RECENT ADVANCES IN THE STUDY OF
FISH EGGS AND LARVAE
M.P. Olivar and J.J. Govoni (eds.)

SCIENTIA MARINA 70S2
October 2006, 7-12, Barcelona (Spain)
ISSN: 0214-8358

MtDNA barcode identification of fish larvae in the southern Great Barrier Reef, Australia

GRAHAM G. PEGG, BILLY SINCLAIR, LEICA BRISKEY and WILLIAM J. ASPDEN.

Marine Molecular Genetics and Biotechnology Group, Central Queensland University, Rockhampton, Queensland, Australia. E-mail: g.pegg@cqu.edu.au

SUMMARY: Planktonic larvae were captured above a shallow coral reef study site on the Great Barrier Reef (GBR) around spring-summer new moon periods (October-February) using light trap or net capture devices. Larvae were identified to the genus or species level by comparison with a phylogenetic tree of tropical marine fish species using mtDNA *HVR1* sequence data. Further analysis showed that within-species *HVR1* sequence variation was typically 1-3%, whereas between-species variation for the same genus ranged up to 50%, supporting the suitability of *HVR1* for species identification. Given the current worldwide interest in DNA barcoding and species identification using an alternative mtDNA gene marker (*cox1*), we also explored the efficacy of different primer sets for amplification of *cox1* in reef fish, and its suitability for species identification. Of those tested, the Fish-F1 and -R1 primer set recently reported by Ward *et al.* (2005) gave the best results.

Keywords: reef fish, mtDNA, HVR1, cox1, DNA barcoding, species identification.

RESUMEN: IDENTIFICACIÓN DE LAS LARVAS DE PECES MEDIANTE MTDNA EN LA PARTE SUR DE LA GRAN BARRERA DE CORAL, AUSTRALIA. – Las larvas estudiadas fueron capturadas en el plancton de una zona coralina somera en la Gran Barrera de Coral en períodos de luna-nueva de la estación primavera-verano (octubre-febrero). Su captura se realizó mediante trampas de luz o redes de plancton. Las larvas fueron identificadas a nivel de género o especie por la comparación de un árbol filogenético de especies de peces tropicales marinas usando datos de la secuencia HVR1 del DNA mitocondrial. El análisis adicional demostró que, para una misma especie, la variación de la secuencia HVR1 era típicamente 1-3%, mientras que entre especies del mismo género la variación fue de hasta 50%, apoyando la conveniencia del uso del HVR1 para la identificación a nivel específico. Dado el interés mundial actual en el "código de barras genético" y en la identificación de especies usando otro marcador genético de DNA mitochondrial, el *cox1*, se exploró también la eficacia de diversos "primers" para la amplificación del *cox1* en peces de los arrecifes, y su conveniencia para la identificación específica. De los "primers" probados, el Fish-F1 y el -R1 *set* recientemente reportado por Ward *et al.* (2005) dieron los mejores resultados.

Palabras clave: peces de coral, mtDNA, HVR1, cox1, DNA identificación específica por código de barras genético.

INTRODUCTION

Australia's Great Barrier Reef (GBR) is recognised internationally for its importance as a tropical marine fish and coral reef biodiversity reserve, with up to 30 percent of the total reef area now protected as 'Green Zones'. In 2004, new fisheries rules in the State of Queensland were introduced to prohibit the taking of all regulated reef fish for nine-day closure periods around the new moon in the spring/early summer months (October-December) as a further measure to improve the chances of successful spawning of *Lutjanus*, *Lethrinus* , *Plectropomus* and other reef fish species important to the commercial and recreational fishery.

The current study was undertaken at an inshore shallow reef in the southern GBR (approximate latitude of 23 degrees south of the equator), where surface water temperature ranges from 18 degrees in winter to 28 degrees in summer. Larvae were cap-

