Sequence comparison of the ribosomal DNA internal transcribed spacers (ITSs) and 5.8S ribosomal gene and their secondary structures in 17 Pteriomorphian bivalves
K Ashokan, M Pillai, S Angadi, D Mundaganur

INTRODUCTION
The internal transcribed spacer (ITSs) of nuclear ribosomal DNA (rDNA) is one of the most extensively sequenced markers, and the region is a component of an rDNA cistron, which consists of 18S, ITS1, 5.8S, ITS2 and 28s. ITSs, exists in several hundred copies in most eukaryotes. They are located in one or several loci and distributed in one or several chromosomes. The nuclear rDNA copies within a genome can be highly homogenous because of concerted evolution of intra and inter chromosomal loci. ITS1 and ITS2 are non coding regions located in rDNA between 18s and rRNA gene respectively (Insune et al., 2003, Jansen et al., 2006, Won and Renner, 2005). DNA sequence of the ITS1 and ITS2 of the rRNA transcription unit have proven useful in resolving phylogenetic relationships for closely related taxa due to their relatively rapid evolution rate (Baldwin, 1992, Mai and Coleman, 1997). Lopez-pinon et al., (2002) used ITSs for the identification of scallop species. Vidigal et al., (2000) used ITS2 sequence to resolve the phylogenetic relationships between among Brachian Biomphalaria species (Mollusca Planorbidae). Navajas et al., (1998) used ITS2 in combination with COI (Cytochrome oxidase sub unit 1) to investigate intra specific variation in Tetranthus urticae and found species-wide homologeneity of ITS2 sequence. Canapa et al., (2003) used 16SrRNA gene to study the phylogeny of Veneridae. The transcripts folding structure of the ITS1 and ITS2 provide some signals that guide the ribosomal coding regions where they are processed into small 5.8S and large ribosomal RNA (Vander Sande et al., 1992). The potential to predict the folding structure has enhanced the role of ITSs particularly ITS2 in phylogenetic studies, since it is important to guide reliable sequence (Michot et al., 1999). The secondary structure can be predicted by many methods
Sequence comparison of the ribosomal DNA internal transcribed spacers (ITSs) and 5.8S ribosomal gene and their secondary structures in 17 Pteriomorphian bivalves

like electron microscopy (Gonzales et al., 1990) chemical and structure probing (Yeh and Lee, 1992,) and computer softwares (e.g. Sfold and mfold) which utilize minimum free energy (Zuker and Steigler, 1981). A highly conserved sequence is situated around a central loop and at the apex of a long stem in the 3’half (Joseph et al., 1999). Due to higher rate of sequence variation of transcribed spacers this may exhibit dramatic size variation and extensive sequence divergence even among moderately distant species (Furlong and Made, 1903). The coding region which has been most widely used in metazoans is the ITS2.

The class bivalves are one of the most important members of most marine and fresh water ecosystems. It includes 6 subclasses. The subclass pteriomorpha includes entirely marine forms. Bivalves are one of the abundant and diverse groups of marine fauna. The phylogeny of mollusca, including bivalves is controversial subject (Sigwat et al., 2007). Attempts are made to substantiate the phylogenetic relationship, between limited members of the species of bivalves (Cheng Han-Liang et al., 2006, Coleman Vacquier, 2002, Ding et al., 2004, He etal, 2005, Insua et al., 2003, Lopez-pinon, 2002, Vidigal et al., 2000).

The overall literature sited shows that ITSs and 5.8S rRNA and secondary structures are widely used to resolve the phylogeny of limited species of bivalves and other organisms in the genus and species level. Utility of ITSs and 5.8S rRNA gene and their secondary structure in the study of relationship between members of higher taxa like subclass level, pteriomorphian subclass, is new. Thus the present investigation was focused on ribosomal DNA internal transcribed spacers and 5.8S rRNA gene sequences and their secondary structures. The sequences of ITS1, ITS2 and 5.8S rRNA and their secondary structures of selected geographically variant pteriomorphian bivalves were comprehensively investigated. Such case studies are relevant in broader phylogenetic contexts and for analyzing the function in ribosomal biogenesis. Since the secondary structure of ITSs region are more conserved than the nucleotide sequences and their sequence analysis helps in understanding molecular evolution and increases the number of models developed in this study can be used for further phylogenetic analysis.

