

Universal cytochrome *b* primers facilitate intraspecific studies in molluscan taxa

T.J.S. Merritt and L. Shi

Department of Biological Sciences, University of South Carolina, Columbia, South Carolina 29208, U.S.A.

M.C. Chase, M.A. Rex, and R.J. Etter

Department of Biological Sciences, University of Massachusetts–Boston, Boston, Massachusetts 02125, U.S.A.

J.M. Quattro*

Program in Marine Science, Baruch Institute, and School of the Environment, Department of Biological Sciences, University of South Carolina, Coker Life Sciences Building, Columbia, South Carolina 29208, U.S.A.

Abstract

We describe the construction of amplification primers designed to target a portion of the mitochondrial cytochrome *b* locus in a variety of molluscan taxa. Combinations of two sets of primers successfully amplified cytochrome *b* from several species of gastropods, bivalves, and cephalopods. Sequence analysis of these amplified products revealed nucleotide diversity in small samples within several of these taxa. We discuss the utility of these primer sets for studies of intraspecific phylogeny in mollusks and potentially other invertebrates.

Introduction

We report here a set of “universal” primers for the polymerase chain reaction (PCR) amplification of a fragment of approximately 430 base pairs (bp) of the mitochondrial cytochrome *b* locus (*Cytb*) developed as part of an ongoing study of deep-sea mol-

lusks. With the growing popularity of sequence data for studies of interspecific and intraspecific phylogeny, the need for primer sets that target gene regions appropriate for taxonomic resolution at various levels is increasing (e.g., Folmer et al., 1994; Cho et al., 1995; Palumbi, 1996). Our goal was to develop primers useful at or below the species level in a variety of molluscan taxa.

It has been shown that the *Cytb* gene is accumulating nucleotide changes at a sufficient rate to resolve phylogenetic relationships among closely related species of snails (Collins et al., 1996; Reid et al., 1996). We demonstrate here that sufficient variability exists within the *Cytb* locus for it to be useful in studies of intraspecific phylogeny. Although the primers reported here were initially designed to amplify *Cytb* in molluscan taxa, we have found that they amplify a variable portion of the cytochrome *b* locus from other invertebrate taxa as well.

Results and Discussion

We designed an original primer set from a CLUSTALW (Higgins et al., 1996) alignment of published cytochrome *b* sequences from *Artemia franciscana* (Perez et al., 1994), *Cepaea nemoralis* (Terrett et al., 1996), *Drosophila yakuba* (Clary et al., 1984), *Katharina tunicata* (Boore and Brown, 1994), *Mytilus edulis* (Hoffmann et al., 1992), *Anopheles gambiae* (Beard et al., 1993), *Ascaris suum* (Okimoto et al., 1992), and *Apis mellifera* (Crozier and Crozier, 1993). Two highly conserved regions of sequence (Table 1) were identified by the computer program PRIMEGEN (O'Hara and Venezia, 1991). Degenerate oligonucleotides were designed to target these regions:

UCYTB144F (438): 5'-TGA GSN CAR ATG TCN TWY TG-3'

UCYTB272R (826): 5'-GCR AAN AGR AAR TAC CAY TC-3'

where degenerate positions are represented by the following ambiguity codes:

N = A | C | G | T; R = A | G; S = C | G; W = A | T; Y = C | T;

* Correspondence should be sent to this author at the above address; tel.: 803-777-3240; fax: 803-777-4002; e-mail: quattro@biol.sc.edu.

