Fisheries Science 64(3), 397-402 (1998)

Intron Length Variation Observed in the Creatine Kinase and Ribosomal Protein Genes of the Swordfish *Xiphias gladius*

Seinen Chow*[1,†] and Haruko Takeyama*[2]

*[1]Pelagic Fish Resources Division, National Research Institute of Far Seas Fisheries, Orido, Shimizu, Shizuoka 424-8633, Japan

*[2]Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184-8588, Japan

(Received November 4, 1997)

Introns may accumulate much higher genetic variation than exons. Universal primers were designed from the conservative nucleotide sequences of exons to amplify the flanking intron. Length variations in the S7 ribosomal protein (RP) gene intron 1 and mitochondrial creatine kinase (CK) gene intron 6 of the swordfish *Xiphias gladius* were found. Single or two banded fragment patterns in each individual were observed by agarose gel electrophoresis. Nucleotide sequence analysis revealed that highly polymorphic fragment patterns observed in the RP gene intron 1 were due to different numbers of a TG repeat (microsatellite). The length of the CK gene intron 6 was dimorphic, in which presence or absence of a 24 bp block was responsible for longer or shorter introns. Additional minor nucleotide insertion/deletions were observed independent of the RP microsatellite and the CK 24 bp block regions. The results of this investigation indicate that introns may be good sources of intraspecific genetic variation for population genetic studies and that the same set of primers can be used to amplify homologous intron regions even among distant species. Further, the conserved exon primed PCR strategy may be useful to prevent appearance of priming site polymorphism (null allele).

Key words: swordfish, intron length polymorphism, fish population genetics

Restriction fragment length polymorphism (RFLP) and nucleotide sequence analyses of mitochondrial DNA (mtDNA) are common techniques for detecting genetic variation within and between fish populations. This is primarily because mtDNA is homoplasmic and present in high copy number per cell, making it easy to purify and analyze. It is believed that the maternally inherited mtDNA is likely to show more differences among populations than nuclear DNA because of its smaller effective population size.[1] MtDNA analyses have been applied to the study of genetic stock structures of highly migratory cosmopolitan fish species, and heterogeneity in haplotype frequencies between samples from different ocean basins was observed in several species such as northern bluefin tuna *Thunnus thynnus*,[2,3] albacore *T. alalunga*[4] *and* swordfish *Xiphias gladius*.[5-8] In the albacore and swordfish, mtDNA genotype frequency distributions of the samples collected in the South Atlantic Ocean were intermediate between those of the North Atlantic and Indo-Pacific.[4,7] However, mtDNA analyses do not have the power to cope with population mixture nor can they verify whether a local sample is from a single Mendelian population.

A more detailed population genetic study requires the investigation of genetic polymorphism in the nuclear DNA. Several DNA-level assays on nuclear genome have been developed, and by incorporating the polymerase chain reaction (PCR), the detection of microsatellites,[9] randomly amplified polymorphic DNAs (RAPDs),[10] and restriction site polymorphisms (RSPs) at anonymous single copy sequences (scnDNA)[11] are becoming conventional. Yet, all these assays based on PCR amplification may have inherent problems leading to genotype artifacts. Especially when anonymous non-coding sequences are targeted, polymorphism at PCR priming sites may result in differential amplification of alleles.[12] Intron-targeted PCR was introduced to find neutral nuclear DNA polymorphism in marine animals.[13-17] Since intron-targeted PCR usually employs primers designed from the conservative exon sequences, polymorphism at priming sites is likely to be rare. Endonuclease digestion of the amplified fragment is usually applied to find codominant alleles,[14,15,17] although in some cases alleles may be distinguished by length difference of the amplified fragments.[13]

We have investigated introns as a source of genetic polymorphism for a fish population genetic study. For amplifying introns, universal primers designed by Palumbi *et al.*[18] targeting the 6th intron of creatine kinase gene were employed. Also, we have found conserved exon sequences in the S7 ribosomal protein gene between fish and amphibian, which allowed us to design universal primers to amplify the flanking intron. In this paper, we report intron length polymorphisms for the creatine kinase and ribosomal protein genes of the swordfish *Xiphias gladius*.

