ELSEVIER

Fish & Shellfish Immunology 23 (2007) 553–566

Fish & Shellfish Immunology

www.elsevier.com/locate/fsi

Adaptive evolution after duplication of penaeidin antimicrobial peptides

Abinash Padhi [a,*], Bindhu Verghese [a], Subhendu K. Otta [b], Binu Varghese [c], Karri Ramu [d]

[a] *Department of Biological Science, University of Tulsa, 600 S. College Ave., Tulsa, OK 74104, USA*
[b] *Baylor College of Medicine, Department of Pediatrics-Hematology/Oncology, 6621 Fannin Street, MC 3-3320, Houston, TX 77030, USA*
[c] *Central Marine Fisheries Research Institute, Vizhinjam Research Centre, Trivandrum, Kerala, India*
[d] *Center for Marine Environmental Studies, Ehime University, Bunkyo Cho 2–5, Matsuyama 790-8577, Japan*

Received 18 September 2006; revised 19 December 2006; accepted 12 January 2007
Available online 20 January 2007

Abstract

Penaeidin antimicrobial peptides in penaeid shrimps are an important component of their innate immune system that provides immunity against infection caused by several gram-positive bacteria and filamentous fungal species. Despite the knowledge on the identification and characterization of these peptides in penaeid shrimps, little is known about the evolutionary pattern of these peptides and the underlying genetic mechanisms that maintain high sequence diversities in the penaeidin gene family. Based on the phylogenetic analyses and maximum likelihood-based codon substitution analyses, here we present the convincing evidence that multiple copies of penaeidins have evolved by gene duplication, and positive Darwinian selection (adaptive evolution) is the likely cause of accelerated rate of amino acid substitutions among these duplicated genes.

While the average ratio of non-synonymous to synonymous substitutions (ω) for the entire coding region of both active domains is 0.9805, few codon sites showed significantly higher ω (3.73). The likelihood ratio tests that compare models incorporating positive selection ($\omega > 1$) at certain codon sites with models not incorporating positive selection ($\omega < 1$), failed to reject ($p = 0$) the evidence of positive Darwinian selection. The rapid adaptive evolution of this gene family might be directed by the pathogens and the faster rate of amino acid substitutions in the N-terminal proline-rich and C-terminal cysteine-rich domains could be due to their direct involvement in the protection against pathogens. When the host expose to different habitats/environment an accelerated rate of amino acid substitutions in both the active domains may also be expected.

Keywords: Antimicrobial peptides; Penaeidin; Gene duplication; Positive Darwinian selection; Penaeids

1. Introduction

Antimicrobial peptides (AMP) are an important component of the innate immune system present in all phyla of the living kingdom and provide immunity against infection caused by pathogens [1,2]. AMPs have been identified and characterized in many vertebrates [3], invertebrates [1] and plants [4]. Due to the lack of adaptive immunity,

* 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 © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.fsi.2007.01.003

invertebrates are dependent on the innate immune system for protection from pathogens [5,6]. The devastating disease outbreaks in many commercially important penaeid shrimps has caused serious setback in the aquaculture industry, therefore, the focus has been now towards a better understanding of penaeid shrimp immunity, which would help in designing efficient strategies for disease control [5]. The identification and characterization of AMPs in commercially important penaeid shrimps is now one of the active areas of research (e.g. refs. [7–25]). Three kinds of AMP have been fully characterized in penaeid shrimps; the penaeidins and the anti-lipopolysaccharide factor (ALF) from haemocytes [1,26] and anionic haemocyanin derived peptides from plasma [27]. Penaeidins are the diverse family of cationic peptides [28] and are classified into four classes (PEN-2, PEN-3, PEN-4 and PEN-5) [7,8,10,12,13,15–20,22,24,25]. Penaeidin consists of two distinct active domains: a proline-rich N-terminal domain (PRD) and cysteine-rich C-terminal domains (CRD) that inhibit the growth of several gram-positive bacterial and filamentous fungal species [13,28]. Multiple isoforms produced by amino acid substitutions and deletions within the PRD and CRD have been reported at the mRNA level for three classes of penaeidins suggesting that this is a highly diverse gene family [22].

Despite the knowledge on the identification and characterization of penaeidins in many penaeid shrimps, the evolutionary pattern of these peptides and the underlying genetic mechanisms that maintain high sequence diversities in the penaeidin gene family is poorly understood. In case of vertebrates, it has been suggested that multiple copies of structurally related AMP sequences are evolved by gene duplication and the rapid functional divergence among these multiple copies could be associated with accelerated rate of amino acid substitutions among the duplicated genes [3]. The accelerated rate of amino acid substitution is mainly the indication of adaptive evolution.

At the genomic level, two types of selective forces shape evolution; one is purifying selection that favours the conservation of existing phenotypes (functionally constrained and also known as purifying or negative selection) and the other one is positive Darwinian selection (adaptive evolution), which favours the fixation of beneficial mutations that lead to evolution of new traits [29]. One of the most powerful approaches to detect positive Darwinian selection is by comparing the number of non-synonymous substitution per non-synonymous site (d_N) with the number of synonymous substitution per synonymous site (d_S) [29,30]. There are three different types of selective pressures that can be detected from d_N and d_S ratio (hereafter referred as ω). Mutation at a codon site that results an amino acid change is known as non-synonymous substitution, whereas in synonymous substitution amino acid remains unchanged. If the protein-coding gene is functionally constrained, then the rate of non-synonymous change will be lower than neutral rate resulting in $\omega < 1$ and the gene is considered to be subjected to strong purifying selection [29]. Alternatively, if non-synonymous mutations are beneficial then the average rate of non-synonymous changes is expected to be higher than the neutral rate resulting in $\omega > 1$, and is considered to be subjected to positive selection. When $\omega = 1$, it indicates neutral evolution. A number of statistical approaches have been developed to detect positive selection at the protein coding genes [31–33]. Maximum likelihood-based codon substitution models [34] that account for variable ω among sites is one of the commonly used (e.g. refs. [31,35,36]) methods to detect positive Darwinian selection.

Here we examined the evolutionary pattern of the multiple copies of penaeidin peptides using the maximum likelihood (ML), Bayesian Inference (BI) and distance based neighbour-joining (NJ) phylogenetic approaches. We also examined the types of selective pressures operating on the codon sites of penaeidin peptides using the ML-based codon substitution models [34] and using the fixed effects likelihood method of Kosakovsky Pond and Frost, [37].

2. Materials and methods

2.1. Phylogenetic analyses

2.1.1. Nucleotide phylogeny

To infer evolutionary relationship among multiple copies of penaeidin peptides representing four classes (PEN-2, PEN-3, PEN-4 and PEN-5) and seven penaeid species, we reconstructed ML, BI and NJ phylogenies based on their respective nucleotide sequence data. A total of 36 nucleotide sequences were retrieved from GenBank ([6,7,11,13,16,18,22,24,25]; Table 1). Sequences were aligned using DAMBE version 4.5.2 [38,39] and BioEdit version 7.0.5.3 [40]. Amino acid alignment of signal peptide and active domains (PRD and CRD) of each sequence were manually edited according to the respective signature subgroups of PEN-2, PEN-3, PEN-4, ([15], http://www.penbase.immunaqua.com/), and PEN-5 [25] using DAMBE. Aligned amino acid sequences were mapped to corresponding codons using DAMBE. Nucleotide sequence data of all the three regions [the signal peptide and both active

Table 1
Protein (DNA) accession number, scientific name, peptide name, amino acid sequence length of the taxon used in the present study

