Next Generation Sequencing (NGS) approach applied to species identification in mixed fishery products

Roberta Piredda[1], Anna Mottola[1]*, Giulia Cipriano[2,3], Roberto Carlucci[2,3], Giuseppina Ciccarese[4], Angela Di Pinto[1]

[1]Department of Veterinary Medicine – University of Bari Aldo Moro – Prov. le Casamassima, km 3 – 70010 Valenzano (Bari), Italy

[2]Department of Biology, University of Bari, Via Orabona 4, Bari, 70125, Italy

[3]CoNISMa, Piazzale Flaminio 9, Rome, 00196, Italy

[4] Zooprophylactic Institute of Puglia and Basilicata, S.S. 7 Ter, Km 59, Zona Industriale, 73012, Campi Salentina, Lecce, Italy

*Corresponding author: Anna Mottola

Department of Veterinary Medicine – University of Bari Aldo Moro – Prov. le Casamassima, km 3 - 70010 Valenzano (Bari) - ITALY

Phone: +390804679878

Fax: +390805443855

e-mail: anna.mottola@uniba.it

22

2

Abstract

Multi-species fish products are a growing market segment driven by several inputs, including urbanization, modern lifestyles and the availability of new technologies. These products are characterized by the absence of species-related morphological traits, as well as by the use of discards or fish waste as ingredients, that cannot be traced using a classical molecular approach. This historic inability to identify ingredients, coupled with a rather lean list of mandatory information required by current legislation, make these products the perfect target for substitution fraud. New molecular tools are now available to overcome this gap, such as the metabarcoding approach which can be used to trace species from complex food matrices. In this study, we used such an approach to sequence a fragment of 16S rRNA mitochondrial gene (~200bp) from 20 multispecies processed seafood products, sold on the Italian market and including breaded, burger and surimi, in order to evaluate mislabeling rates and formulate hypotheses regarding the potential drivers of fraudulent activities. Our results highlighted the presence of 120 marine taxa with an overall mislabeling rate of 45%, including some low-abundance taxa not included in the current Italian legislation, found mainly in the surimi samples. The presence of Tetraodontidae and swine DNA raise concerns regarding consumer safety and protection with regard to ethical or religious issues. The high number of low-frequency taxa confirms the inclusion of discards and fish waste, doubtless positive for sustainable fisheries, but also boosting profits for the fish industry. Such practice, however, without an adequate labelling and traceability system, is a source of risk for consumers and marine ecosystems. In this context, a revision of national and international food safety legislation is needed; indeed, metabarcoding assessments can provide useful information to stakeholders and act as the future operational tool for inspective monitoring.

Keywords: seafood fraud, DNA-based methods, NGS, multispecies fish products, food metabarcoding

1. Introduction

During the last two decades, massive changes have occurred in food habits, in terms of quantity and quality, driven by several inputs such as urbanization, modern lifestyles and the availability of new technologies, shifting consumer preferences toward ready-to-eat or ready-to-cook products (Giusti et al., 2017a; Mottola et al., 2020). Indeed, processed seafood market growth is expected because fish and seafood are relatively cheap and nutritious (high-protein coupled with low fat and calories).

Fish-based ready-to-eat/cook foods cover a wide range of product types featuring different species, food-processing methodologies and final packaging. One of the most popular and cheaper types is surimi (Okazaki & Kimura, 2014), but burgers and breaded fish products are increasing in terms of production and quality thanks to technological improvements, such as the mechanical separation of meat (MSM), which makes it possible to include fish waste without altering the nutritional value of whole fish (Borgogno et al. 2017). After fillets, fish burgers are regarded as the healthiest products, perceived by consumers as a good way of getting children and elderly people to eat fish (Paci et al. 2018). By contrast, breaded products, which are made of many ingredients and generally pre-fried, are perceived by consumers as being artificial and unhealthy. However, children and adolescents often dislike fresh fish products, driving families toward breaded or other processed complex products, as they contain no fish bones and have a less fishy flavour (Husein, 2019).

Food labelling is the most important instrument for informing consumers and providing essential guarantees of safe fish (Di Pinto at al., 2015, 2016). Labels not only safeguard consumer safety, but also help people to differentiate and choose products based on food attributes or such ethical issues as sustainable production, animal health and wellness, health problems or religious laws. However, substitutions of species, whether by accidental or deliberate fraud, are frequent, and seafood ranks among the most commonly counterfeited foodstuffs (Di Pinto at al.,2015, 2016; Marchetti et al., 2020). Fish substitutions are largely driven by a desire for economic gain (with high-priced species

being substituted by cheaper ones), but recently, meta-analyses on mislabelling have highlighted the complexity of the seafood market and multiple context-dependent causes (Donlan et al. 2019). Therefore, a robust labelling and control system is even more important in mixed products that may simultaneously include different fish or mollusc species and, due to the use of several food-processing methodologies, always generate products lacking morphological traits related to the species used as ingredients. Giusti et al. (2017b) generated a table of 89 fish and cephalopod species reported in literature as ingredients used in surimi-based products, whereas no such specific information is available for burger and breaded products. Moreover, all these product categories are well known to include by-product, underutilized, less valuable, non-directly marketable fish species, but also fish waste indicated as a source of sustainable and nutritious food (Palmeira et al., 2016).

In most studies, seafood authentication involves DNA methodologies (Luque & Donlan, 2019); indeed, DNA-barcoding is regarded as the most important policy instrument for species identification (Clark, 2015). However, in the presence of a food matrix containing multiple species, Sanger sequencing produces no useful output, showing only the dominant component and failing to identify the other species present (Paracchini et al. 2019). Even though PCR cloning has previously been used for species identification in mixed products (Galal-Khallaf et al. 2016; Silva, 2019), the High throughput sequencing (HTS), also known as Next Generation Sequencing (NGS), ~~the high throughput sequencing (HTS) metabarcoding approach~~ is the most promising tool for routine analysis of mixed-species food inspection without requiring any previous knowledge about the species (Haynes et al. 2019; Noh et al., 2021). A metabarcoding approach based on 16S rRNA gene amplification has been applied to analyse the microbial communities associated with fish and fish products (Sørensen et al. 2020; Sun et al. 2020; Zhang et al. 2019; Zhuang et al. 2019). Moreover, several studies have tested the metabarcoding approach in different food products, including meat (Cottenet et al. 2020; Xing et al. 2019), dairy products (Ribani et al. 2018), commercial plant products (Bruno et al. 2019), herbal

ha formattato: Tipo di carattere: (Predefinito) Times New Roman, 12 pt

ha formattato: Evidenziato

medicinal (Anthoons et al. 2021), honey (Wirta et al. 2021) and pet food (Palumbo et al. 2020). Few studies have explored the application of NGS on processed seafood but they have shown that metabarcoding has detected species from commercial or artificially-prepared mixtures, even revealing the presence of non-target species (Carvalho et al. 2017; Franco et al., 2021; Giusti et al. 2017b; Ho et al. 2020; Kappel et al. 2017; Noh et al. 2021; Voorhuijzen-Harink et al. 2019).

ha formattato: Evidenziato

Therefore, it is now possible to track ingredients even in the case of multispecies products that were not traceable in the past, providing new elements to understand the complex relationships among different components in the food sector. To this aim, the taxonomic profiles generated in this study using the metabarcoding approach in a wide range of multispecies processed seafood products were analysed to evaluate mislabelling rates and provide a hypothesis of the nature of the potential drivers of fraudulent activities, as well implications and consequences for consumers and impacts on fisheries activities and marine ecosystems.

2. Materials and methods

2.1 Sample collection

A total of 20 processed seafood product samples were collected including three different types: i) ten samples of highly processed surimi-based products (eight frozen, one in brine and one in modified atmosphere packaging); ii) five samples (all frozen) of breaded products (two sticks, two cutlets, and one nugget); iii) five samples of burger (two frozen and three in modified atmosphere packaging). They were purchased from different markets in south-eastern Italy (Apulia region), but the labels indicated different production sites (Table 1 and Full description in Supplementary Table 7). All samples were stored at -20 °C until DNA extraction.

