12-3-2021

Identification of Tuna Species in Raw and Processed Products Using DNA Mini-barcoding of the Mitochondrial Control Region

Jiahleen Roungchun
Chapman University

Amanda M. Tabb
Chapman University

Rosalee S. Hellberg
Chapman University, hellberg@chapman.edu

Recommended Citation

Roungchun, J., Tabb, A. M., & Hellberg, R. S. (2021). Identification of tuna species in raw and processed products using DNA mini-barcoding of the mitochondrial control region. *Food Control, 134*, 108752. https://doi.org/10.1016/j.foodcont.2021.108752

Identification of Tuna Species in Raw and Processed Products Using DNA Mini-barcoding of the Mitochondrial Control Region

Title: Identification of tuna species in raw and processed products using DNA mini-barcoding of the mitochondrial control region

Authors: Jiahleen B. Roungchun[a], Amanda M. Tabb[a], and Rosalee S. Hellberg[a]*

[a]Chapman University, Schmid College of Science and Technology, Food Science Program, One University Drive, Orange, CA 92866

***Corresponding Author**

Rosalee S. Hellberg, Ph.D.

Chapman University

One University Drive

Orange, CA 92866

Phone: (714) 628-2811

E-mail: hellberg@chapman.edu

Abstract

Accurate species identification methods are needed to combat tuna fraud, improve tuna stock regulation, and mitigate health risks associated with mislabeled tuna products. The objective of this study was to conduct a market survey of raw and processed tuna products using a DNA mini-barcoding system based on the mitochondrial control region (CR). A total of 80 samples of raw, dried, and canned tuna products were collected at the retail level for CR mini-barcoding analysis. The samples underwent DNA extraction, polymerase chain reaction (PCR), and DNA sequencing of the 236-bp CR mini-barcode. The resulting sequences were searched against GenBank using the nucleotide Basic Local Alignment Search Tool (BLAST) to determine the species. The study achieved species identification for 100% of the raw samples, 95% of the dried samples, and 50% of the canned samples, for an overall success rate of 86% (69/80 samples). Mislabeling occurred in 11 of the identified samples (16%), including 8 products marketed as raw, dried, or canned yellowfin tuna, 2 samples marketed as dried or canned skipjack tuna, and 1 raw fillet sold as bluefin tuna. Overall, the DNA mini-barcoding system proved to be a promising method in identifying tuna species in both raw and processed samples. However, testing with a secondary marker is required in some cases to resolve instances of possible species introgression. Future research should explore optimization of this method for improved identification of canned tuna samples.

Keywords: canned tuna, DNA barcoding, mini-barcoding, mislabeling, seafood fraud, species identification

1. Introduction

Seafood is a staple of the global food supply, with close to 180 million tonnes produced in 2018 (FAO, 2020). Global seafood production is dominated by finfish, with the most captured groups – small pelagics, gadiformes, tuna and tuna-like species – making up 85% of total production (FAO, 2020). Tuna captures have consistently increased each year, reaching approximately 8 million tonnes in 2019. Around 58% of these tuna captures consisted of skipjack (*Katsuwonus pelamis*) and yellowfin tuna (*Thunnus albacares*) (FAO, 2020). In 2019, the combined landings and imports of fresh and frozen tuna in the U.S. totaled 500 thousand tonnes, 62% of which was used for canning. Canned tuna ranks as the third-most consumed seafood in the U.S. after shrimp and salmon, with per capita annual consumption at 1.0 kg in 2019 (NFI, 2021; NMFS, 2021).

With an increasing amount and variety of seafood in the global marketplace, fraud has become a major concern throughout the seafood supply chain (FAO, 2018; Pardo et al., 2016; Silva et al., 2021). The intentional substitution of fish species for economic gain is a form of seafood fraud that has been challenging to combat due to the similar appearances of various species and fluctuations in quality, supply, and demand of specific seafood products. Tuna, in particular, is susceptible to fraudulent activity due to its high production, growing popularity, and disparate prices between species. For instance, in 2019, the average ex-vessel price for skipjack tuna was $1.21/kg, while the average ex-vessel price for bluefin tuna was $8.36/kg (NMFS, 2021). In addition to the economic deception associated with tuna fraud, there are health risks as well. For example, some types of tuna have elevated levels of mercury and, therefore, at-risk individuals are advised to eat only one serving per week of yellowfin or albacore tuna (*Thunnus alalunga*), and to avoid bigeye tuna (*Thunnus obesus*) (FDA/EPA, 2019). Mislabeling of these

47 tuna species as a lower-mercury fish, such as skipjack tuna, could lead to potentially unsafe
48 levels of mercury exposure in these at-risk consumers. Furthermore, “white tuna” sold at sushi
49 restaurants in the U.S. has frequently been identified as escolar (*Lepidocybium flavobrunneum*),
50 an oily fish that contains high levels of gempylotoxins and can cause gastrointestinal distress in
51 sensitive individuals (FDA, 2020; Lowenstein et al., 2009; Warner et al., 2013). Substitution of
52 closely related tuna species also hinders the effective recording and conservation of certain
53 species, such as the endangered Atlantic bluefin tuna (*Thunnus thynnus*) and the critically
54 endangered Southern bluefin tuna (*Thunnus maccoyii*) (FAO, 2018; Liou et al., 2020; Viñas &
55 Tudela, 2009).

56 Morphological characteristics, such as color, fin shape, and head structure, are often used
57 to distinguish species of whole fish from one another. However, morphological features are
58 removed during processing, making it difficult to visually identify fish species that are similar in
59 appearance (Mitchell & Hellberg, 2016; Pollack et al., 2018). In these instances, analytical
60 methods, such as those based on protein or DNA analysis, are required for the accurate
61 identification of fish species (Silva & Hellberg, 2021). In the United States, the Food and Drug
62 Administration (FDA) has implemented a method based on DNA barcoding for the regulatory
63 identification of fish species (Handy et al., 2011). DNA barcoding differentiates species by
64 analyzing the genetic diversity in standardized sequences of DNA, referred to as DNA barcodes
65 (Hebert et al., 2003). The standard DNA barcoding procedure for fish species identification
66 targets a ~650 base pair (bp) region within the cytochrome *c* oxidase subunit 1 (CO1)
67 mitochondrial gene (Ward et al., 2005). While standard (full-length) DNA barcoding has been
68 effective for identification of raw or minimally processed fish, it is challenging to obtain a full-
69 length sequence in products that have been canned because the DNA is highly fragmented

(Shokralla et al., 2015). One way to overcome the limitations of full DNA barcoding in processed seafood has been the implementation of DNA mini-barcoding systems that use shorter genetic regions (~150-300 bp) to identify fish species (Mitchell & Hellberg, 2016; Pollack et al., 2018; Shokralla et al., 2015).

