GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Homology cloning, sequence characterization, and expression analysis of cDNA encoding electron transfer flavoprotein beta polypeptide in mud crab (*Scylla paramamosain*)

X.L. Hao[1,2], H.F. Yao[2], Y.Z. Cheng[2] and R.X. Wang[2]

[1]College of Life Science, Jilin Normal University, Siping, China
[2]Laboratory for Marine Living Resources and Molecular Engineering, College of Marine Science, Zhejiang Ocean University, Zhoushan, China

Corresponding author: R.X. Wang
E-mail: wangrixin1123@126.com

Genet. Mol. Res. 11 (4): 4316-4322 (2012)
Received November 14, 2011
Accepted June 22, 2012
Published November 12, 2012
DOI http://dx.doi.org/10.4238/2012.November.12.11

ABSTRACT. Electron transfer flavoproteins (ETFs) are αβ-heterodimers found in eukaryotic mitochondria and bacteria. Herein we report a full-length complementary DNA of a mud crab (*Scylla paramamosain*) ETF β subunit (Scpa-ETFB) isolated with a homology cloning strategy. The complete complementary DNA of the Scpa-ETFB contains a 17-nt 5'-untranslated region, a 765-nt open reading frame encoding 254 amino acids, and a 248-nt 3'-untranslated region. The high identity of Scpa-ETFB with ETFB in other organisms indicated that Scpa-ETFB is a new member of the ETFB family. Although the conserved motif associated with flavin adenine dinucleotide binding is absent in Scpa-ETFB, the signature sequences of the ETF superfamily were identified. Using reverse transcriptase polymerase chain reaction, we detected the messenger RNA transcript of Scpa-ETFB in high levels in the tissues of the hepatopancreas, ovary, heart, and muscle. Phylogenetic analysis showed that Scpa-ETFB is most closely related to the ETFB genes of *Caligus rogercresseyi* and *Lepeophtheirus salmonis*.

These results provided basic information for elucidating the molecular mechanism of energy production in the mud crab.

Key words: Mud crab; Electron transfer flavoprotein β subunit; Homology cloning; Energy production

INTRODUCTION

Mud crab (*Scylla paramamosain*; Scpa) is a species of the family Portunidae that inhabits muddy bottoms in brackish water along shoreline, mangroves, and river mouths (Macnae, 1969). In oceans, adult crabs mate inshore and gravid females generally migrate offshore where they spawn (Perrine, 1979). The females then return inshore and can spawn up to 3 times without needing to mate again (Davis, 2003). Owing to its high market demand (Overton et al., 1997) and its advantages of large size and fast growth, mud crab is important for aquaculture and resource conservation. Records of mud crab aquaculture may date back more than 100 years in China (Shen and Lai, 1994), and the artificial culture of this species is currently developing rapidly and contributes significantly to local economies. To date, many immune and reproductive genes have been cloned and have brought new insights to these areas (Liu et al., 2010; Zou et al., 2011); however, the molecular mechanism of energy production remains largely unknown, and few related genes have been reported. Researching mud crab energetics could provide useful information that will improve aquaculture production. Thus, identifying genes associated with energy production in mud crab is important.

Electron transfer flavoprotein (ETF) is a mitochondrial matrix heterodimer with α (ETFA; 30 kDa) and β (ETFB; 28 kDa) subunits. It contains one flavin adenine dinucleotide (FAD) and one adenosine monophosphate molecule per heterodimer (Roberts et al., 1996), which serve as obligatory electron acceptors for at least 11 dehydrogenases found in the matrix of mitochondria or peripherally associated with the matrix side of the inner mitochondrial membrane (Frerman, 1988; Izai et al., 1992). Electrons are passed from the primary dehydrogenases through ETF to ubiquinone via ETF-ubiquinone oxidoreductase, an integral membrane protein containing FAD and a 4Fe4S cluster (Ruzicka and Beinert, 1977; Beckmann and Frerman, 1985). In the present study, we cloned and characterized a complementary DNA (cDNA) encoding the ETFB from mud crab, detected its expression level in the tissue of the hepatopancreas, ovary, heart, and muscle, and examined its phylogenetic relationship with other ETFB genes.