2.2 DNA extraction and sequencing

Genomic DNA was extracted and purified starting from 25 mg aliquots of sample, using the DNeasy Blood and Tissue Kit (QIAGEN, Hilden, Germany) as reported by Marchetti et al. (2020). Negative extraction control (no added tissue) was included to verify the purity of the extraction reagents. The concentration and purity of DNA were established by evaluating the A260 nm/A280 nm ratio using a BioPhotometer D30 filter (Eppendorf, Milan, Italy). The DNA samples were amplified with the primer pairs 16sf-var 5´-CAAATTACGCTGTTAT CCCTATGG-3´ and 16sr-var 5´-GACGAGAAGACCCTAATGAGCTTT-3´ designed by Chapela et al. (2002), targeting a fragment of ~200bp of the 16S ribosomal RNA mitochondrial gene. The region was then sequenced (2 x 150 bp sequencing) on the Illumina NextSeq platform by LGC Genomics GmbH (Berlin, Germany). Raw sequences were deposited in the Sequence Read Archive (SRA) under the BioProject XX.

2.3 Data processing and taxonomic assignment

Paired-end reads were processed using Mothur v. 1.45.2 (Schloss et al., 2009). Contigs between read pairs were assembled and differences in base calls in the overlapping region were solved using ΔQ parameter as described in Kozich et al. (2013). Primer sequences were removed (pdiffs = 3), and no ambiguous bases were allowed; the maximum homopolymer size was 8 bp and reads shorter than 80 bp were removed. The remaining reads were dereplicated (unique haplotypes or OTU at 100%) and screened for chimeras using UCHIME in *de novo* mode (Edgar et al., 2011). OTUs with total abundance ≤ 50 reads were removed and taxonomic assignment of the remained dataset was performed using standalone blast in the blast+ suite (Altschul et al., 1990; Camacho et al., 2009) against the 16S mitochondrial custom database (16S_DB), discarding the assignments with a similarity of < 90%. Reads assigned to the same species in the range 100-98% similarity were merged and considered at species level, and values lower than 98% as genus. In the case of ambiguous assignments (shared sequence among species) the Lowest Common Ancestor (LCA) approach was applied.

The 16S_DB used for assignments was generated by downloading from GenBank (March 2021) 16S reference sequences belonging to all the Eukaryota taxa and then merged with the latest version (July 2020) of the Ribosomal Database Project containing Bacterial 16S references sequences (Cole et al. 2014). The database (fasta file and corresponding taxonomy) is available on request.

2.4 Data analyses

The R package "rfishbase" was used to interface with the FishBase database (www.fishbase.org) in order to access the available information on over 33,000 fish species related to their taxonomy, morphology, biology, ecology, life history traits and so on. We checked synonyms, extracted taxonomy and, in the cases of assignments at species level, we extracted and associated all the information available including conservation status of species, environmental information as well as information about method of fishery, the economic importance of the species and its vulnerability to fishery (https://www.fishbase.de/manual/english/fishbasethe_species_table.htm). Among this information, four categories were used to assess ecological and economic aspects related to the different samples and products: habitat (adapted from Holthus and Maragos, 1995), importance, price category (Sumaila et al. 2007) and vulnerability (Cheung et al. 2005). In order to highlight the ecological habitus of different taxa identified in different samples, they were classified as pelagic, benthopelagic or demersal. Pelagic encompasses all the taxa occurring mainly in the water column, not feeding on benthic organisms and categorized in FishBase as pelagic-neritic, pelagic-oceanic and bathypelagic. Benthopelagic encompasses all taxa living and feeding on or near the bottom, as well as in the water column, whereas demersal includes all taxa living and/or feeding on or near the bottom. This latter category encompasses species categorized in FishBase as bathydemersal, demersal and benthic. For all species categorized in FishBase as reef-associated, a specific check on the species' habits was carried out in order to classify the species into one of the other categories. Finally, where it was not possible to associate any category, not determined (NA) was assigned.

Concerning the importance of the taxa identified, these are classified as highly commercial, medium commercial, minor commercial, no commercial interest, according to the commercial importance and use of the species in fisheries. In all the cases where it was not possible to associate any category, not determined was assigned. According to price category based on the ex-vessel fish price database (Sumaila et al. 2007), taxa were associated to the following categories: very high, high, medium, low and not determined. Finally, all the taxa recorded in different products were classified as high-very high, high, moderate-high-moderate, low-moderate and low class, according to a vulnerability-to-fishery score (Cheung et al. 2005). When it was not possible to obtain a score, the taxon was classified as not determined (NA). In the case of ambiguous taxonomic assignments, a mean vulnerability score was calculated.

Alpha diversity of the mixtures of samples was explored using several descriptors generated using the diverse R package (Guevara et al., 2016): i) richness (R) number of observed taxa in sample; ii) Berger-Parker dominance index (BP dominance) i.e. the proportion of the most abundant taxa compared with the total abundance of taxa in the sample; values for this index range from 0 to 1, in which 0 means no dominance and 1 means total dominance; iii) Pielou evenness index (E) measures if the taxa in the sample had similar proportions; values range from 0-1: samples scoring 1 have equal proportions of species in the sample – samples scoring 0 have very unequal proportions of taxa). For multivariate analysis, Non-metric multiDimensional Scaling (NMDS) based on the Bray-Curtis dissimilarity matrix was generated using the R package vegan version 2.5-5 (Oksanen et al., 2019). Plots were generated using R package ggplot2 (Wickham et al., 2016).

2.5 Mislabelling assessment

Information reported in labels was evaluated based on the current Regulation (EC) No. 1169/2011. Molecular identifications of each sample were compared with the corresponding

187 ingredients according to the label, and non-compliance was evaluated by applying the following
188 criteria: 1) labels did not report the precise term ‘fish’ among the ingredients; 2) the voluntarily
189 declared species did not correspond to that detected by the molecular analysis; 3) detection of molluscs
190 not declared on the label.

191 Molecular identifications were also crossed with the Decree of the Italian Ministry of Agricultural,
192 Food and Forestry Policies (MiPAAF) dated 22th September 2017, in order to evaluate the presence of
193 taxa not included in the list.

194

195 **3. Results**

196 The Illumina sequencing of the 20 samples generated a total of 4,785,851 raw reads and filtering
197 reduced the dataset to 3,819,087 reads. After the removal of OTUs with total abundance ≤ 50 and
198 taxonomic assignment, the final curated dataset included 2,817,061 reads corresponding to 120
199 different marine taxa. The highest number of taxa were found in surimi 102/120 (85%), while the
200 number of taxa contained in breaded and burger amounted to 23/120 (19%) and 36/120 (30%),
201 respectively (Supplementary Material Table1 and Supplementary Table2).

202 *3.1 Taxonomic overview*

203 Most of the taxa (112) found matched with reference sequences with a level of similarity of
204 between 100-98%. Moreover, of the 120 taxa, 73 of them (60%) were unambiguously assigned at
205 species level, 36 (30%) at genus and 11 (10%) at higher level. Overall, the dataset included 15 orders
206 (13 Actinopterygii, one Cephalopoda and one Malacostraca), 49 families (47 Actinopterygii, one
207 Cephalopoda and one Malacostraca) and 71 genera (69 Actinopterygii, one Cephalopoda and one
208 Malacostraca) (Supplementary Table2). At Class level, Actinopterygii were the only or dominant
209 ingredient in all samples with the exception of Burger S26. Summary of distributions in samples at

Order level (Fig. 1) showed Gadiformes dominant in breaded but present in all product types, whereas Clupeiformes were exclusively found in surimi (only exception breaded S13). Salmoniformes were dominant in one burger sample while Perciformes were absent in breaded products.

