ELSEVIER

Fish & Shellfish Immunology 22 (2007) 628–640

Fish & Shellfish Immunology

www.elsevier.com/locate/fsi

Pattern of nucleotide substitution and divergence of prophenoloxidase in decapods

Abinash Padhi [a,*], Bindhu Verghese [a], Ajula Vaid [a], Subhendu K. Otta [b]

[a] *Department of Biological Science, University of Tulsa, 600 South College Avenue, Tulsa, OK 74104, USA*
[b] *Department of Pediatrics—Hematology/Oncology, Baylor College of Medicine, 6621 Fannin Street, MC 3-3320, Houston, TX 77030, USA*

Received 21 June 2006; revised 8 August 2006; accepted 8 August 2006
Available online 17 August 2006

Abstract

Despite the unprecedented development in identification and characterization of prophenoloxidase (proPO) in commercially important decapods, little is known about the evolutionary relationship, rate of amino acid replacement and differential selection pressures operating on proPO of different species of decapods. Here we report the evolutionary relationship among these nine decapod species based on proPO gene and types of selective pressures operating on proPO codon sites. Our analyses revealed that all the nine decapod species shared a common ancestor. The mean percentage sequence divergence at proPO gene was $34.4 \pm 0.6\%$. Pairwise estimates of nonsynonymous to synonymous ratio (ω) for *Homarus americanus–H. gammarus* is greater than one, therefore indicating adaptive evolution (functional diversification) of proPO in these two species. In contrast, strong purifying selection ($\omega < 1$) was observed in all other species pairs. However, phylogenetically closely related decapods revealed relatively higher ω value ($\omega = 0.15 \pm 0.3$) than the distantly related species pairs ($\omega = 0.0075 \pm 0.005$). These discrepancies could be due to higher fixation probability of beneficial mutation in closely related species. Maximum likelihood-based codon substitution analyses revealed a strong purifying selection operating on most of the codon sites, therefore suggesting proPO is functionally constrained (purifying selection). Codon substitution analyses have also revealed the evidence of strong purifying selection in haemocyanin subunits of decapods.

Keywords: Decapod; Prophenoloxidase; Haemocyanin; Codon substitution; Nonsynonymous; Synonymous; Genetic distance

1. Introduction

The prophenoloxidase [proPO: zymogen of phenol oxidase (PO; monophenol, L-dopa: oxygen oxidoreductase, EC1.14.18.1)] plays crucial role in innate immunity of invertebrates, for example, sclerotization of arthropod cuticle, pigmentation, wound healing and humoral immune defence [1]. Due to the lack of adaptive immunity in invertebrates, the focus has been towards enhancing innate immune system in many of the cultured species, especially crustaceans [2–7]. Despite the unprecedented development in identification and characterization of proPO (e.g. [2–7]), little is known about the evolutionary relationship of proPO present in different species of arthropods, especially in decapods

* Corresponding author. Tel.: +1 918 631 3062; fax: +1 918 631 2762.
E-mail address: abinash-padhi@utulsa.edu (A. Padhi).

1050-4648/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.fsi.2006.08.010

except the fact that phylogenetically it belongs to the arthropod haemocyanin super-family [8–10]. Although based on its activity proPO is similar in all species, the expression of proPO is species specific [2–7] and influenced by both intrinsic (cellular) and extrinsic environmental factors [11]. Based on this, one might expect that differential selective pressure is operating on the protein-coding region of proPO during the evolutionary time scale. The selective pressure in the protein-coding gene can be measured by comparing the number of nonsynonymous substitutions per nonsynonymous site (d_N) with the number of synonymous substitutions per synonymous site (d_S) [12]. The mutational change that resulted in a change of amino acid is known as nonsynonymous substitution, whereas in synonymous substitution, amino acid remains unchanged even though there is a change in nucleotide. There are three different types of selective pressures that can be detected from d_N and d_S ratio (hereafter referred as ω). If the protein-coding gene is functionally constrained, that is, if most nonsynonymous mutations are deleterious, then the rate of nonsynonymous change will be lower than neutral rate resulting in $\omega < 1$ and the gene is subject to be under strong purifying selection [13]. If nonsynonymous mutations are beneficial, average rate of nonsynonymous changes is expected to be higher than neutral rate, resulting in $\omega > 1$, indicating functional diversification of the gene and subject to positive selection. The evolution of pseudogenes is attributed to the lack of functional constraint on the protein coding genes and is referred to as neutral evolution ($\omega = 1$).

Many of the adaptive and innate immunogenic genes are reported to be under the influence of positive selection (e.g. [14]). It could be possible that selective pressure might also be operating on the entire coding region of proPO. However, if the entire coding region of proPO is not under the influence of positive selection, it is possible that positive selection might be operating on a few codon sites of proPO. The aim of the present study is to investigate the type of selective pressure operating on codon sites of proPO of nine commercially important decapods using maximum likelihood-based codon substitution analyses [15]. From the proPO nucleotide sequence data of these nine decapods, the evolutionary relationship and degree of genetic divergence among these decapods are also estimated using maximum likelihood, Bayesian, maximum parsimony and distance based phylogenetic methods [16–18].

2. Materials and methods

2.1. Phylogenetic analyses

2.1.1. Amino acid phylogeny

To infer evolutionary relationship between arthropod proPO and haemocyanin, we reconstructed maximum likelihood (ML), Bayesian inference (BI) and neighbour joining (NJ) phylogenies based on the amino acid sequence data. A total of 55 amino acid sequences representing Crustacea and Insecta haemocyanin and proPO genes were retrieved from the GenBank ([3–5,19–48], Table 1). All the amino acid sequences were aligned using DAMBE ver. 4.5.2 [49,50]. The unrooted ML and BI amino acid phylogenies were reconstructed using PHYML ver. 2.4.4 [51] and MrBayes 3.04 [16] programs, respectively. Amino acid ML analyses were performed using an input tree generated by BIONJ [52]. The JTT model of sequence evolution was used in both ML and BI analysis. NJ tree was performed with Poisson correction amino acid model using MEGA 2.0 [17]. The nodal support for NJ tree was estimated with 10,000 bootstrap replicates using MEGA 2.0 [17]. PHYML bootstrap trees were constructed using the same parameters as the ML trees. The amino acid BI analyses were performed by running four simultaneous chains for 1.5×10^6 generations and sampling every 1000 generations. All trees below the observed stationary level were discarded, resulting in a "burn-in" of 15,000 generations. As noted above, the fluctuating value of log likelihood was plotted in Tracer ver. 1.3.1 [53] to verify that convergence was reached. The 50% majority-rule consensus tree was used to calculate the posterior probabilities for each node.

2.1.2. proPO nucleotide phylogeny

A total of 11 published proPO and 2 haemocyanin nucleotide sequences representing nine species of decapods were obtained from the GenBank ([3,5,19,20,30,45], Table 1). As described by Burmester [8], haemocyanin sequences were used as out-group in proPO phylogeny. After alignment of all the 11 proPO sequences, a 2085 base pairs (bp) sequence length was produced. When the out-group sequences were included, the total length of all 13 species became 2178 bp as there were several inserts in the haemocyanin gene. All sequences were aligned using MacClade 4.03 [54] and DAMBE ver. 4.5.2 [49,50].

