BIOTROPIA VOL. 14 NO. 2, 2007 : 22 - 31

PHYLOGENETIC ASSESSMENT OF GIANT CLAMS (*TRIDACNIDAE*) USING PARTIAL AMINO ACIDS SEQUENCES OF CYTOCHROME C OXIDASE I GENE

AGUS NURYANTO[1], DEDI DURYADI[2],
D. SOEDHARMA[3], AND BLOHM D[4]

[1] *Faculty of Biology, Jenderal Soedirman University, Purwokerto*
[2] *Department of Biology, Faculty of Mathematics and Life Sciences, Bogor Agriculture University*
[3] *Faculty of Fisheries and Marine Sciences, Bogor Agriculture University*
[4] *Department of Biotechnology and Molecular Genetics, FB2-UFT, University of Bremen*

ABSTRACT

Phylogenetic assessment of the giant clams need to be improved using various genetic markers since their relaionships are plagued by inconsistency result of several studies. This is especially true for the species under sub-genus *Chametrachea*. Here we studied the phylogeny of the giant clams using amino acid sequence of the mitochondrial COI gene. This study is aimed to assess and better understand the relationships of the giant clams, especially for three species under subgenus *Chametrachea*. The result showed close proximity between *T. crocea* and *T. squamosa* and between *T. maxima* and *T. gigas*. However, amino acids sequences of cytochrome c oxidase I gene was not strong enough to discriminate between *T. crocea* and *T. squamosa*. The majority of mutations were hydrophobic to hydrophilic amino acids.

Key words: phylogenetic assessment, *Chametracea*, amino acid, cytochrome c oxidase I

INTRODUCTION

Phylogenetic relationships of the giant clams are plagued by uncertain results of several studies using different molecular taxonomic character. This is true especially for the species under sub-genus *Chametrachea* (*Tridacna crocea, T. maxima* and *T. squamosa*).

A study using protein electrophoresis data supported the relationships among the members of giant clams as those observed based on morphological character (Benzie and Williams 1998) and showed that *T. crocea* was closely related to *T. maxima* than to *T. squamosa*. However, studies using DNA-based marker showed different pattern of phylogenetic relationships of the giant clams *T. crocea, T. maxima* and *T. squamosa*, whereas relationships between *T. gigas, T. derasa, T. tevoroa, Hippopus hippopus* and *H.*

porcellanus were clear (Benzie and Williams 1998, Maruyama *et al.* 1998, Schneider and O'Foighil 1999). This inconsistence could be due to the difference of the resolving power of the markers. For example protein has a lower mutation rate than DNA-based marker and mutation rate of animal nuclear DNA is lower compared to those of mitochondrial DNA (Sunnuck 2000, Randi 2000). This suggested to further study the phylogeny of giant clam using various genetic markers.

The family of *Tridacnidae*, also known as giant clams, is conspicuous bivalve inhabitants of coral reef across the Indo-Pacific region (Lucas 1988). Total length of the adult individuals of this clam ranges from 15 cm in *Tridacna crocea* to 150 cm in *T. gigas*. There is increasing interest to giant clams, not only due to their beautiful colour pattern but also because of a remarkable decline of natural populations occurred throughout their geographic range due to over-fishing (Lucas 1988) and environmental deterioration (Pandolfi *et al.* 2003). Developing of mariculture method and conservation technology of this clam to overcome the problem has led a paramount research on clam biology (Copland and Lucas 1988). However, their systematic and phylogeny has been given little interest (Schneider and O'Foighil 1999), whereas that information is vital for marine resources conservation and management.

Given that the relationship among the species under sub-genus *Chametrachea* is still ambiguous and that all giant clams are listed as vulnerable species on the CITES red list data book (Wells 1997), is imperative to understand their systematic and phylogeny so that decisions can be made concerning their conservation.