MATERIAL AND METHODS

Three pairs of primers (Table 1) were designed to isolate the full-length cDNA of the ETFB gene. Both 5'-rapid amplification of cDNA ends (RACE) and 3'-RACE were carried out using a Smart RACE cDNA amplification kit (Clontech) according to manufacture instructions. All polymerase chain reaction (PCR) products were resolved via electrophoresis on 1% agarose gels and then purified using a Tiangen Gel Extraction Kit (Tiangen). The purified fragments were ligated into pMD-19T vectors (TaKaRa) and cloned to DH5α cells according to a standard protocol. Positive clones were sequenced via PCR with M13F/13R primers. At least 3 clones were sequenced per fragment using an ABI 3730 automated sequencer with M13F primer. For expression analysis, total RNA was extracted from the tissue of the hepatopancreas, ovary, heart, and muscle using the Trizol reagent (Invitrogen).

First-strand cDNA was synthesized from 2 µg of each total RNA using oligo dT-adaptor primer (TaKaRa). Reverse transcriptase PCR was used to analyze the levels of Scpa-ETFB expression in four tissues. PCR was performed with an initial denaturation step of 3 min at 94°C, and then 30 cycles were run as follows: 30 s denaturation at 94°C, 45 s annealing at 55°C, and a 1-min extension at 72°C, followed by a cycle of 72°C for 5 min. The expression of β-actin was used as an internal control for Scpa-ETFB gene expression analysis. For sequence characterization, putative amino acid sequence alignment was performed using Clustal X (Thompson et al., 1997), and possible motifs were detected. A phylogenetic tree was constructed with the neighbor-joining method of Molecular Evolutionary Genetics Analysis 4.1 (Tamura et al., 2007). The species used are summarized in Table 2. Confidence was determined by bootstrapping 1000 replications.

Table 1. Primers used in this study.

Name	Sequences (5'-3')	Target of amplification
ETFB-3'-RACE-outer-F	ATGGGBGCWGACMGAGGYATCCACGT	For 3'-RACE of Scpa-ETFB
ETFB-3'-RACE-inner-F	TGCCCAACATCATGAAAGCYAAGAAGA	
ETFB-5'-RACE-outer-R	CCTTTGATGCCAGCCTCCCTCACAGGT	For 5'-RACE of Scpa-ETFB
ETFB-5'-RACE-inner-R	GGAGATCAGCTGAGACCACAGCTGGC	
ETFB-RT-F	TTCCAGCCGAGGAGGTTGAG	Expression of Scpa-ETFB
ETFB-RT-R	TGTTGGGCAGGGTGGCATAT	
β-actin-RT-F	TCACCAACTGGGACGACATG	Expression of β-actin
β-actin-RT-R	ATAGCGTGAGGAAGGGCATA	

Table 2. Sequences used in this study (sequences collected from Genbank and Ensembl databases).

Species	Accession No.	Species	Accession No.
Acromyrmex echinatior	EGI65819.1	*Drosophila persimilis*	XP_002013215.1
Ailuropoda melanoleuca	XP_002927967.1	*Drosophila pseudoobscura*	XP_001358978.1
Amphimedon queenslandica	XP_003388745.1	*Drosophila virilis*	XP_002053820.1
Anolis carolinensis	XP_003225134.1	*Drosophila willistoni*	XP_002070200.1
Anopheles darlingi	EFR22327.1	*Equus caballus*	XP_001918034.1
Apis mellifera	XP_393789.2	*Glossina morsitans*	ADD19228.1
Bombus terrestris	XP_003398953.1	*Harpegnathos saltator*	EFN88718.1
Bombyx mori	NP_001040123.1	*Homo sapiens*	NP_001976.1
Bos taurus	NP_001033671.1	*Ixodes scapularis*	XP_002412689.1
Branchiostoma floridae	XP_002609303.1	*Lepeophtheirus salmonis*	ACO11820.1
Caenorhabditis elegans	NP_490973.2	*Loxodonta africana*	XP_003406893.1
Caligus rogercresseyi	ACO10406.1	*Macaca mulatta*	NP_001181636.1
Callithrix jacchus	XP_002762471.1	*Mus musculus*	NP_080971.2
Camponotus floridanus	EFN66681.1	*Nasonia vitripennis*	XP_001608218.1
Ciona intestinalis	XP_002131618.1	*Osmerus mordax*	ACO09178.1
Cricetulus griseus	EGW04637.1	*Pediculus corporis*	XP_002432998.1
Culex quinquefasciatus	XP_001845433.1	*Pongo abelii*	NP_001124625.1
Danio rerio	NP_998163.1	*Rattus norvegicus*	NP_001004220.1
Daphnia pulex	EFX87969.1	*Sus scrofa*	NP_001192208.1
Drosophila ananassae	XP_001955254.1	*Tribolium castaneum*	XP_970257.1
Drosophila grimshawi	XP_001993652.1	*Xenopus* (*Silurana*) *tropicalis*	NP_989154.1
Drosophila melanogaster	AAM52624.1	*Xenopus laevis*	NP_001080044.1
Drosophila mojavensis	XP_002001311.1	Fugu	ENSTRUG00000004992
Medaka	ENSORLG00000015134		

RESULTS AND DISCUSSION

We obtained the full-length mud crab ETFB cDNA using 5'- and 3'-RACE approaches.

