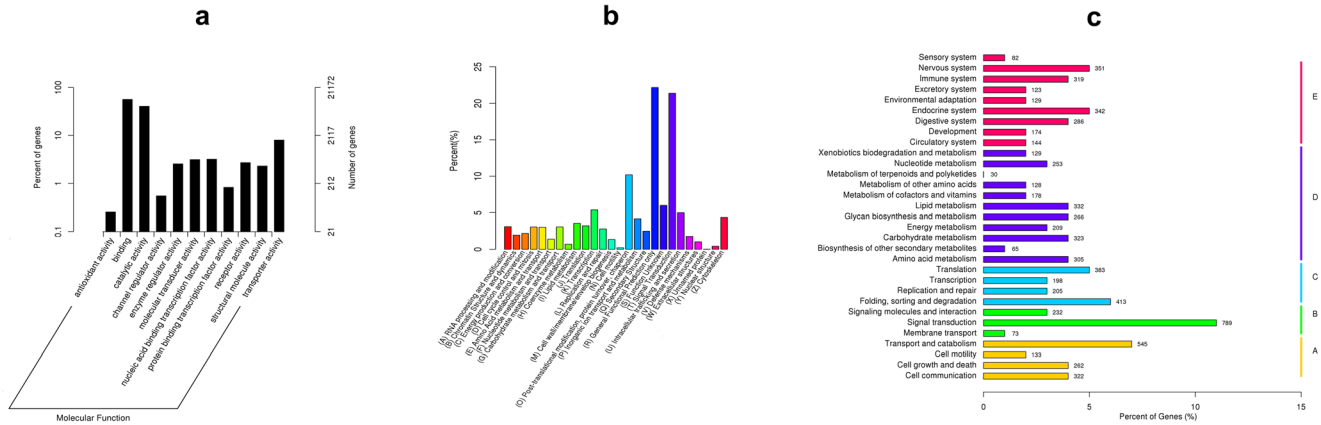

Fig. 2 Gene function annotation based on different databases. **a** Molecular function (MF) categories of the unigenes annotated in GO. **b** Classification of the unigenes annotated in COG. **c** Classification of the unigenes annotated in KEGG (*A*, cellular processes; *B*, environmental information processing; *C*, genetic information processing; *D*, metabolism; *E*, organismal systems)

For DEGs in each pairwise comparison among four samples, GO enrichment analysis was performed to identify the main molecular function these DEGs exercise. Significantly enriched GO terms of MF (corrected *P* value less than 0.05) based on upregulated or downregulated DEGs were exhibited in Supplemental Table S6. With respect to the significantly enriched GO terms based on upregulated DEGs, "calcium ion binding" was shared across the comparisons of WM vs W, WM vs SP, WM vs SB, and SB vs W. Go term of "calcium ion binding" was also enriched from the upregulated DEGs in comparison of SP vs W, although it was not significant. DEGs involved in the GO term of "calcium ion binding" for each pairwise comparison were listed in Table 2.

qRT-PCR Validation

Illumina sequence data were validated by comparing the gene's transcript level estimated from the RNA-Seq data with qRT-PCR results of six selected genes (Supplemental Table S7). The results showed that the expression patterns of *Mitf*, *Notch*, *Myl*, and *Ryr* agreed well between RNA-Seq and qRT-PCR. The expression patterns of other two genes *Scp* and

