Ancient Phylogenetic Separation between Pacific and Atlantic Cephalochordates as Revealed by Mitochondrial Genome Analysis

Authors: Nohara, Masahiro, Nishida, Mutsumi, Manthacitra, Vipoosit, and Nishikawa, Teruaki

Source: Zoological Science, 21(2) : 203-210

Published By: Zoological Society of Japan

URL: https://doi.org/10.2108/zsj.21.203

ZOOLOGICAL SCIENCE **21**: 203–210 (2004)

Ancient Phylogenetic Separation between Pacific and Atlantic Cephalochordates as Revealed by Mitochondrial Genome Analysis

Masahiro Nohara[1,2], Mutsumi Nishida[2*], Vipoosit Manthacitra[3] and Teruaki Nishikawa[4]

[1]*Yokohama R&D Center, HITEC Co. Ltd., 2-20-5 Minamisaiwai, Nishi-ku, Yokohama, Kanagawa 220-0005, Japan*
[2]*Ocean Research Institute, University of Tokyo, 1-15-1 Minamidai, Nakano, Tokyo 164-8639, Japan*
[3]*Faculty of Science, Burapha University, Bangsean, Chonburi 20130, Thailand*
[4]*The Nagoya University Museum, Nagoya University, Chikusa, Nagoya, Aichi 464-8601, Japan*

ABSTRACT—The subphylum Cephalochordata (lancelets) is a relatively small taxonomic group in contrast to the subphyla Urochordata and Vertebrata. As an initial step to determine whether lancelets exhibit small genetic divergence in keeping with their conservative body organization or large genetic variation, four *Branchiostoma* species from the Pacific (*B. belcheri* and *B. malayanum*) and Atlantic (*B. floridae* and *B. lanceolatum*) Oceans were genetically compared using partial mitochondrial DNA sequences of the cytochrome oxidase c subunit I (COI) and 16S ribosomal RNA (16S rRNA) genes. In both genes, large genetic differences were revealed between the Pacific and Atlantic species, as well as within the former. Two maximum-likelihood trees from the COI and 16S rRNA genes showed that the Pacific and Atlantic lancelets were reciprocally clustered into different clades. Furthermore, both gene trees consistently exhibited deep phylogenetic separation between the two oceans. The estimated divergence time suggested that differentiation may have followed the migration of ancestral lancelets from the Pacific to the Atlantic Oceans via the Tethys Sea.

Key words: lancelet, cytochrome c oxidase subunit I, 16S ribosomal RNA, molecular phylogeny

INTRODUCTION

The subphylum Cephalochordata (lancelets), a benthic marine invertebrate taxon is believed to be the sister group of vertebrates (Vertebrata). Accordingly, it has been well studied for clues to the origin of vertebrates, especially from the points of view of developmental biology and physiology (Gee, 1996; Hall, 1998). However, cephalochordates have remained little studied in terms of evolutionary biology, including phylogeny and population genetics. The main reason for such paucity of evolutionary studies may be the occurrence in Cephalochordata of only ca. 29 known living species, all exhibiting poor morphological variation (Poss and Boschung, 1996), in contrast to the great morphological and species diversity of the Urochordata and Vertebrata (*ca.* 2,500 species in the former and *ca.* 45,000 species in the latter) (based on Table 6 in Minelli, 1993). The recent progress of molecular biological techniques has made the exploration of genetic diversities of organisms more straight forward, the genetic analyses aided by modern techniques being effective particularly in evolutionary studies of morphologically similar organisms. Indeed, comparative genetic studies of congeneric animals using molecular markers have revealed large genetic differentiation (*e.g.* Glenn and Avise, 1998), suggesting that their general body organization has remained stable for a long time following phylogenetic splitting. Lancelets have persisted for a long period of time as indicated by fossil records referred to the cephalochordates [for example, *Pikaia* and *Cathaymyrus* from the Lower Cambrian (Shu *et al.*, 1996)]. Therefore, although the animals have shown few morphological changes, it is possible that they have accumulated significant genetic changes at molecular level, should the extant species have had long histories. However, previous molecular studies of two Atlantic species of *Branchiostoma floridae* Hubbs, 1922 and *B. lanceolatum* (Pallas, 1774) showed small genetic differ-

* Corresponding author: Tel. +81-3-5351-6329;
FAX. +81-3-5351-6579.
E-mail: mnishida@ori.u-tokyo.ac.jp

ences in their complete mitochondrial DNA sequences (Boore *et al.*, 1999; Spruyt *et al.*, 1998). The question as to whether or not such small morphological and genetic differentiation in the Atlantic lancelets implies recent diversification of all extant species can be examined by a genetic survey extended to more species.

In the present study, two *Branchiostoma* species from the Pacific Ocean were genetically surveyed by mitochondrial DNA sequences to examine the extent of their genetic differentiation from the two Atlantic species mentioned above, and their divergence time and probable evolutionary history were discussed.