Accession Number	Scientific name	Peptide name	aa[a]	Source
AAX58699 (AY956420)	*Litopenaeus schmitti*	Litsch PEN 4-1	67	[7]
AAK77540 (AF390147)	*Litopenaeus vannamei*	Litvan PEN 4a	67	[11]
AAK83455 (AY039207)	*Litopenaeus setiferus*	Litset PEN 4d	67	[11]
AAK77542 (AF390149)	*Litopenaeus vannemei*	Litvan PEN 4c	67	[11]
ABA63168 (DQ211701)	*Litopenaeus vannemei*	Litvan PEN 4-3	67	[22]
AAP33450 (AY260151)	*Fenneropenaeus chinensis*	Fenchi PEN 3-1	71	[19]
AAX58696 (AY956417)	*Farfantepenaeus paulensis*	Farpau PEN 2-2	73	[7]
AAZ80041 (DQ154152)	*Fenneropenaeus chinensis*	Fenchi PEN 5-2	79	[25]
AAV85945 (AY669323)	*Fenneropenaeus chinensis*	Fenchi PEN	79	GenBank
AAX58695 (AY956416)	*Farfantepenaeus paulensis*	Farpau PEN 2-1	73	[7]
AAZ79334 (DQ153253)	*Fenneropenaeus chinensis*	Fenchi PEN 5-1	79	[25]
AAQ62565 (AY351655)	*Litopenaeus stylirostris*	Litsty PEN 2	72	[6]
AAQ05769 (AF475082)	*Penaeus monodon*	Penmon PEN	74	[16]
ABA63166 (DQ211699)	*Litopenaeus vannamei*	Litvan PEN 2-4	72	[22]
CAA75142 (Y14925)	*Litopenaeus vannamei*	Litvan PEN 2	72	[13]
AAK77539 (AF390146)	*Litopenaeus vannamei*	Litvan PEN 2b	72	[11]
AAX58698 (AY956419)	*Litopenaeus schmitti*	Litsch PEN 2-2	72	[7]
AAX58697 (AY956418)	*Litopenaeus schmitti*	Litsch PEN 2-1	72	[7]
AAK83453 (AY039205)	*Litopenaeus setiferus*	Litset PEN 2d	72	[11]
AAK83450 (AY039202)	*Litopenaeus setiferus*	Litset PEN 3k	75	[11]
AAK83454 (AY039206)	*Litopenaeus setiferus*	Litset PEN 3-1	75	[11]
AAK83452 (AY039204)	*Litopenaeus setiferus*	Litset PEN 3n	75	[11]
AAY33770 (DQ010422)	*Litopenaeus stylirostris*	Litsty PEN 3-2	79	GenBank
AAK83451 (AY039203)	*Litopenaeus setiferus*	Litset PEN 3m	75	[11]
AAQ62566 (AY351656)	*Litopenaeus stylirostris*	Litsty PEN 3	87	[6]
AAK77535 (AF390142)	*Litopenaeus vannamei*	Litvan PEN 3f	82	[11]
AAK77533 (AF390140)	*Litopenaeus vannamei*	Litvan PEN 3d	82	[11]
AAK77532 (AF390139)	*Litopenaeus vannamei*	Litvan PEN 3a	82	[11]
AAK77537 (AF390144)	*Litopenaeus vannamei*	Litvan PEN 3h	82	[11]
AAK77538 (AF390145)	*Litopenaeus vannamei*	Litvan PEN 3i	82	[11]
AAK77536 (AF390143)	*Litopenaeus vannamei*	Litvan PEN 3g	82	[11]
AAK77534 (AF390141)	*Litopenaeus vannamei*	Litvan PEN 3e	82	[11]
CAA75144 (Y14927)	*Litopenaeus vannamei*	Litvan PEN 3b	82	[13]
CAA75145 (Y14928)	*Litopenaeus vannamei*	Litvan PEN 3c	81	[13]
AAK73083 (AF387660)	*Litopenaeus vannamei*	Litvan PEN 3j	81	[11]
ABA63167 (DQ211700)	*Litopenaeus vannamei*	Litvan PEN 3–11	81	[22]
1UEO A[b]	*Litopenaeus vannamei*	Litvan PEN 3	63	[24]

Nomenclature of the peptide names are according to Gueguen et al. [15].
[a] Amino acid sequence length.
[b] Protein Data Bank Code.

domains (PRD and CRD)] were included in the phylogenetic analyses. Hasegawa-Kishino-Yano (HKY) model with gamma distribution shape parameter (G) was the best-fit model selected by hierarchical likelihood ratio test (hLRTs) implemented in Modeltest version 3.5 [41]. PHYML version 2.4.4 [42] was used for ML analyses and MrBayes version 3.04 [43] was used for BI. The resulting trees were drawn using TreeView [44]. Nodal support for ML tree was estimated using 1000 non-parametric 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 Markov chain Monte Carlo to optimize efforts to find peaks in tree-space. Parameters were set as nst = 2, rates = gamma, and one tree was sampled in every 100. Convergence of tree was checked using Tracer version 1.3.1 [45] and resulting trees were used to generate a majority consensus tree with posterior probability values. NJ trees based on Kimura-2-parameter with gamma corrected model were also reconstructed using MEGA version 3.1 [46]. Using the same program nodal supports were estimated with 10 000 bootstrap replicates.

With the exclusion of signal peptides, we also reconstructed ML, BI, and NJ phylogenies and these inferred trees were independently used to tests for selection using PAML program [47]. To account for uncertainty regarding the true tree topology, we repeated the tests for positive selection using trees from ML, BI, and NJ analyses.

2.2. Tests for selection

To account for among-site variations and to test for positive selection on different codon sites and lineages, we performed tests for positive selection using two approaches; (1) maximum likelihood-based codon substitution analyses [34]; and (2) fixed effects likelihood (FEL) method of Kosakovsky Pond and Frost [37].

2.2.1. Pairwise comparison

We estimated the pairwise d_N, d_S and ω for each functional domain (PRD and CRD) using ML approach described by Goldman and Yang [48] implemented in CODEML program of the PAML version 3.15 [47].

2.2.2. ML-based codon substitution analyses

2.2.2.1. Branch model. The branch model allow the ω ratio to vary among branches in the phylogeny, and therefore useful in detecting positive selection operating on a particular lineage [49,50]. Using the inferred phylogenies and the free-ratios model [47], which assumes an independent ω ratio for each branch; we estimated the likelihood score using the CODEML program of PAML ver. 3.15 [47]. Likelihood ratio test (LRT) was used to compare free-ratios with the one-ratio model that assume all the branches have same ω.

2.2.2.2. Site specific models. We estimated parameters under seven different codon substitution models [34] implemented in the CODEML program of the PAML package [47] and their performances were evaluated using likelihood ratio tests (LRTs). LRTs were used to compare models that assume no positive selection ($\omega < 1$) with those that assume positive selection ($\omega > 1$). The seven codon substitution models are: M0 (one-ratio), M1a (nearly neutral), M2a (positive selection), M3 (discrete), M7 (β-distribution; $0 \leq \omega \leq 1$), M8 ($\beta + \omega > 1$: continuous) [34] and M8a ($\beta + \omega = 1$) [51]. The M1a model estimates a single parameter p_0, the frequency of conserved sites with $\omega_0 = 0$ and the remaining sites with frequency p_1 ($p_1 = 1 - p_0$) assuming $\omega_1 = 1$. The M2a model adds a class of positively selected sites with frequency p_2 (where $p_2 = 1 - p_1 - p_0$), with ratio ω_2 estimated from the data. Thus, whereas M1a estimates a single parameter (p_0), M2a estimates three parameters (p_0, p_1, and ω_2). In the M7 model, ω follows a beta distribution such that $0 \leq \omega \leq 1$ and the two parameters (p and q) of the beta distribution are estimated from the data. In the M8 model, a proportion p_0 of sites have ω drawn from the beta distribution. The remaining sites with proportion p_1 are positively selected and have $\omega_1 > 1$. Thus, M7 estimates two parameters (p and q), while M8 estimates four parameters (p, q, p_0, and ω_1). 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 equal to the difference in the number of parameters between models [34].