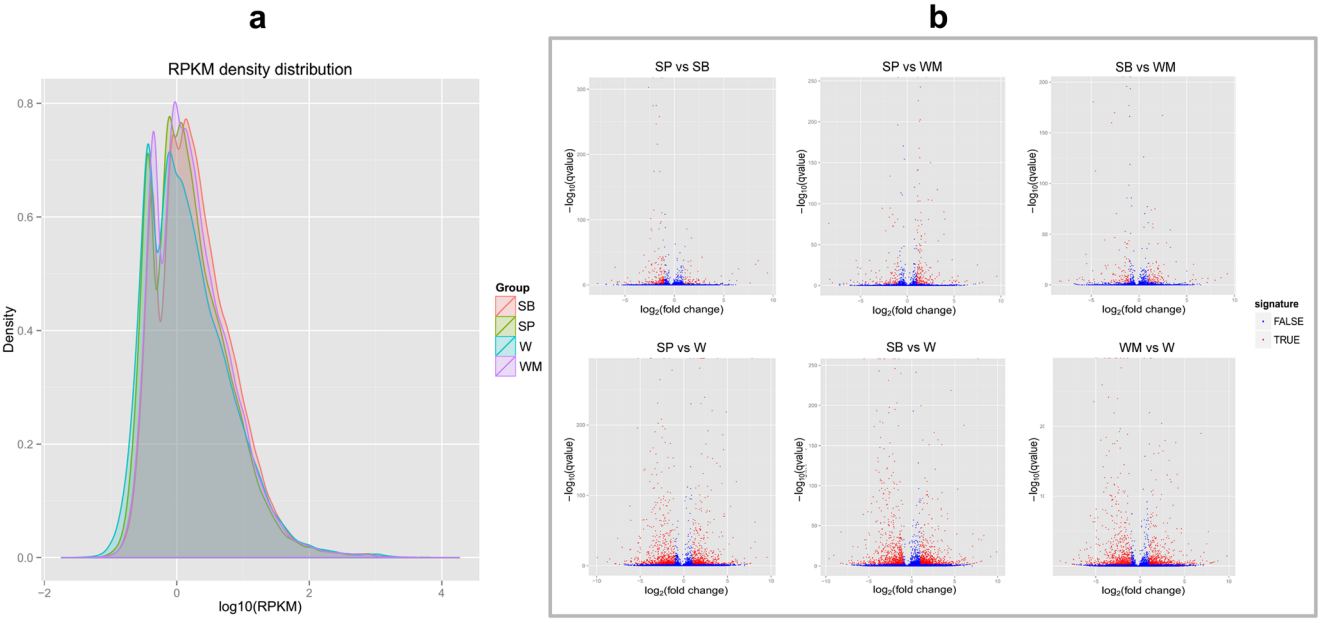

Fig. 3 **a** Density distribution of expression level based on $\log_{10}$ (RPKM), exhibiting the overall difference in expression profile among four DGE libraries. **b** Volcano plots summarizing DEGs in each pairwise comparison according to the gene-specific *t* test. Each point represented a gene which is ranked by fold change and by *t* test. The *red point* was DEG, while the *blue* one was not)

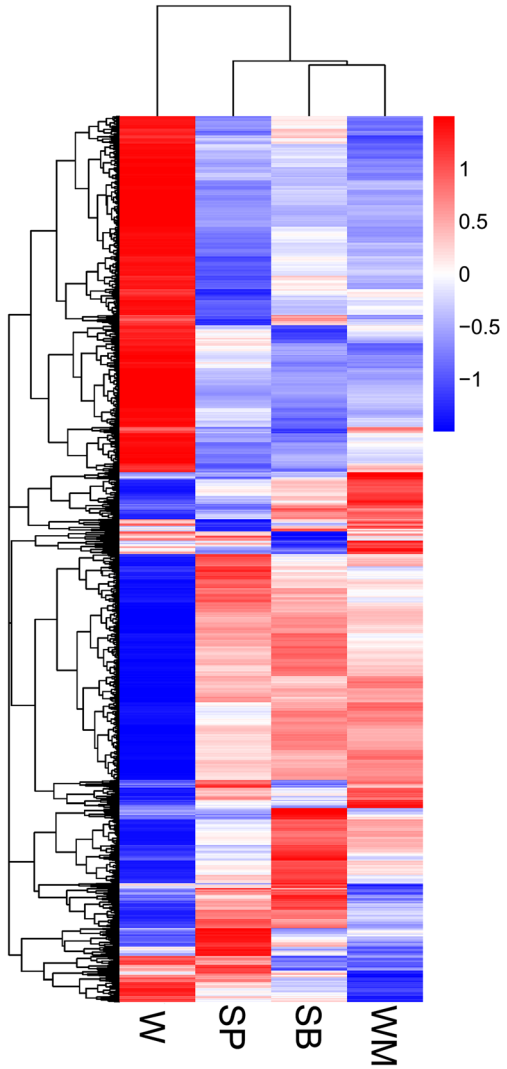

Fig. 4 Hierarchical clustering of DEGs union among four DGEs libraries (*SP*, *SB*, *WM*, and *W*)

Spink showed a similar trend in up- or downregulation exhibited by RNA-seq.

Discussion

Pigmentation has received much attention due to its strong implications for speciation and adaptation. Apart from a few model species, the molecular mechanism of this phenotypic diversification remains largely unknown, especially for the pigmentation in the shell of mollusks, which forms a variety of shell color patterns. In this study, we expect to investigate the genes involved in the shell color polymorphism and get information about the molecular mechanism of shell color determination by comparative transcriptome analyses among clams (*M. meretrix*) with different shell color morph.

