Indian J. Fish., 62(2): 98-103, 2015

Note

DNA barcoding of selected species of pufferfishes (Order: Tetraodontiformes) of Puducherry coastal waters along south-east coast of India

K. KALESHKUMAR, R. RAJARAM, S. VINOTHKUMAR, V. RAMALINGAM AND KHANGEMBAM BRAJAMANI MEETEI*

School of Marine Sciences, Bharathidasan University, Tiruchirappalli - 620 024, Tamil Nadu, India
**Krishi Vigyan Kendra, Bishnupur District, Utlou P. O., Nambol, Manipur - 795 134, India*
e-mail: drrajaram69@rediffmail.com

ABSTRACT

Samples of three species of pufferfishes were collected from Puducherry in the south-east coast of India. Collected fishes were identified as *Chilomycterus reticulatus*, *Arothron hispidus* and *Lagocephalus guentheri* by mitochondrial CO1 gene sequencing. Neighbor-Joining method was used for phylogenetic analysis which confirmed that the analysed species had a dichotomous relationship with its ancestor species. Genetic variation was studied by sequence similarity using the homologous sequences retrieved from BLAST server. The maximum length of open reading frames (ORF) for the three pufferfish species *viz*., *C. reticulatus*, *A. hispidus* and *L. guentheri* were 171, 345 and 465 respectively. Results from a simple gene prediction algorithm revealed that *C. reticulatus*, *A. hispidus* and *L. guentheri* have 7 start codons each. The pairwise genetic distance calculated among the three species ranged from 0.002 to 0.176% for *C. reticulatus,* 0.002 to 0.169% for *A. hispidus* and 0.002 to 0.208% for *L. guentheri*. Genetic distance analysis of mitochondrial DNA sequences is demonstrated as unique to their related species of all pufferfishes. Genetic distance is also used in the comparison of genetic similarity between different species and also in different subspecies of pufferfishes belonging to the order Tetraodontiformes.

Keywords: Genetic distance analysis, Mitochandrial CO1 sequencing, Phylogenetic analysis, Pufferfishes, Tetraodontiformes

In silico studies can identify the genetic diversity and enable fast prediction of gene expression based on sequence comparison within species. Genetic materials are the major source to reveal intra-specific genetic diversity and phenotypic plasticity from intra-specific differences for all kinds of organisms (Fabrice *et al*., 2010). DNA - based identification techniques have been proved to be powerful analytical tools (Ko *et al*., 2013. Morphological information such as body shape, pigmentation and meristic count are inadequate for genus or species level study in fish (Zhang *et al*., 2004; Comi *et al*., 2005; Teletchea, 2009). Hence, molecular level information becomes important for ecological monitoring, environmental impact assessment, formulating fisheries compensation plans, resource management, smuggling prevention and establishing marine protected areas (Moura *et al*., 2008; Valdez-Moreno *et al*., 2010). Presently, DNA barcoding is the standardised and universal method to identify species and in uncovering biological diversity (Hebert *et al*., 2003; 2004).

Mitochondrial DNA (mt DNA) has proven to be a useful molecular marker for evolutionary studies in genetic diversity, because of its predominantly maternal inheritance, relatively rapid base substitution rate, lack of recombination and easy isolation (Avise *et al*., 1987). In particular, the mitochondrial COI gene, can be amplified from most living beings using universal primers and has been proposed as a reference marker for cross - amplification over a wide range of species, genera and families (Hebert *et al*., 2003). The evolution rate of mt DNA and its amino acid sequence is highly conserved to COI gene across the phyla that have been used in various molecular tools in several genetic studies. Finally, it has been proposed as a universal barcode database for all animal species (http://www.barcodinglife.com) and has already been used as an identifier for numerous insect taxa (Hogg and Hebert, 2004; Hufbauer *et al*., 2004). Ajmal Khan *et al.* (2011) barcoded 42 fishes belonging to 32 genera and 23 families while Lakra *et al.* (2011) barcoded 115 marine finfish species belonging to 37 families and 79 genera, from Indian waters.