At genus level, the dataset was dominated by ten genera, nine belonging to the class of Actinopterygii and one to the class of Cephalopoda, representing about 97% of total reads (Fig. 2). The *Gadus* genus represented 45% of total reads with three species (*Gadus chalcogrammus* 41.36%, *Gadus morhua* 3.88% and *Gadus macrocephalus* 0.01%). It was the main component in breaded samples S8, S21, S30 and S33 which made up 99.9% of the reads, but high percentages were also reported within surimi S50 (90%), surimi 41 (59%) and burger S24 (38%). *Thunnus* and *Sardinella* genera showed the same total abundance (~15% of total reads). Species level within *Thunnus* genus was not achieved since the molecular region is shared by four species (*Thunnus tonggol*/*Thunnus maccoyii*/*Thunnus thynnus thynnus*/*Thunnus albacares*) and distribution in the samples highlighted the genus in burgers S12 and S25, in which *Thunnus* was the main component (86% and 95% of reads in sample, respectively). *Sardinella* was exclusively found in surimi with abundance in samples ranging between 94-0.5% and included seven different taxa, three of them assigned at species level (*Sardinella aurita*, *Sardinella albella* and *Sardinella gibbosa*). The fourth genus was *Sepia* (7% of total dataset), belonging to Decapoda, with two low-abundance taxa assigned at species level (*Sepia hierredda* 0.6% and *Sepia pharaonic* 0.01%) plus one more abundant taxa with ambiguous assignment to either *Sepia recurvirostra* or *Sepia madokai* (6.7%). Distribution of *Sepia* was limited to Burger samples, with the *Sepia recurvirostra*/*madokai* complex being the main component (66%) in Burger S26, and *Sepia hierredda* reaching 6% in Burger S25. Each of the remaining genera had total abundance lower than 5%: *Xiphias* (3.50%), *Dicentrarchus* (2.71%), *Engraulis* (2.51%), *Pleuronectes* (2.49%), *Oncorhynchus* (2.40%), *Decapterus* (1.3%). However, some of them were important components within a single sample. *Oncorhynchus*, including three species (*Oncorhynchus keta*, *Oncorhynchus*

nerka and *Oncorhynchus mykiss*), was the dominant component (98%) in burger sample S27 while *Pleuronectes platessa* represented 83% in breaded S13. *Decapterus macarellus* and *Decapterus maruadsi* were dominant (68%) in surimi S32 and *Engraulis ringens* made up 29% and 24% of the components in surimi S36 and S41, respectively. Further, *Xiphias gladius* was the main component in Burger S26 in which it made up 34% of the reads. The other 72 taxa had total read abundances lower than 1%, and were exclusively found in surimi with the exception of four taxa (*Platichthys* spp., Pleuronectidae, *Limanda punctatissima*, *Psettichthys melanostictus*) exclusively present in breaded S13, two taxa (*Lota lota*, *Sparus aurata*) only found in burger S24 and *Melanogrammus aeglefinus* exclusive of Burger S12. Among these 72 taxa, some of them were important components within a single sample: *Merluccius gayi* (33% in surimi S20) and *Merluccius productus* (9% in surimi S51); *Nematalosa japonica* represented ~50% of surimi S50 and seven different taxa within *Nemipterus* genus formed ~29% of surimi S20. *Portunus sanguinolentus* was the only species belonging to Crustacea found in the dataset and its occurrence was restricted to surimi S32 in which its abundance was equal to 0.14%.

The presence of non-marine taxa was traced to Burger S26, S27 containing Bos taurus (0.07% and 0.12%, respectively) and in burger S26, breaded S33 and S8 containing *Sus scrofa* (0.05%, 0.03%, 0.06%, respectively) (Supplementary Table 3).

3.2 Diversity of products

Box plots of quantitative descriptors of alpha diversity showed signals related to the different types of products (Fig. 3). Details of indices in each sample and statistics are reported in Supplementary Table 5 and Supplementary Table 6. In general, richness (number of observed taxa) showed lower and homogeneous values in burger and breaded than in surimi. The Berger-Parker Dominance and the complementary Evenness indices showed breaded as the products dominated by

one ingredient in contrast with surimi, highlighting more equal mixtures. The NMDS plot in Figure 4 included samples and taxa. Samples were grouped by product type and, in NMDS1, values below or above zero split surimi from burger, with breaded in between. The NMDS panels in Figure 5 included only taxa on the same graphical space of samples with taxa being coloured according to the associated descriptors (*Habitus*, *Commercial Importance*, *Price Category* and *Vulnerability to fishery*) summarized in Supplementary Table 1. Surimi products are characterized by pelagic taxa with medium or minor commercial interest and low or low-moderate Vulnerability to fishery while breaded and Burger show a prevalence of demersal taxa with medium or high commercial interest and moderate and high Vulnerability to fishery. In terms of price category, burgers show a clear prevalence of taxa belonging to high and very high categories, in contrast with surimi and breaded products.

3.3 Mislabelling assessment

All samples reported the exact term "fish" and the mandatory information of geographical origin in accordance with Regulation (EC) 1169/2011. Fifteen samples were produced in five different European countries (Germany, Italy, Lithuania, Spain and Poland), four samples were produced in China and one in Vietnam. In most of the cases, labels also contained voluntary declaration of species with breaded 2/5 (40%), burger 5/5 (100%) surimi 6/10 (60%). Overall, using the criteria defined above, analyses of declared ingredients in labels and metabarcoding data revealed mislabelling in 9/20 (45%) samples (Table 1). Based on the kind of products, non-compliances were differently distributed: 2/5 (40%) samples in breaded, 4/5 (80%) in burger and 3/10 (30%) in surimi. Based on the kind of frauds, in most of the cases samples failed to match the voluntarily declared scientific names in the ingredient list (7/9 samples 77.8%) while in 2/9 (22.2%) samples we detected the unreported presence of molluscs. Details of substitutions showed that, in breaded S8, *Merluccius gayi* was substituted with *Gadus chalcogrammus*, while in Burger *Sepiella japonica* was substituted with *Sepia* spp. (S25, S26),

and *Salmo salar* was substituted with *Oncorhynchus* spp. (S27). In the three mislabelled surimi samples, *Nemipterus nemipterus* was substituted with a mix of species dominated by *Sardinella aurita* (85% in S18), in S51 *Merluccius gayi* was substituted with *Merluccius productus* and in S50 *Nemipterus virgatus* was substituted with a mix of species that failed to include the one declared on the label (*Nematalosa japonica* 49.17%, *Lepidotrigla grandis* 18.68%, *Trachurus* spp. 13.71%, *Chelidonichthys kumu*/*spinosus* 12.29%, *Auxis rochei* 3.07%, *Saurida undosquamis* 2.6%). The cases of unreported presence of molluscs showed traces (< 0.05%) of *Sepia* spp. in one Burger (S12) and one breaded (S13) product.

Molecular identifications also highlighted the presence of 21 taxa (17.5% of total taxa) not included in the MiPAAF 2017 Decree, most of which (19) were found to be widespread in seven of the surimi samples. The other two taxa were found in Burger products, with *Sepia hierredda* in S12 and S25 and *Psettichthys melanostictus* in S13 (Supplementary Table1). Yet, with a similarity of 100% we identified traces (< 0.1%) of a Tetraodontidae taxa not assigned at species level (shared sequence *Takifugu fasciatus*/*Lagocephalus wheeleri*/*Lagocephalus lunaris*) in three surimi samples (S17, S18, S34).