Mollusk shell is a natural biomaterial secreted by the mantle (Addadi and Weiner 1985). Like another species of hard clam *Mercenaria mercenaria*, the shell color of *M. meretrix* roots in periostracum pigmentation, which originates in the mantle (Hillman 1961). To obtain an integrated view of the transcriptional events involved in the periostracum pigmentation, we used high-throughput sequencing technologies to analyze the transcripts in the mantle of *M. meretrix*. About 109, 108,128 clean reads were obtained and assembled into 177, 867 transcripts and 85,141 unigenes. This represents the most extensive transcriptional data set for the mantle of hard clams so far available. The functional classification of these transcripts according to the GO database showed that "binding" and "catalytic activity" were the dominant molecular functions involved (Fig. 2a), which was consistent with previous studies in other bivalves (Freer et al. 2014; Joubert et al. 2010; Shi et al. 2013). It is worth noting that a significant proportion of sequences in our mantle transcriptome were implicated in the function classification of "signal transduction" based on both COG and KEGG databases (Fig. 2b, c). Signal transduction occurs when an extracellular signaling molecule activates a specific receptor located on the cell surface or inside the cell and triggers a biochemical chain of events inside the cell in turn (Hlavacek et al. 2003). It is therefore hypothesized that there are complex interactions between cells in mantle, which is potentially involved in the formation of variant shell color patterns of *M. meretrix*.

In this study, digital gene expression (DGE) analysis was applied to identifying the genes potentially involved in the shell color determination of the clam *M. meretrix*, by comparing the gene expression profiles in the mantle among clams with four kinds of shell color morphs. All results including density distribution of expression level (Fig. 3a), volcano plots (Fig. 3b), hierarchical clustering (Fig. 4), and Venn charts (Fig. 5) indicated that W was distinguishable from the other three morphs (SP, SB, and WM). There were more DEGs detected in the comparison between W and each of other three morphs than those detected in the comparison between SP, SB, and WM. SP, SB, and WM were three clam strains obtained by our artificial selection in successive generations, which made them with lower genetic diversity than W. The increased number of DEGs would be derived from the shell color difference between W and any of the other three morphs, and from the other difference owning to genetic background difference, the latter potentially brought in false-positives. Similar phenomenon happened in other reports. A total of 2235 DEGs (479 genes upregulated and 1756 genes downregulated) were detected in black versus white sheep skin, and the sheep were randomly sampled (Fan et al. 2013). A total of 1161 DEGs were detected in *q* mutants versus wild-type silkworm integument (Nie et al. 2014). While much fewer DEGs (200–300) were detected in full-sib *Midas cichlids* with different body color (Henning et al. 2013). Combined with these reports, our data proposed that using the samples with lower genetic diversity was a solution to avoid false-positives for

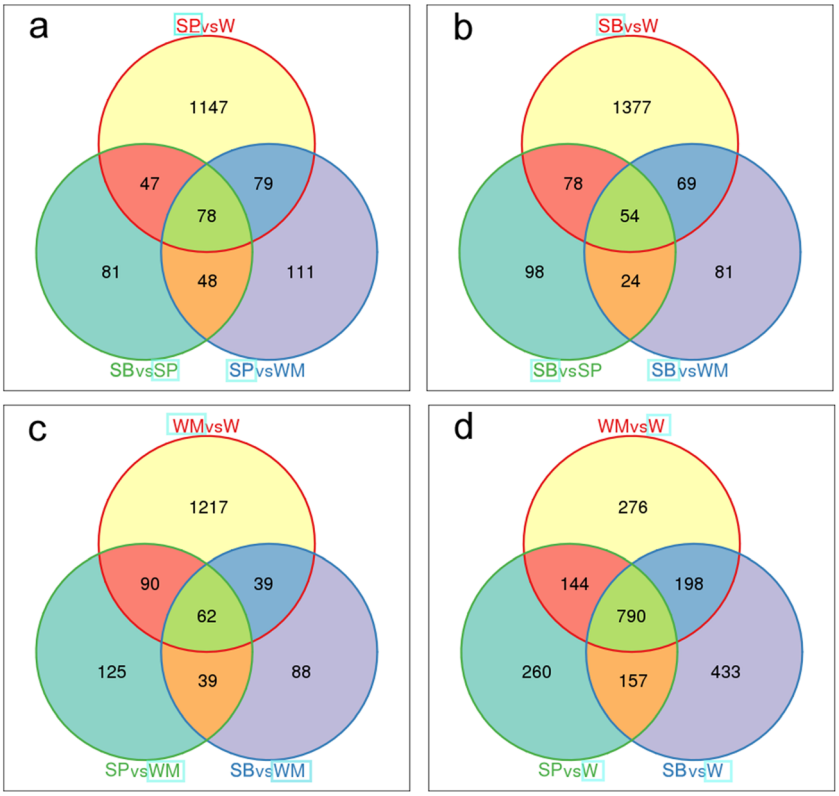

Fig. 5 Venn diagrams for number comparisons of DEGs among four DGE libraries (SP, SB, WM, and W). Shared or specific DEGs were detected in multiple pairwise comparisons, including SP compared with other three samples in pairs (**a**), SB compared with other three samples in pairs (**b**), WM compared with other three samples in pairs (**c**), and W compared with other three samples in pairs (**d**). *SPvsW*, the number of DEGs between SP and W, the same to *SBvsW*, *WMvsW*, *SBvsSP*, *SBvsWM*, and *SPvsWM*. The number of the shared DEGs are in the cross area, while the number of the specific DEGs are in the single area

DEGs detecting. The application of the three clam strains we developed to the DGE analyses makes it possible for mining the real and effective DEGs responsible for the different shell color formation.