Materials and Methods

Fish Samples and DNA Preparation

Swordfish *Xiphias gladius* samples, derived from our

† Address of corresponding author.

laboratory collection, were caught in six oceanic basins (Mediterranean Sea, North Atlantic, South Atlantic, Indian Ocean, Northwest Pacific and Southeast Pacific). Detailed information of these samples is available elsewhere.[7] Crude nucleic acids were extracted from fresh, frozen or ethanol preserved muscle using the standard phenol/chloroform method.

Design of Primers and PCR Amplification

A pair of primers to amplify the 1st intron of the S7 ribosomal protein (RP) gene was designed from the conserved sequences of exons 1 and 2 of puffer fish *Fugu rubripes*[19] and frog *Xenopus laevis*.[20] The sequences of RPEX1F and RPEX2R are 5′-TGGCCTCTTCCTTGGCCGTC-3′ and 5′-AACTCGTCTGGCTTTTCGCC-3′, respectively. Primer sequences to amplify the flanking region of exons 6 and 7 containing the entire intron 6 of the creatine kinase gene (CK) were from Palumbi *et al.*[18] The sequences of the primers designated as CK6F and CK7R are 5′-GACCACCTCCGAGTCATCTC-3′ and 5′-CAGGTGCTCGTTCCACATGA-3′, respectively. The PCR reaction mixture contained 0.2 U *Taq* DNA polymerase (Perkin-Elmer Cetus, Norfolk, USA), 0.2 mM of each dNTP, 1.0 μl 10 × buffer (Perkin-Elmer Cetus, Norfolk, USA), 2 mM $MgCl_2$, 10 pmol of each primer, and 10 to 50 ng template DNA. The volume of the reaction mixture was made up to 10 μl with sterile distilled water. PCR was carried out with an initial denaturation at 95°C for 1.5 min, followed by 30 cycles of amplification (denaturation at 95°C for 1 min, annealing at 60°C for 1 min, and extension at 72°C for two min, with a final extension at 72°C for 10 min). PCR products were electrophoresed through a 2% agarose gel in TBE buffer (50 mM Tris, 1 mM EDTA, and 48.5 mM boric acid) to confirm amplification.

Cloning and Nucleotide Sequence Analysis of the PCR Products

PCR products were directly subjected to cloning without further purification. Cloning was carried out using a TA cloning vector, PUCII (3.9 kbp) (Invitrogen, San Diego, USA). Plasmids isolated from *Escherichia coli* by a general method[21] were digested with *Eco* RI, and those containing target PCR products were screened. The nucleotide sequences of the inserted PCR products were determined by automated sequencing (Applied Biosystems model 373A, USA) using a *Taq* dye terminator cycle sequencing kit (Perkin-Elmer Cetus, Norfolk, USA) with M13 reverse and M13 (-20) forward primers, those used for amplification and internal primers designed from the determined nucleotide sequences. The Clustal V program[22] was used to align the sequences, and the sequences were realigned by eye.