tured around the new moon periods in spring 2004 by light traps and by plankton nets. Because identification of freshly hatched fish larvae to species level is often difficult due to poorly defined morphological characteristics (Leis and Carson-Ewart, 2004), we were interested in investigating DNA analysis methods that would enable unequivocal identification of larval fish or planktonic eggs. Two distinct approaches were to be explored, both of which involved mitochondrial (mtDNA) sequencing. Our group had been using mtDNA *HVR1*(*D-loop*) non-coding sequence to study within-species population diversity for adult reef fish (Aspden *et al.*, 2005) and in this study we sought to compare the suitability of *HVR1* sequence data compared with Cytochrome Oxidase subunit 1 (*cox1*) sequence data for species identification of larvae. The expectation was that lesser within-species variation would be expected for the functional *cox1* gene sequences (approximately 650 bp.), whereas more genetic diversity might be expected for the non-coding *HVR1* gene. The hypothesis to be tested in the study was whether *cox1* sequence data were sufficiently divergent between closely related species to enable unambiguous identification, or whether *HVR1* sequence data gave too much variation to be useful as a specific species identifier. We also wished to test a number of the *cox1* primer sets for fish reported in the literature to determine which primers gave the best results for reef fish. The study was timely, since Hebert *et al.* (2003) had initiated an international Barcode of Life project proposing that the *cox1* gene sequence could be used to differentiate between most animal species, including fishes. Subsequent to initiating this study, examples of the application of the barcoding concept to adult and larval fish identification have been reported by Ward *et al.* (2005) and Steinke *et al.* (2005).

MATERIALS AND METHODS

Larvae capture

Larvae were captured using a combination of light traps and plankton net devices. The light traps (900x300x300 mm) were constructed from translucent plastic boxes held together by cable ties. A waterproof pond light (50 W) was fixed into the base of the trap along with a 500 litre per hour electric bilge pump (Rule Corp, USA), thereby facilitating passage of larvae (and eggs) through a grid into the lower section of the trap. The light and pump were powered by a 12 V battery and the trap was suspended vertically 1 metre below the surface beside the research vessel for 2-hour periods. At other times, small zooplankton nets (50 cm diameter), sometimes incorporating a 10 W light, were let out into the 4 km/h tidal current to 0-1 m depth. A dual bongo plankton net of 70 cm diameter (Ocean Instruments, USA) was towed obliquely at a constant speed of 2 knots from a depth of 15 metres to the surface over 10 minutes. Captured larvae and eggs were stored in 80% ethanol / distilled water at 0°C in the field and later sorted with the aid of microscopy. Captures of individual fish larvae types ranged from a single specimen to many hundreds. Typically, groups (n=5 to 20 or more) of the same species were captured within any collection interval within the sampling period.

Adult fish capture

Representative juvenile and adult fish of known species for this study were caught by line fishing as approved by Queensland Fisheries Regulations. Small pectoral fin clip samples were stored initially in 20% DMSO / saturated brine solution at room temperature, and later washed and transferred into 80% ethanol solution and refrigerated. The field work was sanctioned by Marine Parks Permit G04/12132.1 and CQU Animal Ethics Permit A04/06-160.

DNA sequencing

Total DNA was extracted from larval and adult fish tissue using a standard proteinase K, NaCl/chloroform method (Sambrook *et al.*, 1989). The *cox1* gene (approximately 650 bp in length) located in the mitochondrial genome was amplified using three sets of primers synthesised by Geneworks, Adelaide: FishF1-5'TCAACCAACCACAAAGACATTGGCAC3', FishR1-5'TAGACTTCTGGGTGGCCAAAGAATCA3' (Ward *et al.*, 2005); cichlidBS1(F)-5'GGTCAACAAATCATAAAGATATTGG3', cichlidBS2(R)-5'TAAACTTCAGGTGACCAAAAAATCA3' (Sparks, 2003); and tautog-BS1(F) 5'AGTATAAGCGTCTGGGTAGTC3' tautog-BS2(R)- 5'CCTGCAGGAGGAGGAGAYCC3' (Orbacz and Gaffney, 2000). PCR of the *HVR1* genetic sequence utilised universal primers

L15995 and H16498, previously reported to generate the *HVR1* fragment in parrot fish (Bay *et al.*, 2004).