The shell color morph of a clam was influenced by the amount of overlying pigments, and it is visible that the amount of pigment in the four kinds of shell color morphs was very different from each other. It appears that WM has the largest amount of pigment, SB comes the second, SP has the third, and W has the minimal (nearly none). The results of hierarchical clustering (Fig. 4) conformed to this trend, which exhibited that the global expression profile of DEGs union in WM clustered more closely with SB rather than with SP, and was almost completely opposite to that in W. Hence, it supported that there was a relationship between gene expression and shell pigmentation in the clam *M. meretrix*, and the DEGs detected would give clues on the molecular mechanism of shell color determination.

There are so many DEGs detected, from which we need to pick out the genes of interest. We presented GO enrichment analysis to identify the main molecular functions the DEGs exercise and the genes involved, which would help us understand the potential cause of different shell color morph formation. It is of great interest that there was a shared GO term (calcium ion binding) enriched in the upregulated DEGs across all pairwise comparisons except SB vs SP (Supplemental Table S6). The upregulated DEGs for these comparisons represented the genes with higher expression level in the samples with a deeper shell color (colored degree: WM>SB>SP>W). The genes involved in the shared GO term were listed in Table 2, among which *Notch*-related genes were presented in all comparisons, and thus arguing for the vital role of Notch signaling in the pigmentation of different shell color.

The Notch signaling pathway is an essential cell-cell interaction mechanism, which regulates processes such as cell proliferation, cell fate decisions, differentiation, or stem cell maintenance (Artavanis-Tsakonas et al. 1999). In this pathway, receptors and ligands are single-pass transmembrane proteins with large extracellular domains that consist primarily of epidermal growth factor (EGF)-like repeats (Bray 2006). Although the function of Notch signaling pathway in development has been well studied and widely accepted (Artavanis-Tsakonas et al. 1999), nothing was known on this pathway in the pigmentary system until a few years ago, and the related reports were mainly focus on its role in the hair pigmentation of mammals. In the pigmentary system, members of the Notch signaling pathway are expressed in melanocytes (Moriyama et al. 2006) and seem to be upregulated in melanoma cell lines (Nickoloff et al. 2005). Transgenic and knockout mice were instrumental in analyzing the roles of Notch signaling pathway in the pigmentary system. Moriyama et al. (2006) reported that conditional deletion of *RBP-Jκ* encoding recombination signal binding protein Jκ (an important transcriptional

Table 2 Upregulated DEGs involved in the shared enriched GO term (calcium ion binding) for each pairwise comparison