Results

Nucleotide Sequence Analysis of the Cloned Fragments Amplified by RP Primers

Electrophoretic patterns of the amplified RP fragments are shown in Fig. 1. Highly polymorphic fragment patterns were observed, but all individuals showed either one or two banded phenotypes suggesting that these length variants were in single locus fashion. Four swordfish individuals (one from Japan, two from Peru and one from Brazil) showing two banded phenotypes were selected for cloning, and the nucleotide sequences of the two fragments (two clones) in each individual were determined. The nucleotide sequence of one clone (Brazil 16D) is shown in Fig. 2. Homology investigation between this sequence and *Fugu* and *Xenopus* S7 ribosomal protein genes[19,20] indicated that the amplified fragments consisted of the entire exon 1 and intron 1, and the partial sequence of exon 2. These exons were observed to be highly conserved among the three species. Although the intron was poorly conserved, there remained several common motifs between the swordfish and *Fugu*. A dinucleotide (TG) repeat (microsatellite) was observed at site 667 in the swordfish. Alignment of sequences of eight clones obtained from four swordfish individuals indicated the sequences to be almost identical to one another except for the dinucleotide repeat number (Fig. 2, double underlined) and a few nucleotide insertion/deletions and substitutions in the intron (Fig. 3). Variation in the number of the TG repeat appears to be responsible for the fragment length polymorphism, although minor insertion/deletions and several nucleotide substitutions were observed independent of the TG repeat unit. Deletion of three nucleotides (ATT) was observed at 5′ end of the repeat unit in one clone (Peru 5B). An internal primer (RPSWF) was designed from the conserved central region (see Fig. 2), and several samples were amplified using this internal primer and reverse primer RPEX2R. The amplified fragment patterns were identical, except for the size, with those amplified using RPEX1F and RPEX2R primers, indicating that the length variants were of a single gene locus.

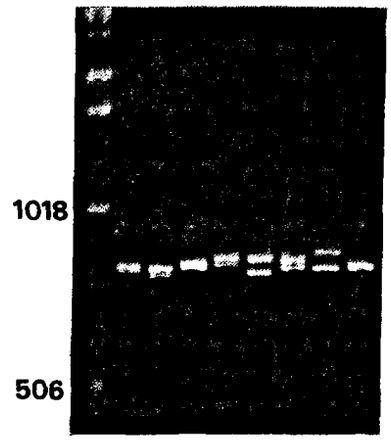

Fig. 1. An agarose gel electrophoresis profile of the amplified S7 ribosomal protein gene (RP) intron 1 of the swordfish (*Xiphias gladius*).

At the left is the molecular marker (GIBCO BRL 1 kb ladder). Highly polymorphic but single or double-banded phenotypes were observed in all individuals.

Nucleotide Sequence Analysis of the Cloned Fragments Amplified by CK Primers

Electrophoretic patterns of the PCR products amplified using the CK6F and CK7R primers are shown in Fig. 4. One or two banded phenotypes were observed in the 600-630 bp size range, and these were tentatively considered as homozygotes and heterozygotes, respectively. Amplifica-

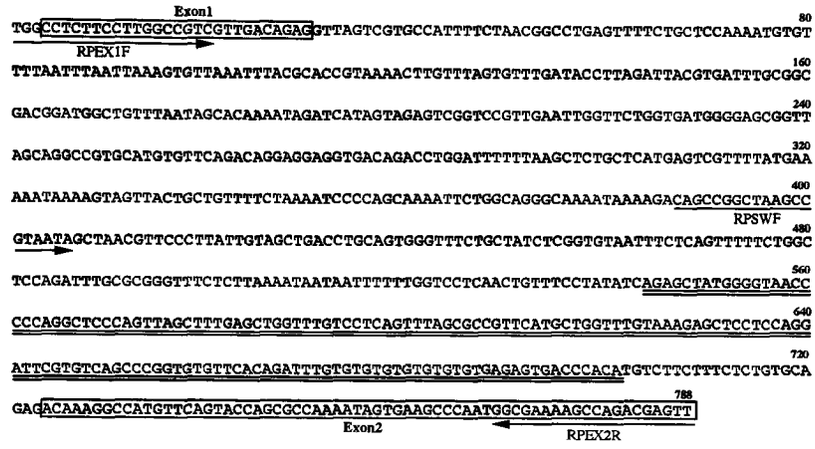

Fig. 2. Nucleotide sequence of a clone (Brazil 16D) containing amplified RP intron 1.
The two exons are boxed. Primers and directions are shown by underlines with arrowheads. Double underlined region containing dinucleotide (TG) repeats was picked up for multiple sequence alignment.