While the combined use of full and mini-barcoding based on COI has been highly successful in identifying fish species, these methods are often inadequate in discriminating closely related tuna species due to low genetic divergences (Lowenstein et al., 2009; Mitchell & Hellberg, 2016; Pollack et al., 2018; Shokralla et al., 2015). The inability to consistently identify the species of these samples combined with ambiguous market labeling prevents a comprehensive assessment of tuna substitution on the commercial market (Hanner et al., 2011). To overcome these challenges, researchers have explored the use of the mitochondrial control region (CR) supplemented with the nuclear first internal transcribed spacer region (ITS1) for introgressed species (Gordoa et al. 2016; Mitchell & Hellberg, 2016; Viñas & Tudela, 2009). For example, Mitchell & Hellberg (2016) developed a mini-barcoding system for the differentiation of canned tuna species targeting a shorter (236-bp) fragment within the CR. In addition to its applications in canned products, this CR mini-barcoding system has shown high potential for use in raw and lightly processed products (Frigerio et al., 2021; Liou et al., 2020).

Despite the potential applications of CR mini-barcoding in a wide variety of tuna products, it has yet to be broadly applied to investigating tuna species identification in the marketplace, especially with regards to dried products. Therefore, the objective of the current study was to apply CR mini-barcoding to the identification of species in raw, dried, and canned tuna products sold on the commercial market.

2. Materials and Methods

2.1 Sample collection

A total of 80 raw and processed tuna products were collected for testing, including 20 raw (or seared) tuna products, 20 raw (or seared) sushi samples, 20 canned tuna products, and 20 dried tuna (plain dried flakes and dried jerky) products (Table S1). For the purpose of this study, seared products (n = 7) were included in the "raw" categories because sampling was conducted on the raw interior of the product. Products were collected from 12 grocery stores and 10 sushi restaurants in Orange County, CA, as well as 3 online retailers. Raw fillets and sushi samples purchased at retail outlets were transported in coolers with ice packs and stored at -80 °C upon arrival at Chapman University. Prior to tissue collection, the frozen samples were thawed in a refrigerator at 4 °C for 24 h. A single piece of tissue (~25 mg) was obtained from the interior of each product using sterile forceps and transferred to sterile 1.5 ml microcentrifuge tubes for DNA extraction. The remaining raw and canned fish portions were stored at -80°C, while the remaining dried fish portions were stored at ambient room temperature (~20°C).

2.2 DNA extraction

The DNeasy Blood and Tissue Kit, Spin-Column protocol (Qiagen, Valencia, CA, USA) was used to extract DNA from all fish samples, according to the manufacturer's instructions. A negative extraction control was included with each batch of samples. Lysis was carried out in a ThermoMixer C (Eppendorf, Hamburg, Germany) set at 56 °C and 300 rpm for 4 h. DNA was eluted in 50 µl of Buffer AE and stored at -20 °C until PCR amplification.

2.3 PCR and sequencing

PCR and DNA sequencing of the CR mini-barcode was carried out as described by Mitchell and Hellberg (2016), with the exception that the primers did not include M13 tails. The

reaction mixture for raw samples included one half of an OmniMix HS PCR bead (Cepheid, Sunnyvale, CA, USA), 0.5 µl of 10-µM forward primer (Table 1), 0.5 µl of 10-µM reverse primer cocktail (Table 1), 3 µl DNA template or negative control, and 21.0 µl molecular grade water for a total volume of 25 µl. The reaction volumes were doubled for amplification of canned and dried samples to allow for sufficient PCR product for gel electrophoresis. Cycling conditions were as follows: 94 °C for 2 min; 35 cycles of 94 °C for 30 s, 49 °C for 40 s, and 72 °C for 1 min; and a final extension of 72 °C for 10 min. Secondary testing with ITS1 was carried out for one sample (R060) that showed a top genetic match to multiple species. DNA was re-extracted from the sample using the procedure described above, with the exception that the lysis step was 12 h and the elution volume was 100 µl. The ITS1 reaction mixture included 12.5 µl of HotStarTaq Plus (2X) (Qiagen), 1 µl of each 10 µM primer (Table 1; Chow et al., 2006; Viñas & Tudela, 2009), 6 µl of DNA template or negative control, and 4.5 µl of molecular grade water for a total reaction volume of 25 µl. Cycling conditions were as described in Viñas and Tudela (2009): 94 °C for 5 min; 35 cycles of touchdown PCR with denaturing at 95 °C for 1 min, initial annealing step of 10 cycles at 65 °C for 1 min with a decrease of 1 °C/cycle followed by 25 cycles of 55 °C for 1 min, and extension at 72 °C for 1 min; and a final extension at 72 °C for 10 min.

A Mastercycler nexus gradient thermal cycler (Eppendorf) was used for thermocycling and a non-template PCR control was included alongside each batch of samples. The primers were synthesized by Integrated DNA Technologies (Coralville, IA, USA). Precast 2.0% E-Gels with ethidium bromide (Invitrogen, Carlsbad, CA, USA) were used to confirm DNA amplification. The gels were loaded with 10 µl of PCR product from raw samples or 20 µl of PCR product from dried or canned samples and run for 30 min with an E-Gel Powerbase

(Invitrogen). A ChemiDoc Imager (Bio-Rad, Hercules, CA, USA) was used to visualize and photograph the results of gel electrophoresis. Any samples that failed the initial round of PCR amplification underwent a repeat DNA extraction with an extended 12 h lysing period and were re-amplified with PCR. Cleanup of PCR products was carried out with ExoSAP-IT (Applied Biosystems, Santa Clara, CA, USA) following the manufacturer's instructions. Bi-directional sequencing of all positive samples was carried out at Eurofins Genomics (Louisville, KY, USA) using a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and a 3730xl DNA Analyzer (Applied Biosystems).

2.4 Species identification

Raw sequence data was assembled and edited using Geneious R7 (Biomatters, Ltd., Auckland, NZ; http://www.geneious.com; Kearse et al., 2012). Quality control (QC) parameters were determined based on Pollack et al. (2018). Samples were considered successfully sequenced if they generated bi-directional sequences that were ≥76% of the target length (236 bp) and had <2% ambiguities or single reads that were ≥76% of the target length and had ≥98% high quality bases (HQ). The resulting consensus sequences were searched against GenBank using the nucleotide Basic Local Alignment Search Tool (BLAST), megablast algorithm, and the top species match was recorded. In order to verify that the species identification was not linked to a single erroneous sequence in GenBank, the top ten sequences in the search results were examined to ensure that multiple sequences were associated with the top species match. Samples that were identified as mislabeled were subjected to another round of DNA extraction, PCR, and DNA sequencing. Samples that showed less than 90% identity to the top species match in GenBank underwent repeat DNA extraction with a 12 h lysing period, PCR amplification, and DNA sequencing.