Table 1
Sequences and species used in the study

Protein (DNA) accession number	Code	Species	aa[a]	Source
Crustacea proPO				
AAD45201 (AF099741)	PENMON1	*Penaeus monodon*	688	[5]
AAM77689 (AF521948)	PENMON2	*Penaeus monodon*	686	GenBank
AAM77690 (AF521949)	PENSEM	*Penaeus semisulcatus*	684	GenBank
AAT73697 (AY655139)	HOMAME	*Homarus americanus*	683	GenBank
CAE46724 (AJ581662)	HOMGAM	*Homarus gammarus*	683	[30]
ABA60740 (DQ182596)	MACROS	*Macrobrachium risenbergii*	671	[3]
ABB59713 (DQ230981)	CANMAG	*Cancer magister*	670	[45]
ABD90511 (DQ435606)	SCYSER	*Scylla serrata*	673	GenBank
BAB83773 (AB065371)	MARJAP1	*Marsupenaeus japonicus*	681	[19]
BAB70485 (AB073223)	MARJAP2	*Marsupenaeus Japonicus*	688	GenBank
CAA58471 (X83494)	PACLEN	*Pacifastacus leniusculus*	706	[20]
Insecta proPO				
AAC27383	ANOGAM	*Anopheles gambiae*	683	[48]
AAC34251	HYPCUN1	*Hyphantria cunea*	681	[41]
AAC34256	HYPCUN2	*Hyphantria cunea*	697	[41]
AAC69182	ANOSTE	*Anopheles stephensi*	686	[26]
AAD45526	SARBUL1	*Sarcophaga bullata*	685	[24]
AAD45527	SARBUL2	*Sarcophaga bullata*	691	[24]
AAF70320	ARMSUB	*Armigeres subalbatus*	684	[25]
AAG02218	AEDAEG1	*Aedes aegypti*	685	[44]
AAG02219	AEDAEG2	*Aedes aegypti*	684	[44]
AAG09304	BOMMOR	*Bombyx mori*	685	[44]
AAK64363	GALMEL	*Galleria mellonella*	683	[38]
AAO22166	ANOCUL	*Anopheles culicifacies*	686	GenBank
AAR84669	MUSDOM	*Musca domestica*	684	GenBank
AAU29555	PLOINT	*Plodia interpunctella*	681	[29]
AAW22859	SPOLIT	*Spodoptera litura*	697	GenBank
AAZ52554	HELARM	*Helicoverpa armigera*	698	GenBank
ABC59699	OSTFUR	*Ostrinia furnacalis*	739	GenBank
BAA75470	TENMOL	*Tenebrio molitor*	684	[34]
BAC15602	HOLDIO	*Holotrichia diomphalia*	684	[31]
NP_001011627	APIMEL	*Apis mellifera*	693	[39]
Crustacea haemocyanin				
AAB22190	PANINT-hc3	*Panulirus interruptus*	661	[40]
AAF59175 (DQ230983)	CYASCA-hc	*Cyamus scammoni*	674	[45]
AAF64305 (AF249297)	CALSAP-hc	*Callinectes sapidus*	676	[23]
AAL27460 (AF431737)	PENMON-hc	*Penaeus monodon*	449	[37]
CAB75960 (AJ272095)	HOMAME-hc1	*Homarus americanus*	672	[32]
CAC69243 (AJ344361)	PALVUL-hc1	*Palinurus vulgaris*	684	GenBank
CAC69244 (AJ344362)	PALVUL-hc2	*Palinurus vulgaris*	684	GenBank
CAC69245 (AJ344363)	PALVUL-hc3	*Palinurus vulgaris*	685	GenBank
CAD56697 (AJ516004)	PALELE-hc4	*Palinurus elephas*	685	GenBank
CAI78901 (AJ937836)	GAMROE-hc1	*Gammarus roeseli*	672	[27]
AAO47336 (AY193781)	PACKEN-hc2	*Pacifastacus leniusculus*	687	[35]
AAM81357 (AF522504)	PACLEN-hc	*Pacifastacus leniusculus*	660	[36]
AAW57889 (AY861676)	CANMAG-hc1	*Cancer magister*	662	[46]
AAW57891 (AY861678)	CANMAG-hc3	*Cancer magister*	669	[45]
AAW57892 (AY861679)	CANMAG-hc4	*Cancer magister*	675	[45]
AAW57893 (AY861680)	CANMAG-hc5	*Cancer magister*	676	[45]
AAA96966 (U48881)	CANMAG-hc6	*Cancer magister*	676	[45]
CAA57880 (X82502)	LITVAN-hc	*Litopenaeus vannamei*	662	[43]
Insecta haemocyanin				
AAC16760	SCHAME-hc	*Schistocerca americana*	674	[42]
CAB89497	EURCAL-hc7	*Eurpelma californicum*	629	[47]

Table 1 (*continued*)

Protein (DNA) accession number	Code	Species	aa[a]	Source
CAC44750	CUPSAL-hc2	*Cupiennius salei*	626	[22]
CAC69246	SCUCOL-hc1	*Scutigera coleoptrata*	656	[33]
CAD68053	NEPINA-hc2	*Nephila inaurata madagascariensis*	628	[21]
CAD87763	PERMAR-hc2	*Perla marginata*	671	[28]

Suffix hc indicates haemocyanin sequences.
[a] Amino acid sequence length.

To infer the evolutionary relationship among proPO of nine decapod species, phylogenetic analyses were carried out using ML, BI and maximum parsimony (MP) methods. A best-fit nucleotide substitution model was selected by Akaike Information Criterion (AIC [55]) implemented in Modeltest 3.5 [56]. MP analysis was conducted using the heuristic search option, implementing stepwise addition with 100 random addition replicates and TBR branch swapping using PAUP* 4.0b10 [18]. PHYML ver. 2.4.4 [51] was used for ML analyses and MrBayes 3.04 [16] was used for BI. Nodal support for the MP and ML trees were estimated using 1000 nonparametric bootstrap replicates. MrBayes was used to conduct a Bayesian approach to phylogenetic inference by running 20 million generations (10,000 burn-in) with four Metropolis coupled MCMC to optimize efforts to find peaks in tree-space. Parameters were set to nst = 6 and rates = invgamma and one tree was sampled in every 100. The convergence was checked using Tracer 1.3.1 [53]. The resulting trees were used to generate a majority consensus tree with posterior probability values.

2.2. Sequence divergence

The uncorrected pairwise distances (*p*-distance) among proPO sequences (9 species) were estimated using MEGA 2.0 [17]. Using the same program, standard errors for the distance estimates were calculated with 1000 bootstrap replicates.

