Food Control 145 (2023) 109473

Contents lists available at ScienceDirect

Food Control

journal homepage: www.elsevier.com/locate/foodcont

DNA sequencing shows that tropical tuna species misidentification can be an underestimated issue in fish landings

Ana Rita Carreiro [a,b], Jaime A. Ramos [a], Vanessa Mata [b,d], Nathalie M. Almeida [a,c], Vítor H. Paiva [a], Ricardo Jorge Lopes [b,d,e,*]

[a] University of Coimbra, MARE - Marine and Environmental Sciences Centre / ARNET – Aquatic Research Network, Department of Life Sciences, Calçada Martim de Freitas, 3004-517, Coimbra, Portugal
[b] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, 4485-661, Vairão, Portugal
[c] Biosfera Cabo Verde, Rua de Moçambique 28, Mindelo, caixa postal 233, São Vicente, Republic of Cabo Verde
[d] BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661, Vairão, Portugal
[e] MHNC-UP, Natural History and Science Museum of the University of Porto, Praça Gomes Teixeira, 4099-002, Porto, Portugal

ARTICLE INFO

Keywords:
Tuna fisheries
Traceability
Genetics
Atlantic Ocean
Cabo Verde

ABSTRACT

Taxonomic identification of food is a key component to the correct implementation of traceability pipelines during fish processing. Genetic methods have increased exponentially our discrimination power at the species level. In this context, sustainable management of tuna fisheries can clearly profit from the implementation of this kind of quality control procedures, although they are more difficult to implement at the earlier stages of tuna processing, from harvest to the arrival of fish landings at a processing plant. Here we provide a case study using tuna landings to a processing plant in Cabo Verde, an important small island developing State, which supplies mainly the European Union region. Using DNA sequencing based species identification of tissue samples, we found a taxonomic miss-assignment of species identity in 33% of the individuals in the three target tuna species, Bigeye (*Thunus obesus*), Yellowfin (*Thunus albacares*), and Skipjack (*Katsuwonus pelamis*). More than half of Yellowfin samples were wrongly assigned as bigeyes or skipjacks. We also found that many bigeyes were wrongly assigned as yellowfins or skipjacks. None of the skipjacks were incorrectly assigned. The proportion of Yellowfin samples incorrectly assigned as bigeye was significantly higher than the samples of bigeye incorrectly assigned as yellowfins. These results clearly prove the need for improved traceability processes, which should encompass the whole processing pipeline, from harvesting to the consumer.

1. Introduction

Traceability of food is an important issue, especially when items are hunted or harvested from natural populations (Badia-Melis et al., 2015; Fanelli et al., 2021). Of special concern are fishery activities, due to the vast diversity of species harvested, the wide spatial range of fishing activities and the large variety of food products available to the consumer (Ruiz-Salmón et al., 2021). Traceability is also a key issue to ensure the sustainable management of fisheries, a goal that is highly acknowledged as crucial for the economic development of many countries, the increase of human populations welfare and the maintenance of biodiversity (Cochrane, 2021; Haas et al., 2021).

Traceability of taxonomic identification and geographical origin is key for fisheries' sustainability. Both can have a large impact on the power of fisheries stakeholders to monitor, identify and mitigate issues in a timely manner (Dimitrakopoulou & Vantarakis, 2021). Concerning taxonomic identification, the incorrect assignment of species identification can have significant impacts on the report of landings and the enforcement of stock quotas. Also, throughout the chain of fish processing and at the consumer level, this can either increase or decrease the value of the harvested fish (Barendse et al., 2019). Additionally, this may bias important information for the assessment of fish consumption impacts, such as the conservation status (e.g., many tuna species are threatened) or health and toxicological parameters (Kumar, 2018;

* Corresponding author. CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, 4485-661, Vairão, Portugal.
E-mail address: riclopes@me.com (R.J. Lopes).

https://doi.org/10.1016/j.foodcont.2022.109473
Received 4 July 2022; Received in revised form 12 October 2022; Accepted 21 October 2022
Available online 26 October 2022

Kusche & Hanel, 2021; Love et al., 2021; Rahmani et al., 2018; Xiong et al., 2019) or the evaluation of the sustainability of the whole industry (Guillotreau et al., 2017).

When traceability measures find taxonomic mislabelling or ambiguities, this can be either an unintentional or deliberate action. At capture, incorrect assignment can arise from the occurrence of many related species, or ambiguous regional common names or fisherman's slang that encompasses many species (Cawthorn & Mariani, 2017). On the other hand, intentional mislabelling can increase the profit margin by increasing the market value of those specimens or allow unsustainable or illegal fisheries to be illegible as certified species.

Methodologies for the traceability of fish products have relied mainly on the genetic assignment based on the retrieval of DNA (mostly mitochondrial) sections with enough discriminant power to distinguish between species (Abdullah & Rehbein, 2014; Amaral et al., 2017; Barendse et al., 2019; Blanco-Fernandez et al., 2021; da Silva et al., 2019; Delpiani et al., 2020; Espiñeira et al., 2009; Fernandes et al., 2021; Helgoe et al., 2020; Hellberg et al., 2019; Kappel & Schröder, 2020; Liu et al., 2021; Minoudi et al., 2020; Yang et al., 2021). This power has been boosted by the continuous increase of reference, validated DNA information from each species, available through genetic information repositories such as NCBI Genbank and BOLD databases (Kappel & Schröder, 2020). To this date, Sanger sequencing is still the most common way to retrieve genetic information (Abdullah & Rehbein, 2014; Chang, Kao, et al., 2021; Chang, Tsai, et al., 2021; Gordoa et al., 2017; Hellberg et al., 2019; Lockley & Bardsley, 2000; McCluney et al., 2019; Servusova & Piskata, 2021; Yao et al., 2020). Although high-throughput sequencing (HTS) technologies, using next generation sequencers (NGS), have been accessible for more than one decade, their use for traceability issues is still not commonly adopted, despite some advantages for this purpose. For example, it may allow for a better discrimination of multiple species in the same sample, be used to survey other aspects simultaneously, such as parasite diversity (Bernatchez et al., 2017; Díaz-Arce et al., 2016; Galimberti et al., 2019) or use environmental DNA approaches to survey meltwater from fishing vessels to obtain signatures of species (Willette et al., 2021).