MATERIALS AND METHODS

DATA SET

ITS1, ITS2 and 5.8S sequences of the 17 pteriomorphian bivalve species belongs to diverse geographical locations that are deposited in Genbank were investigated. The accession numbers, species scientific name, common name and geographical locations are listed in Table 1

SEQUENCE ANALYSIS AND PHYLOGENETIC TREE CONSTRUCTION

Multiple sequence alignment was performed by using MegaAlign of DNA star package using Clustal W method. The trees were produced by Neighbor-joining (NJ) and maximum parsimony (MP) methods using MEGA (Molecular Evolutionary Genetics Analysis, Version.4) (Tamura et al., 2007). The evolutionary history was inferred using the Neighbor-Joining method (Saitou N & Nei M 1987). The optimal tree with the sum of branch length = 15.74528780 is shown. The tree is drawn to scale, with branch lengths in the same units as

Figure 1

Table 1: Species name, geographical distribution and accession numbers of 17 Pteriomorphian bivalve species for genetic study
analyses were conducted in MEGA4 (Tamura et al., 2007). The evolutional history was also inferred using the Maximum Parsimony method (Eck & Dayhoff 1966). The MP tree was obtained using the Close-Neighbor-Interchange algorithm (Eck & Dayhoff 1966) with search level 2 (Eck & Dayhoff 1966, Nei & Kumar 2000) in which the initial trees were obtained with the random addition of sequences (10 replicates). Phylogenetic analyses were conducted in MEGA4 (Tamura et al., 2007).

ESTIMATES OF EVOLUTIONARY DIVERGENCE BETWEEN SEQUENCES.

The number of base substitutions per site from analysis between sequences is shown. All results are based on the pairwise analysis of 17 sequences. Analyses were conducted using the Maximum Composite Likelihood method in MEGA4 (Tamura et al., 2004, Tamura et al., 2007). The motif was identified by using SeSiMCMC (Sequence Similarities by Markov Chain Monte-Carlo) algorithm, which finds DNA motifs of unknown length and complicated structure such as direct repeats or palindromes with variable spacers in the middle in a set of unaligned DNA sequences. It uses an improved motif length estimator and careful Bayesian analysis of the possibility of a site absence in a sequence. The extracted motif was aligned and common motif was constructed by STAMP server (http://www.benoslab.pitt.edu/stamp)

SECONDARY STRUCTURE PREDICTION

The RNA secondary structure for ITSs was predicted by using RNADRAW online software (Christoffersen et al., 1994). RNADRAW predict RNA structure by identifying suboptimal structure using the free energy optimization methodology at a default temperature 37º C. In the current study ITS1, ITS2 and 5.8S rRNA sequences were used separately for RNA structure prediction. The algorithm used in RNADRAW was ported from RNAFOLD program included in the Vienna RNA package. (Hofacker et al., 1994). The dynamic programming algorithm used in RNADRAW was based on the work of Zuker and Stiegler (1981) and uses energy parameters taken from Frier et al., (1986) and Jaeger et al., (1989).

RNA FOLD

The Srieb program in Sfold (Statistical and Rational Design of Nucleic Acids) was used to predict the probable target accessibility sites (loop) for trans-cleaving ribozymes ITS2 (Ding et al., 2004). The prediction of accessibility is based on a statistical sample of the Boltzman ensemble for secondary structures. Here, we assessed the likelihood of unpaired sites for potential ribozyme target. Each mRNA exists as population of different structures. Hence, stochastic approach to the evolution of accessible sites was found appropriate (Christoffersen et al., 1994). The probability profiling approach by Ding and Lawrence (2001) reveals target sites that are commonly accessible for a large number of statistically representative structures in the target RNA. This novel approach bypasses the longstanding difficulty in accessibility evaluation due to limited representation of probable structures and high statistical confidence in predictions. The probability profile for individual bases (W=1) is produced for the region that includes a triple and two flanking sequences of 15 bases each in every site of the selected cleavage triplet (e.g. GUC).

RESULTS

The ribosomal DNA ITSs, 5.8S rRNA gene and rDNA ITS2+5.8S rRNA gene regions are analyzed for its sequence length, GC contents, AT contents, Motifs and secondary structural parameters.