Samples that showed multiple equivalent top matches with the CR mini-barcode underwent phylogenetic analysis in MEGA11 (Tamura et al., 2021). The sequences were aligned using ClustalW with a subset of sequences downloaded from GenBank (accession numbers can be found in Fig. 1). The reference sequences were selected based on their use in previous studies (Cawthorn et al. 2011; Mitchell & Hellberg, 2016; Viñas & Tudela, 2009) and were quality checked for ambiguities (0%) and sequence length (234-236 bp). Genetic divergence was calculated using the Kimura 2-parameter method (Kimura, 1980) and a neighbor-joining (NJ) tree was compiled (Saitou & Nei, 1987). The robustness of the tree was evaluated using bootstrap analysis with 1000 iterations.

3. Results and Discussion

3.1 PCR amplification and DNA sequencing of the CR mini-barcode

Initially, 59 of the 80 samples exhibited PCR amplification success based on gel electrophoresis results. The samples that failed amplification consisted of 17 canned products and 4 dried products. However, following re-extraction with an extended 12 h lysing period and PCR amplification, all 21 re-extracted products showed bands in gel electrophoresis. Based on these results, an extended lysing period is recommended for improved PCR amplification of processed (dried and canned) tuna products. Of the 80 samples successfully amplified, 69 samples generated sequences (Table S1) that passed the QC parameters established by Pollack et al. (2018). The successful sequences had an average sequence length of 235 ± 3 bp, average HQ% of 92.68 ± 17.66%, and 0.30 ± 0.00% ambiguities. According to the QC parameters, raw products yielded the highest quality sequences with an average HQ% of 98.22 ± 2.28%, followed by dried (95.52 ± 6.42%) and canned (65.14 ± 35.44%) products.

All 69 successfully sequenced samples were identified to the species level (Table 2), with the top sequence matches for 67 of these samples showing >90% identity and ≥92% query coverage with BLAST. Due to its greater rate of divergence, the CR typically shows lower levels of intraspecies similarity as compared to other mitochondrial markers, such as COI and cytochrome *b*. For example, Viñas and Tudela (2009) reported average intraspecies divergence levels of 3.5-5.4% (not including introgressed sequences) for CR, as compared to 0.1-0.7% for COI. The two samples in this study with <90% identity to the top species match were canned products with 83.5% and 88.4% identity to yellowfin tuna (R009) and Pacific bluefin tuna (R010), respectively. R009 was labeled as containing yellowfin tuna and R010 was labeled as light tuna. Although these two sequences passed the quality control parameters, they had relatively low quality scores: R009 was 215 bp (91.1% of target length) and had 1.4% ambiguities and 7.9% HQ, while R010 was 224 bp (94.9% of the target length) and had 1.8% ambiguities and 2.2% HQ. As shown in Figure 1, R009 showed the closest phylogenetic relationship to yellowfin tuna sequences (20-22 % divergence), followed by longtail tuna (*Thunnus tonggol*; 24% divergence) and blackfin tuna (*Thunnus atlanticus*; 26% divergence). R010 showed the closest phylogenetic relationship to Pacific bluefin tuna and Pacific-like Atlantic tuna (14.5-15.7% divergence). The relatively weak genetic matches associated with these sequences make it difficult to confidently assign a species identification and these results should be interpreted with caution. For future studies, it is recommended that at least 90% genetic identity be required for a species to be assigned. The repeat analysis for these two samples resulted in sequencing failure for sample R010 and a positive identification for R009 to bigeye tuna, with 100% identity and 100% query coverage.

Within the product categories, all 40 raw samples from grocery stores and sushi restaurants were identified to the species level, 19 out of the 20 dried samples were identified to the species level, and 10 out of 20 canned samples were identified to the species level (Table 2). Species identification rates for raw (100%) and dried (95%) samples were comparable to previous CR mini-barcoding studies, which reported identification rates of 90% in raw tuna fillets and 100% in dried tuna roe (Frigerio et al., 2021; Liou et al., 2020). The species identification rate of 50% for canned tuna samples was slightly higher than the rate of 45% for canned tuna reported by Mitchell and Hellberg (2016). While the canned tuna identification rate was relatively low compared to raw and dried samples, it is a significant improvement compared to previous studies that were unable to amplify DNA extracted from canned tuna samples and/or only identified tuna samples to the genus level using COI mini-barcoding (Armani et al., 2017; Chin et al., 2016; Mitchell & Hellberg, 2016; Pollack et al., 2018; Shokralla et al., 2015). Compared to Mitchell and Hellberg (2016), the current study also showed a higher PCR amplification success rate (100% vs. 49%) for canned tuna samples. The improved PCR amplification success in the current study was attributed to the increased (12 h) lysing period used for DNA extraction for repeat samples, as compared to a 1-3 h lysing period in Mitchell & Hellberg (2016). The difference between the amplification success (100%) and sequencing success (50%) for canned samples in the current study may have been due to the presence of multiple species, which can co-amplify and generate a mixed chromatogram (Shokralla et al., 2015). Indeed, many of the samples that failed the QC sequencing parameters showed multiple, overlapping peaks in the sequencing chromatogram, which may indicate the presence of multiple species.

All samples showed a top genetic match to a single species in GenBank with the exception of R060, which showed equivalent matches to both albacore tuna and bigeye tuna. This sample was a raw sample labeled as albacore tuna. It yielded a high-quality sequence with 100% HQ that matched 2 albacore tuna sequences in GenBank with the same % identity (97.03%) and query coverage (100%) as a singular bigeye tuna sequence (Accession ID KM055385). The secondary genetic matches in GenBank were all to albacore tuna sequences (n = 97). While introgression between albacore and bigeye tuna is a possibility, it has not been reported in previous research. For further clarification, this sample underwent ITS1 sequencing as described in Viñas and Tudela (2009). The results of ITS1 sequencing were similar to those of CR, in which the R060 sequence matched 5 albacore tuna sequences in GenBank with the same % identity (96.52%) and query coverage (100%) but also matched a singular bigeye tuna sequence (Accession ID KM055385). Upon further investigation, it was found that both the CR and ITS1 bigeye sequence entries in GenBank were from the same unpublished research and were linked to the same isolate (isolate (CD-zj-dm). Because ITS1 sequencing would have been expected to differentiate introgressed species, it is likely that this isolate was mistakenly identified and uploaded to GenBank. Previous studies have also noted the possibility of researchers uploading inaccurate reference sequences onto genetic databases (Hanner et al., 2011; Hellberg et al., 2019; Lowenstein et al., 2009; Mitchell & Hellberg, 2016). Moreover, R060 grouped more closely with the albacore tuna clade as compared to the bigeye tuna clade when its CR mini-barcode was further analyzed against reference tuna sequences (Fig. 1). Therefore, considering the combination of BLAST results and phylogenetic analysis, R060 was determined to be albacore tuna. In order to avoid future encounters with erroneous GenBank

sequences, it is recommended that a reference database be created for species identification using CR sequences from authenticated specimens.

