Molecular cloning and characterization of Pif gene from pearl mussel, *Hyriopsis cumingii*, and the gene expression analysis during pearl formation

Rui Zhang¹ · Mengting Qin² · Jie Shi² · Lu Tan² · Jiamin Xu² · Zhenyan Tian² · Yuhui Wu² · Yuxuan Li² · Yitian Li² · Ning Wang²

Received: 18 September 2017 / Accepted: 2 April 2018 / Published online: 9 April 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

In the present study, the Pif gene of the freshwater pearl aquaculture mussel, *Hyriopsis cumingii* (HcPif) was successfully cloned and functionally characterized. The full sequence of HcPif gene consists of 3415 base pairs, which putatively encode two proteins, HcPif90 and HcPif80. A sequence analysis revealed that HcPif contained a von Willebrand factor type A domain and a chitin-binding domain, and shared many functional residues with other Pif homologues. A highly conserved sequence, FKGLDEIELML, at the C-terminus of Pif80s was identified as the key functional site. The corresponding peptide fragment markedly modified the morphology of calcite crystallites in CaCO₃ crystallization assay and might play an essential role in the interactive binding between HcPif80 and CaCO₃. Moreover, real-time PCR results showed that HcPif gene was dominantly expressed in the pearl secreting tissues and its expression changed in response to the different development status of the pearl sac during pearl aquaculture. The gene expression of HcPif was maximum 7 days after mantle grafting and declined to about the control level on day 30. Our in vitro and in vivo experimental data indicated that HcPif gene possessed the inherent characteristics of a nacre formation gene and its expression might faithfully reflect the pearl secretion status of the pearl mussels examined. Our findings may extend the understanding of the biomineralization mechanism of nacre formation and provide a potential biomarker for pearl farming.

Keywords Pif · *Hyriopsis cumingii* · Biomineralization · Pearl aquaculture

Introduction

The molecular mechanism of nacre formation has been investigated intensively because of its fantastical regulation in calcium carbonate (CaCO₃) biomineralization and commercial importance to pearl aquaculture (Yao et al. 2014; Watabe 2013; Nudelman 2015). It is widely accepted that the biogenesis of nacre results from an amazing assembly

of CaCO₃ tablets with organic matrix molecules, such as proteins and chitin, in a specific nanostructure (Takahashi et al. 2004; Bahn et al. 2015). The highly organized structure endows nacre with exceptional mechanical properties and beautiful iridescence to form either hard mollusk shells with superior fracture resistance or precious pearls (Falini et al. 1996). The focus of nacre formation studies is to identify the key organizer and the relevant structural arrangement mechanism underlying the interaction between inorganic CaCO₃ crystals and the organic framework.

The Pif gene is one of the best-studied nacre formation genes and potentially offers a thorough understanding of the organization of biocrystals during nacre biogenesis (Zhao et al. 2014; Liu et al. 2015; Marie et al. 2017). The Pif gene encodes two acidic proteins in pearl oysters, chitin-binding Pif90 and aragonite-binding Pif80, which might associate with one another via disulfide bonds to form a protein complex (Suzuki et al. 2009). Theoretically, the Pif complex has a versatile binding capacity, allowing it to assemble

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-018-1233-z>) contains supplementary material, which is available to authorized users.

✉ Ning Wang
wangning79@ujs.edu.cn

¹ School of Medicine, Jiangsu University,
Zhenjiang City 212013, China

² School of Food and Biological Engineering, Jiangsu
University, Zhenjiang City 212013, China

aragonite, aragonite-binding proteins, and a chitin matrix that nucleates nacre formation in the necessary structure (Kroger 2009). In vitro biomineralization experiments have also demonstrated the activity of Pif in promoting CaCO_3 nucleation on a chitin substrate and shown that the amino acid residues of “DDRK” were important in the interactive binding between Pif80 and aragonite crystals (Bahn et al. 2015; Chang and Evans 2015; Du et al. 2016). However, the functional sites in the Pif are still unclear, since most of previous studies of Pif have been limited to a few marine mollusk species (Joubert et al. 2010; Marie et al. 2012; Suzuki et al. 2013). For instance, the conserved sequence DDRK was only present in the Pifs of shellfish in the genus *Pinctada* (Suzuki et al. 2013). Comparative studies of more Pifs from other nacre-producing animals were highly desirable to determine the precise functional sites of Pif and to better understand their functional roles in nacre formation.

Besides the biomineralization study, a recent in vivo investigation of the Pif gene in *Pinctada margaritifera* revealed that the gene expression level of Pif was strongly associated with the pearl secretion rate in the test oysters. Pif gene may be a useful biomarker of pearl secretion rate and have utility in genetic selection during future breeding programs to produce high-quality pearls for the pearl aquaculture industry (Blay et al. 2016).

The freshwater pearl mussel, *Hyriopsis cumingii* is the most important freshwater pearl aquaculture species in China and is responsible for 80% of Chinese pearl production (Wang et al. 2007; Chen et al. 2017). As far as we know, the Pif gene of this species has not been identified. Therefore, in this study, we undertook the molecular cloning and functional characterization of the Pif gene from *H. cumingii* (HcPif) to investigate the sequence features and its functional mechanism. Both an in vitro CaCO_3 crystallization assay and an in vivo gene expression analysis were used to assess the functional roles of HcPif gene in CaCO_3 biomineralization during nacre formation.