4. Discussion

The general pattern obtained in our study reveal the complex structure and heterogeneity of the ingredients used in the mixed fishery products, with several implications and consequences for consumers, also highlighting the weakness of current labelling legislation and the need for regular monitoring. High level of detail is now available thanks to the new generation of sequencers allowing the simultaneous sequencing of multiple DNA molecules, that are a powerful tool for tracing ingredients in complex matrices that had not been traceable in the past. In this study, the high

304 sequencing depth (almost five million raw reads) and the universality of primers used (15 orders, 49
305 families and 71 genera), reveal an unexpectedly high complexity within the twenty mixtures, that had
306 not previously been reported. Despite the short length (about 200 bp), the mitochondrial 16S rRNA
307 mini-barcode fragment used allowed assignation at species level in 60% of the taxa, with failures
308 including genera such as *Thunnus*, *Scomber* and *Sardinella* for which the difficulty of discrimination to
309 species level even using the full DNA barcoding fragments has previously been reported (Catanese et
310 al., 2010; Chan et al., 2019; Viñas et al., 2009). Beside marine taxa, the good taxonomic coverage
311 provided by the primers includes the capability to trace the presence of ingredients from terrestrial
312 mammals (bovine and swine) that is a critical point for consumer protection based on ethical or
313 religious issues (Xing et al., 2019). However, tests in silico (data not shown) highlighted the difficulty
314 for these primers to amplify chicken, making it unable to trace components such as *eggs* that, even
315 though they were reported on the labels of some samples, were lost in our taxonomic assessment.
316 Although processed fish products shared the labelling system governed by EU Regulation
317 1169/2011, our study shows evidence of strong heterogeneity in structure and composition of products
318 included under the umbrella term of *processed product*. As a general trend, number of species and their
319 evenness are higher in surimi, but this is not a hard and fast rule, so the number of species in the
320 mixtures could differ based on brand and production country (Galal-Khallaf et al., 2016). Among the
321 27 families of fish included in the table generated by Giusti et al. (2017b) for surimi, only 15 are shared
322 with the 46 fish families found as ingredients in our surimi samples, highlighting the huge marine
323 biodiversity included in this type of product, and raising concern for their origin and negative impact
324 on ecosystems. For breaded and burger, our data represent the first overview and suggest a more
325 limited number of ingredients in their matrices, especially in burger.
326 Overall, this study underlines a 45% mislabelling rate, with non-compliance mainly due to
327 incongruences in the voluntary declared scientific names in the ingredient list (78%), and 22% of

328 samples containing the unreported presence of molluscs. Surprisingly, the highest non-compliance
329 value was found in Burger products (80%), a less processed product mainly obtained from fish fillets or
330 from fish pulp and containing lower number of taxa, whereas lower mislabelling rates were reported in
331 breaded fish, (40%) and surimi (30%), usually considered highly processed products. However, it
332 should be taken into account that the mislabelling was the result of a voluntary declaration by
333 manufacturers on the product labels, presumably as a marketing strategy to target an audience of more
334 aware consumers for whom additional detailed information on the product labels plays an important
335 role in a positive perception of safe and healthy food. By contrast, breaded products undergo a higher
336 degree of processing, targeting consumers (like teenagers or young adults) who do not like the taste of
337 fish and are less interested in the details of the fish used as ingredients. In this case, the strategy of
338 omitting species names enables industries to use any species as an ingredient, thus lowering the
339 potential risk of mislabelling and reducing the danger of losing clients.
340 Therefore, although our inferences need to be confirmed with a higher number of samples, the
341 results suggest the labels of processed products, both those with a voluntary declaration or worth
342 omission of species names, cannot be used by consumers as a proxy of product safety, but are probably
343 part of the marketing policies applied in different products by manufacturers.
344 The current legislation governing processed seafood is EU Regulation 1169/2011, the Food
345 Information to Consumer Regulation (FIC), which aims to achieve a high level of consumer health
346 protection and to guarantee their right to information, so that they can make informed choices and
347 prevent practices that risk misleading the consumer. EU Regulation 1169/2011 does not include a list
348 of mandatory information but simply provides for the designation of all fish species in which fish is an
349 ingredient in a food and, provided that the name and presentation of this food do not refer to a specific
350 fish species, it can be designated just with the name of a category such as "fish" rather than with a
351 specific name (Di Pinto et al., 2016). The bland information required in labels is probably related to the

fact that it used to be impossible to verify which species were used as ingredients in the mixtures. In this regard, a future EU consumer protection Regulation could, instead, make a species list mandatory, given that advanced molecular approaches are now available which can successfully trace species in complex matrices. Therefore, the NGS technology, which can reveal deliberate substitution/addition of non-authentic food products, protects and promotes fish products against fraud and species substitution and guarantees accurate food labelling and the legal base requirements, which are so crucial for an innovative food-safety management system.

The mislabeling rates found in our dataset are in line with other studies, showing that surimi contained a 25% presence of molluscs not indicated on the label (Giusti et. al 2017b) and similar overall mislabelling rates within highly processed cod products (Carvalho et. al 2017) and surimi (Giusti et. al 2017b). However, a clear comparison with the mislabelling rates reported in other studies using NGS in processed seafood is somewhat difficult due to the low number of samples usually analysed with this approach.

Only one substitution could be directly linked with economic gain, involving a species from a higher price category substituted with a lower one, as in the case of *Nemipterus nemipterus* substituted with a mix of species dominated by *Sardinella aurita*. Less clear is the substitution of *Nemipterus virgatus* with a mixture including species belonging to different price categories, whereas all the other substitutions involve species with similar commercial value or price, as in the cases of *Merluccius gayi* with *Gadus chalcogrammus*, *Sepiella japonica* with *Sepia* spp., *Merluccius gayi* with *Merluccius productus*, and *Salmo salar* salmon from the Atlantic substituted with *Oncorhynchus* spp. (Pacific salmon).

In addition to mislabelling, samples include low-frequency taxa not reported in the list of ingredients, or even not included in the latest MiPAAF Decree dating back to 2017. These taxa could be related to technological processes for mixing fish products, such as the addition of gelatines to

improve texture and rheological characteristics (i.e. viscosity, texture and stability). Actually, only some taxa found in our samples (*Trachurus* spp., *Lutjanus*, *Priacanthus*, and *Oreochromis niloticus*) are known to be used as an alternative to mammalian gelatine by food industries to overcome ethical and religious limitations (Li et al., 2018; Liu et al., 2007). However, in one burger and two surimi samples, we also found swine DNA, probably added as gelatine obtained from porcine skin. This poses important religious and ethical consequences for consumers with cultural and religious dietary requirements such as Muslim (Halaal) and Jewish (Kashrut) believers who are not allowed to consume pork (Hernández-Briones et al., 2009). By contrast, the presence of bovine DNA in two samples corresponds to the 'cheese' reported on the label.

The low-abundance taxa found involved health implications for consumers due to the traces of Tetraodontidae, containing Tetrodotoxin, a neurotoxin with a paralyzing effect, found in three surimi samples. These taxa, banned from the Europe market (Regulation EC No 854/2004), are currently sold in Asian countries, such as Taiwan or Japan, and mislabelling involving *Lagocephalus* spp. has been reported at a Chinese market in substitution of products commercialized under the generic name Cod (Xiong et al., 2016) or angelfish (Li et al., 2016), but also in products sold in Europe (Armani et al. 2015; Giusti at al. 2018). However, in our study, the samples have different geographic origins (Italy, China, Lithuania) highlighting how international trading and expansion of the supply chain affect traceability and can lead to increases in fraudulent behaviours, even involving species associated with health risks to European consumers (Sameera & Ramachandran, 2016). Moreover, even though information on allergens must be provided to consumers under EU Regulation 1169/2011, in our data, two samples reveal the unreported presence of molluscs among the low frequency taxa, confirming the general concern regarding allergenic taxa. Seafood allergies are the most common food allergies and adverse health issues can occur even in the presence of small percentages of molluscs, and so, in the absence of a clear safe allergen threshold, this information is mandatory under European labeling

400 legislation. By contrast, traces of the potential allergenic Crustacea *Portunus sanguinolentus*, found in
401 a surimi sample, had been correctly reported on the label.

402 Patterns of the substitutions and the high number of low-abundance taxa found in samples, even
403 thought they may be involuntary cases of substitution linked to the presence of morphologically similar
404 traits and / or to the lack of specific training of fisheries operators, confirm that processed seafood is an
405 important channel used by manufacturers to recycle discards and fish waste, generated by the filleting
406 process, including viscera, heads, fins, skins, scales, and bones (Palmeira at al 2016; Rustad et al.,
407 2011). The FAO 2018 report revealed that 35% of global catches are wasted, becoming by-catch or
408 discarded from trawlers where fish are thrown back dead because they are too small or an unwanted
409 species. Of course, the possibility to use such material to produce food for human consumption would
410 help boost profits for the fishing industry. In terms of health, the discarded species are often small, with
411 a lower risk of mercury accumulation than some fish species of higher commercial value (Simeone &
412 Scarpato, 2014), whereas more studies are needed to clarify the impact of microplastics from discarded
413 tissues (Hantoro et al. 2019). Moreover, also other aspects need to be monitored because fish discards,
414 even though they have low or no commercial value, can include rare, endangered or protected as well
415 as toxic species with untraceable impacts on marine ecosystems and consumer health.