Comparison	NR description of genes	Symbol	$\log_2$ (fold_change)
WM vs W	Neurogenic locus notch protein	*Notch*	2.40
	Delta-like protein	*Dll*	3.25
	Ryanodine receptor 3, brain	*Ryr3*	1.46
	Myosin: SUBUNIT = regulatory light chain	*Myl*	1.25
	Sarcoplasmic calcium-binding protein	*Scp*	2.42
	Calmodulin	*Calm*	2.10
	X-box binding protein-like	*Xbpl*	3.21
	Von Willebrand factor D and EGF domains	*Vwde*	3.00
	Serine protease 48 isoform 1	*Prss48*	1.10
	Neurocalcin	*Ncald*	2.78
	Neurotrypsin	*Prss12*	2.41
	Allograft inflammatory factor 1	*Aif1*	1.10
WM vs SP	Neurogenic locus Notch protein	*Notch*	3.85
	Delta-like protein	*Dll*	1.62
	Ryanodine receptor 3, brain	*Ryr3*	1.32
	Myosin:SUBUNIT = regulatory light chain	*Myl*	1.03
	Sarcoplasmic calcium-binding protein	*Scp*	1.84
	Calmodulin	*Calm*	1.81
	X-box binding protein-like	*Xbpl*	1.28
WM vs SB	Neurogenic locus Notch protein	*Notch*	2.59
	Delta-like protein	*Dll*	1.38
	Ryanodine receptor 3, brain	*Ryr3*	1.19
	Myosin:SUBUNIT = regulatory light chain	*Myl*	1.76
	Sarcoplasmic calcium-binding protein	*Scp*	2.55
	Calmodulin	*Calm*	2.75
	X-box binding protein-like	*Xbpl*	1.86
	Troponin C	*Tnnc*	1.24
SB vs W	Neurogenic locus Notch protein	*Notch*	1.44
	Delta-like protein	*Dll*	1.86
	Fibrillin 1	*Fbn1*	1.94
	Scavenger receptor	*Scar*	5.23
	Integrin beta-like protein	*Itgbl*	1.33
	Latrophilin-2	*Lphn2*	1.35
	Latent-transforming growth factor beta-binding protein 4	*Ltbp4*	1.31
	Zonadhesin	*Zan*	1.35
	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase gamma-1	*Plcg1*	1.03
	Von Willebrand factor D and EGF domains	*Vwde*	3.83
	Neurocalcin	*Ncald*	2.61
SP vs W	Neurogenic locus Notch protein	*Notch*	1.68
	Delta-like protein	*Dll*	1.62
	Ryanodine receptor 3, brain	*Ryr3*	0.14
	Myosin:SUBUNIT = regulatory light chain	*Myl*	0.22
	Sarcoplasmic calcium-binding protein	*Scp*	0.58
	Fibrillin-1	*Fbn1*	1.82
	Latent-transforming growth factor beta-binding protein 4	*Ltbp4*	1.99
	Scavenger receptor	*Scar*	1.37
	Tenascin-R-like	*Tnrl*	3.35
	Cadherin-23	*Cdh23*	1.21
	Neurocalcin	*Ncald*	2.64

regulatory factor in Notch signaling pathway) in melanocytes of mice would cause elimination of melanocyte stem cells and thereby resulted in hair graying, which demonstrated the crucial role of Notch signaling in the maintenance of melanocyte stem cells, by preventing apoptosis. Conditional deletion of *Notch1* and *Notch2* in mice could also result in obvious coat color dilution, and this dilution was dose-dependent since it was influenced by the number of intact *Notch1* and *Notch2* alleles. No hair graying was observed when only one *Notch* allele was absent in melanocytes. In contrast, dispersed gray hairs were discernible when two *Notch* alleles are floxed ($Notch1^{flox/+}$, $Notch2^{flox/+}$; $Notch1^{+/+}$, $Notch2^{flox/flox}$; $Notch1^{flox/flox}$, $Notch2^{+/+}$). The coat was nearly white in the absence of both *Notch1* and *Notch2* ($Notch1^{flox/flox}$ and $Notch2^{flox/flox}$) (Schouwey et al. 2007). Data from these reports demonstrated that Notch pathway was essential for the pigment cell homeostasis and had a gene-dosage effect to induce a coat color dilution. Similar function of Notch pathway may be present in the shell pigmentation. Our data revealed that the expression of *Notch* was upregulated in the mantle of clam with the darker shell color. In other words, the lighter the shell color was, the less the *Notch* was expressed. It is therefore proposed that Notch pathway was involved in the shell pigmentation in a gene-dosage dependent pattern.

It was also different in shell color pattern among the four kinds of shell color morphs except for the colored degree. Spatial patterns are specified by cell interactions, so signaling pathways must be involved (French and Brakefield 2004). It has been reported that Notch pathway was involved in the color patterning. For mice, conditional deletion of *Notch1* in the embryonic ectoderm resulted in a mosaic pattern of hair growth (Pan et al. 2004). More direct evidence for the involvement of Notch pathway in body color patterning comes from the studies in the butterfly. Notch-mediated lateral inhibition was implicated in the spatial organization of butterfly wing scales which served the function of color pattern formation (Reed 2004). Moreover, the concentric eyespot is an intriguing

color pattern on the wing of butterfly (Beldade and Brakefield 2002), and eyespot pigment patterns are induced by a signal that originates from a group of focal cells at the center of the eyespot (Nijhout 1991). Reed and Serfas (2004) reported that *Notch* upregulation was an early event in the development of the eyespot pattern across multiple species of butterflies, and in a loss-of-eyespot mutant *Notch* expression was reduced at missing eyespots site. These reports support that Notch was probably involved in the shell color patterning, and it may be an upstream component of the shell color pattern determining process by defining the boundary of pigment occurs.