The sequences of the cDNA were deposited in GenBank with accession No. JN896342. The determined full-length cDNA of the Scpa-ETFB gene is 1030 nt long, including a 5'-untranslated region (UTR) of 17 nt and a 3'-UTR of 248 nt. One typical polyadenylation signal (AATAAA) is located 6 nt upstream of the poly (A) tail, and an mRNA instability motif (ATTTA) involved in the rapid degradation of mRNA is in the 3'-UTR region (Figure 1). A 765-nt open reading frame encodes a protein of 254 amino acids with a predicted molecular weight of 27.8 kDa and a predicted isoelectric point of 6.4. BlastP similarity comparisons clearly indicated that Scpa-ETFB shares high similarity with ETFB genes from other species. The members of the ETFA family reportedly exhibit little identity throughout their N-terminal portions; their C-terminal regions exhibit a high degree of sequence identity.

```
ggccattcaccgatacaATGTCTCTCCGGGTACTTGTTGGCGTCAAGAGAGTGATTGATTATGCCGTGAAGATTCGGGT 79
                  M  S  L  R  V  L  V  G  V  K  R  V  I  D  Y  A  V  K  I  R  V    21
CCGGCCAGACAAGCTTGGGGTGGTGACAGATGGCGTGAAACACTCCATGAACCCTTTCGATGAGATAGCCATTGAGGAG 158
 R  P  D  K  L  G  V  V  T  D  G  V  K  H  S  M  N  P  F  D  E  I  A  I  E  E     47
GCCGTGCGGCTCAAGGAGAAGAAGATTGCAAAGGAGGTGGTGGCTGTGTCATGTGGCCCAACCCAGGCACAGGAGACCA 237
A  V  R  L  K  E  K  K  I  A  K  E  V  V  A  V  S  C  G  P  T  Q  A  Q  E  T  I    74
TCCGCACTGCCCTGGCCATGGGGGCTGACAGAGGTATCCACGTAGAGATTCCAGCCGAGGAGGTTGAGAAACTAGAACC 316
  R  T  A  L  A  M  G  A  D  R  G  I  H  V  E  I  P  A  E  E  V  E  K  L  E  P    100
ATTGCATGTAGCCAAGATTATGGCCAAGCTGGTGGAGCAGGAGAAGGCTGACTTAGTTGTGCTTGGCAAACAGGCTATT 395
 L  H  V  A  K  I  M  A  K  L  V  E  Q  E  K  A  D  L  V  V  L  G  K  Q  A  I     126
GATGATGACTCCAATGCCACTGCCCAGATGACAGCTTCCATCCTTGACTGGCCTCAGGCAACCTTTGCTTCCAAGATTG 474
D  D  D  S  N  A  T  A  Q  M  T  A  S  I  L  D  W  P  Q  A  T  F  A  S  K  I  E   153
AGAAGACTGATGGTGAGCTGCAGGTGACAAGAGAGGTGGATGGAGGTCTGGAGACAATCAAAGTGAAGTTGCCAGCTGT 553
  K  T  D  G  E  L  Q  V  T  R  E  V  D  G  G  L  E  T  I  K  V  K  L  P  A  V    179
GGTCTCAGCTGATCTCCGCCTCAATGAGCCCCGATATGCCACCCTGCCCAACATCATGAAAGCTAAGAAGAAGAAAATT 632
 V  S  A  D  L  R  L  N  E  P  R  Y  A  T  L  P  N  I  M  K  A  K  K  K  K  I     205
GCTAAGATGAAGGCTGCTGACCTTGGAGTAGATACAACTTCTCACTTTGAAGTTTTGGAGGTGGCTGATCCACCTGTGA 711
A  K  M  K  A  A  D  L  G  V  D  T  T  S  H  F  E  V  L  E  V  A  D  P  P  V  R   232
GGGAGGCTGGCATCAAAGTGGAGGATGTGGACACTCTTATTACAAAGTTGAAGGAGACTGGGCGCATTTAGcacagtta 790
  E  A  G  I  K  V  E  D  V  D  T  L  I  T  K  L  K  E  T  G  R  I  *             254
atcggggacgtacagtgtgactgtggttggaggaattgcaccatacctgtacacctgccaccaatcaatgcaatatacc 869
tattttctttgtgggcaagctatccttaatgaatgtgctacttttcagagatttctttaatacacaatctgtattata 948
ttttcttgtacaaaacattctatagaaacaggtttaataaaatatataaaacaaaaaaaaaaaaaaaaaaaaaaaaa 1027
aaa                                                                              1030
```