Five LRTs were conducted. The 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. The third comparison was between M1a, which allows for two site classes ($0 < \omega < 1$, $\omega = 1$) with M2a. The fourth comparison was between a model of beta-distributed selective pressures, that allows for 10 site classes, each with $\omega < 1$ and M8, which has 11 site classes, one of them allowed for $\omega > 1$. The last comparison was between M8 and M8a, in which an additional parameter was constrained to have $\omega = 1$ [51]. In all LRTs good evidence for positive selection is found if the LRT indicates that models that allow for selection (i.e. M2a and M8) are significantly better than their respective null models (M1a and M7) [34].

2.2.2.3. Branch-site models. Considering that positive selection may operate in very short episodes during the evolution of a protein [52] and affect only a few sites along a few lineages in the phylogeny, recently developed likelihood models known as branch-site model [53,54] that allow ω ratios to vary both among lineages and amino acid sites was also used in the analysis. We used branch site model to test for positive selection within each lineage of penaeidin (a, b, and c; Fig. 1) and among the three lineages (lineage d in Fig. 1). The branch-site model (null: model A1) has four parameters and four site classes; 0, 1, 2a, and 2b that account for $0 < \omega_0 < 1$, $\omega_1 = 1$, $\omega_2 > 1$, and $\omega_2 > 1$, respectively, was compared with the alternative model that constrained to have $\omega_2 = 1$ (model A). The LRTs between

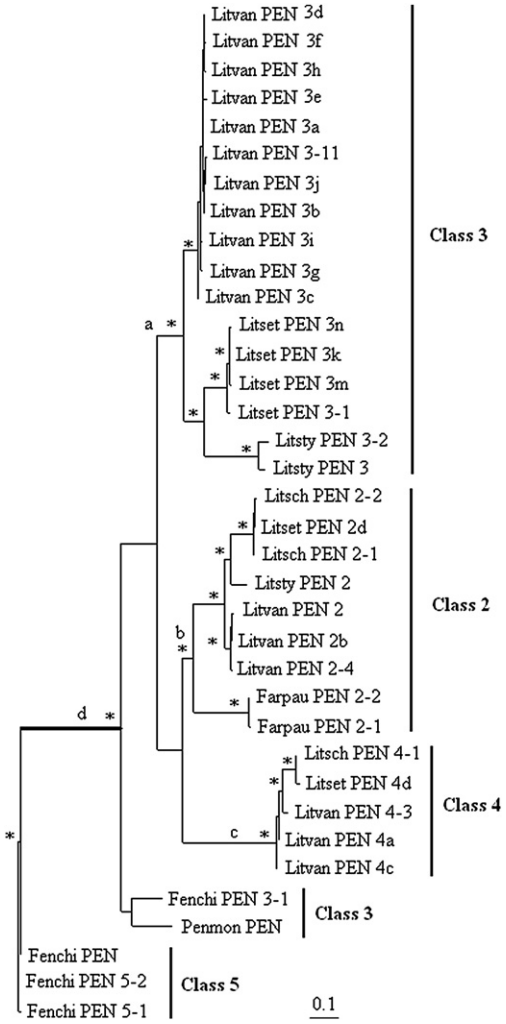


Fig. 1. Phylogenetic relationships among different subunits of penaeidin inferred from their nucleotide sequence data using maximum likelihood approach. Bootstrap values ≥75 are indicated by asterisks. Lineages a, b, c, and d were tested whether positive Darwinian selection is operating in different lineages using the branch-site models implemented in CODEML program of PAML ver. 3.15 [47]. Lineage ''d'' is under the influence of positive Darwinian selection. See Tables 2 and 3 for branch-site models and for their LRTs, respectively.

the two 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 one [53,54].

2.2.3. Fixed effects likelihood method

To identify codons affected by positive and negative selection, we used FEL method available in HyPhy package (http://www.datamonkey.org) [55]. FEL uses the entire alignment to infer model parameters shared by all sites (e.g. branch lengths) and then fits d_S and d_N rates individually at every site. Neutrality of an individual site is tested using the likelihood ratio test [37].

2.3. Protein structure of penaedin

We used 3-dimensional (3D) structure of penaeidin-3 of *Litopenaeus vannamei* (Protein Data Bank number = 1UEO; [24]; http://www.rcsb.org/pdb/) to map the positively selected sites in the 3D structure. The alignment of all classes of penaeidin with 1UEO amino acid residues is shown in Table 4. Positively selected sites are marked

accordingly. The structure was displayed and positively selected sites were located using RasMol V2.7.2.1.1 (http://www.openrasmol.org/software/rasmol/).

3. Results

3.1. Phylogenetic analyses

For the entire coding sequence (signal peptide + PRD + CRD), the HKY model with gamma distribution shape parameter (G = 0.6636) was the best-fit model selected by hLRTs. The transition to transversion ratio (T_i/T_v) and the log likelihood score (−ln L) were 0.8639 and 2137.2756, respectively. The nucleotide base frequencies for A, C, G and T were 0.2122, 0.2884, 0.2453 and 0.2542, respectively. When signal peptides were excluded from the analyses, the K80 model with G (= 1.0185) was the appropriate model selected by hLRTs. The T_i/T_v and the log likelihood score (−ln L) were 0.8647 and 1862.4727, respectively.

The ML tree inferred from the nucleotide sequence data showed the phylogenetic relationships among multiple copies of penaeidins (Fig. 1). Phylogenetic trees inferred from BI, and NJ (trees not shown) are congruent with ML tree. The phylogenetic analyses clearly indicated that each class formed a distinct cluster, thus gene duplication is the likely explanation for divergence of this peptide family. From the phylogenetic analyses, it is also apparent that paralogy seems obvious between PEN-2, PEN-3, and PEN-4 in the genus *Litopenaeus*, however, it is difficult to distinguish orthology from paralogy for PEN-5 and all PEN-5 sequences outside the genus *Litopenaeus* because of unbalanced taxon/copy sampling. Nevertheless, all the trees (ML, BI and NJ) are consistent with the fact that each class formed a distinct cluster, therefore suggesting that multiple copies of this peptide family evolved by gene duplication events.

3.2. Tests for selection

3.2.1. Pairwise comparison

Distribution of pairwise estimates of ω for PRD and CRD are shown in Fig. 2A,B, respectively. Although the average ω for PRD exceeds one (1.272 ± 0.973), the ω value ranged from 0.001 to 6.45. Similarly, the ω for CRD ranged from 0.001 to 94.698, with average 2.814 ± 3.185. Despite the wide range of variation of ω in both domains, our analyses clearly indicate that both domains are under the influence of positive Darwinian selection. To know which codons/amino acid residues are under the positive Darwinian selection, we performed ML-based codon substitution analyses.

3.2.2. ML-based codon substitution analyses

3.2.2.1. Branch model. The tree in Fig. 1 has 69 branches; therefore, 68 additional ω parameters are involved in the free-ratios model. However, comparison of one-ratio model, which assumes the same ω ratio for all lineages with free-ratios model, which assumes an independent ω ratio for each branch failed to suggest the rejection of one-ratio model ($p \geq 0.54$, Table 3), therefore suggesting that the ω ratios are not variable among lineages.