2.3. Pattern of nucleotide substitution and tests for selection

2.3.1. proPO

After alignment gaps were removed, a total of 1830 bp was produced that resulted in 610 codons. *P. monodon* and *M. japonicus* each represented by two individuals were included in pairwise comparison. For codon substitution analyses of proPO gene, we chose nine sequences each representing a single species. The inclusion of either of the individuals of *P. monodon* and *M. japonicus* in codon substitution analyses did not alter the results. The two out-group haemocyanin sequences were excluded from the analyses. Pattern of nucleotide substitution and tests for positive selection were carried out using the maximum-likelihood (ML) approach implemented in PAML 3.14b [57]. In order to account for uncertainty regarding the true topology, we repeated the tests for positive selection using the trees from ML, BI and MP analyses.

We performed two different types of analyses; first, we estimated the pairwise d_N and d_S among all the 11 proPO sequences for the entire coding region and for each functional region [Copper binding site-A (Cu-A); Copper binding site-B (Cu-B) and proteolytic activation site] using a maximum-likelihood (ML) approach described by Goldman and Yang [58] implemented in CODEML in the PAML program. Second, to account for among-site variations and to test for positive selection on different codon sites, we estimated parameters under six different codon substitution models [15] and their performances were evaluated using likelihood ratio tests (LRTs). These six codon substitution models are: M0: one-ratio; M1a: nearly neutral; M2a: positive selection; M3: discrete; M7: beta; and M8: beta + $\omega > 1$: continuous [15]. The LRTs between nested models were conducted by comparing twice the difference in log-likelihood values ($2\ln\Delta l$) against a χ^2-distribution, with degrees of freedom (df) equal to the difference in the number of parameters between models [15]. Four likelihood ratio tests (LRTs) were conducted. First comparison was made between M0, a model that fits a single ω for all sites with M2a which allows three site classes ($0 < \omega < 1$, $\omega = 1$ or $\omega > 1$). The second comparison was between M0 and M3. Third comparison was between M1a, which allows for two site classes ($0 < \omega < 1$, $\omega = 1$) with M2a. The last comparison was between a model of beta distributed selective pressures which allows for 10 site classes, each with $\omega < 1$ (M7) and M8, which has 11 site classes, one of them allowed for $\omega > 1$.

2.3.2. Haemocyanin

To know the types of selective pressure operating on the different subunits of decapod haemocyanin, we performed both pairwise comparison and codon substitution analyses as described for proPO. A total of 17 decapod haemocyanin nucleotide coding sequences [23,27,32,35–37,40,43,45,46] representing 10 species and different subunits were used in the analyses (Table 1).

3. Results

3.1. Phylogenetic analyses

3.1.1. Amino acid phylogeny

The phylogenetic relationship between different groups of proPO and haemocyanin is shown in Fig. 1. The amino acid phylogeny revealed that proPO from Insecta and Crustacea are closely related and their relationship is strongly supported. Insecta and Crustacea proPO each form a separate monophyletic group (Fig.1), therefore allowing us to perform codon substitution analyses for the group in question. The amino acid phylogeny also demonstrated that Crustacea haemocyanin are not closely related to Crustacea proPO, but ancestral to both Insecta and Crustacea proPO.

3.1.2. proPO nucleotide phylogeny

The general time reversible model (GTR) with invariable site model ($I = 0.1126$) and gamma distribution shape parameter G ($= 1.6455$) was the best-fit model selected by AIC. The transition to transversion ratio (T_i/T_v) and the log likelihood score ($-\ln L$) were 1.434 and 19412.7363, respectively. The substitution rate matrix for the data was A ↔ C = 1.8692, A ↔ G = 3.3549, A ↔ T = 1.7322, C ↔ G = 1.4542, C ↔ T = 3.8913, G ↔ T = 1.00. The nucleotide base frequencies for A, C, G and T were 0.2497, 0.2868, 0.2542 and 0.2094, respectively. The inferred proPO phylogenies among nine taxa are shown in Fig. 2A and B. The proPO phylogeny (proPO gene tree) revealed sister-group relationship between *Penaeus monodon* and *Penaeus semisulcatus*, indicating recent divergence at proPO gene. *P. monodon–P. semisulcatus* clade and *Marsupenaeus japonicus* diverged from a common ancestral proPO. Both *Homarus* species showed sister-group relationship with *Pacifastacus leniusculus*. Similarly, *Scylla serrata* and *Cancer magister* showed close relationship in their proPO gene with strong nodal support. Among decapod proPO, *Macrobrachium rosenbergii* emerged from the base of proPO phylogeny; however, the relationships among proPO genes of all nine decapods was strongly supported.

3.2. Sequence divergence

The mean percentage of sequence divergence among nine species at the proPO is 34.4 ± 0.6%. The pairwise uncorrected distance among all nine species are given in Table 2.

3.3. Pattern of nucleotide substitution and tests for selection

3.3.1. ProPO

Pairwise estimates of d_N and d_S among eleven proPO sequences for the entire coding region revealed that the nonsynonymous rates ranged from 0.022 to 0.38, whereas the synonymous rates ranged from 0.018 to 78.63 (Fig. 3A). Pairwise ω-ratios ($\omega = 1.2$) between *H. gammarus* and *H. americanus* indicated evidence of positive selection on the proPO gene of these two species. However, pairwise comparisons among the remaining species indicated strong purifying selection (Fig. 3B). Pairwise comparisons within genus *Penaeus*, between genus *Penaeus* and *Marsupenaeus*, between genus *Scylla* and *Cancer* and between genus *Homarus* and *Pacifastacus* indicated elevated ω ratios (ranged from 0.11 to 0.19) with mean $\omega = 0.15 \pm 0.3$, whereas remaining species pairs indicated reduced ω ratios (ranged from 0.003 to 0.025) with mean $\omega = 0.0075 \pm 0.005$ (Fig. 3B). Distribution of pairwise estimates of ω for Cu-A, Cu-B and proteolytic activation sites of proPO are shown in Fig. 4A, B and C, respectively. Pairwise ω-ratios ($\omega = 1.36$) between *H. gammarus* and *H. americanus* indicated evidence of positive selection on the proteolytic activation site of proPO in these two species. Although Cu-A and Cu-B indicated evidence of purifying selection ($\omega < 1$), the elevated ω values (Cu-A: $\omega = 0.676$; and Cu-B:$\omega = 0.814$) for this species pair indicated less functionally constrained (reduced level of purifying selection) than the remaining species pairs (Fig. 4A and B).