Materials and methods

Experimental animals and tissue sampling

Eighty healthy *H. cumingii* with a full shell length of 8–9 cm were obtained from Weiwang freshwater pearl farm. The collected mussels were cultured for 2 weeks in our laboratory before any experiments. The laboratory culture of the mussels and mantle tissue transplantation surgery were performed as described in our previous report (Zhang et al. 2016). *Hyriopsis cumingii* without mantle tissue transplantation were killed and their tissue anatomy was analyzed. The adductor muscle, digestive tract, foot, gill, mantle, and hemocytes of the mussels were sampled for the analysis of

the gene expression. To examine the gene expression profile of HcPif during pearl formation, we first dissected the grafting donor mantle tissue pieces from recipient mussels 3, 7, and 30 days after transplantation surgery and then carefully excised a thin inner layer (<0.5 mm) of the pearl sac from each donor mantle tissue sample by removing the outer layer of cells with a sharp surgical blade, as reported in the references (McGinty et al. 2012; Zhu et al. 2015). The same batch of donor mantle tissues that had not been implanted was also sampled as control specimens for the differential expression analysis of HcPif gene. Each sample was harvested in triplicates and stored in RNA-later® solution (Ambion, Austin, USA) at -70°C .

Total RNA isolation

Total RNA was extracted from each sample with TRI Reagent® (Sigma, Saint Louis, USA), according to the manufacturer's instruction. The total RNA isolated from each sample was treated with Turbo DNA-free™ (Applied Biosystems, Carlsbad, USA) at 37°C or 1 h to remove any contaminating genomic DNA. The integrity and quantity of the isolated RNA were examined with an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, USA) and a Nanodrop ND-2000 spectrophotometer (Thermo Scientific, Wilmington, USA), respectively.

Molecular cloning and sequencing of HcPif

To clone the HcPif gene, the total RNA isolated from pearl sacs sampled 7 days after mantle tissue allograft was purified with Dynabeads® Oligo (dT) 25 (Ambion, Carlsbad, USA). The purified mRNA was used to synthesize cDNA, which was sequenced with 5-Gb-scale high-throughput sequencing on the Illumina HiSeq™ 2500 system (Illumina, San Diego, USA). After the raw reads were filtered, the raw sequence data were assembled de novo with the CLC Genomics Workbench software 6.0.4 (CLC Inc., Aarhus, Denmark). The assembled sequences were clustered with the CAP3 software to identify nonredundant unigenes (Huang and Madan 1999). One of the unigene sequences showed strong similarity to the Pif gene of *Pinctada maxima* (*e*-value: $9e-34$) in a blastX alignment analysis (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The sequence data were deposited in GenBank under accession number MF496231. The full cDNA sequence was then characterized.

Bioinformatic analysis

The open reading frame search and deduced amino acid sequence of HcPif analysis were performed with the Expasy web server (<http://web.expasy.org/>). The SignalP program (<http://www.cbs.dtu.dk/services/>) was used to predict the

signal peptide on the putative HcPif amino acid sequence. The functional domains of HcPif were assessed at the Interpro web page server (<http://www.ebi.ac.uk/interpro/>). The presence of disulfide bonds in HcPif was predicted with the DiANNA 1.1 program (<http://clavius.bc.edu/clotelab/DiANNA/>) (Kawabata et al. 1996; Ferre and Clote 2005; Zhang 2007). Pairwise and multiple sequence alignments of HcPif were conducted with the Clustal Omega web server (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). The phylogenetic relationships among the Pif homologues were determined by constructing a phylogenetic tree using the neighbor-joining algorithm with the MEGA 6.06 program and bootstrapping of 1000 replicates (Kumar et al. 2004).

Preparation of HcPif80 peptides and CaCO₃ crystallization assay

For the purpose of characterization of the functional sequence of HcPif, the highly conserved sequences close to the C-terminus of putative HcPif80 were chosen to synthesize test peptides for calcium carbonate crystallization experiments. Specifically, the conserved peptide WRTV-VVSLRKDKLKLK (HcPif80-16-basic), the conserved peptide FKGLDEIELML (HcPif80-12-acidic), and the 16 amino acid residues containing HcPif80-12-acidic, SFK-GLDEIELMLDDD (HcPif80-16-acidic), were synthesized by the GeneScript Corporation (GeneScript, Nanjing, China). The crystallization experiments were conducted with a conventional method reported in references (Blank et al. 2003; Wang et al. 2008). Briefly, 100 mM NaHCO₃ was added dropwise to 40 mM CaCl₂ solution and 0.22 µm filters were used to remove the immediate precipitate of CaCO₃ to produce a saturated CaCO₃ solution. Each test peptide was added to the prepared CaCO₃ solution at a final concentration of 1 mg/ml and incubated at 4 °C for 72 h. Meanwhile, a same volume of Tris–HCl buffer (1.0 M Tris–HCl, 200 mM NaCl, pH 7.4) was added to a test saturated CaCO₃ solution as the control treatment. To examine the morphology of CaCO₃ crystals in each test solution, 40 µl of the test solution was dried on aluminum sample stubs in a vacuum desiccator and imaged by a field emission scanning electron microscope of S-4800 II FESEM model (Hitachi high

technologies corporation, Tokyo, Japan) at an accelerating voltage of 5–30 kV.

