A new PCR method based on single nucleotide differences for identification of tuna species in fresh and processed fish products

Master's Thesis

A NEW PCR METHOD BASED ON SINGLE NUCLEOTIDE DIFFERENCES FOR IDENTIFICATION OF TUNA SPECIES IN FRESH AND PROCESSED FISH PRODUCTS

September 2011

Graduate School of Marine Science and Technology
Tokyo University of Marine Science and Technology
Master's Course of Food Science and Technology

César Martínez-Ibarra

Table of contents

Background

Tuna is a major species in fisheries industries, being one of the most traded fish worldwide and accounting for a significant percentage of the global fish catches every year (around 8%, according to estimations by the FAO). There are 8 main species classified into the genus *Thunnus*: *Thunnus thynnus* (Atlantic bluefin tuna), *T. maccoyii* (southern bluefin tuna), *T. orientalis* (Pacific bluefin tuna), *T. alalunga* (albacore), *T. obesus* (bigeye tuna), *T. albacares* (yellowfin tuna), *T. tonggol* (longtail tuna) and *T. atlanticus* (blackfin tuna). With the exception of *T. atlanticus*, the smallest species, which lives in limited areas of the Western Atlantic, all *Thunnus* species are harvested at industrial level. In addition to the authentic *Thunnus* species, *Katsuwonus pelamis* (skipjack tuna), a closely related species, is regarded as commercially relevant, thus being included in the so-called "major tuna species group". This group represents almost 85% of total "tuna" catches (almost 5 million tons in 2009, according to data by the FAO). Tuna species are divided in two major groups, following commercial considerations: Bluefin tunas (including *T. thynnus*, *T. orientalis*, *T. maccoyii* and *T. obesus*) and yellowfin tunas (including *T. albacares*, *T. tonggol*, *T. alalunga* and *K. pelamis*). These categories respond only to the relative market value and main usage of each species and are not based in stringent taxonomical considerations.

Bluefin tunas are big species, surpassing 150 cm of length in adulthood, living in warm to moderately cold waters of all the oceans in the world, and relatively rich in fat. The main destination of these species is the Japanese *sashimi* and *sushi* market. Among bluefin tunas, *Thunnus obesus* represents about 7% of all main tuna catches, while the other bluefin species represent less than 3%. Nevertheless, bluefin tuna are considered major species because of their

high prices, which have been increasing steadily due to a growing demand. Populations of bluefin tunas are naturally small and have been continuously decreasing due to overfishing in the last decades (Safina & Klinger, 2008, Cascorbi et al., 2004). Presently, international organisms (such as the International Commission for Conservation of Atlantic Tuna, ICCAT and others) have been struggling trying to convince governments and fisheries to regulate catch and trade volumes of bluefin tuna, in an effort to stabilize the decreasing populations and ensure the future sustainability of this industry. One of the issues concerning the rational management of bluefin tuna is the need for a method to differentiate them accurately. Although adult specimens of *Thunnus obesus* and *T. maccoyii* are somehow morphologically distinct, *T. orientalis* and *T. thynnus* cannot be effectively differentiated by external appearance, which becomes critical, since these two species are under different environmental pressure. Even more, upon removal of the external characteristics, it becomes virtually impossible to distinguish accurately bluefin tunas (and also some yellowfin tuna species) from each other. Thus, tuna loins and filets are easy target of substitution for cheaper, low-quality tuna or even for other non-tuna species. This kind of malpractice, which initiated among wholesalers and retailers, is also common in sushi and sashimi restaurants of Japan and it is expanding rapidly to Western countries, along with the increasing popularity of this Japanese food (Lowenstein et al., 2009).

Yellowfin tunas are smaller than bluefin tunas and live primarily in warm or tropical oceans. Their populations are also larger and thus, the concern of the sustainability of yellowfin fishing is less serious. Yellowfin tunas account for about 90% of the major tuna catches and their main use is in canned tuna for human and pet consumption. Canned tuna has become a very appreciated food commodity in modern society, since it is a ready-to-eat product; it is a source of high-quality protein and it can be stored for long terms. In later years, canned tuna has gained even more

appraisal as a low-cholesterol source of essential fatty acids. Because of these reasons, it is easy to understand that tuna canning industry is not immune to the malpractices that pervade raw or frozen tuna trade. Due to its inherent characteristics, species identification in canned tuna is even more difficult to accomplish than in raw tissue. Since each country and region has a different legislation regarding which species (and in which proportion) can be used, canned tuna becomes target of speculation and of partial or total substitution of its components. Also, canned tuna is prone to unintended mislabeling. This kind of practices derives in a negative impact on the price and quality of the product as well as on the image and trust of the consumers and thus, they are considered a major issue in the industry (Jacquet & Pauly, 2008). Even more, in recent years, replacement of tuna with some cheaper species has gone beyond being an economic fraud, becoming a danger for consumers: In the United States, there have been reports of intoxications due to consumption of escolar (*Lepidocybium flavorunneum*), a species used as a cheap substitute of white tuna, *T. alalunga*, and that contains waxy esters (gempylotoxin) that cause steatorrhea and other digestive symptoms (Lowenstein et al., 2009)

Accurate identification of species in fresh and frozen as well as in cooked and canned tuna products is essential to ensure fair trade as well as in response to the environmental concerns. As mentioned before, not only the use of morphological features is prone to mistakes and inaccuracies but also it is a method that renders ineffective even after minor processing, where these features are removed. Hence, in order to perform an accurate identification, it is necessary to use physicochemical or biochemical analyses. Although there are a number of methods based on protein analysis for identification of animal species, most of the identification methods developed so far are based invariably on DNA analyses, since DNA provides higher accuracy and resolution, given the natural specificity of nucleic acids. Presently, the amount of literature reporting DNA analysis

methods for identification of tuna species is significant (Unseld et al., 1995; Quinteiro et al, 1998; Mackie et al., 1999; Rehbein et al., 1999; Terol et al., 2002). Nevertheless, most of these methods have structural shortcomings that have precluded them to be readily adopted in practice. The following are some of those disadvantages: 1) Identification of only a limited number of species; 2) identification limited to fresh tissue or on samples not subjected to severe heat processing; 3) use of highly specialized equipment or materials; 4) time- or resources-consuming; 5) complex methods requiring skilled or trained operators. Currently, there is not a single analytical method that identifies all major tuna species and that can be used effectively even in severely heat-treated food products.

Even more, as mentioned previously, in the case of canned tuna, not only identification but also quantification is necessary, in order to detect partial substitutions. Some of the above methods, namely real-time PCR techniques, have been also developed as quantitative methods but the high cost of these techniques avoids their adoption as routine assays.

Thus, the present study aimed to develop a DNA-based method for accurate identification of all main tuna species, that is not only universal, i.e., applicable to a variety of food products regardless their processing, but also based in readily available technology, making it much feasible than other methods developed so far. The result is the method hereby introduced, based in allele-specific PCR amplification, a technique that can be accomplished with easiness and speed and does not require particularly specialized equipment. This method was successfully applied to fresh and frozen products (unheated samples) as well as to canned tuna and other food products (heat-processed samples). In addition to the identification method using allele-specific PCR, a quantification method using quantitative PCR was also designed and tested on mixtures of canned tuna, with accurate and reproducible results.

This research has been presented in the following academic conferences:

- 2010 Fall Academic Conference of the Japanese Society of Fisheries. Kyoto, Japan
- 2011 Genomic in Aquaculture International Symposium. Heraklion, Greece.

Chapter 1. Determination of the complete mitochondrial DNA sequences of *Thunnus spp*., *Katsuwonus pelamis* and *Gasterochisma melampus*

1.1 Introduction

A large amount of the existing methods for identification not only of tuna or fish but of other animals are based on mitochondrial DNA (mtDNA) (Lockley & Bardsley, 2000), although methods based on nuclear DNA (Liu & Cordes, 2004; Chow et al., 2006) as well as in other non-nuclear DNA have also been reported (Appleyard et al., 2001; Carlsson et al., 2004). This is because mtDNA offers some key advantages compared to nuclear DNA: i) Shorter and more simple structure, therefore easier to analyze; ii) facilitated DNA extraction, as a typical animal cell encloses hundreds of copies of mtDNA, compared to only one copy of each nuclear chromosome; iii) mtDNA evolves about 10 times faster than nuclear DNA, showing a higher rate of interspecies variability, thus enabling discrimination even between closely related species. The reason of the fast evolutionary rate of the mtDNA is thought to be due, at least partially, to its increased exposure to mutagenic oxygen radical species generated in the mitochondrial matrix and to a reduced effective population size, due to its maternal inheritance (Broughton & Renau, 2006). The sequencing and characterization of partial and complete mtDNA sequences of a large number of species has been increasing in recent years, as a result of efforts such as the Barcode For Life Initiative. However, there are several species whose mtDNA genome has been determined only partially or even not studied at all. In the case of *Thunnus* species, before this study, the complete mtDNA sequence had been established only for three species (*T. thynnus*, *T. orientalis* and *K. pelamis*) (Manchado et al., 2004; Broughton & Renau, 2006, Takashima et al., 2006). Thus, as a first objective, the complete mtDNA sequences of 9 fish species were investigated: *T. thynnus*, *T. maccoyii*, *T. orientalis*, *T.*

alalunga, T. obesus, T. albacares, T. tonggol, Katsuwonus pelamis and *Gasterochisma melampus*. The latter species, known in English as butterfly tuna or butterfly mackerel, is the only member of the genus *Gasterochisma* and of the subfamily Gasterochismatinae, related to tuna. This species lives roughly in the same area than *T. maccoyii* and both species are often caught together (Warashina, 1972). Unlike other species of the Scombridae family, to which tuna belongs, there were no data about the mtDNA of *G. melampus* and thus, it was necessary to include this species in this part of the study in addition to the main tuna species. For other relevant species of the Scombridae family that do not belong to the genus *Thunnus* or *Katsuwonus*, the complete mtDNA sequences have already been reported and thus, it was not necessary to study them *de novo*.

1.2 Materials and Methods

1.2.1 Samples and materials

Muscle samples of *Thunnus thynnus, T. maccoyii, T. orientalis, T. alalunga, T. obesus, T. albacares, T. tonggol, Katsuwonus pelamis* and *Gasterochisma melampus* for total DNA extraction were obtained from fishery industries or purchased in local markets. Tissue was kept frozen at -30°C prior to use. The information of the studied species as well as their pictures are provided in Table 1-1 and Fig. 1-1, respectively.

1.2.2. Total DNA extraction

Total DNA extraction from muscle tissue was performed using the QuickGene Tissue DNA kit S (Fujifilm) according to the instructions provided by the manufacturer. First, 15 mg of frozen tissue from each of the studied species were weighted in an 1.5-mL microcentrifuge tube, 180 µL of MDT buffer (lysis buffer) as well as 20 µL of EDT solution (proteinase K solution) were added and

mixed gently. Each extract was incubated at 55°C for 2 hours, and then centrifuged for 3 min at 20 000 *g*. The supernatant was collected and 180 µL of LDT buffer (containing guanidine chloride) were added. The mixture was incubated for 10 min at 70°C and 180 µL of 99% ethanol were added. After vortexing for 15 seconds, lysates were spin-downed and transferred for DNA extraction into a Nucleic Acid Isolation System QuickGene 810 (Fujifilm). Total extracted DNA was kept at 4°C.

1.2.3 Design of primers for direct sequencing

Primers for long PCR amplifications as well as internal primers for direct sequencing of all the studied species were designed based on a mtDNA sequence of *Thunnus thynnus* (Accession Number: AB047669) registered in the American National Center for Biotechnology Information (NCBI) as well as on the partial results of the ongoing sequencing. In general, the same primer pairs were used for all studied species but in the case of *G. melampus*, specific primers were necessary sometimes. The complete list of primers for direct sequencing is provided in Table 1-2. Dry primers were purchased from Operon, and reconstituted with TE buffer (pH 8.4) to a concentration of 100 µM. Then, an aliquot was further diluted with TE buffer to a concentration of 20 µM. High-concentration primers (100 µM) were kept as stock at -30° while low-concentration primers (20 µM) were used for PCR amplifications and kept at 4°C.

1.2.4 PCR amplification

Four primer pairs were used separately in long PCR amplifications on total DNA from the studies species, to obtain four main DNA fragments covering to the complete mitochondrial genomes and overlapping each other in a range of about 200 bp. The four main fragments were

subsequently purified and used as templates for a BigDye® Terminator labeling reaction using internal primers as described later.

Long PCR amplification was performed using LA Taq™ DNA polymerase (TaKaRa). Each amplification was carried out in 25 µL of the following reaction mixture: 2 µL of template DNA (average concentration, 20 ng/µL), 1 µL of 20 µM solution of both forward and reverse primer (final concentration, 0.8 µM), 2.5 µL of 10x PCR buffer, 2.5 µL of 25 nM $MgCl_2$ (final concentration, 2.5 mM), 4 µL of dNTP mixture (final concentration, 0.4 mM each), 0.25 µL of LA Taq DNA polymerase and distilled water to complete 25 µL. PCR amplification was performed on a Veriti® Thermal Cycler (Applied Biosystems) using the following program: 1 cycle of 1 min at 94°C; 30 cycles of i) 10 s at 98°C, ii) 5 s at 50°C, iii) 10 min at 68°C; and 1 cycle of 10 min at 72°C. PCR products were kept at 4°C until use.

1.2.5 Electrophoresis

Agarose gel electrophoresis was used to evaluate PCR amplifications. Briefly, 5 µL of PCR product were mixed with 1 µL of 6x dye buffer and loaded in 1% agarose gels containing GelGreen™ (Biotium). Electrophoresis was carried on at 100 V for 30 min in a Multi-Submerger Agarose 42 (Atto Corp.) and gels were observed in a ImageQuant™ LAS-4000 mini transilluminator (GE Healthcare).

1.2.6 Purification of PCR products with ExoSAP-IT

PCR product (5 µL) was mixed with 2 µL of a 10-fold diluted aliquot of ExoSAP-IT™ (GE Healthcare) and incubated for 1 hour at 37°C followed by 15 min at 85°C.

1.2.7 DNA sequencing

Purified products from the long PCR amplification were labeled for direct sequencing using a BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) as follows: 3.5 µL of purified PCR product were mixed with 2.0 µL of 5x sequencing buffer, 1.0 µL of Ready Reaction mix, 1 µL of 2 µM sequencing primer (final concentration, 0.2 µM) and distilled water to complete 10 µL. Labeling reaction was performed in a Veriti® Thermal Cycler (Applied Biosystems) using the following program: 25 cycles of i) 10 s at 96°C, ii) 5 s at 50°C and iii) 4 min at 60°C. After this, reaction mixtures were kept at 4°C until purified.

Purification of labeled DNA fragments was done by ethanol precipitation. Ten µL of labeled DNA fragments were transfer to a 1.5 mL-microcentrifuge tube, 2.5 µL of EDTA 250 mM and 30 µL of 99.9% ethanol were added and the mixture was immediately vortexed. The mixture was left at room temperature for 10 min and then centrifuged for 20 min at 40 000 *g*, at 4°C. After discarding the supernatant, 30 µL of 70% ethanol were gently added and the mixture was centrifuged for 20 min at 40 000*g*, at 4°C. The supernatant was discarded and the pellet was dried over vacuum in the dark for 7 min. Immediately after drying, 15 µL of Hi-Di™ Formamide (Applied Biosystems) was added. Tubes were gently shaken to dissolve the pellet and then incubated for 3 min at 95°C followed by 1 min on ice. After this, the solution was transferred to 96-well plates and prepared for sequencing using a AB3130 Genetic Analyzer (Applied Biosystems).

Sequencing data were processed, aligned and edited using Sequencher™ v4.10 (GeneCodes) in order to assemble the complete mitochondrial DNA sequences. Edited sequences were further analyzed with DNASIS® Pro (Hitachi Solutions) and with MEGA v5.0 (Tamura et al., 2011) for phylogenetic analyses.

1.3 Results and discussion

The complete mitochondrial sequence of 9 fish species was determined and characterized. According with the sequencing strategy, first 4 main segments were amplified (Table 1-2): Segment #1 (3974 bp) , segment #2 (6372 bp), segment #3 (5041 bp) and segment #4 (2851 bp). These 4 segments were used as templates for a BigDye® Terminator labeling reaction using internal primers. The schematic diagram of the mtDNA amplification is shown in Fig. 1-2. Finally, partial sequences were aligned and then the complete sequence was assembled. Except *T.orientalis*, two specimens of each species were analyzed in order to track significant differences among haplotypes. The complete mtDNA sequences were stored in the NCBI Genbank database, under the Accession Numbers shown in Table 1-1.

1.3.1 Description of the mtDNA sequences

The description of regions of the 9 mtDNA sequences is shown in Table 1-3. The length of the mitochondrial genome of the *Thunnus spp* was virtually the same (16527±1 bp), while the sequences of *G. melampus* (16506 bp) and *K. pelamis* (16519 bp) were slightly shorter. These minor changes in the sequence length were caused by deletions located in the major non-coding region (also known as D-Loop or control region). No differences in length were observed between sequences of different specimens of the same species. Each sequence contained 13 protein-coding regions: 7 genes encoding NADH dehydrogenase subunits (*ND1*, *ND2*, *ND3*, *ND4*, *ND4L*, *ND5*, and *ND6*), 3 genes encoding cytochrome oxidase subunits (*COI*, *COII*, and *COIII*), 2 genes encoding ATPase subunits (*ATP6* and *ATP8*) and 1 gene encoding the subunit b of the cytochrome bc1 complex (*cytb*). Each sequence also contained 2 rRNA-coding regions (*12S* and *16S*) as well as

22 tRNA-coding regions (two for leucine, two for serine and one for each of the other 18 amino acids). Two non-coding regions containing respectively the origin of replication of each DNA strand were identified. Most of the coding regions were located on the heavy strand of the mtDNA, while some of the tRNA-coding regions (glutamate, alanine, aspargine, cysteine, tyrosin, one for serine, aspartate, glutamine, and proline) and one protein-coding gene, *ND6*, were located on the light strand. The order of all regions was the same as reported previously for mtDNA of other bony fishes (Noack et al., 1996; Nagase et al., 2005).

The base composition of the heavy strand of the 9 mtDNA sequences, complete and also by region, is depicted in Table 1-4. The overall composition for the 9 sequences was almost identical. The mean composition was: Thymine, 25.2%; cytosine, 29.7%; adenine, 28.3% and guanine, 16.8% In all cases, the content of guanine was relatively lower than average, while adenine and cytosine were present in higher quantities. The mean composition of the major non-coding region (not shown) and of the tRNA-coding regions (T=27.1%; C=30.3%; A= 26.0%; G=16.6%) was similar to the overall composition but the protein-coding regions as well as the rRNA-coding regions differed significantly. On one side, the composition of the rRNA-coding regions tended to be more uniform, with guanine levels closer to average yet with high levels of adenine (T=20.9%; C=26.1%; A=31.6%; G=21.4%). On the other side, the composition of the protein coding regions was somehow similar to the overall composition as a whole (T=27.4%; C=30.5%; A=25.2%; G=16.9%) but within each codon position, very biased patterns were observed. In particular, the guanine content in the third codon position was remarkably low in all species, although it was somehow higher in *G.melampus*. This pattern has been observed in other fish and other vertebrates and it has been associated to a codon usage biased to A- or C- ending codons that favors transcription efficiency (Xia, 1996).

1.3.2 Scanning of species-specific DNA markers

After characterization of the 9 mtDNA sequences, the software MEGA v5.0 was used for phylogenetic analyses over the complete sequences as well as over rRNA- and protein-coding regions by separate. In addition to the sequences obtained in this experiment, sequences of *Thunnus* and *Katsuwonus* species reported in other studies were included for reference, as well as sequences of related Scombridae species (*Auxis rochei*, *A. thazard*, *Euthynnus alleteratus*, and *Scomber japonicus*). Phylogenetic trees were done using a neighbor-joining method with 1000 replicates. The bootstrap condensed trees (cut-off value: 70%) of the complete sequences as well as that of each region are shown in Fig. 1-3. The analysis of the complete mtDNA sequences confirmed the ancestral relations of the mitochondrial DNA of the genus *Thunnus*. The 7 *Thunnus* species were clustered in the same group while *Katsuwonus pelamis* and *Gasterochisma melampus* were placed separately. Within the *Thunnus* group, two subgroups could be readily distinguished: One containing *T. thynnus*, *T. maccoyii*, *T. obesus*, *T. tonggol* and *T. albacares* and the other containing *T. alalunga* and *T. orientalis*. These two subgroups could be observed not only in the phylogenetic tree of the complete mtDNA sequences but also in that of each of the mtDNA regions, confirming previous reports (Chow & Kishino, 1995) about the segregation of the mtDNA of *T. alalunga* and *T. orientalis* respect other tuna fish. The same reports have informed that although the mtDNA of *T. orientalis* is very similar to that of *T. alalunga*, its nuclear DNA is more closely related to that of *T. thynnus*. As mentioned previously, the accurate differentiation between *T. orientalis* and *T. thynnus* is relevant since the latter is subjected to more overfishing stress and its catches are more restricted that those of *T. orientalis*. The current results support the use of mtDNA for identification of these two tuna species, compared with nuclear DNA.

The phylogenetic analyses of each mtDNA region showed that the genealogic relations between *T. albacares*, *T. tonggol*, *T. maccoyii* and *T. thynnus* are not constant but change from one region to another and in some cases it is not possible to differentiate them with enough statistical support. Particularly, the *16S* region, which is used commonly in identification of other fish species, was found to be unable to differentiate accurately all the *Thunnus* species. On the other hand, in the *COI* and *ATP6* regions, which have also been used for differentiation of *Thunnus* species previously, the species could be differentiated clearly through phylogenetic analysis. The *cytb* region, which is also commonly used for discrimination in other fishes, was moderately effective to differentiate all the tuna species, with some low resolution in case of *T.obesus*, *T. maccoyii* and *T. alalunga*. The analyses of other regions revealed that the *12S* and *ATP8* regions, similarly to the *16S* region, are inadequate for discrimination purposes, because of the high degree of homology within *Thunnus* species in these regions. The 12S and 16S regions codify RNA subunits of the ribosomes, which are essential components to sustain the life. Because of its importance, the sequence of the ribosomal RNA (and in consequence, of its parent DNA) is usually very well conserved in all living beings and even more between close species. Thus, the high degree of homology of the *12S* and *16S* regions between *Thunnus* species can be regarded as the rule, rather than as the exception. For the *ATP8* region, further analyses (see below) revealed that the sequence of this region is identical among *Thunnus* species and thus, it cannot be used for differentiation purposes. This extreme case of homology contrasts with other observations reporting that the *ATP8* region tends to differ in other fish species (Broughton & Reneau, 2006). On the other side, phylogenetic trees of the *COIII* region, as well as of other NADH dehydrogenase-coding regions suggested that these regions may be used for accurate differentiation too.

In order to further investigate which regions are better candidates for design of species-specific DNA markers, in addition to phylogenetic analyses, the frequency of interspecies differences (disparity rate) as well as intraspecies differences (polymorphisms) were also investigated. Table 1-5 contains the results of the disparity rate between sequences. In accordance with the phylogenetic trees, *Gasterochisma melampus* was the species showing more disparities, with an average of 10.12% nucleotide substitutions respect to other species. *Katsuwonus pelamis* also showed an high disparity rate of 7.38% respect to *Thunnus* species. These high rates suggest that differentiation of *G.melampus* and *K.pelamis* from *Thunnus* species can be readily accomplished using mtDNA. In contrast, the disparity rate of the complete mtDNA sequence between *Thunnus* species was much lower, ranging from 2.29% (between *T.obesus* and *T.alalunga*) to values as low as 0.55% (between *T.albacares* and *T.maccoyii*). Nevertheless, these results also showed that the disparity rate is not uniform along regions. The house-keeping regions, *12S* and the *16S*, were almost identical in *Thunnus* species and, as mentioned before, the *ATP8* region did not show any difference within these species at all, while other regions showed a high degree of disparity, hence been potential candidates for the design of species-specific DNA markers. The regions with the highest disparity rates among *Thunnus* species were *ND1* (average: 3.57%), *ND2* (2.85%), *ND4* (2.63%), *ND5* (3.39%), *ND6* (3.54%) and *ATP6* (2.72%). Interestingly, the region COI, which was shown to be suitable for discrimination through sequencing and phylogenetic analysis, had a relatively low disparity rate (1.47%), which means that this region might be not so useful for discrimination of *Thunnus* species based on punctual DNA markers. The present results partially agree with previous observations performed in other fish species (Miya & Nishida, 2000) regarding the use of each mtDNA gene for establishing phylogenetic relations. These authors classified the mtDNA genes according with their theoretical capacity of differentiating 8 fish

species (none of them was tuna): Best performance: *ND5*, *ND4*, *COI* and *COIII*; good performance: *COII* and *cytb*; regular performance: *ND3* and *ND2*; poor performance: *ND1* and *ATP6*; and very poor performance: *ND4L, ND6, ATP8*. In turn, the current results, comparing only *Thunnus* species, showed that the *COII* region had poor resolution among these species, while in the *ND1*, *ND6* and *ATP6* regions, all species could be discriminated clearly. This proved that the genetic relations between a group of species cannot be necessarily extrapolated to other species.

The results of the analysis of intraspecies differences are shown in Table 1-6. The control region was the segment with largest rate of polymorphisms, which was expected, since this region is characterized for having zones of high variability in all species (Yu, 2005). Regarding the distribution of polymorphisms in other regions, no discernible pattern was observed within coding regions or within species: Some species showed a high rate of polymorphisms in regions where other species had few or any at all. Since there is no detailed information about the origin of each specimen, the results cannot be associated to geographical considerations. Nevertheless, it is worth to note that the species with the highest rate of polymorphisms, *T. alalunga* (4.62 polymorphic sites per thousand) and *K.pelamis* (6.59 polymorphic sites per thousand) are also two species of tuna species with the widest distribution, living in oceans all around the world. In comparison, *T. tonggol*, which lives in a relatively limited habitat, was the species with less polymorphism, reflecting a more uniform population. The highest rate of polymorphisms was observed in the *cytb* region, since *T. maccoyii* and *T. obesus* exhibited almost 1% of polymorphic sites in this region. However, a further comparison of *cytb* regions of *T. obesus* previously recorded in the NCBI GenBank (not shown) proved that this high rate of polymorphisms is not unusual. Even more, the *cytb* region contains suitable DNA markers, which have been using in other studies for tuna species differentiation (Terol et al., 2002). Therefore, the results of polymorphism rate revealed that,

although there is not a single region that can be used in all species for design of specific DNA markers, there are suitable regions for each species. Also, even regions with relatively high rate of polymorphisms may have sub-regions with low polymorphism rate. Finally, it is worth to mention that the polymorphisms on the nucleotide sequence were not reflected at protein level: The amino acid sequences translated from the respective coding-regions were identical among specimens of the same species as well as almost identical within *Thunnus* species (not shown). This was because the majority of the polymorphisms in the protein-coding regions were silent mutations. In contrast, the comparison of the translated proteins of *Thunnus* species respect to *K.pelamis* and especially to *G.melampus*, showed greater variations as it can be expected from species of different genus.

In conclusion, after determination and characterization of the complete sequence of the mitochondrial genome of 7 *Thunnus* species, *Katsuwonus pelamis* and *Gasterochisma melampus*, and upon comparison among these as well as respect to related species, it was found that the genetic relations of the mitochondrial DNA within the genus *Thunnus* are consistent and provide a good substrate for design of species-specific DNA markers. Due to the great level of homology between species, single nucleotide differences (SND) are the dominant source of genetic variability. It was found that not all the mtDNA regions used traditionally for species identification are suitable for *Thunnus* species and that, in contrast, other regions can contain suitable mtDNA markers for species differentiation.

Table 1-1. Fish species studied in this research

Scientific name	English name	NCBI GenBank accession numbers
Thunnus thynnus	Atlantic Bluefin Tuna	1) GU256522 2) JN086149
Thunnus orientalis	Pacific Bluefin Tuna	1) GU256524
Thunnus maccoyii	Southern Bluefin Tuna	1) GU256523 2) JN086150
Thunnus obesus	Bigeye Tuna	1) GU256525 2) JN086152
Thunnus albacares	Yellowfin Tuna	1) GU256528 2) JN086153
Thunnus alalunga	Albacore	1) GU256526 2) JN086151
Thunnus tonggol	Longtail Tuna	1) HQ425780 2) JN086154
Katsuwonus pelamis	Skipjack Tuna	1) GU256527 2) JN086155
Gasterochisma melampus	Butterfly Tuna	1) HQ425781 2) JN086156

Table 1-2. Complete list of primers used for determination of the mtDNA sequence of 9 fish species. Primers in bold were used in LA PCR, while the others were used as internal primers. Primers exclusive for *G. melampus* are underlined. Starting and ending nucleotides were assigned based on a previously reported mtDNA for *Thunnus thynnus* (AB097669)

NAME	SEQUENCE	bp	START	END	NAME	SEQUENCE	bp	START	END
Segment 1 Length: 3974 bp									
L073	**GCT TGG TCC TGA CTT TAC TGT**	21	73	93	H1068	CAC GGA TGA CTT TTC AGT GTA A	22	1047	1068
L878	AAA TCG GCT CTG AAG TGC GTA	21	878	898	H1710	GTG TAA TGC TTG AGC TTT AAC GC	23	1688	1710
L1393	AAT AGG GCA AAC CCG TCT CTG T	22	1393	1414	H2279	ACA TGA TAT GCC TTG GTG TCT AA	23	2257	2279
16SarL	CGC CTG TTT ATC AAA AAC AT	20	2007	2026	16SbrH	GCG GTG TGA ACT CAG ATC ACG T	22	2602	2623
L2379	GAG TAC TAC CTC CTA CAA CCA A	22	2379	2400	H3242	TTA GAT GCT CAG CCT GAG CCT A	22	3221	3242
L2938	ATT GAA CGA AAA GTA CTA GGC TA	23	2939	2961	**H4047**	**GTT CAT TAG TAA AAG AAG GAT TTT A**	25	4023	4047
Segment 2 Length: 6372 bp									
L3683	**TAT GAG TGC GAG CCT CCT AC**	20	3683	3702	H4559	TGG GTT TGG TTT AGT CCA CCT CA	23	4537	4559
L4352	TAA AAA TTG GGC TAG CCC CAG T	22	4352	4373	H5263	TAA GAG TTT GTA GGA TCG AGG CCT	24	5240	5263
L5104	ATA GTC CAA GGG CCT TCA AAG	21	5104	5124	H6143	TGC AGG GTC GAA GAA GGT T	19	6125	6143
L5895	GCA TCA GTT GAC CTA ACT ATT TTC	24	5895	5918	H6673	AAT AGT GGG AAT CAG TGT ACG AA	23	6651	6673
L6414	TTC CAA CTG GTG TAA AAG TAT TTA	24	6414	6437	H7139	CTC GAC AAG GCA GTG TAA TA	20	7120	7139
L6893	AGA AGC TTT CGC TGC CAA AC	20	6823	6842	H7685	CAA GAG AAG GGA CTG CTC ATG A	22	7664	7685
L7347	AGA TTC CCA AGA AAT CGA AAT CAT	24	7347	7370	H8277	AGT TCG TGG GCA AAT CGG	18	8260	8277
L7868	CCT CGC TAA GAA GCT AAA CCG	21	7868	7888	H8757	AGT AGG AAA AGT AGG GTT GCT	21	8683	8703
L8459	CGA AAC CAA CCA ACA GAA GC	20	8459	8478	H9138	TGG TCA GCA GCC TCC TAG TT	20	9119	9138
L8803	ACA CGC ATA CCA CAT AGT TGA	21	8803	8824	H9396	AAG CCT GTG GCT ACA AAG AA	22	9374	9396
L9043	TCA TCA CCT CTG AAG TCT TCT TC	23	9043	9065	**H10055**	**TAT GGA CTT TAA CCA CAA GTT C**	22	10034	10055

Table 1-2. (continued)

Segment 3	Length: 5041 bp								
L9502	**GCC TGA TAC TGA CAC TTC GTA GAC GT**	26	9502	9527	H10585	TTA TTA ATG GGA GGA GTC AGC A	22	10564	10585
L10123	ATT CCA TCG AAC CCA CCT CCT	21	10123	10143	H11186	AGA CAG ATT GAG CCT GTT AT	20	11167	11186
L10780	GTT TAA ATG CAG GGA CTT ACT T	22	10780	10801	H11716	TTA GCT CGG GTT TGA GGA TA	20	11697	11716
L11413	CCC CTT ATA ACA ACA TGA TGA TT	23	11413	11435	H12266	AGG ATT GAT CAT GTT ACG TAG AG	23	12244	12266
L11413b	TTA TTA CCT CCC TGT TCA ACT G	22			H12266b	ATG TAA CAT GAT GCG AAT TCG	21		
L11934	AAT CCA AGT AGC AGC TAT GCA	21	11934	11954	H12924	AAG GCC AAT TGT TAC TAT CA	20	12905	12924
L12509	TCA TTC TTG CCA TAG CAT GAA T	22	12509	12530	H13526	GTT GTA AGC ATT GGG GTA GGT	21	13506	13526
					H13526b	ATA ATT AGT CCG AGG ATG GTA ACG	24		
L13153	ATC GAA GCA CTA AAC ACA TCA CA	23	13153	13175	H14089	TAG CTG CTG AGC CTT ATC CTG A	22	14068	14089
L13704	CAA TGC CCA GCA AGG TAT AAT	21	13704	13724	**H14543**	**TAG TGT ATT GCG AGG AAT AGT CCT GT**	26	14518	14543
Segment 4	Length: 2851 bp								
L14261	**CCC TAA TAC TAA ACC AAA TAA TAA**	24	14262	14285	H15122	TAG GGG AGA ATA GTG CTA GA	20	15103	15122
L14873	GAG GAG GCT TTT CAG TAG ACA A	22	14873	14894	H15641	ATC CTA GCT TTG GGA GTT AGG	21	15621	15641
L15352	TTC TGA ACC CTT ATT GCA GAC GT	23	15352	15374	H16122	TTA TGC AAG CGT CGA TGA AA	20	16103	16122
L15352b	TTT TGA ACC CTA ATC GCA AAC GT	23	15352	15374	H16122b	TTA ACT TAT GCA AGC GTC GAT	21	16106	16127
L16003	ACG GTT ATT GAA GGT GAG GGA CA	23	16003	16025	**H585**	**AAT TCT ATC AAT GTT TAC GGC T**	22	564	585

Table 1-3. Description of the regions of the complete mtDNA genome of 9 fish species. Numbers indicate the starting and ending nucleotides of each region. *Region allocated in the light strand. **Protein-coding regions where the termination codon is incomplete

Region name	*All Thunnus spp*	*K.pelamis*	*G. melampus*	Region name	*All Thunnus spp*	*K.pelamis*	*G. melampus*
tRNA-Phe	1-68	1-68	1-68	tRNA-Lys	7873-7946	7875-7948	7872-7945
12S rRNA	69-1015	69-1017	69-1015	ATP8	7948-8115	7950-8117	7947-8114
tRNA-Val	1016-1088	1018-1090	1016-1088	ATP6	8106-8789	8108-8791	8105-8788
16S rNA	1089-2780	1091-2781	1089-2780	COIII**	8789-9574	8791-9576	8788-9573
tRNA-Leu	2781-2854	2782-2855	2781-2854	tRNA-Gly	9574-9644	9576-9646	9573-9644
ND1	2855-3829	2857-3831	2855-3829	ND3**	9645-9995	9647-9997	9645-9995
tRNA-Ile	3834-3904	3835-3905	3834-3904	tRNA-Arg	9994-10062	9996-10064	9994-10062
tRNA-Gln*	3904-3974	3905-3975	3904-3974	ND4L	10064-10360	10065-10361	10064-10360
tRNA-Met	3974-4042	3975-4043	3974-4042	ND4	10354-11734	10355-11735	10354-11734
ND2	4043-5089	4044-5090	4043-5089	tRNA-His	11735-11804	11736-11805	11735-11804
tRNA-Trp	5089-5158	5091-5162	5089-5160	tRNA-Ser	11805-11872	11806-11873	11805-11872
tRNA-Ala*	5160-5228	5164-5232	5162-5230	tRNA-Leu	11877-11949	11878-11950	11877-11949
tRNA-Asn*	5231-5303	5235-5307	5232-5304	ND5	11950-13788	11951-13789	11950-13788
Ori-L*	5305-5338	5308-5341	5305-5338	ND6*	13785-14306	13786-14307	13875-14306
tRNA-Cys*	5339-5406	5342-5409	5339-5405	tRNA-Glu*	14308-14376	14308-14376	14307-14375
tRNA-Tyr*	5407-5474	5410-5476	5406-5472	Cytb	14381-15521	14381-15521	14380-15520
COI	5476-7026	5478-7028	5474-7024	tRNA-Thr	15522-15593	15522-15593	15521-15592
tRNA-Ser*	7027-7097	7029-7099	7025-7096	tRNA-Pro*	15593-15662	15593-15662	15592-15661
tRNA-Asp*	7101-7173	7103-7175	7100-7172	D-Loop	15663-16527		
COII**	7182-7872	7184-7874	7181-7871		(15663-16526 *T. maccoyii, T.albacares*) (15663-16528 *T. tonggol*)	15663-16519	15662-16506

Table 1-4. Base composition of mtDNA sequences from 9 fish species (percentage)

Complete mitochondrial genome				
Species	**T**	**C**	**A**	**G**
T. thynnus	25.4	29.5	28.4	16.7
T. maccoyii	25.3	29.5	28.4	16.8
T.orientalis	25.4	29.5	28.4	16.7
T.alalunga	25.4	29.5	28.4	16.7
T.obesus	25.5	29.4	28.4	16.7
T.albacares	25.3	29.6	28.4	16.7
T.tonggol	25.4	29.5	28.5	16.6
K.pelamis	24.6	30.2	28.5	16.7
G.melampus	24.8	30.8	26.9	17.5

rRNA-coding region (12S)				
Species	**T**	**C**	**A**	**G**
T. thynnus	21.1	26.5	30.0	22.4
T. maccoyii	21.0	26.6	30.0	22.4
T.orientalis	21.4	26.5	29.6	22.5
T.alalunga	21.2	26.6	29.8	22.4
T.obesus	21.1	26.5	29.9	22.5
T.albacares	21.0	26.6	29.9	22.5
T.tonggol	21.0	26.6	29.8	22.6
K.pelamis	21.2	27.0	29.4	22.4
G.melampus	21.3	27.0	29.4	22.3

rRNA-coding region (16S)				
Species	**T**	**C**	**A**	**G**
T. thynnus	20.7	25.5	33.6	20.3
T. maccoyii	20.7	25.5	33.5	20.3
T.orientalis	20.7	25.5	33.5	20.3
T.alalunga	20.8	25.4	33.5	20.3
T.obesus	20.7	25.4	33.6	20.2
T.albacares	20.7	25.5	33.6	20.2
T.tonggol	20.7	25.5	33.5	20.3
K.pelamis	20.6	25.5	33.5	20.3
G.melampus	20.8	25.6	33.1	20.5

tRNA-coding regions				
Species	**T**	**C**	**A**	**G**
T. thynnus	27.4	30.1	26.0	16.5
T. maccoyii	26.5	31.0	26.1	16.4
T.orientalis	27.5	30.0	26.0	16.5
T.alalunga	27.4	30.0	26.0	16.5
T.obesus	27.4	30.0	26.0	16.5
T.albacares	27.4	30.0	26.1	16.5
T.tonggol	27.4	30.0	26.1	16.5
K.pelamis	26.1	31.0	26.1	16.8
G.melampus	26.8	30.5	25.7	17.0

Protein-coding regions					*First position of codon*				*Second position of codon*				*Third position of codon*			
Species	**T**	**C**	**A**	**G**	**T**	**C**	**A**	**G**	**T**	**C**	**A**	**G**	**T**	**C**	**A**	**G**
T. thynnus	27.6	30.2	25.4	16.7	21.0	28.3	23.5	27.2	40.7	27.9	17.6	13.8	21.2	34.6	35.3	9.0
T. maccoyii	27.7	30.2	25.4	16.7	21.0	28.3	23.4	27.3	40.7	27.9	17.7	13.8	21.4	34.4	35.1	9.1
T.orientalis	27.6	30.4	25.5	16.6	20.8	28.5	23.5	27.2	40.8	27.9	17.6	13.8	21.2	34.8	35.3	8.7
T.alalunga	27.5	30.4	25.4	16.6	20.8	28.5	23.5	27.3	40.8	27.9	17.6	13.8	21.0	34.9	35.2	8.9
T.obesus	27.7	30.1	25.5	16.7	21.0	28.3	23.5	27.2	40.6	27.9	17.6	13.8	21.5	34.2	35.3	9.0
T.albacares	27.7	30.3	25.4	16.7	20.9	28.3	23.4	27.3	40.7	27.9	17.6	13.8	21.3	34.6	35.2	8.9
T.tonggol	27.7	30.2	25.5	16.6	21.0	28.3	23.6	27.2	40.7	27.9	17.6	13.8	21.6	34.3	35.3	8.9
K.pelamis	26.7	31.3	25.5	16.5	20.6	28.6	23.5	27.3	40.8	27.8	17.6	13.8	18.8	37.6	35.4	8.3
G.melampus	26.9	31.8	23.4	18.0	20.4	29.3	22.8	27.5	40.7	27.9	17.6	14.0	19.4	38.3	29.7	12.4

Table 1-5. Disparity rates between complete mtDNA sequences as well as within regions. Numbers represent nucleotide differences per hundred sites.