SEQUENCE ANALYSIS OF FIRST INTERNAL TRANSCRIBED SPACER

The length of ITS1 sequence ranged from 21bp to 475bp and GC contents from 25.84% to 50.00 % in 17 species (Table.2). Interspecific ITS1 sequences showed remarkable divergence and obvious length polymorphism (Fig.1). The GC contents in ITS1 were higher than AT contents in all species and it is more than double in Mimchlamy sp.NT-2006. Craiostrea ariakensis had the longest ITS1 sequence (475bp) with 49.74 % GC contents, and two dinucleoptide micro satellites (ACAC) and (GGCG) were observed. As well as a repeat sequence of (AT) TTAAAAAAAAAA (Perna picta) and (AA) AAAAAAAAAAA (Perna perna) was found occurring twice in ITS1 regions in these species. Decatopectin radula had the shortest ITS1 sequence (212bp) with 25.84% GC contents. Mimchlamy senatoria, Ostrea circumplcta and Spondylus variance had intermediate ITS1 sequence (246bp, 402bp, 385bp respectively). ITS1 sequence 17 species were aligned. The minimum and maximum pair wise genetic distance in percentage within and among analyzed. The average observed genetic differences among all sequence was 7.156, with greatest differences being detected between the analyzed species is 10.021%-11.085 %.
Sequence comparison of the ribosomal DNA internal transcribed spacers (ITSs) and 5.8S ribosomal gene and their secondary structures in 17 Pteriomorphian bivalves

Figure 2
Fig 1: Alignment between Pteriomorphian Bivalve species

ITS2 SEQUENCES
Among Mimichlamy species the genetic distance ranged from 1.751%-9.009 %. Among Perna species the genetic distance ranged from 0.055%-10.494 %. The other species shows genetic distance variably lower than these genera.

Figure 3
Table 2: Ribosomal DNA ITS1, ITS2 and 5.8S ribosomal DNA gene sequence data in 17 Pteriomorphian bivalve species

We find relatively conserved motifs in ITS1 region across the 17 species of pteriomorphian bivalve species, the consensus sequence (39bp) of these motifs is depicted in Fig 3. The motifs show 11 polymorphic sites, 7 of which were transition and 4 of which were transversion.

Figure 4
Fig 2: Phylogenetic tree Common to ITS1, ITS2 and ITS2+5.8S of 17 pteriomorphian bivalve species

Phylogenetic analysis was conducted with neighbor-joining and maximum parsimony method. All obtained tree showed the same topology and differed only in their supportive values for certain branch (Fig 2).

Figure 5
Fig 3: Consensus sequence (1-39 bp) of motif in ITS1

Figure 6
Fig 4a: Motif 1 in ITS2

Figure 7
Fig 4b: Motif 2 in ITS2
SEQUENCE ANALYSIS OF THE SECOND INTERNAL TRANSCRIBED SPACER

The length of ITS2 sequence ranged from 233bp to 504bp and GC contents from 28.91% to 51.69% (Table 2). The length variation for ITS2 was less variable in all selected species. The contents of GC in ITS2 sequence were higher than AT in all the species studied. Sacostrea kegaki had the longest ITS2 sequence (504bp) with 48.10% GC contents, Michlamy senatoria, Ostrea circumpicta, Huotissa hyotis and Spondylus varius had intermediate length of ITS2 sequence (Table 2). Three common motifs, having sequence: (26)GCAUUGCGCCUUGGACCGUACAACUC,(32)GGACUCCGUCCGUUCUAAUAACAGACCGAUUGCand(50)AAACAGACGCACGGUGGUAUAUGAAUUGCCGU CUAUGUCUCUGUGA respectively were conserved in all classes (Fig 4a-c). ITS2 region form 17 pteriomorphian bivalve species were aligned (Fig 1). The minimum and maximum pair wise genetic distance in percentage within and among the analyzed species shown in Table 4. The average observed genetic distance among all sequences was 3.016, with greatest difference being detected between the analyzed species is 4.014%-4.744%. Among the MImichlamy species the genetic distance ranged from 2.786%- 4.20%. Among the Perna species the genetic distance ranged from 0.377% to 3.646%. The other species shows genetic distance variably lower than these species.