3.2 Mislabeled samples

Out of the 69 samples that were identified, 11 were determined to be mislabeled (Table 3). To avoid the possibility of a single erroneous sequence in GenBank leading to a sample being identified as mislabeled, the top ten sequence matches in the search results were examined. For all 11 samples in Table 3, the top ten sequence matches all belonged to the same species, supporting the original identification. Furthermore, all mislabeled samples underwent DNA extraction, PCR, and sequencing a second time for confirmation. Upon repeat testing, mislabeling was confirmed for 7 of the samples, while one canned sample (R015) failed the second round of PCR amplification and three samples (R018, R044, and R078) yielded different species matches compared to the initial sequencing results (discussed below). The majority (73%) of mislabeled products were advertised as "yellowfin tuna," followed by "skipjack tuna" (18%) and "bluefin tuna" (9%). When comparing across product categories, the highest rate of mislabeling was identified in dried samples (26%), followed by canned samples (20%) and raw samples (5%). Species mixtures were observed in canned and dried samples (Table 3), indicating that future studies should collect multiple subsamples from these types of processed samples and/or that metabarcoding should be conducted.

Five samples labeled as yellowfin tuna were found to be partially or completely substituted with bigeye tuna (Table 3). This type of species substitution was observed across multiple product types, specifically two canned products (R009 and R015), one sushi product (R067), and two dried jerky products (R078, R080). Sample R078 was labeled as only containing yellowfin tuna but was found to include both yellowfin and bigeye tuna after repeat

sequencing. Yellowfin tuna is considered to be lower in mercury and safe to eat once a week for at-risk consumers, but bigeye and Pacific bluefin tuna both have elevated mercury levels and can pose a health risk to vulnerable individuals (FDA/EPA, 2019; Lowenstein et al., 2010). Bigeye and yellowfin tuna inhabit the same tropical waters, look similar to one another, and are difficult to accurately distinguish morphologically without trained personnel (NOAA, 2020). However, the ex-vessel price of bigeye tuna was $6.12/kg in 2019, making it more economically valuable than yellowfin tuna ($1.76/kg) (NMFS, 2021). Therefore, it is possible that bigeye tuna were captured alongside yellowfin tuna and processed into canned and dried products labeled as yellowfin tuna.

Two samples labeled as skipjack tuna (R018 and R044) were found to contain skipjack tuna plus additional tuna species, specifically bigeye tuna (R018, canned) and yellowfin tuna (R044, dried flakes). These mislabeling events introduce a health concern because skipjack tuna is part of the "Best Choices" category for consumers at risk of mercury exposure, meaning that it can be consumed multiple times a week. On the other hand, yellowfin tuna is in the "Good Choices" category for seafood that may be eaten once a week, while consumption of bigeye tuna should be completely avoided by at-risk consumers (FDA/EPA, 2019). With an ex-vessel price of $1.21/kg in 2019, skipjack tuna is a cheaper species compared to yellowfin ($1.76/kg) and bigeye ($6.12/kg) tuna (NMFS, 2021), suggesting that this mislabeling event was not associated with an economic incentive. Skipjack tuna is morphologically distinct from yellowfin and bigeye tuna, but all three species inhabit the same waters and bycatch may occur (NOAA, 2020). Therefore, it is possible that the undeclared species were present due to accidental inclusion.

One sample (R035) was identified to be yellowfin tuna even though it was a raw fillet labeled as "B/F Tuna" (considered to be an abbreviated description for "bluefin tuna"). The

296 sample was priced at $55.10/kg, which is similar to the price of $61.72/kg observed for another
297 raw sample explicitly labeled “bluefin tuna chutoro” and identified as Southern bluefin tuna from
298 the same grocery store (R034, Fig. 2). “Chutoro” refers to a valuable cut of tuna consisting of the
299 upper belly area of the fish and often garners a higher price (Shimose et al., 2018). Interestingly,
300 sample R037 was also from the same grocery store and was labeled as “Tuna Steak” with a price
301 of $28.64/kg. R037 was identified as yellowfin tuna and was almost half the price of the other
302 samples labeled as bluefin tuna, which demonstrates the economic incentive associated with
303 mislabeling yellowfin as bluefin (Hu et al., 2018; Pardo et al., 2018). These factors suggest that
304 the mislabeled sample R035 was indeed intended to be sold as bluefin tuna at a premium price
305 instead of the lower price yielded for yellowfin tuna.
306 Samples R074-R076 were packaged as single strips of dried “yellowfin” tuna and were
307 identified as Pacific bluefin tuna. These three samples were from the same company but
308 contained different flavors, indicating that this mislabeling issue impacted several product lines.
309 The ex-vessel price for Pacific bluefin tuna ($8.36/kg) is significantly higher than that of
310 yellowfin tuna so there does not appear to be an economic incentive for the company to
311 intentionally mislabel their product (NMFS, 2021). However, mislabeling may be a method to
312 conceal instances of IUU and to provide a legal market for these catches (Hu et al., 2018; Warner
313 et al., 2013). Pacific bluefin tuna continue to be an overfished, threatened species and their
314 inclusion in mislabeled products deters conservation efforts to preserve the species. Similar to
315 the results of the current study, substitution of yellowfin tuna with bigeye or bluefin tuna has
316 been reported in numerous studies (Gunther et al., 2017; Hu et al., 2018; Liou et al., 2020; Pardo
317 et al., 2018; Warner et al., 2013).

It is important to note that the CR mini-barcode cannot be used to differentiate among introgressed tuna species. Low levels of introgression (2-3%) have been reported between Pacific and Atlantic bluefin tuna, as well as for albacore DNA introgressed into Pacific or Atlantic bluefin tuna (Alvarado Bremer et al., 1999; Alvarado Bremer et al., 2005; Bayona-Vasquez et al., 2018; Chow & Kishino, 1995; Viñas & Tudela, 2009). Therefore, it is recommended that cases of mislabeling that could have been misidentified due to introgression of tuna species be further analyzed with a secondary marker, such as ITS1 (Mitchell & Hellberg, 2016). In the case of the CR mini-barcode, this refers to instances of a sample labeled as Atlantic bluefin but identified as Pacific bluefin (or vice versa) or a sample labeled as Atlantic or Pacific bluefin but identified as albacore tuna. Because none of these instances was observed in the current study, additional testing with a secondary marker was not deemed necessary. An alternative approach used when testing fresh/unfrozen tuna with the 450-bp CR has been to conduct secondary testing with ITS1 on any samples associated with possible unresolved introgression (i.e., samples identified as albacore), regardless of mislabeling status (Gordoa et al., 2016). While this is a more thorough approach, it is also more time-consuming and costly compared to targeted testing of mislabeled samples. Additionally, there is a lack of well-developed secondary genetic markers for use in resolving introgression in heavily processed tuna samples. While Mitchell & Hellberg (2016) developed a short ITS1 marker specifically for the differentiation of albacore and introgressed Atlantic bluefin tuna, this marker does not allow for differentiation of introgressed Pacific and Atlantic bluefin tuna. Therefore, future studies should consider development of a secondary marker for resolving introgression of the bluefin tuna species in heavily processed products.