Complete genome	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*	*G.melampus*
T.maccoyii	0.77								
T.orientalis	2.21	2.17							
T.alalunga	2.07	2.03	0.64						
T.obesus	1.10	0.87	2.29	2.15					
T.albacares	1.08	0.55	2.24	2.10	0.80				
T.tonggol	1.13	0.84	2.25	2.07	0.95	0.62			
K.pelamis	7.44	7.30	7.50	7.56	7.33	7.28	7.28		
G.melampus	10.26	10.15	10.15	10.14	10.13	10.14	10.12	9.85	
Average (9 species)	3.26	3.08	3.68	3.60	3.20	3.10	3.16	7.69	10.12
Average (excluding *G.melampus*)	2.26	2.08	2.76	2.66	2.21	2.09	2.16	7.38	

12S rRNA	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	0.32							
T.orientalis	0.96	1.07						
T.alalunga	0.64	0.75	0.32					
T.obesus	0.53	0.21	1.07	0.75				
T.albacares	0.42	0.11	0.96	0.64	0.11			
T.tonggol	0.53	0.21	1.07	0.75	0.21	0.11		
K.pelamis	2.84	2.51	2.51	2.62	2.51	2.40	2.51	
Average (8 species)	0.89	0.74	1.14	0.92	0.77	0.68	0.77	2.56
Average (*Thunnus spp*)	0.57	0.44	0.91	0.64	0.48	0.39	0.48	

16S rRNA	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	0.05							
T.orientalis	0.47	0.42						
T.alalunga	0.38	0.33	0.19					
T.obesus	0.14	0.09	0.52	0.33				
T.albacares	0.14	0.09	0.52	0.33	0.09			
T.tonggol	0.09	0.05	0.38	0.28	0.14	0.14		
K.pelamis	2.01	1.96	2.30	2.31	2.06	1.96	2.01	
Average (8 species)	0.47	0.43	0.69	0.59	0.48	0.47	0.44	2.09
Average (*Thunnus spp*)	0.21	0.17	0.42	0.31	0.22	0.22	0.18	

Table 1-5 (continued).

ND1	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	0.83							
T.orientalis	5.13	5.36						
T.alalunga	4.69	4.91	1.24					
T.obesus	2.40	2.62	5.91	5.24				
T.albacares	2.73	2.51	5.14	4.58	2.30			
T.tonggol	2.72	2.72	5.58	4.69	1.77	1.98		
K.pelamis	14.67	14.67	15.59	15.98	15.31	15.03	15.18	
Average (8 species)	4.74	4.80	6.28	5.91	5.08	4.89	4.95	15.20
Average (*Thunnus spp*)	3.08	3.16	4.73	4.23	3.37	3.20	3.24	

ND2	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	0.81							
T.orientalis	3.09	3.16						
T.alalunga	3.02	3.09	1.06					
T.obesus	1.70	1.76	3.03	3.23				
T.albacares	1.89	1.96	3.63	3.56	1.83			
T.tonggol	2.02	2.09	3.77	3.84	1.83	1.12		
K.pelamis	11.61	11.53	12.13	12.21	11.70	11.35	11.79	
Average (8 species)	3.45	3.49	4.27	4.29	3.58	3.62	3.78	11.76
Average (*Thunnus spp*)	2.09	2.14	2.95	2.97	2.23	2.33	2.44	

ND3	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	0.85							
T.orientalis	2.27	2.44						
T.alalunga	2.26	2.44	0.33					
T.obesus	0.34	0.50	1.91	1.90				
T.albacares	0.85	0.00	2.44	2.44	0.50			
T.tonggol	1.19	0.67	2.81	2.80	0.84	0.67		
K.pelamis	10.55	10.53	10.57	10.33	10.08	10.53	10.99	
Average (8 species)	2.61	2.49	3.25	3.22	2.30	2.49	2.85	10.51
Average (*Thunnus spp*)	1.29	1.15	2.04	2.03	1.00	1.15	1.50	

Table 1-5 (continued)

ND4	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	1.72							
T.orientalis	3.39	3.64						
T.alalunga	3.08	3.41	0.81					
T.obesus	1.63	1.42	3.63	3.31				
T.albacares	1.72	0.17	3.73	3.50	1.50			
T.tonggol	1.46	0.72	3.31	3.09	1.50	0.72		
K.pelamis	10.39	10.39	9.93	10.27	10.28	10.62	10.39	
Average (8 species)	3.34	3.07	4.06	3.92	3.32	3.14	3.03	10.32
Average (*Thunnus spp*)	2.17	1.85	3.08	2.87	2.17	1.89	1.80	

NDL4	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	0.48							
T.orientalis	0.36	0.85						
T.alalunga	0.36	0.85	0.00					
T.obesus	0.48	0.48	0.85	0.85				
T.albacares	0.48	0.00	0.85	0.85	0.48			
T.tonggol	0.36	0.12	0.73	0.73	0.36	0.12		
K.pelamis	3.77	3.46	3.30	3.30	3.46	3.46	3.62	
Average (8 species)	0.90	0.89	0.99	0.99	0.99	0.89	0.86	3.48
Average (*Thunnus spp*)	0.42	0.46	0.61	0.61	0.58	0.46	0.40	

ND5	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	1.71							
T.orientalis	3.92	4.04						
T.alalunga	3.75	3.75	0.99					
T.obesus	2.55	2.04	4.09	3.92				
T.albacares	2.49	1.32	4.04	3.86	2.04			
T.tonggol	2.60	2.21	4.26	3.97	2.26	1.32		
K.pelamis	13.19	13.25	13.20	13.52	13.11	13.05	13.45	
Average (8 species)	4.31	4.04	4.93	4.82	4.29	4.02	4.30	13.25
Average (*Thunnus spp*)	2.84	2.51	3.56	3.37	2.82	2.51	2.77	

Table 1-5 (continued)

ND6	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	1.75							
T.orientalis	4.38	4.17						
T.alalunga	5.00	4.79	1.75					
T.obesus	1.55	1.75	5.21	5.84				
T.albacares	1.75	0.00	4.17	4.79	1.75			
T.tonggol	1.75	1.16	4.17	4.79	2.15	1.16		
K.pelamis	18.41	19.23	20.92	22.65	17.59	19.23	18.68	
Average (8 species)	4.94	4.69	6.40	7.09	5.12	4.69	4.84	19.53
Average (*Thunnus spp*)	2.70	2.27	3.98	4.50	3.04	2.27	2.53	

COI	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	0.71							
T.orientalis	1.77	1.84						
T.alalunga	1.57	1.64	0.32					
T.obesus	1.04	1.11	1.64	1.57				
T.albacares	0.91	0.85	1.64	1.57	0.65			
T.tonggol	1.17	1.24	1.97	1.77	1.04	0.52		
K.pelamis	9.67	9.68	9.59	9.59	9.35	9.04	9.20	
Average (8 species)	2.41	2.44	2.68	2.58	2.34	2.17	2.42	9.45
Average (*Thunnus spp*)	1.20	1.23	1.53	1.41	1.18	1.02	1.29	

COII	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	0.29							
T.orientalis	0.87	0.88						
T.alalunga	1.02	1.02	0.14					
T.obesus	0.14	0.14	0.73	0.87				
T.albacares	0.29	0.29	0.87	1.02	0.14			
T.tonggol	0.44	0.44	0.73	0.87	0.29	0.44		
K.pelamis	7.34	7.50	6.84	7.00	7.50	7.34	7.50	
Average (8 species)	1.48	1.51	1.58	1.71	1.40	1.48	1.53	7.29
Average (*Thunnus spp*)	0.51	0.51	0.70	0.83	0.39	0.51	0.53	

Table 1-5 (continued)

COIII	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	0.26							
T.orientalis	3.02	3.02						
T.alalunga	3.02	3.02	0.51					
T.obesus	0.90	0.64	2.89	2.89				
T.albacares	0.38	0.38	2.62	2.62	0.51			
T.tonggol	0.90	0.90	2.62	2.62	0.77	0.51		
K.pelamis	8.93	8.93	9.10	9.09	8.78	8.46	8.77	
Average (8 species)	2.49	2.45	3.40	3.40	2.48	2.21	2.44	8.86
Average (*Thunnus spp*)	1.41	1.37	2.45	2.45	1.43	1.17	1.39	

cytb	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	0.79							
T.orientalis	3.24	3.05						
T.alalunga	2.78	2.59	0.44					
T.obesus	1.42	1.33	3.24	2.78				
T.albacares	1.24	0.79	3.06	2.60	0.88			
T.tonggol	1.69	1.60	3.15	2.69	1.33	1.15		
K.pelamis	12.87	12.33	12.75	12.32	12.42	12.12	11.69	
Average (8 species)	3.43	3.21	4.13	3.74	3.34	3.12	3.33	12.36
Average (*Thunnus spp*)	1.86	1.69	2.70	2.31	1.83	1.62	1.93	

ATP6	*T.thynnus*	*T.maccoyii*	*T.orientalis*	*T.alalunga*	*T.obesus*	*T.albacares*	*T.tonggol*	*K.pelamis*
T.maccoyii	0.88							
T.orientalis	4.54	4.22						
T.alalunga	4.07	4.07	0.73					
T.obesus	2.08	1.78	5.18	5.02				
T.albacares	1.33	1.03	4.38	4.22	1.93			
T.tonggol	1.33	1.33	5.00	4.85	2.23	0.88		
K.pelamis	16.01	16.37	15.63	15.81	15.64	15.63	15.45	
Average (8 species)	4.32	4.24	5.67	5.54	4.84	4.20	4.44	15.79
Average (*Thunnus spp*)	2.37	2.22	4.01	3.83	3.04	2.30	2.60	

*No differences were encountered in the *ATP8* region within *Thunnus* species and thus, that region is not shown.

Table 1-6. Single nucleotide polymorphisms (SNP) detected in the mtDNA sequences. Numbers represent SNPs per thousand sites. No number means there were no differences between specimens in the respective region.

Species (no. of specimens)	**Complete sequence**	***12S***	***16S***	***ND1***	***ND2***	***ND3***	***ND4***	***ND4L***
T.thynnus (4)	2.16	2.47	0.28	5.70	1.11	--	2.66	14.05
T.maccoyii (2)	2.43	3.18	0.85	7.24	0.96	--	0.72	4.47
T.orientalis (2)	1.88	2.12	0.85	--	0.96	--	1.45	--
T.alalunga (3)	4.62	--	1.71	6.19	11.00	1.05	6.35	--
T.obesus (2)	1.39	1.06	--	--	--	--	--	--
T.albacares (2)	0.85	--	0.43	1.03	0.96	--	3.63	1.48
T.tonggol (2)	0.83	1.06	--	5.16	--	--	2.18	--
K.pelamis (3)	6.59	--	3.72	4.13	5.79	3.19	7.84	11.86
G.melampus (2)	3.87	--	--	4.12	0.96	3.18	11.76	--
Average (Thunnus spp.)	2.02	1.41	0.58	3.61	2.49	0.15	2.42	3.33

Species (no. of specimens)	***ND5***	***ND6***	***COI***	***COII***	***COIII***	***ATP6***	***Cytb***	**Control Region**
T.thynnus (4)	1.17	--	0.65	0.72	1.27	--	1.17	10.9
T.maccoyii (2)	2.73	--	--	--	--	1.19	10.6	15.80
T.orientalis (2)	3.28	--	--	--	--	0.00	1.75	17.70
T.alalunga (3)	5.10	4.56	--	1.93	2.55	1.18	1.17	23.00
T.obesus (2)	--	--	--	--	--	3.59	11.47	9.40
T.albacares (2)	--	--	--	--	--	0.59	1.76	8.20
T.tonggol (2)	0.54	--	0.65	--	--	0.59	--	1.20
K.pelamis (3)	7.24	7.64	3.46	1.93	3.41	2.79	5.28	34.90
G.melampus (2)	4.37	3.41	6.49	1.45	1.27	2.39	7.94	5.00
Average (Thunnus spp.)	1.83	0.65	0.18	0.37	0.54	1.02	3.98	12.31

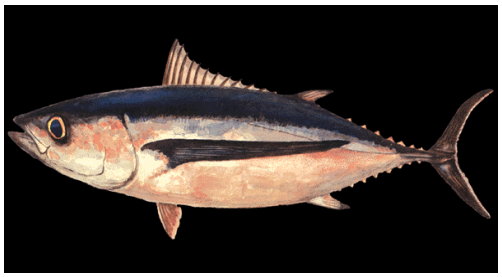
Thunnus thynnus

Thunnus orientalis

Thunnus maccoyii

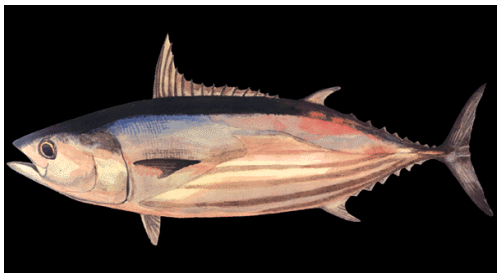
Thunnus obesus

Thunnus albacares



Thunnus alalunga

Thunnus tonggol



Katsuwonus pelamis

Fig. 1-1 Fish species studied in this research.
*Images from FAO Fishstat and www.fishbase.org

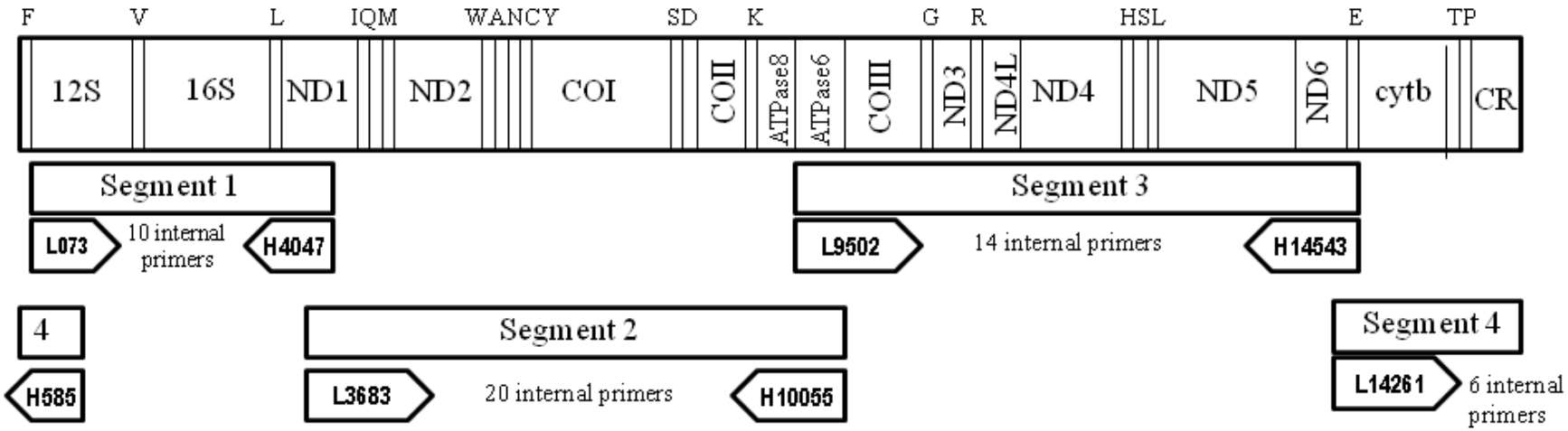


Fig 1-2. Schematic diagram of amplified regions for the sequencing of the complete mtDNA of 9 fish species. For a complete description of the sequencing primers, see Table 1-2.

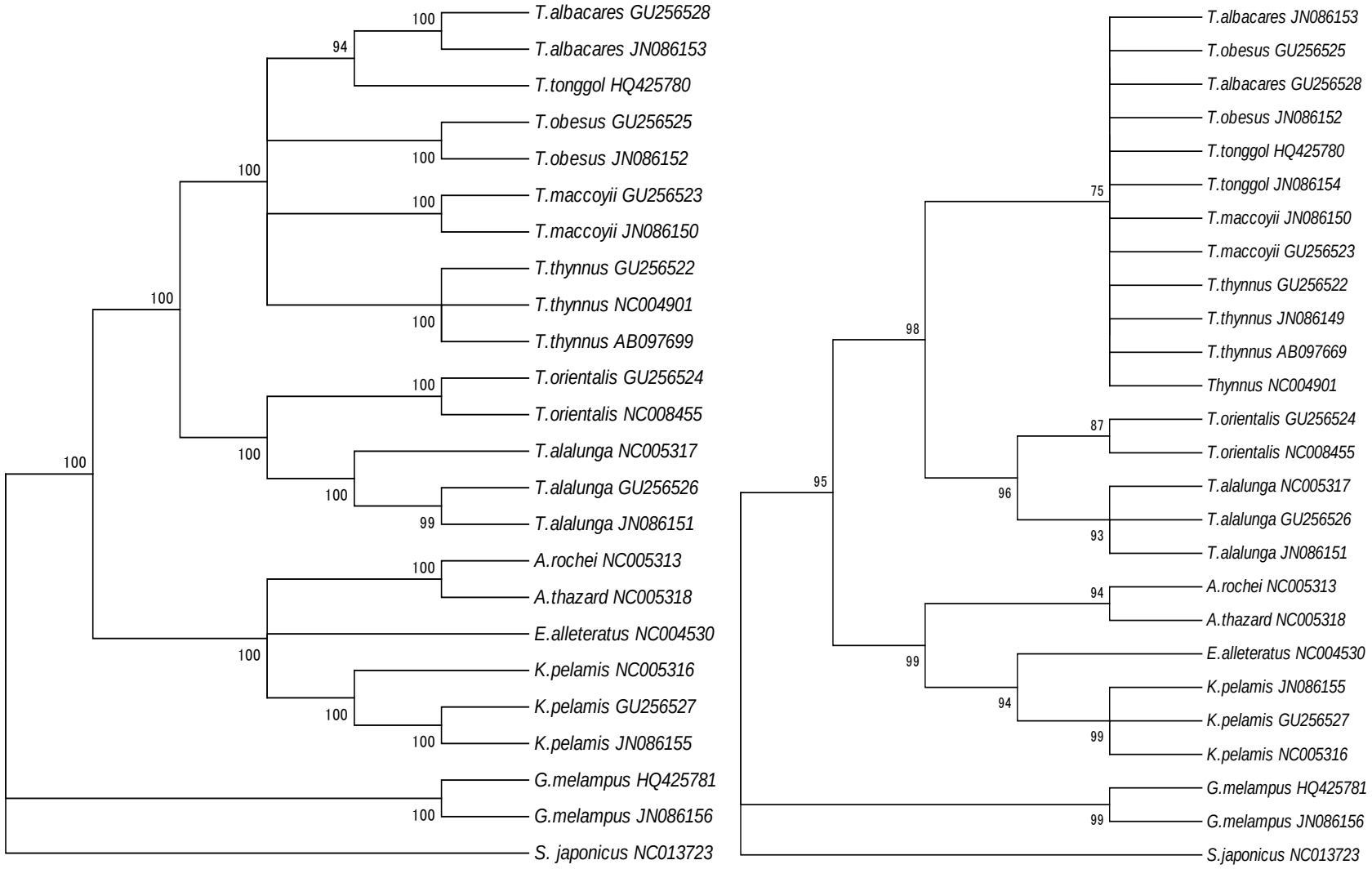


Fig. 1-3. Phylogenetic tree of mtDNA genomes from tuna and related species. **Left.** Complete mtDNA. **Right**. *12S* region

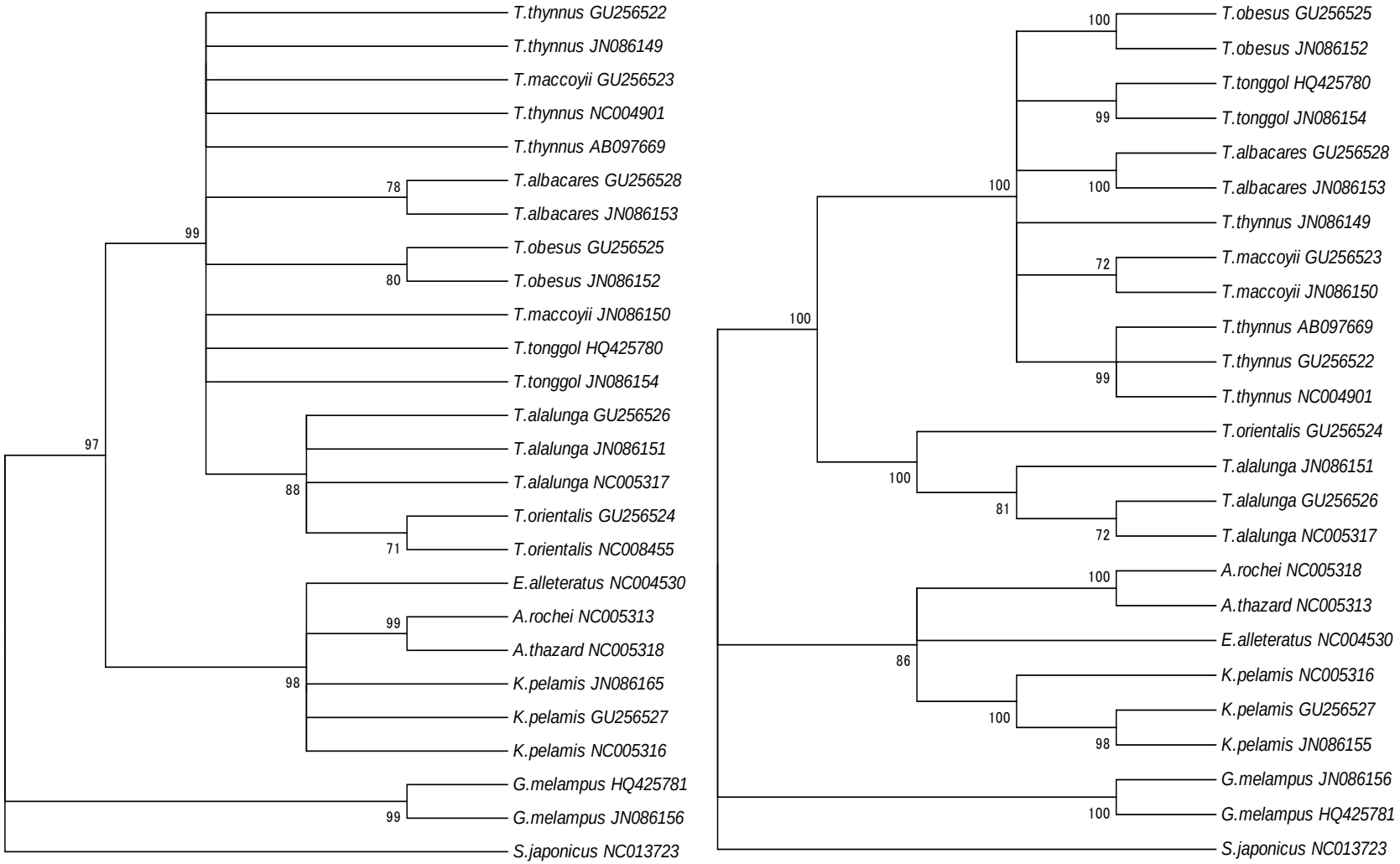


Fig. 1-3. (continued). Left. *16S* region. **Right.** *ND1* region

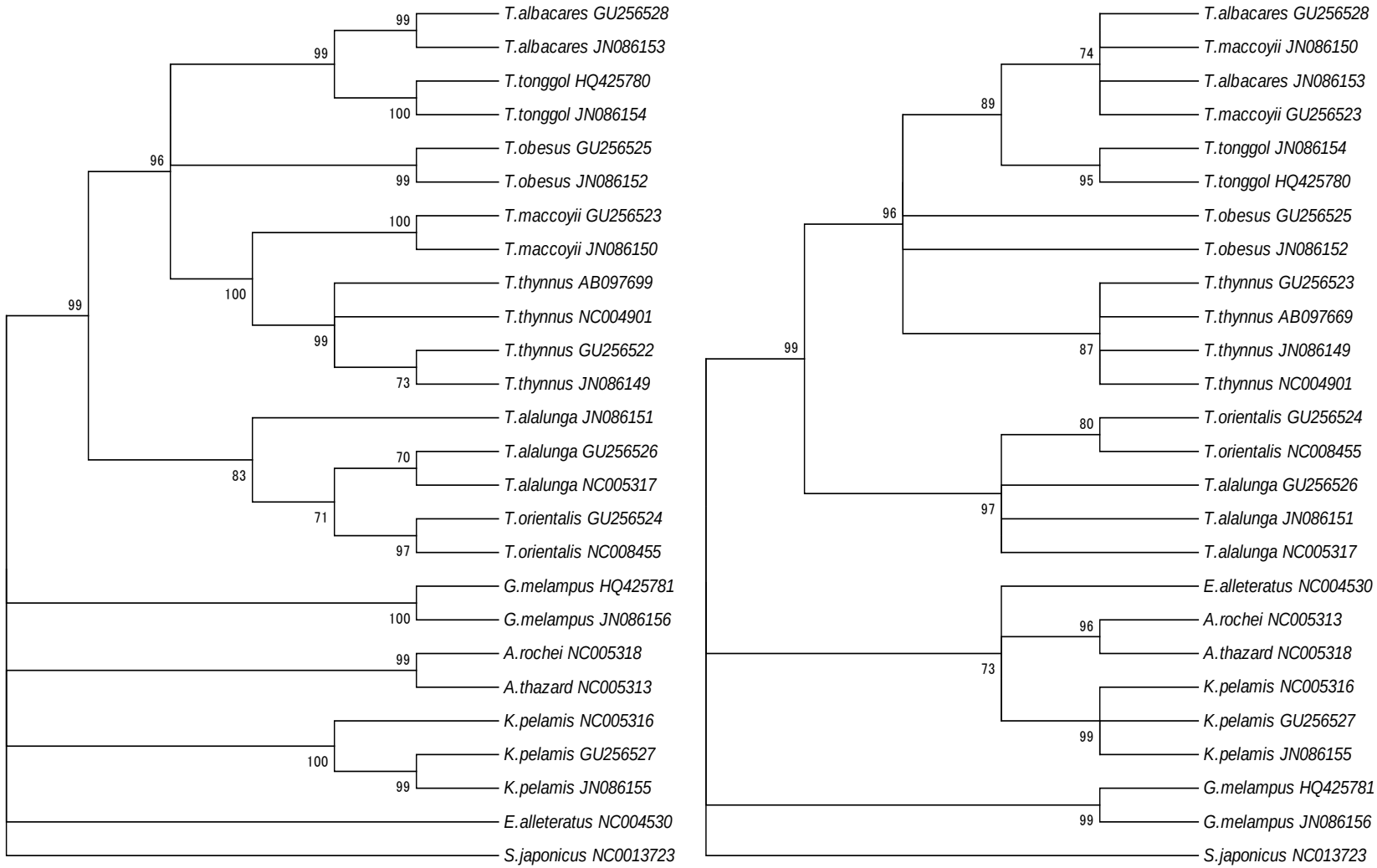


Fig. 1-3. (continued). Left. *ND2* region. **Right.** *ND3* region

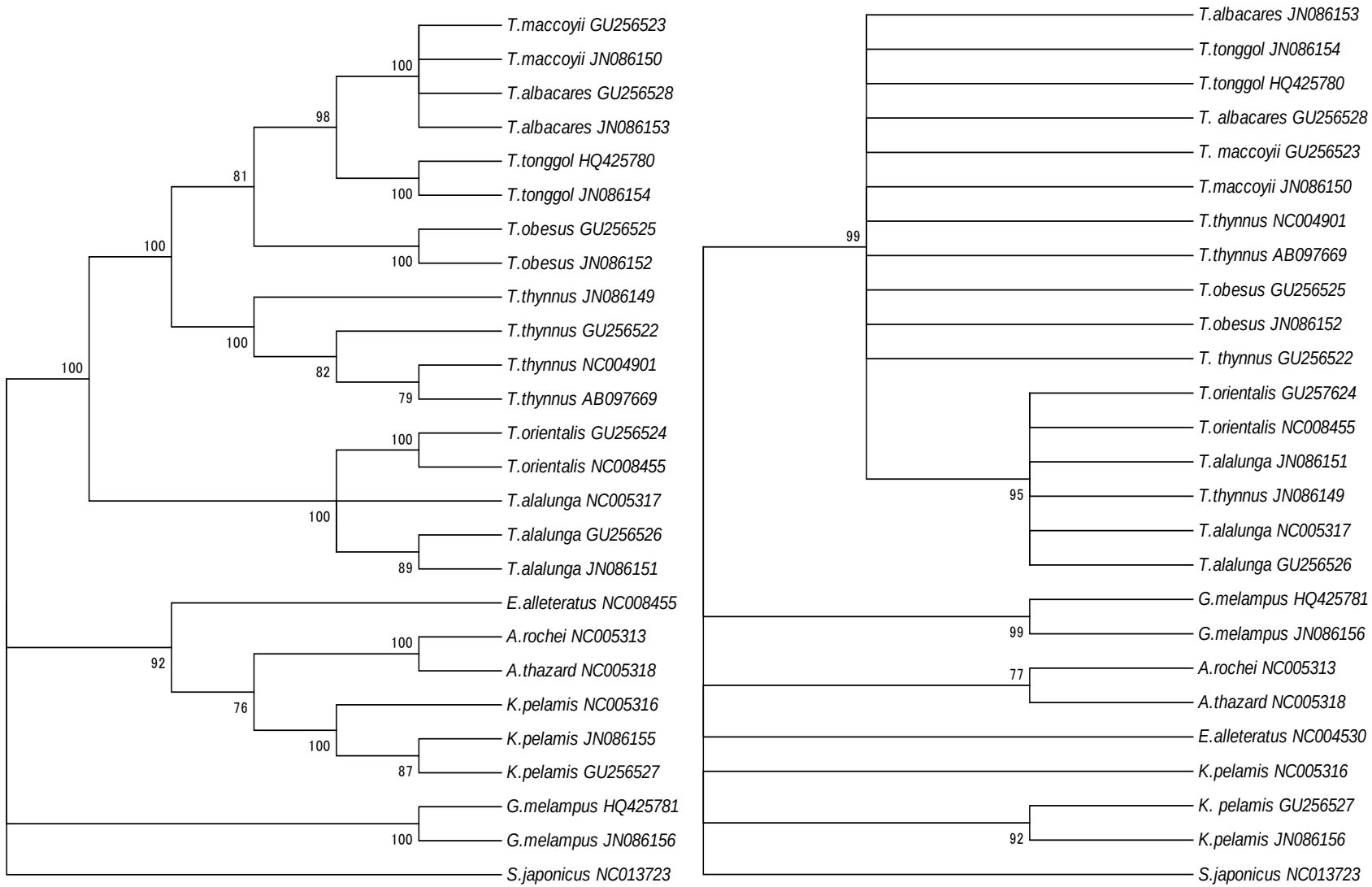
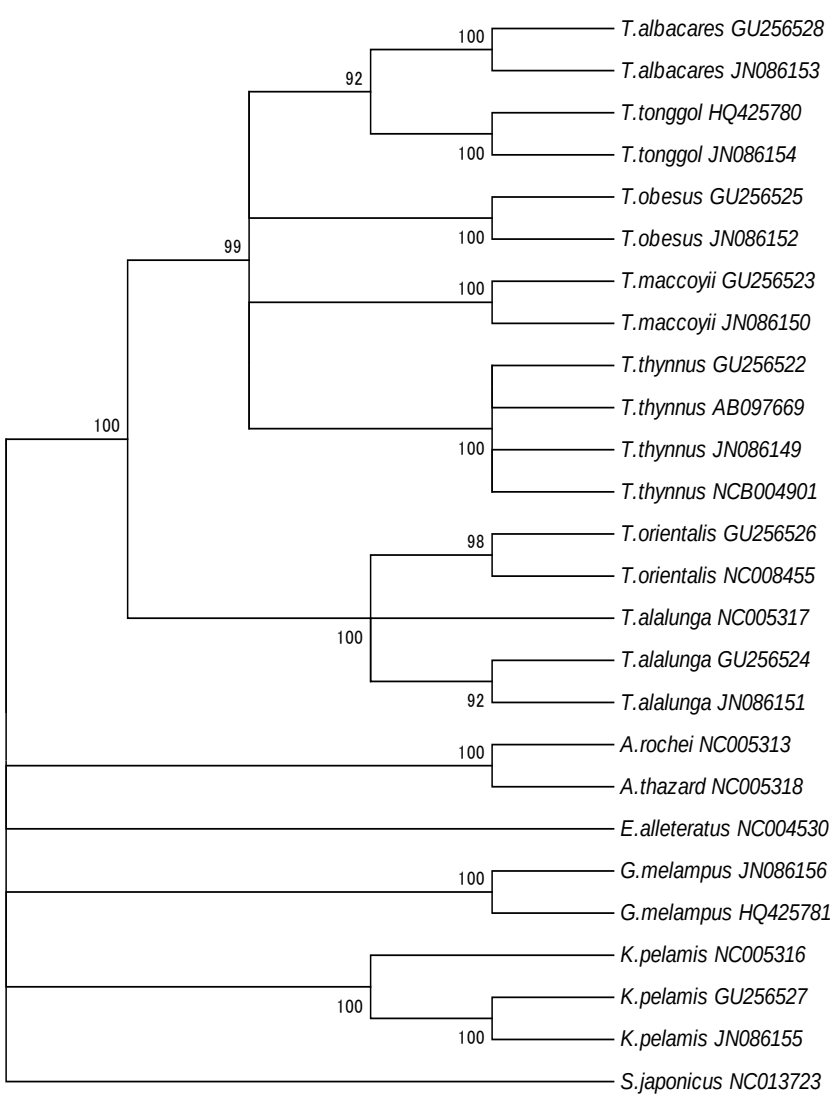


Fig. 1-3. (continued). Left. *ND4* region. **Right.** *ND4L* region



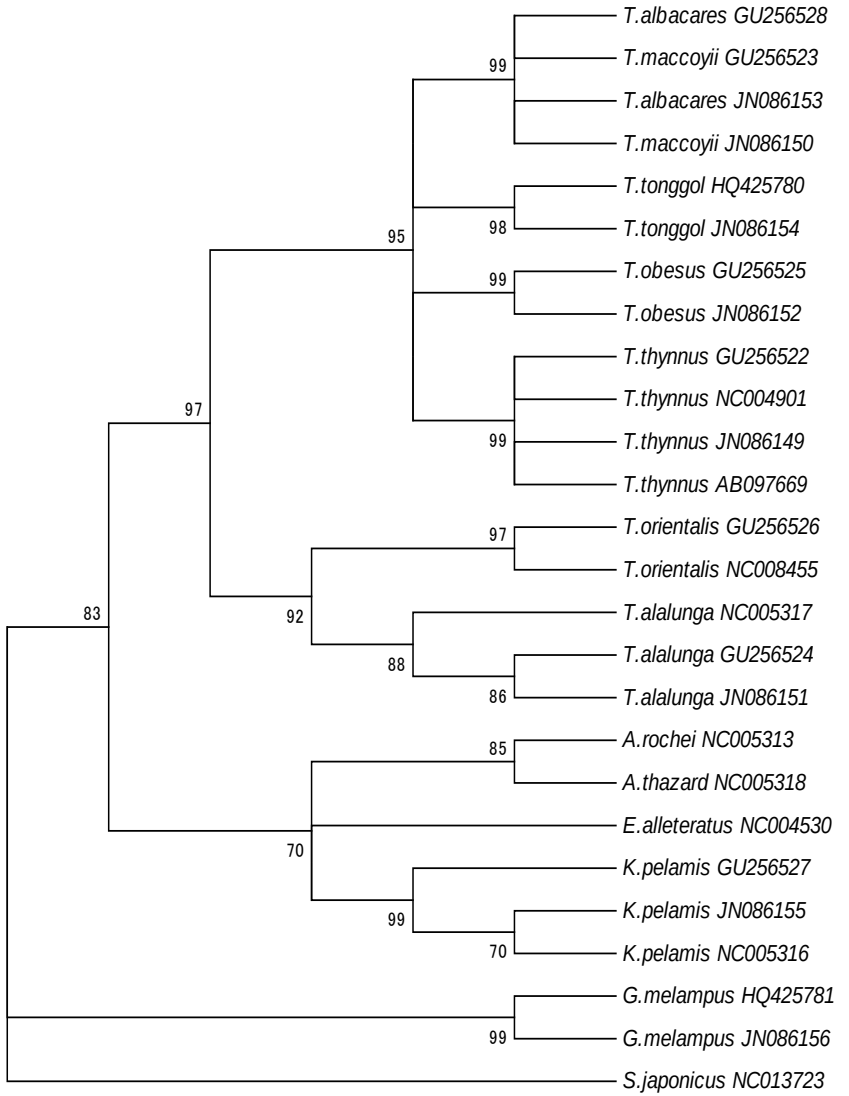
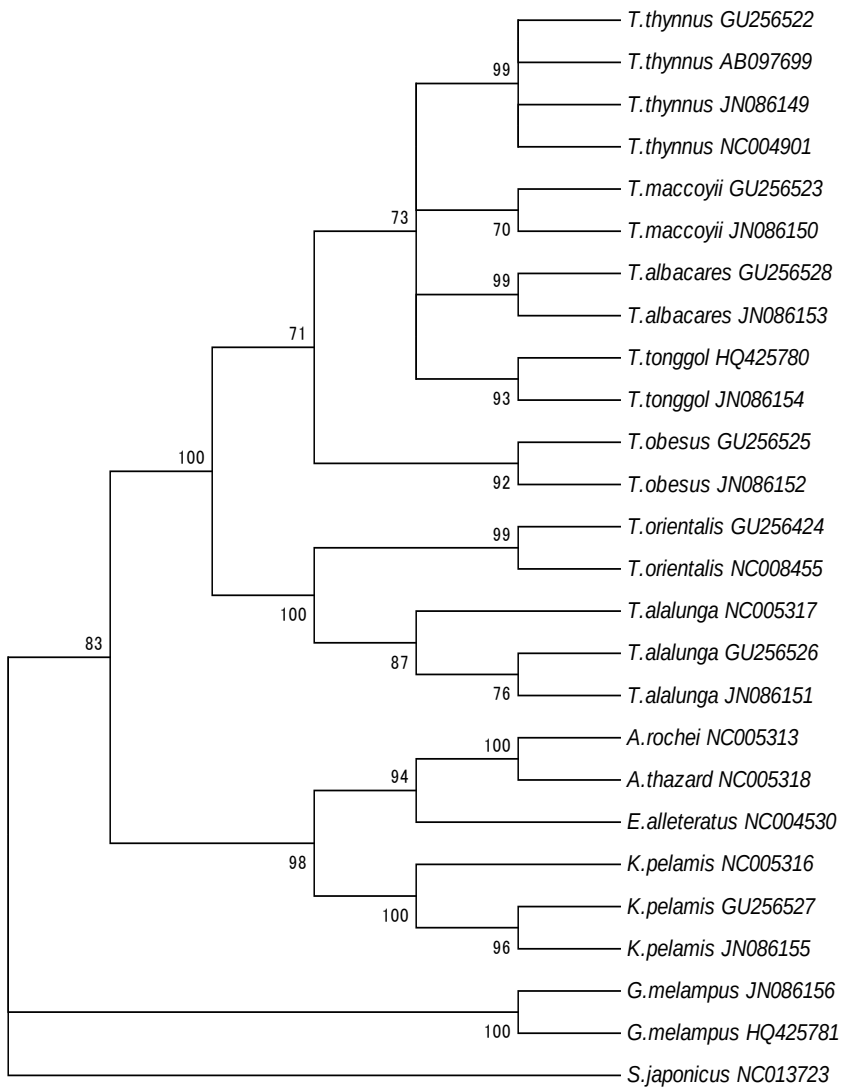