The length of 5.8S rRNA gene was 127bp to 167bp (Table 2). The GC content ranged from 32.73% to 46.67% depending on species. The sequence divergence ranged from 0.000% to 19.319% across 17 species analyzed (Table 4). The 5.8S rRNA gene contain 11 polymorphic sites in these species of which, 7 were transition and 4 were transversion.

PHYLOGENETIC ANALYSIS WITH ITS2 CONTAIN 5.8S rRNA GENE SEQUENCE

Using ITS2 spanning 5.8S rRNA gene, the phylogenetic tree was constructed with Neighbor-Joining method and maximum parsimony method (Fig 2). The topology of both trees was similar. Tree topology of both trees was similar. Tree topology showed that the 17 pteriomorphian bivalve belongs to 4 clads.

In contrast to ITS2, the construction of phylogeny using ITS1 was not applicable because of some disadvantages, such as ITS1 length variation, the presence of tandem repeated sequence, and large number of indels. ITS1 sequence analysis using neighbor-joining and maximum-parsimony method gives poor resolution. Therefore ITS1 was not used for phylogenetic analysis.
AND ITS2 + 5.8S RRNA GENE

Secondary structural features of ITS1, ITS2 and ITS2 + 5.8S were given in the Table 3. The secondary structure of the 17 pteriomorphian species were classified into three classes based on the conserved stem and loops. Class one includes Amusium pleuronectes, Mimichlamys nobilis, Mimichlamys senatoria, Mimichlamys sp.NT-2006, class two includes Perna pecta, Perna canaliculus, Perna perna, Aulacomya atra maoriana and class three includes Crassostrea ariakensis, Dendostrea crenulifera, Ostrea circumpicta, Hyotissa hyotis and Spondylus varius.

![Figure 10](image1)

Figure 10

![Figure 11](image2)

Figure 11

Secondary structure of the remaining species is highly variant. Three common motifs, having sequence

(26)GCAUUGCUCUUUGGACCGUACAACUC,(32)GGACUGACCCGCUUAAAUACAGACCGAUGCCand(50)AUAACAGAGACGCAGCUGGUUAUGAAUUGCGCUAUGUCUGUGA respectively were conserved in all classes. Apart form the common conserved motifs shared among the species that are categorized into different classes, variable regions also do exist. The observed similarities at the secondary level structural level and further reflected at energy level.

![Figure 12](image3)

Figure 12

Fig 6: A secondary structural fold found in common to all 17 Pteriomorphian bivalve species. Identical bases are marked

DISCUSSION

The present study provided information about the nucleotide sequence of ITS1 and ITS2 regions, complete sequence of 5.8S ribosomal DNA gene in 17 species of pteriomorphian bivalves. Their characteristics and variation were demonstrated through sequencing. The length of ITS1 in the family veneridae were longer that in 4 Pectinade scallop with size 209bp to 276bp for ITS1 and 270bp to 294bp for ITS2 (Insua et al., 2003) respectively. The length ranges in ITS1 and ITS2 reported in the present study was one of the largest observed in bivalves. The size of ITS is species-dependent and difference could be significant among species. The largest ITS1, such as that in Ladybird beetle Exocomus quadripustulatus, was as long as 2752bp (Von des Schulenburg et al., 2001) and shortest one only 70bp to 80bp in Acropora species (Odorica and Miller, 1997). Depending on species the length of ITS2 can be twice or more than that
of ITS1 (Dahlgren et al., 2000), both situations, the length of ITS1 were similar to ITS2 or larger than ITS2 were also reported (Cheng et al., 2002, Coleman and Vacquier, 2002). The GC contents ranged from 12.09% to 36.44% for ITS1 and from 28.91% to 51.69% for ITS2 in Pteriomorphian bivalves in the current study. The GC contents in ITS2 is double than that of ITS1. The GC content in the present study is higher than the other bivalve species. For instance, the GC contents in pectinidae scallop were 43% to 49% for ITS1 or 44% to 49% for ITS2 (Insua et al., 2003) and 55.5% for ITS2 in Pear oyster (He et al., 2005). A frequent characteristic of species is a balance between GC contents between ITS1 and ITS2 and this is also occurs in the pteriomorphian bivalve species analyzed, this fact could indicate the coevolution between the two species at the level of base composition.