3.2.2.2. Site specific model. Parameter estimates and log likelihood values under models of variable ω among codon sites and their LRTs are shown in Tables 2 and 3, respectively. Under the simplest model, which allows only a single ω across all codon sites (M0), the ML estimate of $\omega = 0.9805$. This estimate is statistically indistinguishable from $\omega = 1$, the expected value under a completely neutral model of sequence evolution. However, a model allowing for variation among sites (M3) provides a significantly better fit to the data showing that there is variation among sites in the strength of selection. However, M0 is a highly unrealistic model and the M0/M3 comparison thus provides a test for variation in ω among sites rather than variation in strength of selection among sites [34]. A more stringent test for the presence of positive selection is a comparison of models M1a and M2a [34], and between M7 and M8 [56]. The LRT comparing M7 and M8 provides strong evidence for positive selection (Table 3; $p = 0$). Similar results were obtained using models M1a and M2a (Table 3). There is a large degree of overlap in the positively selected sites identified from models M2a and M8. Yang ([47]; in PAML ver. 3.15) reported that the M1a/M2a comparison seems more robust than the M7/M8 pair. If the true null model assumes several classes of conserved sites with $\omega < 1$ as well as

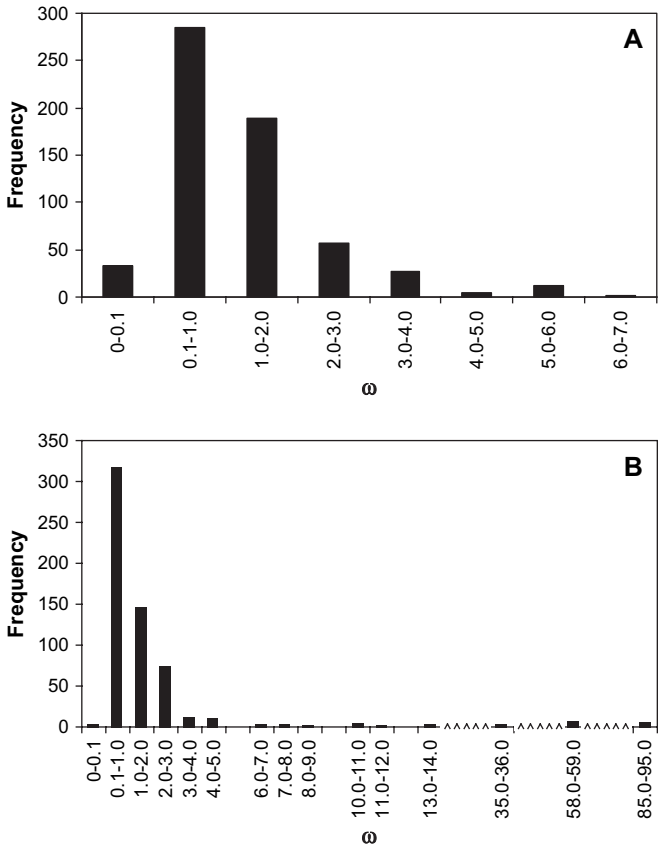


Fig. 2. Distribution of pairwise estimates of ω-ratios among 36 penaeidin coding sequences estimated using maximum likelihood approach [48] implemented in PAML [47]. (A) N-terminal proline-rich domain; and (B) C-terminal cysteine-rich domain.

neutral sites with $\omega = 1$, the M7/M8 comparison may often be significant, and among half of such cases, the ω estimate under M8 will be >1, and will produce false positives ([47]; PAML ver. 3.15). In such cases, the M8a/M8 comparison or M1a/M2a comparison may be more robust than M7/M8 comparison. Nevertheless, our analyses showed that all the models revealed consistent results ($p = 0$; Table 3), indicating evidence of positive Darwinian selection is widespread across the codons of penaeidin. The posterior probability along with ω of each positively selected codon sites identified by M2a model is shown in Fig. 3. Our analyses showed that under M2a model, 9 codon sites were positively selected (8 in PRD and 1 in CRD) with posterior probability ≥0.90, and the ω ranged from 4.075 ± 0.917 to 2.401 ± 1.405 (Fig. 3 and Table 4). Codon sites in PRD and CRD that are under the influence of positive Darwinian selection are shown in Table 4. The codon sites (identified by M2a) predicted to be subjected to positive Darwinian selection also shown on the PDB structure (Fig. 4).

3.2.2.3. Branch-site model. Branch-site models clearly indicate that lineages a (PEN-3), b (PEN-2), and c (PEN-4) are not under the influence of positive Darwinian selection (Tables 2 and 3), however, lineage d that comprise of three lineages (a, b and c) is under the influence of positive Darwinian selection (Tables 2 and 3).

3.2.3. FEL analyses

The positively and negatively selected sites identified by FEL method are shown in Table 5A,B, respectively. There are eight positively and six negatively selected codon sites ($p = 0.1$) were identified. Majority of codon sites that were positively selected (with posterior probability 0.90) by the ML-based codon substitution models were also positively selected by FEL methods. Of the 6 negatively selected codon sites, 1 and 5 sites are in PRD and CRD, respectively. Nevertheless, both ML codon substitution and FEL methods supported the fact that positive Darwinian selection is widespread in penaeidin.

Table 2
Parameter estimates and log-likelihood values under models of variable ω-ratios among sites (site-specific models), branch model (free-ratio), and branch-site models