The *cox1* and *HVR1* fragments were amplified as follows. Each PCR reaction mix totalled 25 μl and included 100 μM dNTP, 7.5 pmol of each primer, 1 mM $MgCl_2$, 2.5 μl 10X Buffer (Promega), 1.0 unit Taq polymerase (Promega), 17.8 μl sterile MQ water and 1 μl of DNA template (ca. 25 ng). Either a Perkin Elmer GeneAmp PCR System 2400, a BIO-RAD iCycler, or a BIO-RAD Mycylcer thermal cycler was used. Standard PCR conditions varied slightly for the 3 sets of *cox1* primers tested in this study. For the tautog and cichlid primers the thermal cycle regime consisted of: 5 min at 95°C for 1 cycle; 1 min at 95°C, 1 min at 47°C, and 1 min at 72°C for 35 cycles; and 5 min at 72°C for 1 cycle. Samples were then held at 4°C until retrieved. The thermal cycle for the Fish-1F and R1 primers consisted of: 2 min at 95°C for 1 cycle; 30 s at 94°C, 30 s at 54°C, and 1 min at 72°C for 35 cycles; and 10 min at 72°C. Samples were then held at 4°C on the thermocycler until collection. For *HVR1* amplification, Touchdown PCR was used. The protocol involved an initial 2 min at 94°C followed by 5 cycles of 30 s at 94°C, 30 s at 45°C and 2 min at 72°C, then 5 cycles with annealing temperature reduced to 43°C, then 25 cycles with annealing temperature reduced to 41°C, followed finally by 10 min at 72°C.

The amplified products and size standards were run on a 1% agarose gel stained with Sybr Green (Astral). The gels were visualised and photographed using a BIO-RAD Gel Documentation Camera. Bands (ca. 400 bp for *HVR1* and ca. 650 bp. for *cox1*) were excised from the gels and placed into separate Eppendorf tubes. The Promega Wizard SV PCR and Gel Cleanup System was used to extract the DNA from the gel. The amplified products were then sequenced using the BigDye® Terminator v3.1 Cycle sequencing kit (Applied Biosystems, Inc) following the thermocycler protocol listed by Applied Biosystems and sequenced both in the forward and reverse directions using an ABI Prism 310 genetic analyser or an AB 3130 Genetic Analyser.

Data analysis

For the major *Lethrinus* and *Lutjanus* species, 20-30 individual fish samples were sequenced. For larvae that were single or few samples, the sequencing was done in both the forward and reverse directions and the sequences were then checked manually to see that the sequence information was entirely consistent from both directions. Any doubtful base calls were checked manually. If necessary, the sequencing run or the entire DNA extraction/amplification/sequencing reaction was repeated to resolve any doubtful data.

Sequences were aligned using Chromas v1.45 (Technelysium Pty. Ltd., Australia) and BioEdit v7.0.4.1 (Ibis Therapeutics, CA., USA) freeware. Genetic distances (corrected) were calculated using the Tamura-Nei model within MEGA v3.1 (Kumar *et al.*, 2004). Phylogenetic trees were calculated using the Neighbour Joining procedure of MEGA v3.1, the model being number of differences and complete deletion for gaps and missing data. The following MtDNA *HVR1* sequences (species, accession number) were accessed from GenBank: *Chlorurus sordidus*, AY392743; *Lutjanus erythropterus*, AY664534; *Pristopomoides multidens*, AF192863; *Siganus vulpinus*, AY057327; *Atherinomorus ogilbyi*, AY026097; *Siganus doliatus*, AY057325; *Engraulis japonicus*, DQ219881.

RESULTS

Within species genetic variation of *HVR1*

Figure 1 shows typical 1-3% within-species variation for mtDNA *HVR1* sequences (approximately 420 bp) for reef snapper species *Lethrinus miniatus* and *Lethrinus laticaudis*, and for red emperor (*Lutjanus sebae*) for groups of fish from the same reef location (n=15-25).

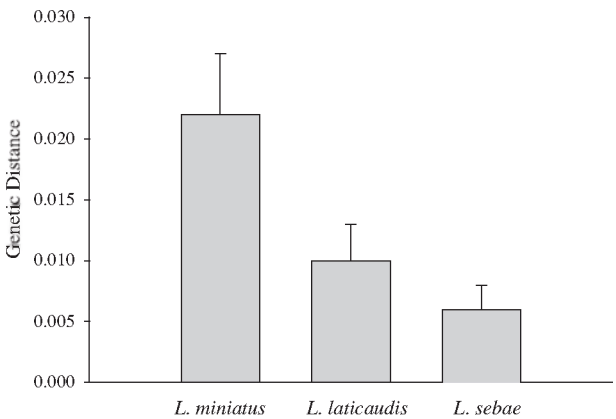

FIG. 1. – Within-species variation for mtDNA *HVR1* sequences for some commercial reef fish species calculated using MEGA v3.1 with Tamura-Nei model (n=15-25 for each group; Means + SEM).

