Species identification in surimibased products using Next Generation Sequencing technologies

Giusti A., Armani A., Sotelo C. G.

The Next Generation Sequencing (NGS) technologies represent the new frontier of the DNA-based analytical methods and they are nowadays largely applied in many research fields. In the food inspection, NGS are especially suitable for species identification in highly processed products composed by a mixture of different species. The lack of the species morphological key features particularly exposes such products at risk for mislabelling and may compromise consumer's safety if toxic or allergic species are involved. Surimi typically represents a multispecies seafood commodity and it can be described as a myofibrillar protein concentrate, essentially represented by minced fish and cephalopod flesh that has been washed repeatedly with water and strained to remove sarcoplasmatic proteins, lipids and other undesirable components. Following straining, it is usually mixed with cryoprotectants and frozen. Surimi is basically an intermediate product used for the preparation of typical ready-to- eat products such as fish sausages, crabsticks and imitation shrimp products, highly marketed and consumed in almost all countries of the world. In this study, a metabarcoding approach based on the Ion Torrent NGS technologies was applied to detect and identify species within surimi-based products. Sixteen commercial products of both EU and non-EU origin were collected. DNA was extracted from all the products. Libraries were prepared by ligating Ion-compatible adapters to amplicons obtained by amplifying each DNA sample using a 16SrRNA universal primer tested on a wide range of fish and cephalopod species. In particular, the used primer pair was proved in a previous study as able to amplify a fragment of ~250-260 (depending on the species) from fish and of 190-200 bp (depending on the species) from cephalopods, both with high inter-species variability. Three equimolar pools of barcoded libraries were prepared: barcoded libraries from samples 1 -6 (Pool 1), samples 7-12 (Pool 2) and samples 13 -16 (Pool 3) were pooled together. Massive DNA clonal parallel amplification and sequencing phases were performed on Ion Chef™ System and Ion PGM™System (Thermo Fisher Scientific), respectively, following the manufacturer protocol. Three Ion 314 v2 BC sequencing chips were loaded (Chip 1 for Pool 1, Chip 2 for Pool 2and Chip 3 for Pool 3). A bioinformatic analysis was performed to evaluate the data quality in term of final number of usable sequences and to obtain the reads taxonomic profile through metabarcoding approach. Overall, 674.983 sequences (78% of total analysed) and 674.826 raw sequences (73% of total analysed) were obtained from Chip 1 and Chip 2, respectively, while Chip 3 was less performant as the 668.081 raw sequences evaluated as usable represented the 58% of all the sequences analysed. Fish DNA was the most represented in all the samples (always greater than 75% and in 75% of the samples greater than 90%); Molluscs DNA was also found in 100% of the samples, in different percentages. In particular, 25% of the samples

contained less than 1% of molluscs' DNA, 50% between 1% and 10%, 12.5% between 10% and 20%, and 12.5% more than 20%. DNA from 13 families, 19 genera and 16 species of fish, and from 3 families, 3 genera and 3 species of cephalopods was found. DNA belonging to the Gadidae family was found in 100% of the samples. The percentage was rather high in most of the cases, exceeding 90% in 50% of the samples, ranging from 70 to 90% in 31.2% of the samples and from 40 to 70% in 12.5% of the samples. Gadidae were poorly represented (<2%) only in one Asian sample of non-EU origin. DNA from Gadus genus was found in 100% of the samples, with the species *Gadus chalcogrammus* identified in 93.75% of the samples, whereas DNA from *Gadus morhua* was detected in two samples. Also, DNA from Arctogadus genus/*Arctogadus glacialis* species, Melanogrammus genus/*Melanogrammus aeglefinus* species and Merlangius genus/*Merlangius merlangus* species was detected in 6.25%, 43.75% and 6.25% of the samples, respectively. DNA belonging to Merluccidae family was found in 25% of the samples, in variable percentages (from ' 1% to over 36%). and only Merluccius genus/*Merluccius merluccius* species was present in that samples. 12.3% of DNA from Nemipteridae family/Nemipterus genus (species identification was not reached) was found in one sample. Variable percentage of DNA of species belonging to other ten families was found in some samples. Overall, the most part of the fish species found in the samples were reported in literature as used for surimi production or they were found in studies aimed at identifying species in such type of products. Unconventional species were instead found in three samples of non-EU origin. Regarding molluscs, DNA from Ommastrephidae family/Todarodes genus/*Todarodes pacificus* species was found in 75% of the samples. DNA belonging to Loliginidae family/Doryteuthis genus/ *Doryteuthis opalescens* species and Architeuthidae family/Architeuthis genus/*Architeuthis dux* species was also found in one sample. Differently from fish, the molluscs species found in the analysed samples did not correspond to those listed as the most used in surimi manufacture. Overall, 37.5% of the samples were found as mislabelled. Among them, 25% voluntary declared a species different from those detected in the analysis and 25% (all of non-EU origin) did not report the presence of molluscs on the label, posing a potential health threat for allergic consumers. The presence of vulnerable species was also proved, such in the case of *Paretroplus maculatus* (Cichlidae family) and *Trichopodus leeri* (Osphoronemidae family). Although further studies are needed to optimize the analytical protocol and larger number of samples are required, the metabarcoding approach based on the Ion Torrent NGS technologies was proved to be suitable for species identification of processed multispecies seafood products. The great and diversified range of species found in the samples, also including vulnerable ones, as well as the often-undeclared presence of potentially allergic risks, strongly remarks the still too weak control system and poorly eco-friendly management around the global

seafood sector, highlighting the necessity to further strengthen the tools aimed at ensuring both consumers and environmental protection.