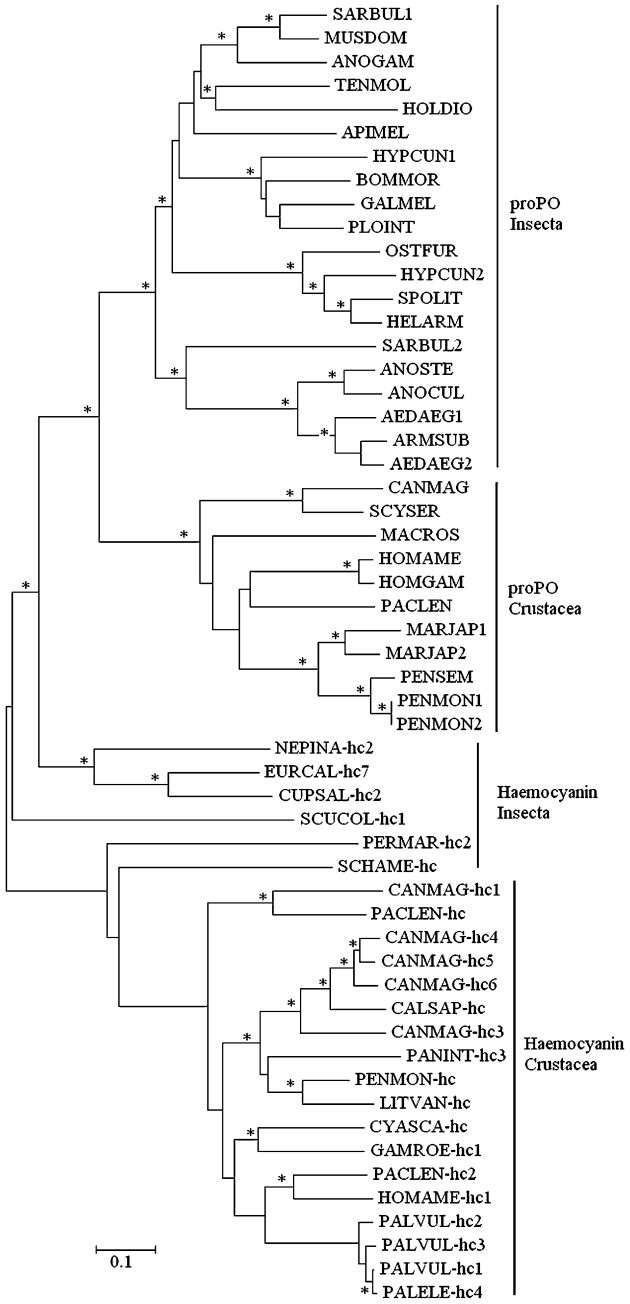

Fig. 1. Phylogenetic relationship between proPO and haemocyanin genes of Insecta and Crustacea based on neighbour joining tree inferred from the amino acid sequence data. Nodal support (ML/BI/NJ) ≥ 80 is indicated by asterisk. See Table 1 for taxon abbreviations.

The model parameters for ω estimated using MP, ML and BI trees were identical. Parameter estimates and log likelihood values under models of variable ω ratios among codon sites of proPO and their likelihood ratio statistic (LRTs) are shown in Tables 3A and 4A, respectively. Although M0−M3 and M0−M2a indicated significant among-site variation ($p = 0$), the LRTs of more stringent models (M1a vs. M2a and M7 vs. M8) failed to support positive selection on codon sites of proPO. However, M8 model detected 11 codon sites (of the total 610) that are positively selected with posterior probabilities greater than >50%, but not significant at 95%.

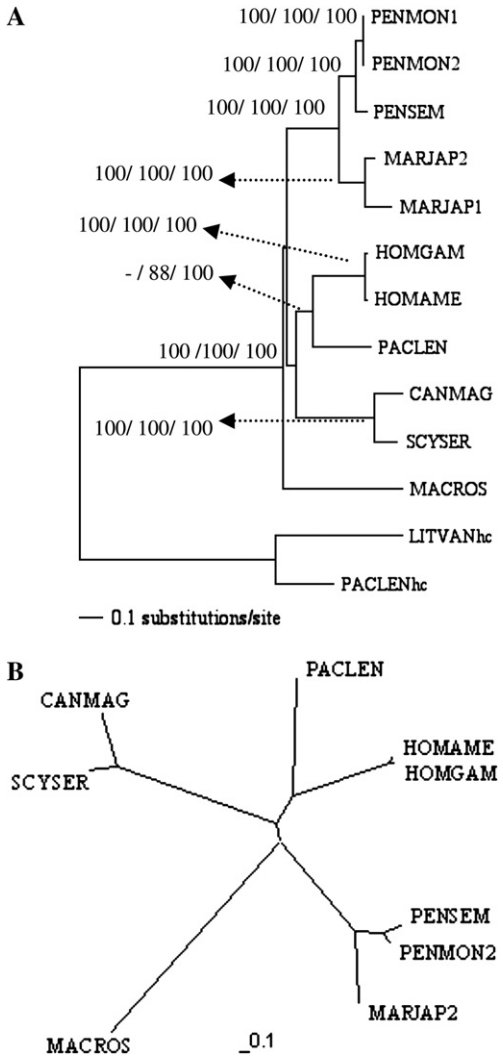

Fig. 2. Phylogenetic relationship among decapod taxa inferred from proPO nucleotide sequence data. (A) ML tree rooted with haemocyanin sequence. Nodal supports (ML/BI/MP) are given. (B) Unrooted phylogram inferred from proPO sequences representing nine species used for codon substitution analyses using PAML.

Table 2
Uncorrected pairwise distances among nine species of decapods analysed in this study

	Species	1	2	3	4	5	6	7	8	9
1.	MARJAP2		0.008	0.009	0.01	0.01	0.011	0.011	0.011	0.010
2.	PENMON2	0.188		0.006	0.009	0.009	0.01	0.01	0.010	0.010
3.	PENSEM	0.199	0.071		0.01	0.01	0.01	0.01	0.010	0.010
4.	HOMGAM	0.378	0.355	0.353		0.003	0.011	0.011	0.010	0.010
5.	HOMAME	0.379	0.355	0.353	0.021		0.01	0.01	0.010	0.010
6.	CANMAG	0.392	0.392	0.386	0.383	0.382		0.008	0.010	0.010
7.	SCYSER	0.4	0.393	0.39	0.37	0.365	0.183		0.010	0.011
8.	PACLEN	0.377	0.353	0.362	0.303	0.3	0.383	0.373		0.011
9.	MACROS	0.402	0.398	0.396	0.396	0.401	0.418	0.414	0.406	

Lower diagonal values are distances and upper diagonal values are standard errors estimated with 1000 bootstrap replicates.