In this study, "signal transduction" was observed as the dominant function classification for the mantle expression data, and Notch signaling pathway may be just one component of a long-range signal for the shell pigmentation. It has been reported that Notch signaling pathway could be activated in a Ca^{2+}-dependent manner. Source of Notch activation was an accumulation of calcium (Raya et al. 2004). Calcium was considered as an important second messenger. Both the receptor (NOTCH) and ligand (DELTA) of Notch pathway contain EGF repeats in tandem, of which some are capable of Ca^{2+} binding (Fehon et al. 1990). Receptor-ligand interaction can initiate Notch pathway and is extremely sensitive to Ca^{2+} concentration (Rand et al. 1997). Elevation of Ca^{2+} concentration can be due to Ca^{2+} release from internal stores, which are located in the sarcoplasmic reticulum (SR). Internal stores provide Ca^{2+} mainly by two types of release channels, the inositol 1,4,5-trisphosphate receptor (IP3R) and the ryanodine receptor (RYR). In this study, as listed in Table 2, the expression of *Ryr* was upregulated across all comparisons except SB vs W. The increase expression of *Ryr* may induce Ca^{2+} release, and thereby determine the shell color via activating Notch signaling. RYR-dependent Ca^{2+} enrichment had been reported to affect laterality through the activation of Notch signaling (Garic-Stankovic et al. 2008). To the comparison of SB vs W, although *Ryr* was not detected to be upregulated, the gene encoding 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase gamma-1 (PLCG1) was present in the list (Table 2). PLCG1 catalyzes the formation of inositol 1,4,5-trisphosphate, which stimulates another channel IP3R and therefore mediates calcium release (Dent et al. 1996). Therefore, calcium signaling process might equally be implicated in shell color formation via activation Notch pathway.

Although there is few information on the mechanism of shell color formation, some contributing factors for the body color formation in other animals has been proposed in the past few years, which affect diverse processes, such as development of pigment-producing cells, synthesis of pigment and its transfer from pigment-producing cells to surroundings (Schouwey and Beermann 2008). Notch combined with calcium signaling, as an upstream component of the shell color-determining process, mainly affects the development of pigment-producing cells, such as maintenance of melanocyte stem cells. Our transcriptome data also revealed the DEGs potentially involved in other two processes, the synthesis of pigment and pigment transfer from pigment-producing cells to surroundings. As listed in Table 2, the expression of *Myl* encoding the protein of myosin regulatory light chain was upregulated across all comparisons except SB vs W. As mentioned above, the increase expression of *Ryr* may induce Ca^{2+} enrichment. Elevated Ca^{2+} encourages the formation of Ca^{2+}-calmodulin complex to activate the myosin light chain kinase to phosphorylate serine 19 of myosin light chain, which in turn removes inhibition of the myosin ATPase. This event is followed by ATP hydrolysis and sliding of myosin on actin filaments to generate cytoskeleton actin-myosin contraction (Somlyo and Somlyo 1994), which is necessary for the trafficking of the pigment organelles where pigment is made (Chang et al. 2012). The distribution of pigment is achieved via pigment organelles transport along cytoskeleton by the mean of contraction from the perinuclear region to dendrites and the subsequent transfer to neighboring environment (e.g., keratinocytes) (Boissy 2003). *Myl* functions in the pigment organelles (e.g., melanosome) translocation, and the significantly different expression of *Myl* observed among clams with shell color variations in this research implied that *Myl* exerted an influence on the pigment distribution which determined the shell color pattern. There were also some DEGs related to the process of the synthesis of pigment detected in our data, such as the gene encodes microphthalmia-associated transcription factor (MITF) (Levy et al. 2006). This transcription factor has been reported to play a pivotal role in melanocyte differentiation through the direct transcriptional control of *Tyr*, *Tyrp1*, and *Dct* genes, encoding three enzymes involved in pigment synthesis. In this study, the expression of *Mitf* in the clam strain WM was significantly higher than that in SB, SP, and W (Supplemental Table S7), which implied that *Mitf* was also involved in the shell color variations.

In conclusion, this study represents the first analysis of the relationship between whole-scale gene expression and shell color variations in the marine bivalve mollusks. Clam strains with different and monotonous shell color morph developed by artificial selection in successive generations were applied to comparative transcriptome analyses, which minimized the false-positives. The results of the comparative transcriptome analyses supported the idea that there was a relationship between gene expression and shell coloration in the clam *M. meretrix*. The possible involvement of Notch combined with calcium signaling pathway in shell color determination is a novel finding in clams. Other DEGs potentially implicated in shell color pigmentation were also discussed. This study sheds light on the color formation mechanism of the shell.

Acknowledgments This work was financially supported by the National Natural Science Foundation of China (31202018) and the Chinese National High-Tech R & D Program (2012AA10A410).