416 The complexity of the components in seafood products makes assessing sustainability far from
417 easy, given that to fisheries and to ecosystems the term sustainability takes on two rather different
418 meanings, in the former referring to stocks while the second focuses on the trophic web as a whole. In
419 assessments of whole specimens or fillets, the inclusion of taxa listed in the Convention on
420 International Trade in Endangered Species (CITES) or Red List of Threatened Species of the
421 International Union for Conservation of Nature (IUCN, www.iucnredlist.org) can be directly linked
422 with the sample, and it is usually an indication of lower sustainability. Furthermore, commercially
423 important species can also be on the IUCN Red List or even considered to be at different levels of risk

on the various IUCN lists at global or local scale, as is the case for *Prionace glauca* (considered NT at global scale but CR in the Mediterranean Sea). Moreover, in processed products, we cannot identify the real source of the DNA (wild or farmed specimens, discards or fish waste) and no such information is mandatory on the label. This makes it difficult to perform a correct and complete evaluation because, even the presence of DNA of a threatened species could be due to the use of discarded tissues, and thus considered in some ways more sustainable.

Overall, the profiles generated suggest that surimi is a more sustainable product than breaded and burger and characterized by taxa with medium or lesser commercial interest and low or low-moderate Vulnerability to fishery. The presence of small pelagic forage fishes, especially belonging to the Sardinella genus, is in line with a previous study which evaluated ecological sustainability in surimi using the trophic level of ten taxa generated by a PCR-cloning approach (Galal-Khallaf et al., 2016). Indeed, the price category profile of taxa used as ingredients confirms surimi to belong to a market segment targeting and attracting consumers for their competitive price. Breaded and burger profiles show a prevalence of demersal taxa with medium or high commercial interest and moderate and high Vulnerability to fishery. Based on the price profile, burger can be considered a luxury processed seafood product, targeting more aware and carefully selected consumers who require a more accurate labelling system, which should include information on the wild or farmed origin of taxa, since some of the species in our Burger samples were probably farmed (*Dicentrarchus labrax*, *Oncorhynchus* spp. and *Thunnus* spp). This information, as well the indication of the inclusion of fish waste, need to be part of the future labelling system for all processed products so as to make it easier to assess its sustainability as well as affording better protection for consumers.

5. Conclusions

Our study shows that NGS metabarcoding is a robust, powerful and reliable methodology able to analyse highly processed multi-species fish products, which make up a growing market segment, due both to the modern lifestyles of consumers and to the profit expectations of the fish industry. Beyond the mislabelling rates, results from samples raise concern due to the high number of taxa contained that are unknowingly eaten by consumers and, in the absence of controls, fished from the oceans. The recycling of discards or fish waste is positive, decreasing the environmental impact of fisheries, but shady areas and potential risks are evident. In this context, the current labelling system is a very weak instrument both when it comes to consumer safety and for conscious choices and may form part of the manufacturers' marketing strategies. An adequate revision of national and international food safety legislations would be a huge challenge for policymakers, but metabarcoding assessments could help both by providing useful information and acting as the operational tool for inspective monitoring.

Acknowledgements

This work was supported by Istituto Zooprofilattico Sperimentale della Puglia e della Basilicata (IZS PB 02/18 RC). A.M was supported by the EU Project PON-AIM (National Operative Program, Attraction and International Mobility), nr. AIM1839894 - Activity 2 - Line 1 funded by Ministry of Education, University and Research (MIUR). R.P. was supported by the Research for Innovation (REFIN) (n.r. 286655ED) initiative co-funded by EU and POR Puglia 2014-2020, Asse Prioritario OT X action 10.4.

Figure captions

Figure 1. Relative read abundance at Order level in samples of breaded, burger and surimi.

Figure 2. Bubble plot. Proportion of taxa reads found in samples at genus level.

Figure 3. Alpha diversity indices. a) Richness; b) Berger Parker dominance; and c) Pielou evenness.

Figure 4. NMDS ordination based on Bray-Curtis matrix. The plot includes the ID samples (blue: surimi; green: breaded; pink: burger) and the ID of taxa in red. Stress 0.175.

Figure 5. NMDS based on Bray-Curtis matrix where only taxa were plotted. Taxa were coloured based on the associated descriptors (Habitus, Commercial importance, Price category and Vulnerability to fishery).

References

Altschul, S.F., Gish, W., Miller, W., Myers, E. W., & Lipman, D.J. (1990). Basic local alignments earch tool. *Journal of Molecular Biology*, *215*, 403–410. https://doi.org/10.1016/S0022 - 2836(05)80360-2.

Anthoons, B., Karamichali, I., Schrøder-Nielsen, A., Drouzas, A. D., de Boer, H., & Madesis, P. (2021). Metabarcoding reveals low fidelity and presence of toxic species in short chain-of-commercialization of herbal products. *Journal of Food Composition and Analysis*, *97*, 103767. https://doi.org/10.1016/j.jfca.2020.103767

Armani, A., Guardone, La Castellana, R., Gianfaldoni, D., Guidi, A., Castigliego, L. (2015). DNA barcoding reveals commercial and health issues in ethnic seafood sold on the Italian market. *Food Control*, *55*, 206-214.

Borgogno, M., Husein, Y., Secci, G., Masi, S., & Parisi, G. (2017). Technological and nutritional advantages of mechanical separation process applied to three European aquacultured species. *LWT - Food Science and Technology*, *84*, 298-305. https://doi.org/10.1016/j.lwt.2017.05.068.

Bruno, A., Sandionigi, A., Agostinetto, G., Bernabovi, L., Frigerio, J., Casiraghi, M., & Labra, M. (2019). Food tracking perspective: DNA Metabarcoding to identify plant composition in complex and processed food products. *Genes*,*10*, 248. doi:10.3390/genes10030248.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. BMC Bioinformatics, 10, 421. https://doi.org/10.1186/1471-2105-10-421.

Carvalho, D. C., Palhares, R. M., Drummond, M. G., & Gadanho, M. (2017). Food metagenomics: Next generation sequencing identifies species mixtures and mislabeling within highly processed cod products. *Food Control*, *80*, 183–186. https://doi.org/10.1016/j.foodcont.2017.04.049.

Catanese, G., Manchado, M., & Infante, C. (2010). Evolutionary relatedness of mackerels of the genus *Scomber* based on complete mitochondrial genomes: Strong support to the recognition of Atlantic *Scomber colias* and Pacific *Scomber japonicus* as distinct species. *Gene*, 35-43.

Chan, A.F.O., Luczon, A.U., Fontanilla, I.K.C., Ong, P.S., Santos, M.D., Willette, D.A., & Quilang, J.P. (2019). DNA barcoding cannot discriminate between *Sardinella tawilis* and *S. hualiensis* (Clupeiformes: Clupeidae). *Mitochondrial DNA Part B*, *4*, 2499-2503.

Chapela, M. J., Sotelo, C. G., Calo-Mata, P., Pérez-Martín, R. I., Rehbein, H., Hold, G. L., Quinteiro, J., Rey-Méndez, M., Rosa, C., & Santos, A. T. (2002). Identification of Cephalopod Species (Ommastrephidae and Loliginidae) in Seafood Products by Forensically Informative Nucleotide Sequencing (FINS). *Journal of Food Science*, *67*, 1672–1676. https://doi.org/10.1111/j.1365-2621.2002.tb08703.x.

Cheung, W.W.L., Pitcher, T.J., & Pauly, D. (2005). A fuzzy logic expert system to estimate intrinsic extinction vulnerabilities of marine fishes to fishing. *Biological Conservation*, *124*, 97–111.

Clark, L. F. (2015). The current status of DNA barcoding technology for species identification in fish value chains. *Food Policy*, *54*, 85–94. https://doi.org/10.1016/j.foodpol.2015.05.005.

Cole, J.R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A., Kuske, C.R., & Tiedje, J. M (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Research*, *42(Database issue)*, D633-D642. doi: 10.1093/nar/gkt1244.