Model	Free Parameters	Parameter estimates	Likelihood scores	Positively selected sites
M0: One-ratio	1	**$\omega = 0.9805$**	−1834.085301	None
Branch model				
M1: Free-ratio	69	Independent ω for each branch	−1801.07791	N/A
Site specific models				
M1a: Nearly neutral	1	$\omega_0 = 0.112$, $\omega_1 = 1$, ($p_0 = 0.34$, $p_1 = 0.66$)	−1787.826202	Not allowed
M2a: Positive selection	3	$\omega_0 = 0.117$, $\omega_1 = 1$, **$\omega_2 = 3.73$**; ($p_0 = 0.29$, $p_1 = 0.45$, $p_2 = 0.26$)	−1767.917763	11, **17**, 20, **21**, **22**, **23**, **24**, **25**, **27**, 28, 29, 30, 31, 41, 49, 57, **64**, 65
M3: discrete[a]	5	$\omega_0 = 0.18$, $\omega_1 = 1.52$, **$\omega_2 = 6.29$**, ($p_0 = 0.35$, $p_1 = 0.48$, $p_2 = 0.17$)	−1766.153099	2, 3, **4, 11, 12, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31**, 32, 33, 34, 35, 36, 37, 38, **40, 41,45**, *46,* **47, 48, 49, 50**, 52, **53, 56, 57, 58, 60, 63, 64, 65**, 66, 67, **69**
M7: β	2	p = 0.301, q = 0.144	−1790.226092	Not allowed
M8: $\beta + \omega_s > 1$	4	$p_0 = 0.69$, $p_1 = 0.31$, p = 0.37, q = 0.25, **$\omega = 3.3$**	−1769.198243	*11*, **16, 17**, 18, 20, **21, 22, 23, 24, 25**, 26, **27**, *28,* **29**, 30, *31,* 32, 33, 34, *41,* 48, 49, 53, *56*, **64**, *65*
M8a: $\beta + \omega_s = 1$	3	$P_0 = 0.37$, $p_1 = 0.63$, p = 1.54, q = 9.56, $\omega = 1$	−1787.425149	17, 20, **21**, **22, 23**, 24, *25*, **27**, 28, *29,* 30, 31, 49, *64*
Branch-site models				
Branch **a**				
Model A	3	$p_0 = 0.34$, $p_1 = 0.66$, $p_{2a} = 0$, $p_{2b} = 0$, $\omega_0 = 0.1$, $\omega_1 = 1$, $\omega_{2a} = 1$, $\omega_{2b} = 1$	−1787.826202	Not allowed
Model A1	4	$p_0 = 0.34$, $p_1 = 0.66$, $p_{2a} = 0$, $p_{2b} = 0$, $\omega_0 = 0.1$, $\omega_1 = 1$, $\omega_{2a} = 1$, $\omega_{2b} = 1$	−1787.826202	44
Branch **b**				
Model A	3	$p_0 = 0.34$, $p_1 = 0.66$, $p_{2a} = 0$, $p_{2b} = 0$, $\omega_0 = 0.1$, $\omega_1 = 1$, $\omega_{2a} = 1$, $\omega_{2b} = 1$	−1787.826202	Not allowed
Model A1	4	$p_0 = 0.34$, $p_1 = 0.66$, $p_{2a} = 0$, $p_{2b} = 0$, $\omega_0 = 0.1$, $\omega_1 = 1$, $\omega_{2a} = 1$, $\omega_{2b} = 1$	−1787.826202	None
Branch **c**				
Model A	3	$p_0 = 0.34$, $p_1 = 0.66$, $p_{2a} = 0$, $p_{2b} = 0$, $\omega_0 = 0.1$, $\omega_1 = 1$, $\omega_{2a} = 1$, $\omega_{2b} = 1$	−1787.826202	Not allowed
Model A1	4	$p_0 = 0.34$, $p_1 = 0.66$, $p_{2a} = 0$, $p_{2b} = 0$, $\omega_0 = 0.1$, $\omega_1 = 1$, $\omega_{2a} = 1$, $\omega_{2b} = 1$	−1787.826202	None
Branch **d**				
Model A	3	$p_0 = 0.26$, $p_1 = 0.49$, $p_{2a} = 0.09$, $p_{2b} = 0.16$, $\omega_0 = 0.095$, $\omega_1 = 1$, $\omega_{2a} = 1$, $\omega_{2b} = 1$	−1786.626898	Not allowed
Model A1	4	$p_0 = 0.2$, $p_1 = 0.4$, $p_{2a} = 0.14$, $p_{2b} = 0.26$, $\omega_0 = 0.093$, $\omega_1 = 1$, **$\omega_{2a} = 54.64$, $\omega_{2b} = 54.64$**	−1776.876147	5, 7, *9,* **11**, 15, 16, **17**, **18**, 19, 20, **21**, 23, **24**, 26, 27, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41, 42, 44, 45, 48, **49**, **50**, **52**, **63**, **65**, 68, 69, 70

Positively selected sites with posterior probability ≥0.95 are in bold, 0.9–0.95 are underlined, 0.8–0.9 in italics, and 0.5–0.8 in plain text. M0: one-ratio ω value is the average for all codon sites, whereas M2a, M3 and M8 ω values are the estimated values for the positively selected codon sites under respective models. See Fig. 1 for the position of branch a, b, c, and d. N/A = Not applicable.

[a] Model detected positively selected sites based on Naive Empirical Bayes (NEB) analysis.

4. Discussion

Phylogenetic analyses of penaedin peptides of penaeid shrimps revealed that regardless of species, multiple copies in each class of penaeidin clustered together with strong nodal support. Thus, suggesting that these peptides of penaeid shrimps are paralogous and evolved by gene duplication events. Previous studies on the molecular evolution of AMPs

Table 3
Likelihood ratio statistics among different models given in Table 2

Comparison	2 Δ*l*	df[a]	*p*
Branch model			
M0 vs M1	66.014782	68	0.546
Site-specific models			
M0 vs M2a	132.3351	2	0.0000
M0 vs M3	135.8644	4	0.0000
M1a vs M2a	39.8169	2	0.0000
M7 vs M8	42.0557	2	0.0000
M8 vs M8a	36.4538	1	0.0000
Branch-site models			
Branch **a**	0.0000	1	1.0000
Branch **b**	0.0000	1	1.0000
Branch **c**	0.0000	1	1.0000
Branch **d**	19.5015	1	0.0000

[a] Degrees of freedom.

in many vertebrates and invertebrates (e.g. refs. [3,35,36,57,58]) also showed strong evidence of gene duplication in AMPs of respective taxonomic groups. Although from the paralogous gene tree, it is difficult to infer the absolute divergences at both synonymous and non-sysnonymous sites since the sequences last shared a common ancestor (which can only be estimated from the orthologous gene tree), tests of selection based on substitution rate ratios (d_N/d_S) do not depend on assumption about orthology [36]. Our phylogenetic analyses suggested evidence of gene duplication in penaeidin, however, it is unclear what recurrent evolutionary forces maintained the high diversity of these peptides at functional level. Previous studies on AMP of vertebrates and invertebrates suggested that gene duplication followed by the accelerated rate of amino acid substitutions among the duplicated genes are the likely cause of rapid functional diversification of these peptides (e.g. refs. [3,35,36,57]). Positive Darwinian selection (positive natural selection) is the likely cause of such accelerated rate of amino acid substitutions (e.g. refs. [3,35,36,57,58]). From the pairwise comparison, it is apparent that positive Darwinian selection is widespread across penaeidins. Therefore, suggesting that the accelerated amino acid substitution rate among the duplicated penaeidin peptides is due to the