References

Adamkewicz L, Castagna M (1988) Genetics of shell color and pattern in the bay scallop *Argopecten irradians*. J Hered 79:14–17

Addadi L, Weiner S (1985) Interactions between acidic proteins and crystals: stereochemical requirements in biomineralization. Proc Natl Acad Sci U S A 82:4110–4114

Artavanis-Tsakonas S, Rand MD, Lake RJ (1999) Notch signaling: cell fate control and signal integration in development. Science 284: 770–776

Bagnara JT, Matsumoto J (2007) Comparative anatomy and physiology of pigment cells in nonmammalian tissues. The Pigmentary System: Physiology and Pathophysiology, Second Edition 11–59

Barbato M, Bernard M, Borrelli L, Fiorito G (2007) Body patterns in cephalopods: "polyphenism" as a way of information exchange. Pattern Recogn Lett 28:1854–1864

Beldade P, Brakefield PM (2002) The genetics and evo–devo of butterfly wing patterns. Nat Rev Genet 3:442–452

Boissy RE (2003) Melanosome transfer to and translocation in the keratinocyte. Exp Dermatol 12:5–12

Bray SJ (2006) Notch signalling: a simple pathway becomes complex. Nat Rev Mol Cell Biol 7:678–689

Cain A (1988) The colours of marine bivalve shells with special reference to *Macoma balthica*. Malacologia 28:289–318

Chang H, Choi H, Joo KM, Kim D, Lee TR (2012) Manassantin B inhibits melanosome transport in melanocytes by disrupting the melanophilin–myosin Va interaction. Pigm Cell Melanoma R 25: 765–772

Croucher PJ, Brewer MS, Winchell CJ, Oxford GS, Gillespie RG (2013) *De novo* characterization of the gene-rich transcriptomes of two color-polymorphic spiders, *Theridion grallator* and *T. californicum* (Araneae: Theridiidae), with special reference to pigment genes. BMC Genomics 14:862

de Hoon MJ, Imoto S, Nolan J, Miyano S (2004) Open source clustering software. Bioinformatics 20:1453–1454

Dent M, Raisman G, Lai FA (1996) Expression of type 1 inositol 1, 4, 5-trisphosphate receptor during axogenesis and synaptic contact in the central and peripheral nervous system of developing rat. Development 122:1029–1039

Fan R et al (2013) Skin transcriptome profiles associated with coat color in sheep. BMC Genomics 14:389

Fehon RG, Kooh PJ, Rebay I, Regan CL, Xu T, Muskavitch MA, Artavanis-Tsakonas S (1990) Molecular interactions between the protein products of the neurogenic loci Notch and Delta, two EGF-homologous genes in *Drosophila*. Cell 61:523–534

Freer A, Bridgett S, Jiang J, Cusack M (2014) Biomineral proteins from *Mytilus edulis* mantle tissue transcriptome. Mar Biotechnol 16:34–45

French V, Brakefield PM (2004) Pattern formation: a focus on notch in butterfly eyespots. Curr Biol 14:R663–R665

Garic-Stankovic A, Hernandez M, Flentke GR, Zile MH, Smith SM (2008) A ryanodine receptor-dependent Ca^{2+} asymmetry at Hensen's node mediates avian lateral identity. Development 135: 3271–3280

Grabherr MG et al (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644–652

Heath D (1975) Colour, sunlight and internal temperatures in the land-snail *Cepaea nemoralis* (L.). Oecologia 19:29–38

Henning F, Jones JC, Franchini P, Meyer A (2013) Transcriptomics of morphological color change in polychromatic *Midas cichlids*. BMC Genomics 14:171

Hillman RE (1961) Formation of the periostracum in *Mercenaria mercenaria*. Science 134:1754–1755

Hlavacek WS, Faeder JR, Blinov ML, Perelson AS, Goldstein B (2003) The complexity of complexes in signal transduction. Biotechnol Bioeng 84:783–794

Houde AE, Endler JA (1990) Correlated evolution of female mating preferences and male color patterns in the guppy *Poecilia reticulata*. Science 248:1405–1408

Innes DJHLE (1977) Inheritance of a shell-color polymorphism in the mussel. J Hered 68:203–204

Joubert C et al (2010) Transcriptome and proteome analysis of Pinctada margaritifera calcifying mantle and shell: focus on biomineralization. BMC Genomics 11:613

Kobayashi T, Kawahara I, Hasekura O, Kijima A (2004) Genetic control of bluish shell color variation in the Pacific abalone, *Haliotis discus hannai*. J Shellfish Res 23:1153–1156

Levy C, Khaled M, Fisher DE (2006) MITF: master regulator of melanocyte development and melanoma oncogene. Trends Mol Med 12: 406–414

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinform 12:323

Liu X, Wu F, Zhao H, Zhang G, Guo X (2009) A novel shell color variant of the Pacific abalone *Haliotis discus hannai* Ino subject to genetic control and dietary influence. J Shellfish Res 28:419–424