TABLE 1. – Between-species *HVR1* mean genetic distances for example reef snapper species calculated using MEGA v3.1 with Tamura-Nei model.

	Lethrinus laticaudis	*Lethrinus miniatus*	*Lutjanus sebae*
Lethrinus miniatus	0.31		
Lutjanus sebae	0.47	0.50	
Lutjanus adetii	0.46	0.48	0.17

Between species genetic variation of *HVR1*

A matrix of the between species mtDNA *HVR1* genetic distances for commercially important reef snapper species *L. laticaudis*, *L. miniatus* and *L. sebae* is shown in Table 1. The data shows between-species *HVR1* sequence differences ranging from 17 to 50% among these three species, demonstrating the potential utility of *HVR1* sequences for species specific markers.

Primer selectivity for marine fish *cox1* sequences

All three primer sets tested for *cox1* amplification yielded amplicons having sizes of approximately 650 bp and for which BLAST analysis using GenBank gave partial matches for fish species cytochrome oxidase, or portions of total mitochondrial DNA gene sequences. However, the FishF1 and R1 primer set for *cox1* (Ward *et al.*, 2005) consistently gave amplicons which yielded longer length and cleaner sequence data than with either of the other two primer sets tested. The amplicons generated from the same DNA sample for *Lutjanus sebae* with FishF1/R1 and tautog-BS1 (F/R) showed only 50% sequence similarity when analysed using Chromas v1.45 and BioEdit v7.0.4.1. As discussed by Ward *et al.* (2005), researchers need to be mindful of pseudo-gene amplification and the tautog primers used might have initiated amplification of shorter nuclear DNA sequences originating from mitochondrial DNA (NUMTs), as discussed by Zhang and Hewitt (1996) and Richly and Lester (2004), or they might have amplified other mitochondrial DNA besides *cox1*.

Species discrimination using *cox1* sequence comparisons

Nearest neighbour analyses of the *cox1* sequences generated in this study were conducted with the barcode of life fish ID database for Australian fishes (available at www.barcodinglife.org), yielding either identification for fish species *cox1* sequences already in the database, or logical closest related species in all cases. This analysis and the recent report by Ward *et al.* (2005) strongly support the contention that *cox1* sequences are highly useful for identification of marine fish species.

DISCUSSION

The question of fish and fish egg identification using DNA analysis methods is being pursued by many laboratories around the world. The use of modern molecular genetics techniques combined with taxonomic expertise provides a very powerful approach to solving existing taxonomic dilemmas, and allows new insights into the relatedness and evolution of fish species.

In a recent study (Aspden *et al.*, 2005), our group utilised *HVR1* sequences to study the population differences for red throat emperor *Lethrinus miniatus* across its distribution along the East and West Australian coasts. In the present study we compared the within species *HVR1* sequence variation for two other commercial reef fish species (*Lethrinus laticaudis* and *Lutjanus sebae*) compared with *L. miniatus* (Fig. 1). With only 1-3% variation within species, but up to 50% difference between species (Table 1), it therefore seemed feasible to explore using *HVR1* sequences to develop a putative taxonomic identification tree for reef fish species. An example partial tree construct using neighbour joining analysis for fish *HVR1* sequences is shown in Figure 2. Fish from the same genus (see the Lethrinid and Lutjanid examples shown) all align closely, suggesting that though *HVR1* is a non-coding genetic sequence, this gene marker is remarkably useful for distinguishing between closely related species. Some of the larvae accessions from the 2004 collection for this study are also shown in Figure 2 and are indicated with an L prefix. As examples, L04 and L11 were closely aligned to *Engraulis japonicus* and are proposed to be the Australian anchovy *Engraulis australis*. The analysis suggests that L09 is a *Lethrinus* species, most likely a larval form of the spangled emperor *Lethrinus nebulosus*, while L03, L08, L13 and L24 are *Siganus* species. Tropical pelagic species such as hardy