Extensive sequence variation and obvious length polymorphism were in both ITS1 and ITS2 regions in all the three family, pectinidae, mytilinidae and ostrioida similar to crustacea Eriocheir formoca (Chen et al., 2002) and other bivalves (Ding et al., 2004). In the present study, the construction of topology in pteriomorphian bivalve species using ITS1 sequence information is not applicable because of the higher length variation, the presence of number of repeated sequences and large number of indels in ITS1. But, it is better in phylogenetic analysis at lower taxonomic levels. In contrast, the interspecific ITS2 were shorter than that of ITS1, providing advantage and convenience in designing primer and sequencing, the study of genetic structure in all the three families of pteriomorphian bivalves. In the present study, three relatively conserved motifs in ITS2 were found in the all the species analyzed. This indicates that these motifs might be involved in certain nucleotide acid-related functions, such as in rRNA processing (Insua et al., 2003). This study also reports two dinucleotide microsatellite (AC) and (GC) in ITS1 and ITS2 regions. These microsatellites can be used as good markers in future studies. Presence of dinucleotide and trinucleotide microsatellites were reported in ITSs sequences in Lasigona (King et al. 1999, Chen, 2006). Inter individual divergences in ITS2 and ITS2 + 5.8S regions were detected in Perna picta, Ostrea circumpicta (0.01%) and Semipallium fulvicostat, Mimichlamys nobilis, Mimichlamys sp.NT-2006, Decatopecten radula, Minnivolva pyxidatus, Perna calaliculus, Perna perna and Aulacomny atra maorian (0.13% to 0.23% in Its2 + 5.8S rRNA gene) and 0.4% in ITS2 region of Perna picta. The sequence cannot be thought a different only if sequence divergence more than 0.9% (Kong et al., 2002). Thus the Ostrea circumpicta, Semipallium fulvicostat, Mimichlamys nobilis, Mimichlamys sp.NT-2006, Decatopecten radula, Minnivolva pyxidatus, Perna calaliculus, Perna perna showed no intra specific variations in ITS2 and ITS2 + 5.8S. The 5.8S ribosomal RNA gene alone is highly conserved across the 17 species studied and had been sequencing length 127bp in ostreidae, 157bp in mytilidae and 167 bp in pectinidae families. Thus it is good marker for species identification at family level, which was reported in other bivalves A.opercularis, M.varia, H.distortus, P.maximus (All pectinadae scallop species) (Insua et al., 2000).

The secondary structure prediction was performed in both ITS1, ITS2 and ITS2 contain partial 5.8S rRNA gene to find out the conserverdness of the sequence in various species of pteriomorphian bivalve species (Table 3). The stems (double stranded paired region) stabilize RNA secondary structure and the number of stems present in each ITS1, ITS2 and ITS2 + 5.8S rRNA is given in Table 2. ITS2 RNA structure form species in ostreidae and spondylidae family shows highest negative free energy ranging -162 kcal to -222.7 kcal. This is followed by mytilidae species and pectinidae species ranging from -128 kcal to -141.01 kcal, indicating the divergence of this family occurred at different periods, with greater stability in ostareae RNA might indicate it evolve before the evolution of pectinidae and mytilidae. Both ostreidae and pectinidae evolved in Ordovician period, but mytilidae evolved in Devonian period (Schneider, 2001). To substantiate the early evolution of ostreidae require further study in molecular level. The same trend in negative free energy was found in ITS1 also, but the variation is higher than ITS2. Visual comparison shows that this is related to the trend in the cladogram given in Fig 2 .This convergence at secondary structural level among species from different geographical isolates may be due to the evolutionary pressure on ITSs to maintain the RNA secondary structure involved in post transcriptional processing of rRNA (Shinohara et al., 1999). Secondary structures predictions for ITS1 and ITS2 regions show that their domain, base pair to form a core region central to several stem features inferring the conserverdness is more important for the proper rRNA folding pattern (Wesson et al., 1992). Fig 5 shows the distribution of loops among different isolates. The segments of ITS2 having score >50 are further probed carefully for target site to assess the likelihood of un-paired segments. Interestingly, the observed
Maximum parsimony) generated forms with similar sequence through both methods (Neighbor-joining and more relationship to ostreidae. Phylogentic analysis of ITS2 species Mimichalmys senatoria, Spondylus varius shows canaliculus, Aulacomya atra maoriana) but pectinidae has close relationship with mytiloidae (Perna picta, Perna pectinidae (Amusium pleuronectes, Mimichlamys nobilis) obtained by sequence analysis revealed that member of (Pectinidae, mytilidae and ostreidae). In this study the tree phylogentic relationship among species from three families 2000). Our study demonstrated that ITS1 provide week length polymorphism analysis (He et al., 2005, Vidgal et al., such as polymer chain reaction and restriction fragment morphological systematics and other molecular techniques, well in agreement with those from analysis based on phylogenetic tree was successfully applied in analyzing of phylogenetic relationship among the Biomphaaria species and among pearl oysters and the conclusions from phylogenetic tree was in agreement with those from analysis based on morphological systematics and other molecular techniques, such as polymer chain reaction and restriction fragment length polymorphism analysis (He et al., 2005, Vidgal et al., 2000). Our study demonstrated that ITS1 provide week phylogentic relationship among species from three families (Pectinidae, mytilidae and ostreidae). In this study the tree obtained by sequence analysis revealed that member of pectinidae (Amusium pleuronectes, Mimichlamys nobilis) has close relationship with mytiloidae (Perna picta, Perna canaliculus, Aulacomya atra maoriana) but pectinidae species Mimichalmys senatoria, Spondylus varius shows more relationship to ostreidae. Phylogentic analysis of ITS2 sequence through both methods (Neighbor-joining and Maximum parsimony) generated forms with similar topology that were very concurrent with the morphological taxonomy proposed previously. Therefore ITS2 sequence characteristics and secondary structure prediction are efficient tools in reconstruction of evolutionary relationship among these organisms, and they can be applied in establishment of species relationship, or reevaluation of the traditional taxonomy.