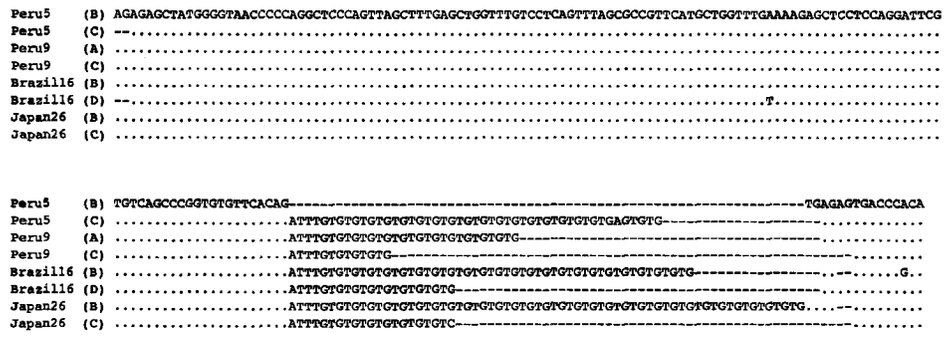

Fig. 3. Aligned partial DNA sequences of RP intron (double underlined region in Fig. 2) of eight clones from four individuals.
The far-left columns indicate locality, individual number and clone (alphabet). The same nucleotides with those of the top reference sequence are shown by dots, while deletions are denoted by dashes. Nucleotide sequences of the entire insert are available in GenBank under the accession numbers of U95515 to U95521.

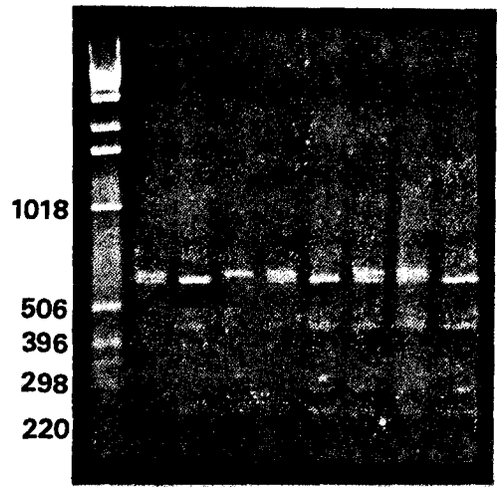

Fig. 4. An agarose gel electrophoresis profile of the amplified mitochondrial creatine kinase gene (CK) intron 6 of the swordfish (*Xiphias gladius*).
At the left is the molecular marker (GIBCO BRL 1 kb ladder). Single or double-banded phenotypes are observed in the 600–630 bp range. Several other fragments are seen with less than 450 bp.

tion of several other fragments less than 450 bp was also observed. Two individuals representing two banded phenotypes in the 600–630 bp size range were selected from each of six localities for cloning these two fragments. The entire nucleotide sequence of one clone (Japan 1A) is shown in Fig. 5. Homology investigation between the swordfish sequence and mouse *Mus musculus* mitochondrial creatine kinase gene[23] indicated that the clone consisted of partial sequences of exons 6 and 7 flanking intron 6. Exons were highly conserved between these species, while very poor homology was observed between the introns. Sequence alignment among twenty four clones revealed a number of nucleotide insertion/deletions. Partial alignments of two regions (Fig. 5, A and B, double underlined) containing insertion/deletions are shown in Fig. 6. Presence or absence of a 24 bp nucleotide block were observed in region A, which appeared to be responsible for the length difference between the longer and shorter introns. A number of nucleotide insertion/deletions were observed in region B, which yielded more size variants with additional 1 to 4 bp differences in each size group. Internal primers (see Fig. 5, CKSWF and CKSWR) were designed to cover these regions. Using these internal primers, PCR amplification was performed for the original template DNA or for the PCR products amplified using CK6F and CK7R primers (nested PCR). The same individuals shown in Fig. 4 were used, and each showed the same pattern as in the earlier ex-