Table 1. Conserved regions of the cytochrome *b* locus used to design degenerate oligonucleotide primers UCYTB144F and UCYTB272R.*

Primer and taxa	Sequence						
	W	G	Q	M	S	F	W
UCYTB144F	5'- T G A	G S N	C A R	A T G	T C N	T W Y	T G -3'
<i>Artemia franciscana</i>	. . G	. G T	. . A	. . G	T C C	T T C	. .
<i>Cepaea nemoralis</i>	. . A	. G A	. . A	. . A	T C T	T A T	. .
<i>Drosophila yakuba</i>	. . A	. G A	. . A	. . A	T C A	T T T	. .
<i>Katharina tunicata</i>	. . A	. G A	. . G	. . A	T C T	T T T	. .
<i>Mytilus edulis</i>	. . G	. G G	. . G	. . A	T C G	T A C	. .
<i>Anopheles gambiae</i>	. . A	. G A	. . A	. . A	T C A	T T T	. .
<i>Ascaris suum</i>	. . G	. C A	. . A	. . G	A G G	T T T	. .
<i>Apis mellifera</i>	. . A	. G A	. . A	. . A	T C A	T A T	. .
UCYTB272R	5'- G A R	T G G	T A Y	T T Y	C T N	T T Y	G C -3'
<i>Artemia franciscana</i>	. . A	. . A	. . T	. . T	C . G	. . T	. .
<i>Cepaea nemoralis</i>	. . A	. . A	. . C	. . T	C . G	. . T	. .
<i>Drosophila yakuba</i>	. . A	. . A	. . T	. . T	T . A	. . T	. .
<i>Katharina tunicata</i>	. . A	. . G	. . T	. . T	C . A	. . C	. .
<i>Mytilus edulis</i> [†]	—	—	—	—	—	—	—
<i>Anopheles gambiae</i>	. . A	. . A	. . C	. . T	T . A	. . T	. .
<i>Ascaris suum</i>	. . G	. . A	. . T	. . T	T . G	. . T	. .
<i>Apis mellifera</i>	. . A	. . A	. . T	. . C	C . A	. . T	. .

* The coding strands are shown in both instances.

[†] The sequence for the UCYTB272R region in *Mytilus edulis* has not been published.

and numbers in parentheses refer to the position of the 3' base in the *Artemia franciscana* sequence.

UCYTB144F and UCYTB272R were used to amplify a 430-bp fragment from several molluscan taxa including *Mercenaria mercenaria*, *Demimacula proxima*, *Mya arenaria*, and *Littorina littorea*. These fragments were cloned and sequenced, then used to design a second set of oligonucleotide primers:

UCYTB151F (458): 5'-TGT GGR GCN ACY GTW
ATY ACT AA-3'

UCYTB270R (820): 5'-AAN AGG AAR TAY CAY
TCN GGY TG-3'

where degenerate positions and numbers in parentheses are as above. This second set of primers, or combinations of the two primer sets, ultimately generated more consistent amplifications from a greater variety of taxa than the original pair. Using various combinations of the four primers, but in particular UCYTB151F and UCYTB270R, we have amplified, and in most cases sequenced, a fragment of the *Cytb* gene from a variety of molluscan and in some cases nonmolluscan taxa (Table 2).

To verify the amplified products as *Cytb*, a minimum of 200 bp was sequenced from most taxa and

compared with other sequences in GenBank (release 100.0). For example, a homology search with the *Loligo pealei* fragment as a query sequence returned a highly significant match with a reported *Cytb* sequence from a gastropod mollusk, *Plicopurpura patula* (Collins et al., 1996). Figure 1 shows a block of amino acid sequence from *L. pealei* and three other representative species with *Artemia franciscana* included as a reference.

Our primary objective was to design primers that would be useful for intraspecific comparisons. Figure 2 shows allelic variation uncovered within a

<i>Artemia franciscana</i>	NDVVQWLWGG FAVDNPFLTR FFFHFHFLIPF LVAGLTMHLL
<i>Solomya velum</i>	KML.E.I... .Y..A..S. .YA...F.. V.I.A.SAL.M
<i>Loligo pealei</i>	?????????? ?????????? ???...L.. VLMAMV.L.F
<i>Acmaea testudinalis</i>	EPI...I... .S.SGV... .YG...I.A.. VLLIFVVG.I
<i>Panaeus duorarum</i>	TEL...I... ..A... ..F... I...AA.I..I
<i>Artemia franciscana</i>	LFLHQSGSNN PLGINANLDK LPEHFYFTI
<i>Solomya velum</i>	...EC... ..V.SDSE. I...I.Y.V
<i>Loligo pealei</i>	...T... ..G... I...Q.YSY
<i>Acmaea testudinalis</i>	.L...EN..S. ...VSS.RA. IR...YST
<i>Panaeus duorarum</i>	..I..T... ..VS.V.. V.SILLYIK

Figure 1. Inferred amino acid sequences from a portion of the cytochrome *b* for several invertebrate taxa. The *Artemia franciscana* sequence is included as a reference. Question marks denote regions of undetermined sequence.