Pufferfishes are appropriate group to study the process of genome size evolution and in particular, smooth pufferfishes (family Tetraodontidae) are the smallest vertebrates with a haploid genome size of ~400 million bp (Neafsey and Palumbi, 2003). Nevertheless, the genetic relationship of scaffold is still unknown owing to the requirement of genetic map of the fish. The use

of pufferfish gene sequences for comparative study with other related species can provide the answer for genome structure and evolution (Kai *et al*., 2003). In this study, we have sequenced and reported the genetic diversity based on sequence information, open reading frames (functional region) and also identified the beginning (start codon) and termination region (polyadenylation signals) of functional mitochondrial COI gene for three pufferfish species *viz*., *Chilomycterus reticulatus*, *Arothron hispidus* and *Lagocephalus guentheri* belonging to the order Tetraodontiformes from Puducherry coastal region along the south-east coast of India.

Pufferfish samples were collected from Puducherry (11°93`N, 79°83`E) and the collected fishes were transported to the laboratory and preserved at -20°C until further analysis. Fishes were identified based on morphological features and tissue samples from the muscle portion of the fishes were dissected and preserved in 95% ethanol for DNA isolation. The mitochondrial DNA was extracted from the tissue following the method of Sambrook *et al*. (2001). DNA amplification was carried out using the specific primer; Forward - FISH F1 (5'- TCAACCAACCACAAAGA CATTGGCAC-3') and Reverse - FISH R1 (5'-TCGACT AATCATAAAGATATCGGCAC-3') (Ward *et al*., 2005). The PCR conditions used were: denaturation step of 95°C for 1 min, followed by 35 cycles of 30 s at 95°C, 30 s at 40°C, and 90 s at 72°C, followed by 10 min at 72°C. PCR products were then purified using Ultra Clean PCR clean up (Mo-Bio) spin columns, and then submitted to Eurofins Genomics India, Bangalore, for sequencing using ABI3730XL (Applied Biosystems). Sequences were verified and manipulated with Sequencer ver. 4.8 (Gene Codes Corporation). Mitochondrial COI gene from the three fish species *viz*., *C. reticulatus*, *A. hispidus* and *L. guentheri* were sequenced and submitted to Barcode database (with accession nos. JX495771, KF422720 and KF442241 respectively) and the length of the COI regions were 631, 636 and 615 bp respectively.

DNA sequences were initially aligned using CLUSTAL X (Thompson *et al*., 1994) and default gap, extension penalties were used followed by manual editing using Seq App 1.9 (Gilbert, 1994). Alignments were adjusted manually following guidance of Odorico and Miller (1997) and Chen *et al.* (2004). A BLAST search was performed in GenBank database for each sequence and the matching homologous sequences were retained for subsequent alignment. Neighbor-Joining method was implemented in MEGA v 5.0 (Kumar *et al*., 2001). BLASTn and BLASTx searches of the NCBI non-redundant protein database were performed for all processed sequences and hits with Expect values (E values) of less than 1×10^{-4} were considered to be homologs to the query sequence. For the differentially expressed sequences, BLASTn searches of the non-human, non-mouse EST database were performed for all sequences while homologs with significant E-values from the BLASTn search were then used as a query in a BLASTx search and the open reading frames and start codons were predicted (Ramalingam *et al*., 2012). Genetic distance calculations implemented in MEGA version were used to estimate sequence divergence (Kumar *et al*., 2001).

In the present study, the nucleotide sequences of three pufferfish species belonging to the families Diodontidae and Tetraodontidae (Order: Tetraodontiformes) *viz*., *Chilomycterus reticulatus*, *Arothron hispidus* and *Lagocephalus guentheri* were initially compared with nucleotide database using BLASTn server which indicated that these organisms acquired maximum identity as compared with other genus. *C. reticulatus* was recorded for the first time in Indian waters and also the sequence information happens to be the first submission in barcode database. Details of barcode submissions of *C. reticulatus*, *A. hispidus* and *L. guentheri* in barcode database are presented in Table 1.

