TELMO JOSÉ RODRIGUES FERNANDES

Molecular-based methodologies for seafood authentication and allergen detection

Porto, 2017

PhD thesis

Molecular-based methodologies for seafood authentication and allergen detection

Thesis submitted to Faculdade de Farmácia da Universidade do Porto for Doctor Degree in Pharmaceutical Sciences - Nutrition and Food Science Specialty

Supervised by:

Dr. Isabel Maria Sousa Gomes Mafra

Prof. Dr. Maria Beatriz Prior Pinto Oliveira

Porto

December, 2017

© Authorised the partial reproduction of this thesis (subject to the approval of the publishers of journals in which the articles were published) only for research purposes through a written declaration of the person concerned that such pledges.

This work was supported by FCT (Fundação para a Ciência e Tecnologia) through project UID/QUI/50006/2013 – POCI/01/0145/FEDER/007265 with financial support from FCT/MEC through national funds and co-financed by FEDER, under the Partnership Agreement PT2020 and by the project NORTE-01-0145-FEDER-000011. T. J. R. Fernandes is grateful to FCT grant (SFRH/BD/93711/2013 financed by POPH-QREN (subsidised by FSE and MCTES).

UNIÃO EUROPEIA

Fundo Europeu
de Desenvolvimento Regional

The work presented in this thesis was performed in the Laboratory of Bromatology and Hydrology, Department of Chemical Sciences, Faculty of Pharmacy, University of Porto.

AKNOWLEDGEMENTS

My special acknowledgements to the following people and institutions:

Dr. Isabel Mafra, for sharing her scientific knowledge and critical spirit during these years, for her availability, supervision and guidance in both experimental and writing processes.

Professor Dr. Beatriz Oliveira, for sharing her scientific knowledge and providing all the necessary conditions that made possible the completion of this thesis.

Dr. Joana Costa for her constant availability to help me, for her guidance and cooperation in the experimental work and writing periods.

Catarina Silva, for her help and cooperation regarding the work of Penaeidae shrimp species discrimination.

Pascoal & Filhos SA, Marfresco and Brasmar Seafood Companies for the kind supply of some fish and crustacean species.

FCT (Fundação para a Ciência e Tecnologia), for the FCT grant (SFRH/BD/93711/2013, financed by POPH-QREN (subsidised by FSE and MCTES) and project UID/QUI/50006/2013 – POCI/01/0145/FEDER/007265 with financial support from FCT/MEC through national funds and co-financed by FEDER, under the Partnership Agreement PT2020 and by the project NORTE-01-0145-FEDER-000011.

ABSTRACT

Fishery products represent a significant market niche that includes a wide range of species with commercial relevance. Since seafood ingestion is remarkably associated with relevant health benefits, its consumption has been increased over the last decades, being Portugal recently recognised as the greatest consumer within the European Union (EU). According to the EU legislation, unprocessed and some processed seafood products must be labelled with the commercial and the scientific name of the species and fish, crustaceans and molluscs must be labelled as allergenic ingredients in foodstuffs. Nevertheless, mislabelling issues still take place nowadays, mainly by deliberate or inadvertent species substitution that may compromise seafood safety and authenticity. The number of reported cases of allergy to seafood has also increased as a result of seafood growing consumption. In particular, fish and crustaceans belong to the eight groups responsible for almost 90% of food allergies. In this sense, to help seafood industry complying with labelling and to assist regulatory authorities for the management of seafood allergens, the development of fast, sensitive and cost-effective analytical methodologies should be addressed. Therefore, the main goal of this work was to develop new, reliable and highly sensitive methods based on polymerase chain reaction (PCR), main real-time PCR for: the assessment of seafood authenticity, namely for the detection and discrimination of relevant groups of fish (Gadidae and Merluccidae) and crustaceans (Penaeidae); and the detection and quantification of fish and crustaceans as potential food allergens.

For authentication purposes, real-time PCR was the key technique that was exploited for species discrimination, targeting DNA barcodes (cytochrome c oxidase subunit I (COI) and cytochrome b (*cytb*)) combined with high resolution melting (HRM) analysis. Three methods were proposed and showed to be powerful tools for the rapid discrimination of genetically close species within the families of Gadidae, Merluccidae and Penaeidae, with high levels of confidence (>99%). The applicability of the methods was effectively demonstrated in the analysis of processed seafood products, allowing verifying a mislabelling incidence ranging from 5 to 30%. For the detection and quantification of trace levels of seafood as potential food allergens, two real-time PCR systems based on the EvaGreen dye and a TaqMan probe were tested and compared. The 16S rRNA gene was chosen as the universal marker for fish and crustacean detection and the preparation of binary model mixtures in a dynamic range of 0.0001 to 50% allowed adequate quantification systems that enabled detecting fish and crustaceans down to 1 mg/kg. In spite of the best performance of the probe system for fish detection, the application of real-time PCR system with EvaGreen dye allowed extending the number of detected fish

species. The quantitative performance of the real-time PCR probe system for shrimp detection was further compared with two kits of Enzyme-Linked Immunosorbent Assays (ELISA), which showed that, at trace amounts of crustaceans, real-time PCR and ELISA were well correlated. However, for the higher contents, both ELISA kits display much higher estimates than those of real-time PCR, suggesting potential cross-reactivity of immunoassays with other food components.

In conclusion, the present work has contributed with new PCR-based approaches exploring different DNA *loci* and using different fluorescence detection chemistries, representing novel and useful tools to verify the authenticity and labelling compliance of seafood products. The proposed highly sensitive tools can contribute to an improved allergen management both at the control laboratories and food industry. Furthermore, the proposed new methods and the results achieved herein can be considered as important advances within the scope of seafood research and seafood allergen detection.

Keywords: real-time PCR, fish, crustaceans, authenticity, allergen.

RESUMO

Os produtos provenientes da pesca representam um nicho de mercado que inclui um vasto conjunto de espécies relevantes a nível comercial. Uma vez que a ingestão de alimentos derivados da pesca está associada a importantes benefícios para a saúde, o seu consumo tem aumentado acentuadamente nas últimas décadas, sendo Portugal reconhecido como o maior consumidor dentro da União Europeia (UE). De acordo com a legislação da UE, os produtos da pesca não processados e alguns produtos processados devem ser rotulados como ingredientes alergénicos em géneros alimentícios. Não obstante, a rotulagem incorreta dos produtos continua a ocorrer hoje em dia, principalmente através da substituição deliberada ou acidental de espécies, a qual pode comprometer a segurança e a autenticidade dos géneros alimentícios provenientes da pesca. O número de casos de alergia descritos ao pescado tem aumentado, também como resultado do aumento do seu consumo. Em particular, os peixes e os crustáceos constituem dois dos oito grupos responsáveis por aproximadamente 90% das alergias alimentares. Assim, de modo a auxiliar a indústria no cumprimento da rotulagem adequada e apoiar as autoridades reguladoras na gestão dos alergénios alimentares, o desenvolvimento de metodologias analíticas rápidas, sensíveis e economicamente rentáveis deve ser assegurado. Deste modo, o principal objetivo deste trabalho foi o desenvolvimento de novos métodos com elevada sensibilidade e robustez baseados na reação em cadeia da polimerase (PCR), sobretudo em PCR em tempo real para: a avaliação da autenticidade de produtos da pesca, nomeadamente na deteção e diferenciação de grupos relevantes de peixes (Gadidae e Merluccidae) e crustáceos (Penaeidae); e a deteção e quantificação de peixe e crustáceos como potenciais alergénios alimentares.