Real-time quantitative PCR analysis

To examine the gene expression of HcPif in various samples, real-time quantitative PCR was conducted as the description in references (Wang et al. 2009). Briefly, 5 µg of total RNA from each test sample was used to synthesize template cDNA. The first-strand template cDNA was reverse transcribed with SuperScript® III First-Strand System (Invitrogen, Carlsbad, USA) according to the kit manual. After cDNA synthesis, real-time quantitative PCR was performed in a CFX-96 Touch™ Real-time PCR Detection System (Bio-Rad, Hercules, USA). Real-time quantitative PCR components were mixed in a total reaction volume of 20:10 µl of 2×SYBR ExTaqII premix (TaKaRa, Dalian, China), the forward and reverse primers (10 µM) and 1 µl diluted template cDNA (about 10 ng). The PCR condition was described previously (Zhang et al. 2016; Mao et al. 2017): pre-incubation at 95 °C for 30 s, followed by 40 cycles of denaturation at 95 °C for 5 s and annealing and extension at 60 °C for 34 s. Dissociation curve analysis (60–95 °C) was implemented to confirm the absence of nonspecific products. The primers used for this study were listed in Table 1. The expression of elongation factor 1 alpha gene (EF1α) was selected as the internal reference. The quantities of transcript were calculated as the relative expression levels of target genes to EF1α gene using the 2^{−ΔΔCt} method (Livak and Schmittgen 2001). Statistical analysis with one-way analysis of variance in the Sigma Plot 10.0 package was used to evaluate the differential transcription of the Pif gene among variant tissue samples. *P* values < 0.05 and < 0.01 were considered as statistically significant and highly statistically significant, respectively.

Results

Sequence analysis of HcPif cDNA

The complete cDNA sequence of HcPif consisted of 3415 base pairs including an open reading frame of 3162 base pairs which encoded 1035 amino acid residues

Table 1 Primers used in the present study

Name	Sequence (5′–3′)
HcPif real-time PCR forward primer	AGTTGCCCAATTACCCAATG
HcPif real-time PCR reverse primer	AGGGACGGGAGGGAAAGG
EF1α real-time PCR forward primer	GGAACCTCCCAGGCAGACTGTGC
EF1α real-time PCR reverse primer	TCAAAACGGGCCGCAGAGAAT

GenBank accession number: GW696134

EF1α elongation factor 1 alpha

Fig. 1 Amino acid sequence alignment of VWA domains from Pif homologues. Most of the referred sequences are from Genbank: Pif from *Pinctada fucata* BAH97338.1; PmPif from *Pinctada margaritifera* BAM66823.1; PmxPif from *Pinctada maxima* BAJ08001.1; PpPif from *Pteria penguin* AB505057.1; BMSP from *Mytilus galloprovincialis* BAK86420.1. The sequences of LG228264, LG232022,

LG233460, and LG236719 are obtained from the genome database of *Lottia Gigantea* (<http://genome.jgi-psf.org/Lotgi1/Lotgi1.home.html>). Asterisks indicate the amino acid residues of MIDAS motif described in the reference (Suzuki et al. 2013). Identical and similar amino acid residues are highlighted in pink and cyan, respectively

Fig. 2 Multiple sequence alignment of chitin-binding domains from Pif homologues. Disulfide bonds expected to be conserved in all Pif homologues are indicated by a line connecting the bonded cysteine

residues. (Abbreviations and sequence sources are same as the illustration in the legend of Fig. 1)

(Supplementary Fig. 1). The first 21 amino acids were predicted to form a signal peptide. A functional domain assessment revealed that HcPif contained a VWA domain and a chitin-binding domain. Two incomplete chitin-binding domains were detected at the downstream of the chitin-binding domain. One dibasic cleavage site was present at the sequence of “RVSKR” located after amino acid position 912. On the basis of the predicted dibasic cleavage site, the deduced proteins of HcPif were inferred to be HcPif90 and HcPif80, two proteins formed after posttranslational modification. The calculated molecular mass and isoelectric point (pI) of HcPif90 were 96 kDa and 4.79, respectively. HcPif80 was smaller and more acidic than HcPif90. The theoretical

molecular mass and pI of HcPif80 were 13 kDa and 4.15, respectively.