Table 1. List of barcoded species available in BOLD

Name	Family	Accession no.	No. of submissions available*
Chilomycterus reticulates	Diodontidae	JX495771	1
Arothron hispidus	Tetraodontidae	KF422720	30
Lagocephalus guentheri	Tetraodontidae	KF442241	13

*Number of submission in barcode database

The base of tree forms a dichotomy leading to two clusters *i.e.,* clade I and clade II (Fig. 1). Clade I consists of two sub clades that carries both *A. hispidus* and *L. guentheri* (Family: Tetraodontidae) while clade II consists of *C. reticulatus* (Family: Diodontidae). The branch length leading to *A. hispidus* clade was similar to the branch length leading to the related species of *A. hispidus*. Similarly, *A. hispidus* and *L. guentheri* clade was highly similar with their respective related species which suggests that *A. hispidus* and *L. guentheri* have unique heterogeneity of mt DNA. The different strains of the species *L. guentheri* are closely related (Santini *et al.,* 2013). Holgroft (2005) reported that the mitochondrial sequences of Balistidae and Monacanthidae are sister groups related to Tetraodontidae and Diodontidae that robustly maintain the monophyletic relation.

A blast search revealed that the segment of control gene of mt COI, has 99% sequence similarity in *C. reticulatus*, *A. hispidus* and *L. guentheri*. The nucleotide and protein sequences of *C. reticulatus*, *A. hispidus* and

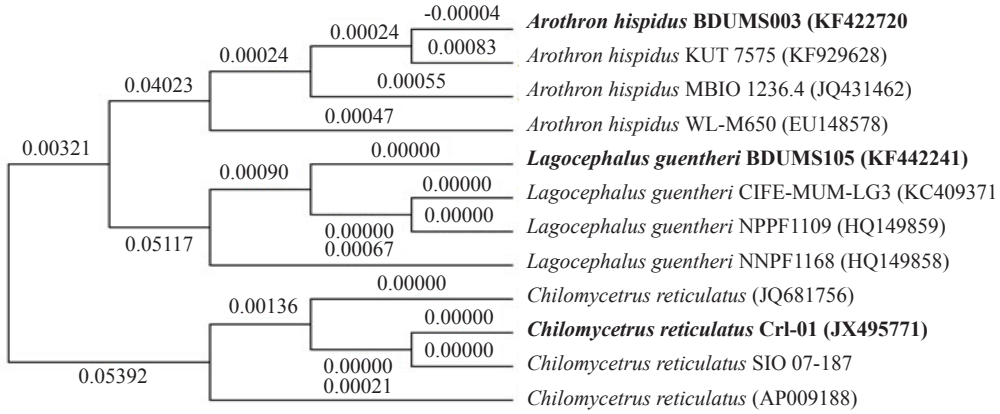

Fig. 1. Phylogenetic analysis of *C. reticulatus* and *L. guentheri* by Neighbor joining method using MEGA v 5.0

L. guentheri were primarily compared and analysed with relevant databases using BLASTn and BLASTx as shown in Table 2, 3 and 4. Maximum identity was observed in different strains of *C. reticulatus* nucleotide sequences (Table 2). Maximum identity of *A. hispidus* was with the nucleotide sequence of *A. hispidus* WL-M650 and *A. hispidus* WL-M651 (Table 3). Maximum identity of *L. guentheri* was observed with the nucleotide sequence of *L. guentheri* CIFE-MUM-LG3 and *L. guentheri* Smith 268.19 (Table 4).

The mt COI sequence of *C. reticulatus* and *A. hispidus* contains four open reading frames (ORFs) and *L. guentheri* has three ORFs. In *C. reticulatus*, the long ORF was obtained between 466-636 bp and 165-335 bp length while in *A. hispidus* the long ORF was between 3 and 347 bp and *L. guentheri* has long ORF between 79 and 543 bp length. The long ORF present in the pufferfishes can encode a functional protein in the cytochrome oxidase I gene. Yamaguchi and Brenner (1997) reported that the F1Aα has an intronless 1.3 kb

Table 2. Annotated maximum sequence identity of nucleotide and protein for *Chilomycterus reticulatus* with related similar species using Blast n and Blast x server