No que concerne a avaliação da autenticidade, a PCR em tempo-real foi a principal técnica explorada para a discriminação de espécies, tendo como alvo *DNA barcodes* (subunidade I da citocromo oxidase (COI) e citocromo b (*cytb*)) combinando com a análise por *melting* de alta resolução (HRM). Foram propostos 3 métodos que demonstraram ser ferramentas úteis para a discriminação rápida de espécies geneticamente relacionadas das famílias Gadidae, Merluccidae e Penaeidae, com um elevado nível de confiança (>99%). A aplicabilidade destes métodos foi efetivamente demonstrada na análise de produtos processados, possibilitando verificar a rotulagem incorreta de alguns produtos numa incidência que variou entre 5 e 30% dos casos. Para a deteção e quantificação de níveis vestigiais de peixe e crustáceos como potenciais alergénios, foram testadas e comparadas duas técnicas de PCR em tempo real baseadas em sistemas de corante EvaGreen e sondas TaqMan. O gene 16S rRNA foi escolhido

como marcador universal para a deteção de peixes e crustáceos, sendo preparadas misturas-modelo binárias num intervalo dinâmico de 0,0001 até 50%, o que permitiu a sua deteção e quantificação até uma concentração de 1 mg/kg. Apesar do melhor desempenho do sistema com sondas, a aplicação da técnica de PCR em tempo real com o corante EvaGreen permitiu aumentar o número de espécies de peixe detetadas. O desempenho a nível quantitativo do sistema de PCR em tempo real com sonda para a deteção de camarão foi posteriormente comparado com dois kits de ensaio de imunoabsorção com ligação enzimática (ELISA), o que demonstrou uma correlação entre os valores obtidos pelos dois tipos de técnicas a nível vestigial. No entanto, no que diz respeito às concentrações mais elevadas, os dois kits ELISA apresentaram valores muito mais elevados do que os verificados por PCR em tempo real, o que sugere a ocorrência de potenciais fenómenos de reatividade cruzada dos ensaios imunoquímicos com outros ingredientes.

Como conclusão, o presente trabalho contribuiu para o desenvolvimento de novas metodologias com base na PCR, explorando diferentes *DNA loci* com uso de diferentes sistemas de deteção por fluorescência, representando ferramentas úteis e inovadoras para a verificação da autenticidade e conformidade da rotulagem de produtos da pesca. As ferramentas propostas com elevada sensibilidade podem contribuir para uma melhor gestão de alergénios alimentares, tanto nos laboratórios de controlo da qualidade assim como na indústria alimentar. Além disso, os novos métodos propostos e os resultados aqui obtidos podem ser considerados avanços importantes no âmbito da investigação sobre produtos da pesca e deteção de alergénios de peixes e crustáceos.

Palavras-chave: PCR em tempo real, peixe, crustáceos, autenticidade, alergénios.

LIST OF PUBLICATIONS AND COMMUNICATIONS

Publications in international peer-review Journals

1. Telmo J. R. Fernandes, Joana Costa, M. Beatriz P. P. Oliveira, Isabel Mafra (**2015**). An overview of fish and shellfish allergens and current methods of detection, *Food and Agricultural Immunology*, 26, 848-869.

2. Telmo J. R. Fernandes, Joana Costa, Isabel Carrapatoso, M. Beatriz P. P. Oliveira, Isabel Mafra (**2017**). Advances on the molecular characterisation, clinical relevance and detection methods of Gadiform parvalbumin allergens, *Critical Reviews in Food Science and Nutrition*, 15, 3281-3296.

3. Telmo J.R. Fernandes, Joana Costa, M. Beatriz P.P. Oliveira, Isabel Mafra (**2017**). DNA barcoding coupled to HRM analysis as a new and simple tool for the authentication of Gadidae fish species, *Food Chemistry*, 230, 49-57.

4. Telmo J.R. Fernandes, Catarina R. Silva, Joana Costa, M. Beatriz P.P. Oliveira, Isabel Mafra (**2017**). High resolution melting analysis of a COI mini-barcode as a new approach for Penaeidae shrimp species discrimination. *Food Control*, 82, 8-17.

5. Telmo J.R. Fernandes, Joana Costa, M. Beatriz P.P. Oliveira, Isabel Mafra (**2018**). COI barcode-HRM as a novel approach for the discrimination of hake species. *Fisheries Research*, 197, 50-59.

6. Telmo J.R. Fernandes, Joana Costa, M. Beatriz P.P. Oliveira, Isabel Mafra (**2018**). Exploiting 16S rRNA gene for the detection and quantification of fish as a potential allergenic food: A comparison of two real-time PCR approaches, *Food Chemistry, 245*, 1034-1041.

7. Telmo J.R. Fernandes, Joana Costa, M. Beatriz P.P. Oliveira, Isabel Mafra (**2017**). A new real-time PCR quantitative approach for the detection of shrimp crustaceans as potential allergens, *Journal of Food Composition and Analysis,* submitted.

8. Telmo J.R. Fernandes, Joana Costa, M. Beatriz P.P. Oliveira, Isabel Mafra (**2017**). A novel quantitative approach for fish detection based on the use of mitochondrial gene markers, Annals of Medicine, accepted for publication.

9. Telmo J. R. Fernandes, M. Beatriz P. P. Oliveira (**2017**). DNA barcoding applied to seafood authentication, *Trends in Food Science and Technology*, submitted.

Book chapters

Joana Costa, Telmo J.R. Fernandes, Caterina Villa, M. Beatriz P.P. Oliveira and Isabel Mafra (**2017**). Advances in Food Allergen Analysis. In Umile Gianfranco Spizzirri and Giuseppe Cirillo (eds.) Food Safety: Innovative Analytical Tools for Safety Assessment, (305–360) © Scrivener Publishing LLC.

Oral communications in Scientific Meetings

1. T.J.R. Fernandes, J. Costa, M.B.P.P. Oliveira, I. Mafra. Clustering of cod-like fish species by DNA barcoding and HRM analysis. 5th MoniQA International Conference 2015, 16-18 September **2015**, Porto, Portugal.

2. Catarina Silva, Telmo J. R. Fernandes, Joana Costa, M. Beatriz P. P. Oliveira, Isabel Mafra. Clustering of shrimp and prawn species by DNA barcoding and HRM analysis. IJUP 16 - Encontro Investigação Jovem da Universidade do Porto, 17-19 February **2016**, Porto, Portugal.

3. T.J.R. Fernandes, J. Costa, M.B.P.P. Oliveira, I. Mafra. COIBar-HRM as a novel approach for the discrimination of hake species. Food Integrity Conference 2017, 10-11 May **2017**, Parma, Italy.

Awards

The following awards were attributed regarding panel and oral communications:

Best panel communication – 2º Simpósio Nacional SPASS, Lisboa, Portugal, November **2015**.

Young Researcher Award – Food Integrity Conference, Parma, Italy, May **2017**.

ACRONYMS

aa – amino acid

AFLP – amplified fragment length polymorphism

BLAST – basic local alignment search tool

BOLD – Barcode of Life Database

bp – base pair

COI – cytochrome oxidase, subunit i

Ct – cycle threshold

Cob – apocytochrome b

Cytb – cytochrome b

DBPCFC – double-blind placebo-controlled food challenge

DNA – deoxyribonucleic acid

EAACI – European academy of allergology and clinical immunology

ELISA – enzyme-linked immunosorbent assay

EU – European Union

FAO – Food and Agriculture Organisation of the United Nations

FINS – forensically informative nucleotide sequencing

FISH-BOL – Fish Barcode of Life Initiative

GMO – genetically modified organisms

HRM – high resolution melting

IEF – isoelectric focusing

IgE – immunoglobulin E

IgG – immunoglobulin G

ITS-1 – nuclear internal-transcribed spacer 1

IUIS – International Union of Immunological Societies

LC-ESI-IT – liquid chromatography with electrospray ionization-ion trap

LC-MS – liquid chromatography mass spectrometry

LFD – lateral flow device

LNA – locked nucleic acid

LOAEL – lowest observed adverse effect level

LOD – limit of detection

LOQ – limit of quantification

MS – mass spectrometry

NCBI – National Center for Biotechnology Information

ND2 – NADH dehydrogenase subunit 2

NGS – next-generation sequencing

NOAEL – no observed adverse effect level
OAS – oral allergy syndrome
OFC – oral food challenge
PAPM – polyphenolic adhesive protein of mussels
PCR – polymerase chain reaction
PEPCK – phosphoenolpyruvate carboxykinase
PDB – Protein Data Bank
PPi – pyrophosphate
PUFA – polyunsaturated fatty acids
QCM – quartz crystal microbalance
RAPD – random amplified polymorphic DNA
RFLP – restriction fragment length polymorphism
RNA – ribonucleic acid
SDS-PAGE – sodium dodecyl sulphate - polyacrylamide gel electrophoresis
SNP – single-nucleotide polymorphism
SPR – surface plasmon resonance
SPT – skin prick test
SPMN-LFIA – superparamagnetic nanoparticle-based lateral flow imunnoassay
SSCP - -single-stranded conformational polymorphism
Ta – temperature of annealing
Tm – temperature of melting
USA – United States of America
UV – ultraviolet radiation
WHO – World Health Organization