According to the multiple sequence alignment with homologous Pifs, HcPif90 was conserved at many essential amino acid residues involved in the VWA domain, the chitin-binding domain and the downstream region of the chitin-binding domain. In the VWA domain, HcPif90 showed a short MADIS motif sequence similar to that of the Pif90 derived from the marine pearl oysters including *Pinctada fucata*, *Pinctada margaritifera*, *Pinctada maxima*, and *Pteria penguin* (Fig. 1). In the chitin-binding domain, HcPif90 displayed the six consensus cysteine residues shared by all the other Pif90s (Fig. 2). In the downstream region of

Fig. 3 Comparison of downstream sequences of the Pif chitin-binding domains. The frame with a single asterisk labels an incomplete chitin-binding domain. The dibasic cleavage sequence is boxed by

broken lines. (Abbreviations and sequence sources are same as the illustration in the legend of Fig. 1)

chitin-binding domain, HcPif90 showed a high similarity with the Pifs of the marine pearl oysters, especially in the incomplete chitin-binding domain and at the dibasic cleavage site (Fig. 3). In the region of HcPif80, the amino acid sequence of HcPif was the shortest one among all the homologues. HcPif80 was composed of 118 amino acid residues and did not contain the repeated sequence DDRK, which were usually present in Pifs from the genus *Pinctada* (Fig. 4).

Of note, HcPif80 shared two highly conserved sequences near its C-terminus with the homologues from marine pearl oysters. One of the conserved sequences of “WRT-VVSLRKDKLKLK” (HcPif80-16 basic) consisted of 16 amino acid residues, which were rich in basic residues and had a theoretical pI of 11.17. Another conserved peptide FKGLDEIELML (HcPif80-12 acidic) contained more

acidic residues, with a theoretical pI of 4.14. The exact 16 amino acid residues at C-terminus of HcPif80, SFK-GLDEIELMLDDD (HcPif80-16 acidic), contained three more aspartates than that of HcPif80-12 and had a theoretical pI of 3.66.

Phylogenetic relationships of Pif homologues

For a better understanding of the evolutionary relationships of homologous Pifs, their amino acid sequences were used for a phylogenetic analysis. A phylogenetic tree was generated using the neighbor-joining method. All the Pifs were classified into three subtrees (Fig. 5). The Pif homologues from marine pearl oysters were clustered on one subtree. HcPif was located close to Pifs derived from *Pterioidea* species and placed more distantly from the Pifs of the genus

```

HcPif 536 -----
Pif 548 DERSRKDYLKAILK--R-----DDR--K--DK--TDDTKGRRIID-----
PmPif 559 DEGGRKGYRKGGRRKNDK-----NGKRTDDRNGKRRDDRIFDDRKGRRITDDRKG
PmxPif 562 DERGRKGYRKGGRRKDDKDRRDDRNGKRRDDRNGKRRDDKIFDDRKGRRITDDKKN
PpPif 550 -----

RNDIIDRRGRRKDDRKDGGRDDGKGRDDRKDNRLDDIKDKNDPMTLISNGEC-DNFELNDCFEP
RRDRIDRRNGRRITDDRKNRRDDRKNRRDDIKDKNRDDKKNADPMTLISNGEC-DNFELHDCFEP
RRDRIDRRNGRRITDDRKD--D-----RRDDIKDNRRDDKKNTPMTLISNGEC-DNFELHDCFEP
-----YAKILQKYN-----FKLPGLIEKLNDADNEMSLSDGECGDDMDLEICAEP

-----HKADSDSSSDNCDSS-----DSKDKN--NK-----
SIAITTGKKSAGFSVTSSEKDEVDLEIDEDKKGYLWDKE--DGFDRNGKDKDRNGDRSDDRRCGYWKKK
SLAITTGKKSAGFSVSSIERDEVDLEVDNKKGYLWNKDKKDKDDKNRRDKDKNGDRITDD-----KK
SLAITTGKKSAGFSVSSSERDEVDLEIDDDKKGYLWNKDKKDKDDKNRRDKDKNGDRITDD-----KK
SIAITANKKKAGFHVCGTRSKDVIDLEDD-----DSKDDNGVNG-PNG-----

--DSDNSDSD--GNSDS-----
DKDDNGKDKDKGDKSDDKKG-----YYWKDKDKDKNGK--KDKRRDKSDDKSLDDIVREIER--
SLDDLVEIER--KSDDKSFDDLVEIERKSDDKSFDDLVEIERKSDDKSLDDLVEIERK
SLDDLVEIEMERS--KIDDK-----D-----
VTDANNNDNPDSS--SSSS-----

-----SD--ESKESGKSAQRTVWVSLRKDKLKLKGNRVAEEFQDSTPLVSYSSLQIGGCSIY
--S-KGNKDDNKDDDDDDKKGKTVSLKINNGHIRGRDDREDK--VLDGDLKTFSGFQIGQGASN
SDDKGNRRKDDDKNDEDDDKKGKTVSLKINNGHIRGRDDREDK--IMDGLKTFSGFQIGQGASN
--NGRGKDDNKNDEDDDKKGKTVSLKINNGHIRGRDDREDK--IMDGLKTFSGFQIGQGASN
-----SSS-ASSLEDDDD-TVKTVSIVNKGKVKVRKDEVED--SLDGLKTFSGIKIGQGAGK

NSFKGLMDETEIMDDD-----
KNFKGYMDEVYTYFCDFGKEADYDEDD-DDDDDDSDENDKNDDKDKGKTDDKDKKDRKDGRRD--K
KNFKGYMDEVYTYFCDFGKEADFDEDDNGDDDDDDDDKD-NDAGDDNK--DDNNNNGRKDD-MNN--
KNFKGYMDEVYTYFCDFGKEADFDEQDDNDDDDDDDDND-NDTGDDNK--DDKNNNGKKDDKNMNGKK
KKFKGYLDEFYTYFCNFGDDDIYGDGED-DSDSDSGSDSD-DDSAAAA SARAQRRARLRALLSCAR--

DGRD--DRKDRRDDRKDDRKGKDDRKDDRKGKDDRKDGRRDVRDDRDRGDKYDKK--DKDNDRLSDKD
D--GRKD--DNNKKDDK--DRSD-KNGCKDD-KD-KDIKD--KFSDKNDKDN--D-A
DDKNNNGKKD--DNNKNDDK--DRSD-KNGCKDD-KD-KDTKD--KFSDKNDKDN--D-A
--GGGS--AAAAAAA--AAAAAGSGGCKFKQFVWGGNG--GGGMMVAGSSNG--G

DRKDVDNDKDDDDNEKLYKRAMKKCDYVNKNVAKWLDKR
DR-DINDNK-----LYRRAMKKCDFVNKNVEKWLDKR
DR-NIDDNK-----LYKRAMKKCDFVNKNVEKWLDKR
GGVMVGSSSG-----GDLGTALYKDFKFKYKAGK