Fig. 1-3. (continued). Left. *ND5* region. **Right.** *ND6* region



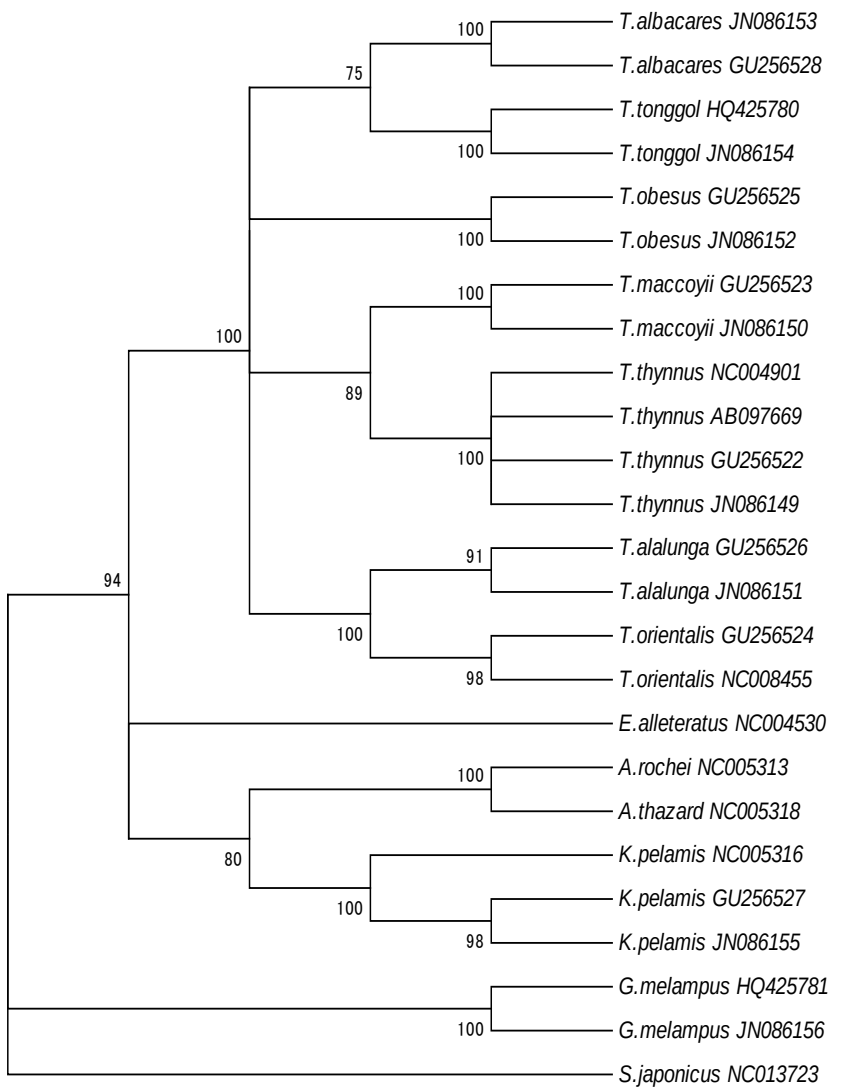
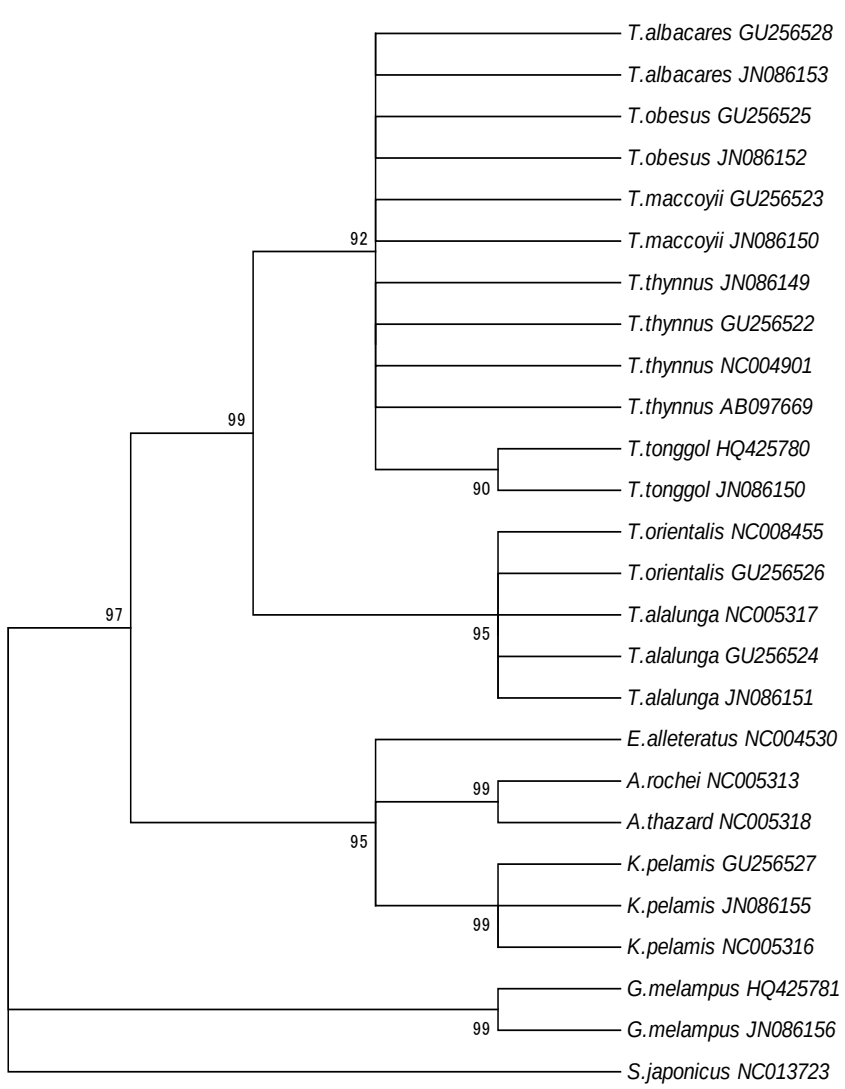


Fig. 1-3. (continued). Left. *ATP6* region. **Right.** *COI* region



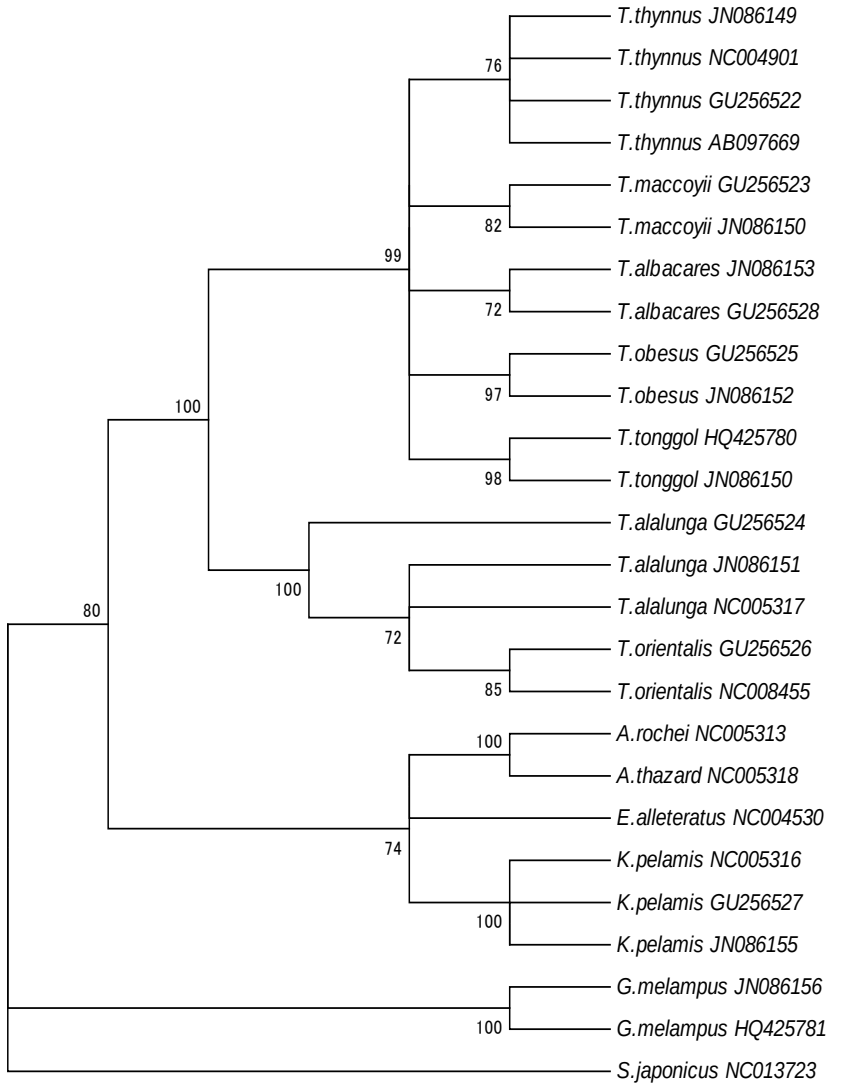


Fig. 1-3. (continued). Left. *COII* region. **Right.** *COIII* region

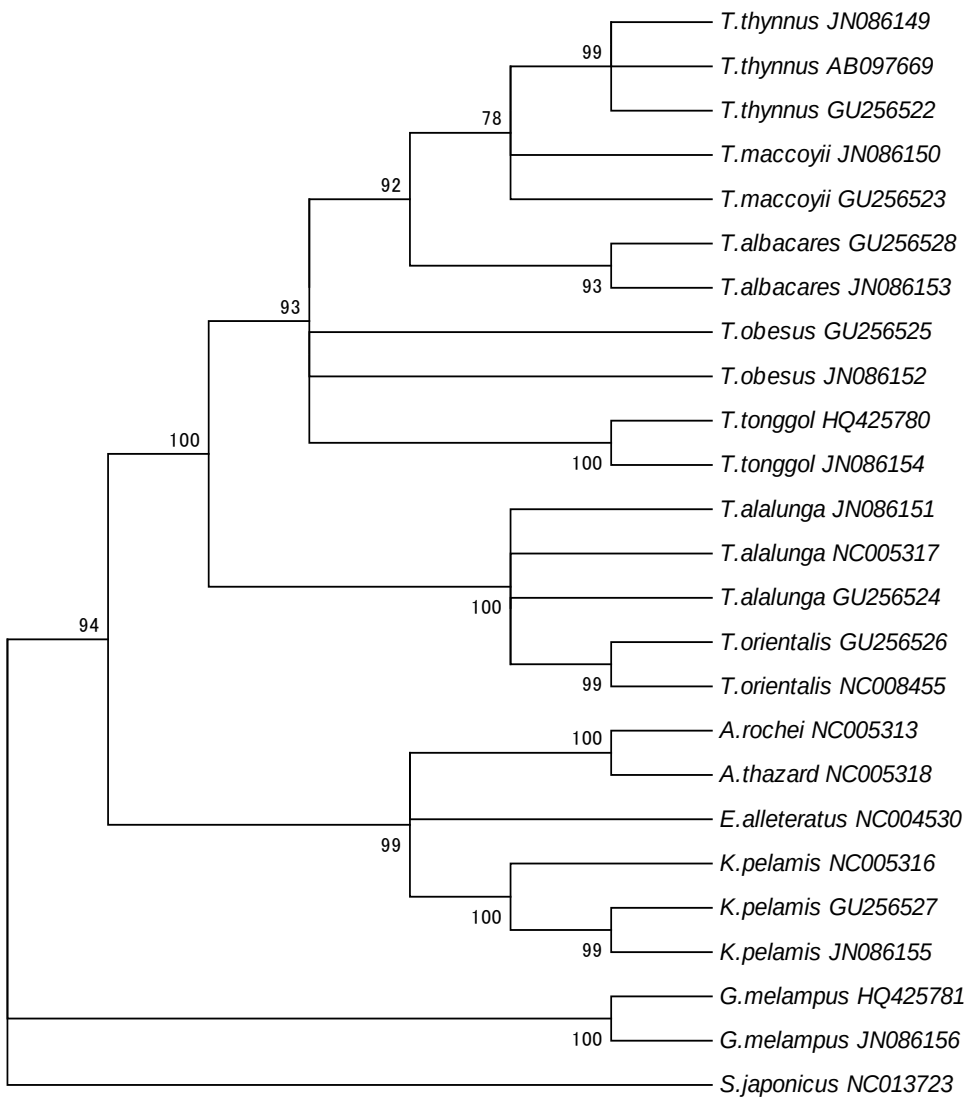


Fig. 1-3. (continued). c*ytb* region

Chapter 2. Development of a PCR-based differentiation method of tuna species in heat-processed and unheated food products

2.1 Introduction

Identification of species based in DNA has the advantage of being particularly accurate, given the intrinsic specificity of the nucleic acids and relative thermostability. In contrast, methods based on proteins, such as isoelectric focusing (Rehbein, 2005) and immunoassays (Asensio et al., 2008), sometimes exhibit problems to distinguish closely related species and, moreover, they become ineffective in heat-processed samples, such as processed food products. For these reasons, DNA has become the standard substrate for authentication and identification. There is a variety of DNA analysis techniques that have been used for species identification, ranging from long-established techniques, like DNA hybridization (Chikuni et al, 1990, Takeyama et. al, 2000), to novel and emerging methods like DNA pyrosequencing (Balitzki-Korte et al., 2005) and MALDI-TOF DNA sequencing (Mazzeo et al., 2008). Nevertheless, the majority of the methods based on DNA analysis for species discrimination in food are based on the polymerase chain reaction (PCR), since it provides efficiency and specificity. In the case of tuna fish, PCR-based methods developed so far include, among others, restriction fragment length polymorphisms (RFLP) (Chow & Inoue, 1993; Lin & Hwang, 2005), single-strand conformational polymorphism (SSCP) (Rehbein et al., 1999), forensically informative nucleotide sequencing (FINS) (Viñas & Tudela, 2009), DNA barcoding (Botti & Giuffra, 2010), nested PCR (Pardo & Pérez-Villareal, 2004), etc. A number of real-time PCR methods, (Terol et al., 2002, Kitaoka et al.,

2009), where the use of specific probes allows identification of the species without post-PCR processing, have also been developed. This variety of methods reflects the commercial and environmental importance of tuna species but nevertheless, none of these methods have fully addressed the issue of identifying quickly and feasibly all major tuna species, this issue becoming therefore the main objective of the present study.

This research used mtDNA as substrate due to the advantages described in the previous chapter. The analysis of the mtDNA sequences confirmed the high degree of homology among *Thunnus* and *Katsuwonus* species and that the dominant source of genetic variability that can be used as DNA marker were single nucleotide differences. In this research, the term "single nucleotide differences" (SND) is used to denote isolated nucleotide ***interspecies*** substitutions occurring in one site of the sequence, in contrast to "single nucleotide polymorphisms" (SNP), which are ***intraspecies*** substitutions. Single nucleotide substitutions have been used as DNA markers for differentiation of species, subspecies and strains of several organisms through different strategies, such as SNP genotyping (Gibson, 2006; Pattemore et al., 2010) and RNAase protection assay (Kitaoka et al., 2008). However, one of the simplest yet effective methods for differentiation of SND and SNP is the use of primers with mismatches in their 3' extreme. A large number of DNA polymerases used commonly for PCR and lacking of 5' exonuclease activity (also called proof-reading activity) are particularly sensitive to mismatches in the 3' extreme of the primer. It is possible to achieve a great specificity to one allele by locating a SND in the 3' extreme of the primer, thus hindering the PCR amplification of mismatched alleles even if the rest f the sequence is identical to that of the target allele. This method, often named as ARMS (Amplification Refractory Mutation System) or also as allele-specific PCR, was first

described by Newton et al. (1989) and since then, it has been used in several clinical (Bottema et al, 1993) and other analytical applications, including food analysis (Hayashi et. al., 2004; Garcés-Claver et al., 2007). Allele-specific PCR has also been applied to identification of some invertebrate marine species (Wang & Guo, 2008, Liu et al., 2011) but so far, it has not been applied to fish species, despite the advantages it offers in terms of specificity and sensibility. In the present study, a new method for discrimination of the major tuna species in unheated as well as in heat-processed food products based on species-specific primer pairs (SSPP) with 3' mismatches was developed, standardized and tested in commercial samples.

2.2 Materials and methods

2.2.1. Samples and materials

For specificity tests, the total DNA extracted from tuna species in the previous chapter for mtDNA sequence determination was used, as well as DNA from the following species: *Scomber japonicus* (chub mackerel), *Xiphias gladius* (swordfish), *Cololabis saira* (Pacific saury), *Engraulis japonicus* (Japanese anchovy), *Pagrus major* (red seabream), *Pleuronectes yokohamae* (marbled flounder), *Salmo salar* (Atlantic salmon) as well as terrestrial animals: *Gallus gallus* (chicken), *Sus scrofa* (domestic swine), *Bos taurus* (common cow), from muscle tissue, and one sample of *Homo sapien*s DNA (human), from blood. For trials on commercial samples, pieces of frozen tuna loins, *sashimi*, *sushi* and other raw tuna products were purchased in local markets. Heat-processed products such as canned tuna, skipjack flakes (*katsuobushi*), tuna pate and others were purchased from local markets and also

obtained from other countries when available. Unheated samples were kept at -30°C until DNA extraction. The complete list of the commercial samples and their information is shown in Table 2-1.

2.2.2 Total DNA extraction

DNA was extracted from samples using the QuickGene Tissue DNA kit S (Fujifilm) as described in the previous chapter. For canned and other heat-processed samples, oil and other cover liquids were removed with absorbent paper when necessary and 30 mg of meat were used.

2.2.3 Measurement of DNA concentration

All determinations of the DNA concentration were carried out in a BioSpec-Nano (Shimadzu Biotech, Japan) according to the instructions of the manufacturer.

2.2.4. Chloroform-phenol DNA extraction

DNA from canned tuna was also extracted using the chloroform-phenol extraction method reported by Asahida et al. (1996) to evaluate the effect of the DNA extraction on the differentiation method. First, 50 mg of canned tuna meat were finely chopped and mixed in a 1.5 mL microtube with 500 µL of TNES-Urea buffer (Tris, 10 mM; NaCl, 125 nM) and 20 µL of proteinase K solution (20 mg/ml) and digested at 55°C overnight. After digestion, 500 µL of phenol/chloroform/isoamylic alcohol (25:24:1) were added and tubes were mixed gently for 10 min. Then, tubes were centrifuge at 5000 *g* for 15 min at 20°C and the upper phase was recollected in a new microtube. Extraction with phenol/chloroform/isoamylic

alcohol (25:24:1) was repeated a second time in the same fashion. Then, 500 µL of chloroform/isoamylic alcohol (24:1) were added, mixed gently for 10 min and centrifuged at 5000 *g* for 15 min at 20°C. The upper phase was recollected and the extraction with chloroform/isoamylic alcohol (24:1) was repeated for a second time in the same fashion. After that, 50 µL of 3M sodium acetate and 1000 µL of 100% ethanol cooled at -20°C were added, the mixture was gently stirred and left at -20°C overnight. Subsequently, the mixture was centrifuged at 5000 *g* for 15 min at 20°C and the supernatant was discarded. The pellet was resuspended in 1000 µL of 70% ethanol cooled at -20°C and centrifuged at 5000 *g* for 15 min at 20°C. The supernatant was discarded. The pellet was dried over vacuum for 7 min and re-dissolved in 200 µL of TE buffer.

2.2.5 Design of specific primers

The comparison of mtDNA sequences and the screening of SND were done using DNASIS® Pro (Hitachi Solutions) on the complete mtDNA of 8 tuna species, except the control region. Primers were designed manually. Dry primers were purchased from Operon and reconstituted with TE buffer (pH 8.4) to achieve a concentration of 100 µM. Then, aliquots were further diluted with TE buffer to a concentration of 20 µM. High-concentration primers (100 µM) were kept in stock at -30°C while low-concentration primers (20 µM) were used for routine experiments and kept at 4°C.

2.2.6 PCR amplification and optimization

All PCR reactions were performed with a Veriti® Thermal Cycler (Applied Biosystems) using ExTaq™ DNA polymerase (TaKaRa). The initial composition of the PCR reaction mixture was as follows: 2.5 µL of 10x PCR buffer, 2 µL of dDNP mixture (final concentration, 0.2 mM each), 1 µL of 20 µM primers (final concentration, 0.8 µM each), 2 µL of DNA template (30 ng), 0.2 µL of Ex Taq DNA polymerase and distilled water to complete 25 µL. Initial PCR amplification conditions were as follows: 1 cycle of 30 s at 98°C and 25 cycles of i) 10 s at 98°C, ii) 30 s at 50°C, iii) 1 min at 72°C. PCR products were kept at 4°C. PCR amplifications were conducted under these initial conditions and changed gradually until attaining the best efficiency and specificity for each primer. The critical parameters that were optimized were: i) Number and length of amplification cycles, ii) denaturing, annealing and extension temperatures, iii) primer concentration.

2.2.7 Detection limit

For trials of detection limit, PCR amplifications using SSPP and TSPP under the optimized conditions were carried out on dilutions of DNA from unheated samples. Total DNA was standardized to a final amount of 30 ng and diluted 1:10 three times with TE buffer to 3 ng, 0.3 ng and 30 pg. DNA from unheated samples was used also for testing short SSPP since DNA from heat-processed products has been fragmented to an unknown extent and thus it would not reflect the real detection limit of the method.

2.2.8 DNA sequencing

For verification of the identity of unheated samples, a mtDNA fragment between the ATP6 and COIII regions (ATCO fragment), reported as suitable for identification of tuna species (Chow & Kishino, 1995), was analyzed by direct sequencing and studied through phylogenetic analysis. Briefly, total DNA from each sample was used as a template in a PCR amplification using ExTaq DNA polymerase. The composition of the PCR reaction as well as the amplification conditions were as described in section 2.2.5 here above. For PCR and for the labeling reaction, the primer pair L8459/H9138, used for the determination of the complete mtDNA sequences (Chapter 1, Table 1-3), was used.

Direct sequencing was also used for characterizing the haplotypes on the surrounding sequence of each SND in unheated samples. The PCR and labeling primers were some of those used during the determination of the complete mtDNA. These haplotypes were recorded in the NCBI GenBank database for future reference.

PCR products were purified, subjected to BigDye® Terminator labeling and sequenced following the procedures described in the previous chapter (sections 1.2.6 and 1.2.7)

2.2.9 RFLP analysis

For verification of the identity of canned samples, a RFLP analysis was carried out according to a method reported previously (Lin & Hwang, 2007). Total DNA (5 µL) extracted from canned samples was used as a template in a PCR amplification on the conditions described here above using the primer pair Cb146L (5'-CCT CGC AAT ACA CTA TAC CCC-3')/ Cb146H (5'-CGA TGT GGA AGT AGA TGC AG-3') and, separately,

the primer pair Cb126L (5'-GCY TYT ACT ACG GYT CYT AC-3')/Cb 126H (5'-CCC CTC AGA ATG ATA TTT GTC C-3'). PCR products were purified with the Exo SAP procedure described previously (Chapter 1, Section 1.2.6) and digested with restriction enzymes: *RsaI* for the Cb126L/H product and *HincII* for the Cb146L/H product. Enzymatic digestions were carried out in 10 µL of a mixture with 20 ng of PCR product, 5U of restriction enzyme, 1 µL of 1% bovine serum albumin and 1 µL of digestion buffer at 37°C for 2 h. The digests were therefore applied to electrophoresis in 3% agarose gels containing GelGreen™ (Biotium) and read in a ImageQuant™ LAS-4000 mini transilluminator (GE Healthcare).

2.3. Results and discussion

2.3.1. Description of the species-specific primer-PCR discrimination system.

As mentioned before, mismatches in the 3' extreme of a primer have a significant influence on the efficiency of PCR. Some authors (Newton et al., 1989; Garcés-Claver et al., 2007) have reported complete specificity by merely introducing 3' mismatches in the primer, but in other cases, additional measures have to be taken in order to ensure specificity (Gibson, 2005). Chun et al. (2007) reported an ARMS method where, in addition to mismatches in the 3' region, they used a "bridge" of inosine inside the primer. Although this system achieved great specificity even in highly similar sequences, it was suggested that the inosine bridge may interfere with the activity of some DNA polymerases. In the present study, in addition to natural mismatches derived from the presence of SND in the 3' extreme of the species-specific primers, one deliberate mismatch was introduced from 2 to 3 bases upstream the SND. Primers designed this way are still able to align the target template, since

between them there is only one mismatch, while in the case of other templates there are two mismatches, one of them in the 3' extreme that preclude amplification. The schematic representation of the double-mismatch mechanism is depicted in Fig. 2-1. In addition, a "*Thunnus* specific primer pair" (TSPP) was also used. The TSPP is a common primer pair without mismatches, designed to amplify only DNA from *Thunnus* species. The role of the TSPP is multiple: i) To act as a positive control of the PCR reaction, ii) to act as an indicator of the state of the DNA template and iii) to act as a "quencher" of excessive DNA polymerase activity.

2.3.2 Design of species-specific primers.

Although the DNA in unheated tuna products such as *sushi* and frozen loins is generally well conserved, in heat-processed products, DNA is usually fragmented. The more intense the heating, the less long DNA sequences remain. Thus, DNA methods applied to food products depend on the identification of short fragments, usually no longer than 200 bp. The kinetics and efficiency of PCR in such small templates are different from that of longer fragments. For this reason, in the present study two different types of species-specific primers (SSPP) were designed: One type to be used in unheated products, generating amplicons of 276-440 bp (long SSPP), and one type to be used in heat-processed products, generating amplicons of 107-161 bp (short SSPP). The mtDNA sequences of 8 tuna species were compared, looking for SND to design SSPP with mismatches in the 3' extreme. Candidate SND were chosen according to the following criteria:

1) Uniqueness: The SND should display one unique allele in one species and different alleles in the other species.

2) Non-polymorphic region. Whenever possible, SNDs located in areas with low intraspecies variability were chosen.

3) Primer design not generating primer dimers.

Following these criteria, 16 primer pairs, 8 short SSPP and 8 long SSPP, were designed. The sequence of all the SSPP is given in Table 2-2 and the corresponding comparison of mtDNA sequences for each short SSPP is shown in Fig. 2-2. Comparisons of long SSPP are not shown, since they are based in the same SND used for short SSPP. Most of the SSPP are based on SND found in the ND4, ND5 and ND6 regions. Only primers for *Thunnus tonggol* were located on the COI region. Whenever it was possible, both the forward and reverse short SSPP were designed to have a SND in the 3' extreme. For long SSPP, only one of the primers has a SND in the 3' extreme, while the other is a normal primer. In the case of the short SSPP of *T. obesus*, additional mismatches were introduced in order to decrease the formation of primer dimers. The forward short SSPP of *Katsuwonus pelamis* is the only primer that is not based on one SND but on a short sequence of 4 nucleotides that was exclusive of *K. pelamis*.

The *Thunnus*-specific primer pair (TSPP) was designed by comparison of the 16S region of several fish species as well as other animals. The 16S region was chosen since it has few polymorphisms and is very similar within *Thunnus* species. The TSPP amplicon is has a length of 89 bp, and can be distinguished clearly from the shortest SSPP product (107 bp, *T. orientalis* short SSPP) when analyzed by electrophoresis. The alignment of mtDNA sequences for design of the TSPP can be seen in Fig. 2-2i. The TSPP does not have mismatches but the reverse primer takes advantage of a deletion of 2 bp that was exclusive

of *Thunnus* species. Since deletions have an impact on specificity even higher than mismatches (Little, 1994), the TSPP can achieve great specificity to *Thunnus* species.

2.3.3 Specificity tests

PCR reactions using SSPP and TSPP simultaneously and by separate were carried out to evaluate specificity. It was shown (Fig. 2-3) that using SSPP without TSPP, was enough to achieve complete selective amplification of the target DNA in some species (*T. albacares* [Fig. 2-3a]; *T. albacares*, *T. obesus,* [Fig. 2-3b]), but in most cases, amplification of other tuna DNA as well as unspecific products were also detected. Remarkably, the majority of the unspecific and untargeted amplifications occurred between species with the highest homology rate: *T. thunnus*/*T. maccoyii* (Fig. 2-3a-II, Fig 2-3b-II), *T. albacares*/*T. tonggol* (Fig. 2-3a-VII), *T. thynnus*/*T. maccoyii*/*T. orientalis* (Fig. 2-3a-I, Fig. 2-3a-III, Fig. 2-3b-I), etc. Although the double mismatch system should theoretically inhibit amplification of alignments between the primer and any DNA template other than the target DNA, in practice, excess of DNA template or DNA polymerase results in partial or marginal amplification. Thus, it was seen that the double mismatch mechanism was not enough to attain complete specificity. In comparison, when using the SSPP and the TSPP simultaneously, effective amplification of the target DNA, with high specificity respect other tuna and other fish and animal DNA, was verified in all cases (Fig. 2-4). The high specificity attained by using two different primer pairs is a result of the interactions between this primers and the DNA polymerase: The DNA polymerase amplifies preferably the TSPP product, as it has no mismatches. The next preferably amplifiable template is the target

SSPP product. With these two high-yielding amplifications in course, the amplification of less favored mismatched and unspecific products is practically avoided and thus, high specificity can be achieved. For the short SSPP of *T. albacores*, marginal amplification of *T. tonggol* DNA was detected sporadically. In this case, increasing the annealing temperature to 60°C (see below) was used to eliminate this problem. All amplifications were conducted under the optimal conditions described in the next section. Specificity trials were further confirmed by investigating the identity of commercial samples through direct sequencing (see below, section 2.3.9).

2.3.4. Optimization

PCR reactions using SSPP and TSPP were carried out to establish the optimal PCR parameters in order to achieve efficiency without compromising specificity: i) Number and length of amplification cycles, ii) PCR temperatures, iii) reactant concentration. The final conditions of the SSPP/TSPP PCR discrimination method are shown in Table 2-3.

It was found that the PCR amplification could proceed readily with a short program of only 25 cycles for long SSPP but that 30 cycles was a better option in case of short SSPP. A larger number of cycles did not improve amplification and in some cases it derived in increased formation of primer dimers and unspecific products. The time length of the denaturalization and extension steps seemed to have little effect in the final amplification so these parameters were set on the minimal time values recommended for PCR amplification of short fragments. In contrast, time and temperature during the annealing step proved to be critical. For long SSPP, an annealing time of 15 s allowed effective amplification but for

short SSPP, it was shown that 20 s rendered better amplifications of the degraded DNA in food products. The profiles of the annealing temperature on the SSPP-PCR are shown in Fig. 2-5. Each primer displays its own range of effective annealing temperatures: The two SSPP of *T. albacares* showed a wide range of effective temperature, been able to amplify the template even at annealing temperatures of 68°C, while the SSPP of *T. obesus* were characterized for having the lowest effective annealing temperature (56°C), which is because these primers had additional mismatches. Since all the SSPP showed efficient amplification in a range of 48°C to 56°C, the annealing temperature was set at 54°C for all species, except for the short SSPP of *Thunnus albacares*. As mentioned previously, for this primer, an annealing temperature of 60°C was required in order to ensure specificity against *T. tonggol*. As a whole, the optimized PCR program is remarkable for been carried out in a relatively short time. Depending on the type of PCR device, the whole reaction can be finished in around 45 min. After the PCR, samples are subjected to electrophoresis for 30 min. This means that the whole process, including a 2-hour digestion during the DNA extraction, can be accomplished in around 4 hours, including preparatory procedures. This analysis time is significantly short compared with other methods, like RFLP, that requires additional time for restriction enzyme digestion (typically from 1 to 2 hours), or sequencing methods (which may take from several hours to days, depending the type of sequencer and the number of samples).

Regarding the primer concentration, it was found that not only the absolute amount of each primer, but also the SSPP/TSPP ratio had a noticeable impact on the efficiency and specificity of the amplification. First, it was seen that the minimal concentration of primers that allowed efficient amplification was 0.3 μM of both TSPP and SSPP (SSPP/TSPP ratio,

1:1). As it can be seen in the example on Fig. 2-6, values lower than (Fig. 2-6a) that derived in loss of specificity, false negatives and low amplification rates both of the SSPP as of the TSPP PCR products. When increasing the concentration of both primers to 0.3 µM (Fig. 2-6b), the TSPP product was generated in a consistent way but the SSPP was still poorly amplified. Ultimately, a SSPP/TSPP ratio of 2:1 (SSPP, 0.6 µM/TSPP, 0.3 µM; Fig. 2-6c) proved to be suitable to achieve efficient amplification of both TSPP and SSPP PCR products. Further increments in the proportion of SSPP/TSPP (Fig. 2-6d) did not produce any significant difference other than a marginal overproduction of primer dimers.

The underlying reasons of the optimization results are undeniably associated to the kinetics of the binding between the two types of primers and the DNA template as well as to interactions between primers. As explained previously, the SSPP achieve specificity through the intelligent use of mismatches between the primers and the DNA templates. Even when the SSPP aligns the target template, there is at least one mismatch near the 3' extreme, which means that the amplification efficiency with SSPP is necessarily inferior to that with TSPP, since the latter has no mismatches with the DNA template. As explained schematically in Fig. 2-7, under similar concentrations of both SSPP and TSPP, the TSPP product is preferentially amplified respect to the SSPP product, thus generating a low signal of the latter (Fig. 2-7a). By adding a double amount of SSPP respect to TSPP, this trend can be overcome (Fig. 2-7b), thus achieving the optimal conditions. Finally, it is worth to mention that other authors (Newton et al., 1989; Little, 1994) have pointed out the importance of the type of mismatch, i.e., how and to which extent is the specificity altered by mismatching one nucleotide or another, by locating the mismatch immediately before the SND or two or three nucleotides upstream, by using deletions or insertions instead of

mismatches, etc. The present research did not deal with these considerations but from now on and in order to improve the method and achieve a better comprehension of the underlying mechanisms, it would be advisable to explore these and other variables.

2.3.5. Detection limit

DNA from unheated samples of the target species was diluted in series and applied to the SSPP/TSPP- PCR discrimination method to test its detection limit. Results are shown in Fig. 2-8. It was observed that in most cases, there was a noticeable amplification with concentrations as low as 30 pg of total DNA. Other quantitative PCR methods, namely real-time PCR techniques, which have been developed for the discrimination of tuna species, have reported detection limits of about 0.1 pg (Kitaoka, et al., 2009). Although compared to such techniques, the current method has a relatively low sensitivity, the results of the analyses of canned tuna and other heat-processed samples (section 2.3.7) showed that, practical identification was feasible even in samples with very low DNA concentration. In addition, it is important to consider that the detection limit is susceptible to change depending on a variety of factors, especially the DNA extraction method. For example, methods for selective extraction of mtDNA would increase the relative amount of mtDNA template, thus allowing relative lower detection limits. Even more, for heat-processed products, it is possible that the total DNA concentration does not correlate with the effective amount of amplifiable DNA. Since it is not possible to evaluate how much the processing will degrade DNA in any particular food system, the present results must be taken only as a reference. As many other analytical methods, application of the present method for routine analysis would require to perform optimization tests of the SSPP/TSPP-PCR discrimination

assay in order to determine its effective parameters for each particular extraction, PCR and electrophoresis system.

2.3.6. Effect of the DNA extraction on the efficiency of the SSPP/TSPP PCR discrimination method

Previous reports (Pardo & Pérez-Villareal, 2004; Chapela et al., 2007) have shown that the DNA extraction method can affect the performance of PCR methods, mostly due to variations in the effective amount of extracted DNA but also because of substances not removed during the extraction that can act as inhibitors of the PCR reaction. In the present study, the QuickGene Tissue DNA kit (Fujifilm) used for DNA extraction method is based in resin columns that adsorb DNA, which is eluted subsequently. Although it is considered that this kind of column methods can achieve great purity, they are not always available. Thus, in order to test the influence of the DNA extraction over the SSPP/TSPP-PCR discrimination method, a phenol-chloroform extraction was used to obtain DNA from canned tuna products, which was then used on the SSPP/TSPP-PCR discrimination method. The concentration of the DNA extracted with this method was somehow comparable to those obtained with QuickGene Tissue DNA kit (average 50 ng/µL in QuickGene vs. average 73 ng/µL in solvent extraction) but with a noticeably decrease in the purity (average OD 280/260 ratio, 2.01 with QuickGene extraction vs. 1.75 with solvent extraction). Nevertheless, the results of the SSPP/TSPP-PCR analysis (not shown) did not differ markedly compared to the analysis of the same samples extracted with the QuickGene Tissue DNA kit.