Here we assessed the phylogeny of giant clams using partial sequences of amino acids of the mitochondrial cytochrome c oxidase I gene. The evolution of this gene is rapid enough to allow the discrimination of not only closely related species, but also phylogeographic groups within single species (Wares and Cunningham 2001). Insertion and deletion (indels) were uncommon in this gene (Hebert *et al.* 2003a). Its third-position nucleotides show a high incidence of base substitution, leading to a rate of molecular evolution that is about three times greater than that of 12S or 16S rDNA (Knowlton and Weigt 1998). However, its amino acid changes more slowly than in cytochrome b or any other mitochondrial gene (Lynch and Jarrell 1993).

On nucleotide level, COI has two important advantages. First, the universal primers for this gene are very robust, enabling recovery of its 5' end from representatives of most animal phyla (Folmer *et al.* 1994). Second, COI appears to possess a greater range of phylogenetic signal than any other mitochondrial gene (Hebert *et al.* 2003a). Moreover, this COI allowed the recognition of genetic divergence of conspecific populations associated with geographic isolation (Bucklin *et al.* 1999, Bucklin *et al.* 2003). It has no recombination processes (Hebert *et al.* 2003b) and has broader phylogenetic sign than alternatives such as cytochrome b (Simmons and Weller 2001). However, the potential of the COI at amino acids level for phylogenetic study is not much known and need to be assessed further.

This study is aimed to assess and better understand the phylogeny of giant clams, especially for three species under subgenus *Chametrachea*.

MATERIALS AND METHODS

Sample collection and DNA isolation

Mantle tissues of *T. crocea* were collected from Pulau Seribu, Spermonde, and Biak, whereas *T. maxima* samples were collected from Padang, Pulau Seribu, Spermonde, Togian Islands and Biak during the field trips in 2004 and 2005 (Figure 1). Tissue samples of *T. gigas* were collected from Togian Islands, while *T. squamosa* samples were collected from the Red Sea in 2004. DNA analysis was carried out at the Department of Biotechnology and Molecular Genetics, FB2-UFT, University of Bremen, Germany

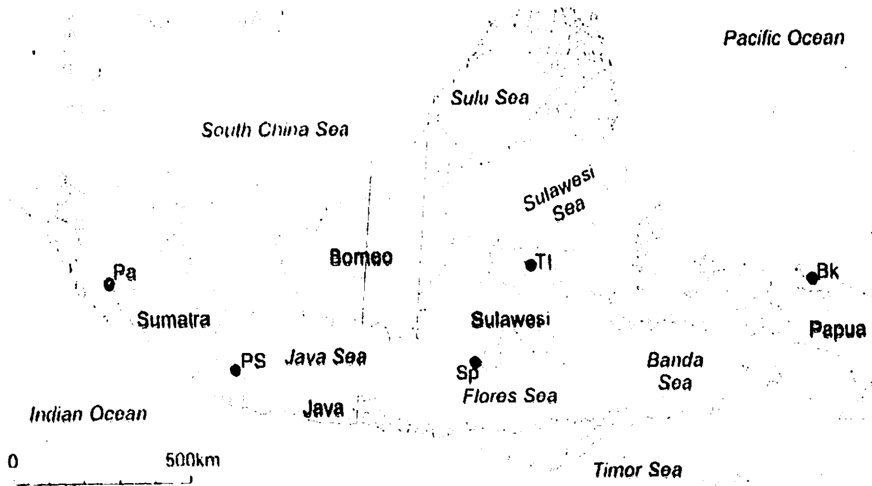

Figure 1 Map of the Indo-Malay Archipelago showing samples sites. Pa = Padang, PS = Pulau Seribu, Sp = Spermonde, TI = Togian Islands, Bk = Biak.

A small piece of mantel tissues were cut off from each specimen under water with the help of forceps and scissors. This sampling method was performed to minimize negative effect of sampling activity to population. Tissue samples were preserved on 96% of ethanol and stored at 4°C until DNA analysis. Total genomic DNA was isolated using Chelex® method following the protocols from Walsh *et al.* (1991).