```

Fig. 4 Amino acid sequence alignment of Pif80 homologues. The sequence used for the peptide synthesis of HcPif80-16 basic is boxed by broken lines. The conserved acidic amino acid residues used for the

synthesis of HcPif80-16 acidic is boxed by lines. (Abbreviations and sequence sources are same as the illustration in the legend of Fig. 1)

Fig. 5 Phylogenetic tree of Pif homologues. The phylogenetic tree is generated by the neighbor-joining method using MEGA software based on amino acid sequence alignment with 1000 replications of bootstrap. The scale bar indicates a branch length of 0.1. (Abbreviations and sequence sources are same as the illustration in the legend of Fig. 1)

Pinctada than from the Pif of *P. penguin*. The blue mussel shell protein (BMSP) and Pifs from *Lottia. gigantean* formed another subtree.

Calcium carbonate crystal modification by HcPif80 peptides

The interactive binding of the HcPif80 peptides to calcium carbonate was examined by a CaCO_3 crystallization assay. In the control groups without peptide addition, calcium carbonate precipitated as calcite crystal with a typical rhombohedral shape, which had six parallelogram-shaped faces (Fig. 6a). The calcite crystal grown with the addition of HcPif80-16 basic did not show visible changes in their morphology (Fig. 6b). By contrast, the addition of HcPif80-12 or HcPif80-16 resulted in considerable modification of the calcite surface. The surface of the HcPif80-12 treated calcite became rough and was in a shape with dense etch pits (Fig. 6c). Compared with HcPif80-12, HcPif80-16 demonstrated a more intensive modification of calcite crystallites. Three corners of HcPif80-16 treated calcites became curved and even lead to an irregular shape of the calcite without the straight sharp edges of rhombohedra (Fig. 6d).

Tissue expression profile of HcPif

Transcription profiles of HcPif in various tissues of *H. cumingii* were determined using real-time quantitative PCR. In mantle tissues, the abundance of HcPif transcripts was

highest, which was more than 500-fold higher than that of other examined tissues. In contrast, adductor muscle, foot, gill, and hemocytes were in very low gene expression level of HcPif. No HcPif transcripts were detected in digestive tract (Fig. 7).

Differential expression of HcPif during pearl sac formation

To understand the association between HcPif gene and pearl formation, the spatiotemporal expression pattern of HcPif gene in the donor mantle tissues during pearl sac formation was investigated by real-time quantitative PCR. There were no significant changes of HcPif transcription in the donor mantle tissues in the 3 days immediately after tissue grafting. By day 7 after tissue grafting, the transcript abundance of HcPif gene in donor mantle was increased dramatically to 26-fold higher than that of the control tissues. At day 30 after transplantation surgery, the expression of HcPif gene had decreased to a level about 1.5-fold higher than that of the control sample (Fig. 8).

Discussion

In this paper, we reported the molecular cloning and functional characterization of the HcPif gene from the freshwater pearl mussels. The sequence analysis suggested the conservation in gene structure between HcPif and its homologues. HcPif gene might act via a similar mechanism of nacre biomineralization as its homologues did. Like the Pif genes from marine pearl oysters, a dibasic cleavage site was located in amino acid position 912, where the encoded HcPif might be separated as two proteins, HcPif90 and HcPif80, after posttranslational modification. The calculated pI of these two proteins indicated they were acid proteins, which were the key elements organizing the higher-order lamellar structure of shells in classical hypothesis of biomineralization (Belcher et al. 1996). Similar to other homologous Pif90s, HcPif90 was conserved in the amino acid residues, which constituted the VWA domain and the chitin-binding domain. The function of the VWA domain in various extracellular matrix proteins was often associated with the regulation of protein–protein interaction (Tuckwell 1999; Whitaker and Hynes 2002). In the genome of *P. fucata*, Pif90 was identified as the unique gene, which contained both a VWA domain and a chitin-binding domain (Suzuki et al. 2013). It had been believed that the VWA domain and chitin-binding domain cooperatively caused Pif90 to bind to chitin microfibrils with other nacre formation proteins such as N16 together to produce an organic framework that allows the formation of the lamellar nacre sheet (Suzuki et al. 2009).