In this context, tuna fisheries are particularly interesting, mainly because they are an important fraction of the global fishery activity (Brill & Hobday, 2017; McCluney et al., 2019), and traceability issues can have specific impacts. First, due to the size and large range of tuna stocks, they comprise many national waters from developed and developing countries as well as international waters, managed by several Regional Fisheries Management Organizations (Seto et al., 2021). Second, tuna products range from raw meat to canned products for human or pet consumption (Guillotreau et al., 2017), and in many of these products it is not possible to visually perform taxonomic identification to the species level (Chang, Tsai, et al., 2021). Third, there are several species that can be misidentified. These comprise true tunas, such as the Atlantic Bluefin (*Thunus thynnus*), Bigeye (*Thunus obesus*), Yellowfin (*Thunus albacares*), Albacore (*Thunus alalunga*) but also other related species, such as the Skipjack (*Katsuwonus pelamis*) and other smaller tunas (Collette & Graves, 2019). Fourth, tuna stocks are targeted from large industrial boats to artisanal fishing, with different traceability methodologies, and therefore many communities, especially in developing countries, dependent on the income of this activity may suffer from identification issues.

However, it is still observed a bias towards the authentication of related tuna species in canned and raw products at the distribution or at the consumer level. Also, taxonomic identification from the boat to the processing industrial plant is not often reported. This is not surprising, since it is much easier to perform this kind of procedure in the countries where the validation laboratories are, and the final products are much easier to find at the consumer level. But, this information is crucial to understand how low in the chain of processing mislabelling can happen. Here we address the probability of misidentification of tunas in fish landings, to understand the magnitude of this issue prior to their arrival at the fish processing plant. This was performed with a sample based on a large variety of fishing vessels that supply a major fish processing plant in Cabo Verde (Fig. 1a). This country, an important small island developing State, is promoting an economy based on the ocean resources and promoted a strong investment in the capacity to prepare fish-based products to overseas markets. Here we used a barcoding genetic method, based on Sanger DNA sequencing of two informative mitochondrial DNA (mtDNA) regions. We also took the opportunity that a parallel HTS analysis of stomach contents of the same individuals was done in other study, for the purpose of describing prey diversity, to understand how species identification could simultaneously be obtained using the same HTS pipeline.

2. Materials and methods

Frozen stomachs from Bigeye (N = 13), Yellowfin (N = 11) and Skipjack (N = 33) individuals, obtained for a related study on tuna diet, were obtained from an industrial fish processing plant in the harbour of Mindelo, Cabo Verde (Fig. 1a). This harbour is a major hub for tuna landings in the Eastern Central Atlantic fishing area, that are mainly exported to the European Union. Samples originated from fishing events in this area, from 23 vessels, ranging from semi-industrial to industrial vessels (Table S1). These individuals were captured from June to September of 2019, frozen and were processed in Mindelo in laboratory conditions during October 2019. Each stomach was thawed, the data on the label was digitised, a sample of muscle from other organs was retrieved from each individual and stored in 96% ethanol. Each stomach was then dissected, its content homogenized and two subsamples per stomach were taken into 2 mL tubes with 96% (v/v) ethanol. All samples were then refrigerated at 4 °C as soon as possible, transported to CIBIO research centre in Portugal and then stored at −20 °C in the laboratory until DNA extraction.

2.1. Sanger sequencing

DNA from the muscle tissue samples was extracted using a silica column-based protocol, EZ-10 Spin Column DNA Gel Extraction Kit (Bio Basic Inc., USA). Two different mtDNA informative regions were amplified with PCR and sequenced. The first region (645bp), from the mtDNA cytochrome *c* oxidase subunit I (COI), was amplified for all samples using the primers LCOI 121 and HCOI 1199 (Paine et al., 2007). The PCR protocol consisted of 5 μL of QIAGEN Multiplex PCR Master Mix (Qiagen, USA), 0.4 μL of each primer (10 nM), 2 μL of DNA template and nuclease-free water for a total volume of 11 μL. The conditions consisted of 15 min at 95 °C, 40 cycles of 30 s at 95 °C, 30 s at 57 °C and 60 s at 72 °C, followed by 10 min at 72 °C. The second region (388bp), from the control region (CR), was used to corroborate COI results for the two *Thunnus* species, and amplified using the primers L15998 and CSBDH (Viñas & Tudela, 2009). The PCR protocol consisted of 5 μL of QIAGEN Multiplex PCR Master Mix (Qiagen, USA), 0.4 μL of each primer (10 nM), 2 μL of DNA template and nuclease-free water for a total volume of 11 μL. The conditions consisted of 15 min at 95 °C, 35 cycles of 45 s at 94 °C, 45 s at 50 °C and 45 s at 72 °C, followed by 10 min at 72 °C.

All PCR products were sequenced on a 3730xl Genetic Analyzer (Applied Biosystems, USA) and results were edited and assembled in Geneious Prime 2022.0.2 (Biomatters, New Zealand). Species assignment was validated using two complementary approaches (Phylogenetic tree and BLAST search) for more robust results. First, for each mtDNA region, sequences were aligned and a phylogenetic tree was done using approximately-maximum-likelihood algorithm in FastTree 2 (Price et al., 2010) to taxonomically assign our sequences. For this approach we used reference sequences for all species of tuna and related species known to occur in our study area. This approach would minimize potential failures of the BLAST approach due to lack of genetic information on data repositories (Hestetun et al., 2020). Only three COI sequences did not have enough quality to be aligned, but partial alignment was