Amplification and sequencing

A fragment of Cytochrome c oxidase I (COI) gene was amplified using tridac-

nid-specific primers (forwards: LCO: 5'-GGG TGA TAA TTC GAA CAG AA-3' and reverse: RCO: 5'-TAG TTA AAG CCC CAG CTA AA-3') which were designed based on *T. crocea* sequences obtained in a preliminary analysis using COI primer from Folmer *et al.* (1994). These primers amplified a fragment of COI gene with approximately 530 bp length (Figure 2). PCR reactions were carried out in a total volume of 50 µl containing approximately10 picogram of DNA template, 1 x PCR buffer, 2 mM of MgCl2, 0.02 µM of each primer, 0.2 mM of each dNTPs and 1 unit Taq polymerase. Thermal cycles were as follows: one cycle at 94 oC for 5 minutes, follows by 35 cycles of 1 minute at 94 °C, 1.5 minutes annealing temperature (at 50 °C for T. *maxima* and *T. gigas*, 43 °C for *T. crocea* and 47 °C for T. squamosa, respectively) and 1 minute at 72 °C for extension. Final extension was carried out at 72 °C for five minutes.

The PCR products were purified using the PeqGOLD cycle-pure kit (Peqlab Biotechnologie GmbH, Erlangen, Germany) and QIA quick purification kit (Qiagen, Hilden, Germany) following the protocol from the manufacturer. Both strands were sequenced using the DyeDeoxy terminator chemistry (PE Biosystem, Foster City) and an automated sequencer (ABI prism 310; Applied Biosystem, Weiterstadt).

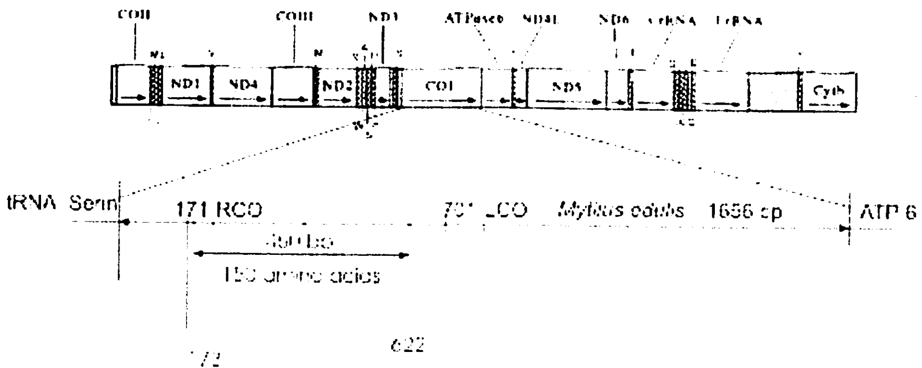

Figure 2 The position of the primers and the fragment of the mitochondrial COI gene on the phylogenetic analysis

Phylogenetic analysis

All sequences were initially aligned and edited manually using Sequences Navigator (version 1.0.1; Applied Biosystem). Multiple sequences alignment was done using ClustalW as implemented in Bioedt ver. 7.0.4.1 (Hall 1999). All sequences were translated into amino acids sequences with Squint software using the third reading frame (http://www.cebl.auckland.ac.nz/bi/index.php).

Molecular phylogenetic analysis was based on 150 amino acids sequences from four species of *Tridacna* (*T. crocea* = 4, *T. maxima* = 5, *T. squamosa* = 1, and *T. gigas*

= 1 individual, respectively), with the sequences *Pavicardium* AF120664, *Ruditapes* AY874536 and *Mytilus* AY484747 as out groups obtained from GenBank. Phylogenetic trees were constructed with MEGA programme version 2 (Kumar *et al.* 2001) using neighbour-joining method. Support for tree branching was based on 1000 non-para metric bootstrap replicates.

RESULTS AND DISCUSSION

We translated a length of 450 bp of mtDNA Cytochrome c oxidase I gene from four species of giant clams (*T. crocea*, 4 individuals; *T. maxima*, 5 individuals; *T. squamosa*, 1 individual and *T. gigas*, 1 individual) and from *Mytilus*, *Ruditapes* and *Pavicardium* into amino acid sequences. The translation of those COI sequences resulted in 150 amino acids sequences. The first codon is starting from nucleotide number one.