TABLE OF CONTENTS

AKNOWLEDGEMENTS viii

ABSTRACT ix

RESUMO xi

LIST OF PUBLICATIONS AND COMMUNICATIONS xiii

ACRONYMS xv

TABLE OF CONTENTS xviii

GENERAL INTRODUCTION 1

The importance of seafood 3

Seafood mislabeling and traceability 5

Allergy to seafood 8

Methodologies for seafood detection and species discrimination 8

Objectives and thesis outline 11

References 13

CHAPTER 1. SEAFOOD SPECIES AUTHENTICITY 21

STATE-OF-THE-ART 23

DNA barcoding applied to seafood authentication 25

EXPERIMENTAL PART 57

DNA barcoding coupled to HRM analysis as a new and simple tool for the authentication of Gadidae fish species 59

High resolution melting analysis of a COI mini-barcode as a new approach for Penaeidae shrimp species discrimination 79

COI barcode-HRM as a novel approach for the discrimination of hake species 101

CHAPTER 2. SEAFOOD ALLERGEN DETECTION 123

STATE-OF-THE-ART 125

An overview on fish and shellfish allergens and current methods of detection 127

Advances on the molecular characterisation, clinical relevance and detection methods of Gadiform parvalbumin allergens 155

EXPERIMENTAL PART 195

Exploiting 16S rRNA gene for the detection and quantification of fish as a potential allergenic food: A comparison of two real-time PCR approaches 197
A new real-time PCR quantitative approach for the detection of shrimp crustaceans as potential allergens 217
FINAL REMARKS 237

GENERAL INTRODUCTION

The importance of seafood

Fish and shellfish represent a relevant economical niche that includes a large range of commercially valuable animal species that are highly consumed. At the moment, it is the most traded food worldwide and one of the most common protein sources for human consumption (almost one quarter of protein intake comes from seafood) [1]. Seafood can be categorised into three different phyla: Chordata (bony or cartilaginous fish), Arthropoda (crustacean shellfish) and Mollusca (molluscan shellfish and cephalopods) [2, 3, 4] (Table 1).

Table 1. Taxonomic classification of common seafood (adapted from [4]).

Phylum	Class	Seafood examples
Chordata	Osteichthyes (bony fish)	Cod, hake, salmon, tuna
	Chrondrichtyes (cartilaginous fish)	Sharks, rays
Arthropoda	Crustaceans	Shrimps, lobsters, crabs
Mollusca	Gastropoda	Snails, abalone
	Bivalves	Mussels, oysters
	Cephalopods	Squid, octopus

In the last 50 years, there has been a remarkable increase in seafood consumption due to many factors, among which the consumer's new perception of their nutritional and health benefits [5]. In 2015, the global production of seafood reached 169 million tonnes, which included 93 million tonnes from capture and 77 million tonnes regarding aquaculture production [6]. Total seafood production (capture and aquaculture) is projected to expand over the upcoming years, reaching 196 million tonnes in 2025 [7].

Regarding fish consumption, European consumers ate more 650 000 tonnes/year in 2014 than in 2013, which represents an increase of 4.5% between this period, being Portugal the country with the highest consume of seafood per capita within the European Union [8]. Globally, China is the top consumer country, surpassing all continent values and the world's average consumption (kg/person/year) [7] (Figure 1). As observed in Figure 1, the trend for the upcoming years is for an increase in fish consumption. China is also the major producer regarding marine fisheries, followed by Indonesia, the United States of America and the Russian Federation [7]. The top-ranked species in terms of global catch was Alaska pollock (*Theragra chalcogramma*). Regarding aquaculture production, China accounted for 45.5 million tonnes in 2014, which corresponds to more

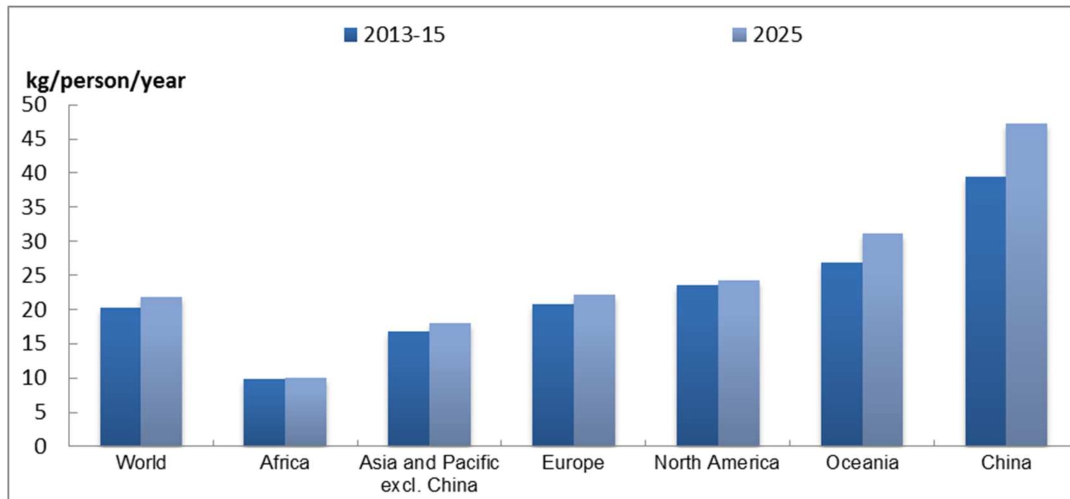


Figure 1. Per capita fish consumption (adapted from [7]).

than 60% of global fish production from aquaculture. Other major producers were India, Vietnam, Bangladesh and Egypt [7]. Perciformes, Salmoniformes, Clupeiformes and Gadiformes are some of the most important fish orders, in terms of economic value, consumption and nutrition of the populations worldwide. The last reached almost 5.9 million tonnes of global marine catch in 2011, comprising some of the most relevant white fish groups, namely the Gadidae and the Merluccidae families. These two families include highly appreciated and consumed fish species, such as codfish (e.g. *Gadus morhua*, *G. macrocephalus*) and hakes (e.g. *Merluccius merluccius*, *M. productus*, *M. hubbsi*) [7].

Crustaceans, such as shrimps, crabs or lobsters, are one the most valued seafood groups all over the world and a major food supply, accounting for 7.4 million tonnes produced in 2015 [8]. About half of all crustaceans on the global seafood market belong to shrimps of the Penaeidae family, which demonstrates the importance of this particular group. *Litopenaeus vannamei*, is the most commonly consumed crustacean species worldwide, accounting with a production of 3.9 million tonnes in 2014 [7, 9, 10].

Despite the substantial variation of seafood composition among different species, there are numerous scientific reports assessing the health benefits associated with the ingestion of seafood nutritional components, namely the long chain omega-3 fatty acids eicosapentaenoic acid (EPA) and docosahexaenoic (DHA), selenium, iodine, potassium vitamins D, A, E and B_{12}, among others [5, 11, 12, 13]. Indeed, the consumption of seafood plays a significant role in the prevention of cardiovascular diseases or cancer and it is also important to beat undernutrition and micronutrient deficiencies [14, 15, 16].

In spite of the continuously increasing consumption and well-studied nutritional advantages, fish and shellfish ingestion entails some risks and hazards regarding human health. Besides the putative contaminants present in seafood, such as methylmercury and

persistent organic pollutants (POP), toxins or pathogens, seafood can pose severe health problems for allergic individuals.

Seafood mislabeling and traceability

The thousands of seafood species present on global supply chains taken together with the pressure to meet the demand for seafood consumption, makes authentication and traceability of seafood products very hard tasks to accomplish. Since most of the captured seafood is exported to other countries, rather than being locally consumed, the opportunities for fraud increase substantially through the supply chains, making it difficult to detect when and where the potential mislabelling phenomenon has occurred. The fraudulent practices are associated to the declared origin of the seafood, the declared weight and the partial or total species substitution [17, 18].