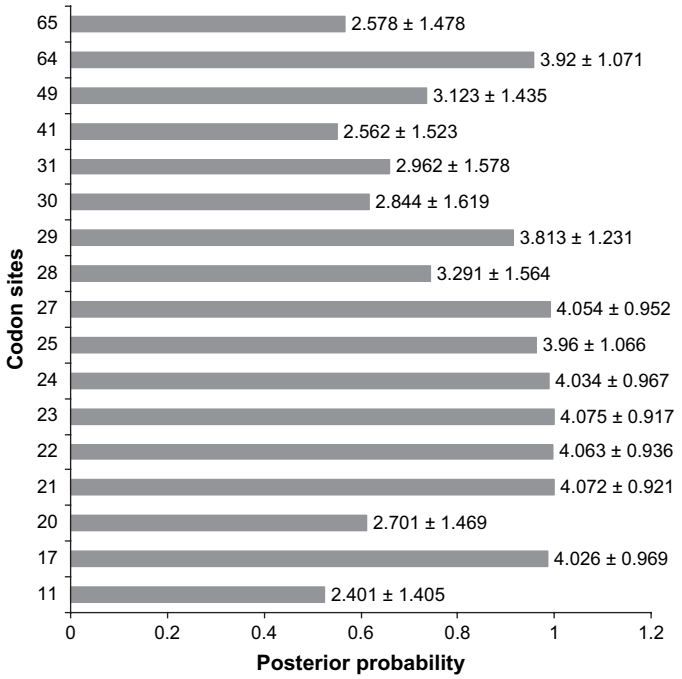
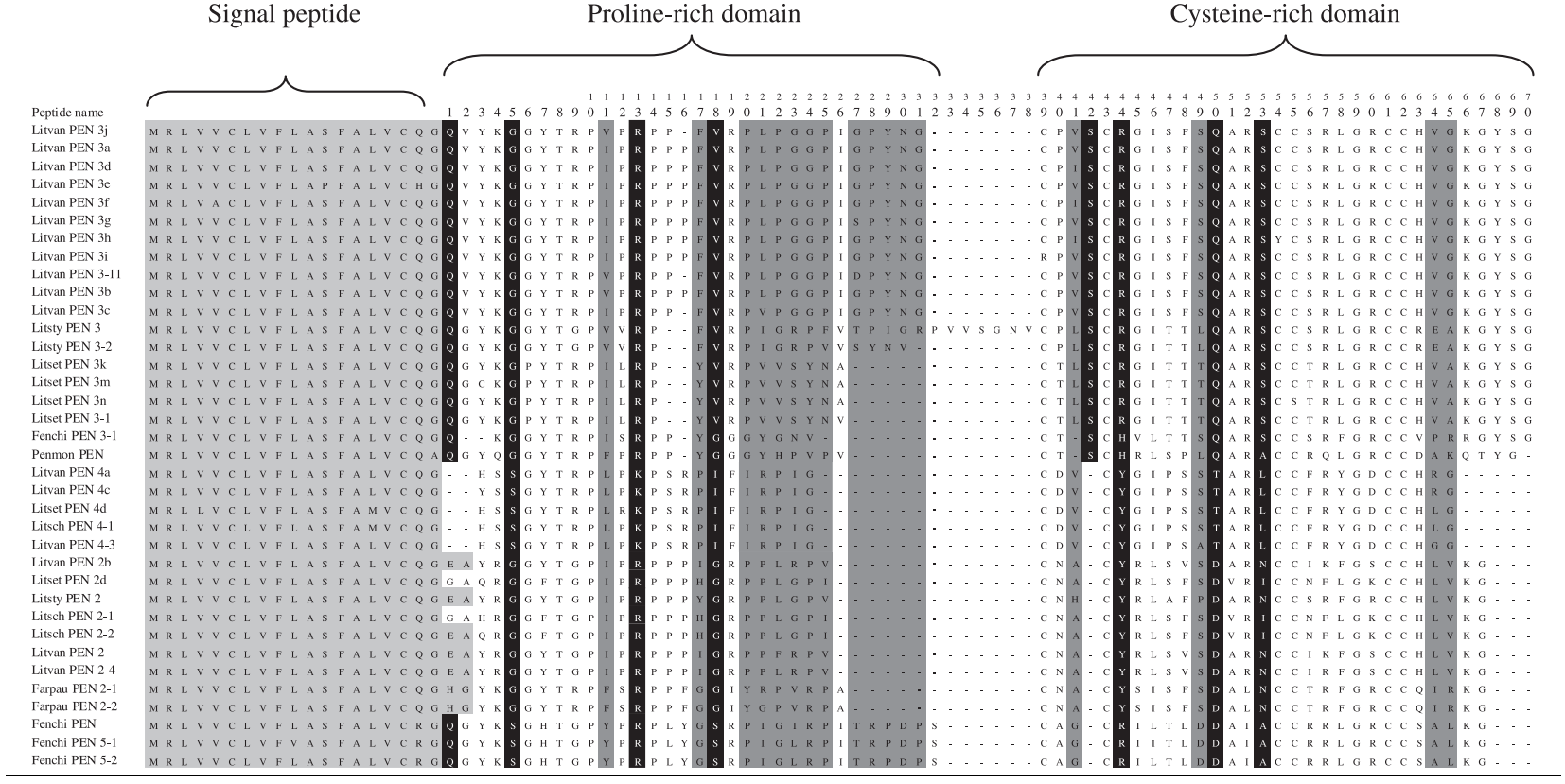


Fig. 3. Posterior probabilities (M2a model) for each codon of penaeidin. The ω values with standard errors are shown for each site.

Table 4
Amino acid sequence comparisons among different antimicrobial peptides (penaeidin) of penaeids



Signal peptides and subgroup signatures are shaded in light grey and black colour, respectively. Positively selected sites on the two domains are shaded in dark grey. The subgroup signatures are identified based on Gueguen et al. [15] and Kang et al. [25].

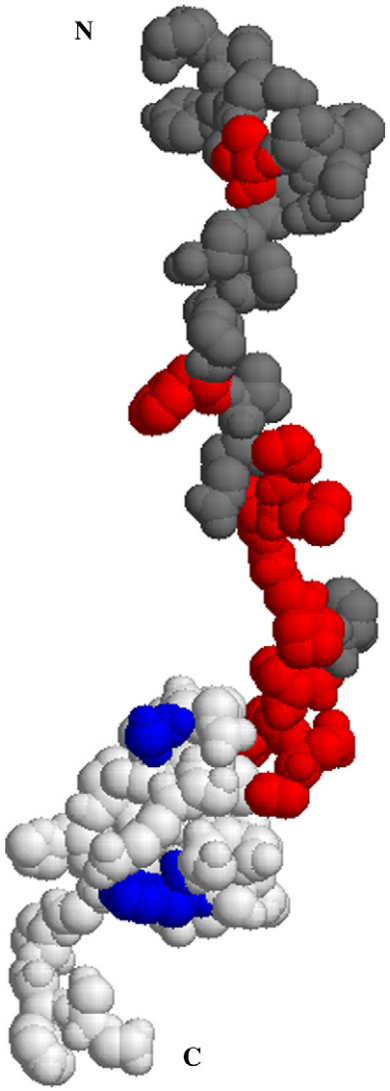


Fig. 4. The three-dimensional structure of penaeidin (PDB number: 1UEO). Proline-rich domain (PRD) and cysteine-rich domain (CRD) are indicated by dark grey and light grey, respectively. Sites coloured red and blue are those sites in PRD and CRD predicted to be under positive Darwinian selection (identified by M2a model, see Table 2), respectively. The structure was displayed using RasMol V2.7.2.1.1 (http://www.openrasmol.org/software/rasmol/). N: N-terminal, C: C-terminal.

positive Darwinian selection. However, the wide variation in ω (Fig. 2A,B) in PRD and CRD further indicated that the entire codon sites are not positively selected, therefore it is apparent that few codon sites in both domains might have experienced accelerated rate of non-synonymous substitutions than that of silent substitutions. Nevertheless, pairwise comparison method indicated the evidence of positive Darwinian selection across the PRD and CRD domains of penaeidins. Many pairwise comparisons in AMP families showed that ω is not significantly different from or even less than one [59] and in some cases, ω is consistently less than one (e.g. refs. [60,61]). However, from the pairwise comparisons, it is difficult to infer which sites are positively or negatively selected. ML based codon substitution models appear to offer a number of advantages over the pairwise comparisons of d_N and d_S among taxa that average over all codon sites and lineages [36,62,63]. Considering this pitfall of pairwise comparison methods, we performed ML based-codon substitution analyses of Yang et al. [34]. Although ML-based codon substitution analyses showed the evidence of positive Darwinian selection in both domains of penaeidin, a relatively more number of positively selected codon sites were observed in PRD than the CRD. It could be possible that CRD is relatively more conserved than the PRD [10,12]; therefore, more amino acid residues in PRD are subjected to natural selection in different environmental

Table 5
Sites selected identified by fixed effects likelihood (FEL) method (0.1% significant level)

Codon	Normalised d_N-d_S	Constrained d_S-d_N	LRT	p-value
(A) Positively selected sites				
11	1.112	1.645	3.301	0.069
12	0.541	0.653	4.217	0.040
27	7.888	8.999	3.199	0.074
48	0.907	1.237	4.134	0.042
49	1.407	1.446	6.463	0.011
53	0.782	0.935	5.902	0.015
56	1.130	1.602	4.461	0.035
64	1.117	1.634	3.532	0.060
(B) Negatively selected sites				
8	−0.936	0.492	9.691	0.0019
51	−0.535	0.488	3.014	0.0826
59	−0.654	0.415	6.294	0.0121
61	−1.117	0.288	7.314	0.0068
66	−1.426	0.617	3.596	0.0579
68	−3.206	0.659	7.974	0.0047

(A) 8 positively selected sites; and (B) 6 negatively selected sites. LRT: Likelihood ratio test.

conditions. Despite the minor discrepancies in the identification of positively selected sites by both ML-based codon substitution models and FEL methods, both methods are consistent with the fact that positive Darwinian selection is widespread across the penaeidin.