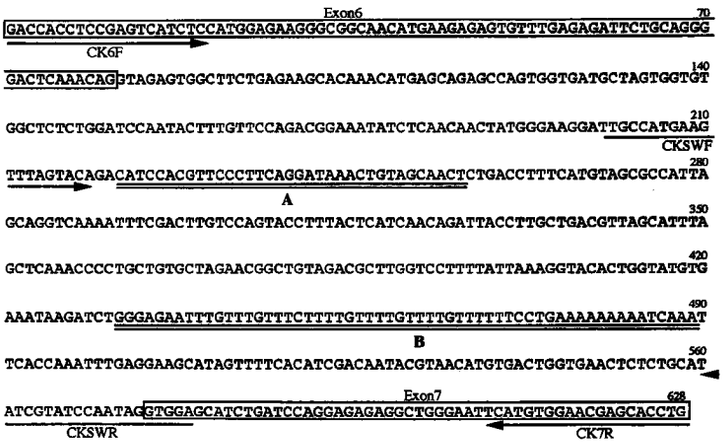

Fig. 5. Nucleotide sequence of a clone (Japan 1A) containing the amplified CK intron 6.
The two exons are boxed. Primers and directions are shown by underlines with arrowheads. Two regions (A and B, double underlined) containing nucleotide insertion/deletions were picked up for multiple sequence alignment.

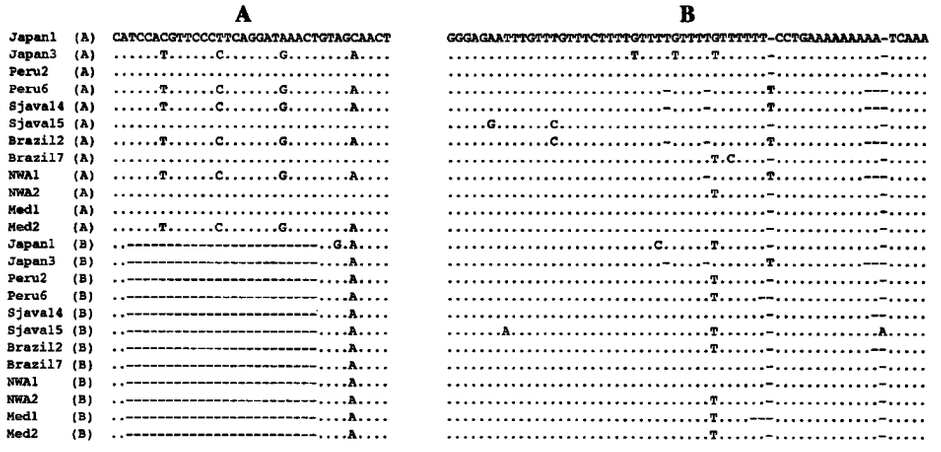

Fig. 6. Aligned partial DNA sequences (regions A and B in Fig. 5) of the CK intron of twenty-four clones from twelve individuals.
The far-left columns indicate locality, individual number and clone (alphabet). The same nucleotides with those of the top reference sequence are shown by dots, while deletions are denoted by dashes. Nucleotide sequences of the entire insert of clones Brazil7, Japan1, Med1, NWA1, Peru6, Sjava15 are available in GenBank under the accession numbers of U75528, 75529, 76547, 76548, U95522 to U95529.