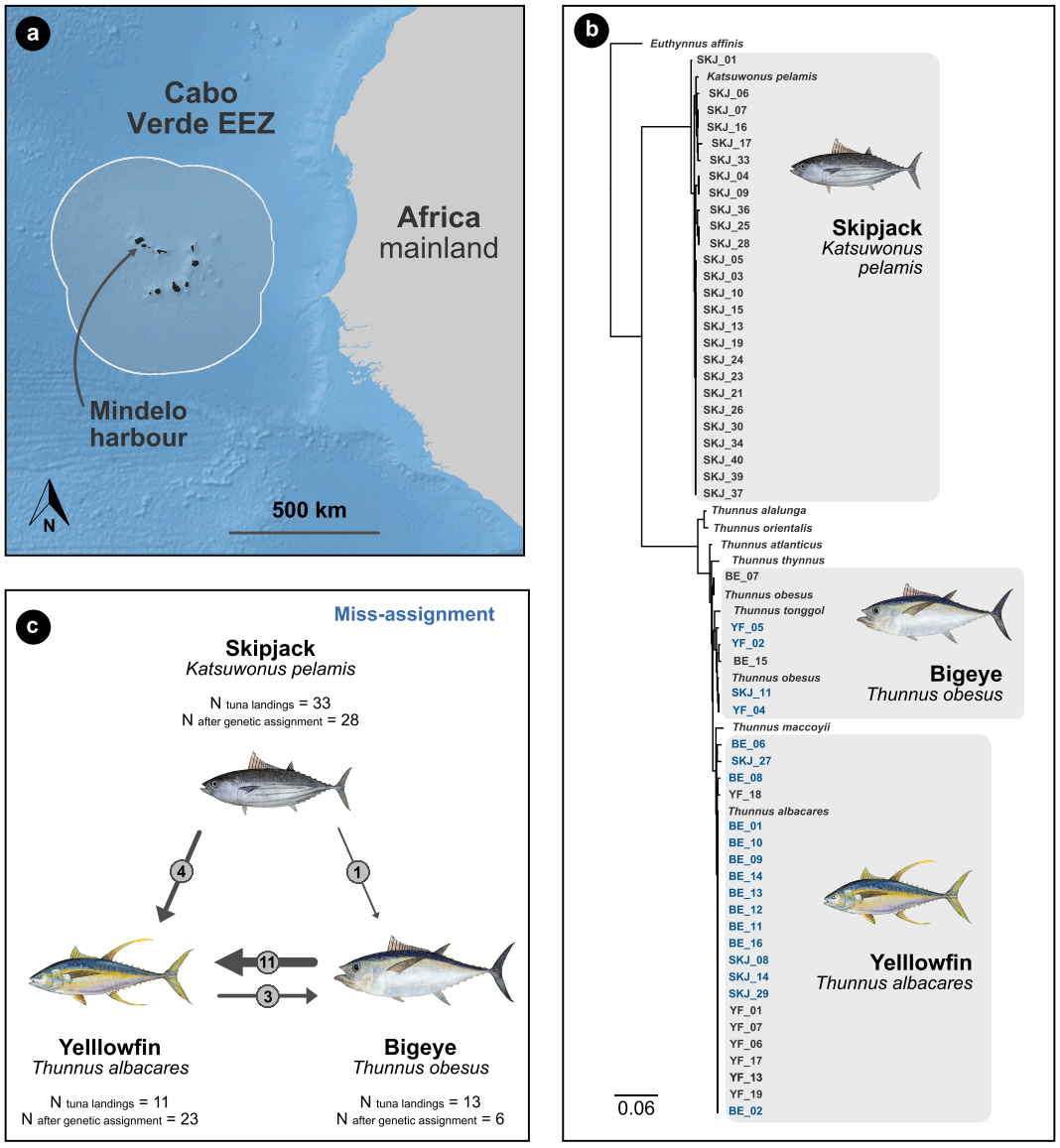

Fig. 1. a) Location of the Cabo Verde Archipelago, its Exclusive Economic Zone (EEZ) and the harbour of Mindelo; b) COI phylogenetic tree (approximately-maximum-likelihood) of the samples and reference sequences (branch lengths are given as substitutions per site). Samples that were incorrectly assigned are highlighted in blue; c) Diagram of changes on sample size before and after the genetic assignment of species identification. Lines show the direction of change of sample identity, while the width of lines is correlated to the number of samples misidentified (the number is also visualized at the middle of each line).

used to confirm the species assignment. Second, each sequence was compared (1000 hits) against the NCBI Genbank Nucleotide online data repository using the BLAST algorithm (Ye et al., 2006), to further confirm the assignment of each sequence to a single species. This approach would also decrease the likelihood of false positives due to introgression issues (Ciezarek et al., 2019).

2.2. HTS sequencing

This methodology was used with the stomach content samples collected with the main purpose of describing the prey diversity. All stomach content samples were extracted on a laboratory equipped for handling non-invasive samples, using the Stool DNA Isolation Kit (Norgen Biotek, Canada). For the purpose of this work, a mitochondrial DNA (mtDNA) fragment was amplified with the primer set 'MiFish-U' (Miya et al., 2015) and we did not used any blocking primer for tuna species. This set amplifies the 12S fragment preferentially from Osteichthyes, thus increasing the probability of detecting DNA from the tuna cells. However, with this primer set it is only possible to discriminate between Katsuwonus and Thunnus genera, since it does not discriminate between Thunnus species due the smaller size of the targeted sequence (313 bp). This issue was already known, but since this primer was chosen mainly for diet discrimination, it is very informative to show how well this primer can identify *Katsuwonus* samples using only stomach contents and at the same time identifying the prey diversity.

Library preparation followed the MiSeq protocol for 16S Metagenomics (Illumina, USA). The PCR protocol consisted of 5 μL of QIAGEN Multiplex PCR Master Mix (Qiagen, USA), 0.2 μL of each primer (10 nM)) 1 μL of DNA template and nuclease-free water for a total volume of 10 μL. The conditions consisted of 15 min at 95 °C, 35 cycles of 30 s at 94 °C, 30 s at 60 °C and 60 s at 72 °C, followed by 10 min at 72 °C. The final library was run in a MiSeq sequencer (Illumina, USA) using a v3 MiSeq reagent kit (Illumina, USA) for an expected average of 12,000 paired end reads per sample.

Bioinformatic analysis started with the alignment of paired-end reads using PEAR (Zhang et al., 2014), discarding both unassembled reads and alignments with overlapping quality score <26. Further

processing of sequencing reads was done using OBITools (Boyer et al., 2016) where reads were assigned to samples. Then primer sequences were removed using 'ngsfilter', allowing a total of four mismatches. Using 'obiuniq' and 'cluster_unoise' from VSEARCH (Rognes et al., 2016), reads were collapsed into exact sequence variants (ESVs) and error-corrected. These were compared against online databases (BOLD and NCBI Nucleotide Database) using the BLAST algorithm (Ye et al., 2006). Haplotypes were assigned to the lowest possible taxonomic level (e.g., family, order, species) for which hits in BLAST, with an identical match, clustered monophyletically. Finally, ESVs that matched Katsuwonus and Thunnus genera were filtered to provide a matrix of reads. Positive identifications were based on the relative proportions of the reads and on a threshold of 100 reads.