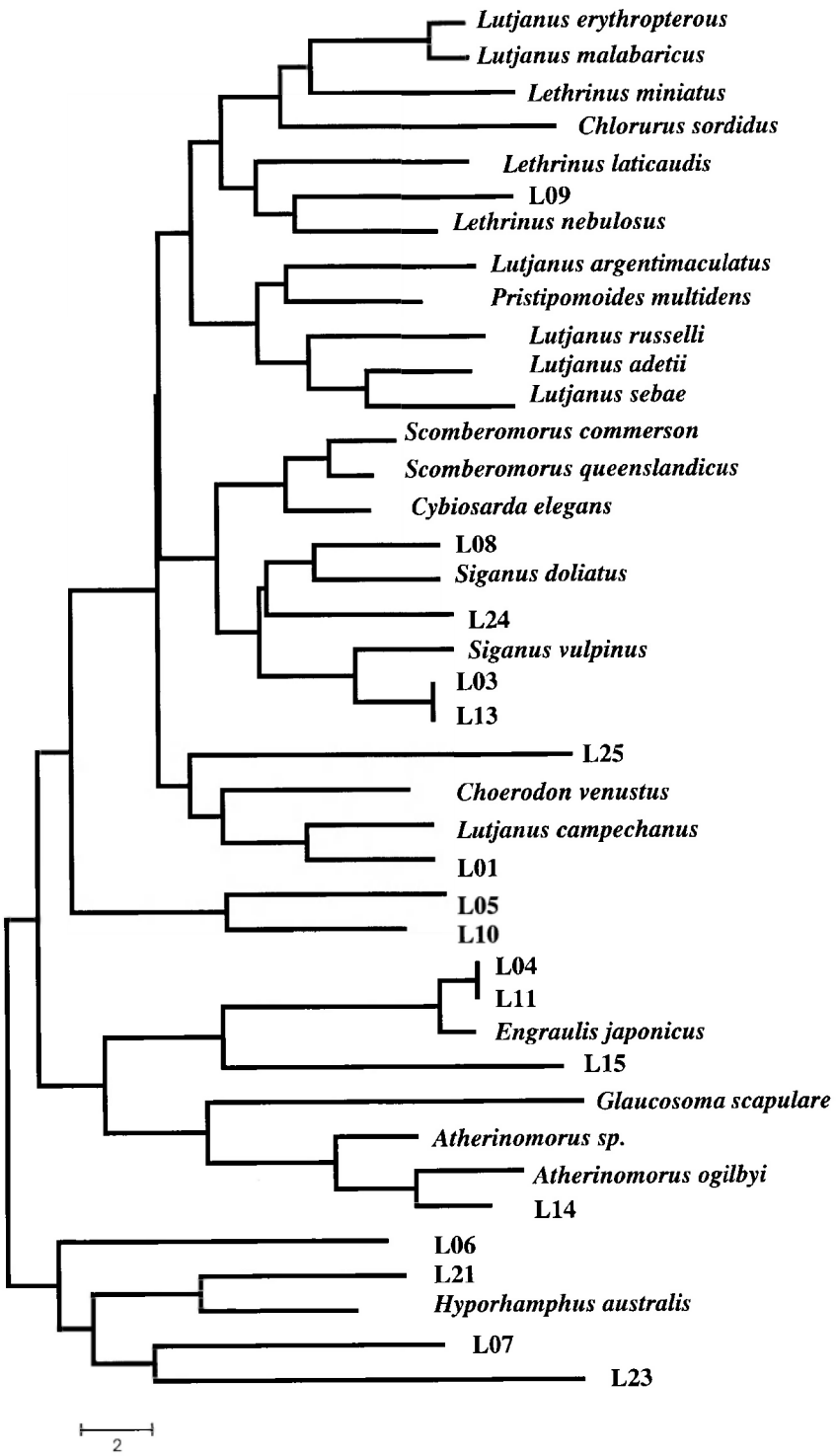

FIG. 2. – NJ topology of larvae by MtDNA *HVR1* sequence comparisons calculated using MEGA v3.1.

heads *Atherinomorus* (L14) are also readily identified by *HVR1* sequence comparisons.