Species	Blast N			Blast X		
	Accession no.	Max. score	Max identity	Accession no.	Max score	Max identity
C. reticulatus	JQ681756.1	1125	99%	AGZ84817.1	289	99%
C. reticulatus	AP009188.1	1120	99%	ACG50200.1	290	99%
C. reticulatus SIO 07-187	HQ010089.1	1016	99%	AEK04469.1	290	99%
C. schoepfii	JQ681758.1	704	87%	AFJ92749.1	290	99%
C. geometricus	JQ681755.1	693	87%	AFJ92700.1	290	99%

Table 3. Annotated maximum sequence identity of nucleotide and protein for *Arothron hispidus* with related similar species using Blast n and Blast x server

Species	Blast N			Blast X		
	Accession no.	Max score	Max identity	Accession no.	Max score	Max identity
A. hispidus	FJ434547.1	1171	100%	ACG50200.1	414	99%
A. hispidus	JQ681763.1	1157	100%	ACG50207.1	413	99%
A. hispidus WL-M651	EU148579.1	1133	100%	AEV56399.1	414	100%
A. hispidus NBE 0677	JQ349778.1	1129	100%	ABO48435.1	413	100%
A. hispidus WL-M650	EU148578.1	1127	99%	ACG50203.1	413	100%

Table 4. Annotated maximum sequence identity of nucleotide and protein for *Lagocephalus guentheri* with related similar species using Blast n and Blast x server

Species	Blast N			Blast X		
	Accession no.	Max score	Max identity	Accession no.	Max score	Max identity
L. guentheri BDUDMS 105	KF442241.1	1136	100%	AGZ92978.1	400	100%
L. guentheri CIFE-MUM-LG3	KC409371.1	1136	100%	AGC25469.1	400	100%
L. guentheri NPPF1109	HQ149859.1	1131	99%	ACM88301.1	399	99%
L. guentheri NPPF 1168	GQ149858.1	1120	99%	AEK02887.1	398	99%
L. guentheri Smith 268.19#3	JF493722.1	1114	99%	ACR44799.1	398	99%

open reading frame encoding a 423 aa protein. F1Aβ has an intronless 1.4 kb ORF encoding a 416 aa protein.

C. reticulatus contains two poly A signals at the nucleotide positions of 421 and 447 while *A. hispidus* has three poly A signals at positions of 201, 425 and 561 whereas *L. guentheri* contains three poly A signal at 211, 433 and 565. The accurate prediction of polyadenylation sites play an important role in understanding transcriptome diversity and regulatory mechansims of gene expression by defining the termination of genes (Wu *et al*., 2012). The pairwise genetic distance computed among the three species ranged from 0.002 to 0.176% for *C. reticulatus,* 0.002 to 0.169% for *A. hispidus* and 0.002 to 0.208% for *L. guentheri* (Table 5, 6 and 7). The lowest genetic distance indicates closely related and highest genetic distance indicates highly diverged status.

Genetic distance analysis of pufferfishes showed close identity with its related species and subspecies. Fish genomes are closely related with human genomes that have potential to identify genes and regulatory regions in humans. Further, pufferfishes especially of the family Tetraodontidae is an important group for studying the phenotypic variation.

Acknowledgements

Authors thank the authorities of Bharathidasan University for infrastructure facilities and acknowledge the Science and Engineering Research Council, Department of Science and Technology, Ministry of Science and Technology, Government of India for providing funds in carrying out this work (Ref. No. SR/FT/LS-050/2009 dated 23 July, 2010).

Table 5. Genetic distance of *C. reticulatus* sequence from related species

Species	1	2	3	4	5	6	7	8
Diodon holocanthus NSMK:P1-000250								
Diodon holocanthus MFC233	0.002							
Diodon holocanthus BIOUG:HLC-15165	0.004	0.005						
Diodon hystrix MBI0319	0.107	0.109	0.104					
Cyclichthys orbicularis ADC09_269.2	0.138	0.141	0.136	0.152				
Chilomycterus reticulatus SIO 07-187	0.112	0.115	0.110	0.141	0.150			
Chilomycterus reticulatus BDUMS001	0.115	0.117	0.112	0.143	0.152	0.002		
Chilomycterus spinosus LBP-35009	0.124	0.126	0.126	0.176	0.158	0.117	0.120	