Figure 3 is a neighbor joining tree showing the phylogeny of giant clams based on 150 amino acids. The tree showed that *T. crocea* and *T. squamosa* are monophyletic taxa, whereas *T. maxima* and *T. gigas* formed another monophyletic group. This dis crimination between the two monopyletic groups are supported by high bootstrap value (100). Different to the analysis based on nucleotide sequences, the placement of *T. squamoasa* was mixed among *T. crocea* specimens and was not clearly discriminated but this pattern was supported by a rather weak bootstrap value (79). The separation of *T. maxima* and *T. gigas* was clear as those observed, although only supported by rather weak bootstrap value of 76 (Figure 3).

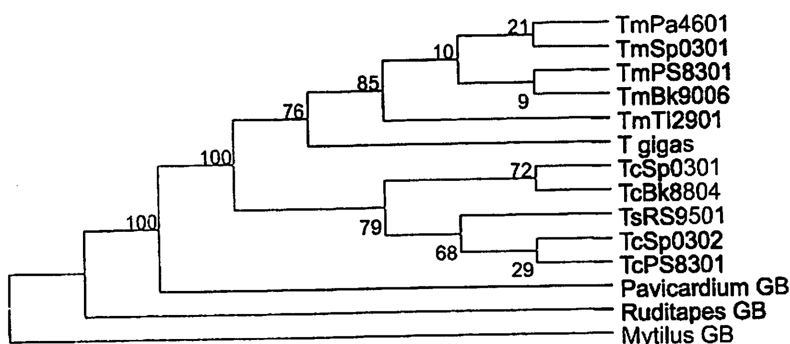

Figure 3 Neighbor joining tree based on 150 amino acids of the mitochondrial COI acording to uncorrected genetic distances (p-distances) and bootstrap analysis with 1000 replicates. Tc = *Tridacna crocea*, Tm = *Tridacna maxima*, Ts = *Tridacna squamosa*, Bk = Biak, Pa = Padang, PS = Pulau Seribu, RS = Red Sea, SP = Spermonde, TI = Togean Islands.

In order to know the differences on the amino acid (Figure 4) level among giant clam species, we performed an analysis of shared amino acid by all possible pairwise of giant clams species only (Table 1). There are differences in certain amino acids site shared among species. On the first site, amino acid glycine (G) is shared by *T. crocea, T. maxima* and *T. squamosa*, while *T. gigas* has serine (S). So glycine has undergone mutation to serine. At the site number seven, all species have isoleucine except two individuals of *T. crocea* which have threonine. It is suggested that the seventh amino acid on *T. crocea* has undergone mutation from isoleucine to threonine (Figure 4).

Table 1 Number of shared amino acid by all possible pairwise of giant clams species

Possible pairwise	Shared amino acids of 150
T. crocea , T. maxima , T. Squamosa, T. gigas	144 (96.00%)
T. maxima, T. gigas	148 (98.67%)
T. crocea, T. squamosa	149 (99.33%)
T. crocea, T. gigas	144 (96.00%)
T. squamosa, T. gigas	145 (96.67%)
T. maxima, T. crocea	146 (97.33%)
T. maxima, T. squamosa	147 (98.00%)

There are certain amino acids shared by *T. crocea* and *T. squamosa* at certain sites. At site number 18 *T. crocea* and *T. squamosa* shared alanine, at site number 88 was amino acid valine, at site number 148 was arginine (R). In contrast, at the same sites, T. maxima and T. gigas shared amino acid serine, methionine, and leucine, respectively (Figure 4). Figure 4 also shows that *T. gigas* has specific amino acid (serine) which can be found at the first and the 75th sites.

It was difficult to amplify COI gene from all giant clam mantle tissue using universal primer from Folmer *et al.* (1994). To overcome this problem, we tried to de sign other pair of universal primer (COI-F and COI-R) based on the COI sequences of several bivalves available on GenBank. Then we designed a pair of tridacnid-specific primers (LCO and RCO) which were designed based on *T. crocea* sequences obtained in a preliminary analysis using of COI primer from Folmer *et al.* (1994). However, the problem was still unresolved until the end of the study. These difficulties could be due to a high proportion of polysaccharides in their tissues. Polysaccharides can inhibit the activity of some enzymes such as polymerase (Skolov 2000).