Seafood mislabelling is a particular sensitive issue because human health can be threatened by exposing consumers to contaminants or allergens related with a particular species of fish, crustaceans or molluscs [3, 17]. Fraudulent labelling of seafood species has also been pointed out as one of the causes of wrong exploitation in several seafood groups and can, ultimately, contribute to unintentionally overexploit some species [19]. Species substitution has been broadly reported by both the media and the scientific community. When deliberate substitution takes place, it is usually carried out by the replacement of valuable species with less expensive ones (economically motivated adulteration), such as the replacement of codfish species (e.g. *G. morhua*) by pollack (e.g. *T. chalcogramma*) [20, 21].

A crucial aspect of seafood market regulation is the correct identification of seafood when it is captured, imported and finally sold to consumers at retail markets or restaurants. The identification based on visual analysis is often very difficult or impossible to be accurately accomplished due to the phenotypic similarities within closely related seafood species and to different presentations after processing and packaging, which often result in the loss of the morphological characteristics (fillets, tails, loins, compressed blocks, seafood sticks, surimi, etc. [22, 23]. These issues raise the challenge of developing efficient authentication methodologies for seafood species identification. In particular, DNA-based approaches can circumvent the problems associated with morphological identification and are currently indispensable tools to check the accuracy of seafood labelling [18, 24]

According to the European Union (EU) legislation, unprocessed and some processed seafood products (e.g. smoked products, salted) must be labelled with the commercial and the scientific name of the species, fishing gear category, net weight, food operator,

identification mark, production method, catch area and storage conditions, while in other products, such as, canned, composite products and breaded products, the label must contain the name of the food, net quantity, food operator, identification mark, list of ingredients, "best before"/"use by date" and storage conditions, being no other information compulsory [25] Furthermore, Regulation (EU) No 1169/2011 demands the obligatory labelling of a list of allergenic foods, in which fish, crustaceans and molluscs are included. These food components should be highlighted from the listed ingredients, regardless of their quantity [26].

Although food product traceability is well established in the EU and supported by proper labelling legislation, a variety of common names may be used for a given species and *vice versa*, even within the same jurisdiction. This situation is even more complex due to the fact that jurisdictions regarding the application of market names are not harmonised among countries, in part because of different language usage, which challenges accurate labelling at a global scale [18]. For instance, the Portuguese legislation [27] states that *G. morhua*, *G. macrocephalus* and *G. ogac* must be named as Atlantic, Pacific and Greenland cod, respectively, or under the commercial designation of cod or codfish ("bacalhau"). On the other hand, the Decree of the Italian Ministry of Agricultural, Food and Forestry Policies [28] dated 31 January 2008 establishes that codfish ("baccalà") designation can be exclusively used when *G. macrocephalus* and *G. morhua* are present in foods [20]. Crossing this information with general European legislation, it is not clear whether it is correct or not to call "cod" to only two or three species in food on the European market.

According to the most recent literature regarding seafood authenticity assessment, mostly based on DNA barcoding approaches, codfish, hake, salmon, tuna and shrimps are amongst the most studied (Table 2). All the reports resumed in Table 2 describe mislabelling cases. The extensive study conducted by Bréchon et al. (2016) reports 100% of accurately labelled codfish products in Norway, Iceland, Netherlands and Belgium. Nevertheless, the same work reported the occurrence of codfish mislabelling in United Kingdom, Denmark, Sweden, Canada and Estonia, mostly due to the partial or total inclusion of haddock (*Melanogrammus aeglefinus*), Alaska pollock (*T. chalcogramma*) and Argentine hake (*M. hubbsi*) [29].

The levels of mislabelling ranged from 6% or less [29, 30, 31, 32, 33, 34] to the alarming value of 86%, observed in frozen products sold as sablefish (*Anoplopoma fimbria*) in China market, mostly substituted by Patagonian toothfish (*Dissostichus eleginoides*) and Antarctic toothfish (*Dissostichus mawsoni*) [35].

Table 2. Examples of recent scientific articles regarding seafood authenticity surveys.

Seafood samples/ species groups	Number of samples	Mislabelling (%)	Location	Reference
Packaged frozen fishery products (e.g. Gadidae, Merluccidae)	120	5	Italy	[31]
Convenience seafood products (Gadidae, Merluccidae)	54	35	Italy	[36]
Processed seafood products including fish (e.g. Salmonidae), surimi, sushi, bivalves (e.g. Ostreidae)	62	16	Malaysia	[37]
Frozen seafood: fish (e.g. Gadidae, Merluccidae), crustaceans (e.g. Penaeidae and Solenoceridae), bivalves (e.g. Mytilidae)	60	19	Portugal	[38]
Products commercialised as cod (e.g. Gadidae, Tetraodontidae)	52	60	China	[39]
Frozen products sold as sablefish (e.g. Anoplopomatidae, Nototheniidae)	42	86	China	[35]
Processed fish (e.g. Salmonidae) and crustacean products (e.g. Penaeidae)	100	22	India	[40]
Codfish products (e.g. Gadidae)	43-53 (each country)	2 19 4 0 0 7 59 0 0	UK Denmark Sweden Norway Iceland Canada Estonia Netherlands Belgium	[29]
Fishery products including fish (e.g. Merluccidae), cephalopods (e.g. Sepiida), crustaceans (e.g. Penaeidae)	277	22	Italy	[41]
Processed cephalopods products including cuttlefish, octopus, squid	95	2	China	[34]
Salmonid products (e.g. Salmonidae)	111	6 24	Alaska/Canada Spain	[33]
Sushi seafood (e.g. Salmonidae, Scombridae)	185	3	Italy	[30]
Fish (e.g. Ophidiidae, Scombridae, Gadidae, Merluccidae)	255	17	Brazil	[42]
Typical sushi fish (e.g. Scombridae, Salmonidae)	364	47	USA	[43]
Fish (e.g. Salmonidae, Gadidae), crustaceans (e.g. Penaeidae)	118	6	Germany	[32]

Allergy to seafood

The growing consumption of seafood at a global scale has also resulted in a clear increase in the number of reported cases of allergy to seafood. Adverse reactions to fish and shellfish might occur in sensitised individuals after ingestion, direct contact or inhalation of their cooking odours or fumes, induced by immunoglobulin E (IgE), although non-immune-mediated reactions may also occur, presenting similar symptoms. The most frequent clinical manifestations of seafood allergy include oral allergy syndrome, gastrointestinal symptoms, urticarial/angioedema and life-threatening reactions of anaphylaxis [3, 44, 45, 46, 47].

Fish and crustaceans belong to the eight groups responsible for almost 90% of food allergies (“big eight”). The estimation of the actual prevalence of food allergies is still currently uncertain and is probably overestimated [48]. Nevertheless, overall data suggest that 1-10% of the general population suffers from food allergies [49], being children the most affected group [50]. Shellfish allergy is estimated to affect 0.5-2.5% of the general population, while fish allergy prevalence ranges from 0.1 to 0.4%, although these numbers can vary significantly among different countries or regions [9, 51, 52]. There is about 75% of chance of a sensitised individual being cross-reactive to another shellfish species, which decreases to 50% in the case of second contact [9, 53].

The major seafood allergens are parvalbumins and tropomyosins, two high molecular weight proteins that are heat stable and highly resistant to food cooking. Several occurrences of these proteins as IgE reactive molecules have been already reported, being officially named and recognised by the IUIS Allergen Nomenclature Sub-committee [54]. Furthermore, different proteins, such as enolases and aldolases were already described as IgE reactive in fish species, while, arginine kinase, myosin light chain, sarcoplasmatic calcium-binding protein and troponin C have been reported as new allergens of crustaceans [44].

Methodologies for seafood detection and species discrimination

The development of analytical methodologies to verify label compliance and to control the presence of fish and shellfish as allergenic ingredients is nowadays of utmost importance for consumer protection. Due to the lack of harmonised and standardised official methodologies, there is a need for the development and implementation of reliable, cost-effective and simple tools for addressing seafood authentication [15, 55]. The ability of a methodology to correctly trace a specific fish or shellfish species depends on the target species, target analyte, analyte extraction and food matrix, which often explains

differences in method sensitivity. In the particular case of seafood species identification of allergen detection, analytical methods may either target proteins (allergenic or marker proteins), or target DNA markers. The availability of specific antibodies or primers is determinant for the selection of the analytical method based on protein or DNA analysis, respectively [56, 57].