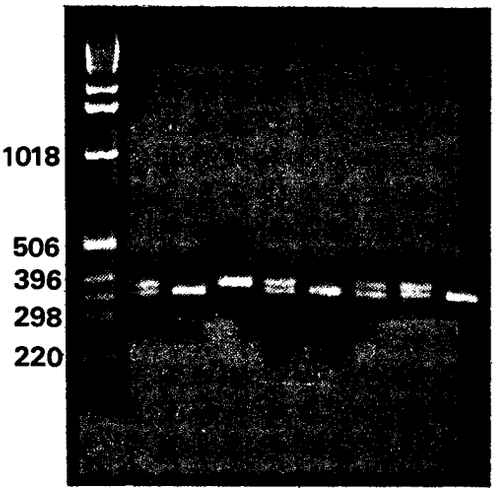

Fig. 7. An agarose gel electrophoresis profile of the CK intron amplified using internal primers (CKSWF and CKSWR, see Fig. 5).
At the left is the molecular marker (GIBCO BRL 1 kb ladder). When nested PCR was performed, 1/100 to 1/1000 dilution of the 1st PCR products were used. The same individuals shown in Fig. 4 were used, and identical patterns were obtained.

periment, but with smaller fragment sizes, as expected (Fig. 7). Further, no amplification of extra fragments was observed, indicating that the length variants observed were from a single gene locus.

Discussion

In the frog *Xenopus laevis*, six introns of the S7 ribosomal protein gene encode small nucleolar RNA (snoRNA U17).[20] Cecconi *et al.*[19] found that four of the six introns of the *Fugu* homolog shared several conserved motifs with those of *Xenopus*, and they appeared to maintain the conserved shape of the snoRNA secondary structure. Cecconi *et al.*[19] suggested that the first two introns in the *Fugu* S7 ribosomal protein gene were pseudogenes, because no snoRNA secondary structure was observed and the rRNA complementary sequences (rRCS) were absent. No rRCS was observed in the swordfish intron 1, and the computer generated secondary structure did not show any matches with those of *Fugu* and *Xenopus*. Hence, it is very likely that intron 1 of the swordfish S7 ribosomal protein

gene is a diverged pseudocopy of snoRNA U17 as in *Fugu* and that the pseudogene might have allowed the microsatellite expansion. Microsatellites in introns are not unusual. There have been a number of such reports using the length variation as a genetic marker.[24-26] Usually fragment length difference in a microsatellite locus is assumed to result from differences in the number of repeat units. As demonstrated in the present study, however, minor nucleotide insertion/deletions may exist independent of the repeat unit. Thus, fragment analysis of microsatellite loci should proceed with caution.

PCR-based nuclear DNA assays for detecting genetic polymorphism may be accompanied by two main problems which may lead to erroneous conclusions in studies of population genetics. One is amplification of multiple paralogous loci, resulting in heterozygote excess. The other is differential amplification of alleles within a locus due to polymorphism at a PCR priming site, resulting in homozygote excess. Although it is necessary to confirm whether the PCR-based variants are inherited in a Mendelian fashion, as demonstrated by laboratory crosses of mussels,[27] that is usually very difficult for most natural populations of marine fishes. Amplification of multiple paralogous loci can be largely circumvented by increasing reaction stringency.[28] Reaction stringency used in this study may be high enough to eliminate this factor. However, fragment amplification from multiple paralogous loci was obvious when universal CK primers were used, and our preliminary PCR attempts for other species such as the northern bluefin tuna *Thunnus thynnus*, southern bluefin tuna *T. maccoyii*, blue marlin *Makaira mazara*, flounder *Paralichthys olivaceus* and crucian carp *Carassius auratus* also amplified multiple fragments even at high stringency (annealing temperature of 60–62°C). Fortunately, the fragments from the paralogous loci appeared not to impede genotype scoring in the swordfish. Identical fragment patterns obtained by both universal exon primers and internal primers are also indicative of fragment length variation within a single locus. Agreement of population genotype frequencies with Hardy-Weinberg equilibrium will be another criterion for the Mendelian behavior of the nuclear gene markers found in this study. The strategy employing highly conserved exon sequences to design the semi-universal PCR primers also considerably reduces the risk for differential amplification of alleles. Furthermore, the conserved exon primers have the additional advantage of amplifying homologous gene regions among distant fish species, as described above.