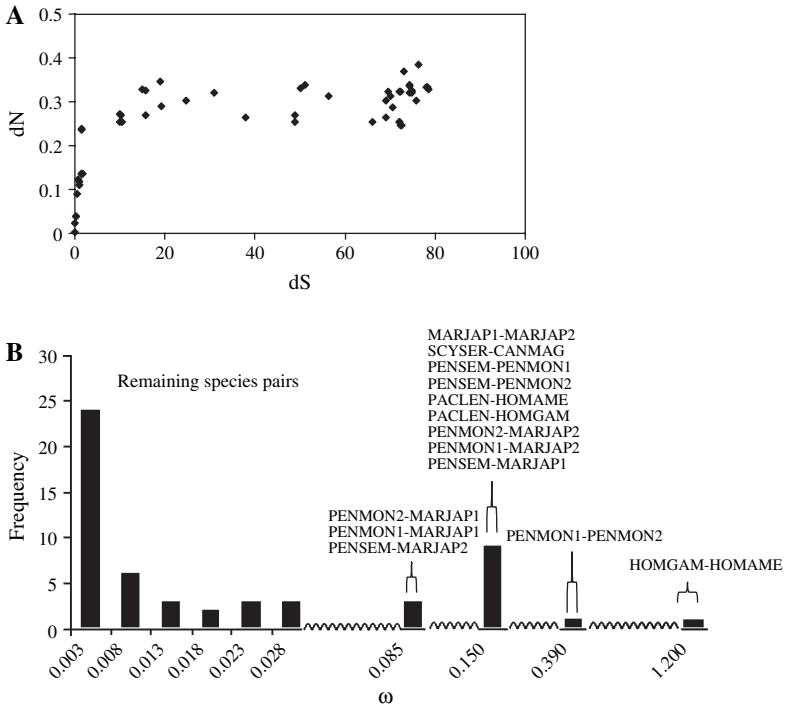

Fig. 3. (A) Estimates of d_N and d_S inferred from pairwise comparison among proPO gene. (B) Distribution of pairwise estimates of ω-ratios among proPO gene inferred from 11 taxa using maximum likelihood approach [58] implemented in PAML [57].

3.3.2. Haemocyanin

Pairwise estimates of d_N and d_S among 17 haemocyanin sequences for the entire coding region revealed that the nonsynonymous rates ranged from 0.0069 to 0.4399, whereas the synonymous rates ranged from 0.0078 to 86.6702 (Fig. 5A). Pairwise ω-ratios ($\omega = 1.25$) between *Palinurus elephas*-hc4 and *P. vulgaris*-hc3 indicated evidence of positive selection on the haemocyanin subunits of this species pair. However, pairwise comparisons among the remaining species indicated strong purifying selection (Fig. 5B).

Parameter estimates and log likelihood values under models of variable ω ratios among codon sites of haemocyanin and their likelihood ratio statistic (LRTs) are shown in Tables 3B and 4B, respectively. M0–M3 and M0–M2a showed significant among-site variation ($p = 0$). Although the LRT of M1a vs. M2a indicated evidence of positive selection across the codon sites of haemocyanin subunits, LRT of more stringent models (M7 vs. M8) failed to support positive selection on codon sites of haemocyanin.

4. Discussion

The topological placement of genus *Penaeus* and *Marsupenaeus* in proPO phylogeny is consistent with mitochondrial DNA phylogeny [59,60]. However, the topological placement of penaeids and *Macrobrachium* in proPO gene is not consistent with the gene tree inferred from the combined data of 16S, 18S, 28S and histone (H3) genes [61]. This discrepancy could be associated with the differential evolutionary rates of each gene in different species. In contrast, the divergence of two *Homarus* species in proPO phylogeny is consistent with their morphological classification as well as with the phylogenetic analyses of 16S, 18S, 28S and H3 genes [61]. Similarly, the sister-group relationship between *Homarus* (lobster) and *Pacifastacus* (Cray fish) in proPO gene tree is consistent with their topological placement based on morphology (species tree) and molecules (gene tree) [61]. Both crab species (*Scylla* and *Cancer*) also revealed a similar sister-group relationship in proPO gene tree. The tree revealed that proPO for all the nine decapod species analysed in this study diverged from a common ancestor. Despite the ambiguities in divergence time and evolutionary relationship between proPO and haemocyanin [8,10], it is unequivocal that both genes belong to the arthropod haemocyanin super-family.

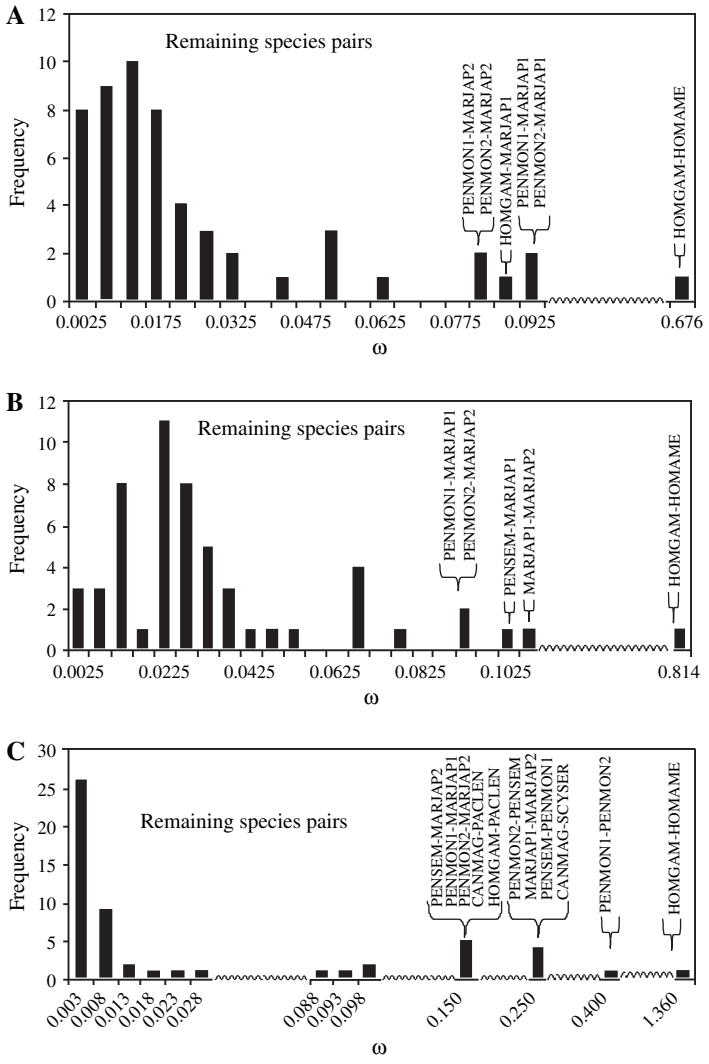

Fig. 4. Distribution of pairwise estimates of ω-ratios for different regions of proPO gene inferred from 11 taxa using maximum likelihood approach [58] implemented in PAML [57]. (A) Cu-binding site A; (B) Cu-binding site B; (C) proteolytic activation site.