Authentication and allergen detection approaches of seafood may rely on protein analysis by enzyme-linked immunosorbent assays (ELISA), isoelectric focusing, lateral flow devices (LFD) and liquid chromatography with mass spectrometry (MS) detection [24, 44, 58, 59, 60]. The DNA-based methods have been suggested to be the most powerful and reliable approaches to identify species in processed seafood products where morphological features are no longer present [61, 62]. There are some well-known advantages of using DNA as target molecules, such as their high thermal stability and resistance to food processing in comparison with proteins.

Particularly, the methods based on polymerase chain reaction (PCR) can provide highly sensitive and specific tools for species identification and allergen detection in seafood [58, 60]. Both conventional PCR and real-time PCR methods are considered as useful approaches for seafood analysis, offering several advantages such as low-cost, simplicity and species detection at trace levels [63]. Several PCR-based methods have been applied to seafood species identification and allergen analysis, which have been combined with DNA barcoding, forensically informative nucleotide sequencing (FINS) and restriction fragment length polymorphism (PCR-RFLP). Recently, several methods based on real-time PCR (probe-based assays or DNA binding dyes), high resolution melting (HRM) analysis and digital PCR, among others, have also been applied to seafood species detection [15, 24, 58, 64]. Concerning seafood allergen analysis, although PCR-based approaches are considered indirect methods of detection, i.e. targeting a DNA marker sequence and not an allergenic protein, they are especially useful to analyse highly processed foods. PCR-based approaches rely on the amplification of sequences from genes encoding allergenic proteins (e.g. parvalbumin, tropomyosin) or mitochondrial DNA regions (e.g. 16S rRNA, cytochrome oxidase subunit I (COI), cytochrome b) that are able to discriminate selected species or to detect/quantify particular groups of organisms in one system, which usually results in more sensitive assays [21, 24, 58, 65, 66, 67].

Real-time PCR has been extensively used in recent years in food analysis, being particularly advantageous due to its quantification capacity [60, 61]. Besides, no post-PCR analysis is required to monitor the amplification process since the measurement of fluorescence intensity allows the number of amplicons to be detected in real-time after each cycle [61]. Currently, the fluorescence signal is provided by two main detection chemistries: hydrolysis probes (e.g. TaqMan and MGB-TaqMan) and double-stranded

DNA (dsDNA) intercalating molecules (e.g. SYBR Green I and EvaGreen dyes) [68]. The hydrolysis probes are designed to bind to a specific region of the target DNA, having a donor fluorescent moiety at the 5′-end and an acceptor fluorescent moiety at the 3′-end that quenches the fluorescence emitted from the donor molecule due to their close proximity. During the extension phase, the bound hydrolysis probe is degraded by the 5′–3′-exonuclease activity of DNA polymerase, generating fluorescence from the donor. The intercalating dyes bind to the minor groove of dsDNA, being the fluorescence increased and measured in the extension phase of each cycle. Given that nonspecific products and primer-dimers can be formed during the amplification process, a melting curve analysis is highly recommended to check the specificity of the amplified fragments [68]. The most commonly used dye is SYBR Green, though some limitations can be highlighted, namely the limited dye stability and possibility of PCR inhibition. EvaGreen is a third generation dsDNA binding dye that offers several advantages, such as being less inhibitory to PCR than SYBR Green I, which enables its use under saturating conditions to generate enhanced fluorescent signals. Besides, EvaGreen is well suited for HRM [68, 69].

HRM analysis detects the variations between sequences based on the melting properties of the dsDNA. The melting temperature monitoring, allows the detection of single base variations between two dsDNA by a shift in melting temperature and a difference of the curve shape. A statistical treatment of HRM software enables the melting curves to systematically assign samples into different clusters, allowing a suitable data manipulation [64, 70]. This approach has proved to be simple, cost-effective and rapid, being applied as a diagnostic tool in several areas, with a significant relevance in species identification and differentiation [71]. Particularly, when HRM targeting a DNA barcode is applied for species authentication, the unknown species is often identified by its comparison to reference melting curves [21, 65]. This approach of HRM combined with DNA barcoding represents a suitable advanced alternative tool for seafood species identification, allowing to overcome some drawbacks associated with sequencing and enabling the reduction of costs per sample and the time of analysis. These technologies should be considered as proper tools for routine food industry analysis [24].

Objectives and thesis outline

Considering that in the EU, the unprocessed seafood products should be labelled with both common and scientific names of the species, among other relevant data, and that fish, crustaceans and molluscs must be declared and highlighted on the label, there is a need for reliable and cost-effective tools to verify labelling compliance. In this context, this PhD thesis intended to address this need, proposing new DNA-based methods as accurate, sensitive and cost-effective tools to differentiate and detect seafood species at trace levels. Bearing this in mind, the main goals of this PhD project regarded the development of PCR-based methods, mostly based on real-time PCR for:

- The assessment of seafood authenticity, namely for the detection and discrimination of relevant groups of fish and crustaceans;
- The detection and quantification of fish and crustaceans as potential food allergens.

To accomplish the main goals, specific objectives were established throughout the development of this project:

- Identification of suitable DNA barcodes for the identification and discrimination of seafood species, using fish species from the order of Gadiformes, particularly Gadidae and Merluccidae families, and crustacean (shrimp) species from the Penaeidae family as case studies;
- Development of novel real-time PCR systems coupled with HRM analysis to combine with DNA barcoding, aiming at differentiating three groups of closely related species from the Gadidae, Merluccidae and Penaeidae families;
- Validation of the barcoding-HRM systems by sequencing and their application to verify labelling compliance of processed seafood products;
- Identification of appropriate DNA regions as universal markers for fish and crustacean detection;
- Development of two new quantitative real-time PCR systems targeting universal sequences specific and highly sensitive for fish and crustacean detection as potential allergens;
- Comparison of the performance of real-time PCR systems using two fluorescence chemistries, namely a hydrolysis probe and the EvaGreen dye, to detect and quantify fish, as case study;
- Evaluation of the performance of the quantitative real-time PCR systems, considering the recommended guidelines for these assays, and their validation in terms of precision and trueness using blind mixtures;

- Application of the two new quantitative real-time PCR systems to detect fish and crustaceans at trace levels in processed foods and comparison of the applicability performance with ELISA kits targeting crustacean species.

This PhD thesis was structured in two main chapters regarding the two established goals, namely seafood species authentication and seafood allergen detection. The two main chapters were proceeded by a general introduction, describing the global relevance of seafood market, current labelling and authenticity issues, an overview of prevalence of seafood allergy and most important allergens, and a brief review on the available methodologies for seafood authentication and allergen detection.

Chapter 1 introduces the first topic of this PhD thesis, mostly centred on fish and crustacean species authentication. It includes a state-of-the-art section concerning the application of DNA barcoding on seafood authentication and most relevant methodologies based on these standardised DNA regions. The review is followed by an experimental part composed of three research articles concerning the development of novel real-time PCR systems coupled with HRM analysis and DNA barcoding, which were designed to target and discriminate different seafood species from three fish and crustacean important families, namely Gadidae, Merluccidae and Penaidae.

Chapter 2 concerns the second main topic of this thesis, introducing seafood as a source of allergens and the methods for allergen detection. The state-of-the-art is presented in two review articles, one general overview on fish and shellfish allergens and their detection, and a second extended review on the molecular characterisation, clinical relevance and methodologies for fish allergen detection, focusing on the parvalbumins from the order of Gadiformes. The experimental part includes two research papers regarding: (i) the development of quantitative real-time PCR systems for fish detection, comparing the fluorescence chemistries using hydrolysis probes and EvaGreen dye; (ii) the development of a quantitative real-time PCR system with a hydrolysis probe to detect shrimp crustaceans and comparison of its applicability to analyse processed foods with two ELISA kits for crustacean detection.

The final section includes an overall conclusion of this work and future perspectives, with a particular emphasis on the major and new achievements of this PhD project on seafood authentication and allergen detection. The research work of the present thesis was performed at the Laboratory of Bromatology and Hydrology, Department of Chemical Sciences, Faculty of Pharmacy, University of Porto, following the research that has been conducted as part of the REQUIMTE-LAQV.

References

[1] Tahergorabi R, Jaczynski J. Seafood proteins and human Health. In: Bibus DM, editor. Fish and fish oil in health and disease prevention. San Diego: Academic Press; 2016. p. 323-31.