4. Conclusions

Accurate methods for the reliable detection of tuna species in food products facilitate the long-term goals of detecting seafood fraud, providing transparency in the marketplace, advancing tuna conservation efforts, and protecting consumer health. This study investigated the efficacy of a CR mini-barcoding method to identify a variety of raw and processed tuna products. Species identification was achieved for the majority of products tested, with high success rates in raw and dried products. Limited success was observed for canned products, likely due to DNA degradation and the presence of multiple species within some products. Metabarcoding should be considered in future studies to overcome the challenges of identifying multiple unknown species in a single product. Overall, 16% of identified samples were determined to be mislabeled, with mislabeling occurring most frequently in dried products, followed by canned and raw products. Most of the mislabeled products were marketed as yellowfin tuna, followed by skipjack and bluefin tuna. These products were likely mislabeled for a variety of reasons, including accidental inclusion due to by-catch, provision of a legal market for IUU catches, and substitution of more expensive species with a cheaper species for economic gain. Overall, this research shows the utility of the CR mini-barcoding method for the detection of species in raw and processed tuna products. However, it is important to consider the need for secondary markers to resolve instances of possible introgression among tuna species. Because GenBank is a public database that may contain erroneous sequence information, a reference database should be created for species identification using CR sequences from authenticated specimens. Future research should examine optimization of the method to further improve identification of species in canned tuna samples.

Acknowledgements

This work was supported by internal funding from Chapman University. The authors would like to thank Ashley Whelpley for laboratory support.

Supplementary Material. Table S1: Detailed product information, consensus sequences, and species determinations for the 80 samples collected in this study.

References

Alvarado Bremer, J.R., Naseri, I., Ely, B. (1999). Heterogeneity of northern bluefin tuna populations. *ICCAT Coll. Vol. Sci. Pap. 49,* 127–129.

Alvarado Bremer, J.R., Viñas, J., Mejuto, J., Ely, B., Pla, C. (2005) Comparative phylogeography of Atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. *Mol Phylogenet Evol, 36,* 169–187.

Armani, A., Tinacci, L., Lorenzetti, R., Benvenuti, A., Susini, F., Gasperetti, L., Ricci, E., Guarducci, M., & Guidi, A. (2017). Is raw better? A multiple DNA barcoding approach (full and mini) based on mitochondrial and nuclear markers reveals low rates of misdescription in sushi products sold on the Italian market. *Food Control*, *79*, 126–133. https://doi.org/10.1016/j.foodcont.2017.03.030.

Bayona-Vásquez, N.J., Glenn, T.C., Uribe-Alcocer, M., Pecoraro, C., & Díaz-Jaimes, P. (2018). Complete mitochondrial genome of the yellowfin tuna (*Thunnus albacares*) and the blackfin tuna (*Thunnus atlanticus*): notes on mtDNA introgression and paraphyly on tunas. *Conservation Genet Resour, 10*, 697–699. https://doi.org/10.1007/s12686-017-0904-0

384 Cawthorn, D.-M., Steinman, H.A., Witthuhn, R.C. (2011). Establishment of a mitochondrial
385 DNA sequence database for the identification of fish species commercially available in
386 South Africa. Mol Ecol Resour, *11*, 979–991.
387 Chin, T. C., Adibah, A. B., Hariz, Z. A. D., & Azizah, M. N. S. (2016). Detection of mislabelled
388 seafood products in Malaysia by DNA barcoding: Improving transparency in food
389 market. *Food Control*, *64*, 247–256. https://doi.org/10.1016/j.foodcont.2015.11.042.
390 Chow, S., Kishino, H. (1995). Phylogenetic relationships between tuna species of the genus
391 *Thunnus* (Scombridae: Teleostei): inconsistent implications from morphology, nuclear
392 and mitochondrial genomes. *J Mol Evol, 41,* 741–748.
393 Chow, S., Nakagawa, T., Suzuki, N., Takeyama, H., Matsunaga, T. (2006). Phylogenetic
394 relationships among *Thunnus* species inferred from rDNA ITS1 sequence. *J Fish Biol,*
395 *68,* 24–35.
396 FAO. (2018). *Overview of Food Fraud in the Fisheries Sector*. *Fisheries and Aquaculture*
397 *Circular No. 1165*, 1–21.
398 FAO. (2020). *The State of World Fisheries and Aquaculture 2020. Sustainability in Action.*
399 Rome. https://doi.org/10.4060/ca9229en.
400 FDA. (2020). *Fish and fishery products: Hazards and controls guidance*.
401 https://www.fda.gov/food/seafood-guidance-documents-regulatory-information/fish-and-
402 fishery-products-hazards-and-controls. Accessed July 3, 2020.
403 FDA/EPA. (2019). *Advice about eating fish: For women who are or might become pregnant,*
404 *breastfeeding mothers, and young children*. https://www.fda.gov/food/consumers/advice-
405 about-eating-fish. Accessed July 3, 2020.

406 Frigerio, J., Marchesi, C., Magoni, C., Saliu, F., Ballabio, D., Consonni, V., Gorini, T., De
407 Mattia, F., Galli, P., & Labra, M. (2021). Application of DNA mini-barcoding and
408 infrared spectroscopy for the authentication of the Italian product "bottarga." *LWT*, *139*,
409 110603. https://doi.org/10.1016/j.lwt.2020.110603.
410 Gordoa, A., Carreras, G., Sanz, N., Viñas, J. (2017). Tuna species substitution in the Spanish
411 commercial chain: a knock-on effect. *PLoS ONE, 12(1),* e0170809.
412 https://doi.org/10.1371/journal.pone.0170809
413 Gunther, B., Raupach, M. J., & Knebelsberger, T. (2017). Full-length and mini-length DNA
414 barcoding for the identification of seafood commercially traded in Germany. *Food*
415 *Control*, *73*, 922–929. https://doi.org/10.1016/j.foodcont.2016.10.016.
416 Handy, S. M., Deeds, J. R., Ivanova, N. V., Hebert, P. D. N., Hanner, R. H., Ormos, A., Weigt,
417 L. A., Moore, M. M., & Yancy, H. F. (2011). A single-laboratory validated method for
418 the generation of DNA barcodes for the identification of fish for regulatory compliance.
419 *Journal of Aoac International*, *94*(1), 201–210.
420 Hanner, R., Becker, S., Ivanova, N. V., & Steinke, D. (2011). FISH-BOL and seafood
421 identification: Geographically dispersed case studies reveal systemic market substitution
422 across Canada. *Mitochondrial DNA*, *22*, 106–122.
423 https://doi.org/10.3109/19401736.2011.588217.
424 Hebert, P. D. N., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications
425 through DNA barcodes. *Proceedings of the Royal Society B-Biological Sciences*,
426 *270*(1512), 313–321. https://doi.org/10.1098/rspb.2002.2218.