Table 6. Genetic distance of *A. hispidus* sequence from related species

Species	1	2	3	4	5	6	7	8
Arothron immaculatus NBFGR:1102b								
Arothron nigropunctatus BLOUG:HLC-12417	0.075							
Arothron stellatus CIFE-MUM-AS2	0.077	0.079						
Arothron hispidus BDUMS003	0.072	0.083	0.052					
Arothron hispidus CIFE-MUM-AH1	0.081	0.089	0.053	0.012				
Arothron hispidus MBIO1236.4	0.074	0.085	0.053	0.002	0.013			
Canthigaster solandri MBIO949.4	0.169	0.165	0.154	0.144	0.154	0.142		
Canthigaster janthinoptera ECOMAR:NBE1142	0.167	0.165	0.154	0.144	0.154	0.142	0.032	

Table 7. Genetic distance of *L. guentheri* sequence from related species

Species	1	2	3	4	5	6	7	8	9
Sphoeroides karkudineus LBP-41589									
Sphoeroides naphelus SMSA7320	0.133								
Lagocephalus wheeleri SIFU1	0.174	0.151							
Lagocephalus spadiceus FSCS302-00	0.177	0.147	0.012						
Lagocephalus guentheri BDUMS006	0.183	0.170	0.077	0.075					
Lagocephalus guentheri NPPF1168	0.181	0.168	0.075	0.073	0.002				
Dactyloptera orientalis WL-M573	0.174	0.166	0.071	0.069	0.008	0.007			
Lagocephalus larigatus HRCP:46919	0.183	0.165	0.108	0.106	0.103	0.102	0.096		
Myripristis adusta MBIO1182-4	0.185	0.208	0.184	0.182	0.184	0.182	0.177	0.214	

The lowest genetic distance is highlighted and the farthest genetic distance is marked in bold

Reference

Ajmal Khan, S., Prasanna Kumar, C., Lyla, P. S. and Murugan, S. 2011. Identifying marine fin fishes using DNA barcodes. *Curr. Sci.,* 101(9): 1152-1154.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T. and Neigel, J. E. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.,* 18: 489-522.

Chen, C. A., Chang, C. C., Wei, N. V., Chen, C. H., Lein, Y. T., Lin, H. E., Dai, C. F and Wallace, C. C. 2004. Secondary structure and phylogenetic utility of the ribosomal internal transcribed spacer 2 (ITS2) in scleractinian corals, *Zool. Stud.,* 43: 759–771.

Comi, G., Iacumin, L., Rantsioua, K., Cantoni, C. and Cocolin, L. 2005. Molecular methods for the differentiation of species used in production of cod-fish can detect commercial frauds. *Food Control.,* 16: 37–42.

Fabrice, G., Irene, T. B. and Sandra, L. 2010. Plant resource use strategies: the importance of phenotypic plasticity in response to a productivity gradient for two subalpine species. *Ann. Botany,* 106: 637–645.

Gilbert, D. C. 1994. Seq App 1.9. *A biological sequence editor and analysis program for Macintosh computers*. Available *via* anonymous ftp at ftp.bio.indiana.edu.

Hebert, P. D. N., Cywinska, A., Ball, S. L. and de Waard, J. R. 2003. Biological identifications through DNA barcodes. *Proc. R. Soc. Biol. Sci. B.,* 270: 313–321.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H. and Hallwachs, W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator. Proc. Natl. Acad. Sci. USA,* 101: 14812–14817.

Hebert, P. D. N., Ratnasingham, S. and deWaard, J. R. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Biol. Sci.,* 270: S96-S99.

Hogg, I. D. and Hebert, P. D. N. 2004. Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. *Can. J. Zool.,* 82: 749 – 754.

Holcroft, N. I. 2005. A molecular analysis of the interrelationships of tetraodontiform fishes (Acanthomorpha: Tetraodontiformes). *Mol. Phylogenet. Evol*., 34: 525–544.

Hufbauer, A., Bogdanowicz, S. M. and Harrison, R. G. 2004. The population genetics of a biological introduction: mitochondrial DNA and microsatellite variation in native and introduced populations of *Aphidius ervi*, a parasitoid wasp. *Mol. Ecol.,* 13: 337– 348.