2.3. Statistical analysis

To test if the proportion of false identifications was higher in a given tuna species, a chi-square two-sample test for equality of proportions with Yates continuity correction was done on the number of Bigeye and Yellowfin tunas correctly and incorrectly assigned using the "prop.test"

Table 1
The initial and genetic assignment of species identification for samples of the three studied tuna species and the results from COI and CR (control region) sequencing with Sanger (BLAST query cover and identity percentage) and 12S sequencing with HTS (number of reads for each genera). An asterisk was added to the HTS reads when the total number of reads for both genera did not allow identification and the letter M added in brackets when both genera had a number of reads higher that the threshold for identification (100 reads). Two asterisks were added when a BLAST query was not conclusive.

Sample ID	Initial Assignment	Genetic Assignment	Result	Sanger sequencing - COI		Sanger sequencing - CR		HTS sequencing	
				Query Cover	Identity %	Query Cover	Identity %	Katsuwonus (N reads)	Thunnus (N reads)
YF_05	*Thunnus albacares*	*Thunnus obesus*	Incorrect	100	99.84	100	95.10	3	3359
SKJ_14	*Katsuwonus pelamis*	*Thunnus albacares*	Incorrect	100	100	100	99.74	17	6041
YF_06	*Thunnus albacares*	*Thunnus albacares*	Correct	100	100	100	99.74	8	1943
BE_11	*Thunnus obesus*	*Thunnus albacares*	Incorrect	100	100	100	98.45	2	3201
BE_12	*Thunnus obesus*	*Thunnus albacares*	Incorrect	100	100	100	99.48	0	4668
YF_18	*Thunnus albacares*	*Thunnus albacares*	Correct	100	99.84	100	98.97	4 *	28 *
BE_08	*Thunnus obesus*	*Thunnus albacares*	Incorrect	100	99.84	100	98.97	1	723
BE_10	*Thunnus obesus*	*Thunnus albacares*	Incorrect	100	100	100	98.71	3	3422
SKJ_26	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			2860	18
SKJ_17	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	99.69			3959	3
SKJ_20	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			4027	6
SKJ_29	*Katsuwonus pelamis*	*Thunnus albacares*	Incorrect	100	100		**	11	3831
YF_16	*Thunnus albacares*	*Thunnus albacares*	Correct	100	100	100	98.97	4	111
BE_09	*Thunnus obesus*	*Thunnus albacares*	Incorrect	100	100	100	98.45	13 *	1 *
BE_13	*Thunnus obesus*	*Thunnus albacares*	Incorrect	100	100	100	99.23	2	1078
BE_16	*Thunnus obesus*	*Thunnus albacares*	Incorrect	100	100	100	98.97	3	2267
SKJ_15	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			7 *	1 *
SKJ_10	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			3566	10
SKJ_28	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	99.84			2365 (M)	4442 (M)
SKJ_19	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			2534	12
SKJ_30	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			6594	7
SKJ_24	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			3276	4
SKJ_16	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			4524	5
SKJ_23	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			3911	1
BE_14	*Thunnus obesus*	*Thunnus albacares*	Incorrect	100	100	100	99.48	4	956
BE_15	*Thunnus obesus*	*Thunnus obesus*	Correct	100	99.69	100	94.33	4378 (M)	792 (M)
YF_04	*Thunnus albacares*	*Thunnus obesus*	Incorrect	100	100	96	94.07	2	3692
YF_02	*Thunnus albacares*	*Thunnus obesus*	Incorrect	100	99.84	100	93.06	3	2758
SKJ_02	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			3916	1
BE_06	*Thunnus obesus*	*Thunnus albacares*	Incorrect	100	99.69	100	99.48	5	2051
YF_13	*Thunnus albacares*	*Thunnus albacares*	Correct	100	100	100	100	3	3629
SKJ_39	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			2839	5
SKJ_36	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	99.84			3231	12
SKJ_04	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	99.69			4757	6
SKJ_40	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			408	7
SKJ_37	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			3498	3
SKJ_33	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	99.84			3618	29
SKJ_09	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	99.69			10112	2
SKJ_21	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			4582	7
SKJ_07	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			8030	3
SKJ_11	*Katsuwonus pelamis*	*Thunnus obesus*	Incorrect	100	100	–	–	4	6559
YF_01	*Thunnus albacares*	*Thunnus albacares*	Correct	100	100	100	99.23	6	243
SKJ_01	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	99.84			2351	5
YF_07	*Thunnus albacares*	*Thunnus albacares*	Correct	100	100	100	99.48	9	1641
SKJ_13	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			7291	5
SKJ_25	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	99.84			5138	3
YF_19	*Thunnus albacares*	*Thunnus albacares*	Correct	100	100	100	98.71	1	2518
SKJ_34	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			3605	10
BE_07	*Thunnus obesus*	*Thunnus obesus*	Correct	100	100	100	99.74	4	1559
SKJ_27	*Katsuwonus pelamis*	*Thunnus albacares*	Incorrect	100	100	100	99.23	12	5200
SKJ_06	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			6447	4
SKJ_05	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			7085	6
SKJ_08	*Katsuwonus pelamis*	*Thunnus albacares*	Incorrect	100	100	–	–	3	3528
BE_02	*Thunnus obesus*	*Thunnus albacares*	Incorrect	100	100	100	99.74	9	2928
BE_01	*Thunnus obesus*	*Thunnus albacares*	Incorrect	100	100	100	99.23	10	1313
YF_17	*Thunnus albacares*	*Thunnus albacares*	Correct	100	100	100	98.45	2	164
SKJ_03	*Katsuwonus pelamis*	*Katsuwonus pelamis*	Correct	100	100			3876	3

command in R 4.1.3 statistical computing environment (R Core Team, 2022).