Figure 1. Complete nucleotide and deduced amino acid sequence of Scpa-ETFB. The termination codon (TAG) was marked with an asterisk. The motif associated with mRNA instability (ATTTA) and the poly-adenylation signal sequences were shadowed.

In contrast to those of ETFA family proteins, the complete multiple alignments of the proteins of the ETFB family showed an approximately equal degree of sequence similarity throughout their lengths (Figure 2). A putative signature sequence (REXDGGLEX$_6$PXVX$_2$VDLRLNXPRYATLPNIMKAKKK; see Figure 2, positions 167-207) was detected in the Scpa-ETFB protein, which differs slightly from previously reported ETFB signature sequences (Tsai and Saier, 1995). Many nucleotide-binding proteins, including flavoproteins, share a consensus sequence, GXGXXG, that is preceded by a short hydrophobic β strand (Wierenga and Hol, 1983; McKie and Douglas, 1991; Baker et al., 1992). In humans, the ETFB subunit has an approximation to this motif in the sequence KXGXXG. However, Figure 2 shows that of these residues, only the last G in the motif (positions 213-218) is conserved. Furthermore, no preceding hydrophobic β-strand-forming segment was observed. Tsai and Saier (1995) have reported a similar result, in which the GXGXXG or KXGXXG motif was absent in the multiple alignments of ETFB proteins of microorganism. These results suggested that this region in Scpa-ETFB is therefore unlikely to represent part of FAD-binding sites, and these sites could be examined

through biochemical approaches. By contrast, the fully conserved sequence, AXGXXG, is found in the ETFB proteins (Tsai and Saier, 1995).

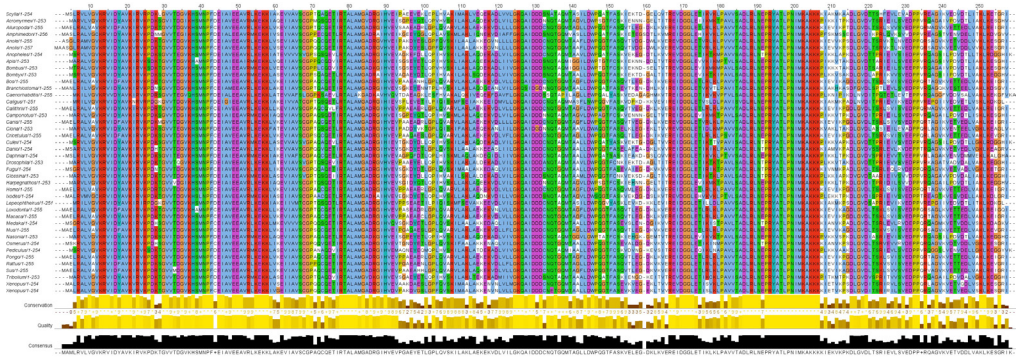

Figure 2. Multiple alignment of the deduced amino acid sequence of Scpa-ETFB with other known ETFB cDNA from various species.

The Scpa-ETFB mRNA levels in the tissues of the hepatopancreas, ovary, heart, and muscle were measured using reverse transcriptase PCR. As shown in Figure 3, Scpa-ETFB was strongly expressed in these 4 tissues.

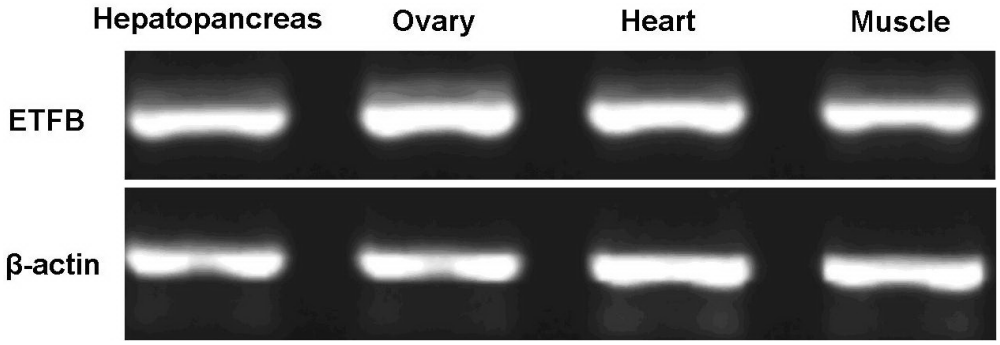

Figure 3. Expression analysis of Scpa-ETFB in hepatopancreas and ovary. As a positive control for RT-PCR, β-actin was amplified to determine the concentration of template.