Fig. 6 Scanning electron microscopy images of calcite crystallites grown in the presence of our test solution. **a** peptide buffer only; **b** 1 mg/ml HcPif80-16 basic; **c** 1 mg/ml HcPif80-12 acidic; **d** 1 mg/ml HcPif80-16 acidic. Scale bar indicates a length of 5 μ m

A phylogenic analysis revealed a close relationship between HcPif and the Pifs of marine pearl oysters. The four Pifs of *L. gigantean* were grouped into a monophyletic clade indicating a high frequency of gene duplication and the rapid evolution in Pif genes. Similarly, the presence of two incomplete chitin-binding domains in HcPif90 also supported this view point. Gene duplication could be one of the main causes of the variable sequence lengths of the Pif genes.

In contrast to the conserved HcPif90, HcPif80 contained novel sequences. Compared with its homologues,

HcPif80 did not contain any duplicated domains or repeated sequences. The short sequence of HcPif80 simplified the identification of the functional sites in Pif80 homologous proteins. A sequence alignment of the Pif80s identified two highly conserved sequences in the C-terminal end of HcPif80.

The functional importance of the conserved HcPif80 sequences was examined in a CaCO_3 crystallization assay. In the control group, calcite was the dominant stable polymorph of CaCO_3 displaying the standard rhombohedral shape with straight edges, sharp corners and large macrosteps, which

Fig. 7 Tissue distribution of HcPif transcripts is analyzed by real-time PCR in pearl mussels. Each value represents the relative expression level of HcPif gene in *M* mantles, *AM* adductor muscles, *F* feet, *G* gills, *DG* digestive glands, and *HE* hemocytes from one pool of three animals. The values are normalized to those of EF-1 α . Each value is the mean of triplicates and standard deviation is represented by the error bar. Significant differences from the mantle tissue sample are indicated with two asterisks at $P < 0.01$. *ND* not detected

made it convenient to detect any morphological changes. Both HcPif80-12-acidic and HcPif80-16-acidic markedly modified the morphology of calcite crystallites, whereas HcPif80-16-basic did not induce any visible changes of calcite morphological structure. This suggested that the conserved acidic amino sequence FKGLDEIELML could play an essential role in the interactive binding between HcPif80 and calcium carbonate. Compared with HcPif80-12-acidic, HcPif80-16-acidic showed a stronger modification effect. This was attributed to three more aspartates, since polyanionic protein residues might favor the binding of proteins to positively charged calcium ions (Baeuerlein 2007). In fact, it had been proved addition of an aqueous aspartic acid solution was able to result in calcite crystallites to display curved corners and etch pits on crystal surfaces. The binding of aspartate to calcite surface steps changed the step-edge free energy, giving rise to direction-specific binding energy and leading to the morphological modification of calcite crystallites (Orme et al. 2001). Furthermore, similar modification was also reported in the studies of other shellfish nacre formation protein studies. It had been proved that the aspartate and asparagine (Asx)-rich proteins from abalone nacre could inhibit step motion at the corners of calcite crystallites and flatten the calcite rhombus (Fu et al. 2005). Besides the intact acidic nacre formation proteins, peptide fragments of Pif80 from *P. fucata* demonstrated an interactive binding to calcite, which effectively roughened calcite surface and caused the crystallites to adopt an irregular shape (Du et al. 2016). Therefore, it is reasonable to assume that the presence of conserved Asx rich peptide fragments allows variant

Fig. 8 Time-course expression analysis of HcPif gene in donor mantle tissues after mantle allograft surgery. The level of HcPif gene expression is determined by real-time PCR. The value relative to the control mantle sample is expressed as the mean of triplicates and standard deviation is represented by the error bar. Significant differences from the control donor mantle tissues are indicated with one asterisk at $P < 0.05$ and two asterisks at $P < 0.01$

acid nacre proteins like HcPif to regulate calcium carbonate crystal nucleation.

Besides the in vitro biomineralization data, our further in vivo gene expression study also indicated a tight association between HcPif and nacre deposition for pearl aquaculture. In tissue distribution analysis, the dominant expression of HcPif gene in pearl mussel mantle tissue suggested that HcPif gene might participate in mollusk calcium carbonate biomineralization, since mollusk mantle tissue was the unique place where calcium and bicarbonate ions could reach a supersaturated state for crystallization to occur (Giles et al. 1995). Moreover, mantle pieces of *H. cumingii* acquired after successful transplantation surgery developed into pearl sacs and produced pearls. The high expression of the HcPif gene in pearl sac samples indicated HcPif was highly demanded during pearl secretion. The differential gene expression pattern of HcPif in pearl sacs at different sampling times indicated the pearl secretion was in a different nacre deposition rate as time progressed. It took at least 7 days for the grafted mantle to become compatible to recipient mussels and capable of promoting pearl production at a high rate. At 30 days after mantle grafting, gene expression of HcPif declined to about the control level. This suggested the demand for pearl secretion in the grafted mantle pieces returned to a level similar to that in normal mantle tissues. Pearl formation is a defense mechanism of mollusk against potentially threatening irritants such as parasites inside the shell. A pearl sac is created to seal off the irritants. Thirty days are usually sufficient for the transplanted mantle pieces to develop into pearl sacs and complete the enclosure of

the inserted nuclei (Norton et al. 2000). Therefore, we proposed that the declining expression of HcPif from day 7 to 30 was associated with reduced immune responses and indicated the accomplishment of pearl sac development.