MATERIALS AND METHODS

Samples

The four *Branchiostoma* species considered here are *B. belcheri* (Gray, 1847), *B. malayanum* Webb, 1956, *B. floridae* and *B. lanceolatum*. Samples of *B. malayanum* and *B. belcheri* were collected from the West Pacific (Fig. 1), the former from Ko Khang Kao Island, Gulf of Thailand, in November 1999 and the latter from Awajishima Island, Central Japan in July 1999. *Epigonichthys lucayanus* (Andrews, 1893) was sampled from Kuroshima Island, SW Japan, in September 2000 and used as an outgroup in the present phylogenetic analysis. All specimens were fixed and preserved in 70% or 99.5% ethanol until analysis.

Data for *B. floridae* collected from Florida, USA and *B. lanceolatum* from Roscoff, France were obtained from the DNA Data Bank of Japan, accession numbers for the species being AF098298 (Boore *et al.*, 1999) and Y16474 (Spruyt *et al.*, 1998), respectively.

DNA preparation, PCR amplification and sequencing

Tissue from the posterior part of each specimen was digested for twelve hours with proteinase K (10 mg/ml) in a lysis buffer [10 mM Tris-HCl, pH 8.0; 2 mM EDTA; 1% SDS (w/v)]. Total DNA of each lancelet was isolated from the digested tissue solution using a standard phenol-chloroform method and ethanol precipitation (Sambrook and Russell, 2001). The isolated DNA was resuspended with TE buffer (10 mM Tris-HCl, pH 8.0; 2 mM EDTA).

The middle to posterior part of the 16S ribosomal RNA (16S rRNA) gene and two parts of the cytochrome oxidase c subunit I (abbreviated as COI) gene on the lancelet mitochondrial genome were amplified with the following primers: L2188 (5'-AGTGGGC-CTAAAAGCAGCCA-3') and H2716i (5'- AAGTTTTATAGGGTCT-TATCGTC-3'; Kitaura *et al.*, 1998) for *ca.* 450 bp of the middle part of the 16S rRNA gene; L2510i (5'-CGCCTGTTTAACAAAAACAT-3'; Palumbi *et al.*, 1991) and H 3058 (5'-TCCGGTCTGAACTCAGAT-CACGTA-3') for ca. 550 bp of the posterior part of the 16S rRNA gene; LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3'; Folmer *et al.*, 1994) and H6609 (5'-ACTTCAGGGTGACCAAAAAAYCA-3'; Shikatani and Nishida, unpublished) for 558 bp of the anterior part of the COI gene; L6631 (5'-TGRTTTTTTGGTCACCCTGAAGT-3'; Shikatani and Nishida, unpublished) and H7227 (5'-CATGTAGTG-TATGCATCAGGGTARTC-3'; Nishida *et al.*, 1998) for 414 bp of the posterior part of the COI gene. The polymerase chain reaction (PCR) was carried out in a 15 μl volume containing TaKaRa Ex Taq™ buffer (2 mM Tris-HCl, pH 8.0; 2 mM $MgCl_2$; 10 mM KCl; 0.01 mM EDTA; 0.1 mM DTT; 0.05% Tween® 20; 0.05% Nonidet P-40®; 5% glycerol), 0.5 units TaKaRa Ex Taq™ polymerase, 2.5 mM each dNTP mixture, 0.5 μM each primer and 10–20 ng template DNA on a thermal cycler (GeneAmp® PCR System 9700, Applied Biosystems) for 30–35 cycles, with the following thermal profile: preheating at 95°C for 2 min, denaturation at 95°C for 15 seconds, annealing at 45°C for 15 seconds and extension at 72°C for 30 seconds.