[2] Prester L. Seafood allergy, toxicity, and intolerance: A review. J Am Coll Nutr. 2016; 35: 271-83.

[3] Sicherer SH. Epidemiology of food allergy. J Allergy Clin Immunol. 2011;127: 594-602.

[4] Wong GWK. Detection and control of fish, shellfish and molluscs as food allergens In Flanagan, S, editor. Handbook of food allergen detection and control. Woodhead Publishing; 2015. p. 379-89.

[5] Rittenschober D, Stadlmayr B, Nowak V, Du J, Charrondiere UR. Report on the development of the FAO/INFOODS user database for fish and shellfish (uFiSh) – Challenges and possible solutions. Food Chem. 2016; 193: 112-20.

[6] FAO, Food and Agriculture Organization of the United Nations. Fisheries and aquaculture statistics 2015, FAO Yearbook. Rome, Italy. Available at: http://www.fao.org/3/a-i7989t.pdf (Last accessed 18 December 2017).

[7] FAO, Food and Agriculture Organization of the United Nations. The state of world fisheries and aquaculture 2016, contributing to food security and nutrition for all. Rome, Italy. 200 pp.

[8] EUMOFA. Directorate-general for maritime affairs and fisheries of the European Commission, European Market Observatory for Fisheries and Aquaculture Products (EUMOFA) (2016). EUMOFA The EU Fish Market.

[9] Khora SS. Seafood-Associated Shellfish Allergy: A Comprehensive Review. Immunol Invest. 2016; 45: 504-30.

[10] Pedrosa M, Boyano-Martínez T, García-Ara C, Quirce S. Shellfish allergy: a comprehensive review. Clinic Rev Allerg Immunol. 2015; 49: 203-16.

[11] Almeida I, Fernandes TJR, Guimarães BMR, Oliveira MBPP. Omega-3 dietary intake: review on supplementation, health benefits and resources sustainability. In Wasim Khan, editor. Omega-3 fatty acids chemistry, dietary sources and health benefits. New York: Nova Science Publishers, 2014. p.191-211.

[12] Oehlenschlager J. Seafood: nutritional benefits and risk aspects. International J Vita Nutr Res. 2012; 82: 168-76.

[13] Sampels S. The effects of processing technologies and preparation on the final quality of fish products. Trends Food Sci Technol. 2015; 44: 131-46.

[14] Hosomi R, Yoshida M, Fukunaga K. Seafood consumption and components for health. Glob J Health Sci. 2012; 4: 72-86.

[15] Leal MC, Pimentel T, Ricardo F, Rosa R, Calado R. Seafood traceability: current needs, available tools, and biotechnological challenges for origin certification. Trends Biotechnol. 2015; 33: 331-6.

[16] Nunes ML, Bandarra NM, Batista I. Health benefits associated with seafood consumption. Handbook of Seafood Quality, Safety and Health Applications: Wiley-Blackwell; 2010. p. 367-79.

[17] Angers A, Ballin NZ, Hofherr J, Kagkli DM, Lievens A, Maquet A, Martinsohn JT, Paracchini V, Petrillo M, Puertas-Gallardo A. Enhancing fish species identification using novel markers and emerging technologies. JRC Technical Reports. 2017, 1-21.

[18] Naaum AM, Warner K, Mariani S, Hanner RH, Carolin CD. Seafood mislabeling incidence and impacts. In Hanner, R, Naaum A, editors. Seafood authenticity and traceability. San Diego: Academic Press; 2016. p. 3-26.

[19] Muñoz-Colmenero M, Klett-Mingo M, Díaz E, Blanco O, Martínez JL, Garcia-Vazquez E. Evolution of hake mislabeling niches in commercial markets. Food Control. 2015; 54: 267-74.

[20] Di Pinto A, Di Pinto P, Terio V, Bozzo G, Bonerba E, Ceci E, Tantillo G. DNA barcoding for detecting market substitution in salted cod fillets and battered cod chunks. Food Chem. 2013; 141: 1757-62.

[21] Fernandes TJR, Costa J, Oliveira MBPP, Mafra I. DNA barcoding coupled to HRM analysis as a new and simple tool for the authentication of Gadidae fish species. Food Chem. 2017; 230: 49-57.

[22] Abdullah A, Rehbein H. The differentiation of tuna (family: Scombridae) products through the PCR-based analysis of the cytochrome b gene and parvalbumin introns. J Sci Food Agric. 2016; 96: 456-64.

[23] Naaum AM, Hanner RH. An Introduction to DNA-Based tools for seafood identification. In Hanner, R, Naaum A, editors. Seafood authenticity and traceability. San Diego: Academic Press; 2016. p. 99-111.

[24] Fernandes TJR, Oliveira MBPP, Mafra I. An updated review on DNA barcodes for seafood authentication. 2017, submitted.

[25] Regulation (EC) No. 1379/2013 of 11 December 2013 on the common organisation of the markets in fishery and aquaculture products, amending Council Regulations (EC) No 1184/2006 and (EC) No 1224/2009 and repealing Council Regulation (EC) No 104/2000. 354, p. 1-21.

[26] Regulation (EU) No 1169/2011 of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No. 1924/2006 and (EC) No. 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004, Official Journal of the European Union, L304, p. 18-63.

[27] Decreto-lei n° 25/2005. Ministério da Agricultura, Pescas e Florestas. Diário da República, I série, no 20, p. 696–703.

[28] MiPAAF. Decree of the Italian Ministry of Agricultural, Food and Forestry Policies of 31 January 2008. Italian name for fish species of commercial interest.

[29] Bréchon AL, Hanner R, Mariani S. A systematic analysis across North Atlantic countries unveils subtleties in cod product labelling. Marine Policy. 2016; 69: 124-33.

[30] Armani A, Tinacci L, Lorenzetti R, Benvenuti A, Susini F, Gasperetti L, Ricci E, Guarducci M, Guidi A. Is raw better? A multiple DNA barcoding approach (full and mini) based on mitochondrial and nuclear markers reveals low rates of misdescription in sushi products sold on the Italian market. Food Control. 2017; 79: 126-33.

[31] Di Pinto A, Mottola A, Marchetti P, Bottaro M, Terio V, Bozzo G, Bonerba E, Ceci E, Tantillo G. Packaged frozen fishery products: species identification, mislabeling occurrence and legislative implications. Food Chem. 2016; 194: 279-83.

[32] Günther B, Raupach MJ, Knebelsberger T. Full-length and mini-length DNA barcoding for the identification of seafood commercially traded in Germany. Food Control. 2017; 73: 922-9.

[33] Muñoz-Colmenero M, Juanes F, Dopico E, Martinez JL, Garcia-Vazquez E. Economy matters: A study of mislabeling in salmon products from two regions, Alaska and Canada (Northwest of America) and Asturias (Northwest of Spain). Fish Res. 2017; 195: 180-5.

[34] Wen J, Tinacci L, Acutis PL, Riina MV, Xu Y, Zeng L, Ying X, Chen Z, Guardone L, Chen D, Sun Y, Zhao J, Guidi A, Armani A. An insight into the Chinese traditional seafood market: Species characterization of cephalopod products by DNA barcoding and phylogenetic analysis using COI and 16SrRNA genes. Food Control. 2017; 82: 333-42.

[35] Xiong X, Guardone L, Cornax MJ, Tinacci L, Guidi A, Gianfaldoni D, Armani A. DNA barcoding reveals substitution of Sablefish (*Anoplopoma fimbria*) with

Patagonian and Antarctic Toothfish (*Dissostichus eleginoides* and *Dissostichus mawsoni*) in online market in China: How mislabeling opens door to IUU fishing. Food Control. 2016; 70: 380-91.

[36] Ferrito V, Bertolino V, Pappalardo AM. White fish authentication by COIBar-RFLP: Toward a common strategy for the rapid identification of species in convenience seafood. Food Control. 2016; 70: 130-7.

[37] Chin Chin T, Adibah AB, Danial Hariz ZA, Siti Azizah MN. Detection of mislabelled seafood products in Malaysia by DNA barcoding: Improving transparency in food market. Food Control. 2016; 64: 247-56.

[38] Harris DJ, Rosado D, Xavier R. DNA barcoding reveals extensive mislabeling in seafood sold in Portuguese supermarkets. J Aqua Food Pro Technol. 2016; 25: 1375-80.