Optimum fragment size for analysis appears to be different for the variant type of the target sequence. Restriction site polymorphisms (RSPs) may be more frequently found in longer fragments, while length difference may be more accurately estimated in shorter fragments. The exon-primed intron targeted PCR strategy employed in the present study might be more suited for RSP investigation, because the amplified fragment is usually larger than the optimum size for fragment size analysis. When intron length variants are found, subsequent design of internal primers closer to insertion/deletion or microsatellite regions to amplify a shorter fragment may be profitable for accurate size determination.[28] However, a dilemma is that the internal primers within an intron sequence may again be a cause of priming site polymorphism. To avoid this, it is necessary to investigate the nucleotide sequence variation in the intron in order to design conservative internal primers. Or, we might have to wait development of novel techniques to accurately determine the size of relatively long fragments.

Acknowledgments The authors want to thank Dr. R. D. Ward (CSIRO Division of Fisheries, Tasmania, Australia) for reviewing the manuscript.

References

1) M. Nei and F. Tajima: DNA polymorphism detectable by restriction endonucleases. *Genetics*, **97**, 145–163 (1981).
2) S. Chow and S. Inoue: Intra- and interspecific restriction fragment length polymorphism in mitochondrial genes of *Thunnus* tuna species. *Bull. Natl. Res. Inst. Far Seas Fish.*, **30**, 207–225 (1993).
3) S. Chow and H. Kishino: Phylogenetic relationships between tuna species of the genus *Thunnus* (Scombridae: Teleostei): Inconsistent implications from morphology, nuclear and mitochondrial genomes. *J. Mol. Evol.*, **41**, 741–748 (1995).
4) S. Chow and H. Ushiama: Global population structure of albacore (*Thunnus alalunga*) inferred by RFLP analysis of the mitochondrial ATPase gene. *Mar. Biol.*, **123**, 39–45 (1995).
5) J. R. Alvarado Bremer, A. J. Baker, and J. Mejuto: Mitochondrial DNA control region sequences indicate extensive mixing of swordfish (*Xiphias gladius*) populations in the Atlantic Ocean. *Can. J. Fish. Aquat. Sci.*, **52**, 1720–1732 (1995).
6) J. R. Alvarado Bremer, J. Mejuto, T. W. Greig, and B. Ely: Global population structure of the swordfish (*Xiphias gladius* L.) as revealed by analysis of the mitochondrial DNA control region. *J. Mar. Biol. Ecol.*, **197**, 295–310 (1996).
7) S. Chow, H. Okamoto, Y. Uozumi, Y. Takeuchi, and H. Takeyama: Genetic stock structure of the swordfish (*Xiphias gladius*) inferred by PCR-RFLP analysis of the mitochondrial DNA control region. *Mar. Biol.*, **127**, 359–367 (1997).
8) P. E. Rosel and B. A. Block: Mitochondrial control region variability and global population structure in the swordfish, *Xiphias gladius*. *Mar. Biol.*, **125**, 11–22 (1996).
9) J. L. Weber and P. E. May: Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.*, **44**, 388–396 (1989).
10) J. G. K. Williams, A. R. Kubelik, K. J. Livak, J. A. Rafalski, and S. V. Tingey: DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.*, **18**, 6531–6535 (1991).
11) S. A. Karl and J. C. Avise: PCR-based assays of Mendelian polymorphisms from anonymous single-copy nuclear DNA: techniques and applications for population genetics. *Mol. Biol. Evol.*, **10**, 342–361 (1993).
12) J. M. Pemberton, J. Slate, D. R. Bancroft, and J. A. Barrett: Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Mol. Ecol.*, **4**, 249–252 (1995).
13) H. B. S. M. Corte-Real, D. R. Dixon, and P. W. H. Holland: Intron-targeted PCR: a new approach to survey neutral DNA polymorphism in bivalve populations. *Mar. Biol.*, **120**, 407–413 (1994).
14) R. Gross, P. Schlee, H. Stein, and O. Rottmann: Detection of allelic variation within the growth hormone gene in common bream using heteroduplex analysis. *J. Fish Biol.*, **48**, 1283–1287 (1996).
15) S. A. Karl, B. W. Bowen, and J. C. Avise: Global population genetic structure and male-mediated gene flow in the green turtle (*Chelonia mydas*): RFLP analyses of anonymous nuclear loci. *Genetics*, **131**, 163–173 (1992).
16) E. P. Lessa: Rapid surveying of DNA sequence variation in natural populations. *Mol. Biol. Evol.*, **9**, 323–330 (1992).
17) P. Schlee, H. Fuchs, J. Blusch, T. Werner, O. Rottmann, and H. Stein: Genetic polymorphism in the intron of the growth hormone gene of the bleak. *J. Fish Biol.*, **48**, 1275–1277 (1996).
18) S. Palumbi, A. Martin, S. Romano, W. O. McMillan, L. Stice, and G. Grabowski: The Simple Fool's Guide to PCR, Version 2.0, Uni-