From the phylogenetic analyses of proPO, it appeared that all decapods share a common ancestor. However, it is unclear whether this evolutionary relationship is strongly influenced by any types of selection pressure across the coding region of proPO—in other words, whether the divergence that we see among decapods is functionally constrained or diversified. Differential expression of proPO in different species [2–7] of decapods could either be due to adaptive evolution (rapid fixation of beneficial mutations; d_N is expected to be higher than d_S) or it could be by chance due to the stochastic nature of the mutational process (d_S is expected to be higher than d_N). In the latter case, the probability of deleterious mutation being wiped out from proPO-coding region is higher than that of fixation of beneficial mutations and proPO might be accumulating more silent substitutions than that of nonsynonymous substitutions and evolving in a nearly neutral fashion [13]. With the exception of *gammarus–americanus* species pairs, all the other species pairs are shown to have $\omega < 1$, indicating strong purifying selection. The greater ω value ($\omega > 1$) between two species of *Homarus* indicated that probabilities of fixation of advantageous mutations are higher than the wiping out of the deleterious mutations. The possible explanations for such a scenario are: (1) these two species diverged recently from their common ancestor and have relatively small effective population size (N_e), therefore the impact of genetic drift is more significant in fixation of advantageous mutations with a probability (P) greater than $1/2N_e$ ($P = 1/2N_e$, if

Table 3
Parameter estimates and log-likelihood values under models of variable ω-ratios among sites: (A) proPO; (B) haemocyanin

Model	Free parameters	Parameter estimates	Likelihood scores	Positively selected sites[a]
A. proPO				
M0: one-ratio	1	$\omega = 0.1228$	−12531.68266	
M1a: nearly neutral	1	$\omega = 0.088$, $\omega_1 = 1$, ($p_0 = 0.77$, $p_1 = 0.23$)	−12322.95109	
M2a: positive selection	3	$\omega_0 = 0.088$, $\omega_1 = 1$, $\omega_2 = 1$, ($p_0 = 0.77$, $p_1 = 0.17$, $p_2 = 0.063$)	−12322.95109	
M3: discrete	5	$\omega_0 = 0.014$, $\omega_1 = 0.17$, $\omega_2 = 0.58$, ($p_0 = 0.41$, $p_1 = 0.43$, $p_2 = 0.16$)	−12199.00992	
M7: beta	2	$p = 0.48$, $q = 2.28$	−12199.99655	
M8: beta + $\omega_s > 1$	4	$p_0 = 0.99$, $p_1 = 0.01$, $p = 0.5$, $q = 2.54$, $\omega = 1.7$	−12199.03585	16, 34, 38, 88, 110, 131, 248, 407, 437, 479, 530
B. Haemocyanin				
M0: one-ratio	1	$\omega = 0.1356$	−18001.68746	
M1a: nearly neutral	1	$\omega_0 = 0.083$, $\omega_1 = 1$, ($p_0 = 0.72$, $p_1 = 0.28$)	−17500.7268	
M2a: positive selection	3	$\omega_0 = 0.083$, $\omega_1 = 1$, $\omega_2 = 8.6$; ($p_1 = 0.72$, $p_1 = 0.275$, $p_2 = 0.005$)	−17493.8713	413**, 510, 513**, 515, 535
M3: discrete	5	$\omega_0 = 0.087$, $\omega_1 = 0.138$, $\omega_2 = 0.525$, ($p_0 = 0.38$, $p_1 = 0.38$, $p_2 = 0.24$)	−17275.85504	
M7: beta	2	$p = 0.397$, $q = 1.74$	−17269.2122	
M8: beta + $\omega_s > 1$	4	$p_0 = 0.996$, $p_1 = 0.004$, $p = 0.397$, $q = 1.74$, $\omega = 305.19$	−17269.41499	413*, 444, 482, 510, 513*, 515, 534, 545, 629

*Sites with posterior probability ≥95%, **sites with posterior probability ≥99%.
[a] Sites with a posterior probability >50% of having $\omega > 1$.

$\omega = 1$); (2) the two species apparently diverged very recently from a small effective population size and accumulated more number of advantageous mutations [13]; or (3) both explanations are also possible. Nevertheless, the high ω-ratio ($\omega > 1$) between *H. americanus* and *H. gammarus* clearly indicates positive selection is operating and therefore indicates rapid functional diversification (adaptive evolution) of proPO gene in these two species. However, comparison between phylogenetically more closely related species revealed a relatively higher ω value, whereas comparison between distantly related species revealed a very low ω value. Although both groups indicated strong purifying selection operating on proPO coding regions, discrepancies between the two groups could be due to the differential rate of fixation of beneficial mutations and wiping of deleterious mutations [13]. Therefore, during the evolutionary time scale more synonymous changes accumulated than the nonsynonymous changes.

From the pairwise estimates of ω ratio, it is difficult to detect whether positive selection is operating on a few codon sites when more than 90% of codon sites are under the influence of strong purifying selection (for example, our data show that most of the species pairs revealed a strong purifying selection on the proPO gene). In such a situation,

Table 4
Likelihood ratio statistics (LRTs) among different model given in Table 3A and B

Comparison	$2\ln\Delta l$	df	p
A. proPO			
M0 vs. M2a	417.463	2	0.00
M0 vs. M3	665.345	4	0.00
M1a vs. M2a	0	2	1.00
M7 vs. M8	1.9214	2	0.38
B. Haemocyanin			
M0 vs. M2a	1015.6323	2	0.000
M0 vs. M3	1451.6649	4	0.000
M1a vs. M2a	13.710994	2	0.001
M7 vs. M8	0.40559	2	0.816

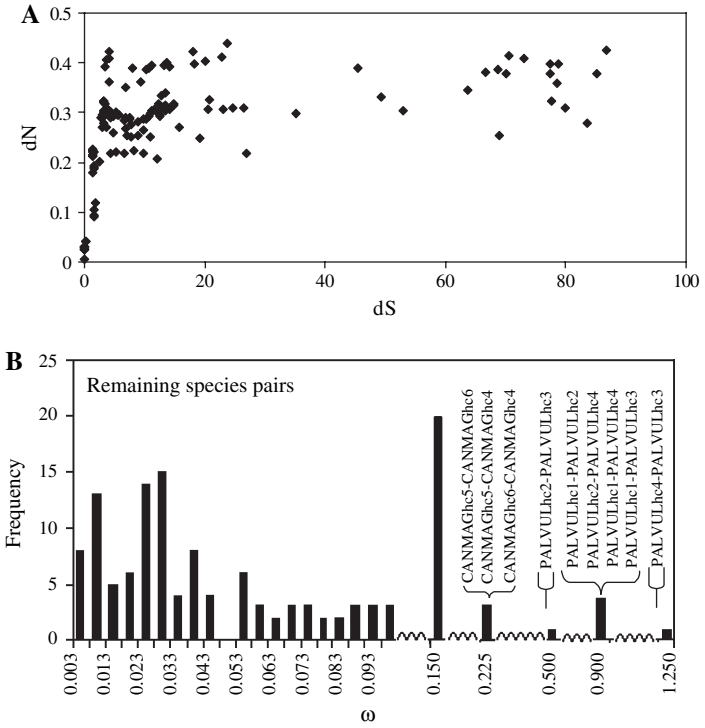

Fig. 5. (A) Estimates of d_N and d_S inferred from pairwise comparison among haemocyanin sequences. (B) Distribution of pairwise estimates of ω-ratios among haemocyanin genes inferred from 17 taxa using maximum likelihood approach [58] implemented in PAML [57].

pairwise estimates of the ω value will obviously be <1 [62] and differential selection pressure operating on each codon site is underestimated. Maximum likelihood based analyses under different models of codon substitutions is more powerful in detecting positive selection among codon sites [15]. Although M0−M3 and M0−M2a supported the fact of variation among codon sites of proPO gene, this is not supported by M1a−M2a. The M8 model detected 1.8% of codon sites that are positively selected with posterior probabilities greater than 50%, but not significant at 95%. However, this result is supported neither by M1a−M2a (nearly neutral vs. positively selected model) nor by a more stringent model (M7 vs. M8), which is more powerful in detecting positive selection. These lines of evidence suggested no sign of a significant level of positive selection on proPO genes of decapods. Thus, the proPO coding region is functionally constrained and evolving in a nearly neutral fashion.