Cottenet, G., Blancpain, C., Chuah, P. F., & Cavin, C. (2020). Evaluation and application of a next generation sequencing approach for meat species identification. *Food Control*, *110*, 107003. https://doi.org/10.1016/j.foodcont.2019.107003.

Di Pinto, A., Marchetti, P., Mottola, A., Bozzo, G., Bonerba, E., Ceci, E., Bottaro, M., & Tantillo, G. (2015). Species identification in fish fillet products using DNA barcoding. *Fisheries Research*, *170*, 9–13.

Di Pinto, A., Mottola, A., Marchetti, P., Bottaro, M., Terio, V., Bozzo, G., Bonerba, E., Ceci, E., & Tantillo, G. (2016). Packaged frozen fishery products: species identification, mislabeling occurrence and legislative implications. *Food Chemistry*, *194*, 279–283.

Donlan, C. J., & Luque, G. M. (2019). Exploring the causes of seafood fraud: A meta-analysis on mislabeling and price. *Marine Policy*, *100*, 258–264. https://doi.org/10.1016/j.marpol.2018.11.022.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, *27*,2194–200.

Franco, C.M., Ambrosio, R.L., Cepeda, A., Anastasio, A. (2021). Fish intended for human consumption: from DNA barcoding to a next-generation sequencing (NGS)-based approach. *Current Opinion in Food Science*, *42*, 86-92. https://doi.org/10.1016/j.cofs.2021.05.005.

ha formattato: Italiano (Italia), Evidenziato
ha formattato: Evidenziato
ha formattato: Italiano (Italia), Evidenziato
ha formattato: Evidenziato
ha formattato: Italiano (Italia), Evidenziato
ha formattato: Evidenziato
ha formattato: Italiano (Italia), Evidenziato
ha formattato: Evidenziato
ha formattato: Tipo di carattere: Corsivo, Evidenziato
ha formattato: Evidenziato
ha formattato: Tipo di carattere: Corsivo, Evidenziato
ha formattato: Evidenziato

Galal-Khallaf, A., Ardura, A., Borrell, Y.L., & Garcia-Vazquez, E. (2016). Towards more sustainable surimi? PCR-cloning approach for DNA barcoding reveals the use of species of low trophic level and aquaculture in Asian surimi. *Food Control, 61*, 62e69.

Giusti, A., Armani, A., & Sotelo, C. G. (2017b). Advances in the analysis of complex food matrices: Species identification in surimi-based products using Next Generation Sequencing technologies. *PLoS ONE*, *12*, 1–18. https://doi.org/10.1371/journal.pone.0185586.

Giusti, A., Ricci, E., Guarducci, M., Gasperetti, L., Davidovich, N., Guidi, A., & Armani, A. (2018). Emerging risks in the European seafood chain: Molecular identification of toxic *Lagocephalus* spp. in fresh and processed products. *Food Control, 91*, 311-320.

Giusti, A., Tinacci, L., Sotelo, C. G., Marchetti, M., Guidi, A., Zheng, W., & Armani, A. (2017a). Seafood Identification in Multispecies Products: Assessment of 16SrRNA, *cytb*, and COI Universal Primers' Efficiency as a Preliminary Analytical Step for Setting up Metabarcoding Next-Generation Sequencing Techniques. *Journal of Agricultural and Food Chemistry*, *65*, 2902–2912. https://doi.org/10.1021/acs.jafc.6b05802.

Guevara, M.R., Hartmann, D., & Mendoza, M. (2016). Diverse: an R Package to AnalyzeDiversity in Complex Systems. *The R Journal, 8/2*, 60-78.

Hantoro, I., Löhr, A.J., Van Belleghem, F.G.A.J, Widianarko, B., & Ragas, A.M.J. (2019). Microplastics in coastal areas and seafood: implications for food safety. *Food Additives & Contaminants: Part A, 36*, 674-711.

Haynes, E., Jimenez, E., Pardo, M. A., & Helyar, S. J. (2019). The future of NGS (Next Generation Sequencing) analysis in testing food authenticity. *Food Control, 101*, 134–143. https://doi.org/10.1016/j.foodcont.2019.02.010.

Hernández-Briones, A., Velázquez, G., Vázquez, M., Ramírez, J.A. (2009). Effects of adding fish gelatin on Alaska pollock surimi gels. *Food Hydrocolloids*, *23*, 2446-2449. https://doi.org/10.1016/j.foodhyd.2009.07.002.

Ho, J.K.I., Puniamoorthy, J., Srivathsan, A., & Meier, R. (2020). MinION sequencing of seafood in Singapore reveals creatively labelled flatfishes, confused roe, pig DNA in squid balls, and phantom crustaceans. *Food Control*, *112*, 107144. https://doi.org/10.1016/j.foodcont.2020.107144.

Holthus, P.F., & Maragos, J.E. (1995). Marine ecosystem classification for the tropical island Pacific. In: Marago, J.E., Peterson, M.N.A., Eldredge, L.G., Bardach, J.E., Takeuchi, H.F. (Eds.), Marine and Coastal Biodiversity in the Tropical Island Pacific Region, East-West Center, Honolulu, pp. 239–278.

Husein Y. (2019) *Application of the Mechanical Separation Process in Different Fish Species for the Development of a New Product Based on Fish*. PhD Thesis, University of Firenze, Italy.

Kappel, K., Haase, I., Käppel, C., Sotelo, C. G., & Schröder, U. (2017). Species identification in mixed tuna samples with next-generation sequencing targeting two short cytochrome b gene fragments. *Food Chemistry*, *234*, 212–219. https://doi.org/10.1016/j.foodchem.2017.04.178

Kozich, J.J., Westcott, S.L., Baxter, N.T. Highlander, S.K., Schlosset, P.D. (2013). Development of a dualindex sequencing strategy and curation pipeline for analysing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied* and *Environmental Microbiology*, *79*, 5112–5120.

Li, Q., Gui, P., Huang, Z., Feng, L., Luo, Y. (2018) Effect of transglutaminase on quality and gel properties of pork and fish mince mixtures. *The Journal of Texture Studies*, *49*, 56-64. doi: 10.1111/jtxs.12281. Epub 2017 Jun 28. PMID: 28581064

578 Li, Z., Chen, Z., Gao, J., Wang, L., Chen, X., & Zhao, Y. (2016). The complete mitochondrial
579 genome of the altum angelfish *Pterophyllum altum* (Pellegrin, 1903). *Mitochondrial DNA. Part A,*
580 *DNA mapping, sequencing, and analysis*, *27*, 3705-3706. doi: 10.3109/19401736.2015.1079866.
581 Liu, R., Zhao, S.-m. Xiong, S.-b. Xie, B.-j., Liu, H.-m. (2007). Studies on Fish and Pork Paste
582 Gelation by Dynamic Rheology and Circular Dichroism. *Journal of Food Science*, *72*, E399-E403,
583 https://doi.org/10.1111/j.1750-3841.2007.00470.x
584 Luque, G. M., & Donlan, C. J. (2019). The characterization of seafood mislabeling: A global meta-
585 analysis. *Biological Conservation*, *236*, 556–570. https://doi.org/10.1016/j.biocon.2019.04.006
586 Marchetti, P., Mottola, A., Piredda, R., Ciccarese, G., & Di Pinto, A. (2020). Determining the
587 Authenticity of Shark Meat Products by DNA Sequencing. *Foods*, *9*, 1194.
588 https://doi.org/10.3390/foods9091194.
589 Mottola, A., Marchetti, P., Ciccarese, G., Terio, V., Sinisi, C., & Di Pinto A. (2020). *Arcobacter*
590 species detection in Italian composite foods. *LWT - Food Science and Technology*, *134*, 110161.
591 Noh, E. S., Lee, M.-N., Kim, E. M., Nam, B.-H., Noh, J. K., Park, J. Y., Kim, K.-H., & Kang, J.-H.
592 (2021). Discrimination of raw material species in mixed seafood products (surimi) using the next
593 generation sequencing method. *Food Bioscience*, *41*, 100786.
594 https://doi.org/10.1016/j.fbio.2020.100786.
595 Okazaki, E., & Kimura, I. (2014). *Frozen surimi and surimi-based products*. In I. S. Boziaris (Ed.),
596 Seafood processing technology, quality and safety (pp. 209e235). UK: John Wiley & Sons, Ltd.
597 Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R.,
598 Ottan, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., & Szoecs, E. (2019). Community Ecology
599 Package. https://cran.r-project.org.