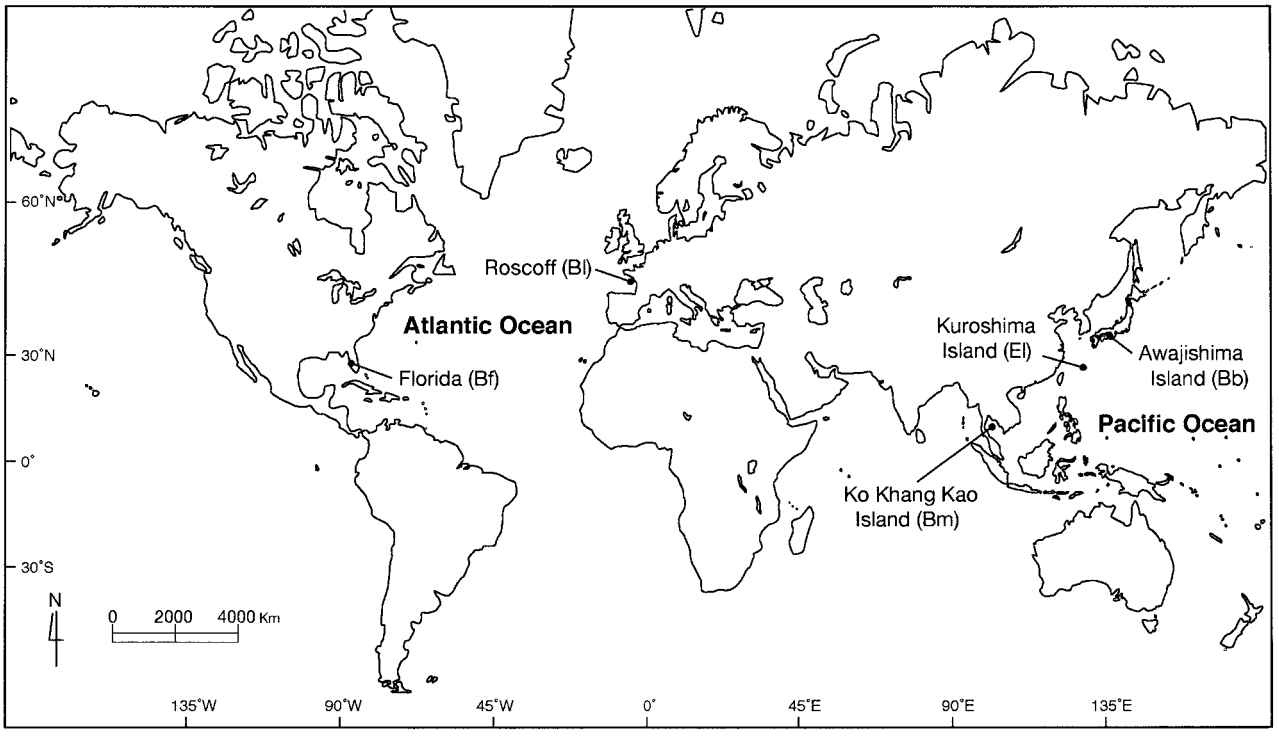

Fig. 1. Sampling localities of lancelets. Bb, Bm, Bf, Bl and El after locality names indicate *Brachiostoma belcheri, B. malayanum, B. floridae, B. lanceolatum,* and *Epigonichthys lucayanus*, respectively. Sampling localities for *B. floridae* and *B. lanceolatum* are approximations following Boore *et al.* (1999) and Spruyt *et al.* (1998), respectively.

Before sequencing the two genes, the double-stranded DNA obtained through PCR was purified with the usb™ PCR Product Pre-Sequencing Kit (USB) composed of exonuclease I and shrimp alkaline phosphatase. Direct sequencing of the purified double-stranded DNA using the BigDye™ Terminator Cycle Sequencing FS Ready Reaction Kit v.2.0 (Applied Biosystems) was performed on an ABI PRISM® 377 DNA Sequencer (Applied Biosystems) or a 310 Genetic Analyzer (Applied Biosystems). DNA sequence data newly determined for *B. belcheri, B. malayanum* and *E. lucayanus* are available from DDBJ/EMBL/GenBank (accession numbers shown in Appendix 1).

Data analysis

Partial sequences of lancelet 16S rRNA and COI genes from the five species (including the outgroup) were primarily aligned with Clustal X (Thompson *et al*., 1997), and then inspected and corrected by eye. Some parts in the aligned sequences of the 16S rRNA gene, totaling 36 sites, were completely excluded in all the present analyses because of their alignment ambiguity (Appendix 2). On the other hand, sequences from the COI gene were unambiguously aligned, allowing all sites to be used in the analyses of that gene. The public domain MEGA ver. 2.1 program (Kumar *et al*., 2001; available at http://www.megasoftware.net/) was used for counting the numbers of transitions and transversions, and for calculating pairwise genetic distances between the lancelets using Kimura's (1980) two-parameter model. Taking gap sites in the 16S rRNA gene into consideration, evolutionary distances in the gene were estimated with pairwise-deletion option of the MEGA program. Phylogenetic relationships among the five lancelets, based on the DNA sequences of the two genes, were respectively inferred with PAUP* 4.0b10 (Swofford, 2001), using maximum-likelihood (abbreviated as ML) method (Felsenstein, 1981) under the HKY 85 model (Hasegawa *et al*., 1985). An exhaustive search was performed to find ML trees for the respective genes. The ratio of transition to transversion (Ts/Tv) was estimated with PAUP* simultaneously with the search for the ML trees. The robustness of each branching point in the gene trees was examined using the bootstrap method (Felsenstein, 1985) with 10,000 replications. In the phylogenetic analysis for the 16S rRNA gene, open sites were treated as missing data.

RESULTS

Sequence divergence among the Pacific and Atlantic lancelets

Table 1 shows large sequence differences between the Pacific and Atlantic species in both the 16S rRNA and COI genes [net difference in the former represented around 150 sites (19.2–20.6%), in the latter around 180 sites (17.9–19.4%)]. The averaged genetic distance between the Pacific and Atlantic species was 0.200 in the 16S rRNA gene and 0.245 in the COI gene (Table 2). The two Pacific lancelets differed significantly from each other in both genes, in contrast to close similarity between the two Atlantic species.