Table 2. Thirteen species from three classes of mollusk and two classes of arthropod in which a combination (indicated) of *Cytb* primers amplified a product of expected size.

Phylum	Class	Order	Genus/species	Primer set
Mollusca	Bivalvia	Nuculoida	<i>Yoldia limatula</i>	151F–272R*
			<i>Deminucula proxima</i>	144F–270R*
		Solemyoidea	<i>Solemya velum</i>	151F–270R*
		Myoidea	<i>Mya arenaria</i>	144F–272R*
		Veneroidea	<i>Mercenaria mercenaria</i>	144F–272R*
	Cephalopoda	Teuthoidea	<i>Loligo pealei</i>	151F–272R*
			<i>Illex illecebrosus</i>	151F–272R*
			<i>Nautilus pompilius</i>	151F–270R
			<i>Littorina littorea</i>	144F–272R*
			<i>Acmaea testudinalis</i>	151F–272R*
Gastropoda	Mesogastropoda	<i>Penaeus duorarum</i>	151F–270R*	
		<i>Penaeus aztecus</i>	151F–270R*	
Arthropoda	Malacostraca	Decapoda	<i>Microarthridion littorale</i>	151F–270R*
	Maxillopoda	Harpacticoida		

* Denotes a taxon in which direct sequencing was used to verify the amplified product as *Cytb*.

sample of seven individuals of the protobranch bivalve *Deminucula proxima* (Figure 2A) and 10 individuals of the long-finned squid *Loligo pealei* (Figure 2B). All seven *D. proxima* individuals harbored unique haplotypes that were defined by mutations at nine nucleotide positions (Figure 2A). All mutations were transitions, and all involved C-to-T changes. Eight of the nine mutational events occurred at third positions, one mutation occurred at the first position of a leucine codon, and all changes were synonymous. Haplotype diversity was also observed at the *Cytb* locus in a small sample of long-finned squid (*Loligo pealei*; Figure 2B). Synonymous changes at two nucleotide positions defined three haplotypes within a sample of 10 individuals.

We have also observed intraspecific haplotype diversity within small samples of other molluscan taxa including short-finned squid (*Illex illecebrosus*) and tortoise-shell limpet (*Acmaea testudinalis*) (data not shown). However, no nucleotide variation was observed within small samples (six or seven individuals) of the spoon clam (*Yoldia limatula*) or the veil clam (*Solemya velum*). Although not extensively tested, we have found that our primers amplify this same region of *Cytb* from a variety of nonmolluscan taxa including shrimp (*Penaeus duorarum* and *P. aztecus*) and copepods (*Microarthridion littorale*) (Table 2). Considerable sequence variation has been observed across individuals of *M. littorale* (N. Schizas and J.M. Quattro, unpublished data). In contrast, we have not observed any sequence variation within or between samples of pink shrimp (*P. duorarum*) from South

Carolina and Florida (J.M. Quattro, unpublished data).

We have designed a set of “universal” primers that facilitates the amplification of a polymorphic mitochondrial locus in a variety of molluscan taxa. Surveys of sequence variation in several molluscan species show the utility of these primers for population-level studies in at least some taxa. Our initial work with nonmolluscan groups is promising and suggests that these primer sets might show broad use for population-level studies in diverse invertebrate taxa. Given the variety of taxa in which these primers have been shown to work, and the amount of sequence variability across individuals within species, these primers show potential for broad utility in systematic studies at or below the species level.