In consistent with the results of previous studies on the molecular evolution of AMPs in vertebrates and invertebrates (e.g. refs. [3,35,36,57]), the present study also present convincing evidence that positive selection on penaeidin, the AMP gene family of penaeid shrimps, is common and taxonomically widespread. However, the question is why selection appears to act differently on different loci/amino acid residues. Although both domains are under the influence of positive Darwinian selection, the differences in the number of positively selected amino acid residues in both domains could be attributed to the structural organization of the respective domains. For example, the N-terminal proline rich residues of penaeidin are less conserved, whereas the cysteine rich C-terminal domain is stabilized by three conserved disulphide bonds [10,12,24], therefore, more number of positively selected amino acid residues in PRD is expected. The high sequence variability that have generated by gene duplication events followed by the positive Darwinian selection among the PRD of different classes of penaeidin may have resulted in variation in target specificity and gain in antimicrobial function [10,12]. On contrary, the CRD, which is relatively conserved, might have a tandem or synergistic role [10,12].

It is possible that since these mature peptides (PRD and CRD) are used to protect against the pathogens and filamentous fungal species [13,28], their adaptive evolution will be caused by several factors. Like MHC receptors, immunoglobulins, defensins and AMPs [36,57,64–66], penaeidin may be under selection directed by the evolution of pathogens. The specific pathogens driving selection certainly vary among hosts, which could also result in different patterns of evolution [36]. Some hosts might also be co-evolving with pathogens that are under selection to resist their defenses. An array of resistance mechanisms to AMPs are known in pathogens, some of which are involved in a single gene product [67,68]. In some cases, derived microbial strains are more resistant than the wild type [69], which is consistent with the hypothesis that positive selection on microbial genomes can result in increased resistance to AMPs. However, it could be more likely that resistance is easy to evolve and happens frequently. On the other hand, penaeidin peptides might attack their targets in such a way that evolving resistance is not possible without synchronized changes at many microbial genes. Selection on the penaeidin mature peptides might also occur when hosts enter new habitats or environments and are forced to adapt to completely different pathogens not previously encountered (e.g. during species introduction). In such a circumstance, mature peptides that involve in protection against pathogens evolve for the particular microbial biota that these species has to encounter in the new environment/habitat.

We suggest that gene duplication is the most possible mechanism that generated multiple copies of penaeidin. The accelerated rate of amino acid substitutions among these duplicated genes is due to the influence of positive Darwinian selection. Like other immunogeneic genes (MHC, defensins, immunoglobulins), the evolution of this AMP family

(penaeidin) of penaeid shrimps might also be directed by the pathogens. Therefore, the mature peptides PRD and CRD that directly involve in the protection against pathogens eventually accumulate new functions.

Acknowledgements

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

References

[1] Bachère E, Gueguen Y, Gonzalez M, de Lorgerli J, Garnier J, Romestand B. Insights into the anti-microbial defense of marine invertebrates: the penaeid shrimps and the oyster *Crassostrea gigas.* Immunol Rev 2004;198:149–68.
[2] Boman HG. Antibacterial peptides: basic facts and emerging concepts. J Intern Med 2003;254:197–215.
[3] Nicolas P, Vanhoye D, Amiche M. Molecular strategies in biological evolution of antribacterial peptides. Peptides 2003;24:1669–80.
[4] Castro MS, Fontes W. Plant defense and antimicrobial peptides. Protein Peptide Lett 2005;12:11–6.
[5] Bachère E. Penaeidins, antimicrobial peptides of shrimp: a comparison with other effectors of innate immunity. Aquaculture 2000;191:71–88.
[6] Muňoz M, Vandenbulcke F, Garnier J, Gueguen Y, Bulet P, Saulnier D, et al. Involvement of penaeidins in defense reactions of the shrimp *Litopenaeus stylirostris* to a pathogenic vibrio. Cell Mol Life Sci 2004;61:961–72.
[7] Barracco MA, de Lorgeril J, Gueguen Y, Bachere E. Molecular characterization of penaeidins from two Atlantic Brazilian shrimp species, *Farfantepenaeus paulensis* and *Litopenaeus schmitti*. FEMS Microbiol Lett 2005;250:117–20.
[8] Chen JY, Pan CY, Kuo CM. Molecular cloning and sequencing of shrimp (*Penaeus monodon*) penaeidin-5 cDNA. Fish Shellfish Immunol 2004;16:665–70.
[9] Chiou TT, Wu JL, Chen TT, Lu JK. Molecular cloning and characterization of cDNA of penaeidin-like antimicrobial peptide from tiger shrimp (*Penaeus monodon*). Mar Biotechnol 2005;7:119–27.
[10] Cuthbertson BJ, Bullesbach EE, Fievet J, Bachère E, Gross PS. A new class (penaeidin class 4) of antimicrobial peptides from the Atlantic white shrimp (*Litopenaeus setiferus*) exhibits target specificity and an independent proline-rich-domain function. Biochem J 2004;381:79–86.
[11] Cuthbertson BJ, Shepard EF, Chapman RW, Gross PS. Diversity of the penaeidin antimicrobial peptides in two shrimp species. Immunogenetics 2002;54:442–5.
[12] Cuthbertson BJ, Yang Y, Bachere E, Bullesbach EE, Gross PS, Aumelas A. Solution structure of synthetic penaeidin-4 with structural and functional comparisons with penaeidin-3. J Biol Chem 2005;280:16009–18.
[13] Destoumieux D, Bulet P, Loew D, Van Dorsselaer A, Rodriguez J, Bachere E. Penaeidins, a new family of antimicrobial peptides isolated from the shrimp *Penaeus vannamei* (Decapoda). J Biol Chem 1997;272:28398–406.
[14] Destoumieux D, Munoz M, Bulet P, Bachère E. Penaeidins, a family of antimicrobial peptides from penaeid shrimp (Crustacea, Decapoda). Cell Mol Life Sci 2000;57:1260–71.
[15] Gueguen Y, Garnier J, Robert L, Lefranc MP, Mougenot I, de Lorgeril J, et al. PenBase, the shrimp antimicrobial peptide penaeidin database: sequence-based classification and recommended nomenclature. Dev Comp Immunol 2006;30:283–8.
[16] Ho S-H, Chao Y-C, Tsao H-W, Sakai M, Chou H-N, Song Y-L. Molecular Cloning and Recombinant Expression of Tiger Shrimp *Penaeus monodon* Penaeidin. Fish Pathol 2004;39:15–23.
[17] Hu SY, Huang JH, Huang WT, Yeh YH, Chen MHC, Gong HY, et al. Structure and function of antimicrobial peptide penaeidin-5 from the black tiger shrimp *Penaeus monodon*. Aquaculture 2006;260:61–8.
[18] Kang C-J, Wang J-X, Zhao X-F, Xiang J-H. Molecular cloning of the cDNA encoding a mature antimicrobial peptide from Chinese shrimp: *Fenneropenaeus chinensis*. Shandong Da Xue Xue Bao 2002;37:554–9.
[19] Kang CJ, Wang JX, Zhao XF, Yang XM, Shao HL, Xiang JH. Molecular cloning and expression analysis of the Ch-penaeidin, an antimicrobial peptide from Chinese shrimp, *Fenneropenaeus chinensis*. Fish and Shellfish Immunol 2004;16:513–25.
[20] Muňoz M, Vandenbulcke F, Gueguen Y, Bachère E. Expression of penaeidin antimicrobial peptides in early larval stages of the shrimp *Penaeus vannamei*. Dev Comp Immunol 2003;27:283–9.
[21] Muňoz M, Vandenbulcke F, Saulnier D, Bachère E. Expression and distribution of penaeidin antimicrobial peptides are regulated by haemocyte reactions in microbial challenged shrimp. Eur J Biochem 2002;269:2678–89.
[22] O'Leary NA, Gross PS. Genomic structure and transcriptional regulation of the penaeidin gene family from *Litopenaeus vannamei*. Gene 2006;371:75–83.
[23] Supungul P, Klinbunga S, Pichyangkura 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.
[24] Yang Y, Poncet J, Garnier J, Zatylny C, Bachère E, Aumelas A. Solution structure of the recombinant penaeidin-3, a shrimp antimicrobial peptide. J Biol Chem 2003;278:36859–67.
[25] Kang CJ, Xue JF, Liu N, Zhao XF, Wang JX. Characterization and expression of a new subfamily member of panaeidin antimicrobial peptides (penaeidin 5) from *Fenneropenaeus chinensis*. Mol Immunol 2007;44:1535–43.
[26] Somboonwiwat K, Marcos M, Tassanakajon A, Klinbunga S, Aumelas A, Romestand B, et al. Recombinant expression and antimicrobial activity of anti-lipopolysaccharide factor (ALF) from the black tiger shrimp *Penaeus monodon*. Dev Comp Immunol 2005;29:841–51.
[27] Destoumieux-Garzon D, Saulnier D, Garnier J, Jouffrey C, Bulet P, Bachere E. Crustacean immunity: antifaungal peptides are generated from the C-terminus of shrimp haemocyanin in response to microbial challenge. J Biol Chem 2001;276:47070–7.