In conclusion, HcPif gene was highly conserved in the functional motifs involved in nacre biomineralization. The acidic peptide FKGLDEIELML was an important functional site of Pif80 homologous genes and suggested a universal control of CaCO_3 crystallization by Pif80. Besides, HcPif gene was always highly transcribed in pearl secreting tissues of pearl mussels and its expression level was timely altered in response to the different development status of pearl sacs during pearl aquaculture. Therefore, it is reasonable to infer that HcPif gene had the inherent characteristics of a nacre formation gene and could be a reliable biomarker for pearl farming.

Acknowledgements This study was supported by the Natural Science Foundation of Jiangsu province (Grant no BK20130496), Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant no 13KJB240001), Scientific Research Foundation for Returned Scholars, Ministry of Education of China and Startup Foundation for Advanced Talents of Jiangsu University (Grant no 13JDG006). Dr. Rui Zhang was supported by the National Natural Science Foundation of China (Grant no 81300673), Startup Foundation for Advanced Talents of Jiangsu University (Grant no 13JDG004), and Cultivation Project for Young Core Teacher of Jiangsu University.

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interest regarding the publication of this article.

References

- Bauerlein E (2007) Handbook of biomineralization. Wiley, Weinheim, pp 273–286
- Bahn S, Jo B, Hwang B, Choi Y, Cha H (2015) Role of Pif97 in nacre biomineralization: in vitro characterization of recombinant Pif 97 as a framework protein for the association of organic-inorganic layers in nacre. *Cryst Growth Des* 15:3666–3673
- Belcher AM, Wu XH, Christensen RJ, Hansma PK, Stucky GD (1996) Control of crystal phase switching and orientation by soluble mollusc-shell proteins. *Nature* 381:56–58
- Blank S, Arnoldi M, Khoshnavaz S, Treccani L, Kuntz M, Mann K, Grathwohl G, Fritz M (2003) The nacre protein perlucin nucleates growth of calcium carbonate crystals. *J Microsc* 212:280–291
- Blay C, Parrad S, Cabral P, Aihio V, Ky C (2016) Correlations between cultured pearl size parameters and PIF-177 biomarker expression in *Pinctada margaritifera* families reared in two contrasting environments. *Estuar Coast Shelf Sci* 182:254–260
- Chang EP, Evans JS (2015) Pif97, a von Willebrand and peritrophin biomineralization protein, organizes mineral nanoparticles and creates intracrystalline nanochambers. *Biochemistry* 54:5348–5355
- Chen X, Liu X, Bai Z, Zhao L, Li J (2017) HcTyr and HcTyp-1 of *Hyriopsis cumingii*, novel tyrosinase and tyrosinase-related protein genes involved in nacre color formation. *Comp Biochem Physiol B Biochem Mol Biol* 204:1–8
- Du Y, Chang H, Yang S, Huang S, Tsai Y, Huang JJ, Chan JCC (2016) Study of binding interaction between Pif80 protein fragment and aragonite. *Sci Rep* 6:30883
- Falini G, Albeck S, Weiner S, Addadi L (1996) Control of aragonite or calcite polymorphism by mollusk shell macromolecules. *Science* 271:67–69
- Ferre F, Clote P (2005) Disulfide connectivity prediction using secondary structure information and diresidue frequencies. *Bioinformatics* 21:2336–2346
- Fu G, Valiyaveetil S, Wopenka B, Morese DE (2005) CaCO_3 biomineralization: acidic 8-kDa proteins isolated from aragonitic abalone shell nacre can specifically modify calcite crystal morphology. *Biomacromolecules* 6:1289–1298
- Giles R, Manne S, Mann S, Morse DE, Stucky GD, Hansma PK (1995) Overgrowth of aragonite on molluscan nacre examined by atomic force microscopy. *Bio Bull* 188:8–15
- Huang X, Madan A (1999) CAP3: a DNA sequence assembly program. *Genome Res* 9:868–877
- Joubert C, Piquemal D, Marie B, Manchon L, Pierrat F, Zanella-Cleon I, Cochenne-Laureau N, Gueguen Y, Montagnani C (2010) Transcriptome and proteome analysis of *Pinctada margaritifera* calcifying mantle and shell: focus on biomineralization. *BMC Genom* 11:613
- Kawabata S, Nagayama R, Hirata M, Shigenaga T, Agarwala KL, Saito T, Cho J, Nakajima H, Takagi T, Iwanaga S (1996) Tachycitin, a small granular component in horseshoe crab hemocytes, is an antimicrobial protein with chitin-binding activity. *J Biochem* 120:1253–1260
- Kroger N (2009) The molecular basis of nacre formation. *Science* 325:1351–1352
- Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 5:150–163
- Liu C, Li S, Kong J, Liu Y, Wang T, Xie L, Zhang R (2015) In-depth proteomic analysis of shell matrix proteins of *Pinctada fucata*. *Sci Rep* 5:17269
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta\text{CT}}$ method. *Methods* 25:402–408
- Mao Y, Jiang B, Peng Q, Liu W, Lin Y, Xie D, He X, Li S (2017) Cloning and characterization of WRKY gene homologs in Chiehqua (*Benincasa hispida* Cogn. var. Chieh-qua How) and their expression in response to fusaric acid treatment. *3 Biotech* 7:86
- Marie B, Joubert C, Tayale A, Zanella-Cleon I, Belliard C, Piquemal D, Cochenne-Laureau N, Marin F, Gueguen Y, Montagnani C (2012) Different secretory repertoires control the biomineralization processes of prism and nacre deposition of the pearl oyster shell. *Proc Natl Acad Sci USA* 109:20986–20991
- Marie B, Arivalagan J, Matheron L, Bolbach G, Berland S, Marie A, Marin F (2017) Deep conservation of bivalve nacre proteins highlighted by shell matrix proteomics of the Unionoida *Elliptio complanata* and *Villosa lienosa*. *J R Soc Interface* 14:20160846
- McGinty EL, Zenger KR, Jones DB, Jerry DR (2012) Transcriptome analysis of biomineralisation-related genes within the pearl sac: host and donor oyster contribution. *Mar Genom* 5:27–33
- Norton JH, Lucas JS, Turner I, Mayer RJ, Newnham R (2000) Approaches to improve cultured pearl formation in *Pinctada margaritifera* through use of relaxation, antiseptic application and incision closure during bead insertion. *Aquaculture* 184:1–17
- Nudelman F (2015) Nacre biomineralisation: a review on the mechanisms of crystal nucleation. *Semin Cell Dev Biol* 46:2–10
- Orme CA, Noy A, Wierzbick A, McBride MT, Grantham M, Teng HH, Dove PM, DeYoreo JJ (2001) Formation of chiral morphologies through selective binding of amino acids to calcite surface steps. *Nature* 411:775–779