3. Results

For all samples it was possible to reach a correct identification using either Sanger or HTS sequencing. In the Sanger approach there was concordance between the results of both mtDNA markers, using phylogenetic trees and BLAST queries. In all Sanger sequences the identity percentage was higher than 99.69% (COI) and 93.06% (CR), always with a 100% query cover (except for one sequence = 96%) (Fig. 1b and Table 1). Only in one sample (SKJ_29) we found a discordance, since it was genetically assigned as Yellowfin by the Sanger and HTS approaches except for the BLAST algorithm in the CR segment. This algorithm found a better match to one Skipjack and two Atlantic bluefin tuna (*Thunnus thynnus*) sequences, followed by 845 Yellowfin sequences matches, all with high identity percentages. The fact that all other methods support the identification as Yellowfin, and the lack of other matches with Skipjack and Bluefin sequences, allow us to infer that this outlier may be most likely due to the miss-identification of these original Genbank sequences and/or evidence of introgression of Yellowfin DNA into Bluefin genome. In the HTS approach, three samples failed to amplify enough reads to identify the genera and in two other samples the two tunid genera were found (Table 1), indicating most likely the occurrence of an event of predation rather than a contamination between samples.

We found that a miss-assignment occurred in 33% (19 out of 57) of the individuals (Fig. 1c). More than half of yellowfin samples (15 out of 23) were wrongly assigned as bigeyes (11 samples) or skipjacks (4 samples). We also found that four out of six bigeyes were wrongly assigned as yellowfins (3 samples) or skipjacks (1 sample). None of the skipjacks (28 samples) were incorrectly assigned. The proportion of wrongly assigned samples in the batch of samples identified as bigeyes by fishermen (11 out of 13) was significantly higher than the same proportion in the batch of samples identified as yellowfins by fishermen (3 out of 11) ($\chi^2_1 = 5.874$, $P = 0.008$). Miss-identification occurred in 12 vessels, all industrial boats. However, sample sizes were highly biased towards industrial vessels (semi-industrial = 8 samples vs industrial = 33 samples). Miss-identified samples per vessel ranged from one (7 out of 19) to three (6 out of 19) samples.

4. Discussion

Our results show that miss-identification can occur at the very beginning of the food processing pipeline, even prior to reaching a processing plant. Due to the highly variable nature of fisheries, traceability of tuna species and transversal quality control procedures may be very hard to implement at this stage, increasing the likelihood of detecting miss-identifications. For instance, vessels can be either artisanal, industrial, or semi-industrial and can use different strategies (e.g., baited longlines, handlines, purse seine). Also, vessels are either national or belong to foreign countries (mostly Spain) and they may fish in different areas in the Atlantic Eastern Central zone, from inside the exclusive economic zone (EEZ) of Cabo Verde to international waters off west Africa, in front of Mauritania and Senegal (Table S1). To minimize this effect, we strongly advise to focus on procedures of quality control regarding species identification when fish landings arrive, since tuna landings are concentrated in a small number of harbours. Mindelo (Cabo Verde), Abidjan (Ivory Coast) and Dakar (Senegal) account for the majority of landings in 2019 (Thorpe et al., 2022), although transhipment at sea may also occur.

We also found that this miss-identification is transversal to multiple fish landings, but we did not find clear evidence that this pattern was intentional, considering that our data is scattered among many vessels, resulting in a small sample size per vessel. Skipjack tuna catches are the highest in the Central East Atlantic, followed by Yellowfin and Bigeye tunas (Thorpe et al., 2022). There is evidence that Skipjack purse seine with fish aggregating devices (dFADs) is associated with tuna bycatch of juvenile Yellowfin and Bigeye (McCluney et al., 2019), and we only found true tunas mistakenly assigned as skipjacks and not the other way around. On the other hand, we did find directionality towards miss-assigning species identification towards higher valued species, since we found a higher percentage of yellowfins (lower value) assigned as bigeyes (higher value).

From a conservation perspective, it is also important to highlight that the Bigeye tunas have been internationally assessed as vulnerable (Collette et al., 2021c), while Yellowfin and Skipjack tuna populations have been evaluated as non-threatened (Collette et al., 2021a; 2021b). These classifications status may then regulate fishing quotas, both in volume and time of harvesting. It is therefore important to assess in the future the effect of these underestimations on the assessment of Bigeye tuna viability. Also, from a toxicological angle, Skipjack tuna have some of the lowest heavy metal concentrations (e.g., mercury) among these tuna species, due to their life history parameters and smaller maximum size (Romero et al., 2021). The mislabelling of other tunas as Skipjack may obscure the patterns of transference to humans of these toxic substances.

Concerning methodologies, species assignment based on Sanger sequencing still proved to be the most logistically feasible method. But it is opportune to evaluate the use of emerging methodologies (especially HTS pipelines) that are now available. HTS sequencing proved to be also effective, but with a lower success rate. However, we did not make any effort to optimize the protocol for this purpose, and increasing the number of extraction and/or PCR replicates and the number of reads per sample may help to ensure that each sample does amplify enough predator DNA. On the other hand, many NGS sequencers are optimized to run many samples at the same time (especially Illumina sequencing), an issue that may impair the use of this technology for small projects. But these sequencers allow to obtain multiple information (diet, predator species identification, and other data, such as parasite composition) from the same tissue samples, thus decreasing costs and increasing informative data. In the case of small projects, other HTS technologies, especially based on Nanopore sequencing, may prove to be as effective, both in time and cost per sample, with a lower dependence on equipment and specialized facilities. Another genetic methodology, loop-mediated isothermal amplification (LAMP) may also prove to be important as a cost and time efficient method in specific cases, where the optimization time is worthwhile (Xiong et al., 2021; Xu et al., 2022).

In conclusion, we expose in this paper the need to ensure better traceability of food products from the source to the consumer, highlighting the magnitude of misidentification at the very beginning of the whole process, at the fishing vessels.

Funding

This work received financial and logistic support for fieldwork campaigns, from the project Alcyon – Conservation of seabirds from Cabo Verde, coordinated by BirdLife International and funded by the MAVA Foundation [MAVA17022; http://en.mava-foundation.org], through its strategic plan for West Africa (2017–2022). This study also benefitted from funding by the strategic program of MARE, financed by FCT [UID/MAR/04292/2020], through national funds. ARC and VAM were supported by Fundação para a Ciência e Tecnologia (FCT; SFRH/BD/139019/2018 and 2020.02547.CEECIND, respectively). RJL was supported by a FCT - Transitory Norm contract [DL57/2016/CP1440/CT0006].