Experimental Procedures

PCR amplification

Total DNA was isolated from fresh, frozen, or alcohol-preserved samples using a QIAamp Tissue kit following instructions of the manufacturer (Qiagen). We typically used 5 µl (approximately 100 ng) of DNA extract as template in a 50-µl PCR reaction. For the squid species, 2.5 µl of purified DNA was treated with 5 µl of GeneReleaser (Bioventures, Inc.), and the resulting mixture was used for PCR amplification. Each 50-µl reaction consisted of 5 µl of 10× reaction buffer (100 mM Tris, pH 8.3; 25 mM MgCl₂; 500 mM KCl; 0.1% Triton X-100), 5 pmol of forward and reverse primer, 200 µM each dNTP, and 1 unit *Taq* DNA polymerase (Promega). Typi-

A.

```

Dp5 AAG ATC GTA GTT CAG TGA GTT TGA GSA GGG TTT GCG GTT GAA AAC GCA ACT TTA ACT 954
Dp1 ... ..C ..T ..
Dp3 ... ..C ..
Dp4 ... ..C ..
Dp6 ... ..C ..
Dp2 ... ..C ..
Dp7 ... ..C ..

Dp5 CGA TTT TTT AGT TTT CAT TTT TTT TTT CCA TCC GCG ATT GCT GCG TTT AGA GTC TTA 1008
Dp1 ... ..C ..
Dp3 ... ..C ..
Dp4 ... ..C ..
Dp6 ... ..C ..
Dp2 ... ..C ..
Dp7 ... ..C ..

Dp5 CAT ATT TTA TTT TTA CAT GAG ACT GGG TCT AAT AAT CCT TTG GGG ATG AAT AGA GAT 1066
Dp1 ... ..C ..
Dp3 ... ..C ..
Dp4 ... ..C ..
Dp6 ... ..C ..
Dp2 ... ..C ..
Dp7 ... ..C ..

Dp5 AGT GAA AAA ATT CCT TTT CAT ATT TAT TT 1095
Dp1 ... ..C ..
Dp3 ... ..C ..
Dp4 ... ..C ..
Dp6 ... ..C ..
Dp2 ... ..C ..
Dp7 ... ..C ..

```

B.

```

Lp1 GTT TTC ATT TTT TAT TAC CTT TTG TTC TGA TGG GAA TAG TGA TAT TAC ATT TTT TGT 1020
Lp4 ... ..C ..
Lp8 ... ..C ..

Lp1 TTC TTC ATC AGA CCG GTA GTA ATA ATC CTT TAG GTA TTA ATG GTA ATT TAG ATA AAA 1077
Lp4 ... ..C ..
Lp8 ... ..C ..

Lp1 TTC CAT TTC ATC AAT ACT ATA GTT ATA AGG ATT TAA TGG GGT TTT TTG TTA TGT TGT 1134
Lp4 ... ..C ..
Lp8 ... ..C ..

Lp1 TGT TAT TGA TGG AAA TTA GAA TGT TAT TTT CTA ATG CTT TAG GGG ATT CTC GAA 1188
Lp4 ... ..C ..
Lp8 ... ..C ..

```

Figure 2. Intraspecific diversity at the *Cytb* locus in select molluscan taxa: (A) *Deminucula proxima* and (B) *Loligo peali*. Sequences are displayed as codons; numbers to the right of the reference sequence denote the corresponding position within the *Cytb* gene of *Artemia franciscana*.

cally, reactions were amplified through 40 cycles under the following conditions: an initial denaturation for 4 min at 94°C, 1 min at 94°C, 1 min at 48°C, 2 min at 72°C, followed by a final extension step at 72°C for 6 min.

DNA sequencing

In some instances amplification products were cloned directly into commercial preparations of T-vector (Promega). However, most templates could be sequenced from a single round of amplification. Occasionally templates provided too little product from a single amplification. In these cases a sample of the first amplification product was diluted 1:100, and a 2- μ l aliquot was used as template for a secondary amplification under identical PCR conditions. Single-stranded PCR product for sequencing was obtained through solid-phase methods (Salminen, 1992). Dideoxy sequencing reactions were performed manually using Sequenase (version 2.0, U.S. Biochemical).