Neighbour-joining tree (Figure 3) showed the monophyly of giant clam. Al though *T. crocea* specimens were collected from geographically structured populations they formed a monophyletic group. The same pattern was observed on *T. maxima* specimens as well. This indicates that geographic distance rather showed geographic variations than formed a separate monophyletic group within species. In other hand, high genetic diversity on COI gene was sufficient to reveal geographically structured populations of those organisms.

```
173/1
TcSp0301   GDNSNSTAWP ASWLQNNALY NVIVTTHALI MIFFMVMPVM MGGFGNWLVP
TcSp0302   ......I... .......... .......... .......... ..........
TcPS8301   ......I... .......... .......... .......... ..........
TcBk8804   .......... .......... .......... .......... ..........
TmPa4601   ......I... .......S.. .......... .......... ..........
TmPS8301   ......I... .......S.. .......... .......... ..........
TmSp0301   ......I... .......S.. .......... .......... ..........
TmTI2901   ......I... .......S.. .......... .......... ..........
TmBk9006   ......I... .......S.. .......... .......... ..........
TsRS9501   ......I... .......... .......... .......... ..........
T_gigas    S.....I... .......S.. .......... .......... ..........

TcSp0301   LMMVMPDMHF PRLNNLSFWF VPNAFFLLGV SGFVEGGVGA GWTIYPPLTS
TcSp0302   .......... .......... .......... .......... ..........
TcPS8301   .......... .......... .......... .......... ..........
TcBk8804   .......... .......... .......... .......M.. ..........
TmPa4601   .......... .......... .......... .......M.. ..........
TmPS8301   .......... .......... .......... .......M.. ..........
TmSp0301   .......... .......... .......... .......M.. ..........
TmTI2901   .......... .......... .......... .......M.. ..........
TmBk9006   .......... .......... .......... .......M.. ..........
TsRS9501   .......... .......... .......... .......... ..........
T_gigas    .......... .......... ....S..... .......M.. ..........

                                                            622/150
TcSp0301   IDFLSDPSMD LAIFSLHLGG ASSIAASLNF ASTVANMRHQ KRGFHKIRCF
TcSp0302   .......... .......... .......... .......... ..........
TcPS8301   .......... .......... .......... .......... ..........
TcBk8804   .......... .......... .......... .......... ..........
TmPa4601   .......... .......... .......... .......... .......L..
TmPS8301   .......... .......... .......... .......... .......L..
TmSp0301   .......... .......... .......... .......... .......L..
TmTI2901   .......... .......... .......... .......... .......L..
TmBk9006   .......... .......... .......... .......... .......L..
TsRS9501   .......... .......... .......... .......... ..........
T_gigas    .......... .......... .......... .......... .......L..
```

Figure 4 Alignment of 150 amino acid sequences of the mitochondrial DNA cytochrome c oxidase I gene. Dots indicate amino acids identity to the amino acid sequence of Tridaacna crocea from Spermonde. 173 is indicating nucleotide position compared to COI gene of Mytilus. 1 is indicating the first amino acid within the fragment.

T. crocea and *T. squamosa* belong to a monophyletic group, *T. maxima* and *T. gigas* formed another monophyletic group. The separation of both monophyletic groups is supported by high bootstrap value (100, Figure 3). However, the placement of *T. squamoasa* was mixed among *T. crocea* specimens and was not clearly discriminated (Figure 3). Unclear separation between *T. crocea* and *T. squamosa* could be due to a silent mutation that occurred on the amino acid number 18 (Alanine). On the amino acids level, the third codon position for that amino acid was C (cytosine) for *T. crocea* while for *T. squamosa* was T (thymine). According to Hebert *et al.* (2003a), COI amino acids sequences was sufficient to reliable assignment and the placement of organism to higer taxonomic categories. However, it shows sufficient nucleotide sequence diversity to

enable animal discrimination even on the level of species identification (Hebert *et al.* 2003b). The separation of *T. maxima* and *T. gigas* was clear as those observed, alhtough was only supported by rather weak bootstrap value of 76 (Figure 3).