CRediT authorship contribution statement

Ana Rita Carreiro: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Roles, Writing – original draft. **Jaime A. Ramos:** Conceptualization, Funding acquisition, Methodology, Writing – review & editing. **Vanessa Mata:** Formal analysis,

Investigation, Methodology, Writing – review & editing. **Nathalie M. Almeida:** Formal analysis, Investigation, Methodology, Writing – review & editing. **Vítor H. Paiva:** Conceptualization, Funding acquisition, Methodology, Writing – review & editing. **Ricardo Jorge Lopes:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Visualization, Roles.

Declaration of competing interest

None.

Data availability

Sanger sequencing data was shared as supplementary files.

Acknowledgments

We would like to thank the generous help from one fisheries processing plant in Mindelo, Cabo Verde for their assistance with the collection of the data. We also thank the "Universidade Técnica do Atlântico – UTA" for allowing the use of laboratory facilities and especially the support given by Profs. Evandro Lopes, Rui Freitas and Corrine Almeida. Species illustrations based on public domain illustrations from NOAA FishWatch.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodcont.2022.109473.

References

Abdullah, A., & Rehbein, H. (2014). Authentication of raw and processed tuna from Indonesian markets using DNA barcoding, nuclear gene and character-based approach. *European Food Research and Technology, 239*(4), 695–706. https://doi.org/10.1007/s00217-014-2266-0

Amaral, C. R. L., Maciel, V., Goldenberg-Barbosa, R., Silva, D. A., Amorim, A., & Carvalho, E. F. (2017). Tuna fish identification using mtDNA markers. *Forensic Science International: Genetics Supplement Series, 6*, e471–e473. https://doi.org/10.1016/j.fsigss.2017.09.193

Badia-Melis, R., Mishra, P., & Ruiz-García, L. (2015). Food traceability: New trends and recent advances. A review. *Food Control, 57*, 393–401. https://doi.org/10.1016/j.foodcont.2015.05.005

Barendse, J., Roel, A., Longo, C., Andriessen, L., Webster, L. M. I., Ogden, R., & Neat, F. (2019). DNA barcoding validates species labelling of certified seafood. *Current Biology, 29*(6), R198–R199. https://doi.org/10.1016/j.cub.2019.02.014

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M. I., Beacham, T. D., Maes, G. E., Martinsohn, J. T., Miller, K. M., Naish, K. A., Ovenden, J. R., Primmer, C. R., Young Suk, H., Therkildsen, N. O., & Withler, R. E. (2017). Harnessing the power of genomics to secure the future of seafood. *Trends in Ecology & Evolution, 32*(9), 665–680. https://doi.org/10.1016/j.tree.2017.06.010

Blanco-Fernandez, C., Garcia-Vazquez, E., & Machado-Schiaffino, G. (2021). Seventeen years analysing mislabelling from DNA barcodes: Towards hake sustainability. *Food Control, 123*, Article 107723. https://doi.org/10.1016/j.foodcont.2020.107723

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: a unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources, 16*(1), 176–182. https://doi.org/10.1111/1755-0998.12428. https://doi.org/10.1111/1755-0998.12428.

Brill, R. W., & Hobday, A. J. (2017). Tunas and their fisheries: Safeguarding sustainability in the twenty-first century. *Reviews in Fish Biology and Fisheries, 27*(4), 691–695. https://doi.org/10.1007/s11160-017-9500-3

Cawthorn, D.-M., & Mariani, S. (2017). Global trade statistics lack granularity to inform traceability and management of diverse and high-value fishes. *Scientific Reports, 7*(1), Article 12852. https://doi.org/10.1038/s41598-017-12301-x

Chang, C.-H., Kao, Y.-T., Huang, T.-T., & Wang, Y.-C. (2021). Product authentication using two mitochondrial markers reveals inconsistent labeling and substitution of canned tuna products in the Taiwanese market. *Foods, 10*(11), 2655. https://doi.org/10.3390/foods10112655

Chang, C.-H., Tsai, M.-L., Huang, T.-T., & Wang, Y.-C. (2021). Authentication of fish species served in conveyor-belt sushi restaurants in Taiwan using DNA barcoding. *Food Control, 130*, Article 108264. https://doi.org/10.1016/j.foodcont.2021.108264

Ciezarek, A. G., Osborne, O. G., Shipley, O. N., Brooks, E. J., Tracey, S. R., McAllister, J. D., Gardner, L. D., Sternberg, M. J. E., Block, B., & Savolainen, V. (2019). Phylotranscriptomic insights into the diversification of endothermic Thunnus tunas. *Molecular Biology and Evolution, 36*(1), 84–96. https://doi.org/10.1093/molbev/msy198

Cochrane, K. L. (2021). Reconciling sustainability, economic efficiency and equity in marine fisheries: Has there been progress in the last 20 years? *Fish and Fisheries, 22*(2), 298–323. https://doi.org/10.1111/faf.12521. https://doi.org/10.1111/faf.12521.

Collette, B. B., Boustany, A., Fox, W., Graves, J., Juan Jorda, M., & Restrepo, V. (2021a). Katsuwonus pelamis. Retrieved 06 October 2022 from https://dx.doi.org/10.2305/IUCN.UK.2021-2.RLTS.T170310A46644566.en.

Collette, B. B., Boustany, A., Fox, W., Graves, J., Juan Jorda, M., & Restrepo, V. (2021b). Thunnus albacares. Retrieved 06 October 2022 from https://dx.doi.org/10.2305/IUCN.UK.2021-2.RLTS.T21857A46624561.en .

Collette, B. B., Boustany, A., Fox, W., Graves, J., Juan Jorda, M., & Restrepo, V. (2021c). Thunnus obesus. Retrieved 06 October 2022 from https://dx.doi.org/10.2305/IUCN.UK.2021-2.RLTS.T21859A46912402.en.

Collette, B. B., & Graves, J. (2019). *Tunas and billfishes of the world*. Johns Hopkins University Press.