2.3.7. Protocol for testing of commercial products

Using the results provided here above, the following protocol was developed.

Identification and discrimination of tuna species in food products using SSPP/TSPP-PCR amplification.

1. Sample preparation.

1.1 For unheated product, fresh or frozen fish tissue (*sushi*, *sashimi*, loins, fillet, etc.), and others, remove fat and extraneous material (rice, vegetables, etc.).

1.2. For heat-processed products (canned tuna, *katsuobushi*, *etc.*), remove cover liquids, oil and extraneous materials (herbs, vegetables, etc.). For oily samples, wash the excess of oil with a volume of chloroform: phenol (50:50) and dry the sample over vacuum.

2. DNA extraction. Conduct the DNA extraction using a commercial kit, extraction with chloroform/phenol or any other similar protocol for total DNA extraction of for selective extraction of mtDNA. Measuring the DNA concentration is recommended in order to adjust it in the PCR reaction.

3. SSPP/TSPP-PCR amplification

3.1 Prepare PCR mixtures using different SSPP according to the species that have to be analyzed. For unheated products, use the respective long SSPP. For heat-processed products or products were DNA fragmentation is suspected, use the respective short SSPP. The PCR reaction composition is given in Table 2.3

3.2 Perform the PCR amplification. For unheated products: 1 cycle at 98°C, 30 s; 25 cycles at i) 98°C, 10 s, ii) 54°C, 15 s, iii) 72°C, 20 s. For heat-processed products: 1

cycle at 98°C, 30 s; 30 cycles at i) 98°C, 10 s, ii)54°C (60°C in case of analyzing *Thunnus albacares*), 20 s; 72°C, 30 s.

3.3 Analyze PCR products in electrophoresis in SYBR Green-stained 2.5%-3.0% agarose gel. Other electrophoresis systems can be used too, but SYBR Green is recommended due to its sensitivity.

4. Interpretation of results.

4.1 According to the results of the electrophoresis, there are 4 possible cases and interpretations:

A. No TSPP product (89 bp) and no SSPP product (characteristic length for each species, refer to Table 2-2). Interpretation: There is not DNA from tuna species in the sample or it is present in a concentration below the detection limit of the method.

B. TSPP product appears at 89 bp. No SSPP product. Interpretation: There is DNA of a *Thunnus* species but not of the target species. Repeat the test with a different SSPP.

C. No TSPP product but the SSPP product appears. Interpretation: Possible error during the preparation of the PCR reaction. Possible DNA from an unknown species having annealing sites for the SSPP and generating a PCR product of the same length that the SSPP amplicon. Inconclusive results.

D. TSPP product and SSPP product. Interpretation: The sample contains DNA from the target tuna.

A flowchart of the analysis method is shown in Fig. 2-9. For quantitative analysis of samples, see Chapter 3.

The SSPP/TSPP-PCR discrimination method was applied under the protocol here above described to analyze commercial products. The results are shown in Fig. 2-10 and summarized in Table 2-4.

It was possible to determine the identity of unheated samples (Fig. 2-10a) and no discrepancies were found respect to the expected species. During the extraction, it was observed that in fat-rich samples, the concentration of the total DNA tended to be less than average but after standardizing the amount of total DNA template (30 ng), identification could be done without problem. Most of the analyzed samples were muscle tissue but in a couple of cases, a fermented Japanese product made of fish gut was also analyzed. The product, called *shuto*, was included as not-heated as it is not cooked. The fermented guts are kept under a cover liquid made of salt and condiments and for this reason, the final DNA concentration was lower than average. Even so, DNA identification on gut tissue could be done readily, which shows that the SSPP/TSPP-PCR discrimination method can be used in DNA from a variety of sources other than muscle, provided that the extraction method is suitable for them.

For the analysis of heat-processed products (Fig. 2-10b), total DNA template was standardized and used as 20 ng whenever possible but in some cases, the extracted DNA was too diluted to achieve this level. In such cases, 5 μL of DNA template were used, regardless its concentration. By using the SSPP/TSPP-PCR method, identification was achieved in the majority of the cases. Particularly, all canned products processed under normal conditions (assuming a commercial sterilization of 121°C for 45 min or 116°C for 75 min), as well as in some food products such as tuna crocket, *maguro kakuni* (tuna seasoned meat), *katsuobushi* (dried skipjack flakes) and tuna pate, could be analyzed satisfactorily. In one

case (sample #45), two different species were detected. In other two cases (samples #5, #9), the species reported in the label was not detected but in turn, a different one was found. The later samples may be considered cases of mislabeling or even of intentional replacement. There were heat-processed products where identification could not be accomplished: A sample of a frozen prepared food of tuna and egg (sample #52), a sample of chicken and tuna *tsukune* (a type of crocket, sample #53), and a sample of dried tuna *kamaboko* (dry paste of fish meat, sample #61). The failure on the identification of these products is associated to impairment of the DNA extraction due to the presence of flour and other components, yielding very low amounts of total DNA (<2 ng/µL, compared to the average 25 ng/µL for processed products). Extraction with phenol-chloroform did not result in higher DNA concentrations and discrimination of DNA extracted with this method was not possible either.

On the other side, analyses of canned tuna in oil and in a variety of cover liquids could also be accomplished despite occasional low DNA concentrations, which is a great advantage respect to other methods which have reported poor results or failure because of oil and condiments in the food matrix. Taken altogether, the results show that although the SSPP/TSPP-PCR discrimination method still has limitations imposed by the characteristics of the food matrix and by the processing conditions, it is a method that can be applied to a wide variety of foods processed under standard conditions with efficiency and accuracy.

2.3.8 Verification of the SND and analyses of haplotypes

All the unheated commercial samples were subjected to direct sequencing with the corresponding primers used during the sequencing of the complete mtDNA, as shown in

Table 2-5, in order to confirm the presence of the SND and to evaluate polymorphisms in the adjacent areas. The comparison of sequences for each species and the summary of haplotypes can be seen in Fig. 2-11 and Table 2-6, respectively. Although there were a few single substitutions detected among all species, most of these can be considered random spontaneous mutations in isolated individuals. However, in the case of *T. obesus* and *K. pelamis*, a significant number of transition-type substitutions were found (3.84% and 5.41% of the total studied DNA sequence, respectively). In a lesser extent, substitutions were also relatively frequent in specimens of *T. alalunga* and *T. albacares* (2.01% and 1.09%, respectively). These results suggest the existence of subpopulations with particular haplotypes within the same species. Specifically, the intraspecies variability of *T. obesus* that is implied in the results of this research agrees with the presence of subpopulations of *T. obesus* that have been described worldwide (Chow el al., 2003). This is supported by the fact that *K. pelamis*, *T. alalunga* and *T. albacares*, all of them species widely distributed around the world and with different subpopulations typified, also showed a relative large number of substitutions. Some of the detected polymorphic sites in these species are comparatively close to the chosen SND and in the case of the reverse SSPP of *K.pelamis*, there is one polymorphic site inside the primer sequence. Even so, since all samples of *T. obesus, T. alalunga*, *T. albacares* and *K.pelamis* could be identified accurately, it can be regarded that these polymorphic sites do not affect the efficiency of the method. For the remaining species, the results showed that the SND used in the SSPP/TSPP-PCR discrimination method display the characteristic haplotype in all specimens and are located in areas with low intraspecies variability, and thus they can be considered as stable DNA markers.

2.3.9 Verification of unheated commercial samples by direct sequencing

In order to verify the accuracy of the SSPP/TSPP-PCR discrimination method, unheated samples were identified by direct sequencing using a fragment of 940 bp from the flanking sequence between the ATP6 and the COIII regions as reported by Chow & Inoue (1993). The phylogenetic tree obtained from the sequencing confirmed the identity of all unheated samples (Fig. 2-12), thus validating the accuracy of the SSPP/TSPP-PCR discrimination method. Incidentally, the same phylogenetic analysis showed the existence of 2 subpopulations of *Thunnus obesus*, which have been reported previously in studies of the same ATP6-COIII area (Chow et al., 2003).

2.3.10 Analysis of heat-processed samples by RFLP

Canned tuna and other heat-processed samples cannot be analyzed by direct sequencing; since this kind of samples do not contain DNA sequences long enough to be analyzed with this method reliably. Thus, for verification of the SSPP/TSPP-PCR discrimination method in canned samples, a RFLP method reported previously was attempted. Unfortunately, this method was developed only for *Thunnus thynnus*, *T. alalunga*, *T. albacares*, *T. obesus* and *Katsuwonus pelamis*, so analysis of the other 3 species could not be done. Even more, initial PCR amplification using Cb126 and Cb146 primers was unsuccessful in samples of *T. obesus* and in many samples of *T. albacares* and *T. alalunga*. The reasons of this failure seem to be associated with the low concentrations of DNA template altogether with a low affinity of the primers for the selected target region. Ultimately, the RFLP analysis could be conducted only in 14 samples out of 68. The results

of these 14 RFLP analyses, shown in Fig. 2-13, agreed with the results obtained with the SSPP/TSPP-PCR discrimination method.

2.3.11 General conclusions

A method for discrimination of tuna species in food products based in the allele-specific PCR was developed and characterized. The method consists in the simultaneous use of two primers: A species-specific primer pair (SSPP) and a *Thunnus*-specific primer pair (TSPP). The method displayed great specificity and it could be applied successfully to the analysis of a variety of tuna products, both raw and processed. The method was standardized and it was shown that its detection limit was comparable to that of more sophisticated techniques. Compared to direct sequencing, the SSPP/TSPP-PCR method was as accurate as this method but faster, as it does not need further purifications or time-consuming labeling reactions. Also, the present method could be applied to mixtures of species, that usually cannot be analyzed by direct sequencing. Unlike RFLP analysis, the SSPP/TSPP-PCR discrimination method can be applied to all the main commercial tuna species and their mixtures and it is also technically easier to perform, as it does not require post-PCR treatment other than a conventional electrophoresis. Compared to other methods that require minimum post-PCR processing, such as SSCP, the current method also exhibits advantages in terms of reproducibility. Even more, the SSCP/TSPP-PCR discrimination method can be regarded as more resistant to the effect of spontaneous intraspecies mutations. Considering that the present method relies on the presence or absence of only one key SND in the 3' extreme of the primer, regardless the changes of the sequence between primers, it is highly unlikely that the precise key SND mutates in any individual. In contrast, methods such as

RFLP and SSCP can produce inconsistent results if even only one mutation occurs in any site of the sequence of the PCR product, as this mutation can generate or delete restriction sites or modify critically the electrophoresis migration pattern, thus altering the performance of RFLP or SSCP methods, respectively.

The use of the TSPP as internal control allows detecting any abnormality in the PCR performance. The use of two separate sets of primers provides flexibility to the analyst: For heat-processed products, only short SSPP can be used but for raw, frozen or slightly heated samples, both short and long SSPP can be used. The advantage of the long SSPP is that PCR amplification using these primers is, to some extent, more refractive to changes in the PCR conditions, such as slight changes in the reactant concentration and variations in the temperature program due to the PCR device. Also, long SSPP generate fragments that can be easily distinguished from the TSPP product, thus enabling the use of agarose gels of lower concentrations and shorter electrophoresis, allowing modest but significant savings of resources and time.

Finally, it is worth to indicate that the SSPP/TSPP-PCR method can be used to analyze other species. Because of the high homology rate among *Thunnus* species, multiplex PCR, i.e., the simultaneous use of two or more SSPP, is not possible under the conditions of the current method, as hybridizations may take place, generating false positives. In turn, by extending the method to other species with DNA sequences different enough to that of the *Thunnus* spp, it may be possible to identify them at the same time. This may be the case, for example, of the differentiation between escolar (*Lepidocybium flavobrunneum*) and white tuna (*T. alalunga*) or even for the detection of cross contamination of chicken or salmon in canned tuna.

Table 2-1a. List of commercial unheated tuna samples.

Thunnus thynnus	
N01	Frozen loin, Croatia (cultured)
N02	Frozen loin, Croatia (cultured)
N03	Frozen loin, Croatia (cultured)
N04	*Sushi*, Ireland (wild)
N05	*Sushi*, Turkey (cultured)
N06	*Sushi*, Greece (cultured)
N07	*Sushi*, Morocco (cultured)
N08	*Sushi*, Cultured
N09	*Sushi*, Cultured
N10	*Sushi*, Cultured
N11	*Sushi*, Spain (cultured)
Thynnus maccoyii	
S01	*Sashimi*, Australia (wild)
S02	Frozen loin, Australia (wild)
S03	*Sushi*, Australia (wild)
S04	*Sashimi*, Australia (wild)
S05	*Sashimi*, New Zealand (wild)
S06	*Sashimi*, Indonesia (wild)
S07	Frozen loin, Tasmania (wild)
Thunnus orientalis	
P01	*Sashimi*, Mexico (wild)
P02	*Sashimi*, Mexico (wild)
P03	*Sushi*, Korea (wild)
P04	*Sashimi*, Kagoshima, Japan (cultured)
P05	*Sashimi*, Ishikawa, Japan (cultured)
P06	*Sashimi*, Tottori, Japan (cultured)
Katsuwonus pelamis	
K01	*Tataki*, Japan
K02	*Tataki*, Japan
K03	*Sashimi*, Japan
K04	*Sashimi*, Korea
K05	*Tataki*, Japan
K06	*Tataki*, Taiwan
K07	*Shuto* (fermented guts), Japan

Thunnus obesus	
X01	*Sashimi*, Korea
X02	Frozen loin, Pacific Ocean
X03	*Sashimi*, Japan
X04	*Sashimi*, Atlantic Ocean (Taiwan)
X05	*Sashimi*, Atlantic Ocean (Taiwan)
X06	*Sashimi*, Indian Ocean (Taiwan)
X07	*Sashimi*, Indian Ocean (Taiwan)
X08	*Sashimi*, Indian Ocean (Japan)
X09	*Sashimi*, Indian Ocean (Taiwan)
X10	*Sashimi*, Indian Ocean (Taiwan)
X11	*Sashimi*, Indian Ocean (Taiwan)
X12	*Sashimi*, Japan
X13	*Sashimi*, Japan
X14	*Sashimi*, Japan
X15	*Sushi*, Japan
X16	*Sashimi*, Japan
X17	*Sushi*, Japan
X18	*Sashimi*, Japan
X19	*Sashimi*, Japan
Thunnus albacares	
Y01	*Sashimi*, Chiba, Japan
Y02	*Sashimi*, Miyagi, Japan
Y03	Frozen loin, Shizuoka, Japan
Y04	*Sashimi*, Indian Ocean
Y05	*Sashimi*, Australia
Y06	*Sashimi*, Indonesia
Y07	*Shuto* (fermented guts)
Thunnus alalunga	
B01	*Sashimi*, Shizuoka, Japan
B02	*Sashimi*, Chiba, Japan
B03	*Sashimi*, Pacific Ocean
B04	*Sashimi*, Pacific Ocean
B05	*Sashimi*, Pacific Ocean
B06	*Sashimi*, Okinawa, Japan

Table 2-1b. List of commercial not heat-treated samples.

Labeled as *T. albacares*	
1	Seachicken (Japan)
2	Light tuna in Canola Oil (Japan)
3	Light tuna in Water (Japan)
4	Tuna in water "Herdez" (Mexico)
5	Tuna in water "Ancla" (Mexico)
6	Tuna in water "Nair" (Mexico)
7	Tuna in water "Dolores" (Mexico)
8	Tuna in water "Tuny" (Mexico)
9	Tuna Salad "Marina Azul" (Mexico)
10	Canned Tuna "Hanamasa" (Japan)
11	Tuna in water "Imperial" (Holland)
Labeled as *T. tonggol*	
12	Tuna in water "Tonnikaala" (Estonia)
13	*T. tonggol* in water*
Labeled as *T. maccoyii*	
14	Southern Bluefin*
Labeled as *T. thynnus*	
15	Atlantic Bluefin*
Labeled as *K.pelamis*	
16	Mild Seachicken in Oil (Japan)
17	Tuna & Double beans (Japan)
18	Tuna in oil (Japan)
19	Tuna in oil "Blik" (Latvia)
20	Tuna in water "Cloverleaf" (Thailand)
21	Tuna & Corn "Ihada" (Japan)
22	Flavored tuna "Maruha" (Japan)
23	Tuna in water "Toonjin moot" (Holland)
24	Tuna in water "Tonnikaala" (Estonia)
25	Light tuna (Japan)
26	Flavored skipjack (Japan)
27	Non oil tuna (Japan)
28	Tuna pate (Japan)
29	Skipjack flakes (Japan)
30	Skipjack flakes (*katsuobushi*) (Japan)
31	Skipjack flakes (*katsuobushi*) (Japan)

Labeled as *T. alalunga*	
32	Seachicken "Fancy" in oil (Japan)
33	White tuna SSK (Japan)
34	Canned tuna "Carrefour" (France)
35	Canned albacore "Gold Seal" (France)
36	Tuna in water "Geisha" (Japan)
Labeled as *T. orientalis*	
37	Pacific Bluefin*
Labeled as *T. obesus*	
38	Tuna Steak (Japan)
39	Bigeye in water (Japan)
40	Bigeye in oil (Japan)
41	*T. obesus* in water*
42	*T. obesus* in oil*
Undeclared species	
43	Tuna meal "Hagoromo" (Japan)
44	Tuna burguer, frozen food (Japan)
45	Tuna in water "Per Krestok" (Russia)
46	Tuna & veggies "Maruha" (Japan)
47	Light tuna in Chunks (Japan)
48	Tuna in oil "Kyokuyo" (Japan)
49	Tuna in oil "Morska" (Russia)
50	Tuna in water "Stucken" (Holland)
51	Light tuna "Honino" (Japan)
52	Tuna & Egg Frozen food (Japan)
53	Tuna *Tsukune* (Japan)
54	Tuna in water "Red eagle" (Taiwan)
55	Tuna and Fermented Bamboo (Taiwan)
56	Tuna in Spring water (Taiwan)
57	PetLife Sardine and Tuna (Japan)
58	Tuna in mirin souce (Japan)
59	Tuna in red wine (Japan)
60	Tuna *kakuni* (Japan)
61	Tuna *kamaboko* (Japan)
62	Petfood "AIXIA" (Japan)
63	Tuna Aizu (Tuna and chicken)
64	Light Tuna Petfood "MIAU" (Japan)
65	Tuna in olive oil "San Cusumano" (Italia)
66	Tuna in olive oil "LUKUS" (Italia)
67	Tuna in olive oil "Consorcio" (Spain)
68	Tuna in olive oil "As do mar" (Italia)

*Simulation of canning by autoclaving at 121°C during 40 min.

Table 2-2 Species-specific primers and *Thunnus*-specific primers..

Name	Sequence	Primer Length	Amplicon length	Annealing temperature	mtDNA region
	Thunnus thynnus				
TNN5-02L	TTT TTT TAT AAC CGA GTC GGG GAT ATC GCG	30	110	≤64°C	ND5
TNN5-02H	GAA GTT TTT AGC GGT TAC GAA CAT TTG C	28			
TNN5-01L	ATA ACC GAG TCG GGG ATA TCG CG	23	440	≤60°C	
TNN5-H12924	AAG GCC AAT TGT TAC TAT CA	20			
	Thunnus maccoyii				
TMN5-01L	TAT CTT CTC ACT GCT AGC ATA CCA C	25	161	≤64°C	ND5
TMN5-01H	TAA TAG CTT CGG CAC CCT CGT AC	23			
TMN5-01L	TAT CTT CTC ACT GCT AGC ATA CCA C	25	276	≤64°C	
TMN5-H12266	AGG ATT GAT CAT GTT ACG TAG AG	23			
	Thunnus orientalis				
TPN6-04L	GGT ACC CAA GAA ATC TCG TCT	21	107	≤60°C	ND6
TPN6-01H	ACT TGA GGT AGT CCA GCT ATT GAG C	25			
TPN6-L13704	CAA TGC CCA GCA AGG TAT AAT	21	363	≤60°C	
TPN6-01H	ACT TGA GGT AGT CCA GCT ATT GAG C	25			
	Thunnus alalunga				
TAN5-03L	ATG TTC CAA CTC TTT ATT GGG TGA	24	131	≤60°C	ND5
TAN5-03H	ATA TCC CTC ACT CGG TTG TTC	21			
TAN5-L12182	CCT TTG ATA TCA ATA TTA GCT AG	23	322	≤56°C	
TAN5-03H	ATA TCC CTC ACT CGG TTG TTC	21			
	Thunnus obesus				
TON4-03L	AGG CTT TAC GGA GGC CCA TAC	21	142	≤56°C	ND4
TON4-03H	AAG TGG CAG TAC TAT CTG CAG G	22			
TON4-03L	AGG CTT TAC GGA GGC CCA TAC	21	440	≤56°C	
TON4-H11716	TTA GCT CGG GTT TGA GGA TA	20			
	Thunnus albacares				
TYN4-02L	TCC TCA TGA TTG CCC ACG GAA TT	23	133	≤68°C	ND4
TYN4-02H	CAT GTT GTT ATA AGG GGC ATC	21			
TYN4-02L	TCC TCA TGA TTG CCC ACG GAA TT	23	420	≤68°C	
TYN4-H11716	TTA GCT CGG GTT TGA GGA TA	20			
	Thunnus tonggol				
TLC1-01L	ACT TGC AAC CCT TCA CGA G	19	110	≤60°C	COI
TLC1-01H	TGG CTA GGA CAA TAC CTG TTA CT	23			
TLC1-01L	ACT TGC AAC CCT TCA CGA G	19	387	<60°C	
TLC1-H6830	GTG TTT CAA AGG GTG TAG GGA	21			
	Katsuwonus pelamis				
KPN5-02L	CTT CTC ATC ACC TCT AAC ATC GTT C	25	137	≤56°C	ND5
KPN5-03H	TGT TTG CTT GTT AGG GAG TCT	21			
L13153	ATC GAA GCA CTA AAC ACA TCA CA	23	347	≤64°C	
KPN5-04H	TGT TTG CTT GTT AGG GAG CCT	21			
	Thunnus-specific primer pair (TSPP)				
ThGen 3L	AGC TAC CCT TGC CCT TCT A	19	98	≤64°C*	16S
ThGen 3H	AAG GCC CTT GGA CTA TGT	18			

*Intentional mismatches are shaded. SND are underlined

Table 2-3. Final conditions of the SSPP/TSPP-PCR identification method

Reaction composition	DNA template	2 µL (ca. 1 ng)
	Buffer 5x ($MgCl_2$)	2.5 µL (2.5 mM)
	dNTP	1.0 µL (0.1 mM each)
	SSPP, 20µM	0.75 µL (final, 0.6 µM)
	TSPP, 20µM	0.375 µL (final, 0.3 µM)
	ExTaq DNA polymerase	0.2 µL
	Nuclease-free H_2O	To 25 µL
PCR program Long SSPP	1) 1 cycle	98°C, 30 s
	2) 25 cycles	98°C, 10 s
		54°C, 15 s
		72°C, 20 s
PCR program Short SSPP	1) 1 cycle	98°C, 30 s
	2) 30 cycles	98°C, 10 s
		54°C, 20 s *
		72°C, 30 s

*In the case of the short SSPP for *Thunnus albacares*, the annealing temperature is 60°C

Table 2-4a. Results of the identification of commercial unheated samples with the SSPP/TSPP-PCR discrimination method

Sample	Identification	Sample	Identification	Sample	Identification
N01	*T. thynnus*	P06	*T. orientalis*	X17	*T. obesus*
N02	*T. thynnus*	B01	*T. alalunga*	X18	*T. obesus*
N03	*T. thynnus*	B02	*T. alalunga*	X19	*T.obesus*
N04	*T. thynnus*	B03	*T. alalunga*	Y01	*T. albacares*
N05	*T. thynnus*	B04	*T. alalunga*	Y02	*T. albacares*
N06	*T. thynnus*	B05	*T. alalunga*	Y03	*T. albacares*
N07	*T. thynnus*	B06	*T. alalunga*	Y04	*T. albacares*
N08	*T. thynnus*	X01	*T. obesus*	Y05	*T. albacares*
N09	*T. thynnus*	X02	*T. obesus*	Y06	*T. albacares*
N10	*T. thynnus*	X03	*T. obesus*	Y07	*T. albacares*
N11	*T. thynnus*	X04	*T. obesus*	T01	*T. tonggol*
S01	*T. maccoyii*	X05	*T. obesus*	T02	*T. tonggol*
S02	*T. maccoyii*	X06	*T. obesus*	T03	*T. tonggol*
S03	*T. maccoyii*	X07	*T. obesus*	T04	*T. tonggol*
S04	*T. maccoyii*	X08	*T. obesus*	K01	*K.pelamis*
S05	*T. maccoyii*	X09	*T. obesus*	K02	*K.pelamis*
S06	*T. maccoyii*	X10	*T. obesus*	K03	*K.pelamis*
S07	*T. maccoyii*	X11	*T. obesus*	K04	*K.pelamis*
P01	*T. orientalis*	X12	*T. obesus*	K05	*K.pelamis*
P02	*T. orientalis*	X13	*T. obesus*	K06	*K.pelamis*
P03	*T. orientalis*	X14	*T. obesus*	K07	*K.pelamis*
P04	*T. orientalis*	X15	*T. obesus*		
P05	*T. orientalis*	X16	*T. obesus*		

*For the description of the samples, see Table 2-1a

Table 2-4b. Results of the identification of commercial heat-processed samples with the SSPP/TSPP-PCR discrimination method

#	Expected	Detected	#	Expected	Detected	#	Expected	Detected
1	*T.albacares*	*T.albacares*	24	*K.pelamis*	*K.pelamis*	47	---	*K.pelamis*
2	*T.albacares*	*T.albacares*	25	*K.pelamis*	*K.pelamis*	48	---	*T. albacares*
3	*T.albacares*	*T.albacares*	26	*K.pelamis*	*K.pelamis*	49	---	*K.pelamis*
4	*T.albacares*	*T.albacares*	27	*K.pelamis*	*K.pelamis*	50	---	*K.pelamis*
5	*T.albacares*	***K.pelamis***	28	*K.pelamis*	*K.pelamis*	51	---	*T. albacares*
6	*T.albacares*	*T.albacares*	29	*K.pelamis*	*K.pelamis*	52	---	**NA**
7	*T.albacares*	*T.albacares*	30	*K.pelamis*	*K.pelamis*	53	---	**NA**
8	*T.albacares*	*T.albacares*	31	*K.pelamis*	*K.pelamis*	54	---	*K.pelamis*
9	*T.albacares*	***K.pelamis***	32	*T. alalunga*	*T. alalunga*	55	---	*T.alalunga*
10	*T.albacares*	*T.albacares*	33	*T. alalunga*	*T. alalunga*	56	---	*T. albacares*
11	*T. albacares*	*T.albacares*	34	*T. alalunga*	*T. alalunga*	57	---	*K.pelamis*
12	*T.tonggol*	*T.tonggol*	35	*T. alalunga*	*T. alalunga*	58	---	*T. albacares*
13	*T.tonggol*	*T.tonggol*	36	*T. alalunga*	*T. alalunga*	59	---	*T. albacares*
14	*T.maccoyii*	*T.maccoyii*	37	*T. orientalis*	*T. orientalis*	60	---	*T. albacares*
15	*T. thynnus*	*T. thynnus*	38	*T.obesus*	*T.obesus*	61	---	**NA**
16	*K.pelamis*	*K.pelamis*	39	*T.obesus*	*T.obesus*	62	---	*T. albacares*
17	*K.pelamis*	*K.pelamis*	40	*T.obesus*	*T.obesus*	63	---	*T. albacares*
18	*K.pelamis*	*K.pelamis*	41	*T.obesus*	*T.obesus*	64	---	*T.alalunga*
19	*K.pelamis*	*K.pelamis*	42	*T.obesus*	*T.obesus*	65	---	*T. albacares*
20	*K.pelamis*	*K.pelamis*	43	---	*T. albacares*	66	---	*T. albacares*
21	*K.pelamis*	*K.pelamis*	44	---	*T. albacares*	67	---	*T. albacares*
22	*K.pelamis*	*K.pelamis*	45	---	***K. pelamis*** ***T. albacares***	68	---	*T. albacares*
23	*K.pelamis*	*K.pelamis*	46	---	*T. albacares*			

Sample number. Refer to Table 2-1b
NA Not achieved. The species could not be identified with the method.

Table 2-5. Generic primers used for verification of SND in the SSPP/TSPP-PCR discrimination method

Species	Primer
Thunnus thynnus	L11934/H12924
Thunnus maccoyii	L11934/H12924
Thunnus orientalis	L13704/H14543
Thunnus alalunga	L11934/H12924
Thunnus obesus	L10780/H11716
Thunnus albacares	L10780/H11716
Thunnus tonggol	L5895/H6673
Katsuwonus pelamis	L13153/H14089

Table 2-6. Haplotypes characterized in commercial samples of not heat-treated tuna products during the verification of SND in the SSPP/TSPP-PCR discrimination method

Species	NCBI GenBank accession number of the haplotype (Number of specimens)	Polymorphic sites*
Thunnus thynnus	JN635374 (3) JN635375 (2) JN635376 (1) JN635377 (2) JN635378 (2) JN635379 (1)	5/830 (0.60)
Thunnus maccoyii	JN635370 (1) JN635371 (2) JN635372 (2) JN635373 (1)	5/859 (0.58)
Thunnus orientalis	JN635380 (5) JN635381 (1)	2/746 (0.27)
Thunnus alalunga	JN635363 (2) JN635364 (1) JN635365 (1) JN635366 (1) JN635367 (1)	17/842 (2.01)
Thunnus obesus	JN635390 (2) JN635391 (1) JN635392 (1) JN635393 (2) JN635394 (1) JN635395 (2) JN635396 (1) JN635397(2) JN635398 (1) JN635399 (1) JN635400 (3) JN635401 (2)	28/729 (3.84)
Thunnus albacares	JN635382 (1) JN635383 (1) JN635384 (1) JN635385 (1) JN635386 (1) JN635387 (1)	8/734 (1.09)
Thunnus tonggol	JN635368 (3) JN635369 (1)	1/649 (0.15)
Katsuwonus pelamis	JN635354 (1) JN635355 (1) JN635356 (1) JN635357 (1) JN635358 (1) JN635359 (1) JN635360 (1)	39/721 (5.41)

* The first figure indicates the number of polymorphic sites, i.e., the sites where two or more different nucleotides were found in specimens of the same species. The second figure indicates the total length in base pairs of the amplified DNA fragment. The number in parenthesis indicates the percentage of polymorphic sites

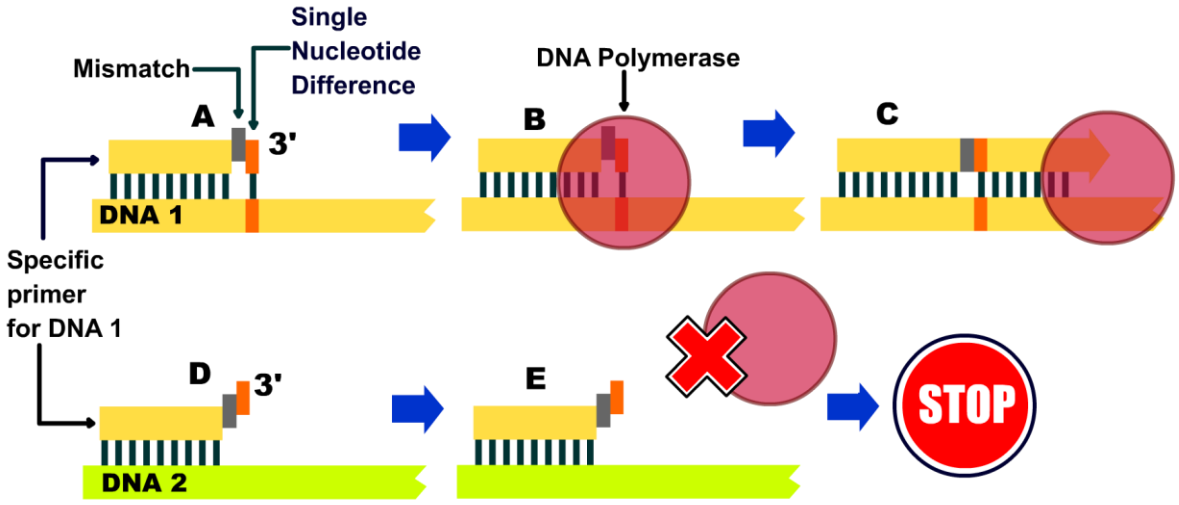


Fig 2-1. Double-mismatch mechanism used to give specificity to the SSPP. **A**. The SSPP binds to the target DNA (DNA 1). **B**. The DNA polymerase recognizes the alignment. **C**. The target DNA is amplified. **D**. The primer aligns imperfectly to a different DNA template (DNA 2). **E**. The DNA polymerase does not bind and amplification of the extraneous template is inhibited.

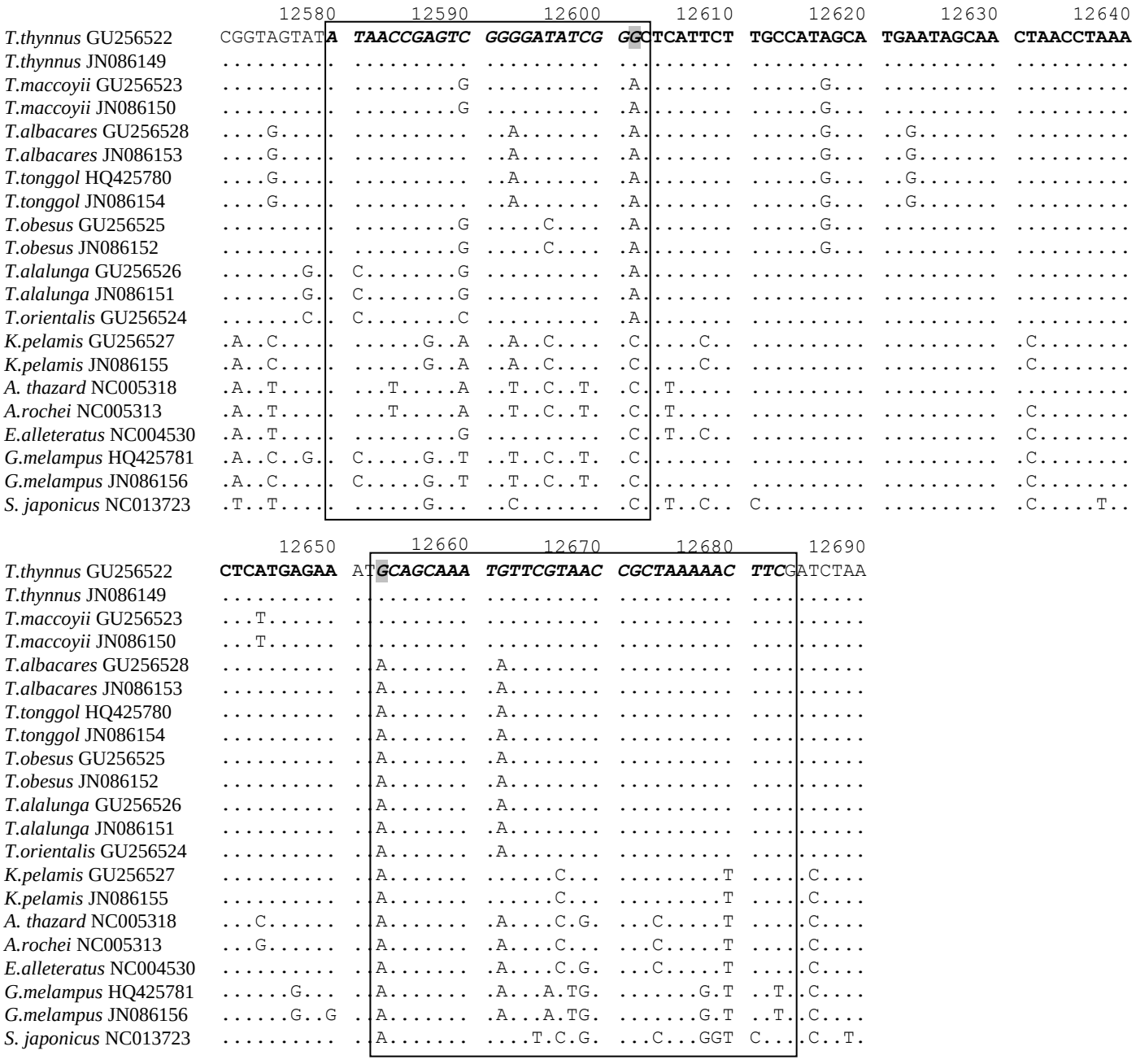


Fig. 2-2a. Alignment of mtDNA sequences showing the SND used for design of the short SSPP for *Thunnus thynnus*. In italics, the SSPP sequence. Shaded, the SND at the 3' extreme of the SSPP

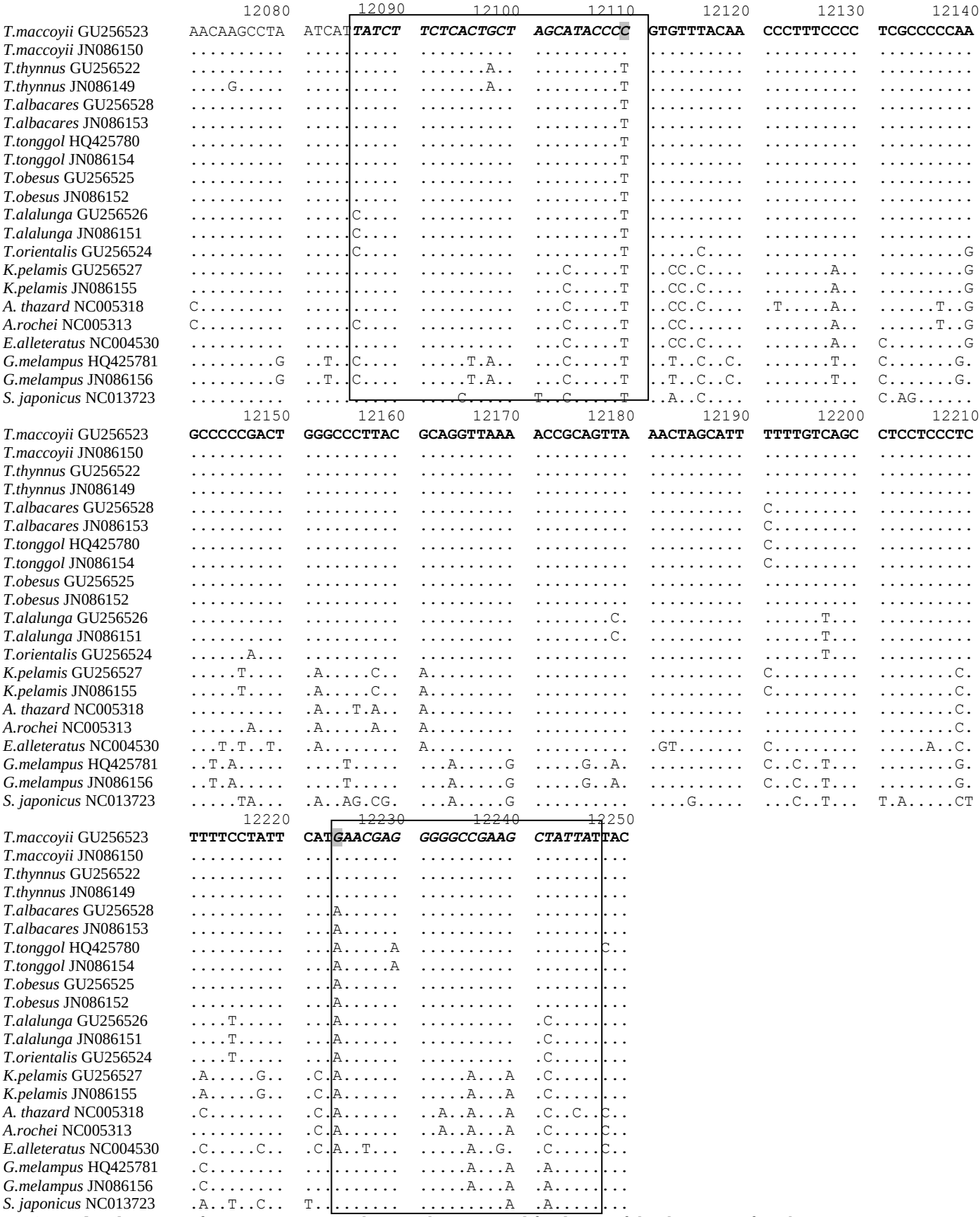


Fig. 2-2b. Alignment of mtDNA sequences showing the SND used for design of the short SSPP for *Thunnus maccoyii*. In italics, the SSPP sequence. Shaded, the SND at the 3' extreme of the SSPP