Paci, F., Danza, A., Del Nobile, M.A., & Conte, A. (2018). Consumer acceptance and willingness to pay for a fresh fish-burger: A choice experiment, *Journal of Cleaner Production*, *172*, 3128-3137. https://doi.org/10.1016/j.jclepro.2017.11.095.

Palmeira, K. R., Másrico, E. T., Guerra, M. L., Lemos, M., Adam, C., & Conte Junior, C. A. (2016). Ready-to-eat products elaborated with mechanically separated fish meat from waste processing: challenges and chemical quality. *CyTA - Journal of Food*, *14*, 227-238. https://doi:10.1080/19476337.2015.1087050.

Palumbo, F., Scariolo, F., Vannozzi, A., & Barcaccia, G. (2020). NGS-based barcoding with mini-COI gene target is useful for pet food market surveys aimed at mislabelling detection. *Scientific Reports*, *10*, 1–8. https://doi.org/10.1038/s41598-020-74918-9.

Paracchini, V., Petrillo, M., Lievens, A., Kagkli, D. M., & Angers-Loustau, A. (2019). Nuclear DNA barcodes for cod identification in mildly-treated and processed food products. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment*, *36*, 1–14. https://doi.org/10.1080/19440049.2018.1556402.

Ribani, A., Schiavo, G., Utzeri, V. J., Bertolini, F., Geraci, C., Bovo, S., & Fontanesi, L. (2018). Application of next generation semiconductor based sequencing for species identification in dairy products. *Food Chemistry*, *246*, 90–98. https://doi.org/10.1016/j.foodchem.2017.11.006.

Rustad, T., Storrø, I., Slizyte, R. (2011). Possibilities for the utilisation of marine by-products. *International Journal of Food Science & technology*, *46*, 2001-2014.

Sameera, S., & Ramachandran, A. (2016). Chain of custody of fish marketed in cochin city: Safety and quality issues. *International Journal of Management*, *7*, 279–292.

621 Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski,
622 R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn,
623 D.J., Weber, C.F. (2009). Introducing mothur: Open-source, platform-independent, community-
624 supported software for describing and comparing microbial communities. *Applied and Environmental*
625 *Microbiology*, *75*, 7537–7541.

626 Silva, A. (2019). *Use of PCR cloning combined with DNA barcoding to identify fish in a mixed-*
627 *species product*. Master's thesis, Chapman University. https://doi.org/10.36837/chapman.000065.

628 Simeone, M., & Scarpato, D. (2014). Low commercial value fish. How can we increase its
629 consumption? *Agricultural Economics Review*, *15*, 43-59.

630 Sørensen, J. S., Bøknæs, N., Mejlholm, O., & Dalgaard, P. (2020). Superchilling in combination
631 with modified atmosphere packaging resulted in long shelf-life and limited microbial growth in
632 Atlantic cod (*Gadus morhua* L.) from capture-based-aquaculture in Greenland. *Food Microbiology*, *88*,
633 103405. https://doi.org/10.1016/j.fm.2019.103405.

634 Sumaila, U.R., Marsden, A.D., Watson, R., & Pauly, D. (2007). A Global Ex-vessel Fish Price
635 Database: Constructionand Applications. *Journal of Bioeconomics*, 9, 39–51.
636 https://doi.org/10.1007/s10818-007-9015-4.

637 Sun, X., Hong, H., Jia, S., Liu, Y., & Luo, Y. (2020). Effects of phytic acid and lysozyme on
638 microbial composition and quality of grass carp (*Ctenopharyngodon idellus*) fillets stored at 4 °C. *Food*
639 *Microbiology*, *86*, 103313. https://doi.org/10.1016/j.fm.2019.103313.

640 Viñas, J., & Tudela, S. (2009). A validated methodology for genetic identification of tuna species
641 (genus *Thunnus*). *PLOS one*, *4*, e7606.

642 Voorhuijzen-Harink, M. M., Hagelaar, R., van Dijk, J. P., Prins, T. W., Kok, E. J., & Staats, M.
643 (2019). Toward on-site food authentication using nanopore sequencing. *Food Chemistry*: *X*, *2*, 100035.
644 https://doi.org/10.1016/j.fochx.2019.100035.

645 Wickham, H., Chang, W., & Studio, R. (2016). Ggplot2: An Implementation of the Grammar of
646 Graphics. https://CRAN.R-project.org/package=ggplot2.

647 Wirta, H., Abrego, N., Miller, K., Roslin, T., & Vesterinen, E. (2021). DNA traces the origin of
648 honey by identifying plants, bacteria and fungi. *Scientific Reports*, *11*, 1–14.
649 https://doi.org/10.1038/s41598-021-84174-0.

650 Xing, R-R., Wanga, N., Hua, R-R., Zhanga, J-K., Hana, J-X., & Chena, Y. (2019). Application of
651 next generation sequencing for species identification in meat and poultry products: A DNA
652 metabarcoding approach. *Food Control* , *101*, 173–179.
653 https://doi.org/10.1016/j.foodcont.2019.02.034.

654 Xiong, X., Guardone, L., Cornax, M. J., Tinacci, L., Guidi, A., Gianfaldoni, D., & Armani, A.
655 (2016). DNA barcoding reveals substitution of Sablefish (*Anoplopoma fimbria*) with Patagonian and
656 Antarctic Toothfish (*Dissostichus eleginoides* and *Dissostichus mawsoni*) in online market in China:
657 How mislabeling opens door to IUU fishing. *Food Control*, *70*, 380–391.
658 https://doi.org/10.1016/j.foodcont.2016.06.010.

659 Zhang, J., Li, Y., Liu, X., Lei, Y., Regenstein, J. M., & Luo, Y. (2019). Characterization of the
660 microbial composition and quality of lightly salted grass carp (*Ctenopharyngodon idellus*) fillets with
661 vacuum or modified atmosphere packaging. *International Journal of Food Microbiology*, *293*, 87–93.
662 https://doi.org/10.1016/j.ijfoodmicro.2018.12.022.

Zhuang, S., Li, Y., Jia, S., Hong, H., Liu, Y., & Luo, Y. (2019). Effects of pomegranate peel extract on quality and microbiota composition of bighead carp (*Aristichthys nobilis*) fillets during chilled storage. *Food Microbiology*, *82*, 445–454. https://doi.org/10.1016/j.fm.2019.03.019.

Table 1. Description of samples and summary of mislabelling assessment (in bold). The column Ingredients includes potential sources of detectable DNA. Molecular identification includes taxa with abundances >1%. Taxa with lower abundances were included in the following cases: taxa were declared as ingredient, non-marine taxa, allergenic or toxic taxa.