Delpiani, G., Delpiani, S. M., Deli Antoni, M. Y., Covatti Ale, M., Fischer, L., Lucifora, L. O., & Díaz de Astarloa, J. M. (2020). Are we sure we eat what we buy? Fish mislabelling in buenos aires province, the largest sea food market in Argentina. *Fisheries Research, 221*, Article 105373. https://doi.org/10.1016/j.fishres.2019.105373

Díaz-Arce, N., Arrizabalaga, H., Murua, H., Irigoien, X., & Rodríguez-Ezpeleta, N. (2016). RAD-seq derived genome-wide nuclear markers resolve the phylogeny of tunas. *Molecular Phylogenetics and Evolution, 102*, 202–207. https://doi.org/10.1016/j.ympev.2016.06.002

Dimitrakopoulou, M.-E., & Vantarakis, A. (2021). Does traceability lead to food authentication? A systematic review from A European perspective. *Food Reviews International, 1–23*. https://doi.org/10.1080/87559129.2021.1923028

Espiñeira, M., Gonzalez-Lavín, N., Vieites, J. M., & Santaclara, F. J. (2009). Development of a method for the identification of scombroid and common substitute species in seafood products by FINS. *Food Chemistry, 117*(4), 698–704. https://doi.org/10.1016/j.foodchem.2009.04.087

Fanelli, V., Mascio, I., Miazzi, M. M., Savoia, M. A., De Giovanni, C., & Montemurro, C. (2021). Molecular approaches to agri-food traceability and authentication: An updated review. *Foods, 10*(7). https://doi.org/10.3390/foods10071644

Fernandes, T. J. R., Amaral, J. S., & Mafra, I. (2021). DNA barcode markers applied to seafood authentication: An updated review. *Critical Reviews in Food Science and Nutrition, 61*(22), 3904–3935. https://doi.org/10.1080/10408398.2020.1811200

Galimberti, A., Casiraghi, M., Bruni, I., Guzzetti, L., Cortis, P., Berterame, N. M., & Labra, M. (2019). From DNA barcoding to personalized nutrition: The evolution of food traceability. *Current Opinion in Food Science, 28*, 41–48. https://doi.org/10.1016/j.cofs.2019.07.008

Gordoa, A., Carreras, G., Sanz, N., & Viñas, J. (2017). Tuna species substitution in the Spanish commercial chain: A knock-on effect. *PLoS One, 12*(1), Article e0170809. https://doi.org/10.1371/journal.pone.0170809

Guillotreau, P., Squires, D., Sun, J., & Compeán, G. A. (2017). Local, regional and global markets: What drives the tuna fisheries? *Reviews in Fish Biology and Fisheries, 27*(4), 909–929. https://doi.org/10.1007/s11160-016-9456-8

Haas, B., Mackay, M., Novaglio, C., Fullbrook, L., Murunga, M., Sbrocchi, C., McDonald, J., McCormack, P. C., Alexander, K., Fudge, M., Goldsworthy, L., Boschetti, F., Dutton, I., Dutra, L., McGee, J., Rousseau, Y., Spain, E., Stephenson, R., Vince, J., … Haward, M. (2021). The future of ocean governance. *Reviews in Fish Biology and Fisheries*. https://doi.org/10.1007/s11160-020-09631-x

Helgoe, J., Oswald, K. J., & Quattro, J. M. (2020). A comprehensive analysis of the mislabeling of Atlantic cod (Gadus morhua) products in Spain. *Fisheries Research, 222*, Article 105400. https://doi.org/10.1016/j.fishres.2019.105400

Hellberg, R. S., Isaacs, R. B., & Hernandez, E. L. (2019). Identification of shark species in commercial products using DNA barcoding. *Fisheries Research, 210*, 81–88. https://doi.org/10.1016/j.fishres.2018.10.010

Hestetun, J. T., Bye-Ingebrigtsen, E., Nilsson, R. H., Glover, A. G., Johansen, P.-O., & Dahlgren, T. G. (2020). Significant taxon sampling gaps in DNA databases limit the operational use of marine macrofauna metabarcoding. *Marine Biodiversity, 50*(5), 70. https://doi.org/10.1007/s12526-020-01093-5

Kappel, K., & Schröder, U. (2020). Difficulties in DNA barcoding-based authentication of snapper products due to ambiguous nucleotide sequences in public databases. *Food Control, 118*, Article 107348. https://doi.org/10.1016/j.foodcont.2020.107348

Kumar, G. (2018). Mercury concentrations in fresh and canned tuna: A review. *Reviews in Fisheries Science & Aquaculture, 26*(1), 111–120. https://doi.org/10.1080/23308249.2017.1362370

Kusche, H., & Hanel, R. (2021). Consumers of mislabeled tropical fish exhibit increased risks of ciguatera intoxication: A report on substitution patterns in fish imported at frankfurt airport, Germany. *Food Control, 121*, Article 107647. https://doi.org/10.1016/j.foodcont.2020.107647

Liu, C. J. N., Neo, S., Rengifo, N. M., French, I., Chiang, S., Ooi, M., Heng, J. M., Soon, N., Yeo, J. Y., Bungum, H. Z., Ota, K., Koul, A. A., Poh, Y. H., & Wainwright, B. J. (2021). Sharks in hot soup: DNA barcoding of shark species traded in Singapore. *Fisheries Research, 241*, Article 105994. https://doi.org/10.1016/j.fishres.2021.105994

Lockley, A. K., & Bardsley, R. G. (2000). Novel method for the discrimination of tuna (Thunnus thynnus) and bonito (Sarda sarda) DNA. *Journal of Agricultural and Food Chemistry, 48*(10), 4463–4468. https://doi.org/10.1021/jf000387p

Love, D. C., Nussbaumer, E. M., Harding, J., Gephart, J. A., Anderson, J. L., Asche, F., Stoll, J. S., Thorne-Lyman, A. L., & Bloem, M. W. (2021). Risks shift along seafood supply chains. *Global Food Security, 28*, Article 100476. https://doi.org/10.1016/j.gfs.2020.100476

McCluney, J. K., Anderson, C. M., & Anderson, J. L. (2019). The fishery performance indicators for global tuna fisheries. *Nature Communications, 10*(1), 1641. https://doi.org/10.1038/s41467-019-09466-6

Minoudi, S., Karaiskou, N., Avgeris, M., Gkagkavouzis, K., Tarantili, P., Triantafyllidou, D., Palilis, L., Avramopoulou, V., Tsikliras, A., Barmperis, K., & Triantafyllidis, A. (2020). Seafood mislabeling in Greek market using DNA barcoding. *Food Control, 113*, Article 107213. https://doi.org/10.1016/j.foodcont.2020.107213

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., Minamoto, T., Yamamoto, S., Yamanaka, H., Araki, H., Kondoh, M., & Iwasaki, W. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species. *Royal Society Open Science, 2* (7), Article 150088. https://doi.org/10.1098/rsos.150088

Paine, M., McDowell, J., & Graves, J. (2007). Specific identification of Western Atlantic Ocean scombrids using mitochondrial DNA cytochrome c oxidase subunit I (COI) gene region sequences. *Bulletin of Marine Science, 80*, 353–367.