While encouraged by these preliminary data proving the usefulness of *HVR1* sequence comparisons for species identification, the international Barcode of Life project initiated by Hebert *et al.* (2003) suggested that another mitochondrial gene sequence, namely cytochrome oxidase subunit 1 (*cox1*), be used as the gene marker of choice for species discrimination. As it is a structural gene encoding a functional respiratory chain enzyme, it might be expected that *cox1* sequence data would

show fewer between-species differences, and this might compromise discrimination between closely related species. For this study, the efficacy of three different published primer pairs for amplification of *cox1* in fish tissue samples was compared, namely FishF1, FishR1 (Ward *et al.*, 2005); cichlidBS1(F), cichlidBS2(R) (Sparks, 2003); and tautog-BS1(F), tautog-BS2(R) (Orbacz and Gaffney, 2000). While all of the primer sets studied yielded PCR amplicons of approximately 650 bp from fish tissue DNA extracts with varying success, the FishF1, FishR1 recently reported by Ward *et al.* (2005) for DNA barcoding of Australian fishes gave the most consistent results across a range of species tested and are therefore recommended to other workers (the authors thank Dr Ward for a pre-publication copy of his paper which assisted this study). Future workers in this field are advised to consider some of the *cox1* DNA sequence data available in the public domain with caution, since some of the reported sequences may in fact be of genomic origin (Ward *et al.*, 2005; Zhang and Hewitt 1996; Richly and Lester 2004), or of other mitochondrial origin. For all fish species tested in this study (including all commercially important *Plectropomus*, *Epinephelus*, *Lethrinis*, and *Lutjanus* examples), the *cox1* sequences generated using the FishF1, FishR1 primer set gave reproducible sequences for each species that allowed ready species discrimination.

CONCLUSIONS

We conclude that the DNA Barcode approach for fish identification appears valid and that while *HVR1* or *cox1* mtDNA sequence data both appear useful for this purpose, *cox1* should be used in future studies as the marker of choice since a large international database for fish identification is presently being constructed. The *cox1* amplification and sequencing method allows identification of planktonic larval fish and fish eggs through comparison to DNA from authenticated adult fish specimens, and therefore provides a major new advance for fisheries biologists, taxonomists and fisheries regulators.

REFERENCES

Aspden, W.J., G.G. Pegg, L. van Heerwerden, S.J. Newman, and W. Sinclair. – 2005. Population Genetics of Red Throat Emperor of Australian East and West Coasts and Norfolk Island Based on Mitochondrial DNA Sequence. *The 7th Indo Pacific Fish Conference*, Taipei, Taiwan. Abstract S1-07: p. 15.

Bay, L.K., J.H. Choat, L. van Heerwerden and D.R. Robertson. – 2004. High genetic diversities and complex genetic structure in an Indo-pacific tropical fish species (*Chlorurus sordidus*): evidence of an unstable evolutionary past? *Mar. Biol.*, 144: 757-767.

Hebert, P.D.N., A. Cywinska, S.L. Ball and J.R. de Waard. – 2003. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B.*, 270: 313-322.

Kumar S., K. Tamura, and M. Nei. – 2004. MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment. *Brief. Bioinform.*, 5: 150-163.

Leis, J.M. and B.M Carson-Ewart (eds.) – 2004. *The larvae of Indo-Pacific coastal fishes*. 2nd edition, Leiden; Boston: Brill.

Orbacz, E.A and P.M. Gaffney. – 2000. Genetic structure of tautog (*Tautoga onitis*) populations assayed by RFLP and DGGE analysis of mitochondrial and nuclear genes. *Fish. Bull.*, 98: 336-344.

Richly, E. and D. Leister. – 2004. NUMTs in sequenced eukaryotic genomes. *Mol. Biol. Evol.*, 21: 1081-1084.

Sambrook, J., E.F. Fritch, and T. Maniatis. – 1989. *Molecular Cloning : a Laboratory Manual*, 2nd edition, Cold Spring Harbour Laboratory press, Cold Spring Harbour.

Sparks, J.S. – 2004. Molecular phylogeny and biogeography of the Malagasy and South Asian cichlids (Teleostei: Perciformes: Cichlidae). *Mol. Phylogenet. Evol.*, 30: 599-614.

Steinke, D., M. Vences, W. Salzburger, and A. Meyer. – 2005. TaxI: a software tool for DNA barcoding using distance methods. *Phil. Trans. R. Soc. B.*, 360: 1975-1980.

Ward, R.D., T.S. Zemlak, B.H. Innes, P.R. Last, and P.D. Hebert. – 2005. DNA Barcoding of Australia's fish species. *Phil. Trans. R. Soc. B.*, 360: 1847-1857.

Zhang, D-X. and G.M. Hewitt. – 1996. Nuclear integrations: challenges for mitochondrial DNA markers. *Trends Ecol. Evol.*, 11: 247-251.

Received February 9, 2005. Accepted November 21, 2005.
Published online September 28, 2006.