A phylogenetic tree generated using the neighbor-joining method resolved 2 major clades (invertebrate and vertebrate), suggesting that these genes were derived from a common ancestral gene. Scpa-ETFB formed a cluster with known non-vertebrate ETFB and is most closely related to the ETFB genes of *Caligus rogercresseyi* and *Lepeophtheirus salmonis* (Figure 4).

Taken together, the present study isolated a full-length cDNA of the mud crab ETFB gene and analyzed its sequence characterizations. These results will facilitate future research of the molecular mechanism of energy production in mud crab.

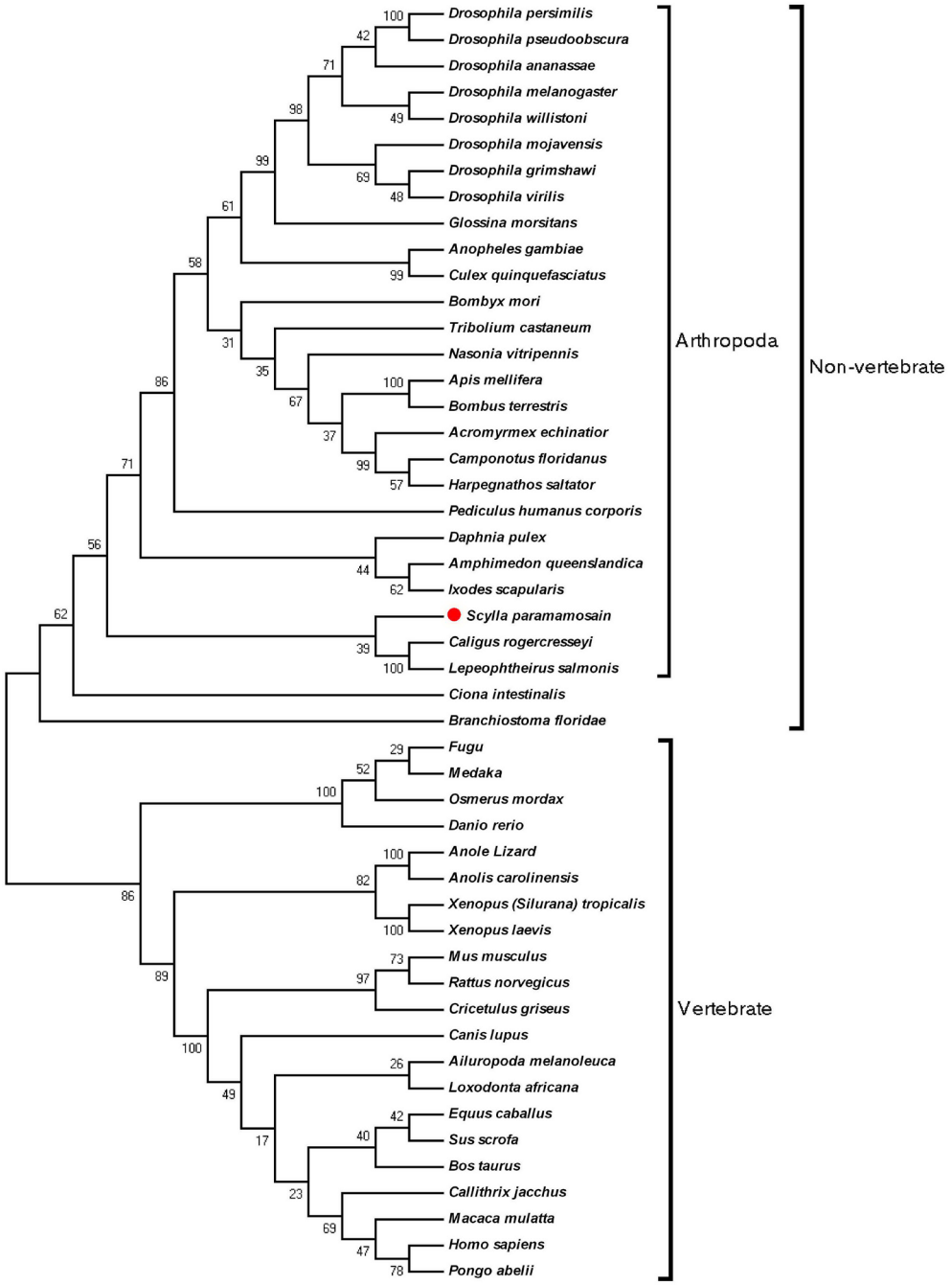

Figure 4. Neighbor-joining phylogenetic tree of the ETFB gene from *Scylla paramamosain* and other species. The position of Scpa-ETFB was marked.

ACKNOWLEDGMENTS

Research supported by the Important Science and Technology Specific Projects of Zhejiang Province (#2009C12080, #2011C12011).

REFERENCES

Baker PJ, Britton KL, Rice DW, Rob A, et al. (1992). Structural consequences of sequence patterns in the fingerprint region of the nucleotide binding fold. Implications for nucleotide specificity. *J. Mol. Biol.* 228: 662-671.

Beckmann JD and Frerman FE (1985). Electron-transfer flavoprotein-ubiquinone oxidoreductase from pig liver: purification and molecular, redox, and catalytic properties. *Biochemistry* 24: 3913-3921.

Davis JA (2003). Development of Hatchery Techniques for the Mud Crab *Scylla serrata* (Forskål) in South Africa. PhD Thesis, Ghet University, Ghent.

Frerman FE (1988). Acyl-CoA dehydrogenases, electron transfer flavoprotein and electron transfer flavoprotein dehydrogenase. *Biochem. Soc. Trans.* 16: 416-418.

Izai K, Uchida Y, Orii T, Yamamoto S, et al. (1992). Novel fatty acid beta-oxidation enzymes in rat liver mitochondria. I. Purification and properties of very-long-chain acyl-coenzyme A dehydrogenase. *J. Biol. Chem.* 267: 1027-1033.

Liu HP, Chen RY, Zhang M and Wang KJ (2010). Isolation, gene cloning and expression profile of a pathogen recognition protein: a serine proteinase homolog (Sp-SPH) involved in the antibacterial response in the crab *Scylla paramamosain*. *Dev. Comp. Immunol.* 34: 741-748.