Table 1. Sequence differences (transitions / transversions) between pairs of *Branchiostoma* species and *Epigonichthys lucayanus* (outgroup) in 773 bp of the 16S rRNA gene (above diagonal) and 972 bp of the COI gene (below diagonal). Proportions of transitional and transversional changes in both genes shown in parentheses.

Species	*B. belcheri*	*B. malayanum*	*B. floridae*	*B. lanceolatum*	*E. lucayanus*
B. belcheri		60 / 64	77 / 71	80 / 70	96 / 108
		(.078 / .083)	(.100 / .092)	(.103 / .091)	(.124 / .140)
B. malayanum	101 / 63		79 / 79	81 / 78	108 / 103
	(.104 / .065)		(.102 / .102)	(.105 / .101)	(.140 / .133)
B. floridae	111 / 74	105 / 69		7 / 2	90 / 90
	(.114 / .076)	(.108 / .071)		(.009 / .003)	(.116 / .116)
B. lanceolatum	115 / 74	109 / 69	10 / 0		94 / 87
	(.118 / .076)	(.112 / .071)	(.010 / .000)		(.122 / .113)
E. lucayanus	110 / 89	137 / 90	119 / 87	118 / 87	
	(.113 / .092)	(.141 / .093)	(.122 / .090)	(.121 / .090)	

Table 2. Pairwise genetic distances in *Branchiostoma* species and *Epigonichthys lucayanus* determined from the 16S rRNA (above diagonal) and COI (lower diagonal) genes using Kimura's (1980) two-parameter model, calculated with MEGA ver. 2.1 (Kumar *et al*., 2001). Ts/Tv values, estimated with PAUP*, for the former and latter genes were 1.30 and 1.70, respectively.

Species	*B. belcheri*	*B. malayanum*	*B. floridae*	*B. lanceolatum*	*E. lucayanus*
B. belcheri		.149	.197	.204	.278
B. malayanum	.193		.188	.191	.289
B. floridae	.223	.207		.011	.233
B. lanceolatum	.228	.213	.010		.236
E. lucayanus	.245	.285	.254	.253	

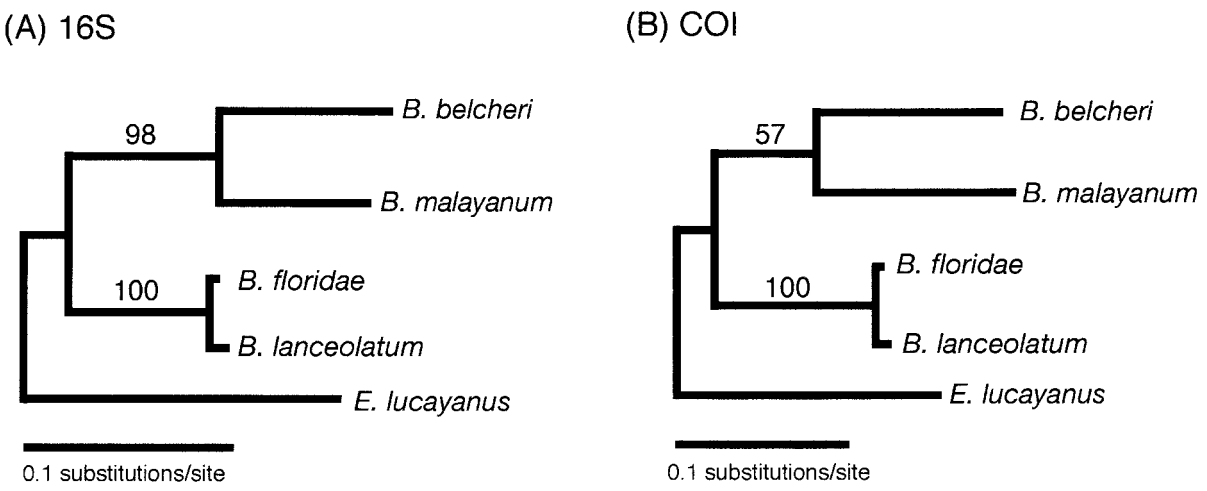

Fig. 2. Maximum-likelihood trees of the 16S rRNA (A) and COI (B) genes based on HYK85 model (Hasegawa *et al.*, 1985) in the genes (estimated parameters for 16S gene: -Ln likelihood=2350, Ts/Tv=1.30; for COI gene: -Ln likelihood=3110, Ts/Tv=1.70). Values on branches of gene trees indicate bootstrap probability with 10,000 replications. Horizontal bar under each tree shows 0.1 substitutions per site.

Phylogenetic relationships

Two ML trees estimated from the 16S rRNA and COI genes showed that the Pacific and Atlantic species pairs were reciprocally clustered into two different clades, supported by high bootstrap probabilities (Fig. 2). Furthermore, the depth of phylogenetic separation between the Pacific and Atlantic lancelets proved to be notable in both gene trees. Contrasting branching patterns in the two gene trees were found between the groups of Pacific and Atlantic species, the phylogenetic splitting between the former being significantly deeper than that between the latter. The interspecific relationship seen in the ML trees was consistent with those found in maximum-parsimony and neighbor-joining trees constructed for the two genes (data not shown).