[39] Xiong X, Guardone L, Giusti A, Castigliego L, Gianfaldoni D, Guidi A, Armani A. DNA barcoding reveals chaotic labeling and misrepresentation of cod (鳕, Xue) products sold on the Chinese market. Food Control. 2016; 60: 519-32.

[40] Nagalakshmi K, Annam P-K, Venkateshwarlu G, Pathakota G-B, Lakra WS. Mislabeling in Indian seafood: An investigation using DNA barcoding. Food Control. 2016; 59: 196-200.

[41] Guardone L, Tinacci L, Costanzo F, Azzarelli D, D'Amico P, Tasselli G, Magni A, Guidi A, Nucera D, Armani A. DNA barcoding as a tool for detecting mislabeling of fishery products imported from third countries: An official survey conducted at the Border Inspection Post of Livorno-Pisa (Italy). Food Control. 2017; 80: 204-16.

[42] Carvalho DC, Guedes D, da Gloria Trindade M, Coelho RMS, de Lima Araújo PH. Nationwide Brazilian governmental forensic programme reveals seafood mislabelling trends and rates using DNA barcoding. Fish Res. 2017; 191: 30-5.

[43] Willette DA, Simmonds SE, Cheng SH, Esteves S, Kane TL, Nuetzel H, Pilaud N, Rachmawati R, Barber PH. Using DNA barcoding to track seafood mislabeling in Los Angeles restaurants. Conserv Biol. 2017: 31, 1076-85.

[44] Fernandes TJR, Costa J, Oliveira MBPP, Mafra I. An overview on fish and shellfish allergens and current methods of detection. Food Agri Immunol. 2015; 26: 848-69.

[45] Giuffrida MG, Villalta D, Mistrello G, Amato S, Asero R. Shrimp allergy beyond Tropomyosin in Italy: clinical relevance of Arginine Kinase, Sarcoplasmic calcium binding protein and Hemocyanin. Eur Ann Allergy Clin Immunol. 2014; 46: 172-7.

[46] Kamath SD, Rahman AMA, Voskamp A, Komoda T, Rolland JM, O'Hehir RE, Lopata AL. Effect of heat processing on antibody reactivity to allergen variants and

fragments of black tiger prawn: A comprehensive allergenomic approach. Mol Nutr Food Res. 2014; 58: 1144-55.

[47] Kobayashi A, Tanaka H, Hamada Y, Ishizaki S, Nagashima Y, Shiomi K. Comparison of allergenicity and allergens between fish white and dark muscles. Allergy. 2006; 61: 357-63.

[48] Nwaru BI, Hickstein L, Panesar SS, Muraro A, Werfel T, Cardona V, Dubois AEJ, Halken S, Hoffmann-Sommergruber K, Poulsen LK, Roberts G,Van Ree R, Vlieg-Boerstra BJ, Sheikh A. The epidemiology of food allergy in Europe: A systematic review and meta-analysis. Allergy: Eur J Allergy Clin Immunol. 2014; 69: 62-75.

[49] Chafen JJS, Newberry SJ, Riedl MA, Bravata DM, Maglione M, Suttorp MJ, Sundaram, V.,Paige NM, Towfigh A, Hulley BJ, Shekelle PG. Diagnosing and managing common food allergies: A systematic review. J Am Med Assoc. 2010; 303: 1848-56.

[50] Sicherer SH, Sampson HA. Food allergy. J Allergy Clin Immunol. 2010; 125: 116-25.

[51] CODEX STAN 1–1985. Amended in 1991, 1999, 2001, 2003, 2005, 2008 and 2010 regarding the general standard for the labelling of pre-packaged foods. FAO/WHO Standards, Off Codex Stand.Rome, Italy. Available at:http://www.codexalimentarius.org/standards/list-of-standards. (Last accessed 18 December 2017)

[52] Kuehn A, Swoboda I, Arumugam K, Hilger C, Hentges F. Fish allergens at a glance: Variable allergenicity of parvalbumins, the major fish allergens. Front Immunol. 2014; 5.

[53] Tsabouri S, Triga M, Makris M, Kalogeromitros D, Church MK, Priftis KN. Fish and shellfish allergy in children: Review of a persistent food allergy. Pediatr Allergy Immunol. 2012; 23: 608-15.

[54] ALLERGEN, official site for the systematic allergen nomenclature. Available at: http://www.allergen.org/ (Last accessed 18 December 2017).

[55] Griffiths AM, Sotelo CG, Mendes R, Pérez-Martín RI, Schröder U, Shorten M, Silva, Helena A, Verrez-Bagnis V, Mariani S. Current methods for seafood authenticity testing in Europe: Is there a need for harmonisation? Food Control. 2014; 45: 95-100.

[56] Espiñeira M, Santaclara FJ. The Use of Molecular Biology Techniques in Food Traceability. In Espiñeira M, Santaclara FJ, editors. Advances in Food Traceability Techniques and Technologies: Cambridge: Woodhead Publishing; 2016. p. 91-118.

[57] Sathe SK, Liu C, Zaffran VD. Food Allergy. Ann Rev Food Sci Technol. 2016; 7: 191-220.

[58] Costa J, Fernandes TJR, Villa C, Oliveira MBPP, Mafra I. Advances in food allergen analysis. In Spizzirri U.G., Cirillo G, editors. Food Safety: Innovative Analytical Tools for Safety Assessment. Beverly: Scrivener Publishing LLC (Wiley). 2017. p. 305-60.
[59] Lago FC, Alonso M, Vieites JM, Espiñeira M. Fish and Seafood authenticity-species identification. In Boziaris IS, editor. Seafood Processing: Technology, Quality and Safety. New Jersey: John Wiley & Sons, Ltd. 2014, p. 419-52.
[60] Mafra I, Ferreira IMPLVO, Oliveira MBPP. Food authentication by PCR-based methods. Eur Food Res Technol. 2008; 227: 649-65.
[61] Salihah NT, Hossain MM, Lubis H, Ahmed MU. Trends and advances in food analysis by real-time polymerase chain reaction. J Food Sci Technol. 2016; 53: 2196-209.
[62] Scarano D, Rao R. DNA Markers for food products authentication. Diversity. 2014; 6: 579.
[63] Applewhite L, Larkin P, Naaum AM. Species identification using other tools. In Hanner, R, Naaum A, editors. Seafood authenticity and traceability. San Diego: Academic Press; 2016. p. 133-48.
[64] Lo Y-T, Shaw P-C. DNA-based techniques for authentication of processed food and food supplements. Food Chem. 2018; 240: 767-74.
[65] Fernandes TJR, Silva CR, Costa J, Oliveira MBPP, Mafra I. High resolution melting analysis of a COI mini-barcode as a new approach for Penaeidae shrimp species discrimination. Food Control. 2017; 82: 8-17
[66] Wilwet L, Jeyasekaran G, Shakila RJ, Sivaraman B, Padmavathy P. A single enzyme PCR-RFLP protocol targeting 16S rRNA/tRNAval region to authenticate four commercially important shrimp species in India. Food Chem. 2018; 239: 369-76.
[67] Zagon J, Schmidt J, Schmidt AS, Broll H, Lampen A, Seidler T, et al. A novel screening approach based on six real-time PCR systems for the detection of crustacean species in food. Food Control. 2017; 79: 27-34.
[68] Navarro E, Serrano-Heras G, Castaño MJ, Solera J. Real-time PCR detection chemistry. Clin Chim Acta. 2015; 439: 231-50.
[69] Druml B, Cichna-Markl M. High resolution melting (HRM) analysis of DNA – Its role and potential in food analysis. Food Chem. 2014; 158: 245-54.
[70] McGlauflin MT, Smith MJ, Wang JT, Young SF, Chen N, Lee YC, Pascal C, Seeb LW, Stevens J, Seeb JE. High-resolution melting analysis for the discovery of novel single-nucleotide polymorphisms in rainbow and cutthroat trout for species identification. Trans Am Fish Soc. 2010; 139: 676-84.

[71] Meistertzheim A-L, Héritier L, Lejart M. High-Resolution Melting of 18S rDNA sequences (18S-HRM) for discrimination of bivalve's species at early juvenile stage: application to a spat survey. Mar Biol. 2017; 164: 133.