Compared to Schneider and O'Foighil's study with 16S rRNA, in our NJ: 1) *T. crocea*, *T. maxima*, *T. squamosa* and *T. gigas* belong to monophyletic group compared to the outgroup, 2) within giant clams group, *T. crocea* and *T. squamosa* belong to monophyletic group (100 NJ bootstrap) and were sister taxa thus formed sister group to *T. maxima* and *T. gigas*, whereas in Schneider and O'Foighil (1999), all *Chametrachea* are placed in the same monophyletic group. This difference could be due to *T. derasa*, *H. hippopus* and *H. porcellanus* which were not included in the study. Therefore, we could not resolve the separation of *Chametrachea* sub-genus into two groups clearly and put *T. gigas* and *T. maxima* in one monophyletic group. At least our study supports the result from Schneider and Foighil (1999) and Roa-Quiaoit (2005) by showing that *T. crocea* and *T. squamosa* have close affinity. Close proximity between *T. crocea* and *T. squamosa* was also supported by the number of shared amino acids (99.33%, Table 1). In contrast, on the second monophyletic group, close relationship between *T. maxima* and *T. gigas* was supported by 98.67% amino acids similarity (Table 1).

Close relationship between *T. crocea* and *T. squamosa*, and *T. maxima* and *T. gigas* implied that within each group they should be managed in a similar way in regard of their conservation. However, to avoid a mistake on the conservation management of those giant clams due to unclear taxonomic status as indicated in this study, further study using longer amino acids sequence or other molecular markers are necessary. However, the result of this study is expected to enrich and to improve the COI sequences database on the GenBank database, especially for giant clams species. A COI database could serve as the basis for a global bio-identification system (GBS) for animals. The GBS would aim for comprehensive taxonomic coverage of just a single gene (Hebert *et al.* 2003b).

As additional information, it seemed that only small numbers of mutations were found on amino acid level (6 mutations). Among six mutation events with reference sequence T. crocea from Spermonde (TcSP0301), one mutation was from polar (hydrophilic) to polar amino acid (G to S), four mutations occurred from hydrophobic (non polar) to hydrophilic amino acids (T to I, A to S, F to S, and V to M, respectively) and one mutation from positively charged to naturally charged hydrophobic amino acid (R to L) or vice versa depending on the reference sequence. However, it appeared that type of mutations (substitutions, e.g. synonymous or synonymous mutations) played important rules in branching pattern on the phylogenetic tree rather than alteration of chemical state (e.g. hydrophobic and hydrophilic) of amino acids.

CONCLUSIONS

According to our result T. crocea and T. squamosa were closely related and formed monophyletic group. Together with *T. gigas*, *T. maxima* they were placed on the other monophyletic group and formed sister group to *T. crocea* and *T. squamosa*. These implied that T. crocea and T. squamosa should be managed in similar ways as well *T. maxima* and *T. gigas*. The majority of mutations were from hydrophobic to hydrophilic amino acid. Type of mutations played an important rule on branching pattern rather than chemical state of amino acids. This study enriched the taxonomic character which confirmed the result of previous study using nucleotide sequences. However, amino acid marker was not strong enough to discriminate *T. crocea* and *T. squamosa.*

ACKNOWLEDGEMENT

We would like to thank to the institutions and individuals who have made our study possible, the German Federal Ministry of Education and Research (BMBF, grant no. 03F0390B), which funded the project "Molecular Genetics as a Tool for the Management of Marine Ornamentals in Sulawesi (Indonesia)" in the frame work of SPICE (Science for the Protection of Indonesian Coastal Marine Ecosystems); German Academic Exchange Service (DAAD) for supporting A. Nuryanto.

REFERENCES

Benzie, J.A.H. and S.T. Williams. 1998. Phylogenetic relationships among giant clam species (Mollusca: Tridacnidae) determined by protein electrophoresis. Mar. Biol. 132:123-133.

Bucklin, A., Guarnieri M, Hill RS, Bentley AM, and S. Kaartvedt. 1999. Taxonomic and systematic assessment of planktonic copepods using mitochondrial COI sequence variation and competitive, species-specific PCR. Hydrobiologia 401: 239-254.