DISCUSSION

Phylogenetic relationships among the examined lancelets

The present survey appears to be the first providing comparable genetic data, thus allowing an estimation of phylogenetic relationships. The data showed the two Pacific species, *B. belcheri* and *B. malayanum*, to be significantly differentiated from the two Atlantic *B. floridae* and *B. lanceolatum.* At present, 6 and 14 species, respectively, are known to inhabit the Pacific and Atlantic Oceans (Poss and Boschung, 1996). Although it is still uncertain if the other extant lancelets conform invariably to the same Pacific and Atlantic lineages, the present result and our preliminary analysis including another Atlantic *Branchiostoma* species (Nohara *et al.*, unpublished data) suggest a possibility that Pacific and Atlantic *Branchiostoma* have diversified independently, following their phylogenetic separation.

Divergence times have been estimated from molecular data from many organisms, especially vertebrates (*e.g.* Avise, 2000). Whereas the divergence time of lancelets is difficult to estimate directly from fossil data because of the paucity of lancelet fossils, an alternative approach using the "molecular clock" of other animals may help an estimation, in spite of the variability of evolutionary rates among animals. The molecular clock determined from the COI and cytochrome b genes of shark mtDNA, estimated by Andrew *et al.* (1992), appears to be suitable for estimating the divergence time of lancelets since sharks can be regarded as the closest relatives of lancelets among all the animals that have been studied for an evolutionary rate of mtDNA. The evolutionary rate in shark mtDNA may give a reasonable estimate of divergence time, judging from its consistency with those in perciform (Cantatore *et al.*, 1994) and anguilliform (Lin *et al.*, 2001) fish mtDNA. According to Andrew *et al.*'s (1992) estimation, the divergence times within the Pacific species and between the Pacific and Atlantic lancelets were estimated at 97.7 million years ago (abbreviated as Mya) and 112 Mya, respectively (Table 3). These divergence-time estimates may be reasonable, indicating differentiation of the Atlantic ancestral lancelets from the ancestors inhabiting the ancient Pacific as having occurred after the beginning of the formation of the Atlantic, probably following migration from the ancient Pacific to the developing

Table 3. Divergence times (T) between *Branchiostoma* species estimated from Andrew *et al.*'s (1992) molecular clock for cytochrome c oxidase subunit I and cytochrome b genes in shark mtDNA (7.1×10^{-10} transversions/site•year). K indicates corrected proportion of transversions between compared sequences: $K=0.5\bullet\log_e(1/(1-2Q))$, where Q is an observed proportion of transversions. Estimate between Pacific and Atlantic species is the averaged value for four pairs of the species.

Branching point	K	T (Mya)
B. floridae and *B. lanceolatum*	0.0000	–
B. belcheri and *B. malayanum*	0.0694	97.7
Pacific and Atlantic species	0.0796	112

ancient Atlantic via the Tethys seaway. Closure of the Tethyan corridor may have played a crucial part in deep phylogenetic splitting between the Indo-Pacific and Atlantic relatives in some marine organisms as well as the Atlanctic lancelets, as suggested by the previous molecular phylogenetic studies for loliginid squids (Anderson, 2000) and eels (Aoyama *et al*., 2001).

On the other hand, the present results also suggest very recent speciation of the two Atlantic lancelets, although the divergence time could not be calculated owing to the lack of transversions in the COI gene. Possible recent genetic differentiation in species on European and North American sides of the Atlantic Ocean has been demonstrated in fishes including mackerel, *Scomber scombrus* (Scoles *et al*., 1998), bluefish, *Pomatomus saltatrix* (Goodbred and Graves, 1996) and capelin, *Mallotus villosus* (Dodson *et al*., 1991). These findings suggest that a recent geographic event (*e.g*. glacial period) has played some role in bringing about such genetic differentiation.

Disparity between morphological similarity and molecular phylogeny

From the morphological point of view, *B. belcheri* appears to be more closely related to the Atlantic species than to *B. malayanum* (Table 4). For example, mean myotomes in *B. belcheri, B. floridae, B. lanceolatum* and *B. malayanum* number 62–67, 59–60, 58–63 and 52, respectively. However, the phylogenetic affinities indicated by the 16S rRNA and COI gene trees were not coincident with morphological similarity among the *Branchiostoma* species examined because of the apparent reciprocal monophyly of the Pacific and Atlantic species. Clearly, much remains to be learned in the pattern and process of morphological changes in the lancelets.

Table 4. Variation of mean numbers of total myotomes, dorsal fin-chambers and preanal fin-chambers in previously studies for the four species of *Branchiostoma*, based on Table 2 of Poss and Boschung (1996). Mean ranges in *B. malayanum* are not given due to poor data.