- Suzuki M, Saruwatari K, Kogure T, Yamamoto Y, Nishimura T, Kato T, Nagasawa H (2009) An acidic matrix protein, Pif, is a key macromolecule for nacre formation. *Science* 325:1388–1390
- Suzuki M, Iwashima A, Kimura M, Kogure T, Nagasawa H (2013) The molecular evolution of the Pif family proteins in various species of mollusks. *Mar Biotechnol* 15:145–158
- Takahashi K, Yamamoto H, Onoda A, Doi M, Inaba T, Chiba M, Kobayashi A, Taguchi T, Okamura T, Ueyama N (2004) Highly oriented aragonite nanocrystal–biopolymer composites in an aragonite brick of the nacreous layer of *Pinctada fucata*. *Chem Commun* 8:996–997
- Tuckwell D (1999) Evolution of von Willebrand factor A (VWA) domains. *Biochem Soc Trans* 27:835–840
- Wang G, Yuan Y, Li J (2007) SSR analysis of genetic diversity and phylogenetic relationships among different populations of *Hyriopsis cumingii* from the five lakes of China. *J Fish China* 12:12–18
- Wang N, Lee Y, Lee J (2008) Recombinant perlucin nucleates the growth of calcium carbonate crystals: molecular cloning and characterization of perlucin from disk abalone, *Haliotis discus discus*. *Comp Biochem Physiol B Biochem* 149:354–361
- Wang N, Kinoshita S, Riho C, Maeyama K, Nagai K, Watabe S (2009) Quantitative expression analysis of nacreous shell matrix protein genes in the process of pearl biogenesis. *Comp Biochem Physiol B Biochem Mol Biol* 154:346–350
- Watabe S (2013) The importance of total genome databases in research on Akoya pearl oyster. *Zool Sci* 30:781–782
- Whittaker CA, Hynes RO (2002) Distribution and evolution of von Willebrand/integrin A domains: widely dispersed domains with roles in cell adhesion and elsewhere. *Mol Biol Cell* 13:3369–3387
- Yao HB, Ge J, Mao LB, Yan YX, Yu SH (2014) 25th anniversary article: artificial carbonate nanocrystals and layered structural nanocomposites inspired by nacre: synthesis, fabrication and applications. *Adv Mater* 26:163–187
- Zhang J (2007) Disulfide-bond reshuffling in the evolution of an ape placental ribonuclease. *Mol Biol Evol* 24:505–512
- Zhang R, Wang M, Xia N, Yu S, Chen Y, Wang N (2016) Cloning and analysis of gene expression of interleukin-17 homolog in triangle-shell pearl mussel, *Hyriopsis cumingii*, during pearl sac formation. *Fish Shellfish Immunol* 52:151–156
- Zhao M, He M, Huang X, Wang Q (2014) A homeodomain transcription factor gene, PfMSX, activates expression of Pif gene in the pearl oyster *Pinctada fucata*. *PLoS One* 9:e103830
- Zhu W, Fan S, Huang D, Liu B, Bi X, Yu D (2015) Highly expressed EGFR in pearl sac may facilitate the pearl formation in the pearl oyster, *Pinctada fucata*. *Gene* 566:201–211