CONCLUSIONS

The present study shows two contrasting aspects of ITSs regions i.e. the general trend of variability among the species as well as the conservedness between few species. Surprisingly the species displaying the conservedness belong to different geographical locations with diverse ecological conditions. Our study implies that ITSs region though have less selective pressure than the ribosomal regions but still evolve slower than the intergenic spacers, indicating that some selective pressure do exist on them, probably from the constraint to maintain the RNA secondary structure required for post-transcriptional processing and are more species specific than geographically influenced. Several common structural folds were shared among the selected pteriomorphian bivalve species for maintaining functional equivalents. Identifying the homologous regions and reconstructing their evolution increase the traits available for the phylogenetnic analysis. The present studies indicate that the class pectinidae shows more closeness to mytiloidae than to ostreidae with monophyletic relationship among the species. It also shows that ITS2 is more powerful, than ITS1, tool with their secondary structural analysis to reconstruct phylogenetnic tree.

ACKNOWLEDGEMENTS

The authors thanks to PVP College principal Dr.AShok.V.Babar to provide us broadband internet facility and various soft wares required for the present study and valuable advise to complete the work in time.

References


r-3. Cheng H L, Meng X P, Hong J J, Dong Z G & Chen S Y 2006. Sequence analysis of the ribosomal DNA internal
transcribed spacers and 5.8S ribosomal RNA gene in representatives of the clam family veneridae (Mollusca: Bivalvia). Journal of Shellfish Research. 1. 2001-2010


r-14. Furlong JC & Maden BE. 1983. Pattern of major divergence between the internal transcribed spacers of ribosomal DNA in Xenopus borealis and Xenopus levis, and of minimal divergence within ribosomal coding regions. EMBO. J. 1: 443-448.


Phylogenetic relationships among Brazilian Biomphalaria species (Mollusca Planorbidae) based upon analysis of ribosomal ITS-2 sequence. Parasitol. 121:611-620.


Sequence comparison of the ribosomal DNA internal transcribed spacers (ITSs) and 5.8S ribosomal gene and their secondary structures in 17 Pteryomorphian bivalves

Author Information

K.V. Ashokan, Ph.D.
Department of biological science P.V.P.Colleg. Kavathe Mahankal Sangli, Maharashtra, India

M.M. Pillai, MSc, PhD
Departments of Biotechnology, KIITs Engineering College, Gokul Shirgaon, Kolhapur, Maharashtra, India

S.M. Angadi, MPhil
Departments of Zoology, Kasturbha Walchand College, Sangli, Maharashtra, India

D.S. Mundaganur, PhD
Departments of Zoology, Willingdon College, Sangli, Maharashtra, India