Species	Total myotomes	Dorsal fin-chambers	Preanal fin-chambers
B. belcheri	62–67	240–316	58–81
B. malayanum	52	201	53
B. floridae	59–60	286–307	41–47
B. lanceolatum	58–63	212–275	33–62

ACKNOWLEDGEMENTS

We appreciate the assistance of Dr. Yoshio Ojima, Japan Marine Science College, and Mr. Nopadon Kakhai, Kung Krabaen Bay Royal Development Study Centre, Thailand, for collecting specimens of lancelets. This study was supported in part by Grant-in-Aids from the Ministry of Education, Science, Sports and Culture of Japan to TN (#11640694 and #13440253).

REFERENCES

Anderson FE (2000) Phylogeny and historical biogeography of the loliginid squids (Mollusca: Cephalopoda) based on mitochondrial DNA sequence data. Mol Phylogenet Evol 15: 191–214

Andrew PM, Naylor GJP, Palumbi SR (1992) Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. Nature 357: 153–155

Andrews A (1893) An undescribed acraniate: *Asymmetron lucayanum*. Studies from the Biological Laboratory, Johns Hopkins University 5: 213–247

Aoyama J, Nishida M, Tsukamoto K (2001) Molecular phylogeny and evolution of the freshwater eel, genus *Anguilla*. Mol Phylogenet Evol 20: 450–459

Avise JC (2000) Phylogeography: The history and formation of species. Harvard University Press, Cambridge

Boore JL, Daehler LL, Brown, W (1999) Complete sequence, gene arrangement, and genetic code of mitochondrial DNA of the Cephalochordata *Branchiostoma floridae* (Amphioxus). Mol Biol Evol 16: 410–418

Cantatore P, Roberti M, Pesole G, Ludovico A, Milella F, Gadaleta MN, Saccone C (1994) Evolutionary analysis of cytochrome b sequences in some Perciformes: evidence for a slower rate of evolution than in mammals. J Mol Evol 39: 589–597

Dodson JJ, Carscadden JE, Bernatchez L, Colombani F (1991) Relationship between spawning mode and phylogeographic structure in mitochondrial DNA of North Atlantic capelin *Mallotus villosus.* Mar Ecol Prog Ser 76: 103–113

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17: 368–376

Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783–791

Folmer O, Black M, Hoen W, Litz R and Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome oxidase subunit 1 from diverse metazoan invertebrate. Mol Biol Biotechnol 3: 294–299

Gee H (1996) Before the Backbone: Views on the origin of the vertebrates. Chapman & Hall, London

Glenn C, Avise JC (1998) A comparative summary of genetic distances in the vertebrates from mitochondrial cytochrome b gene. Mol Biol Evol 15: 1481–1490

Goodbred CO, Graves JE (1996) Genetic relationship among geographically isolated populations of bluefish (*Pomatomus saltatrix*). Mar Freshwater Res 47: 347–355

Hall BK (1998) Evolutionary developmental biology. 2nd ed, Chapman & Hall, London

Hasegawa M, Kishino H, Yano T (1985) Dating the human-ape split by a molecular clock of mitochondrial DNA. J Mol Evol 22: 160–174

Hubbs C (1922) A list of lancelets of the world with diagnoses of five new species of *Branchiostoma*. Occas Pap Mus Zool Univ Michigan (105): 1–16

Kimura M (1980) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. J Mol Evol 16: 111–120

Kitaura J, Yamamoto G, Nishida M (1998) Genetic variation in populations of diamond shaped squid *Thysanoteuthis rhombus* as examined by mitochondrial sequence analysis. Fish Sci 64: 538–542

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: Molecular Evolutionary Genetics Analysis software. Bioinformatics 17: 1244-1245

Lin Y-S, Poh Y-P, Tzeng C-S (2001) A phylogeny of freshwater eels inferred from mitochondrial genes. Mol Phylogenet Evol 20: 252–261

Minelli A (1993) Biological Systematics. Chapman & Hall, London

Nishida M, Ohkawa T, Iwata Y (1998) Methods of analysis of genetic population structure with mitochondrial DNA markers. Fish Genet Breed Sci 26: 81–100 (in Japanese)

Palumbi S, Martin A, Romano S, McMillan WO, Stice L, Grabowski G (1991) The simple fool's guide to PCR. Dept of Zool and Kewalo Mar Lab, Univ of Hawaii, Honolulu

Poss SG, Boschung HT (1996) Lancelets (Cephalochordata: Branchiostomidae): How many species are valid? Isr J Zool 42: S13–S66

Sambrook J, Russell DW (2001) Preparation and analysis of eukaryotic genomic DNA. In "Molecular Cloning Vol 1. 3rd ed", Cold Harbor Laboratory Press, New York pp 6.4–6.11

Scoles DR, Collette B, Graves JE (1998) Global phylogeography of mackerels of the genus Scomber. Fish Bull 96: 823–842