versity of Hawaii, Honolulu, 1991, pp. 1–23.
19) F. Cecconi, C. Crosio, P. Mariottini, G. Cesareni, M. Giorgi, S. Brenner, and F. Amaldi: A functional role for some *Fugu* introns larger than the typical short ones: the example of the gene coding for ribosomal protein S7 and snoRNA U17. *Nucleic Acids Res.*, **24**, 3167–3172 (1996).
20) P. Mariottini, C. Bagni, A. Francesconi, F. Cecconi, M. J. Serra, Q. M. Chen, F. Loreni, F. Annesi, and F. Amaldi: Sequence of the gene coding for ribosomal protein S8 of *Xenopus laevis*. *Gene*, **132**, 255–260 (1993).
21) J. Sambrook, E. F. Fritsch, and T. Maniatis: Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, chapter 1, Cold Spring Harbor, New York, 1989, pp. 21–52.
22) D. G. Higgins, A. J. Bleasby, and R. Fuchs: CLUSTAL V: Improved software for multiple sequence alignment. *CABIOS*, **8**, 189–191 (1992).
23) K. Steeghs, W. Peters, M. Bruckwilder, H. Croes, D. Van Alewijk, and B. Wieringa: Mouse ubiquitous mitochondrial creatine kinase: gene organization and consequences from inactivation in mouse embryonic stem cells. *DNA Cell Biol.*, **14**, 539–553 (1995).
24) A. Catto, A. M. Carter, J. H. Barrett, M. Stickland, J. Bamford, J. A. Davies, and P. J. Grant: Angiotensin-converting enzyme insertion/deletion polymorphism and cerebrovascular disease. *Stroke*, **27**, 435–440 (1996).
25) S. Kochling, J. T. den Dunnen, B. Dworniczak, and J. Horst: Two polymorphic dinucleotide repeats in intron 44 of the dystrophin gene. *Hum. Genet.*, **95**, 475–477 (1995).
26) B. K. Saha: Typing of murine major histocompatibility complex with a microsatellite in the class II Eb gene. *J. Immunol. Methods*, **194**, 77–83 (1996).
27) H. B. S. M. Corte-Real, P. W. H. Holland, and D. R. Dixon: Inheritance of a nuclear DNA polymorphism assayed in single bivalve larvae. *Mar. Biol.*, **120**, 415–420 (1994).
28) M. P. Hare, S. A. Karl, and J. C. Avise: Anonymous nuclear DNA markers in the American oyster and their implications for the heterozygote deficiency phenomenon in marine bivalves. *Mol. Biol. Evol.*, **13**, 334–345 (1996).