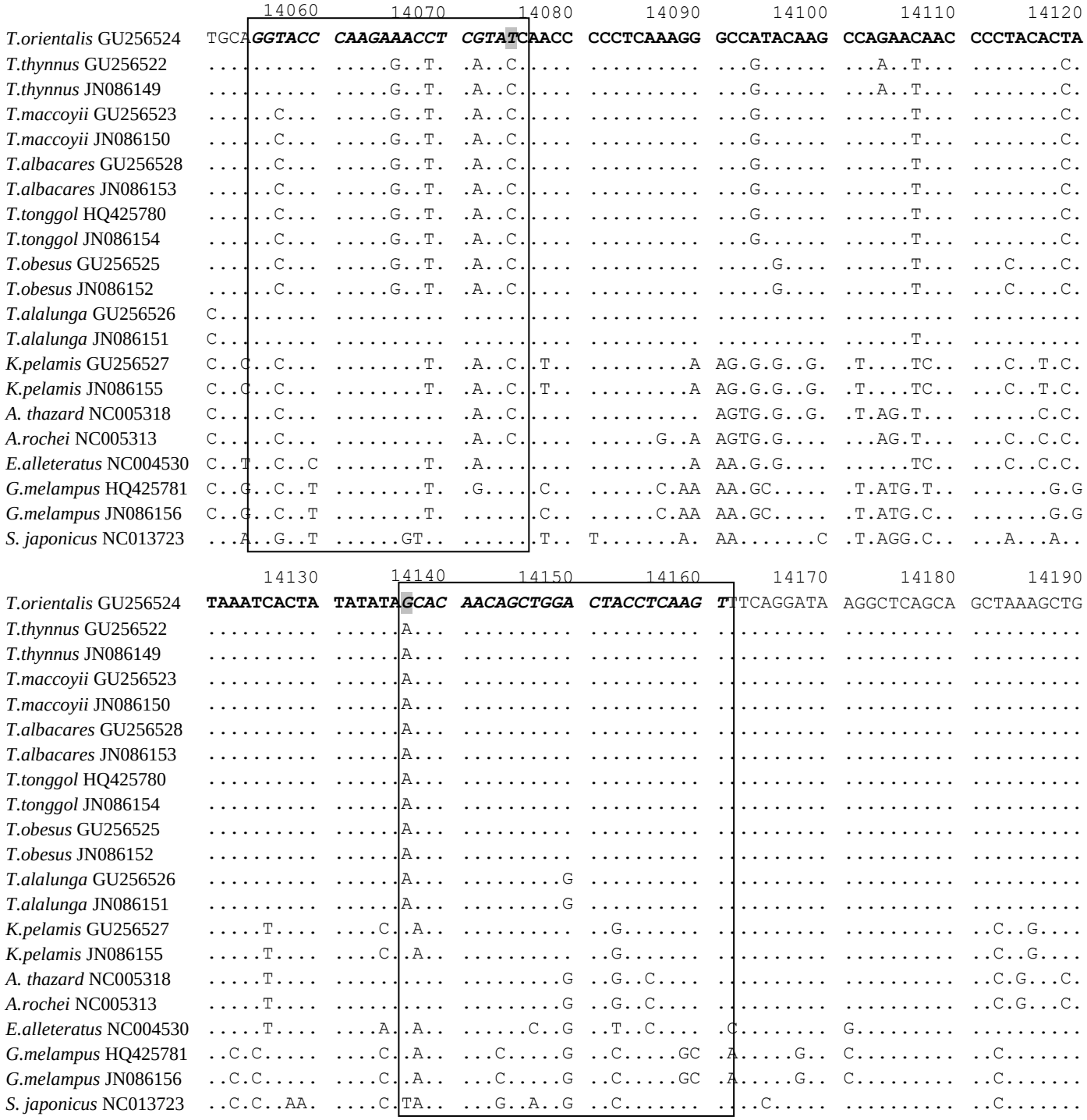


Fig. 2-2c. Alignment of mtDNA sequences showing the SND used for design of the short SSPP for *Thunnus orientalis*. In italics, the SSPP sequence. Shaded, the SND at the 3' extreme of the SSPP

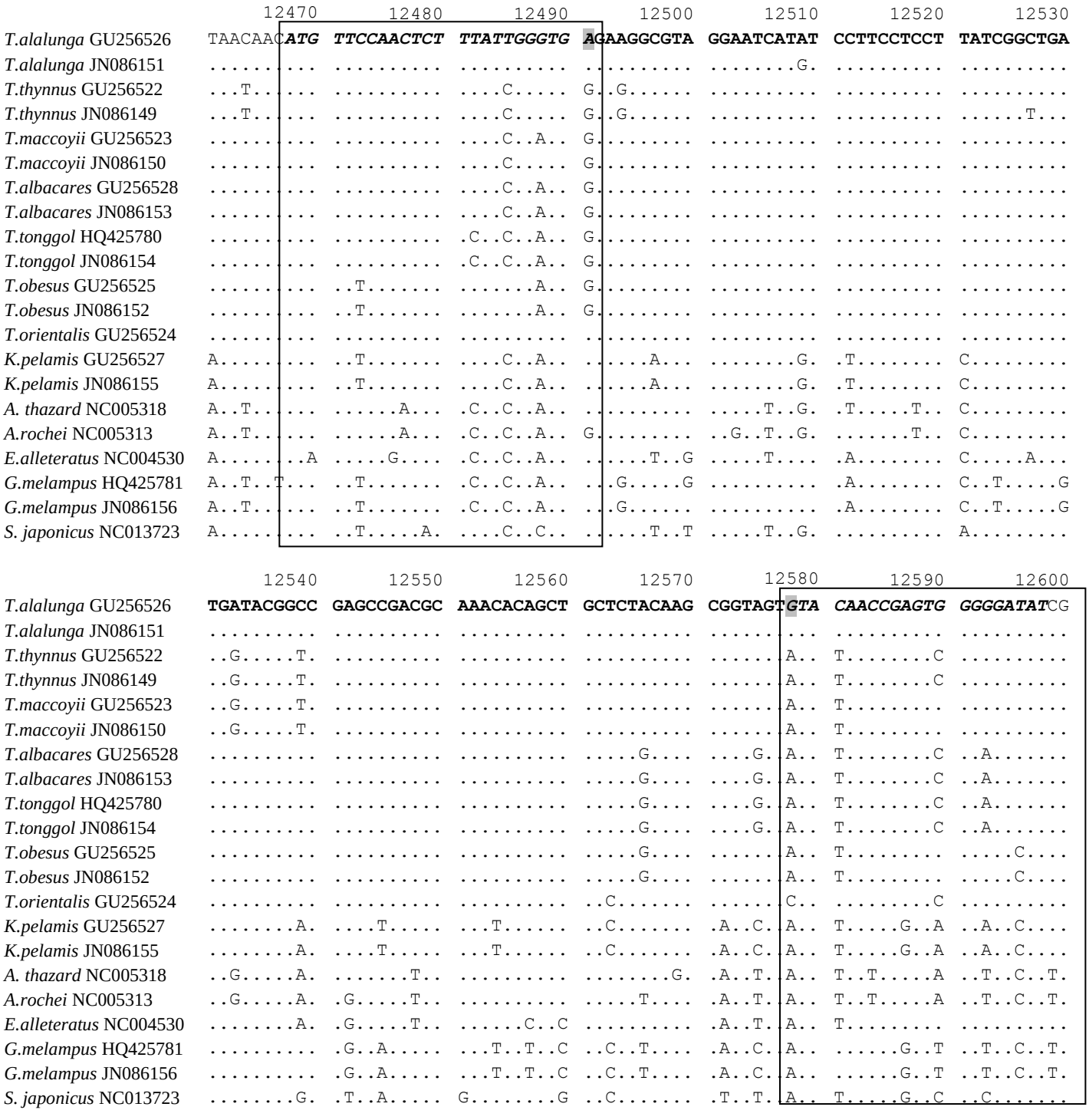


Fig. 2-2d. Alignment of mtDNA sequences showing the SND used for design of the short SSPP for *Thunnus alalunga*. In italics, the SSPP sequence. Shaded, the SND at the 3' extreme of the SSPP

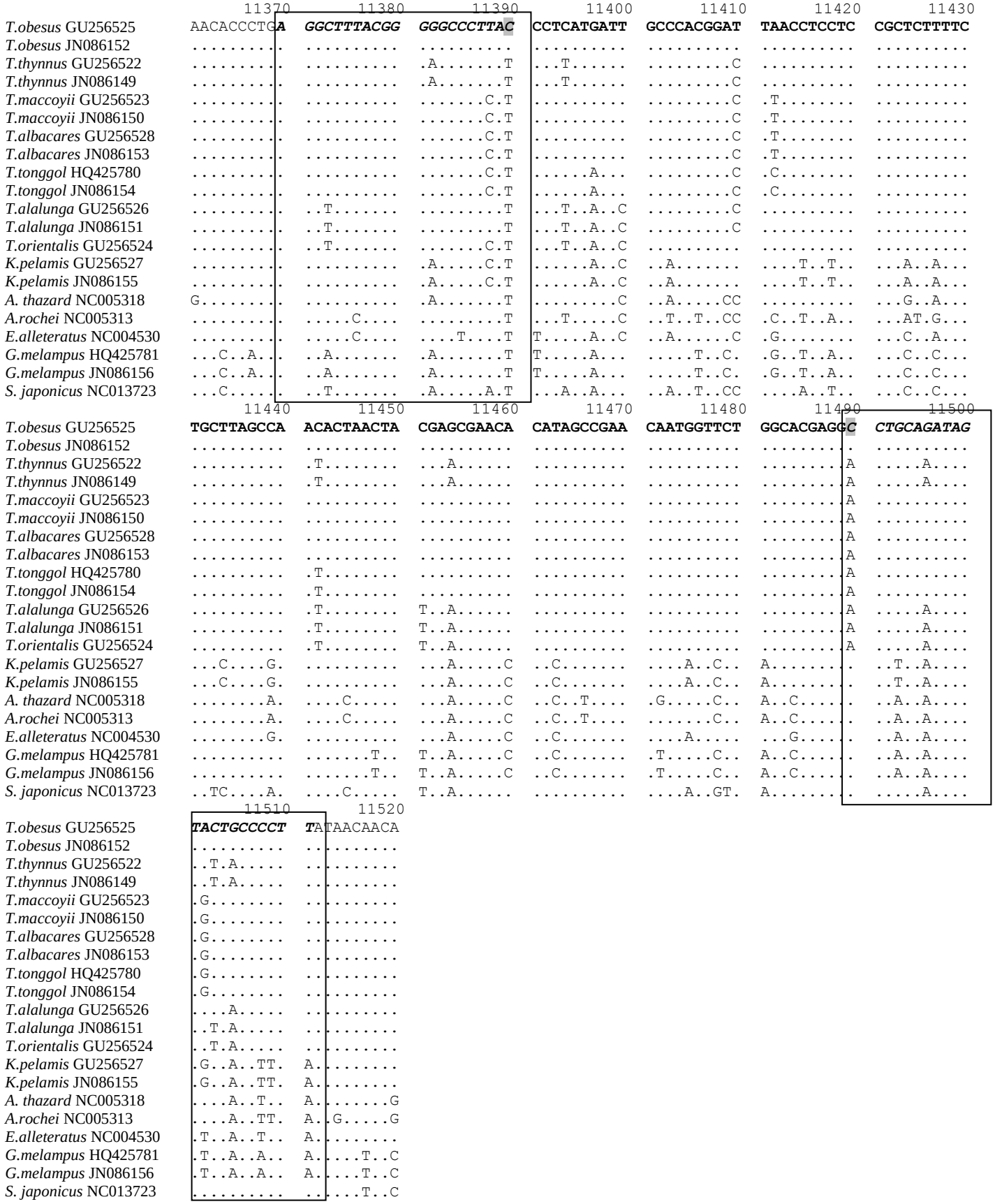


Fig. 2-2e. Alignment of mtDNA sequences showing the SND used for design of the short SSPP for *Thunnus obesus*. In italics, the SSPP sequence. Shaded, the SND at the 3' extreme of the SSPP

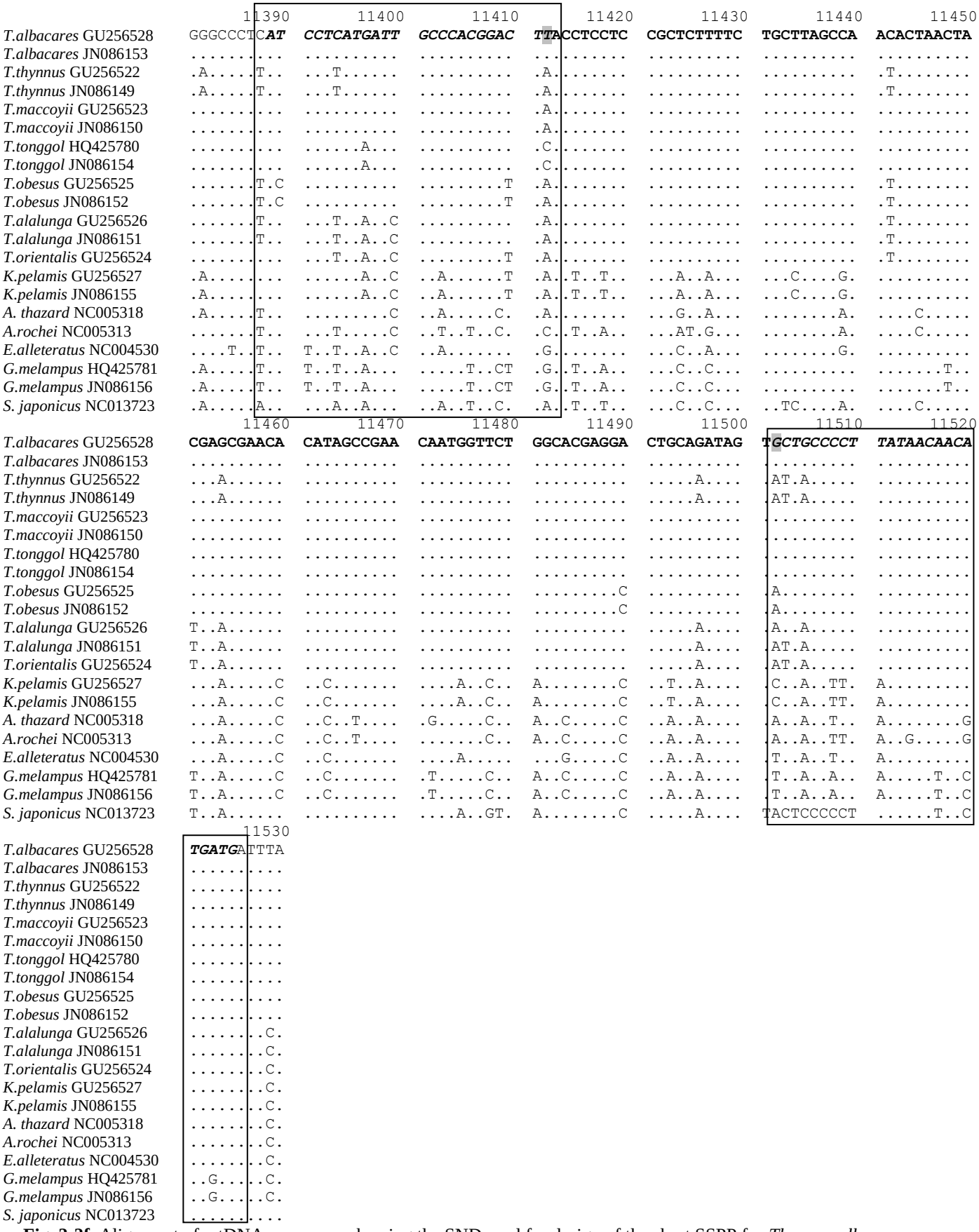


Fig. 2-2f. Alignment of mtDNA sequences showing the SND used for design of the short SSPP for *Thunnus albacares*. In italics, the SSPP sequence. Shaded, the SND at the 3' extreme of the SSPP

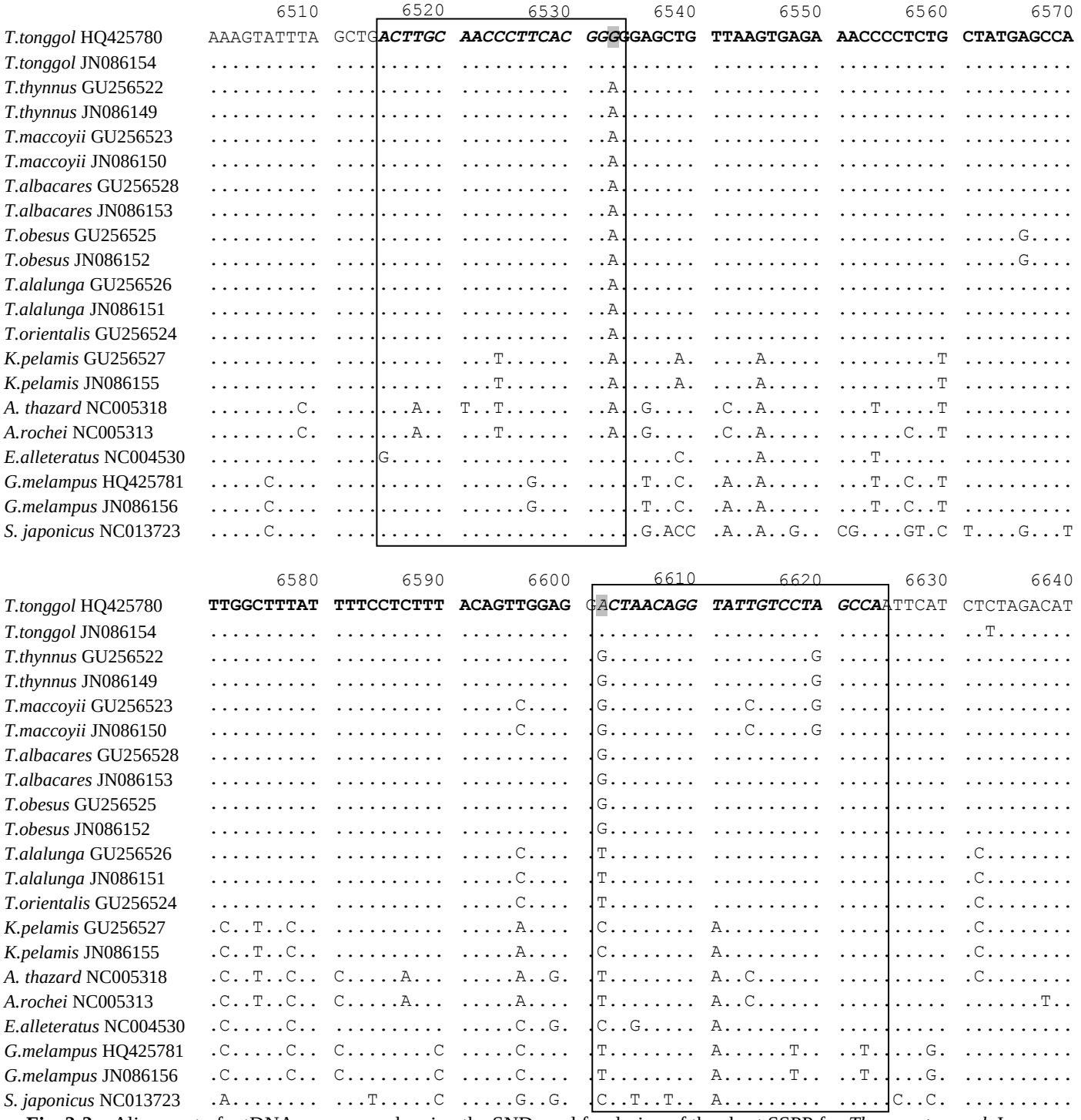


Fig. 2-2g. Alignment of mtDNA sequences showing the SND used for design of the short SSPP for *Thunnus tonggol*. In italics, the SSPP sequence. Shaded, the SND at the 3' extreme of the SSPP

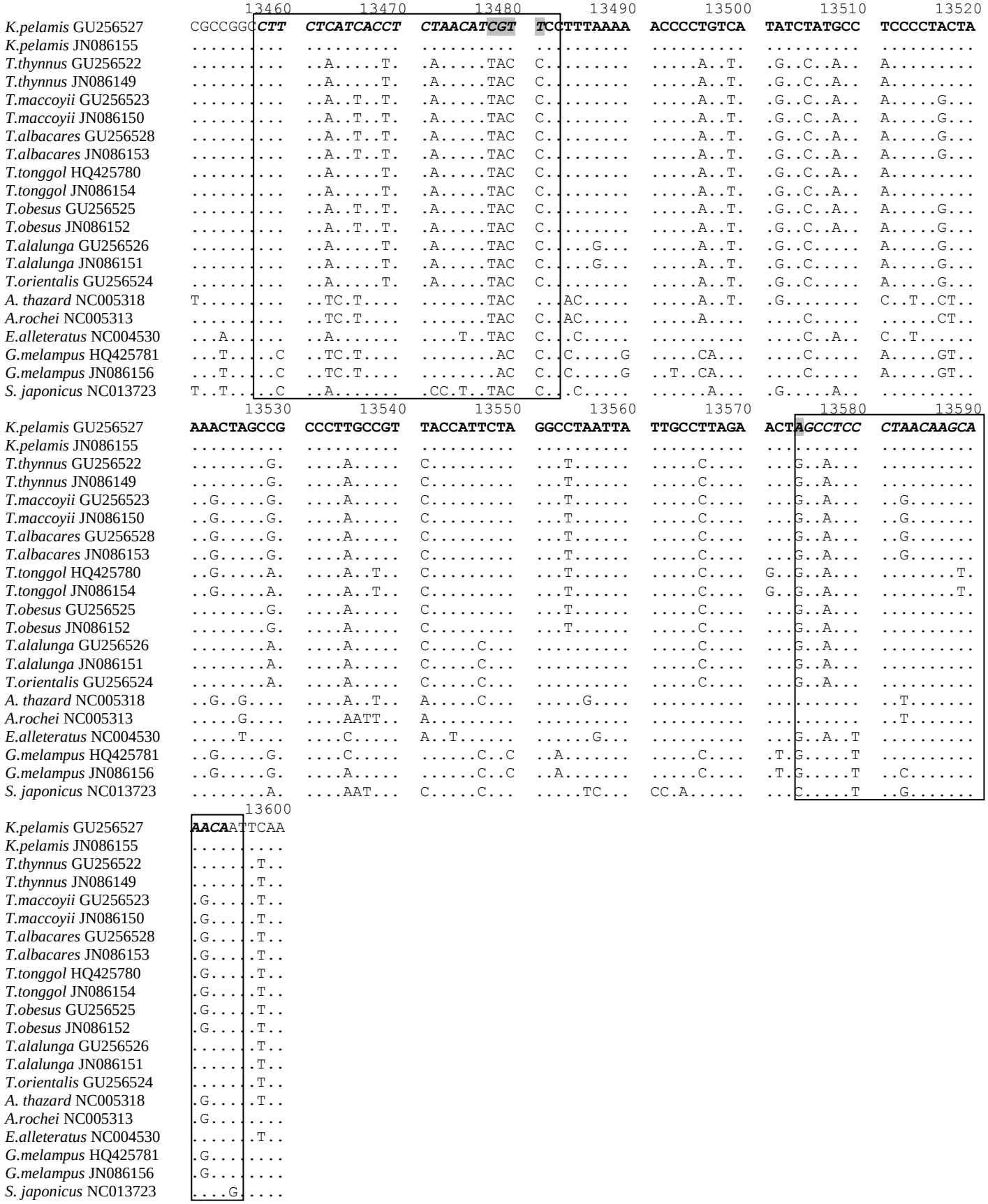


Fig. 2-2h. Alignment of mtDNA sequences showing the region used for design of the short SSPP for *Katsuwonus pelamis*. In red, the SSPP sequence. Shaded, the SND at the 3' extreme of the SSPP

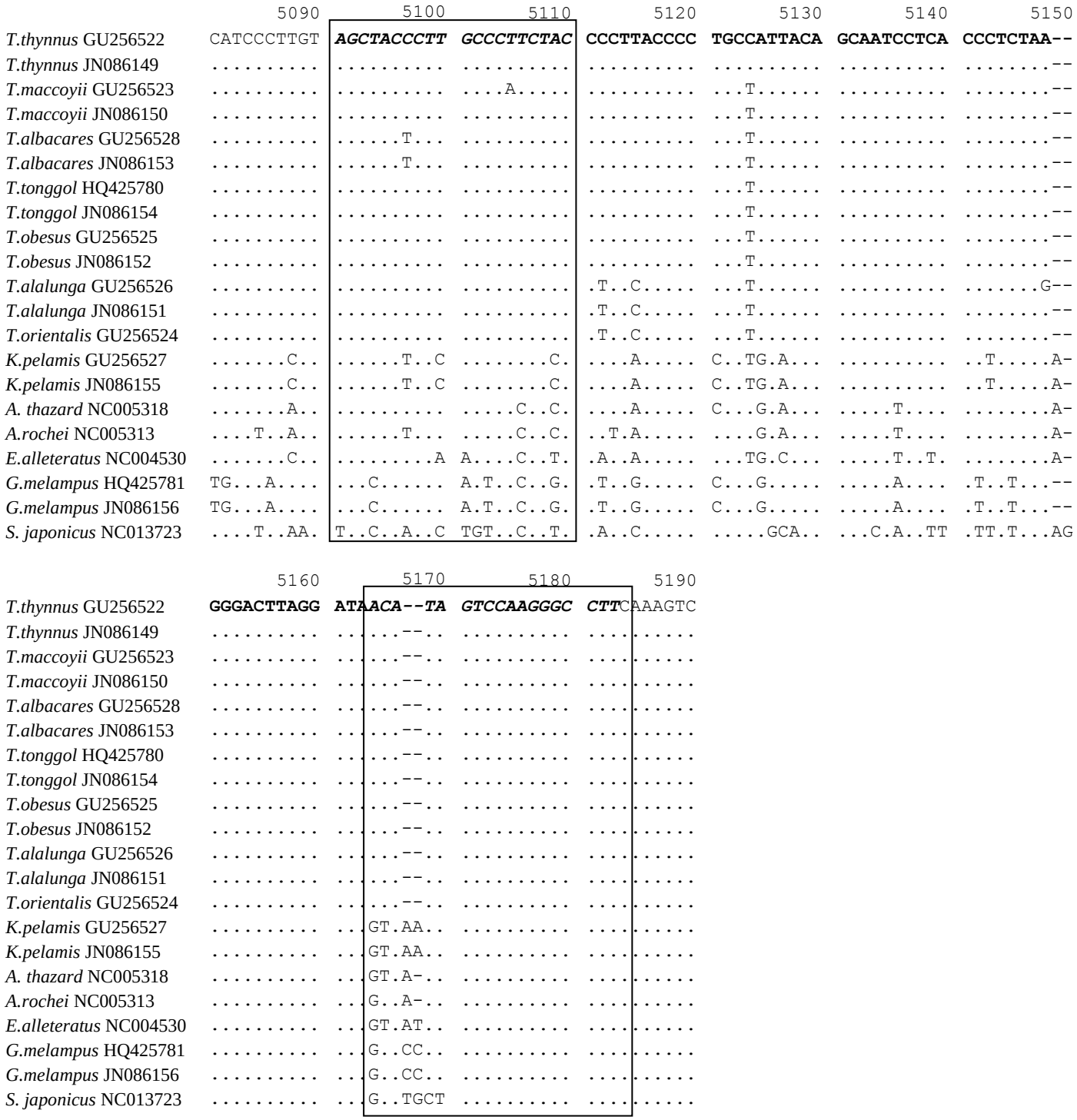


Fig. 2-2i. Alignment of mtDNA sequences showing the region used for design of the TSPP. In italics, the TSPP sequence. Lines represent nucleotides absent in the sequence.

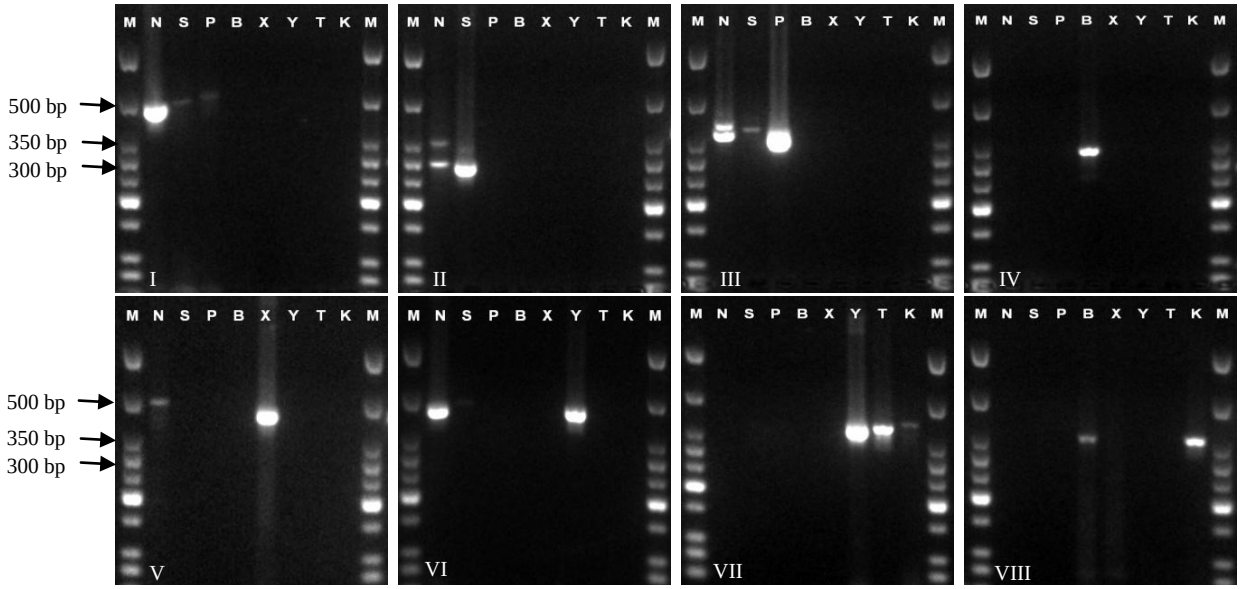


Fig. 2-3a. Specificity of the SSPP/TSPP-PCR discrimination method using long SSPP alone. Each PCR was done with 30 ng of total DNA extracted from unheated tuna tissue.
SSPP used in each electrophoretogram: I. *Thunnus thynnus*; II. *T. maccoyii*; III. *T. orientalis*; IV. *T. alalunga*; V. *T. obesus*; VI. *T. albacares*; VII. *T. tonggol*; VIII. *Katsuwonus pelamis*.
DNA template: N. *T. thynnus*; S. *T. maccoyii*; P: *T. orientalis*; B. *T. alalunga*; X. *T. obesus*; Y. *T. albacares* ; T. *T. tonggol*; K. *K. pelamis*

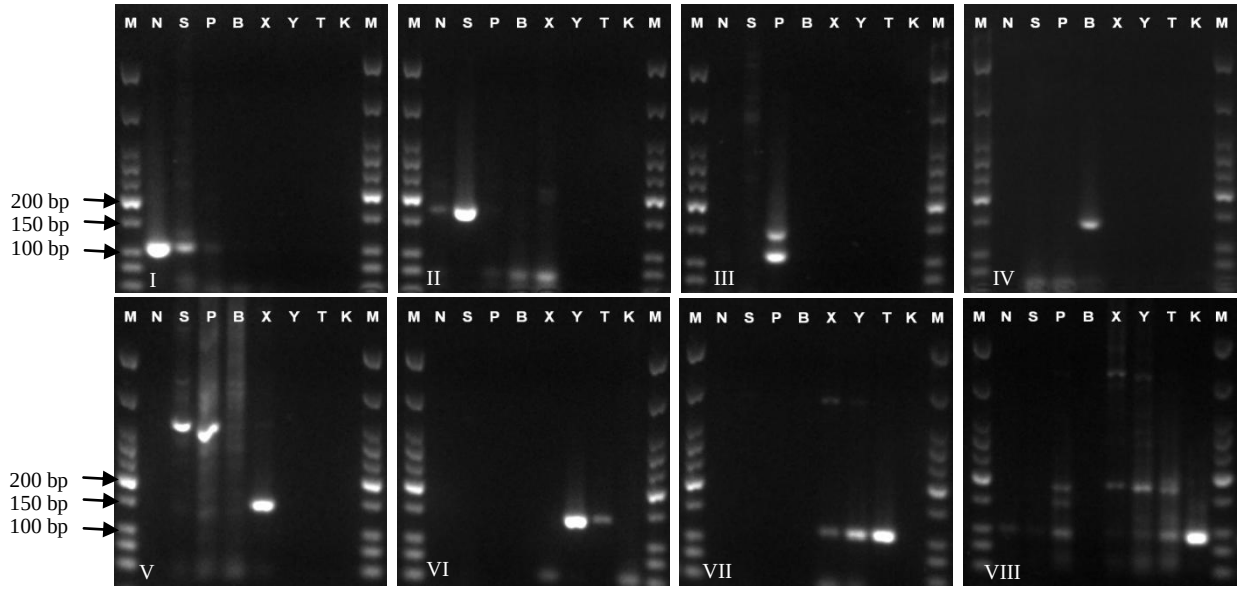


Fig. 2-3b. Specificity of the SSPP/TSPP-PCR discrimination method using short SSPP alone. Each PCR was done with 30 ng of total DNA extracted from unheated tuna tissue.
SSPP used in each electrophoretogram: I. *Thunnus thynnus*; II. *T. maccoyii*; III. *T. orientalis*; IV. *T. alalunga*; V. *T. obesus*; VI. *T. albacares*; VII. *T. tonggol*; VIII. *Katsuwonus pelamis*.
DNA template: N. *T. thynnus*; S. *T. maccoyii*; P: *T. orientalis*; B. *T. alalunga*; X. *T. obesus*; Y. *T. albacares* ; T. *T. tonggol*; K. *K. pelamis*

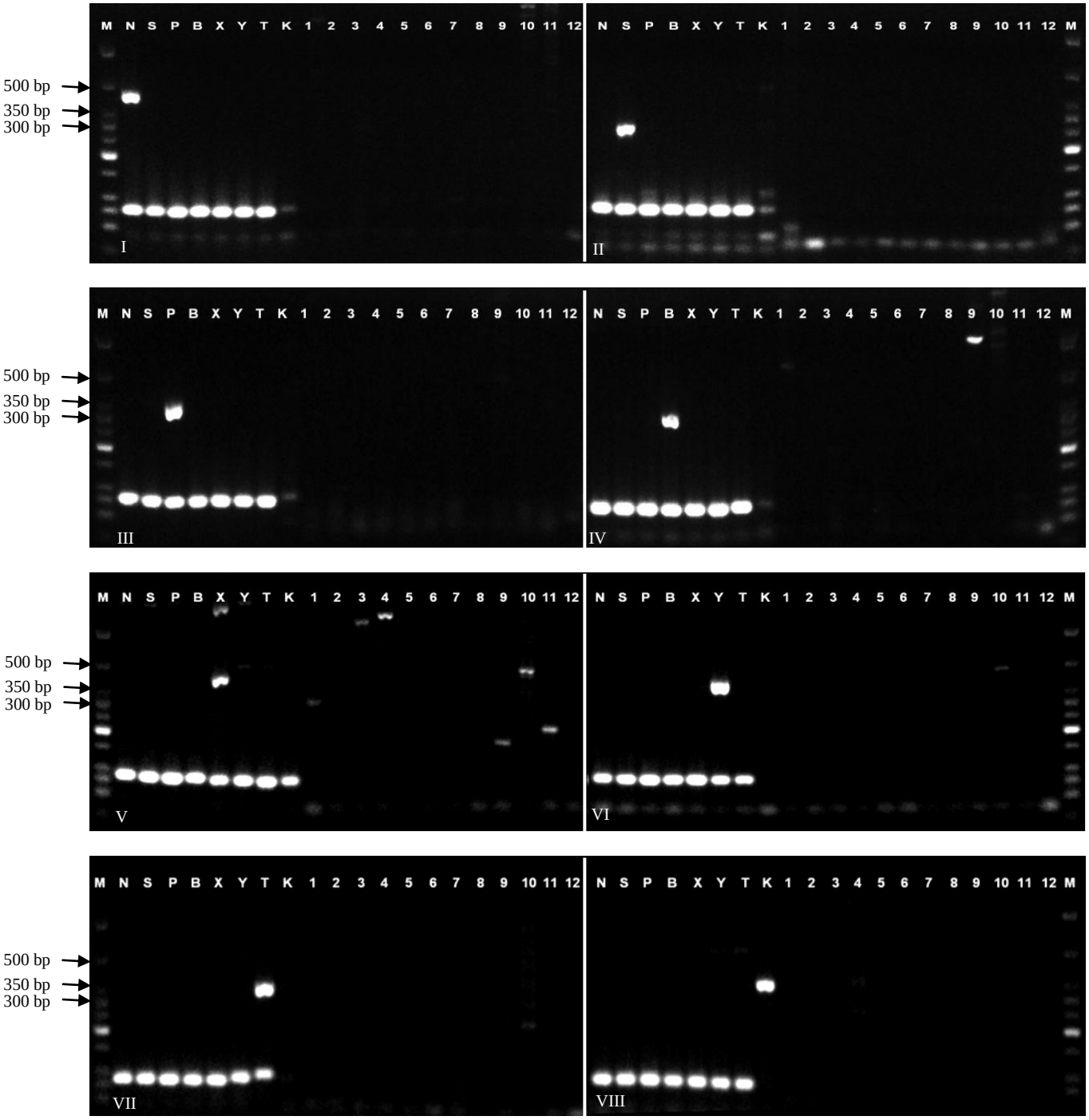


Fig. 2-4a. Specificity of the SSPP/TSPP-PCR discrimination method using long SSPP and TSPP. Each PCR was done with 30 ng of total DNA extracted from unheated tuna tissue as well as from different fish and animals. The lower band at 89 bp corresponds to the TSPP PCR product. The upper band corresponds to the respective SSPP PCR product. **SSPP used in each electrophoretogram**: I. *Thunnus thynnus*; II. *T. maccoyii*; III. *T. orientalis*; IV. *T. alalunga*; V. *T. obesus*; VI. *T. albacares*; VII. *T. tonggol*; VIII. *Katsuwonus pelamis*. **DNA template:** N. *T. thynnus*; S. *T. maccoyii*; P: *T. orientalis*; B. *T. alalunga*; X. *T. obesus*; Y. *T. albacares* ; T. *T. tonggol*; K. *K. pelamis*; 1. *Gasterochisma melampus*; 2. *Scomber japonicus*; 3. *Xiphias gladius*; 4.*Cololabis saira*; 5. *Engraulis japonica*; 6. *Pagrus major*; 7. *Pleuronectes yokohamae*; 8. *Gallus gallus*; 9. *Sus scrofa*; 10; *Bos taurus*;11. *Homo sapiens*; 12. Water