Bucklin, A., B.W. Frost, J. Bradford-Grieve, L.D. Allen, and N.J. Copley. 2003. Molecular systematic and phylogenetic assessment of 34 calanoid copepod species of the Calanidae and Clausocalanidae. Mar. Biol. 142: 333-3343.

Copland, J.W. and J.S. Lucas. 1988. Giant Clams in Asia and the Pacific. Australian Centre for International Agricultural Research, Canberra.

Folmer, O., M. Black, R. Lutz and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from metazoan invertebrates. Molec. Mar. Biol. and Biotech. 3: 294-299.

Hall, B.G. 2001. Phylogenetic Trees Made Easy: A How-to Manual for Molecular Biologist. Sinauer Associates Inc. Publisher, Massachusetts.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series. 41:95-98.

Hebert, P.D.N., A. Cywinska, S.L. Ball and J.R. de Waar. 2003. Biological identification through DNA barcodes. Proc. R. Soc. London B. 270:313-321.

Hebert, P.D.N., R. Ratnasingham and J.R. de Waar. 2003b. Barcoding animal life: cytochrome c oxidase subunit I divergences among closely related species. Proc. Royal Soc. London B 270: 96-99.

Knowltom, N. and L.A. Weigt. 1998. New dates and new rates for divergence across the Isthmus of Panama. Proc.Royal Soc. London B 265: 2257-2263

Kumar, S., K. Tamura, I.B. Jakobsen and M. Nei. 2001. MEGA2: Molecular Evolutionary Genetics Analysis software. Bioinformatics (In press).

Lucas, J.S. 1988. Giant Clams: Description, Distribution and Life History. In: Copland JW, Lucas JS, editor. Giant Clams in Asia and the Pacific. Australian Centre for International Agricultural Research. Canberra. p. 21-32.

Lynch, M. and P.E. Jarrell. 1993. A method for calibrating molecular clocks and its application to animal mitochondrial DNA. Genetics 135: 1197-1208.

Maruyama, T., M. Ishikura, S. Yamazaki and S. Kanai. 1998. Molecular phylogeny of zooxanthellate bivalves. Biol. Bull. 195:70-77.

Pandolfi, J.M., R.H. Bradbury, E. Sala, T.P. Hughes, K.A. Bjorndal, R.G. Cooke, D. McArdle, L. McClenachan, M.J.H Newman, P. Gustavo, R.R. Warner and J.B.C. Jackson. 2003. Global trajectories of the long-term decline of coral reef ecosystem. Sci. 301(5635): 955-958.

Randi, E. 2000. Mitochondrial DNA. In: Baker JA: Molecular Methods in Ecology. Blackwell Science, Toronto.

Roa-Quiaoit, H.A.F. 2005. The ecology and culture of giant clams (Tridacnidae) in the Jordanian sector of the Gulf of Aqaba, Red Sea. Dissertation. University of Bremen, Bremen.

Schneider, J.A. and D. O'Foighil. 1999. Phylogeny of giant clams (Cardiidae: Tridacnidae) based on partial mitochondrial 16S rDNA gene sequences. Molec. Phylogenet. Evol. 13 (1):59-66.

Simmons, R.B., S.J. Weller. 2001. Utility and evolution of cytochrome b in insects. Molec. Phylogenet. Evol. 20: 196-210.

Skolov, E,P. 2000. An improve method for DNA isolation from mucopolysaccharide-rich molluscan tissues. J. Molec. Stud. 66: 573-575.

Sunnucks, P. 2000. Efficient genetic marker for population biology. TREE.15 (5):199-203.

Walsh, P.S., D.A. Metzger and R. Higushi. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques. 10 (4): 506-513.

Wares, J.P. and C.W Cunningham. 2001. Phylogeography and historical ecology of the North Atlantic intertidal. Evolution Journal 12: 2455-2469.

Wells, S. 1997. Giant clams: status, trade and mariculture, and the role of CITES management. IUCN, Switzerland and Cambridge.