[28] Destoumieux D, Bulet P, Strub JM, Van Dorsselaer A, Bachère E. Recombinant expression and range of activity of penaeidins, antimicrobial peptides from penaeid shrimp. Eur J Biochem 1999;266:335–46.
[29] Kimura M. The neutral theory of molecular evolution. New York: Cambridge University Press; 1983.
[30] Hughes AL, Nei M. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc Natl Acad Sci USA 1989;86:958–62.
[31] Ford MJ. Applications of selective neutrality tests to molecular ecology. Mol Ecol 2002;11:1245–62.
[32] Nielsen R. Statistical tests of neutrality in the age of genomics. Heredity 2001;86:641–7.
[33] Yang Z, Bielawski JP. Statistical methods for detecting molecular adaptation. Trends Ecol Evol 2000;15:496–503.
[34] Yang Z, Nielsen R, Goldman N, Pederson AM. Codon-substitution models for heterogeneous selection pressures at amino acid sites. Genetics 2000;155:431–49.
[35] Bulmer MS, Crozier RH. Duplication and diversifying selection among termite antifungal peptides. Mol Biol Evol 2004;21:2256–64.
[36] Tennessen JA. Molecular evolution of animal antimicrobial peptides: widespread moderate positive selection. J Evol Biol 2005;18:1387–94.
[37] Kosakovsky Pond S, Frost SD. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol 2005;22:1208–22.
[38] Xia X. Data analysis in molecular biology and evolution. Boston: Kluwer Academic Publishers; 2000.
[39] Xia X, Xie Z. DAMBE: Data analysis in molecular biology and evolution. J Heredity 2001;92:371–3.
[40] Hall TA. BioEdit: a user friendly biological sequence alignment editor and analyses program for windows 95/98/N.T. Nucl Acids Symp Ser 1999;41:95–8.
[41] Posada D, Crandall KA. Modeltest: testing the model of DNA substitution. Bioinformatics 1998;14:817–8.
[42] Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 2003;52: 696–704.
[43] Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 2001;17:754–5.
[44] Page RDM. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 1996;12:357–8.
[45] Rambout A., Drummond A.J. Tracer 1.3 [computer program] Available: <http://evolve.zoo.ox.ac.uk/software.html>; 2004.
[46] Kumar S, Tamura K, Nei M. MEGA3.1: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 2004;5:150–63.
[47] Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci 1997;13:555–6.
[48] Goldman N, Yang Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol 1994;11:725–36.
[49] Yang Z, Nielsen R. Synonymous and non-synonymous rate variation in nuclear genes of mammals. J Mol Evol 1998;46:409–18.
[50] Yang Z. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol Biol Evol 1998;15:568–73.
[51] Swanson WJ, Nielsen R, Yang Q. Pervasive adaptive evolution in mammalian fertilization proteins. Mol Biol Evol 2003;20:18–20.
[52] Gillespie JH. The causes of molecular evolution. Oxford: Oxford University Press; 1991.
[53] Zhang J, Nielsen R, Yang Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol 2005;22:2472–9.
[54] Yang Z, Wong WS, Nielsen R. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol 2005;22:1107–18.
[55] Kosakovsky Pond S, Frost SD. Datamonkey:rapid detection of selective pressure on individual sites of codon alignment. Bioinformatics 2005;21:2531–3.
[56] Anisimova M, Bielawski JP, Yang Z. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. Mol Biol Evol 2001;18:1585–92.
[57] Duda Jr TM, Vanhoye D, Nicolas P. Roles of diversifying selection and coordinated evolution in the evolution of Amphibian antimicrobial peptides. Mol Biol Evol 2002;19:858–64.
[58] Semple CAM, Rolfe M, Dorin JR. Duplication and selection in the evolution of primate β-defensin genes. Genome Biol 2003;4:R31.
[59] Morrison GM, Semple CAM, Kilanowski FM, Hill RE, Dorin JR. Signal sequence conservation and mature peptide divergence within subgroups of the murine β-defensin gene family. Mol Biol Evol 2003;20:460–70.
[60] Boniotto M, Antcheva N, Zelezetsky I, Tossi A, Palumbo V, Verga Falzacappa MV, et al. A study of host defense peptide β-defensin 3 in primates. Biochem J 2003;374:707–14.
[61] Del Pero M, Boniotto M, Zuccon D, Cervella P, Spano A, Amoroso A, et al. β-defensin 1 gene variability among non-human primates. Immunogenetics 2002;53:907–13.
[62] Akashi H. Within and between-species DNA sequence variation and the "footprint" of natural selection. Gene 1999;238:39–51.
[63] Crandall KA, Kelsey CR, Imamichi H, Lane HC, Salzman NP. Parallel evolution of drug resistance in HIV: failure of non-synonymous/synonymous rate ratio to detect selection. Mol Biol Evol 1999;16:372–82.
[64] Hughes AL. Rapid evolution of immunoglobulin superfamily C2 domain expressed in immune system cells. Mol Biol Evol 1997;14:1–5.
[65] Hughes AL, Yeager M. Coordinated amino acid changes in the evolution of mammalian defensins. J Mol Evol 1997;44:675–82.
[66] Ota T, Sitnikova T, Nei M. Evolution of immunoglobulin variable gene segments. Curr Top Microbiol Immunol 2000;248:221–45.
[67] Andreu D, Rivas L. Animal antimicrobial peptides: an overview. Biopolymers 1998;47:415–33.
[68] Yeaman MR, Yount NY. Mechanisms of antimicrobial peptide action and resistance. Pharmacol Rev 2003;55:27–55.
[69] Thevissen K, Warnecke DC, Francois IEJA, Leipelt M, Heinz E, Ott C, et al. Defensins from insects and plants interact with fungal glucosylceramides. J Biol Chem 2004;279:3900–5.