Shu DG, Morris SC, Zhang X-L (1996) *A Pikaia*-like chordate from the Lower Cambrian of China. Nature 384: 157–158

Swofford DL (2001) PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods) Version 4. Sinauer Associates, Sunderland, Massachusetts

Spruyt N, Delarbre C, Gachelin G, Laudet V (1998) Complete sequence of the amphioxus (*Branchiostoma lanceolatum*) mitochondrial genome: relation to vertebrates. Nucleic Acids Res 26: 3279–3285

Thompson JD, Gibson T, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24: 4876–4882

Webb JE (1956) A note on the lancelets of Singapore, with a description of a new species of *Branchiostoma*. Proc Zool Soc Lond 127: 119–123

(Received May 6, 2003 / Accepted November 5, 2003)

Appendix 1. Accession numbers for sequences determined in the present study.

Species	Region	Accession No.
B. belcheri	16S rRNA	AB105142
	COI anterior part	AB105136
	COI posterior part	AB105137
B. malayanum	16S rRNA	AB105143
	COI anterior part	AB105138
	COI posterior part	AB105139
E. lucayanus	16S rRNA	AB105144
	COI anterior part	AB105140
	COI posterior part	AB105141

Appendix 2. Aligned sequences of 809 bp, including indels, of the 16S rRNA gene for five lancelets. Hyphens and dots indicate a gap and identity to first sequence, respectively. Accession numbers for five sequences of *B. belcheri*, *B. malayanum B. floridae*, *B. lanceolatum* and *E. lucayanus* are AB105142, AB105143, AF098298, Y16474 and AB105144, respectively.

```
                1                                                  50
B. belcheri     TTTAGTATGG TAGTAAATAA CTCTTG--T- AGAGGATCTT AT--GTTAAA
B. malayanum    .......A.. ..T....C.. ..T.-.GG.- .A..A..T.. ..........
B. floridae     ......T... ..T.G..... .....A---- ...A---T.. ..........
B. lanceolatum  ......T... ..T.G..... .....A---- ...A---T.. ..........
E. lucayanus    .....CT... ..T...G... TC...AA-.T G.GA--CA.C ..CT......
               51                                                 100
B. belcheri     ATGCGTAGAT --AGAGAAAA ACTTAGACAC GAG------- TA-AATTTTA
B. malayanum    ........G. TT..-T.--- -....AT..T .T.-TGTAGG ..C...A-..
B. floridae     ........G. T-.ACT..-C T....A-... ---ATGTAAG ..-.......
B. lanceolatum  ........G. T-.ACT..-C T....A-... ---ATGTAAG ..-.......
E. lucayanus    .......... TC.ACTGT.T T.C-.AT-.T T..-TGTGT- ......A...
              101                                                 150
B. belcheri     TGCTT---GA TAGGTATATA CAAGAAAGTA GAACTAT-AA GAGTTAATCT
B. malayanum    A....TAG.. --T...A... .....T.T-- ---T.T.G.. .-........
B. floridae     ----.CTTT.- --TT--G.T .TG..T.A-. ...T-----. T-..G..G..
B. lanceolatum  ----.CTTT.- --TT--G.T .TG..T.A-. ...T-----. T-..G..G..
E. lucayanus    .A--CTAC-- --TT...G.C ......CC-- ---------G .C-..CTGA..
              151                                                 200
B. belcheri     GACATGAGAG TGTGGTTA-A GTAAATTAAT TAGGGAATAG GAACTCGGCA
B. malayanum    .........A ..CTT...A. ..T..A...C ..A.......... ..........
B. floridae     A...G..... .....A..G- ...T.CA... .....GGA.. ..........
B. lanceolatum  A...G..... ...AA..G-. ...T.CA... .....GGA.. ..........
E. lucayanus    ....AA..CA .A.AT.GGG. A.T.T.AT.C ......G-.. ..........
              201                                                 250
B. belcheri     AATCTTAAGC TCGCCTGTTT AACAAAAACA TCGCCTTCAG ATTCAA--TA
B. malayanum    ...A...... .......... .......... .......... G-----AA..
B. floridae     ..CA.G.... .......... .......... .......... T...C-----..
B. lanceolatum  ..CAGG.... .......... .......... .......... T...C-----..
E. lucayanus    ...AA..... ..........C .......... ....T..T.. ----GC--..
              251                                                 300
B. belcheri     TTGGAGGTCT GGTCTGCCCA GTGTAATTAA TTAACGGCCG CGGTATTCTG
B. malayanum    .......... .......... .....G.-.. .......... .....C...
B. floridae     ..AA...... ..........G .........- .........- ...-......
B. lanceolatum  ..AA...... ..........G .........- .......... ..........
E. lucayanus    ..AA.A.... TA.......G .....---.- .......... .....CT..
              301                                                 350
B. belcheri     ACTGTGCAAA GGTAGCATAA TCACTTGCCC TTTAAATGGG GGAATGTATG
B. malayanum    .......... .......... .......... .......... ..........
B. floridae     .......... .......... .......... .......... ..........
B. lanceolatum  .......... .......... .......... .......... ..........
E. lucayanus    .......T.. .......... .......... .......... ..........
              351                                                 400
B. belcheri     AATGATTAGA CGAGGTTTTT ACTGTCTCTT CCTTAT-AAA TTGAGATTAC
B. malayanum    .......C.. ..........G .......T.. .T...GT... ..........
B. floridae     ..C....... .......CG. .....T.C.. .C..A-...C .......T..
B. lanceolatum  .......... .......CG. .....TCC.. .C..A-...C .......T..
E. lucayanus    .......... ...A....AA G.....T.C. ....TAT..- .C......T.
              401                                                 450
B. belcheri     AATACTTGTG AAAATGCGGG TAAGGTAA-T AAAGGACGAG AAGACCCTAT
B. malayanum    .T..TC.... .........A .....---.G ..T.......... .....G..
B. floridae     .G.GTC.... .......... C.T..-CTG. ..T.......... .....G..
B. lanceolatum  .G.GTC.... .......... C.T..-CTG. ..T.......... .....G..
E. lucayanus    .TCG...... .......AA. C.GACACT-C .......... .....G..
```