Consistent with the results of proPO, our results demonstrated that the coding regions of haemocyanin subunits are also under the influence of strong purifying selection. The different subunits of haemocyanin that might have evolved by gene duplication events appeared to be functionally more constrained.

Acknowledgements

We thank the University of Tulsa for providing facilities to carry out this work. We thank two anonymous reviewers for their suggestions on the earlier version of the manuscript.

References

[1] Cerenius L, Söderhäll K. The prophenoloxidase-activating system in invertebrates. Immunol Rev 2004;198:116–26.
[2] Lai CY, Cheng W, Kuo CM. Molecular cloning and characterisation of prophenoloxidase from haemocytes of the white shrimp, *Litopenaeus vannamei*. Fish Shellfish Immunol 2005;18:417–30.
[3] Liu CH, Tseng DY, Lai CY, Cheng W, Kuo CM. Molecular cloning and characterisation of prophenoloxidase cDNA from haemocytes of the giant freshwater prawn, *Macrobrachium rosenbergii*, and its transcription in relation with the moult stage. Fish Shellfish Immunol 2006;21:60–9.

[4] Lu KY, Huang YT, Lee HH, Sung HH. Cloning the prophenoloxidase cDNA and monitoring the expression of proPO mRNA in prawns (*Macrobrachium rosenbergii*) stimulated in vivo by CpG oligodeoxynucleotides. Fish Shellfish Immunol 2006;20:274–84.
[5] Sritunyalucksana K, Cerenius L, Soderhall K. Molecular cloning and characterization of prophenoloxidase in the black tiger shrimp, *Penaeus monodon*. Dev Comp Immunol 1999;23:179–86.
[6] Sung HH, Chang HJ, Her CH, Chang JC, Song YL. Phenoloxidase activity of hemocytes derived from *Penaeus monodon* and *Macrobrachium rosenbergii*. J Invertebr Pathol 1998;71:26–33.
[7] Supungul P, Klinbunga S, Pichyanngkura R, Jitrapakdee S, Hirono I, Aoki T, et al. Identification of immune-related genes in hemocytes of black tiger shrimp (*Penaeus monodon*). Mar Biotechnol 2002;4:487–94.
[8] Burmester T. Molecular evolution of the arthropod hemocyanin superfamily. Mol Biol Evol 2001;18:184–95.
[9] Fujimoto K, Okino N, Kawabata S-I, Iwanaga S, Ohnishi E. Nucleotide sequence of the cDNA encoding the proenzyme of the phenoloxidase $A_1$of *Drosophila melanogaster*. Proc Natl Acad Sci USA 1995;92:7769–73.
[10] Hughes AL. Evolution of the arthropod prophenoloxidase/hexamerin protein family. Immunogenetics 1999;49:106–14.
[11] Le Moullac G, Haffner P. Environmental factors affecting immune responses in crustacean. Aquaculture 2000;191:121–31.
[12] Kimura M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. Nature 1977;267:275–6.
[13] Kimura M. The neutral theory of molecular evolution. New York: Cambridge University press; 1983.
[14] Ford MJ. Applications of selective neutrality tests to molecular ecology. Mol Ecol 2002;11:1245–62.
[15] Yang Z, Nielson R, Goldman N, Pederson AMK. Codon substitution models for heterogeneous selection pressure at amino acid sites. Genetics 2000;155:431–49.
[16] Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 2001;17:754–5.
[17] Kumar S, Tamura K, Jakobsen IB, Nei M. MEGA2: molecular evolutionary genetics analyses software. Bioinformatics 2001;17:1244–5.
[18] PAUP*, Swofford DL. Phylogenetic Analysis Using Parsimony (and other methods) ver. 4.0.10b. Sunderland, MA: Sinauer Associates; 2002.
[19] Adachi K, Hirata Y, Nagai K, Fujio A, Sakaguchi M. Hemocyanin-related reactions induce blackening of freeze-thawed prawn during storage. In: Sakaguchi M, editor. More efficient utilization of fish and fisheries products. Amsterdam: Elsevier; 2004. p. 317–30.
[20] Aspan A, Huang TS, Cerenius L, Soderhall K. cDNA cloning of prophenoloxidase from the freshwater crayfish *Pacifastacus leniusculus* and its activation. Proc Natl Acad Sci USA 1995;92:939–43.
[21] Averdam A, Markl J, Burmester T. Subunit sequences of the 4 x 6-mer hemocyanin from the golden orb-web spider, *Nephila inaurata*. Eur J Biochem 2003;270:3432–9.
[22] Ballweber P, Markl J, Burmester T. Complete hemocyanin subunit sequences of the hunting spider *Cupiennius salei*: recent hemocyanin remodeling in entelegyne spiders. J Biol Chem 2002;277:14451–7.
[23] Brouwer M, Syring R, Hoexum Brouwer T. Role of a copper-specific metallothionein of the blue crab, *Callinectes sapidus*, in copper metabolism associated with degradation and synthesis of hemocyanin. J Inorg Biochem 2002;88:228–39.
[24] Chase MR, Raina K, Bruno J, Sugumaran M. Purification, characterization and molecular cloning of prophenoloxidases from *Sarcophaga bullata*. Insect Biochem Mol Biol 2000;30:953–67.
[25] Cho WL, Liu HS, Lee CH, Kuo CC, Chang TY, Liu CT, et al. Molecular cloning, characterization and tissue expression of prophenoloxidase cDNA from the mosquito *Armigeres subalbatus* inoculated with *Dirofilaria immitis* microfilariae. Insect Mol Biol 1998;7:31–40.
[26] Cui L, Luckhart S, Rosenberg R. Molecular characterization of a prophenoloxidase cDNA from the malaria mosquito *Anopheles stephensi*. Insect Mol Biol 2000;9:127–37.
[27] Hagner-Holler S, Kusche K, Hembach A, Burmester T. Biochemical and molecular characterisation of hemocyanin from the amphipod *Gammarus roeseli*: complex pattern of hemocyanin subunit evolution in Crustacea. J Comp Physiol B, Biochem Syst Environ Physiol 2005;175:445–52.
[28] Hagner-Holler S, Schoen A, Erker W, Marden JH, Rupprecht R, Decker H, et al. A respiratory hemocyanin from an insect. Proc Natl Acad Sci USA 2004;101:871–4.
[29] Hartzer KL, Zhu KY, Baker JE. Phenoloxidase in larvae of *Plodia interpunctella* (Lepidoptera: Pyralidae): Molecular cloning of the proenzyme cDNA and enzyme activity in larvae paralyzed and parasitized by *Habrobracon hebetor* (Hymenoptera: Braconidae). Arch Insect Biochem Physiol 2005;59:67–79.
[30] Hauton C, Hammond JA, Smith VJ. Real-time PCR quantification of the in vitro effects of crustacean immunostimulants on gene expression in lobster (*Homarus gammarus*) granular haemocytes. Dev Comp Immunol 2005;29:33–42.
[31] Kim MS, Baek MJ, Lee MH, Park JW, Lee SY, Soderhall K, et al. A new ester-type serine protease cleaves a masquerade-like protein during prophenoloxidase activation in *Holotrichia diomphalia* larvae. J Biol Chem 2002;277:39999–40004.
[32] Kusche K, Burmester T. Molecular cloning and evolution of lobster hemocyanin. Biochem Biophys Res Commun 2001;282:887–92.
[33] Kusche K, Hembach A, Hagner-Holler S, Gebauer W, Burmester T. Complete subunit sequences, structure and evolution of the 6 x 6-mer hemocyanin from the common house centipede, *Scutigera coleoptrata*. Eur J Biochem 2003;270:2860–8.
[34] Lee HS, Cho MY, Lee KM, Kwon TH, Homma K, Natori S, et al. The pro-phenoloxidase of coleopteran insect, *Tenebrio molitor*, larvae was activated during cell clump/cell adhesion of insect cellular defense reactions. FEBS Lett 1999;444:255–9.
[35] Lee SY, Lee BL, Soderhall K. Processing of crayfish hemocyanin subunits into phenoloxidase. Biochem Biophys Res Commun 2004;322:490–6.
[36] Lee SY, Lee BL, Soderhall K. Processing of an antibacterial peptide from hemocyanin of the freshwater crayfish *Pacifastacus leniusculus*. J Biol Chem 2003;278:7927–33.
[37] Lehnert S, Johnson S. Expression of hemocyanin and digestive enzyme messenger RNAs in the hepatopancreas of the Black Tiger Shrimp *Penaeus monodon*. Comp Biochem Physiol B, Biochem Mol Biol 2002;133:163.
[38] Li D, Scherfer C, Korayem AM, Zhao Z, Schmidt O, Theopold U. Insect hemolymph clotting: evidence for interaction between the coagulation system and the prophenoloxidase activating cascade. Insect Biochem Mol Biol 2002;32:919–28.