Macnae W (1969). A general account of the fauna and flora of mangrove swamps and forests in the Indo-West-Pacific region. *Adv. Mar. Biol.* 6: 73-103.

McKie JH and Douglas KT (1991). Evidence for gene duplication forming similar binding folds for NAD(P)H and FAD in pyridine nucleotide-dependent flavoenzymes. *FEBS Lett.* 279: 5-8.

Overton JL, Macintosh DJ and Thorpe RS (1997). Multivariate analysis of the mud crab *Scylla serrata* (Brachyura: Portunidae) from four locations in Southeast Asia. *Mar. Biol.* 128: 55-62.

Perrine D (1979). The Mangrove Crab on Ponape. Marine Resources Division, Ponape, Eastern Caroline Islands.

Roberts DL, Frerman FE and Kim JJ (1996). Three-dimensional structure of human electron transfer flavoprotein to 2.1-A resolution. *Proc. Natl. Acad. Sci. U. S. A.* 93: 14355-14360.

Ruzicka FJ and Beinert H (1977). A new iron-sulfur flavoprotein of the respiratory chain. A component of the fatty acid beta oxidation pathway. *J. Biol. Chem.* 252: 8440-8445.

Shen Y and Lai Q (1994). Present Status of Mangrove Crab (*Scylla serrata* (Forskål)) Culture in China. The ICLARM Quarterly. *Naga* 17: 28-29.

Tamura K, Dudley J, Nei M and Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596-1599.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, et al. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876-4882.

Tsai MH and Saier MH Jr (1995). Phylogenetic characterization of the ubiquitous electron transfer flavoprotein families ETF-alpha and ETF-beta. *Res. Microbiol.* 146: 397-404.

Wierenga RK and Hol WG (1983). Predicted nucleotide-binding properties of p21 protein and its cancer-associated variant. *Nature* 302: 842-844.

Zou Z, Zhang Z, Wang Y, Han K, et al. (2011). EST analysis on the gonad development related organs and microarray screen for differentially expressed genes in mature ovary and testis of *Scylla paramamosain*. *Comp. Biochem. Physiol. Part D Genomics Proteomics* 6: 150-157.