	451				500
B. belcheri	TGAGCTTTTA	AG-CTAAACT	ATAGTACAGG	TTAAG-TTAA	GTAA---TAA
B. malayanum	....	..-T...TTA	.AG...T.A.	...GC-....	....-A-..-
B. floridae	.......G-.	..A...T.TA	T...CGTCT.	...TAA.A..	A...C--...
B. lanceolatum	.......G-.	..A...C.TA	T...CGTCT.	...GAA.A..	A...--T...
E. lucayanus	....T..C.G	..-.GT...A	.A......TA	....C-.A..	-...CAC..G

	501				550
B. belcheri	-TTTACAATT	AAAACGAGTA	ATTTTTTGTA	GGCTTTTTGG	CTGGGGTGGC
B. malayanum	-....TT..C	....TAT...	T.....AT..	AT........	....
B. floridae	-..-.TC.CC	...-.AT.CG	TAA..AAACG	.T.....-..	....
B. lanceolatum	-..-.TC.CC	...-.AT.CG	TAA..AAACG	.T.....-..	....
E. lucayanus	A.--.A..-C.	T.-GT.CCT.	AAG.AAA-GC	.T....--..	....

	551				600
B. belcheri	AAGCAAAGAT	ATTAAGCTTT	GTTGTAGTAT	ACATT-TTGT	ATTTCTAGAT
B. malayanum	........GA	...G......	.C.AG.A..G	GA...-...G	T.CGGG....
B. floridae	........GA	....	.--A..T..A	.-...-.-AC	....T...GC
B. lanceolatum	........GA	....	.--A..T..A	.--..-.-AC	....T...GC
E. lucayanus	.C.......A	....	CC----...A	.T...C...A	.G..AA...C

	601				650
B. belcheri	AATATGTATC	TATAAAT-AA	TTAAATT--G	ATCCGTTAAA	ATAGAACGAT
B. malayanum	GG...AC...	A..G...T..	......A--.	.....C..G-	T..-.G....
B. floridae	TG.T.A.GC.	A.GG...--.	....-.ACA.	.....C...-	---..G....
B. lanceolatum	TG.T.A.GC.	A.GG...--.	....-.ACA.	.....C...-	---..G....
E. lucayanus	TCCT.AGG..	A..-..CCC-	..C--.ACC.	.C...C..--	C.--.G....

	651				700
B. belcheri	TAAAAGAATA	AGTTACCACA	GGGATAACAG	CGTAATTCTT	TTTGAGAGCT
B. malayanum	...T......	....	....	...T....C.	....
B. floridae	...T..TA..	....	....	....	....
B. lanceolatum	...T..TA..	....	....	......C...	C
E. lucayanus	.C....CC..	.A........	....	...T......	....

	701				750
B. belcheri	CGAATTGACA	AAGGAGTTTG	CGACCTCGAT	GTTGGATCAA	GATTCCTAGC
B. malayanum	.T........	G.........	....	....	.........T
B. floridae	.AT.......	..A.G.....	....	....	........AT
B. lanceolatum	.AT.......	..A.G.....	....	....	........AT
E. lucayanus	.A........	..A.......	....	....	..C....G.T

	751				800
B. belcheri	GGTGTAGCAG	CTGTTACGGG	TTTGCCTGTT	CGGCGATTAA	TATCTTACGT
B. malayanum	....	..A...T...	....	....	A.........
B. floridae	....	..A...T...	....	....	A.........
B. lanceolatum	....	..A...T...	....	....	A.........
E. lucayanus	....C.A.C.	T.ACC.....	....	....	A.........

	801	809
B. belcheri	GATCTGAGT	
B. malayanum	....	
B. floridae	....	
B. lanceolatum	....	
E. lucayanus	....	