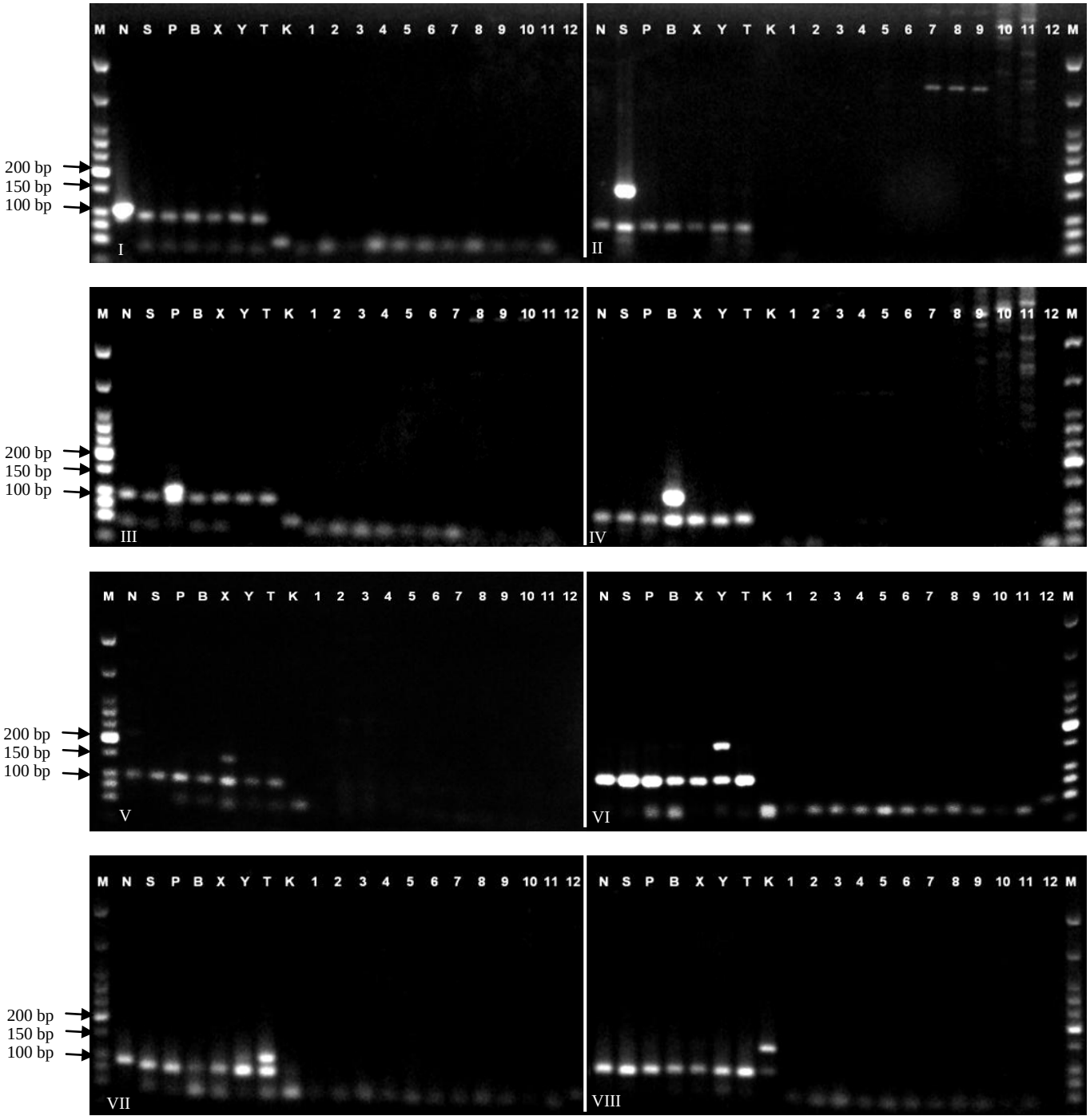


Fig. 2-4b. Specificity of the SSPP/TSPP-PCR discrimination method using long SSPP and TSPP. Each PCR was done with 30 ng of total DNA extracted from unheated tuna tissue as well as from different fish and animals. The lower band at 89 bp corresponds to the TSPP PCR product. The upper band corresponds to the respective SSPP PCR product. **SSPP used in each electrophoretogram**: I. *Thunnus thynnus*; II. *T. maccoyii*; III. *T. orientalis*; IV. *T. alalunga*; V. *T. obesus*; VI. *T. albacares*; VII. *T. tonggol*; VIII. *Katsuwonus pelamis*. **DNA template:** N. *T. thynnus*; S. *T. maccoyii*; P: *T. orientalis*; B. *T. alalunga*; X. *T. obesus*; Y. *T. albacares* ; T. *T. tonggol*; K. *K. pelamis*; 1. *Gasterochisma melampus*; 2. *Scomber japonicus*; 3. *Xiphias gladius*; 4.*Cololabis saira*; 5. *Engraulis japonica*; 6. *Pagrus major*; 7. *Pleuronectes yokohamae*; 8. *Gallus gallus*; 9. *Sus scrofa*; 10; *Bos taurus*;11. *Homo sapiens*; 12. Water

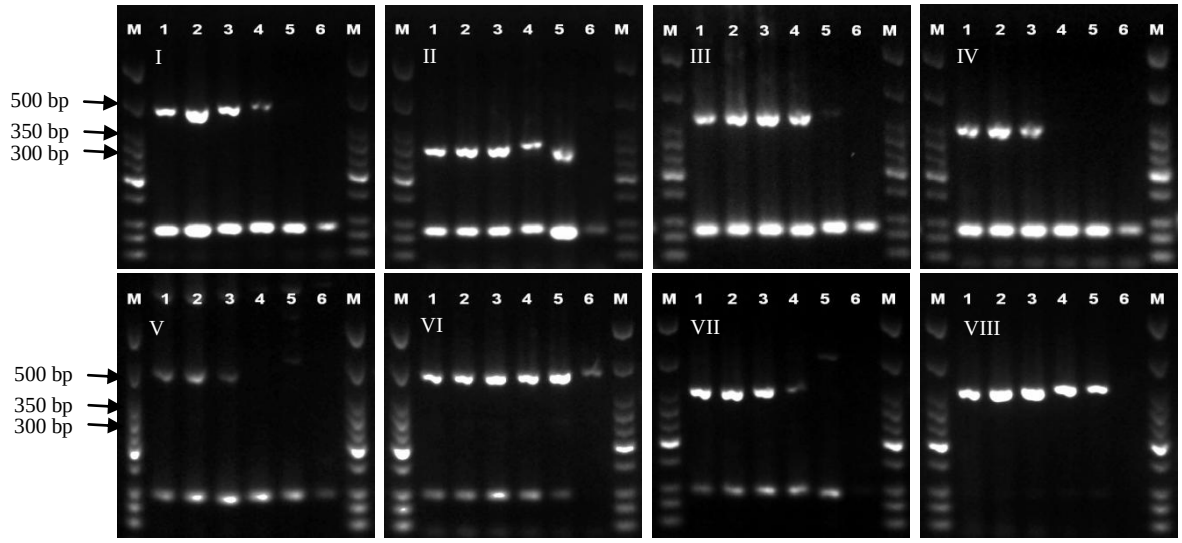


Fig. 2-5a. Effect of annealing temperature in the PCR amplification using long SSPP and TSPP. Each PCR was done with 30 ng of total DNA extracted from unheated tuna tissue. The lower band at 89 bp corresponds to the TSPP PCR product. The upper band corresponds to the respective SSPP PCR product.
Applied SSPP: I. *Thunnus thynnus*; II. *T. maccoyii*; III. *T. orientalis*; IV. *T. alalunga*; V. *T. obesus*; VI. *T. albacares*; VII. *T. tonggol*; VIII. *Katsuwonus pelamis*. **Annealing temperature: 1.** 48°C; **2.** 52°C; **3.** 56°C; **4.** 60°C, **5.** 64°C, **6.** 68°C

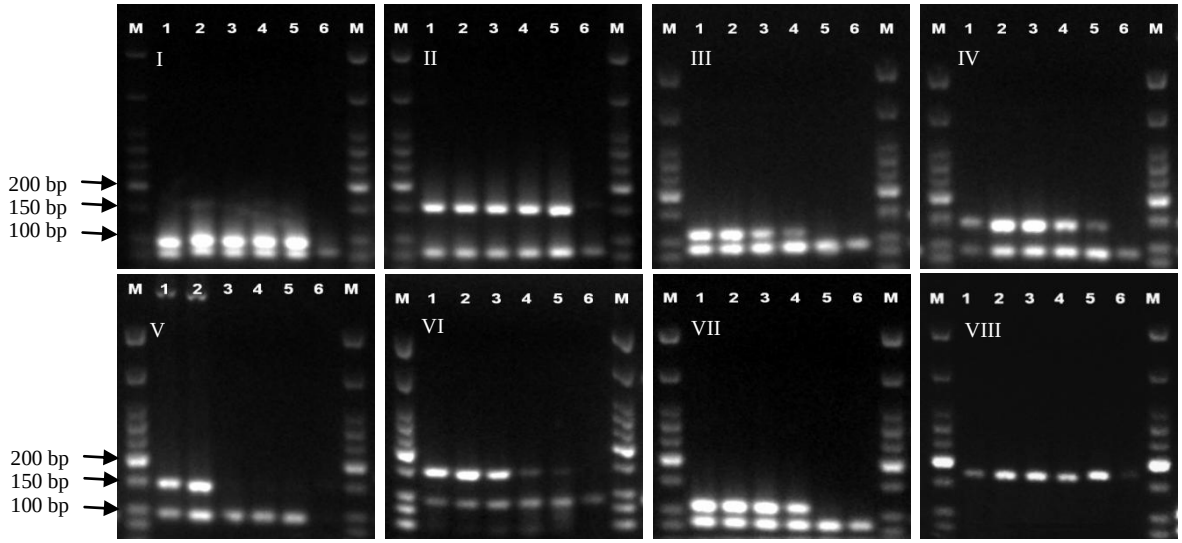


Fig. 2-5b. Effect of annealing temperature in the PCR amplification using short SSPP and TSPP. Each PCR was done with 30 ng of total DNA extracted from unheated tuna tissue. The lower band at 89 bp corresponds to the TSPP PCR product. The upper band corresponds to the respective SSPP PCR product.
Applied SSPP: I. *Thunnus thynnus*; II. *T. maccoyii*; III. *T. orientalis*; IV. *T. alalunga*; V. *T. obesus*; VI. *T. albacares*; VII. *T. tonggol*; VIII. *Katsuwonus pelamis*. **Annealing temperature: 1.** 48°C; **2.** 52°C; **3.** 56°C; **4**. 60°C, **5.** 64°C, **6.** 68°C

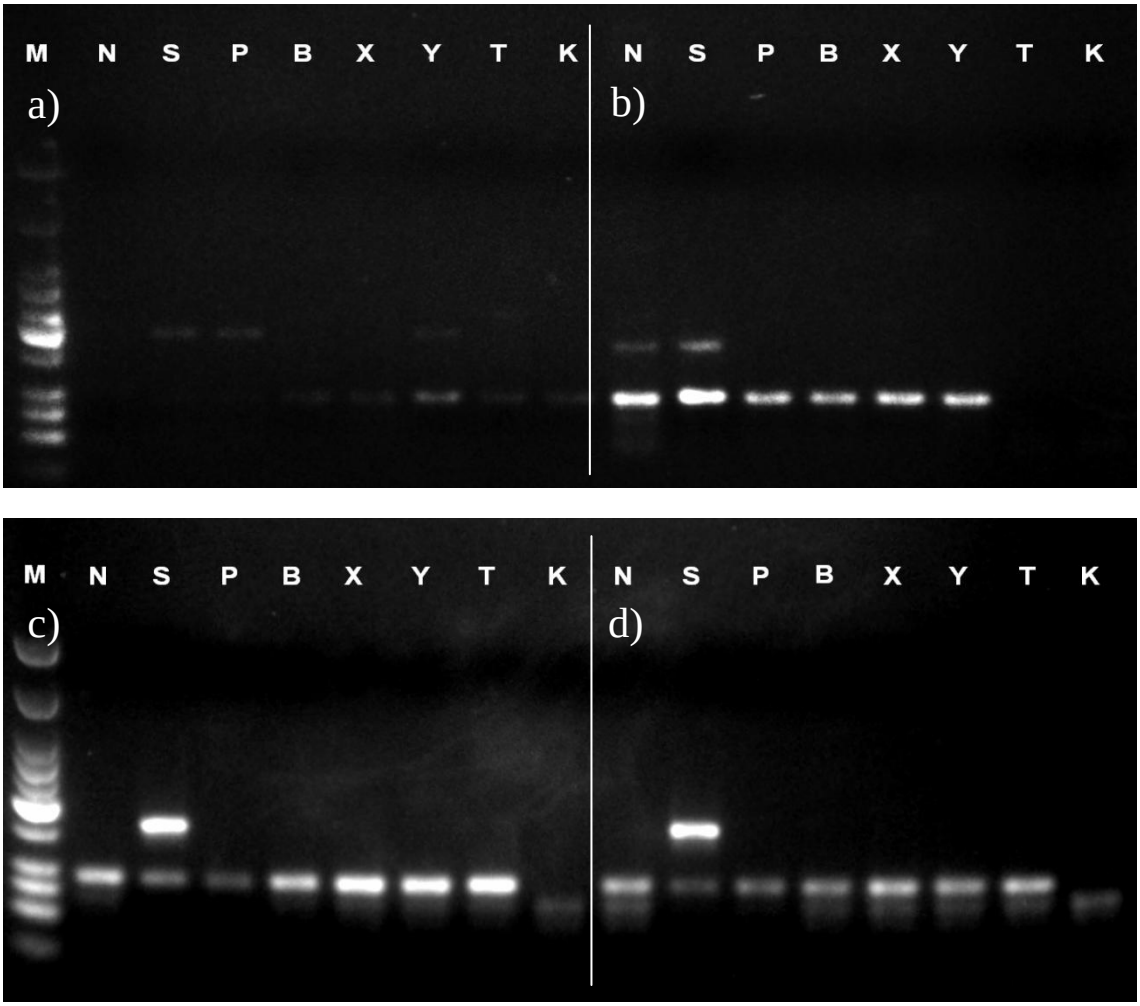


Fig. 2-6. Representative results of the effect of primer concentration on the efficiency of the SSPP/TSPP PCR discrimination method. a) TSPP= 0.15 µM, SSPP= 0.15 µM; b) TSPP= 0.3 µM; SSPP= 0.3 µM; c) TSPP= 0.3 µM, SSPP= 0.6 µM; d) TSPP= 0.3 µM, SSPP= 1.2 µM. M = DNA marker, N= *Thunnus thynnus*, S= *T. maccoyii*, P= *T. orientalis*, B= *T. alalunga*, X= *T. obesus*, Y= *T. albacares*, T= *T. tonggol*, K= *Katsuwonus pelamis*. SSPP for *T. maccoyii* were used in this example.

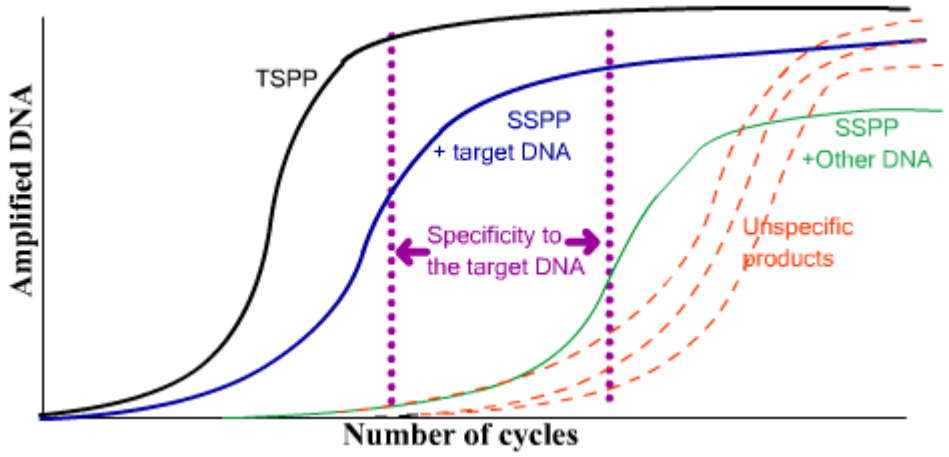


Fig 2-7a. Amplification with SSPP/TSPP (1:1). The SSPP signal is weak and delayed.

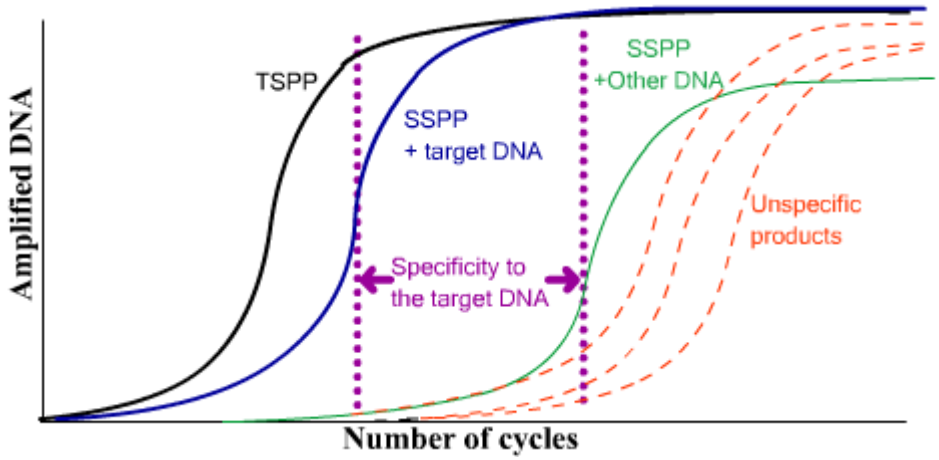


Fig 2-7b. Amplification under the optimal conditions. Adapted from Gibson, 2005

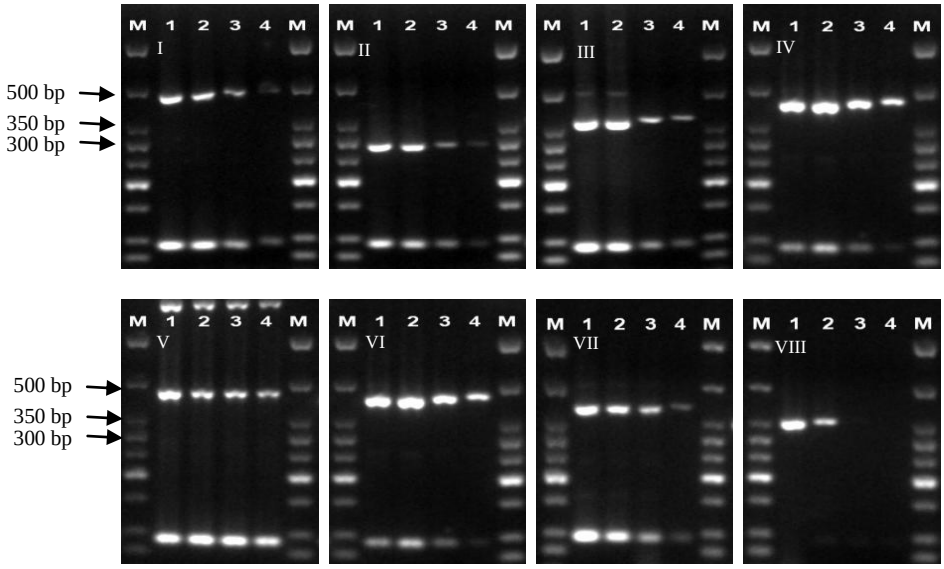


Fig. 2-8a. Detection limit using long SSPP and TSPP on DNA from unheated tuna tissue.
Species: I. *Thunnus thynnus*; II. *T. maccoyii*; III. *T. orientalis*; IV. *T. alalunga*; V. *T. obesus*; VI. *T. albacares*; VII. *T. tonggol*; VIII. *Katsuwonus pelamis*. **Total DNA**: 1. 30 ng; 2. 3 ng; 3. 0.3 ng; 4. 30 pg

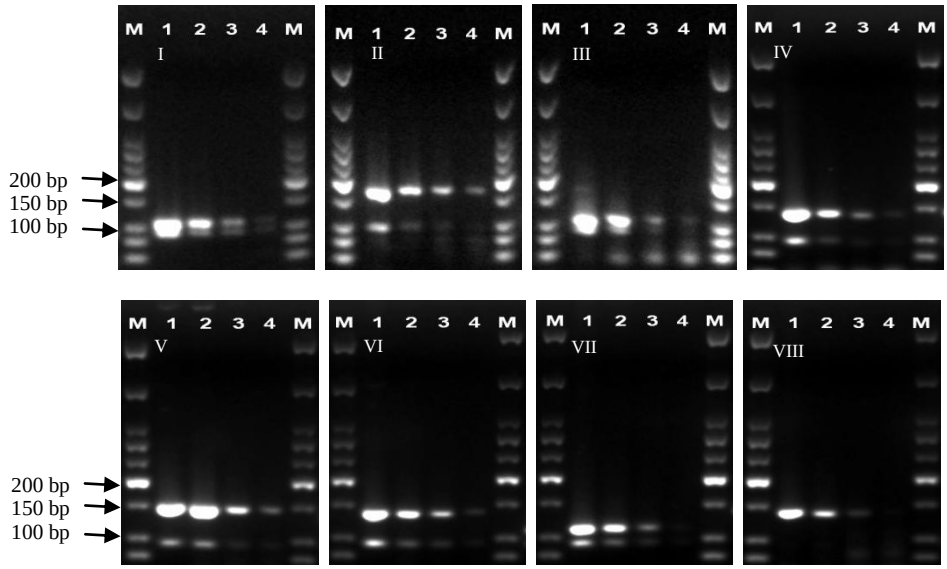
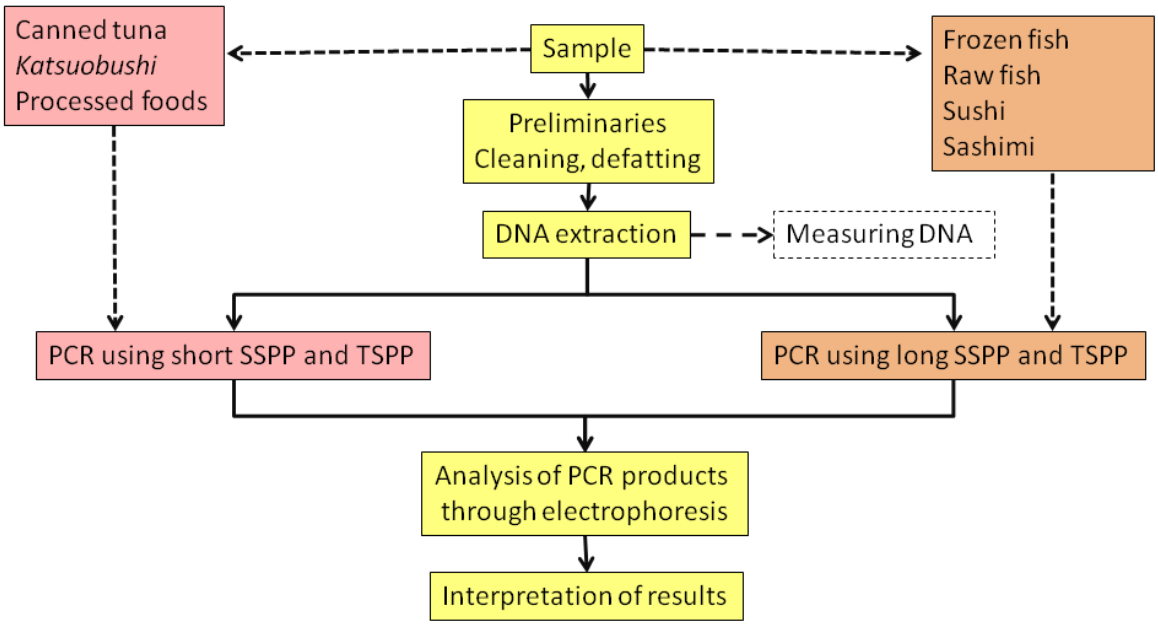


Fig. 2-8b. Detection limit using short SSPP and TSPP on DNA from unheated tuna tissue.
Species: I. *Thunnus thynnus*; II. *T. maccoyii*; III. *T. orientalis*; IV. *T. alalunga*; V. *T. obesus*; VI. *T. albacares*; VII. *T. tonggol*; VIII. *Katsuwonus pelamis*. **Total DNA**: 1. 30 ng; 2. 3 ng; 3. 0.3 ng; 4. 30 pg



1	×TSSP product	×SSPP product	No tuna or below detection limit
2	✓TSSP product	×SSPP product	Another tuna species detected
3	×TSSP product	✓SSPP product	Inconclusive
4	✓TSSP product	✓SSPP product	Target tuna present in the sample

Fig. 2-9. Flow chart of the SSPP/TSPP-PCR method for discrimination of tuna species in food products

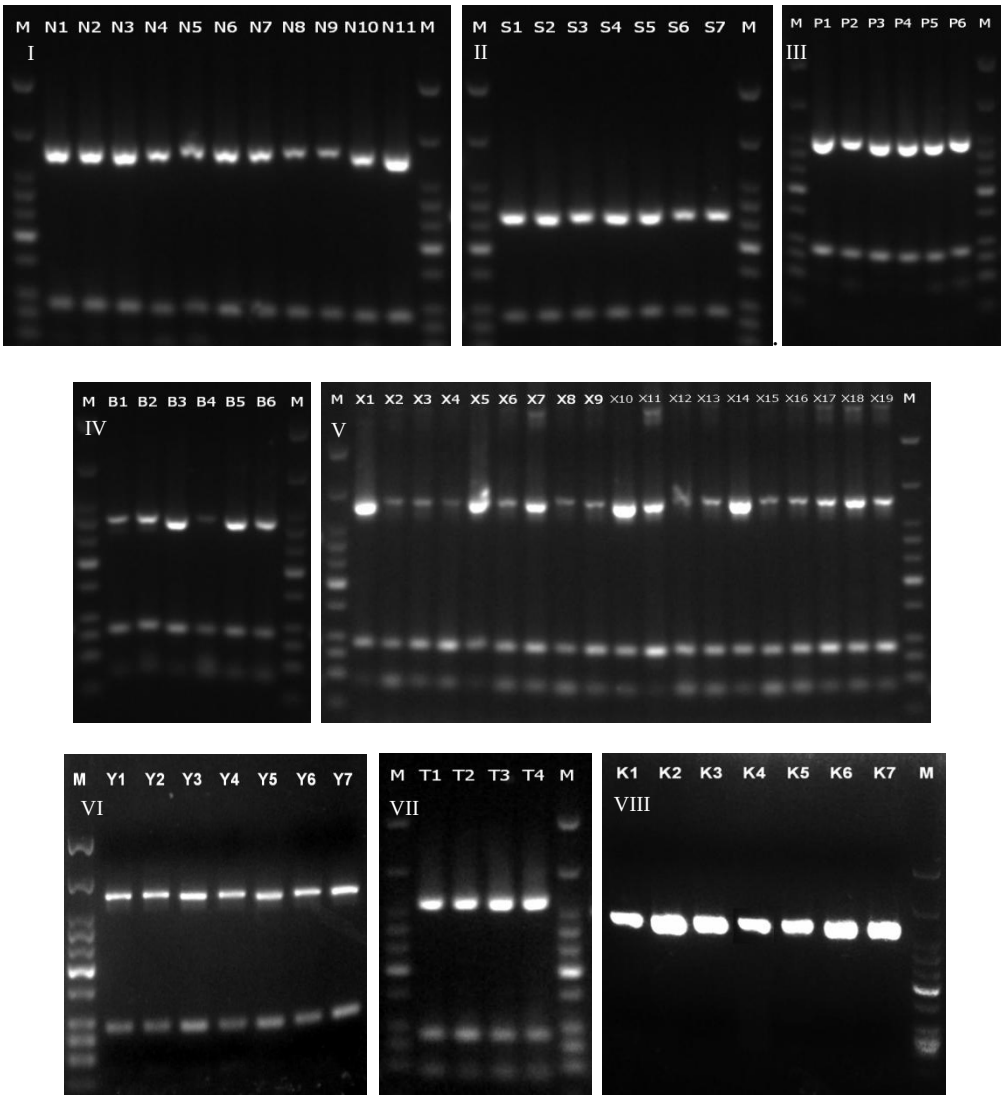


Fig. 2-10a. Identification of commercial samples of unheated tuna products using the SSPP/TSPP-PCR method. The description of samples is given in Table 2-1a. I. *T. thynnus*; II. *T. maccoyii*; III. *T. orientalis*; IV. *T. alalunga*; V. *T. obesus*; VI. *T. albacares*; VII. *T. tonggol*; VIII. *K.pelamis.*

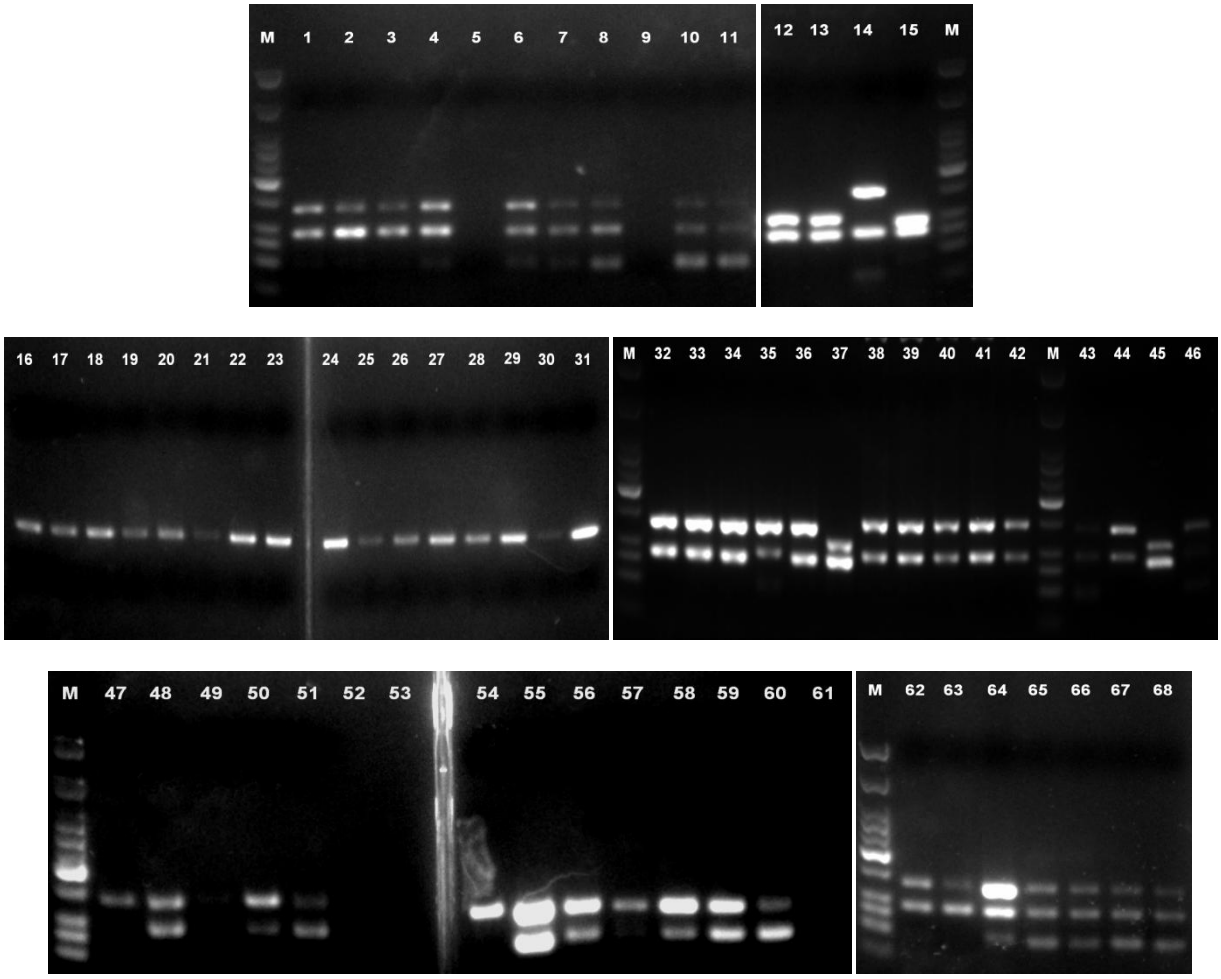


Fig. 2-10b. Identification of commercial samples of heat-processed tuna products using the SSPP/TSPP-PCR method. The description of samples is given in Table 2-1b.