427 Hellberg, R. S., Isaacs, R. B., & Hernandez, E. L. (2019). Identification of shark species in
428 commercial products using DNA barcoding. *Fisheries Research*, *210*, 81–88.
429 https://doi.org/10.1016/j.fishres.2018.10.010.
430 Hu, Y., Huang, S. Y., Hanner, R., Levin, J., & Lu, X. (2018). Study of fish products in Metro
431 Vancouver using DNA barcoding methods reveals fraudulent labeling. *Food Control*, *94*,
432 38–47. https://doi.org/10.1016/j.foodcont.2018.06.023.
433 Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S.,
434 Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., &
435 Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software
436 platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12),
437 1647–1649. https://doi.org/10.1093/bioinformatics/bts199.
438 Liou, P., Banda, A., Isaacs, R. B., & Hellberg, R. S. (2020). Labeling compliance and species
439 authentication of fish fillets sold at grocery stores in Southern California. *Food Control*,
440 *112*, 107137. https://doi.org/10.1016/j.foodcont.2020.107137.
441 Lowenstein, J. H., Amato, G., & Kolokotronis, S. O. (2009). The real maccoyii: Identifying tuna
442 sushi with DNA barcodes—Contrasting characteristic attributes and genetic distances.
443 *Plos One*, *4*(11). https://doi.org/10.1371/journal.pone.0007866.
444 Lowenstein, J. H., Burger, J., Jeitner, C. W., Amato, G., Kolokotronis, S. O., & Gochfeld, M.
445 (2010). DNA barcodes reveal species-specific mercury levels in tuna sushi that pose a
446 health risk to consumers. *Biology Letters*, *6*(5), 692–695.
447 https://doi.org/10.1098/rsbl.2010.0156.

448 Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions
449 through comparative studies of nucleotide sequences. *Journal of Molecular*
450 *Evolution,* 16, 111-120.
451 Mitchell, J. K., & Hellberg, R. S. (2016). Use of the mitochondrial control region as a potential
452 DNA mini-barcoding target for the identification of canned tuna species. *Food Analytical*
453 *Methods*, *9*(10), 2711–2720. https://doi.org/10.1007/s12161-016-0460-3.
454 NFI. (2021). *Top 10 List for Seafood Consumption*. National Fisheries Institute.
455 https://aboutseafood.com/about/top-ten-list-for-seafood-consumption/. Accessed January
456 6, 2020.
457 NMFS. (2021). *Fisheries of the United States, 2019*.
458 https://www.fisheries.noaa.gov/resource/document/fisheries-united-states-2019-report.
459 NOAA. (2020). *Species Directory*. NOAA Fisheries. https://www.fisheries.noaa.gov/species-
460 directory. Accessed March 24, 2021.
461 Pardo, M. A., Jimenez, E., & Perez-Villarreal, B. (2016). Misdescription incidents in seafood
462 sector. *Food Control*, *62*, 277–283. https://doi.org/10.1016/j.foodcont.2015.10.048.
463 Pardo, M. A., Jimenez, E., Vidarsson, J. R., Olafsson, K., Olafsdottir, G., Danielsdottir, A. K., &
464 Perez-Villareal, B. (2018). DNA barcoding revealing mislabeling of seafood in European
465 mass caterings. *Food Control*, *92*, 7–16. https://doi.org/10.1016/j.foodcont.2018.04.044.
466 Pollack, S. J., Kawalek, M. D., Williams-Hill, D. M., & Hellberg, R. S. (2018). Evaluation of
467 DNA barcoding methodologies for the identification of fish species in cooked products.
468 *Food Control*, *84*, 297–304. https://doi.org/10.1016/j.foodcont.2017.08.013.
469 Saitou, N., & Nei, M. (1987). The neighbor-joining method: A new method for reconstructing
470 phylogenetic trees. *Molecular Biology and Evolution*, *4*(4), 406–425.

Shimose, T., Kanaiwa, M., & Tanabe, T. (2018). Influence of the fat content on the auction price of Pacific bluefin tuna Thunnus orientalis at Japanese fish markets. *Fisheries Research*, *204*, 224–230. https://doi.org/10.1016/j.fishres.2018.03.001.

Shokralla, S., Hellberg, R. S., Handy, S. M., King, I., & Hajibabaei, M. (2015). A DNA mini-barcoding system for authentication of processed fish products. *Scientific Reports*, *5*. https://doi.org/10.1038/srep15894.

Silva, A. J., Hellberg, R. S., & Hanner, R. H. (2021). Seafood fraud. In R. S. Hellberg, K. Everstine, & S. A. Sklare (Eds.), *Food Fraud: A Global Threat with Public Health and Economic Consequences* (pp. 109–137). Academic Press/Elsevier.

Silva, A. J., & Hellberg, R. S. (2021). DNA-based techniques for seafood species authentication. In F. Toldra (Ed.), *Advances in Food and Nutrition Research* (Vol. 95, pp. 207–255). Academic Press/Elsevier. https://doi.org/10.1016/bs.afnr.2020.09.001.

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA 11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution, 38(7),* 3022-3027, https://doi.org/10.1093/molbev/msab120

Viñas, J., & Tudela, S. (2009). A validated methodology for genetic identification of tuna species (Genus Thunnus). *Plos One*, *4*(10). https://doi.org/10.1371/journal.pone.0007606.