Acknowledgments

We thank P. Gaffney, W.J. Jones, K. Oswald, N. Schizas, G. Street, and R. Wetzler for valuable comments on an earlier version of the manuscript. N. Schizas provided some unpublished *Cytb* sequences. The Nautilus sample was kindly provided by the Marine Biomedical Institute of the University of Texas Medical Branch. This research was supported, in part, by grants from the National Science Foundation (OCE-9402855), Environmental Protection Agency (R 825439-01-0), and the Cooperative Institute for Fisheries Molecular Biology (FISHTEC; NOAA/NMFS [RT/F-1]).

References

- Beard, C.B., Hamm, D.M., and Collins, F.H. (1993). The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. *Insect Mol Biol* 2:103–104.
- Boore, J.L., and Brown, W.M. (1994). Complete DNA sequence of the mitochondrial genome of the black chiton, *Katharina tunicata*. *Genetics* 138:423–443.
- Cho, S., Mitchell, A., Regier, J.C., Mitter, C., Poole, R.W., Friedlander, T.P., and Zhao, S. (1995). A highly conserved nuclear gene for low-level phylogenetics: *Elongation Factor-1* recovers morphology-based tree for heliothine moths. *Mol Biol Evol* 12:650–656.
- Clary, D.O., Wahleithner, J.A., and Wolstenholme, D.R. (1984). Sequence and arrangement of the genes for cytochrome *b*, URF1, URF4L, URF4, URF5, URF6 and five tRNAs in *Drosophila* mitochondrial DNA. *Nucleic Acids Res* 12:3747–3762.
- Collins, T.M., Frazer, K., Palmer, A.R., Vermeij, G.J., and Brown, W.M. (1996). Evolutionary history of northern hemisphere *Nucella* (Gastropoda, Muricidae): molecular, morphological, ecological, and paleontological evidence. *Evolution* 50:2287–2304.
- Crozier, R.H., and Crozier, Y.C. (1993). The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genomic organization. *Genetics* 133:97–117.
- Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for the amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 3:294–299.
- Higgins, D.G., Thompson, J.D., and Gibson, T.J. (1996). Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 266:383–402.
- Hoffman, R.J., Boore, J.L., and Brown, W.M. (1992). A novel mitochondrial genome organization for the blue mussel, *Mytilus edulis*. *Genetics* 131:397–412.
- Okimoto, R., Macfarlane, J.L., Clary, D.O., and Wolstenholme, D.R. (1992). The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*. *Genetics* 130:471–498.
- O'Hara, P.J., and Venezia, D. (1991). PRIMEGEN, a tool for designing primers from multiple alignments. *Comput Appl Biosci* 7:533–534.
- Palumbi, S.R. (1996). Nucleic acid II: the polymerase chain

- reaction. In: Hillis, D.M., Moritz, G., Mable, B.K. (eds.). *Molecular Systematics*. Sunderland, Mass.: Sinauer Associates, 205–247.
- Perez, M.L., Valverde, J.R., Batuecas, B., Amat, F., Marco, R., and Garesse, R. (1994). Speciation in the *Artemia* genus: mitochondrial DNA analysis of bisexual and parthenogenetic brine shrimps. *J Mol Evol* 38:156–168.
- Reid, D.G., Rumbak, E., and Thomas, R.H. (1996). DNA, morphology and fossils: phylogeny and evolutionary rates of the gastropod genus *Littorina*. *Phil Trans R Soc Lond B Biol Sci* 351:877–895.
- Salminen, M. (1992). Rapid and simple solid phase direct sequencing of in vivo HIV-1 quasispecies. *AIDS Res Hum Retroviruses* 8:1747–1756.
- Terrett, J.A., Miles, S., and Thomas, R.H. (1996). Complete DNA sequence of the mitochondrial genome of *Cepaea nemoralis* (Gastropoda: Pulmonata). *J Mol Evol* 42:160–168.