[39] Lourenco AP, Zufelato MS, MM B, Simoes ZL. Molecular characterization of a cDNA encoding prophenoloxidase and its expression in *Apis mellifera*. Insect Biochem Mol Biol 2005;35:541–52.
[40] Neuteboom B, Jekel PA, Beintema JJ. Primary structure of hemocyanin subunit c from *Panulirus* interruptus. Eur J Biochem 1992; 206:243–9.
[41] Park DS, Shin SW, Kim MG, Park SS, Lee WJ, Brey PT, et al. Isolation and characterization of the cDNA encoding the prophenoloxidase of fall webworm, *Hyphantria cunea*. Insect Biochem Mol Biol 1997;27:983–92.
[42] Sanchez D, Ganfornina MD, Gutierrez G, Bastiani MJ. Molecular characterization and phylogenetic relationships of a protein with potential oxygen-binding capabilities in the grasshopper embryo. A hemocyanin in insects? Mol Biol Evol 1998;15:415–26.
[43] Sellos D, Lemoine S, Van Wormhoudt A. Molecular cloning of hemocyanin cDNA from *Penaeus vannamei* (Crustacea, Decapoda): structure, evolution and physiological aspects. FEBS Lett 1997;407:153–8.
[44] Taft AS, Chen CC, Li J, Christensen BM. Molecular cloning of two prophenoloxidase genes from the mosquito *Aedes aegypti*. Insect Mol Biol 2001;10:97–103.
[45] Terwilliger NB, Ryan MC. Functional and phylogenetic analyses of phenoloxidases from Brachyuran (*Cancer magister*) and Branchiopod (*Artemia franciscana, Triops longicaudatus*) crustaceans. Biol Bull 2006;210:38–50.
[46] Terwilliger NB, Ryan MC, Towle D. Evolution of novel functions: cryptocyanin helps build new exoskeleton in *Cancer magister*. J Exp Biol 2005;208:2467–74.
[47] Voit R, Feldmaier-Fuchs G, Schweikardt T, Decker H, Burmester T. Complete sequence of the 24-mer hemocyanin of the tarantula *Eurypelma californicum*. Structure and intramolecular evolution of the subunits. J Biol Chem 2000;275:39339–44.
[48] Lee WJ, Ahmed A, della Torre A, Kobayashi A, Ashida M, Brey PT. Molecular cloning and chromosomal localization of a prophenoloxidase cDNA from the malaria vector *Anopheles gambiae*. Insect Mol Biol 1998;7:41–50.
[49] Xia X. Data analysis in molecular biology and evolution. Boston: Kluwer; 2000. p. 276.
[50] Xia X, Xie Z. DAMBE: Data analysis in molecular biology and evolution. Journal of Heredity 2001;92:371–3.
[51] Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 2003;52:696–704.
[52] Gascuel O. BIONJ: An improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol 1997;14:685–95.
[53] Rambout A, Drummond AJ. Tracer Version 1.3.1. Available from, http://evolve.zoo.ox.ac.uk/2003.
[54] Maddison WP, Maddison D. MacClade. Sunderland, MA: Sinauer Associates; 2001.
[55] Huelsenbeck JP, Crandall KA. Phylogeny estimation and hypothesis testing using maximum likelihood. Annu Rev Ecol Syst 1997;28: 437–66.
[56] Posada D, Crandall KA. Modeltest: testing the model of DNA substitution. Bioinformatics 1998;14:817–8.
[57] Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci 1997;13:555–6.
[58] Goldman N, Yang Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol 1994;11:725–36.
[59] Baldwin JD, Bass AL, Bowen BW, Clark Jr WH. Molecular phylogeny and biogeography of the marine shrimp *Penaeus*. Mol Phylogenet Evol 1998;10:399–407.
[60] Lavery S, Chan TY, Tam YK, Chu KH. Phylogenetic relationships and evolutionary history of the shrimp genus *Penaeus s.l.* derived from mitochondrial DNA. Mol Phylogenet Evol 2004;31:39–49.
[61] Porter ML, Pérez-Losada M, Crandall KA. Model-based multi-locus estimation of decapod phylogeny and divergence times. Mol Phylogenet Evol 2005;37:355–69.
[62] Jansa SA, Lundrigan BL, Tucker PK. Tests for positive selection on immune and reproductive genes in closely related species of the Murine genus *Mus*. J Mol Evol 2003;56:294–307.