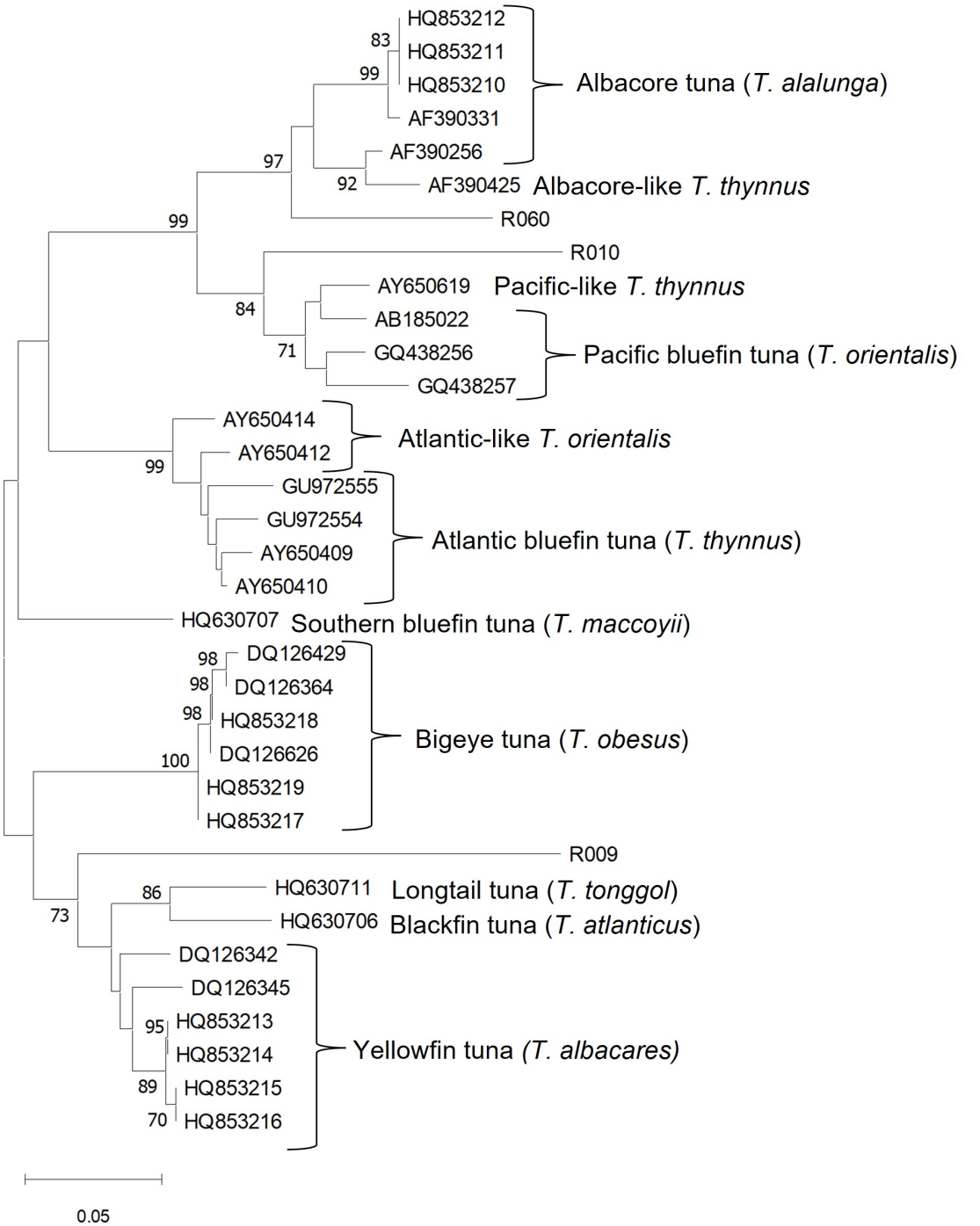
Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *360*(1462), 1847–1857. https://doi.org/10.1098/rstb.2005.1716.

Warner, K., Timme, W., Lowell, B., & Hirshfield, M. (2013). *Oceana Study Reveals Seafood Fraud Nationwide*. Oceana.

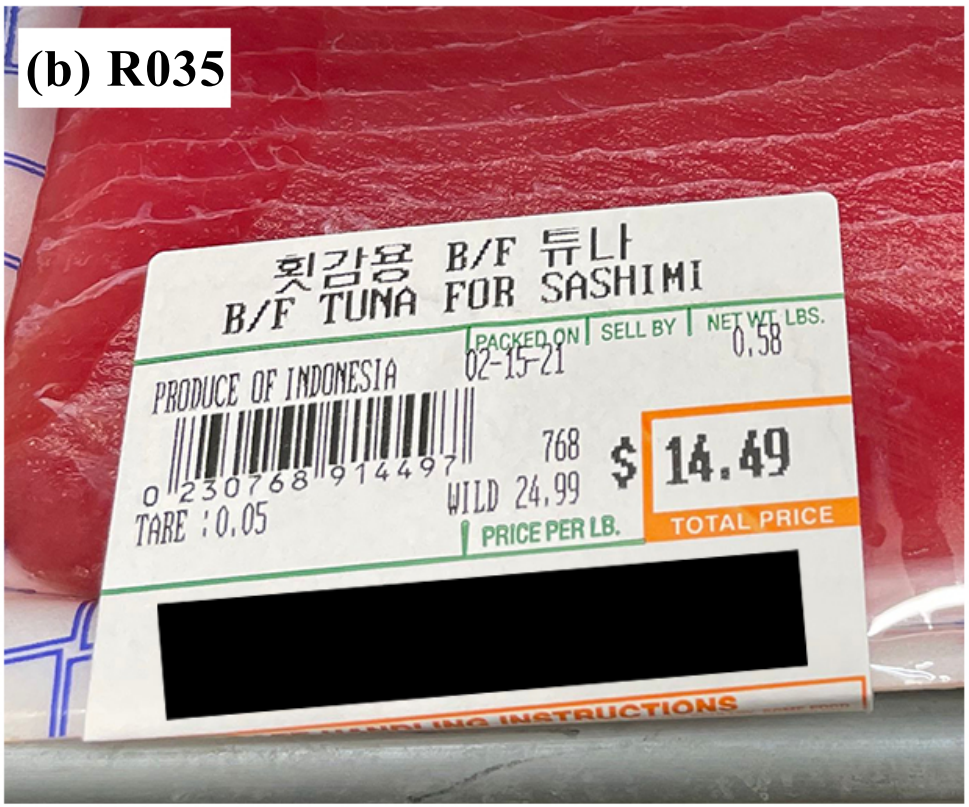
Figure captions

Figure 1. Neighbor-joining tree showing samples R009, R010, and R060 and reference sequences for each *Thunnus* species targeted by the CR mini-barcode. GenBank accession numbers are shown for all reference sequences. The Kimura 2-parameter method was used to calculate genetic distances and bootstrap analysis was conducted with 1000 replicates. The numbers at the nodes represent bootstrap values greater than 70%.

Figure 2. Product labels for raw samples (a) R034, (b) R035, and (c) R037 which were all purchased from the same grocery store. Note: R034 was labeled as "bluefin tuna chutoro" on the reverse side of the package and was identified as Southern bluefin tuna. R035 and R037 were identified as yellowfin tuna.