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25:402–408

Mitton JB (1977) Shell color and pattern variation in *Mytilus edulis* and its adaptive significance. Chesapeake Sci 18:387–390

Mori R, Wang Q, Danenberg KD, Pinski JK, Danenberg PV (2008) Both beta-actin and GAPDH are useful reference genes for normalization of quantitative RT-PCR in human FFPE tissue samples of prostate cancer. Prostate 68:1555–1560

Moriyama M et al (2006) Notch signaling via Hes1 transcription factor maintains survival of melanoblasts and melanocyte stem cells. J Cell Biol 173:333–339

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5:621–628

Nickoloff BJ, Hendrix M, Pollock PM, Trent JM, Miele L, Qin J-Z (2005) Notch and NOXA-related pathways in melanoma cells. J Invest Dermatol Symp Proc 10:95–104

Nie H et al (2014) Transcriptome analysis of integument differentially expressed genes in the pigment mutant (quail) during molting of silkworm, *Bombyx mori*. PLoS One 9:e94185

Nijhout HF (1991) The development and evolution of butterfly wing patterns. Smithsonian series in comparative evolutionary biology (USA)

Pan Y, Lin M-H, Tian X, Cheng H-T, Gridley T, Shen J, Kopan R (2004) γ-Secretase functions through Notch signaling to maintain skin appendages but is not required for their patterning or initial morphogenesis. Dev Cell 7:731–743

Parkash R, Kalra B, Sharma V (2009) Impact of body melanisation on contrasting levels of desiccation resistance in a circumtropical and a generalist *Drosophila* species. Evol Ecol 24:207–225

Protas ME, Patel NH (2008) Evolution of coloration patterns. Annu Rev Cell Dev Biol 24:425–446

Rand MD, Lindblom A, Carlson J, Villoutreix BO, Stenflo J (1997) Calcium binding to tandem repeats of EGF-like modules. Expression and characterization of the EGF-like modules of human Notch-1 implicated in receptor-ligand interactions. Protein Sci 6: 2059–2071

Raya Á et al (2004) Notch activity acts as a sensor for extracellular calcium during vertebrate left–right determination. Nature 427: 121–128

Reed RD (2004) Evidence for Notch-mediated lateral inhibition in organizing butterfly wing scales. Dev Genes Evol 214:43–46

Reed RD, Serfas MS (2004) Butterfly wing pattern evolution is associated with changes in a Notch/Distal-less temporal pattern formation process. Curr Biol 14:1159–1166

Rodgers GM, Kelley JL, Morrell LJ (2010) Colour change and assortment in the western rainbowfish. Anim Behav 79:1025–1030

Scheil AE, Hilsmann S, Triebskorn R, Köhler H-R, Kapitel (2013) Shell colour polymorphism, injuries and immune defense in three helicid snail species, *Cepaea hortensis*, *Theba pisana* and *Cornu aspersum maximum*. Results in Immunol 3:73–78

Schouwey K, Beermann F (2008) The Notch pathway, hair graying and pigment cell homeostasis. Histol Histopathol 23:609–619

Schouwey K, Delmas V, Larue L, Zimber-Strobl U, Strobl LJ, Radtke F, Beermann F (2007) Notch1 and Notch2 receptors influence progressive hair graying in a dose-dependent manner. Dev Dynam 236: 282–289

Shi M, Lin Y, Xu G, Xie L, Hu X, Bao Z, Zhang R (2013) Characterization of the Zhikong scallop (*Chlamys farreri*) mantle transcriptome and identification of biomineralization-related genes. Mar Biotechnol 15:706–715

Sivka U, Snoj A, Palandačić A, Sušnik Bajec S (2013) Identification of candidate genes involved in marble color pattern formation in genus *Salmo*. Comp Biochem Phys D 8:244–249

Sokolova I, Berger VJ (2000) Physiological variation related to shell colour polymorphism in White Sea *Littorina saxatilis*. J Exp Mar Biol and Ecol 245:1–23

Somlyo AP, Somlyo AV (1994) Signal transduction and regulation in smooth muscle. Nature 372:231–236

Wang X, Liu B, Xiang J (2009) Cloning, characterization and expression of ferritin subunit from clam *Meretrix meretrixin* different larval stages. Comp Biochem Phys B 154:12–16

Wang C, Wachholtz M, Wang J, Liao X, Lu G (2014) Analysis of the skin transcriptome in two Oujiang color varieties of common carp. PLoS One 9:e90074

Winkler F, Estevez B, Jollan L, Garrido J (2001) Inheritance of the general shell color in the scallop *Argopecten purpuratus* (Bivalvia: Pectinidae). J Hered 92:521–525

Young MD, Wakefield MJ, Smyth GK, Oshlack A (2010) Method Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol 11: R14