Kai, W., Kikuchi, K., Fujita, M., Suetake, H., Fujiwara, A., Yoshiura, Y., Ototake, M., Venkatesh, B., Miyaki, V. and Suzuki, Y. 2003. A genetic linkage map for the tiger pufferfish, *Takifugu rubripes*. *Genetics,* 171: 227–238.

Ko, H. L., Wang, Y. T., Chiu, T. S., Lee, M. A., Leu, M. Y., Chang, K. Z., Chen, W. Y. and Shao, K. T. 2013. Evaluating the accuracy of morphological identification of larval fishes by applying DNA barcoding. *PLoS One,* 8(1): e53451.

Kumar S., Tamura K., Jakobsen I. B. and Nei, M. 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics,* 17: 1244 -1245.

Lakra, W. S., Verma, M. S., Goswami, M., Lal, K. K., Mohindra, V., Punia, P., Gopalakrishnan, A., Ward, R. D., and Hebert, P. 2011. DNA barcoding of Indian marine fishes. *Mol. Ecol. Resour.,* 11: 60-71.

Moura, T., Silva, M. C., Figueiredo, I., Neves, A., Munoz, P. D. Coelho, M. M. and Gordo, L. S. 2008. Molecular barcoding of north-east Atlantic deep-water sharks: species identification and application to fisheries management and conservation. *Mar. Freshw. Res.,* 59: 214–223. doi: 10.1071/mf07192

Neafsey, D. E. and Palumbi, S. R. 2003. Genome size evolution in pufferfish: A comparative analysis of diodontid and tetraodontid pufferfish genomes. *Genome Res.*, 13: 821-830.

Odorico, D. and Miller, D. 1997. Variation in the ribosomal internal transcribed spacers and 5.8S rDNA among five species of *Acropora* (Cnidaria; Scleractinia): patterns of variation consistent with reticulate evolution. *Mol. Biol. Evol.,* 14: 465-473.

Ramalingam, V., Rajaram R. and Suresh, V. 2012. Secondary structure prediction of Scleractinia Corals: A proteomic approach. *Ind. J .Geo-Mar. Sci.,* 42(4): 503-509.

Sambrook, J., Maccallum, P. and Russel, D. 2001. *Molecular cloning: A laboratory manual*, 3rd edn., Cold Springs Harbour Press*.,* NY, ISBN 0-87969-577-3: 2344.

Santini, F., Sorenson, L. and Alfaro, M. E. 2013. A new phylogeny of tetraodontiform fishes (Tetraodontiformes, Acanthomorpha) based on 22 loci. *Mol. Hylogenet. Evol.,* 69: 177–187.

Teletchea, F. 2009. Molecular identification methods of fish species: reassessment and possible applications. *Rev. Fish. Biol. Fish.,* 19: 265–293.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL X: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.,* 22: 4673-4680.

Valdez-Moreno, M., Vasquez-Yeomans, L., Elias-Gutierrez, M., Ivanova, N. V. and Hebert, P. D. N. 2010. Using DNA barcodes to connect adults and early life stages of marine fishes from the Yucatan Peninsula, Mexico: potential in fisheries management. *Mar. Freshw. Res.,* 61: 665–671.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R. and Hebert, P. D. N. 2005. DNA barcoding Australia's fish species. *Biol. Sci.,* 360: 1847-1857.

Wu, X., Ji, G. and Zeng, Y. 2012. *In silico* prediction of mRNA poly(A) sites in *Chlamydomonas reinhardtii*. *Mol. Genet. Genomics,* 287(11-12): 895-907.

Yamaguchi, F. and Brenner, S. 1997. Molecular cloning of 5-hydroxytryptamine (5-HT) type 1 receptor genes from the Japanese puffer fish, *Fugu rubripes*. *Gene*, 191: 219-233.

Zhang, J. B., Huang, L. M. and Huo, H. Q. 2004. Larval identification of *Lutjanus* Bloch in Nansha coral reefs by AFLP molecular method. *J. Exp. Mar. Biol. Ecol.*, 298: 3–20.

Date of Receipt : 21.07.2014
Date of Acceptance : 27.12.2014