HQ853212
HQ853211
HQ853210
AF390331
AF390256
AF390425
R060
R010
AY650619
AB185022
GQ438256
GQ438257
AY650414
AY650412
GU972555
GU972554
AY650409
AY650410
HQ630707
DQ126429
DQ126364
HQ853218
DQ126626
HQ853219
HQ853217
R009
HQ630711
HQ630706
DQ126342
DQ126345
HQ853213
HQ853214
HQ853215
HQ853216
83
99
97
92
99
84
71
99
98
98
98
100
86
73
95
89
70
Albacore tuna (T. alalunga)
Albacore-like T. thynnus
Pacific-like T. thynnus
Pacific bluefin tuna (T. orientalis)
Atlantic-like T. orientalis
Atlantic bluefin tuna (T. thynnus)
Southern bluefin tuna (T. maccoyii)
Bigeye tuna (T. obesus)
Longtail tuna (T. tonggol)
Blackfin tuna (T. atlanticus)
Yellowfin tuna (T. albacares)
0.05


(a) R034
튜나뱃살/주토로
TUNA BELLY (ZU-TORO)
PACKED ON
SELL BY
NET WT. LBS.
PRODUCE OF INDONESIA
02-02-21
0.54
0 230101 715118
101
WILD 27.99
$ 15.11
TARE : 0.05
PRICE PER LB.
TOTAL PRICE
SAFE HANDLING INSTRUCTIONS
THIS PRODUCT WAS PREPARED FROM INSPECTED AND PASSED MEAT AND/OR POULTRY. SOME FOOD PRODUCTS MAY CONTAIN BACTERIA THAT COULD CAUSE ILLNESS IF THE PRODUCT IS MISHANDLED OR COOKED IMPROPERLY. FOR YOUR PROTECTION, FOLLOW THESE SAFE HANDLING INSTRUCTIONS.
KEEP REFRIGERATED OR FROZEN. THAW IN REFRIGERATOR OR MICROWAVE.
KEEP RAW MEAT AND POULTRY SEPARATE FROM OTHER FOODS. WASH WORKING SURFACES (INCLUDING CUTTING BOARDS), UTENSILS, AND HANDS AFTER TOUCHING RAW MEAT OR POULTRY.
COOK THOROUGHLY.
KEEP HOT FOODS HOT. REFRIGERATE LEFTOVERS IMMEDIATELY OR DISCARD.


(b) R035
횟감용 B/F 튜나
B/F TUNA FOR SASHIMI
PACKED ON
SELL BY
NET WT. LBS.
PRODUCE OF INDONESIA
02-15-21
0.58
0 230768 914497
768
WILD 24.99
$ 14.49
TARE : 0.05
PRICE PER LB.
TOTAL PRICE


(c) R037
튜나스테이크
TUNA STEAK
PACKED ON
SELL BY
NET WT. LBS.
PRODUCE OF PHILIPPINES
02-15-21
0.83
0 230647 710783
647
WILD 12.99
$ 10.78
TARE : 0.02
PRICE PER LB.
TOTAL PRICE

Table 1. Primers used in this study. The CR reverse primers were combined into a primer cocktail using a ratio of 1:1.

Primer set	Primer name	Primer direction	Primer sequence (5'-3')	Target fragment length	Reference
CR mini-barcode	Tuna CR_F	Forward	GCAYGTACATATATGTAAYTACACC	236 bp	Mitchell & Hellberg (2016)
	Tuna CR_R1	Reverse	CTGGTTGGTRGKCTCTTACTRCA		
	Tuna CR_R2	Reverse	CTGGATGGTAGGYTCTTACTGCG		
ITS1	ITS1_Full_F	Forward	TCCGTAGGTGAAACCTGCGG	594-656 bp	Chow et al. (2006); Viñas & Tudela (2009)
	ITS1_Full_R	Reverse	CGCTGCGTTCTTCATCG		

Table 2. Summary of the species identification results for the 80 tuna products analyzed in this study with the CR mini-barcode.

Category	Labeled species	Sample size	Identified to species level	Mislabeled samples	% of samples identified in each category
Raw, grocery store	Albacore	3	3	0	100
	Yellowfin	8	8	0	
	Bluefin	5	5	1	
	Tuna	4	4	0	
Raw, sushi restaurant	Albacore	8	8	0	100
	Yellowfin	1	1	1	
	Bluefin	1	1	0	
	Tuna	10	10	0	
Dried	Yellowfin	12	12	5	95
	Skipjack	7	6	1	
	Tuna	1	1	0	
Canned	Albacore	5	3	0	50
	Yellowfin	4	3	2	
	Skipjack	6	1	1	
	Light tuna	5	3	0	
Combined	N/A	80	69	11	86

Table 3. Summary of the 11 mislabeled tuna products reported in this study. Top species matches were determined using BLAST.

Mislabeled sample ID	Category, product type	Expected species based on product label	Species determination	% Identity for top species match in GenBank	Accession ID and E-value for top species match in GenBank	Price paid (US $/kg)
R009	Canned, chunk style in oil	Yellowfin tuna (*T. albacares*)	Yellowfin tuna[b] (*T. albacares*); bigeye tuna (*T. obesus*)	83.49%; 100%	KC165917 (2e-46); LC497955 (2e-117)	23.27
R015	Canned, in oil with sun-dried tomato	Yellowfin tuna (*T. albacares*)	Bigeye tuna[c] (*T. obesus*)	99.57%	LC497998 (6e-116)	40.41
R018	Canned, chunk style in water	Skipjack tuna (*K. pelamis*)	Bigeye tuna[c] (*T. obesus*); skipjack tuna (*K. pelamis*)	98.7%; 90.78%	LC497998 (2e-111); JF752257 (9e-74)	14.01
R035	Raw (grocery store), block for sashimi	Bluefin tuna[a] (*T. maccoyii/ orientalis/thynnus*)	Yellowfin tuna (*T. albacares*)	99.57%	KJ535766 (2e-116)	55.10
R067	Raw (restaurant), sushi	Yellowfin tuna (*T. albacares*)	Bigeye tuna (*T. obesus*)	99.57%	LC497998 (6e-116)	211.64
R044	Dried, smoked flakes	Skipjack tuna (*K. pelamis*)	Yellowfin tuna[c] (*T. albacares*); skipjack tuna (*K. pelamis*)	99.15%; 98.73%	KJ535745 (7e-115); KP669063 (1e-113)	87.57
R074	Dried, jerky strip	Yellowfin tuna (*T. albacares*)	Pacific bluefin tuna[d] (*T. orientalis*)	98.73%	AB933628 (1e-112)	117.65
R075	Dried, jerky strip	Yellowfin tuna (*T. albacares*)	Pacific bluefin tuna[d] (*T. orientalis*)	98.72%	AB933628 (3e-113)	117.65
R076	Dried, jerky strip	Yellowfin tuna (*T. albacares*)	Pacific bluefin tuna[d] (*T. orientalis*)	98.72%	AB933628 (1e-112)	117.65
R078	Dried, jerky	Yellowfin tuna (*T. albacares*)	Bigeye tuna[c] (*T. obesus*); yellowfin tuna (*T. albacares*)	99.58%; 98.73%	LC498002 (6e-116); LC585308 (3e-113)	164.67
R080	Dried, jerky	Yellowfin tuna (*T. albacares*)	Bigeye tuna (*T. obesus*)	99.57%	LC498057 (3e-113)	164.67

[a] Labeled as "B/F Tuna" (considered to be an abbreviated description for bluefin tuna)
[b] Less than 90% genetic identity to top species match
[c] Repeat testing failed to confirm species
[d] The CR mini-barcode is unable to differentiate *T. orientalis* from introgressed Pacific-like *T. thynnus.* This form of introgression has been reported at levels of 2-3% (Viñas & Tudela).