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One, 5*(3), Article e9490. https://doi.org/10.1371/journal.pone.0009490

R Core Team. (2022). R: A language and environment for statistical computing. In *R foundation for statistical computing*. https://www.R-project.org/.

Rahmani, J., Fakhri, Y., Shahsavani, A., Bahmani, Z., Urbina, M. A., Chirumbolo, S., Keramati, H., Moradi, B., Bay, A., & Bjørklund, G. (2018). A systematic review and meta-analysis of metal concentrations in canned tuna fish in Iran and human health risk assessment. *Food and Chemical Toxicology, 118*, 753–765. https://doi.org/10.1016/j.fct.2018.06.023

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). Vsearch: A versatile open source tool for metagenomics. *PeerJ, 4*, Article e2584. https://doi.org/10.7717/peerj.2584

Romero, J., Catry, P., Hermida, M., Neves, V., Cavaleiro, B., Gouveia, L., & Granadeiro, J. P. (2021). Tunas off northwest Africa: The epipelagic diet of the Bigeye and Skipjack tunas. *Fisheries Research, 238*, Article 105914. https://doi.org/10.1016/j.fishres.2021.105914

Ruiz-Salmón, I., Laso, J., Margallo, M., Villanueva-Rey, P., Rodríguez, E., Quinteiro, P., Dias, A. C., Almeida, C., Nunes, M. L., Marques, A., Cortés, A., Moreira, M. T., Feijoo, G., Loubet, P., Sonnemann, G., Morse, A. P., Cooney, R., Clifford, E., Regueiro, L., … Aldaco, R. (2021). Life cycle assessment of fish and seafood processed products – a review of methodologies and new challenges. *The Science of the Total Environment, 761*, Article 144094. https://doi.org/10.1016/j.scitotenv.2020.144094

Servusova, E., & Piskata, Z. (2021). Identification of selected tuna species in commercial products. *Molecules, 26*(4). https://doi.org/10.3390/molecules26041137

Seto, K., Galland, G. R., McDonald, A., Abolhassani, A., Azmi, K., Sinan, H., Timmiss, T., Bailey, M., & Hanich, Q. (2021). Resource allocation in transboundary tuna fisheries: A global analysis. *Ambio, 50*(1), 242–259. https://doi.org/10.1007/s13280-020-01371-3

da Silva, C. F., Daneluz, C. M., Camacho-Oliveira, R. B., do Prado, F. D., Foresti, F., Rodrigues, C. E., & Porto-Foresti, F. (2019). DNA Barcode reveals mislabelling in the identification of marine fish swimming bladders for commercialization. *Forensic Science International, 299*, 41–43. https://doi.org/10.1016/j.forsciint.2019.03.001

Thorpe, A., Hermansen, O., Pollard, I., Isaksen, J., Failler, P., & Touron-Gardic, G. (2022). Unpacking the tuna traceability mosaic – EU SFPAs and the tuna value chain. *Marine Policy, 139*, Article 105037. https://doi.org/10.1016/j.marpol.2022.105037

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

Willette, D. A., Navarrete-Forero, G., Gold, Z., Lizano, A. M. D., Gonzalez-Smith, L., & Sotil, G. (2021). Characterizing industrial and artisanal fishing vessel catch composition using environmental DNA and satellite-based tracking data. *Foods, 10* (6). https://doi.org/10.3390/foods10061425

Xiong, X., Xu, W., Guo, L., An, J., Huang, L., Qian, H., Cui, X., Li, Y., Cao, M., Xiong, X., Ying, X., & Wang, L. (2021). Development of loop-mediated isothermal amplification (LAMP) assay for rapid screening of skipjack tuna (Katsuwonus pelamis) in processed fish products. *Journal of Food Composition and Analysis, 102*, Article 104038. https://doi.org/10.1016/j.jfca.2021.104038

Xiong, X., Yuan, F., Huang, M., Lu, L., Xiong, X., & Wen, J. (2019). DNA barcoding revealed mislabeling and potential health concerns with roasted fish products sold across China. *Journal of Food Protection, 82*(7), 1200–1209. https://doi.org/10.4315/0362-028X.JFP-18-514

Xu, W., Li, Q., Cui, X., Cao, M., Xiong, X., Wang, L., & Xiong, X. (2022). Real-time loop-mediated isothermal amplification (LAMP) using self-quenching fluorogenic probes: The application in skipjack tuna (Katsuwonus pelamis) authentication. *Food Analytical Methods, 15*(3), 658–665. https://doi.org/10.1007/s12161-021-02159-1

Yang, Q.-Q., Qian, Z.-X., Ye, Z.-H., Zhou, A.-N., Zhao, X.-X., Zhang, P.-J., Liu, G.-F., & Yu, X.-P. (2021). Widespread mislabeling of nonnative apple snails (ampullariidae: Pomacea) as native field snails (Viviparidae: Cipangopaludina) on the Chinese food markets. *Aquaculture, 530*, Article 735756. https://doi.org/10.1016/j.aquaculture.2020.735756

Yao, L., Lu, J., Qu, M., Jiang, Y., Li, F., Guo, Y., Wang, L., & Zhai, Y. (2020). Methodology and application of PCR-RFLP for species identification in tuna sashimi [10.1002/fsn3.1552]. *Food Sciences and Nutrition, 8*(7), 3138–3146. https://doi.org/10.1002/fsn3.1552

Ye, J., McGinnis, S., & Madden, T. L. (2006). Blast: Improvements for better sequence analysis. *Nucleic Acids Research, 34*(suppl_2), W6–W9. https://doi.org/10.1093/nar/gkl164

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: A fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics, 30*(5), 614–620. https://doi.org/10.1093/bioinformatics/btt593