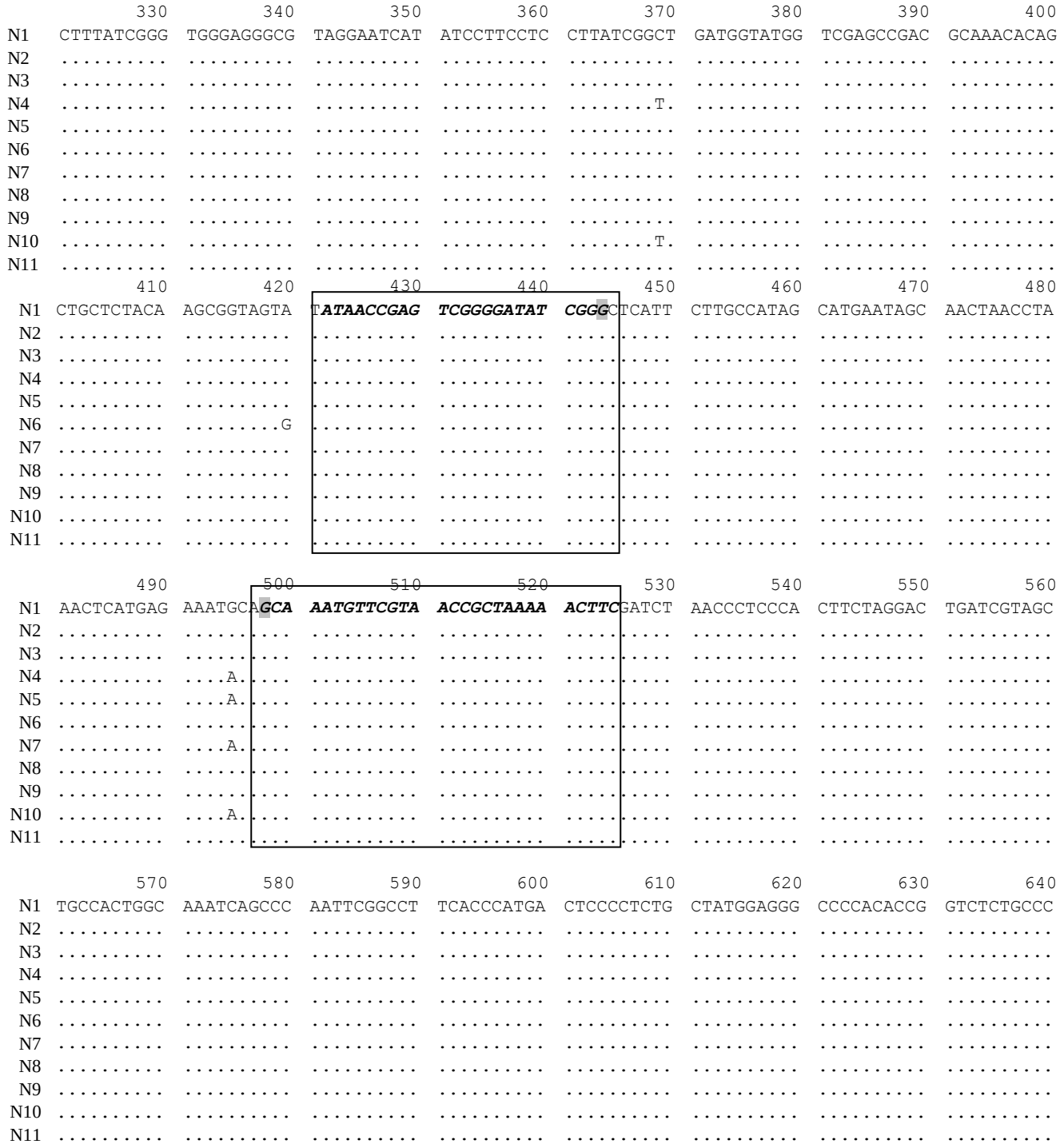


Fig. 2-11a. Alignment of 11 mtDNA sequences of *Thunnus thynnus* showing the position of the SSPP (in italics) and the presence of polymorphisms and other substitutions in adjacent regions

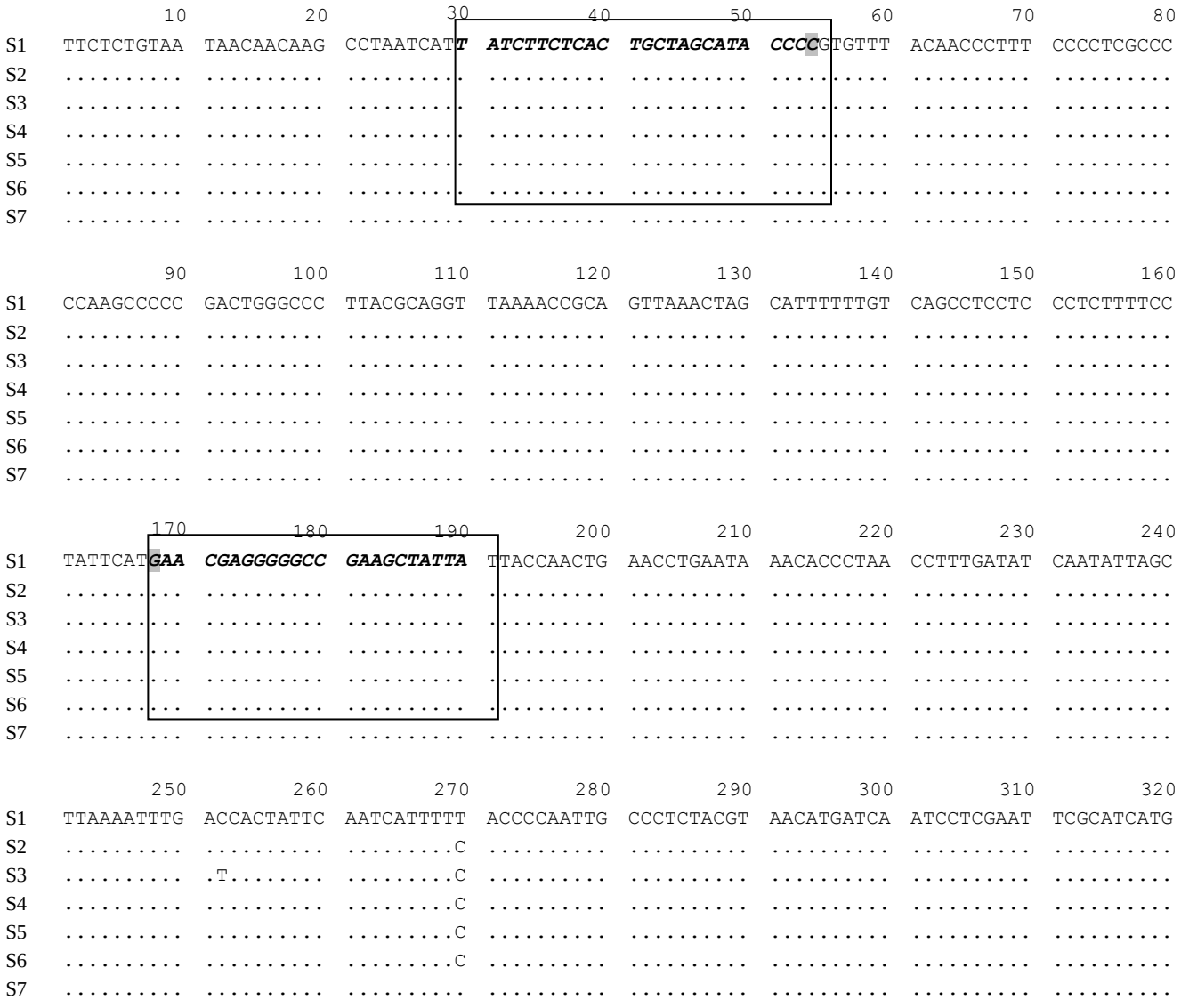


Fig. 2-11b. Alignment of 7 mtDNA sequences of *Thunnus maccoyii* showing the position of the SSPP (in italics) and the presence of polymorphisms and other substitutions in adjacent regions

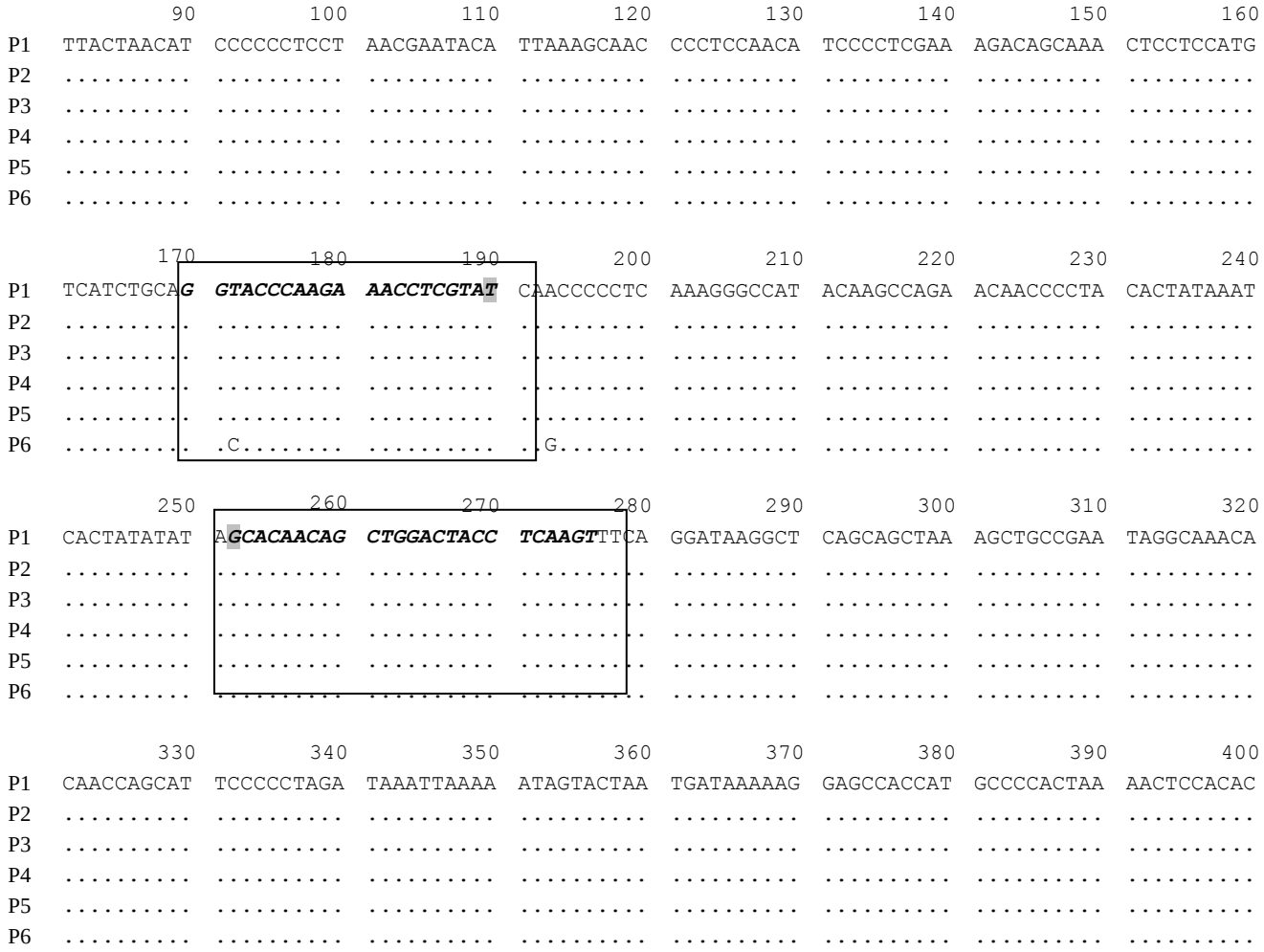


Fig. 2-11c. Alignment of 6 mtDNA sequences of *Thunnus orientalis* showing the position of the SSPP (in italics) and the presence of polymorphisms and other substitutions in adjacent regions

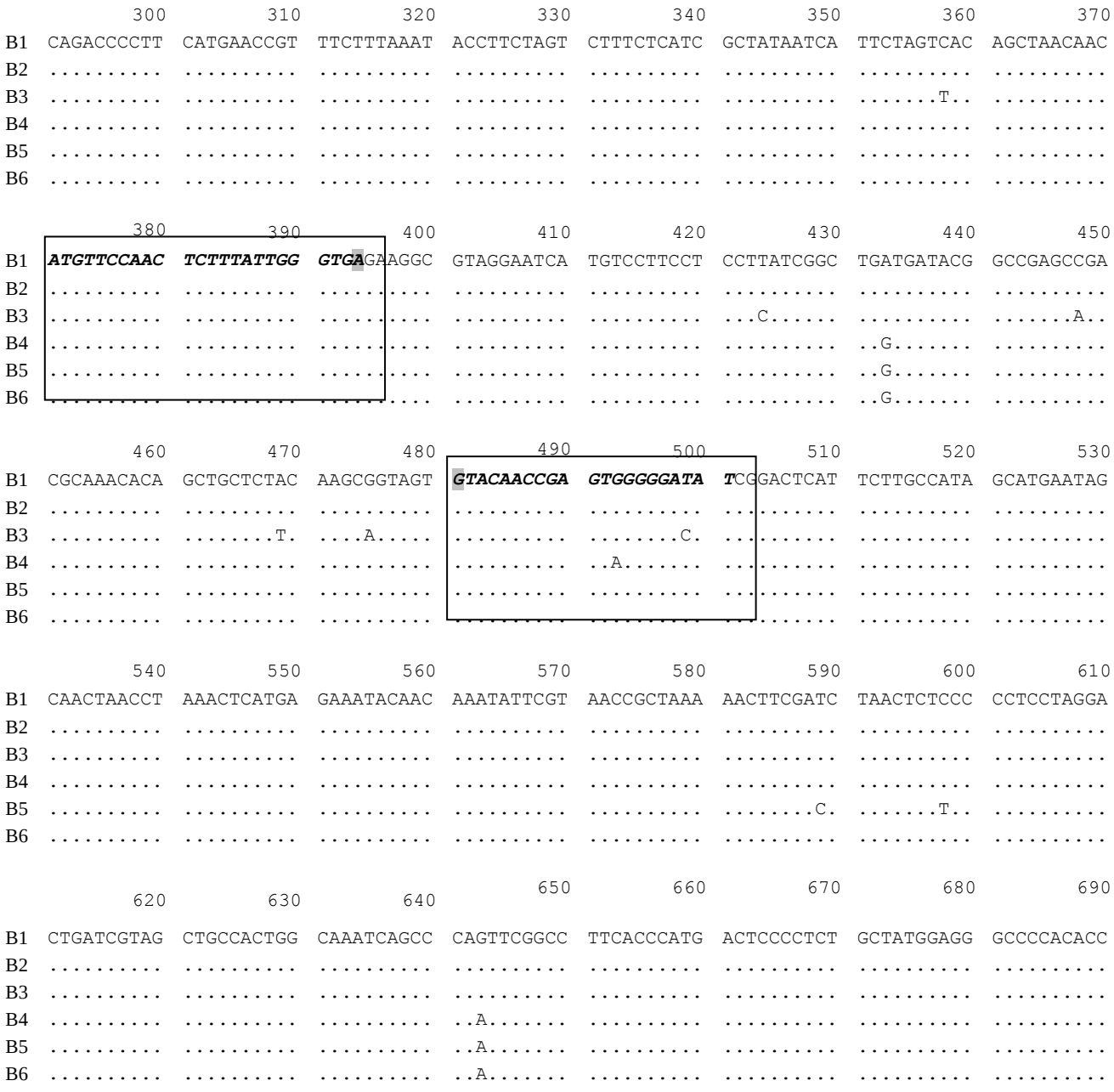


Fig. 2-11d. Alignment of 6 mtDNA sequences of *Thunnus alalunga* showing the position of the SSPP (in italics) and the presence of polymorphisms and other substitutions in adjacent regions

Fig. 2-11e. Alignment of 19 mtDNA sequences of *Thunnus obesus* showing the position of the SSPP (in italics) and the presence of polymorphisms and other substitutions in adjacent regions

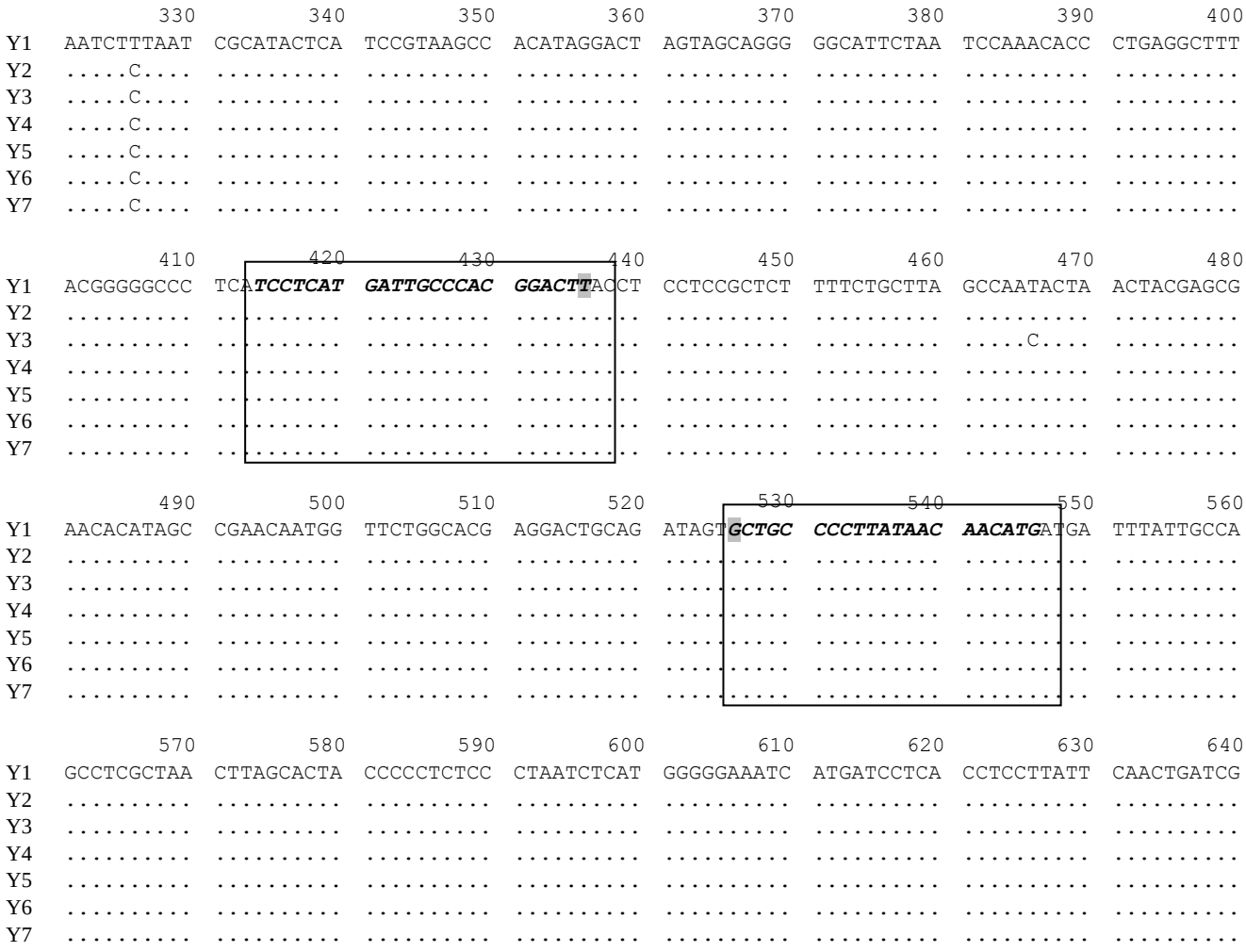


Fig. 2-11f. Alignment of 7 mtDNA sequences of *Thunnus albacares* showing the position of the SSPP (in italics) and the presence of polymorphisms and other substitutions in adjacent regions

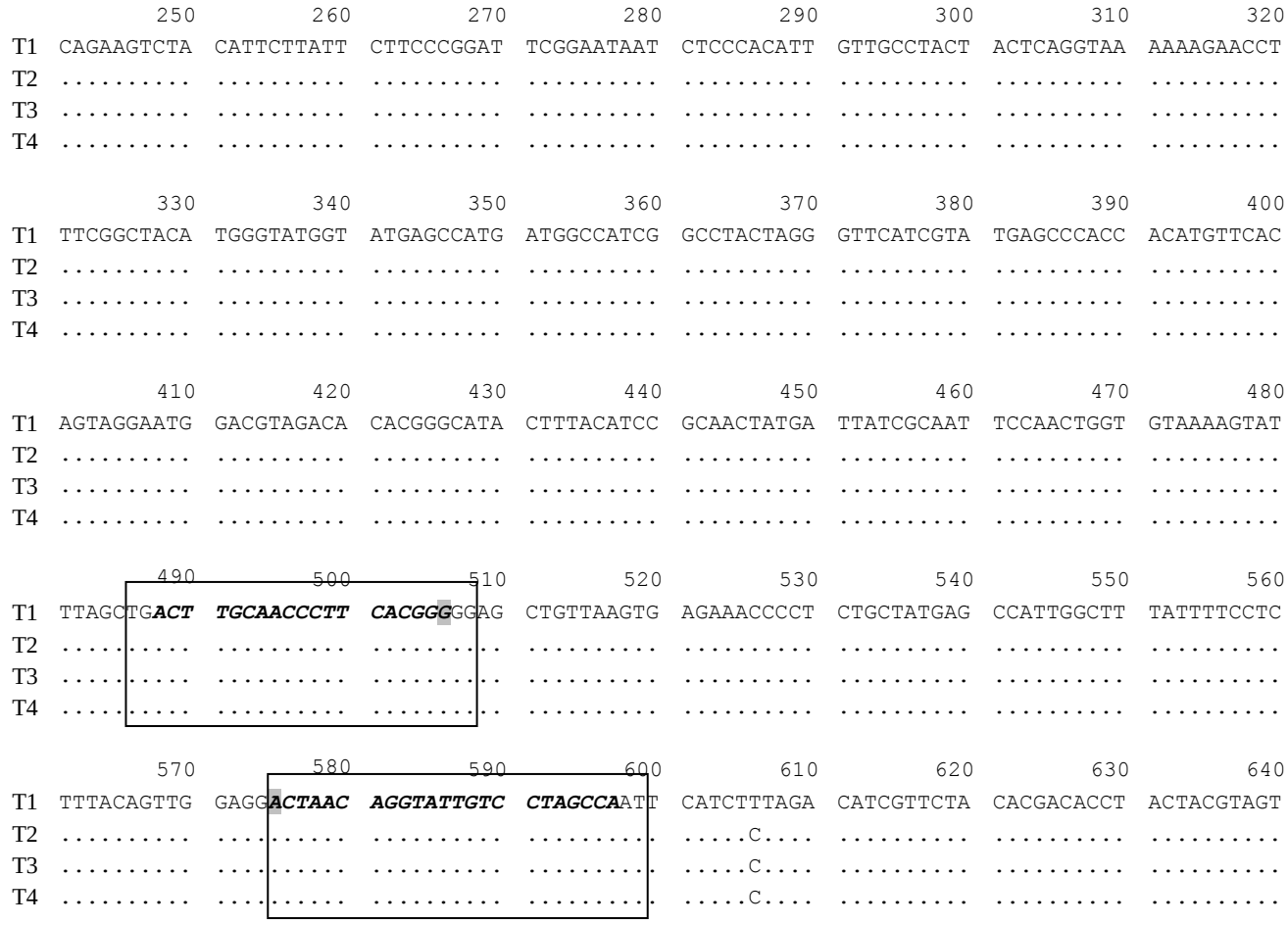


Fig. 2-11g. Alignment of 4 mtDNA sequences of *Thunnus tonggol* showing the position of the SSPP (in italics) and the presence of polymorphisms and other substitutions in adjacent regions

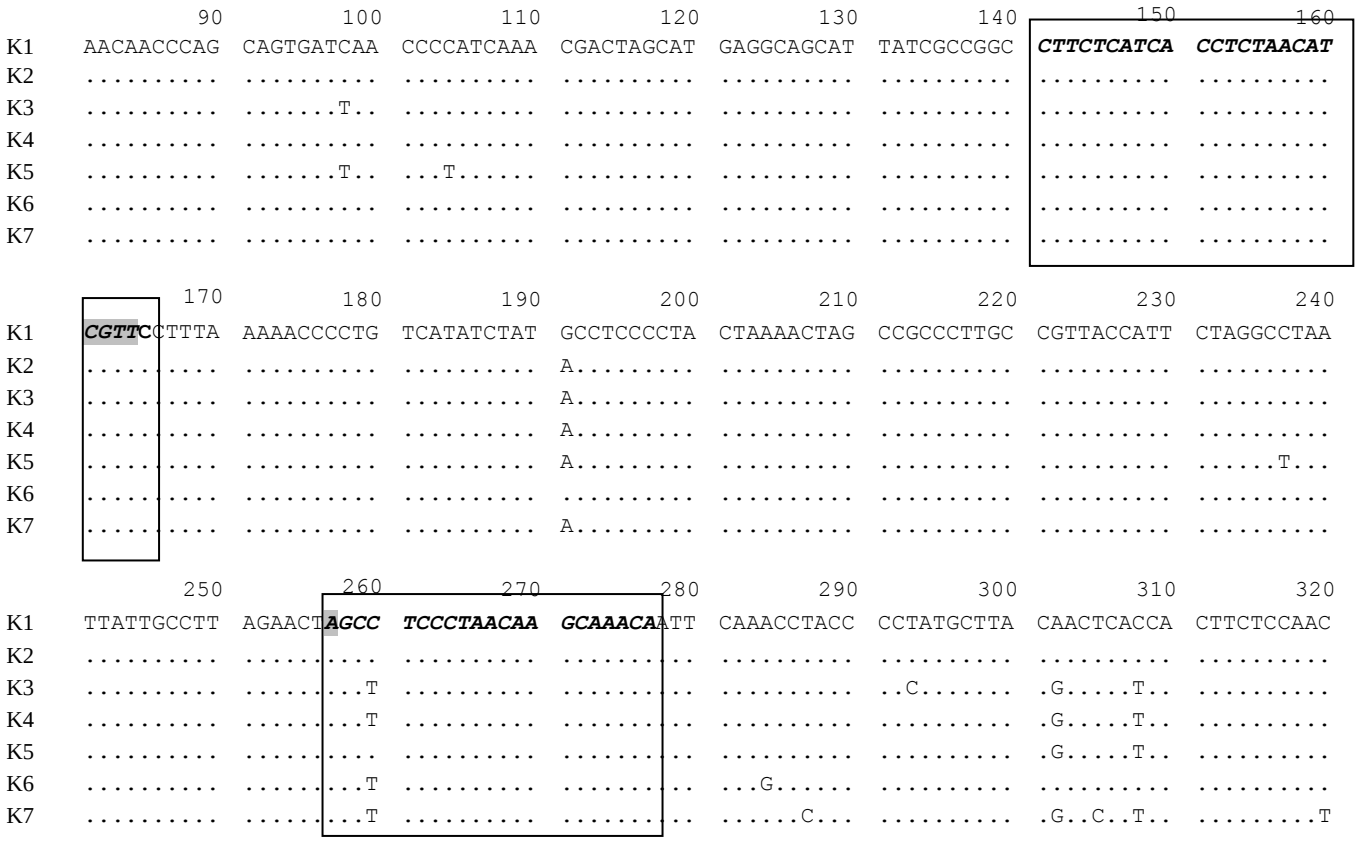


Fig. 2-11h. Alignment of 7 mtDNA sequences of *Katsuwonus pelamis* showing the position of the SSPP (in italics) and the presence of polymorphisms and other substitutions in adjacent regions

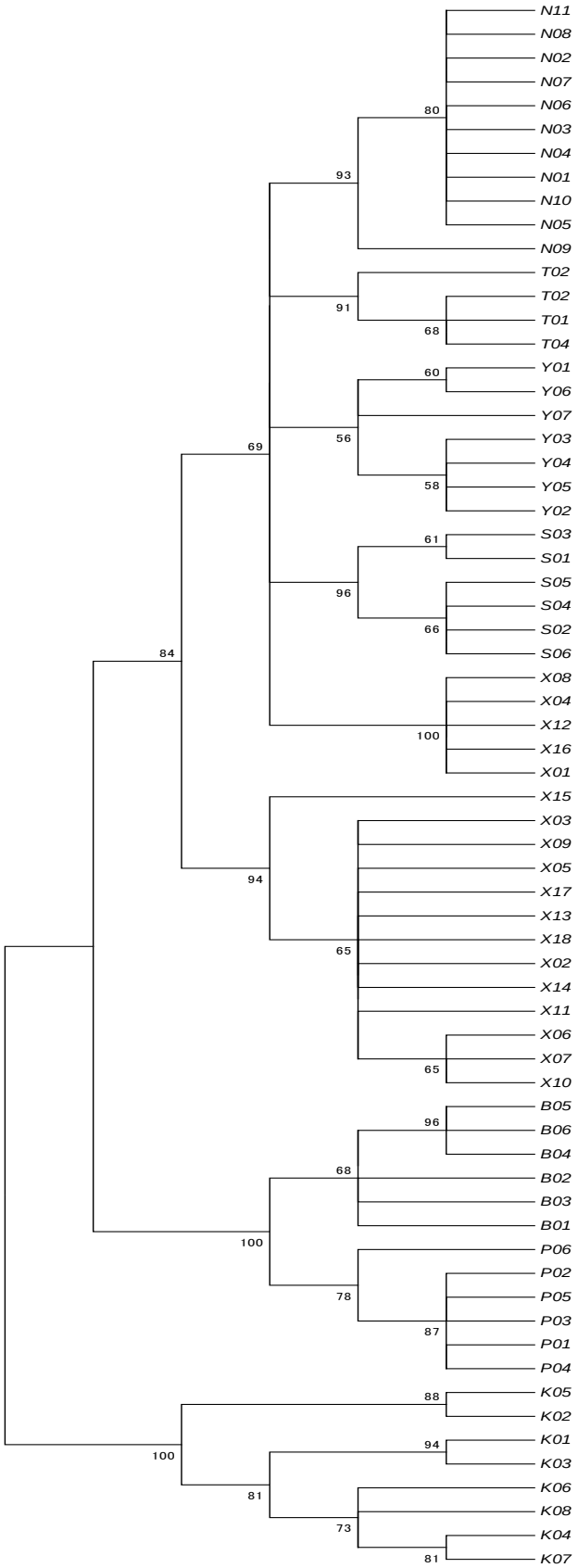


Fig. 2-12. Phylogenetic analysis of the ATP6-COIII region for validation of the SSPP/TSPP-PCR method. Description of each sample if given in Table 2-1a

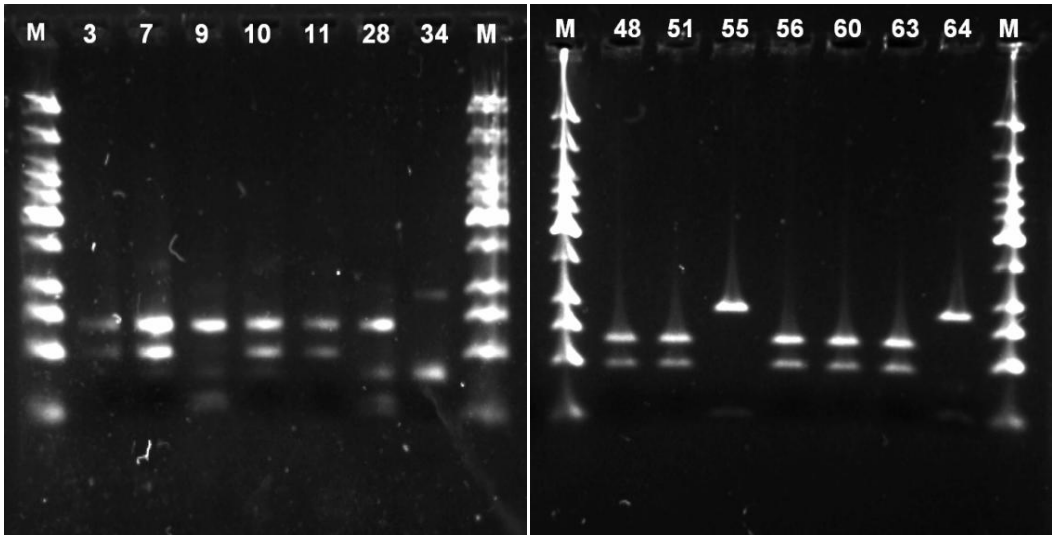


Fig. 2-13. Results of the RFLP method applied to samples of canned tuna. Only 14 of 68 samples could be analyzed satisfactorily. For description of each lane, see Table 2-1b

Chapter 3. Use of a competitive PCR method for quantitative analysis of tuna

3.1 Introduction

Fraudulent replacement of species in food products can be done by complete substitution but also by partial substitution. Partial replacement has the advantage of not altering the color and flavor of the product as much as complete replacement would do (Mackie et al., 1999). On the other hand, unintentional mixture of extraneous species due to misidentification, accidental contamination, insufficient sorting, etc., is also frequent and may not be completely avoided not even with the best manufacture practices. Finally, although some legislations explicitly indicate the use of particular species in some products (for example, according to the Code of Federal Regulations of the United States and the Regulation 1536/92 of the European Union, products labeled as 'white tuna' can only contain *T. alalunga* meat), sometimes certain level of admixture is allowed. For these reasons, in order to distinguish between accidental admixture and intentional replacement, identification and discrimination of species are not always sufficient and quantification analysis becomes necessary. Although real-time PCR methods for quantification in food products (including tuna) have been reported previously (Kesmen et al., 2009; Benedetto et al., 2011), the technical complexity as well as the cost of these methods makes it necessary to develop easier and cheaper alternatives. In this chapter, a method based in competitive PCR for quantitative analysis of mixtures of tuna is described. In Japan, as well as in many other countries, the species most commonly used for canning are *Thunnus albacares*, *T. alalunga* and *Katsuwonus pelamis* and thus, the quantification method was developed on

these three species. However, the principles applied to them can be easily extended to any other tuna species by synthesizing the respective competitor DNA.

3.2 Materials and methods

3.2.1. Samples and materials

Samples of frozen tissue and canned tuna described in previous chapter, included one commercial canned tuna where two species were detected, were used in this experiment too. For mixtures, frozen raw tissue or meat from canned tuna was weighed, finely chopped and mixed in the respective proportions, and kept at -30°C until use.

3.2.2 Total DNA extraction

DNA was extracted from tissue using the QuickGene Tissue DNA kit S (Fujifilm) as described in previous chapters.

3.2.3 Synthesis of competitor DNA

The 16S region of the mtDNA sequence of *Bos taurus* (common cow) was used as base for the synthesis of competitive DNA for analysis of *T. alalunga*, *T. albacares* and *K.pelamis*. DNA from *B. taurus* was used in order to ensure that, excepting the priming region, the sequence of the competitor DNA would be different enough from the sequence of *Thunnus* species. Three primer pairs, one for each species, were designed. Primers were purchased from Operon and prepared as described in previous chapters. These primers were used in a PCR reaction under standard conditions (see Chapter 2) on total DNA of *Bos taurus*. The PCR products were applied to electrophoresis and then excised from the

agarose gel using a HiYield™ Gel Extraction Kit (Real Biotec Corporation). Briefly, the portion of the gel containing the PCR product was excised with a lancet, cut in slices and put into a Eppendorf tube. Then, 500 µL of lysis buffer containing guanidine chloride were added and tubes were incubated at 45°C for 15 min, shaking manually until dissolving completely the gel. After this, the slurry was transferred to the column provided in the kit and centrifuged 30 s at 1000 *g*. The filtrate was discarded, 300 µL of washing buffer were added to the column and it was centrifuged once again 30 s at 10 000 *g*. The filtrate was discarded and the column was dried by centrifugation, 2 min at 10 000 *g*. Finally, 30 µL of elution buffer were added and 2 min later the column was placed in a new Eppendorf tube and centrifuged for 2 min at 10 000 *g* to recover the PCR product. Competitive DNA prepared and purified this way was kept at 4°C until use. Further 1:10 dilutions of competitive DNA were done with TE buffer.

3.2.4. Competitive PCR amplification

All PCR reactions were performed in a Veriti® Thermal Cycler (Applied Biosystems) using ExTaq™ DNA polymerase (TaKaRa). The composition of the PCR reaction was as follows: 2.5 µL of 10x PCR buffer, 1 µL of dDNP mixture (final concentration, 0.1 mM each), 0.75 µL of 20 µM SSPP primer (final concentration, 0.6 µM each), 2 µL of DNA template (60 ng), 1 µL of the respective dilution of competitive DNA, 0.2 µL of Ex Taq DNA polymerase and distilled water to complete 25 µL. PCR amplification conditions were as follows: 1 cycle of 30 s at 98°C and 25 cycles of i) 10 s at 98°C, ii) 15 s at 54°C (60°C for assays of *T. albacares*), iii) 20 s at 72°C. Afterwards, PCR products were applied to electrophoresis in 2.5% agarose gel stained with GelGreen™

(Biotium), with 100 V of electric current for 30 min and read in a ImageQuant™ LAS 4000-mini transilluminator (GE Healthcare). Light densities of each band were measured and used to perform the concentration analysis.

3.3. Results and discussion

3.3.1. Description of the competitive PCR quantification system

The principle of the competitive PCR is the use of a DNA sequence that bears the same priming regions as the target DNA, so both will compete as substrates of the DNA polymerase. Since the priming region is the same in target DNA and competitor DNA, thus eliminating the influence of mispriming, steric strain, etc, the effective amplification rate will depend only on the initial concentration of both templates. By repeating the same PCR analysis with decreasing amounts of competitor and measuring the relative change of light density of both PCR products, it is possible to determine the equivalence point and therefore the initial concentration of target DNA. The schematic mechanism of competitive PCR is explained in Fig. 3-1. Some of the main advantages of competitive PCR are: i) it can be done using conventional equipment, unlike real-time PCR; ii) it provides an internal control of the PCR reaction, as the competitor DNA will be affected by any inhibitor that may have been carried during DNA extraction; iii) it can be used in mixtures of DNA by using specific primers.

3.3.2 Synthesis of competitor DNA

Competitive DNAs for the quantitative analysis of *T. alalunga*, *T. albacares* and *K. pelamis* were prepared using mtDNA of *B. taurus* (common cow) as starting template. First,

a sequence of the complete mtDNA of *B. taurus* stored in the NCBI GenBank (Accesion Number, AF492351) was used to design three primer pairs, (TYN4-Uni L/H for *T. albacares*, TAN5-Uni L/H for *T. alalunga*, and KPN5-Uni L/H for *K.pelamis*). Each primer consists in two sections: The 3'- half is a sequence that binds the mtDNA of *B. taurus*. The 5'- half is the sequence of the respective tuna short SSPP (Fig. 3-2). By using these primer pairs in a PCR reaction on total DNA of *B. taurus* as template, three hybrid PCR products were generated, each one containing a sequence of the respective tuna species in both extremes and an internal sequence of *B. taurus*. The PCR products were excised from the agarose gel and purified, and their concentration was measured. The final concentration of the three competitors was: 3.50 ng/µL for *T.albacares*, 4.05 ng/µL for *T. alalunga*, and 5.73 ng/µL for *K.pelamis*. The DNA competitors were further diluted with TE buffer in series of 1:10 and used in the respective PCR amplifications. The length of the competitor was 99 bp for *T. albacares*, 101 bp for *T. alalunga* and 100 bp for *K.pelamis*.

3.3.3. Competitive PCR

Although *T. albacares*, *T. alalunga* and *K. pelamis* are the tuna species most commonly used for canned tuna, the use and proportion of each species in different products varies depending on the country. In America, for example, products labeled as "white tuna" cannot include any species but *T. alalunga*, while products labeled as "light tuna" can contain any species of tuna without limitation. Nevertheless, *K. pelamis* is the species with the lowest prize and largest relative abundance, for which it has become the most used species for illegal substitution in canned tuna. For this reason, the quantification method of these three species was designed to determine the content of *K.pelamis* in *T.*

albacares and in *T. alalunga* products. First, raw tissue of one *Thunnus* species was mixed in different proportions with raw tissue of *K. pelamis* and total DNA was extracted. This DNA was used as template for two series of competitive PCR reactions, one using SSPP and competitive DNA for the *Thunnus* species and other using the SSPP and the competitive DNA for *K. pelamis*. The same was done for the other *Thunnus* species. The summary of results of both analyses is shown in Table 3-1. A representative electrophoretogram of the competitive PCR results as well as the respective light density plots are also shown in Fig. 3-3. From the measurement of total extracted DNA it was seen that, although the amount of tuna tissue used for extraction was the same in all cases (15 mg), the final amount of extracted DNA was not constant: Total DNA extracted in the mixtures of *T. albacares* and *K.pelamis* ranged from 11.64 ng/µL (99:1 mixture) to 25.24 ng/µL (50:50 mixture). Similarly, total DNA extracted in mixtures of *T. alalunga* and *K.pelamis* ranged from 9.11 ng/µL (99:1 mixture) to 27.05 ng/µL (50:50 mixture). The maximum concentrations of DNA were obtained in the samples with highest proportion of *Thunnus* tissue but this is a mere coincidence, as no relation between the proportion of the two species and the final DNA concentration was found. The variable results of final DNA concentration proved that the extraction method is not quantitative. However, since the DNA proportion of both species remains constant regardless the overall yield of the extraction, it is possible to establish the relative proportion even if the extraction is not quantitative. Under the conditions of the current method, it was possible to establish accurately a proportion of *Thunnus*/*Katsuwonus* DNA reflecting the original proportion of tissue with up to a level of 5% of *K.pelamis*. At lower proportions, it was still possible to identify the presence of *K. pelamis* but the quantified amount did not reflect the original

proportion either in mixtures with *T. alalunga* or with *T. albacares*. Thus, the quantification limit under the present conditions can be established as 5% of *K. pelamis* in *Thunnus* tissue. This limit is enough to differentiate accurately between accidental admixture and intentional replacement, since it is considered that contaminations can be controlled to a below that level with good manufacture practices (Wolf & Lüthy, 2001). The analysis of the 95:5 mixture of *T. albacares* and *K. pelamis* as well as the analysis of the 90:10 mixture of *T. alalunga* and *K. pelamis* were performed 4 times to assess the reproducibility of the method. As a result, it was confirmed that the method can be performed repeatedly producing reproducible results.

The next, mixtures of DNA from the three species (*T. albacares* and *K. pelamis* or *T. alalunga* and *K. pelamis*), taken from canned tuna, were prepared and analyzed in a similar manner as with raw tissue. For DNA extraction, 30 mg of canned tuna were used. The total DNA was used as a template for analysis of the three species with the competitive PCR-SSPP method. In addition, a sample of a commercial tuna product (*kakuni*, cubes of seasoned tuna meat) that had been analyzed with the SSPP/TSPP-PCR discrimination method, where both *T.albacares* and *K.pelamis* had been detected (Sample #45, see Chapter 2) was also analyzed, in order to determine the content of these two species. Results are summarized in Table 3-2 as well as in Fig. 3-4. As in the case of unheated samples, it was seen that the DNA extraction was not quantitative. The total DNA extracted from canned tuna samples was lower than in case of unheated tissue, as it was expected, considering degradation due to thermal processing. The range of extracted total DNA in mixtures of canned tissue of *T. albacares* and *K. pelamis* was 4.22 to 5.17 ng/μL and in mixtures of *T. alalunga* and *K.pelamis* was 3.67 to 7.85 ng/μL. As in the case of unheated

samples, the quantitative analysis of heat-processed samples was accurate to a level of 5% of the second species (*K.pelamis*) and results were also significantly reproducible. The analysis of the commercial sample (tuna seasoned cubes, sample #60), revealed a proportion of 2.5% of *K.pelamis* and 97.5% of *T. albacares*. This level of *K.pelamis* falls below the quantification limit hereby established and thus, it is not possible to determine the accurate content of *K.pelamis*. Even so, as it is less than 5%, it is reasonable to assume that this level of *K.pelamis* is rather an incidental rather than intentional admixture.