Sample ID	Production site	Type of product	Packaging	Commercial denomination	(Main) Ingredients	Declared species	**Molecular Identification**	**Mislabeling**
S8	Italy	Breaded	Frozen	Cooled breaded fish cutlets	Hake (47%). May contain trace of crustacea, molluscs	*Merluccius gay*	*Gadus chalcogrammus* 99.9%, ***Sus scrofa* 0.06% (*)**	Yes (2)(*)
S13	Germany	Breaded	Frozen	Sea cutlets	Fish (47%)	ND	*Pleuronectes platessa* 82.54%, *Gadus chalcogrammus* 6.9%, *Anchoa nasus* 6.65%, *Pleuronectes* spp. 1%, ***Sepia* spp. 0.01% (3)**	Yes (3)
S21	Poland	Breaded	Frozen	Fish sticks	Fish (56%). May contain crustacea, eggs, milk, molluscs	ND	*Gadus chalcogrammus* 61.01%, *Gadus morhua* 38.94%	No
S30	Germany	Breaded	Frozen	Fish sticks	Minced fish (65%)	*Theragra chalcogramma*	*Gadus chalcogrammus* 99.99%	No
S33	Germany	Breaded	Frozen	Fish nuggets - Funny fish	Minced fish (48%)	ND	*Gadus chalcogrammus* 99.92%, ***Sus scrofa* 0.03% (*)**	No (*)
S17	Italy	Surimi	Frozen	Surimi stiks	Surimi (fish) 45%, egg whites, (contain crustaceans). May contain trace of molluscs	ND	*Sardinella* spp. (1) 73.49%, *Gadus chalcogrammus* 12,24%, *Hirundichthys marginatus* 2.44%, *Sardinella* spp. (3) 2.41%, *Sardinella* spp. (2) 2.01%, *Merluccius productus* 1.23%, *Engraulis ringens* 1.16%	No (**)
S18	China	Surimi	Frozen	Surimi stiks	Surimi (fish-*Nemipterus*-*Nemipterus*) 40%, crab extract, crab flavour, egg whites	*Nemipterus*-*Nemipterus*	*Sardinella aurita* 85.71%, *Decapterus maruadsi* 6.71%, *Sardinella* spp. (1) 2.68%, *Selar crumenophthalmus* 2.43%, *Rastrelliger kanagurta* 1.05%	Yes (2)
S19	China	Surimi	Frozen	Surimi Crab Claw	Surimi (fish: *Chelidonichthys kumu*) 40%, crab claws, crab flavour. Contains fish, egg, crab	*Chelidonichthys kumu*	*Sardinella* spp. (1) 57.39%, *Sardinella aurita* 33.68%, *Carangidae* 1.81, *Sardinella* spp. (3) 1.67, *Sardinella* spp. (2) 1.02%, *Chelidonichthys* spp. 0.02%	No (**)

S20	Vietnam	Surimi	Frozen	Surimi stiks-shrimp flavour	Surimi (fish: *Nemipterus nemurus* and *Nemipterus japonicus*) 45%, egg whites, shrimp flavour	*Nemipterus nemurus* and *Nemipterus japonicus*	*Merluccius gayi* 31.86%, *Parupeneus heptacanthus* 10.40%, *Priacanthus macracanthus* 9.35%, *Nemipterus spp. (1)* 9%, *Pomadasys maculatus* 7.29%, *Nemipterus spp. (2)* 7.04%, *Nemipterus mesoprion* 4.57%, *Nemipterus spp. (3)* 3.06%, *Nemipterus bathybius* 2.52%, *Upeneus* spp. 2.60%, *Pristipomoides multidens* 2.32%, *Pterois volitans* 2.26%, *Iniistius* spp. 1.48%, *Sillago ingenuua* 1.23%	No (**)
S32	China	Surimi	Frozen	Surimi Crab Claw	Surimi (fish: *Chelidonichthys kumu*) 40%, crab claws, crab flavour, Contains fish, egg, crab	*Chelidonichthys kumu*	*Decapterus maruadsi* 37.40%, *Decapterus macarellus* 31.07%, *Gadus chalcogrammus* 5.79%, Lutjanus spp. 5.72%, Carangidae 5.43%, *Nematalosa japonica* 3.45%, Ariomma luridum 1.89%, *Trachurus* spp. 1.52%, *Chelidonichthys* spp. 0.21%, *Portunus sanguinolentus 0.14%*	No
S34	Lithuania	Surimi	In brine	Surimi exotic salad	Surimi 62% (fish:21%), egg whites	ND	*Sardinella* spp. 46.38, *Sardinella albella* 18.42%, *Dussumieria elopsoides* 7.74%, Rastrelliger kanagurta 6.91%, *Carangidae* 4.45%, *Caranx* spp. 3.30%, *Merluccius gayi* 2.43%, *Scomber* spp. 1.71%, *Engraulis ringens* 1.81%, *Sardinella* spp. (2) 1.24%, *Sardinella* spp. (1) 1.16%	No (**)
S36	Lithuania	Surimi	Frozen	Surimi Claw - Crab flavour	Surimi 78% (fish: 33%), white egg, crab, crustacea, fish, molluscs	ND	*Engraulis ringens* 28.47%, *Sardinella* spp. 25.29%, *Sardinella* spp. 23.45%, *Sardinella gibbosa* 8%, *Hirundichthys marginatus* 5.34%, *Gadus chalcogrammus* 3.05%, *Thryssa* spp. 2.36%;	No (**)

S41	Spain	Surimi	Frozen	Frozen surimi stiks	Surimi (fish) 40%, crab fravour, egg yolk, egg whites. May contain milk, celery, molluscs trace	ND	*Gadus chalcogrammus* 59.09%, *Engraulis ringens* 23.65%, *Sardinella* spp. (1) 6.19%, *Sardinella* spp. (2) 6.03%, *Sardinella albella* 3.10%, *Merluccius productus* 1.34%	No (**)
S50	China	Surimi	Frozen	Surimi Claw - Crab flavour	Surimi 36% (fish - *Nemipterus virgatus*), crab flavour, crab claws, Allergens: eggs, shellfish, fish	*Nemipterus virgatus*	*Nematalosa japonica* 49.17%, *Lepidotrigla grandis* 18.68%, *Trachurus* spp. 13.71%, *Chelidonichthys* spp. 12.29%, *Auxis rochei* 3.07%, *Saurida undosquamis* 2.60%	Yes (2)
S51	Lithuania	Surimi	Modified atmosphere	Surimi slices	Surimi 42% (fish - *Merluccius gayi* and *Gadus chalcogrammus*) egg whites,crab fravour. May contain milk, molluscs trace	*Merluccius gayi* and *Gadus chalcogrammus*	*Gadus chalcogrammus* 90.29%, *Merluccius productus* 9.42%	Yes (2)
S12	Italy	Burger	Frozen	Frozen tuna burger	Yellow fin tuna (40%), Atlantic cod, eggs whites	*Thunnus albacares* and *Gadus morhua*	*Thunnus spp. 86.20%, Gadus morhua 11.78%, Gadus chalcogrammus* 1.73%, ***Sepia hierredda* 0.03% (3) (**)**	Yes (3) (**)
S24	Italy	Burger	Frozen	Frozen sea bass burger	Sea bass (45%) (*Dicentrarchus labrax*), Atlantic cod (*Gadus morhua*), eggs whites, potato flakes	*Dicentrarchus labrax* and *Gadus morhua*	*Dicentrarchus labrax* 61.61%, *Gadus morhua* 38.18%	No
S25	Italy	Burger	Defrost, modified atmosphere	Tuna burger	Yellow fin tuna (63%), Japanese spineless cuttlefish, cheese. May contains trace of molluscs, eggs	*Thunnus albacares* and *Sepiella japonica*	*Thunnus spp. 94.15%,* ***Sepia hierredda* 5.83% (**),** *Sepia* spp. 0.01%	Yes (2) (**)
S26	Italy	Burger	Defrost, modified atmosphere	Swordfish burger	Swordfish 58%, Japanese spineless cuttlefish,cheese. May contains trace of molluscs, eggs	*Xiphias gladius* and *Sepiella japonica*	*Sepia* spp. 65.52%, *Xiphias gladius* 34.35%, ***Bos taurus* 0.07% (*), *Sus scrofa* 0.05% (*)**	Yes (2) (*)
S27	Italy	Burger	Defrost, modified atmosphere	Salmon burger	Salmon 58%, South African hake, cheese. May contains trace of molluscs, eggs	*Salmo salar* and *Merluccius capensis* or *Merluccius paradoxus*	*Oncorhynchus keta 96%, Oncorhynchus nerka 1.95%, Sepia* spp. 1.42%, *Merluccius paradoxus 0.22%,* ***Bos taurus 0.12% (*)***	Yes (2) (*)

1) labels did not report the precise term 'fish' among the ingredients
2) the voluntarily declared species did not correspond to that detected by the molecular analysis
3) detection of molluscs not declared on the label
(*) non-marine taxa

674 (**) The analysis revealed the presence of species not included on the list of agreed commercial designations for use,
675 contained in MiPAAF Decree 2017
676 ND: not declared