During the quantification of DNA from heat-processed tuna, the amount of total original DNA template measured was around one order of magnitude inferior to that of unheated samples. Considering that the amount of total DNA only decreased to about one-half to one-third of the total DNA extracted from not heat-treated samples, it can be seen that there is not a lineal relation between the change in extracted DNA and the change in the effective amplifiable amount of target template. This implies that DNA extracted from food products with relatively low concentrations of total DNA may be analyzed satisfactorily despite such low concentrations, provided that there is enough target template. In the experiments with DNA from heat-processed products, the degradation of DNA was also reflected on the light density plot. Compared to the light density curve of unheated samples (Fig. 3.3), the one from canned tuna (Fig. 3-4) had a smoother slope and they reached a plateau at lower light densities. It is possible that in samples with highly degraded DNA, the light density curves become too soft to measure accurately the equivalence point respect to the competitor DNA. In such cases, it may be necessary to increase the initial amount of template, regardless the amount of measured total DNA

3.3.4 Conclusions

The quantitative method developed in this research was applied successfully to the analysis of three tuna species. The method exhibited good reproducibility and accuracy and could be applied even in canned products. Unlike the SSPP/TSPP-PCR discrimination method, that requires two different primer pairs to ensure specificity, the simultaneous use of a competitive DNA in the quantitative PCR method allows to use only the SSPP, as the competitor will provide the "buffer" substrate required to avoid excessive DNA polymerase on non-target templates.

Although other quantitative methods, especially those based in real-time PCR, exhibit much lower detection limits, of an order of 0.1 pg (Kesmen et al., 2009), those high levels of sensitivity are of little significance for practice purposes and, in turn, such sensitive methods could become inaccurate when analyzing higher concentrations of DNA. Benedetto et al. (2011) reported that although their real time PCR method could detect up to 0.2 pg of fish DNA in flour, it became inaccurate when applied to samples with >2.5% of fish meat in flour. In comparison, the method here described can detect accurately presence of *K. pelamis* in *Thunnus* products accurately to a level of 5%, which is a proper limit for discrimination between intentional admixture and contamination. Also, since it is based in conventional PCR, it can be performed quickly in a large number of analysis facilities without the need of expensive equipments or reactants and it is susceptible to be standardized for routine assays.

Table 3-1a. Results of the quantitative analysis of mixtures of raw tissue of *T.albacares* and *K.pelamis* using competitive PCR and species-specific primers

T. albacares/*K.pelamis* (weight basis)	Total DNA (ng)*	Total template DNA (fg/ng)	Content of *T.albacares* (%)	Content of *K.pelamis* (%)
50:50	50.48	2.138	50.6	49.4
75:25	39.94	2.308	76.1	23.9
90:10	48.60	1.866	89.8	10.2
95:5 (1)	39.64	2.586	93.1	6.9
95:5 (2)		2.743	95.0	5.0
95:5 (3)		2.436	94.6	5.4
95:5 (4)		2.309	94.9	5.1
95:5 Average**	**39.64**	**2.519**	**94.4 ±0.87**	**5.6 ±0.87**
99:1	23.28	2.792	96.5	3.5

* Indicated the amount of DNA used in the competitive PCR assay. The volumen of DNA template in all assays was 2 µL and thus, the total DNA concentration (extracted from 15 mg of muscle tissue in all cases) can be calculated by dividing the total DNA amount by 2.

**The average values include a confidence interval with α=0.05

Table 3-1b. Results of the quantitative analysis of mixtures of tissue of *T.alalunga* and *K.pelamis* using competitive PCR and species-specific primers

T. alalunga/*K.pelamis* (weight basis)	Total DNA (ng)*	Total template DNA (fg/ng)	Content of *T.alalunga* (%)	Content of *K.pelamis* (%)
50:50	54.10	1.722	53.7	46.3
75:25	31.42	3.060	77.6	22.4
90:10 (1)	27.57	4.425	90.8	9.2
90:10 (2)		5.304	88.6	11.4
90:10 (3)		4.174	88.5	11.5
90:10 (4)		4.344	89.7	10.3
90:10 Average**	**27.57**	**4.56**	**88.55±1.05**	**11.45±1.05**
95:5	28.01	5.404	95.0	5.0
99:1	18.23	3.254	99.6	0.4

* Indicated the amount of DNA used in the competitive PCR assay. The volumen of DNA template in all assays was 2 µL and thus, the total DNA concentration (extracted from 15 mg of muscle tissue in all cases) can be calculated by dividing the total DNA amount by 2.

** The average values include a confidence interval with α=0.05

Table 3-2a. Results of the quantitative analysis of mixtures of canned meat of *T.albacares* and *K.pelamis* using competitive PCR and species-specific primers

T. albacares/*K.pelamis* (weight basis)	Total DNA (ng)	Total template DNA (fg/ng)	Content of *T.albacares* (%)	Content of *K.pelamis* (%)
90:10 (1)		0.253	89.46	10.54
90:10 (2)	10.34	0.435	89.88	10.12
90:10 (3)		0.202	89.65	10.35
90:10 Average*	**10.34**	**0.295**	**89.66±0.24**	**10.34±0.24**
95:5 (1)		0.192	94.23	5.77
95:5 (2)	8.45	0.430	95.87	4.13
95:5 (3)		0.230	95.45	4.55
95:5 Average*	**8.45**	**0.284**	**95.18±0.96**	**4.82±0.96**
99:1 (1)		0.634	98.10	1.90
99:1 (1)	9.34	0.225	97.68	2.32
99:1 (1)		0.662	95.88	4.12
99:1 Average*	**9.34**	**0.507**	**97.22±1.33**	**2.78±1.33**

* The average values include a confidence interval with α=0.05

Table 3-2b. Results of the quantitative analysis of mixtures of canned meat of *T.alalunga* and *K.pelamis* using competitive PCR and species-specific primers

T. alalunga/*K.pelamis* (weight basis)	Total DNA (ng)	Total template DNA (fg/ng)	Content of *T.alalunga* (%)	Content of *K.pelamis* (%)
90:10 (1)		0.385	90.65	9.35
90:10 (2)	15.71	0.753	89.32	10.68
90:10 (3)		0.844	90.35	9.65
Average*	**15.71**	**0.661**	**90.11±0.70**	**9.89±0.70**
95:5 (1)		0.255	95.04	4.96
95:5 (2)	13.53	0.206	95.33	4.67
95:5 (3)		0.645	95.17	4.83
Average*	**13.53**	**0.369**	**95.18±0.30**	**4.82 ±0.30**
99:1 (1)		0.112	96.14	3.86
99:1 (1)	7.34	0.345	97.48	2.52
99:1 (1)		0.382	97.87	2.13
Average*	**7.34**	**0.280**	**97.16±1.03**	**2.84±1.03**

* The average values include a confidence interval with α=0.05

Table 3-2c. Results of the quantitative analysis of tuna seasoned meat for detection of *T.albacares* and *K.pelamis* using competitive PCR and species-specific primers

Total DNA (ng)	Total template DNA (fg/ng)	Content of *T.alalunga* (%)	Content of *K.pelamis* (%)
	0.152	96.52	3.48
18.2	0.255	98.41	1.59
	0.196	97.50	2.50
Average:	**0.201**	**97.48**	**2.52**

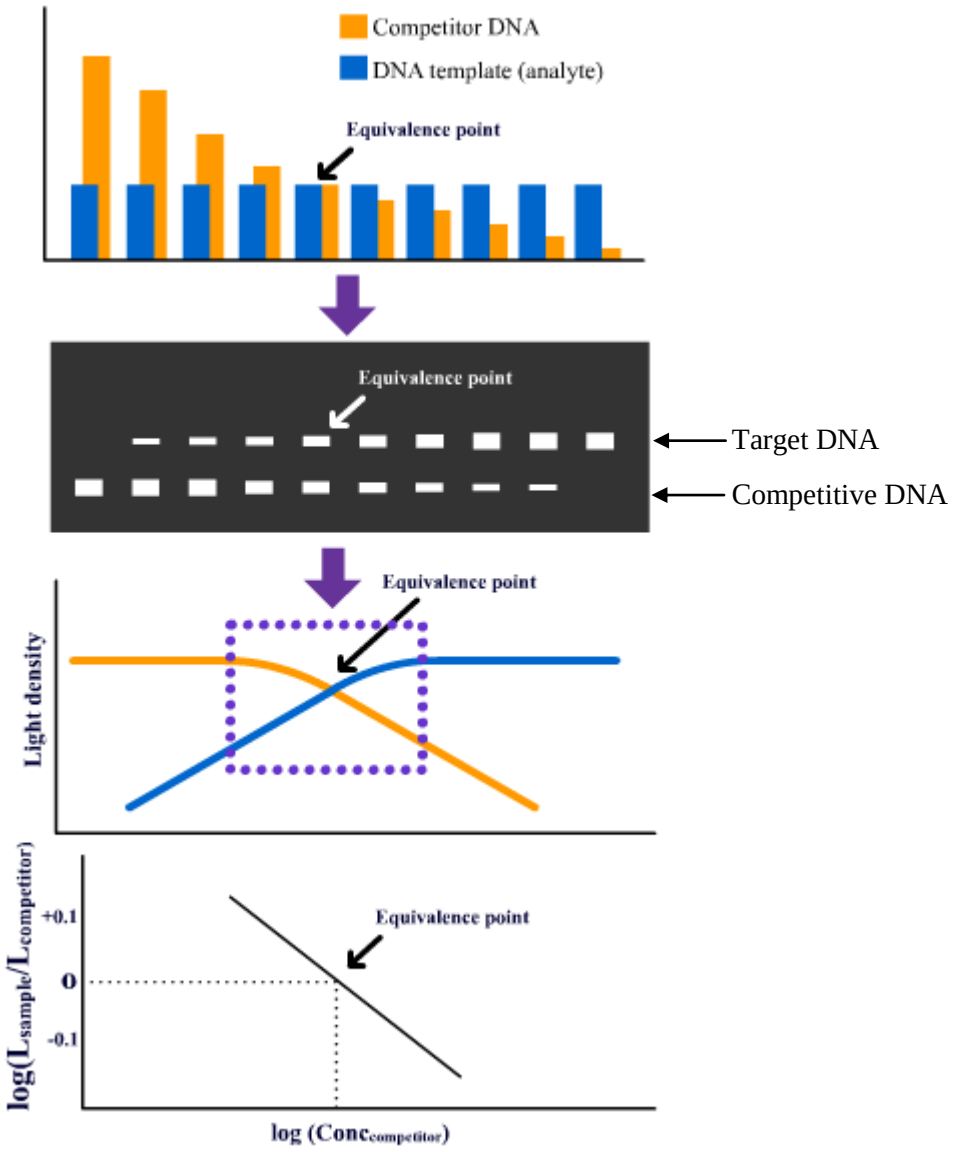


Fig 3.1 Schematic representation of the principles of the competitive PCR for quantitative analysis. A series of PCR amplifications with decreasing amounts of competitive DNA are done and according to the chance of light density in the electrophoretogram, the content of DNA template is determined.

Thunnus albacares

	2460	2470	2480	2490	2500	2510	2520	2530	2540
Bos taurus	*CCGCTATCAA AGGTTCG***TTT GTTCAACGAT TAAAGTCCTA CGTGATCTGA GTTCAGACCG GAGTAATCCA GGTCGGTTTC TAT***CTATTAC**								
Primer TYN4-Uni L/H	**TCCT CATGATTGCC CACGGAATTT GTTCAACGAT TAAAG**						**GTAATCCA GGTCGGTTTC TATCGATGCC CCTTATAACA ACATG**		
Competitor DNA	**TCCT CATGATTGCC CACGGAATTT GTTCAACGAT TAAAGTCCTA CGTGATCTGA GTTCAGACCG GAGTAATCCA GGTCGGTTTC TATCGATGCC CCTTATAACA ACATG**								

Thunnus alalunga

	2460	2470	2480	2490	2500	2510	2520	2530	2540
Bos taurus	*CCGCTATCAA AGGTTC***GTTT GTTCAACGAT TAAAGTCCTA CGTGATCTGA GTTCAGACCG GAGTAATCCA GGTCGGTTTC *TAT***CTATTAC**								
Primer TAN5-Uni L/H	**A TATCCCTCAC TCGGTTGTTT GTTCAACGAT TAAAG**						**GTAATCCA GGTCGGTTTC AATAAAGAGT TGGAACAT**		
Competitor DNA	**A TATCCCTCAC TCGGTTGTTT GTTCAACGAT TAAAGTCCTA CGTGATCTGA GTTCAGACCG GAGTAATCCA GGTCGGTTTC AATAAAGAGT TGGAACAT**								

Katsuwonus pelamis

	2460	2470	2480	2490	2500	2510	2520	2530	2540
Bos taurus	*CCGCTATCAA AGGT***TCGTTT GTTCAACGAT TAAAGTCCTA CGTGATCTGA GTTCAGACCG GAGTAATCCA GGTCGGTTTC TATC***TATTAC**								
Primer KPN5-Uni L/H	**CTTCT CATCACCTCT AACATCGTTT GTTCAACGAT TAAAG**						GTAATCCA GGTCGGTTTC **TATCAGCCTC CCTAACAAGC AAACA**		
Competitor DNA	**CTTCT CATCACCTCT AACATCGTTT GTTCAACGAT TAAAGTCCTA CGTGATCTGA GTTCAGACCG GA**GTAATCCA GGTCGGTTTC **TATCAGCCTC CCTAACAAGC AAACA**								

Fig. 3-2. Partial sequence of the 16S region of mtDNA of *Bos taurus*, used as template for the synthesis of competitor DNAs for the quantitative analysis of *Thunnus albacares*, *T. alalunga*, and *Katsuwonus pelamis*, as well as the sequence of the respective primers and the sequence of the final competitor DNA. Underlined, the sequence of the original short SSPP for tuna species. In italics, parts of the original *Bos taurus* mtDNA sequence that were not included in the final competitor DNA

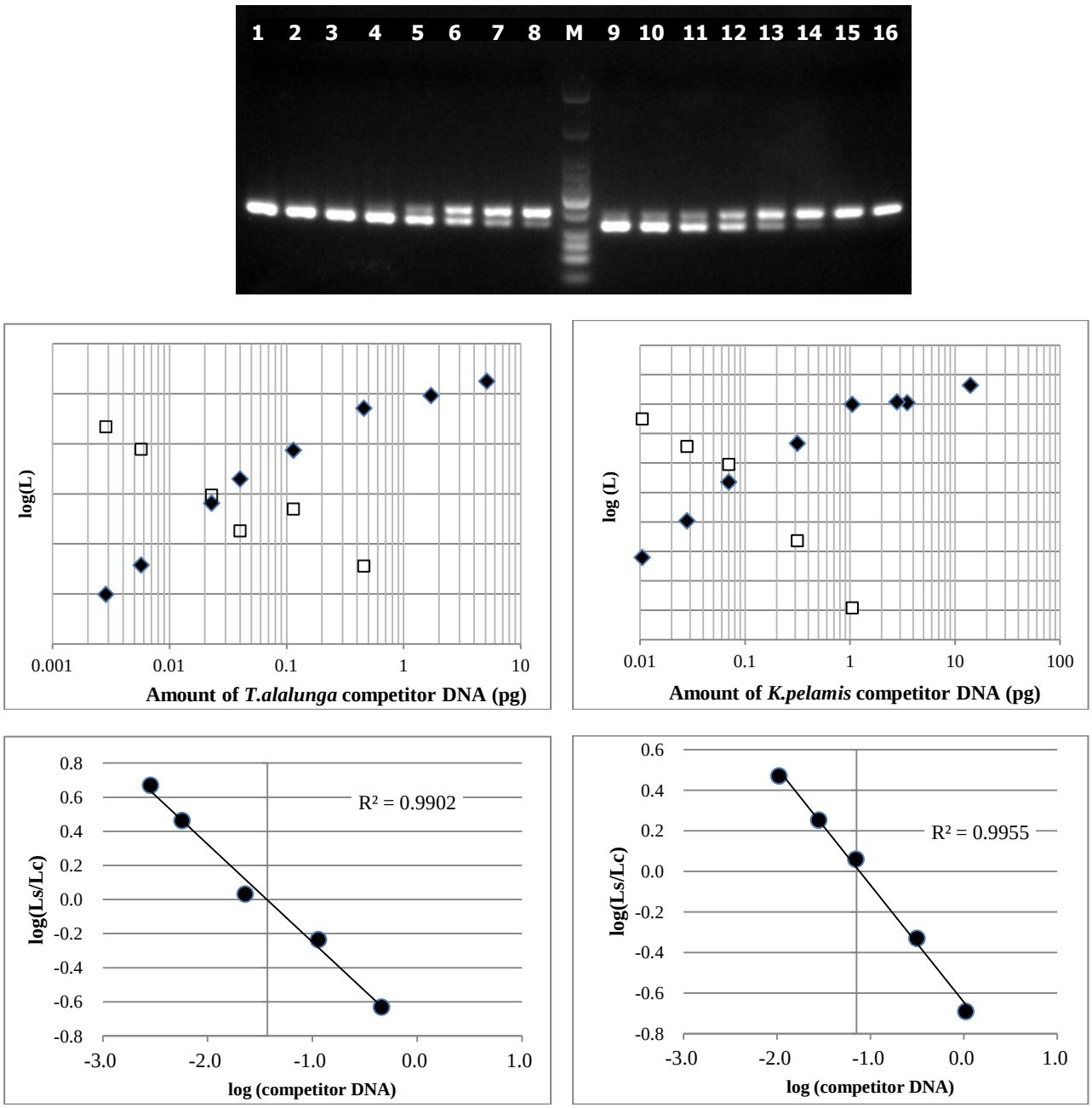


Fig. 3-3. Representative results of the competitive PCR-SSPP method for quantification of tuna in mixtures of *Thunnus alalunga* and *Katsuwonus pelamis*. Left: Light density analysis of *T. alalunga*. Right. Light density analysis of *K. pelamis* ◆Light density of the competitor PCR product. ☐ Light density of the analyte PCR product. Amount of *T. alalunga* competitor DNA: **1**. 14 pg; **2**. 3.5 pg; **3.** 2.8 pg; **4.** 1.05 pg; **5.** 315 fg; **6.** 70 fg; **7.** 28 fg; **8.** 10.5 fg. Amount of *K. pelamis* competitor DNA: **9.** 229 fg **10.** 57.3 fg; **11.** 45.8 fg; **12.** 17.2 fg; **13.** 5.1 fg; **14.** 1.1 fg; **15.** 458 ag; **16.** 171 ag

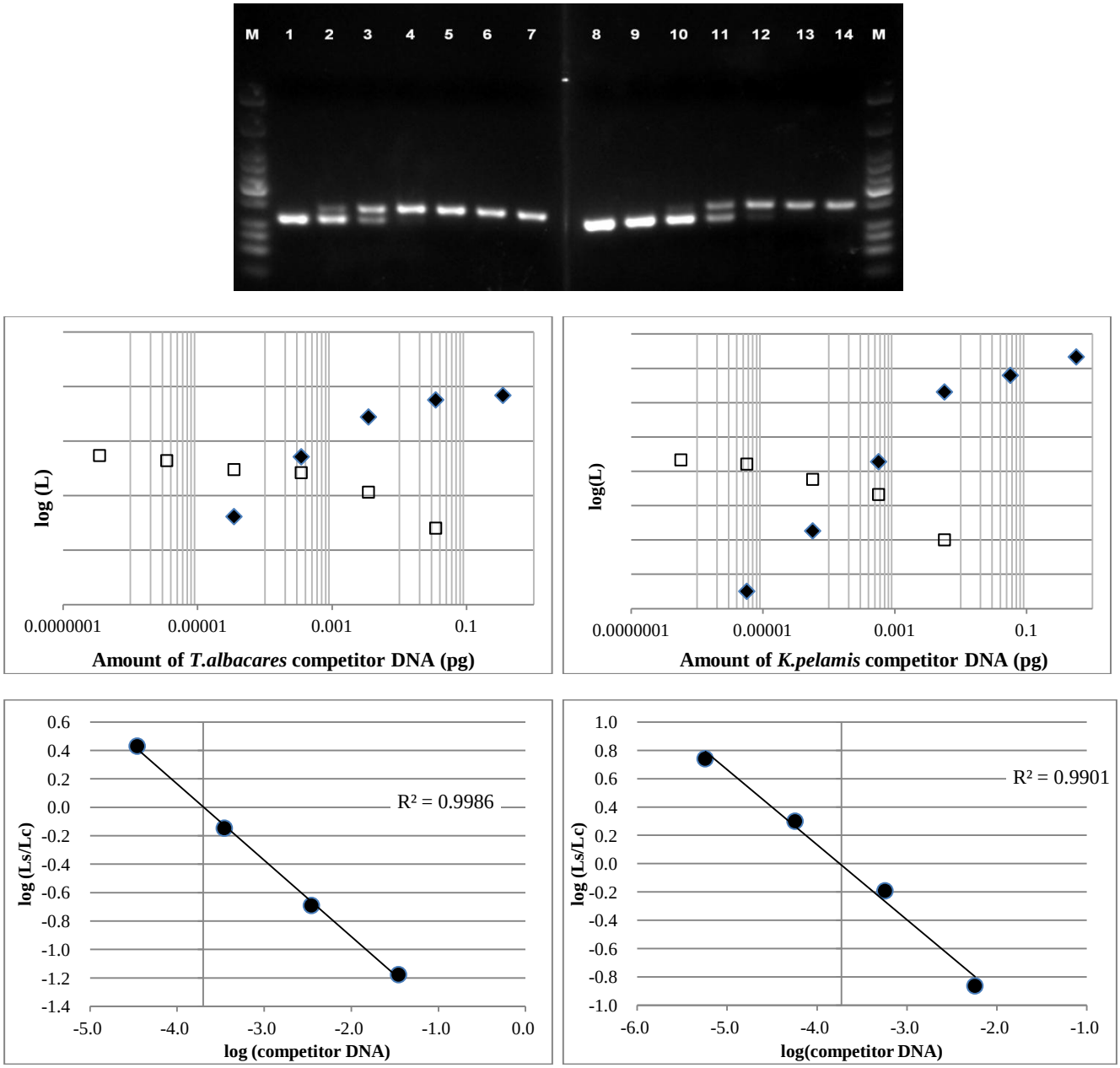


Fig. 3-4. Representative results of the competitive PCR-SSPP method for quantification of *Thunnus albacares* and *Katsuwonus pelamis* in a sample of canned tuna. Left: Light density analysis of *T.albacares*. Right. Light density analysis of *K.pelamis* ◆Light density of the competitor PCR product. ☐ Light density of the analyte PCR product. Amount of *T. albacares* competitor DNA: **1**. 350 pg; **2**. 35 pg; **3.** 3.5 pg; **4.** 0.35 pg; **5.** 35 fg; **6.** 3.5 fg; **7.** 0.35 fg. Amount of *K. pelamis* competitor DNA: **8.** 573 pg **9.** 57.3 pg; **10.** 5.73 pg; **11.** 573 pg; **12.** 57.3 fg; **13.** 5.73 fg; **14.** 573 ag

Chapter 4. General conclusions

In this research, the complete mitochondrial DNA (mtDNA) sequences of 9 fish species, including *Thunnus* species of relevant commercial importance (*T. thynnus*, *T. maccoyii*, *T. orientalis*, *T. alalunga*, *T. obesus*, *T. albacares*, *T. tonggol*), *Kastuwonus pelamis* and *Gasterochisma melampus*, were determined and investigated in order to find suitable polymorphisms and other genetic features to be used as DNA markers for species discrimination. As a result, the phylogenetic relations of these species based on mtDNA could be corroborated and it was established that, despite the high degree of homology within *Thunnus* species, there were genetic features, namely single nucleotide differences (SND) that could be potential candidate for DNA markers.

The next, an allele-specific PCR method was developed to identify 8 different fish species (7 *Thunnus* species plus *K. pelamis*). This method is based in the use of species specific primers constructed on SND located in the mtDNA along with a universal primer in a conventional PCR amplification. PCR products are then analyzed with electrophoresis. Depending on the type of sample, two different sets of species specific primers were developed: i) For raw, frozen and other unheated tuna foods and ii) for heat-processed tuna products. The method was optimized and standardized and then tested in a variety of samples. It was seen that the method was highly specific. The accuracy of the method is comparable to that of sequencing techniques but it can be performed with conventional PCR techniques, thus being a faster and less expensive that these techniques. Likewise, the technical simplicity of the method makes it more practical than other techniques, such as like RFLP, as it can be carried out faster and less probability of operational mistakes and

drawbacks. Furthermore, the current method can be applied to all main tuna species, unlike all of the methods developed previously. Finally, the method also displayed a remarkable efficiency, being able to identify accurately species in samples that typically have been difficult to analyze, such as canned products. Because of these advantageous characteristics, it is considered that the species-specific PCR method developed in this research has potential to be used in practical applications. However, in order to consolidate and improve the method, there are a few issues that could be addressed: i) To study the effect of the mismatch type and position; ii) to conduct a more comprehensive analysis of a variety of samples, including key information such as origin; iii) exploring the possibility of lowering the detection limit by using more refined DNA extraction methods and modifying the primers to optimize its affinity to the template without losing specificity. In particular, regarding the second point, it is worth to note that the principles used in the present method could be extended and applied for classification of tuna species according to their geographic origin, which would be a great asset the for research and commerce of tuna species.

On the other hand, a method based in competitive PCR for quantification of tuna in mixtures was also developed. This method proved to be effective to detect and quantify levels of *K. pelamis* in *T. albacares* or *T. alalunga* samples, detecting admixture levels of 5%, thus being regarded as a good tool for quantification of extraneous samples and discrimination between accidental contamination and intended replacement. The three species that can be detected with this competitive PCR method are by far the most used in the industry of canned tuna but from now on, it would be desirable to further investigate how to detect and quantify other tuna species, such as *T. tonggol*, *T. obesus* and even bluefin tuna species.

References

Appleyard, SA; Grewe, PM; Innes, BH; Ward RD Population structure of yellowfin tuna (*Thunnus albacares*) in the western Pacific Ocean, inferred from microsatellite loci. *Marine Biol.* (2001) **139**, 383-393.

Asahida, T; Kobayashi, T; Saito, K; Nakayama, I Tissue preservation and total DNA extraction from fish at ambient temperature using buffers containing high concentration of urea. *Fish. Sci.*(1996) **62**, 727-730.

Asensio L; González I; García, T; Martín, R Determination of food authenticity by enzyme-linked immunosorbent assay (ELISA). *Food Control* (2008) **19**, 1-8.

Benedetto, A; Abete, MC; Squadrone, S Towards a quantitative application of real-time PCR technique for fish DNA detection in feedstuffs. *Food Chem.* (2011) **126**, 1436-1442.

Bottema CDK; Sarkar, G; Cassady, JD; Li, S; Dutton, CM; Sommer SS. Polymerase chain reaction amplification of specific alleles: A general detection method of detection of mutations, polymorphisms and haplotypes. *Methods in Enzymol.* (1993) **218**, 388-402.

Botti, S; Giuffra, E. Oligonucleotide indexing of DNA barcodes: Identification of tuna and other scombrid species in food products. *BMC Biotechnology* (2010) **10**, http://biomedcentral.com/1472-6750/10/60.

Broughton, RE; Reneau, PC Spatial covariation of mutation and non-synonymous substitution rates in vertebrate mitochondrial genomes. *Mol. Biol. Evol*. (2006) **23 (8)**, 1516-1524.

Carlsson, J; McDowell, JR; Díaz-Jaimes, P; Carlsson, JEL; Boles, SB; Gold, JR; Graves, JE Microsatellite and mitochondrial DNA analyses of Atlantic bluefin tuna (*Thynnus thynnus thynnus*) population structure in Mediterranean Sea. *Mol. Ecol.* (2004) **13**, 3345-3356.

Cascorbi, A; Chabot H. Seafood Watch Report. Tunas. Volume II. Atlantic Bluefin Tuna. (2004). Monterrey Bay Aquarium.

Chapela, MJ; Sotelo, CG; Pérez-Martin, RI; Pardo, MA; Pérez-Villareal, B; Gilardi, P; Riese, J. Comparison of DNA extraction methods from muscle of canned tuna for species identification. *Food Control* (2007) **18**, 1211-1215.

Chow, S; Inoue, S Intra and interspecific restriction fragment length polymorphism in mitochondrial genes of *Thunnus* tuna species. *Bull. Nat. Res. Inst. Far Seas Fish* (1993) **30,** 207-224.

Chow, S; Kishino, H Phylogenetic relations between tuna species of the genus *Thunnus* (Scombridae: Teleostei): Inconsistent implications from morphology, nuclear and mitochondrial genomes. *J. Mol. Evol.* (1995) **41**, 741-748.

Chow, S; Nakagawa, T; Suzuki, H; Takeyama, H; Matsunaga, T Phylogenetic relations among *Thunnus* species inferred from rDNA ITS1 sequence. *J. Fish Biol.* (2006) **68** Suppl. A, 24-35

Chow, S; Nohara, K; Tanabe, T; Ito, T; Tsuji, S; Nishikawa, Y; Ueyanagi, S; Uchikawa, K Genetic and morphological identification of larval and small juvenile tunas (Pisces: Scombridae) caught by a mid-water trawl in the western Pacific. *Bull.Fish. Res. Agen.* (2003) **8**, 1-14.

Chow, S; Okamoto, H; Miyabe, N; Hiramatsu, K; Barut, N Genetic divergence between Atlantic and Indo-Pacific stocks of big eye tuna (*Thunnus obesus*) and admixture around South Africa. *Mol. Ecol.* (2000) **9**, 221-227.

Chun, JY; Kim, KJ; Hwang, IT, Kim, YJ Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of CYP2C19 gene. *Nucleic Acids Res.* (2007) **35**, e40 doi:10.1093/nar/gkm051.

FAO Fisheries and Aquaculture Information and Statistics Service – FISHSTAT http://www.fao.org/fishery/en.

Garcés-Claver, A; Fellman, SM; Gil-Ortega, R; Jahn, M; Arnedo-Andrés, MS Identification, validation and survey of a single nucleotide polymorphism (SNP) associated with pungency in *Capsicum* spp. *Theor. Appl. Genet.* (2007) **115**, 907-916.

Gibson, NJ The use of real-time PCR methods in DNA sequence variation analysis *Clin. Chim. Acta* (2006) **363**, 32-47.

Hayashi, K; Hashimoto, N; Daigen, M; Ashikawa, I Development of PCR-based SNP markers for rice blast resistance genes at the Piz locus. *Theor. Appl. Genet.* (2004) **108**, 1212-1220.

Infante, C; Catanese, G; Ponce, M; Manchado, M Novel method for the authentication of frigate tunas (*Auxis thazard* and *Auxis rochei*) in commercial canned products, *J. Agric. Food Chem.* (2004) **52**, 7435-7443.

Jacquet, JL; Pauly, D Trade secrets: Renaming and mislabeling of seafood. *Marine Policy* (2008) **32**, 309-318.

Kesmen, Z; Gulluce, A; Sahin, F; Yetim, H Identification of meat species by TaqMan-based real-time PCR assay. *Meat Science* (2009) **82**, 444-449.

Kitaoka, M; Okamura, N; Ichinose, H; Goto, M Detection of SNP in fish DNA: Application of the fluorogenic ribonuclease protection (FRIP) assay for the authentication of food contents *J. Agric. Food Chem.* (2008) **56**, 6246-6251.

Lin, WF; Hwang, DF Application of PCR-RFLP analysis on species identification of canned tuna. *Food Control* (2007) **18**, 1050-1057.

Little, S Amplification-Refractory Mutation System (ARMS) analysis of point mutations. In: Current Protocols in Human Molecular Genetics (1994), John Wiley & Sons, New York. pp 9.8.1-9.8.12.

Liu, WD; Li, HJ; Bao, XB; He, CB; Li WJ The first set of EST-derived single nucleotide polymorphism markers for Japanese scallop, *Patinopecten yessoensis*. *J. World Aquaculture Soc.* (2011) **42**, 456-461.

Liu, ZJ; Cordes, JF DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* (2004) **238,** 1-37.

Lockley, AK; Bardsley, RG DNA-based method for food authentication. *Trends in Food Sci. & Nutr*. (2000) **11**, 67-77.

Lowenstein, JH; Amato, G; Kolokotronis, SO The real maccoyii; Identifying tuna sushi with DNA barcodes – Contrasting characteristic attributes and genetic distances. *PLoS ONE* (2009) **4** (11): e7866.doi:10.1371/journal.pone.0007866.

Mackie, IM; Pryde, SE; Gonzales-Sotelo, C; Medina, I; Pérez-Martin, R; Qunteiro, J; Rey-Mendez, M, Rehbein, H Challenges in identification of species of canned fish. *Trends Food Science Tech.* (1999) **10**, 9-14.

Manchado, M; Catanese, G; Infante, C Complete mitochondrial DNA sequence of the Atlantic bluefin tuna *Thunnus thynnus*. *Fish. Sci.*(2004) **70**, 68-73.

Mazzeo MF; De Giulio B; Guerreiro, G; Ciarcia G; Malorni, A; Russo, GL; Siciliano, RA Fish authentication by MALDI-TOF mass apectrometry. *J. Agric. Food Chem.* (2008) **56**, 11071-11076.

Nagase, M; Aimi, T; Suginaka, K; Kitamoto, Y; Morinaga, T Complete mitochondrial DNA sequence of the Japanese flying fish *Cypselurur hiraii*. *Fish. Sci*. (2005) **71**, 914-923.

Newton, CR; Graham, A; Heptinstall, LE; Powell, SJ; Summers, C; Kalsheker, N; Smith, JC; Markham, AF Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS) *Nucleic Acids Res.* (1989) **17**, 2503-2516.

Noeack, K; Zardoya, R; Meyer, A The complete mitochondrial DNA sequence of the bichir (*Polypterus ornatipinnis*), a basal ray-finned fish: Ancient establishment of the consensus vertebrate gene order. *Genetics* (1996) **144**, 1165-1180.

Pardo MA; Pérez-Villareal, B Identification of comercial canned tuna species by restriction site analysis of mitochondrial DNA products obtained by nested primer PCR. *Food Chem.* (2004) **86**, 143-150.

Pattemore, JA; Rice, N; Marshall, DF, Waugh, R; Henry, RJ Cereal variety identification using MALDI-TOF mass spectrometry SNP genotyping *J. Cereal Sci.* (2010) **52**, 356-361.

Quinteiro, J; Sotelo CG; Rehbein, H; Pryde, SE; Medina, I, Rey-Méndez, M; Mackie, IM Use of mtDNA direct polymerase chain reaction (PCR) sequencing and PCR-restriction fragment length polymorphism methodologies in species identification of canned tuna. *J. Agric. Food Chem.* (1998) **46**, 1662-1669.

Rehbein, H; Etienne, M; Hattula, T; Knudsen, LB; Jessen, F; Luten, JB; Bouquet, W; Mackie, IE; Ritchie, AH; Martir, R; Mendes, R Influence of variation in methodology on the

reliability of the isoelectric focusing method of fish species identification. *Food Chem.* (1995) **52**, 193-197.

Rehbein, H; Mackie IM; Pride, S; Gonzales-Sotelo, C; Medina, I; Pérez-Martin, R; Quinteiro, J; Rey-Mendez, M Fish species identification in canned tuna by PCR-SSPP: validation by a collaborative study and investigation of intra-species variability of the DNA patterns. *Food Chem.* (1999) 64, 263-268.

Safina, C; Klinger, DH Collapse of bluefin tuna in the western Atlantic. *Conservation Biol.* (2008) **22**, 243-246.

Takashima, Y; Morita, T; Tamashita, M Complete mitochondrial DNA sequence of Atlantic horse mackerel *Trachurus trachurus* and molecular identification of two commercially important species *T. trachurus* and *T. japanicus* using PCR-RFLP. *Fish. Sci.* (2006) **72,** 1054-1065.

Takeyama, H; Tsuzuki, H; Chow, S; Nakayama, H; Matsunaga, T Discrimination between Atlantic and Pacific subspecies of Northern bluefin tuna (*Thunnus thynnus*) by magnetic-capture hybridization using bacterial magnetic particles. *Mar. Biotechnol.* (2000) **2**, 309-313.

Tamura, K; Peterson, D; Peterson, N; Stecher, G; Nei, M; Kumar, S MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. *Mol. Biol. Evol*. (in press). doi: 10.1093/molbev/msr121.

Terol, J; Mascarell, R; Fernández-Pedrosa, V; Pérez-Alonso, M Statistical validation of the identification of tuna species: Bootstrap analysis of mitochondrial DNA sequences. *J. Agric. Food Chem.* (2002) **50,** 963-969.

Unseld, M, Beyermann, B; Brandt, P; Hiesel, R Identification of the species origin of highly processed meat products by mitochondrial DNA sequences. *PCR Methods Appl.* (1995) **4**, 241-243.

US Code of Federal Regulations. 21CFR161.190 http://www.law.cornell.edu/cfr/

Viñas, J, Tudela, S A validated methodology for genetic identification of tuna species (genus *Thunnus*) PLoS ONE (2009) **4** (10): e7606. doi: 10.1371/journal.pone.0007606.

Wang, H; Guo, X Identification of *Crassostrea ariakensis* and related oysters by multiplex species-specific PCR. *J. Shellfish Res.* (2008) **27**, 481-487.

Warashina, I; Hisada, K Geographical distribution and body length composition of two tuna-like fishes, *Gasterochisma melampus* Richardson and *Allothunnus fallai* Serventy, taken by Japanese tuna longline fishery. *Bull. Far Seas Fish. Res. Lab.* (1972) **6**, 12-19

Wolf, C; Lüthy, J Quantitative competitive PCR for quantification of porcine DNA. *Meat Science* (2001) **57**, 161-168.

Xia, X Maximizing transcription efficiency causes codon usage bias. *Genetics* (1996) **144**, 1309-1320.

Yu, J; Kong, X; Yu, Z; Zhou, L Comparative analysis of mitochondrial control region sequence from three flatfish species (Pleuronectidae) *J. Ocean Univ. China*. (2005) **4**, 80-84.

Acknowledgements

This research would not have been possible without the support of many persons who helped me during its conduction. More than anyone else, I want to express my gratitude to Dr. Shoichiro Ishizaki, who had the patience to be my supervisor during these three years, as well as to Dr. Yuji Nagashima, who accepted me in the Laboratory of Marine Biomaterials and Functional Biochemistry as a foreign student. They not only provided me with the resources to carry out high-level research but also did their best to make my stay in Japan a pleasant experience beyond the lab walls. My most sincere thanks to both of them. I also want to thank them as well as to Dr. Kuniyoshi Shimakura for their time in reviewing this work.

To all the students of the Laboratory of Marine Biomaterials and Functional Biochemistry, I want to thank them for being sincere fellows to me and for providing me aid and advice whenever I required it. I hope that our time together was as useful for them as it was for me, an opportunity to contribute to our mutual growth.

This work was carried out in the frame of the MEXT Scholarship program. I am deeply in debt with Japan and its people for bringing me the opportunity of coming to this country and experiencing its culture and lifestyle. The memories of the friends I made here will always be in my heart.

Finally, I want to thank my mother and sister. Either near or far, in Mexico or Japan, we know we are there for each other.