CHAPTER 1. SEAFOOD SPECIES AUTHENTICITY

STATE-OF-THE-ART

DNA barcoding applied to seafood authentication,
Trends in Food Science and Technology, **2017**, submitted.

EXPERIMENTAL PART

DNA barcoding coupled to HRM analysis as a new and simple tool for the authentication of Gadidae fish species
Food Chemistry, **2017**, 230, 49-57.

High resolution melting analysis of a COI mini-barcode as a new approach for Penaeidae shrimp species discrimination.
Food Control, **2017**, 82, 8-17.

COI barcode-HRM as a novel approach for the discrimination of hake species.
Fisheries Research, **2018**, 197, 50-59.

STATE-OF-THE-ART

DNA barcoding applied to seafood authentication

Trends in Food Science and Technology, **2017**, submitted.

ELSEVIER

Trends in Food Science & Technology , submitted

TRENDS IN FOOD SCIENCE & TECHNOLOGY

Review

DNA barcoding applied to seafood authentication

Telmo J.R. Fernandes, M. Beatriz P.P. Oliveira, Isabel Mafra*

REQUIMTE-LAQV, Faculdade de Farmácia, Universidade do Porto, Rua de Jorge Viterbo Ferreira, 228, 4050-313 Porto, Portugal

*Corresponding author: Tel: +351 220428640. Fax: +351 226093390. E-mail: isabel.mafra@ff.up.pt

Abstract

Background: The world's seafood supply and trade has remarkably increased in the last decades, as well as the potential for species substitution on the commercial market. Indeed, seafood safety and authenticity became central issues, crucial for the identification of improper labelling of fish and shellfish in processed foods.

Scope and Approach: To detect and prevent mislabelling issues, species identification using DNA barcodes has been widely used as effective molecular markers. Therefore, this review intends to present the current status on the application of DNA barcoding to seafood species authentication. In this regard, the barcode regions, reference databases and methodologies are described, while applications are listed and summarised.

Key Findings and Conclusions: Cytochrome c oxidase subunit I (COI) gene has been the preferential targeted DNA region in animal species, including fish and shellfish, though other mitochondrial (*cytb*, 12S rRNA, 16S rRNA) and nuclear genes have been used. DNA barcode sequencing based on Sanger's method is the most used approach for seafood authentication. In recent years, noteworthy progresses have been done towards DNA barcoding strategies, involving next generation sequencing and high resolution melting analysis applied to seafood species discrimination.

Keywords: DNA barcodes; fish; shellfish; DNA markers; next generation sequencing; high resolution melting.

Introduction

Fish and shellfish species are widely consumed as foods due their nutritional valuable and sensorial characteristics, being presently the most traded food commodities in the world. There is a wide range of scientific reports describing the health benefits associated to seafood ingestion, such as the prevention of cardiovascular diseases and cancer, which are mainly related to the high contents of omega-3 polyunsaturated fatty acids (PUFA) (Larsen et al., 2011).

Over the last 50 years, there has been a steady increase of seafood capture and aquaculture production related with the remarkable growth on seafood consumption worldwide (Angers et al., 2017). In fact, European Union (EU) consumers ate, on average, 1 kg more of fish in 2014 than in 2013, which represented an consumption increase of almost 650,000 tonnes (+4.5%) between the two years (EUMOFA, 2016). The notable expand of seafood market has been followed by a growing number of cases of mislabelling and species substitution, raising a recurrent problem in seafood product value chains (Angers et al., 2017; Clark, 2015). Actually, the potential for fraud, particularly, the substitution of high-valued seafood species lower cost ones has increased over the years. Seafood mislabelling is a specific sensitive issue because human health might be threatened by exposing consumers to hidden or undeclared species related to the presence of contaminants, such as toxins, pathogens, parasites, chemicals or allergens (Cohen et al., 2009; Sicherer et al., 2011). Misidentification of seafood species at capture might also be responsible for mislabelling and has also been pointed out as one of the causes of overexploiting certain species (Muñoz-Colmenero et al., 2015).

Whenever it is possible, morphological characteristics may be used to identify seafood species. Species-specific features such as size, texture and coloration can be used to differentiate species, often helped by guides with photographs, drawings and descriptions of what can be visually determined. Nevertheless, discrimination of fish or shellfish species based on visual examination often requires expert training and, when morphological characteristics are removed, which frequently occurs upon food processing, visual identification of species becomes a difficult or impossible task to accomplish (Muñoz-Colmenero et al., 2015; Naaum & Hanner, 2016).

In the EU, for the control and reduction of economic fraud, through species substitution and/or mislabelling, and for consumer protection the Regulation (EU) No 1379/2013 establishes the mandatory information of fishery and aquaculture products for human consumption, which must be labelled with the commercial and scientific names, fishing and production methods, catch area and the fishing gear. Additional, the Regulation (EU) No 1379/2013, on the provision of food information to consumers, establishes the

mandatory labelling of fish, crustaceans and molluscs, regardless of their amount because on being considered as potential allergenic foods to sensitised individuals.

Therefore, to verify the authenticity of seafood products and the compliance with the labelling legislation, several analytical methodologies have been developed and reported in the literature in the last decades. DNA-based methods have provided powerful tools for seafood species differentiation, being widely implemented regarding authentication purposes (Naaum & Hanner, 2016). Particularly, DNA-barcoding has been considered an important policy instrument for species identification, addressing the referred authenticity issues (Clark, 2015). DNA barcoding is considered an increasingly fashionable and novel concept, creating confidence in enhanced biodiversity assessment. It has been argued as an accurate, rapid, and inexpensive approach for species identification, which has caught the attention of taxonomists, geneticists, and evolutionary biologists around the world (Hellberg et al., 2016). DNA barcoding proposes that short sequences can be used for species discrimination. While several studies use DNA sequences from a variety of genes for species discrimination, DNA barcoding suggests a single gene to be used as the basis of a global identification system for animals. For most species groups, the standard DNA barcode is a region of ~650 bp of the cytochrome c oxidase subunit I (COI) gene, which is relatively conserved within species, showing sufficient variation among species to allow for differentiation (Hebert et al., 2003; Ward et al., 2009). However, DNA barcodes targeting other mitochondrial genes and nuclear DNA (nDNA) have also showed their effectiveness for molecular species differentiation (Ardura et al., 2013).

In this review, the current status on the application of DNA barcoding to seafood species authentication is described, focusing on suitable barcode *loci*, reference databases and methodologies, from the classical Sanger to the promising next generation sequencing (NGS) and high resolution melting (HRM) approaches.

Barcode regions

The concept of DNA barcode was firstly proposed and developed by Paul Hebert and his colleagues (Hebert et al., 2003) that named short nucleotide stretches with sequence variation as “barcodes”. Nevertheless, before that, these short nucleotide regions had already been used for species authentication although the actual concept of DNA barcode was not yet established. For instance, forensically informative nucleotide sequence (FINS) proposed by Bartlett & Davidson (1992) is based on the combination of DNA sequencing with phylogenetic analysis and genetic distance calculation using DNA barcodes (Hellberg et al., 2016). Contrarily to DNA barcoding, in FINS there is the requirement of using previously authenticated species material for further methodology development.

In DNA barcoding after DNA extraction, the amplification of specific DNA regions by polymerase chain reaction (PCR) is followed by sequencing of the produced amplicon. The main difference in comparison with FINS consists on the final comparison of the DNA sequence of the amplicons with a database of reference samples (Hellberg & Morissey, 2011) (Figure 1). Therefore, it is possible to assess the species recognition or possibly found new species barcodes if there is not any match with the database (Angers et al., 2017). A 648 bp mitochondrial COI gene (also referred as the Folmer region) was primarily suggested as a standard DNA barcode region for metazoans, being also tested in birds species with a success rate of 98-100% (Hebert et al., 2003; Hebert et al., 2004). Along the past 15 years, COI was settled as a universal barcode of kingdom Animalia (Scarano & Rao, 2014). Nonetheless, several DNA barcodes concerning other mitochondrial/nDNA *loci* have been established as suitable genetic markers, allowing the identification and differentiation of numerous seafood species. As it was previously mentioned, the nucleotide sequence of an appropriate DNA barcode should be easily amplified across a large range of related individuals and polymorphic enough to discriminate among species, yet showing low intraspecies variability (Hellberg & Morissey, 2011; Lago et al., 2014).

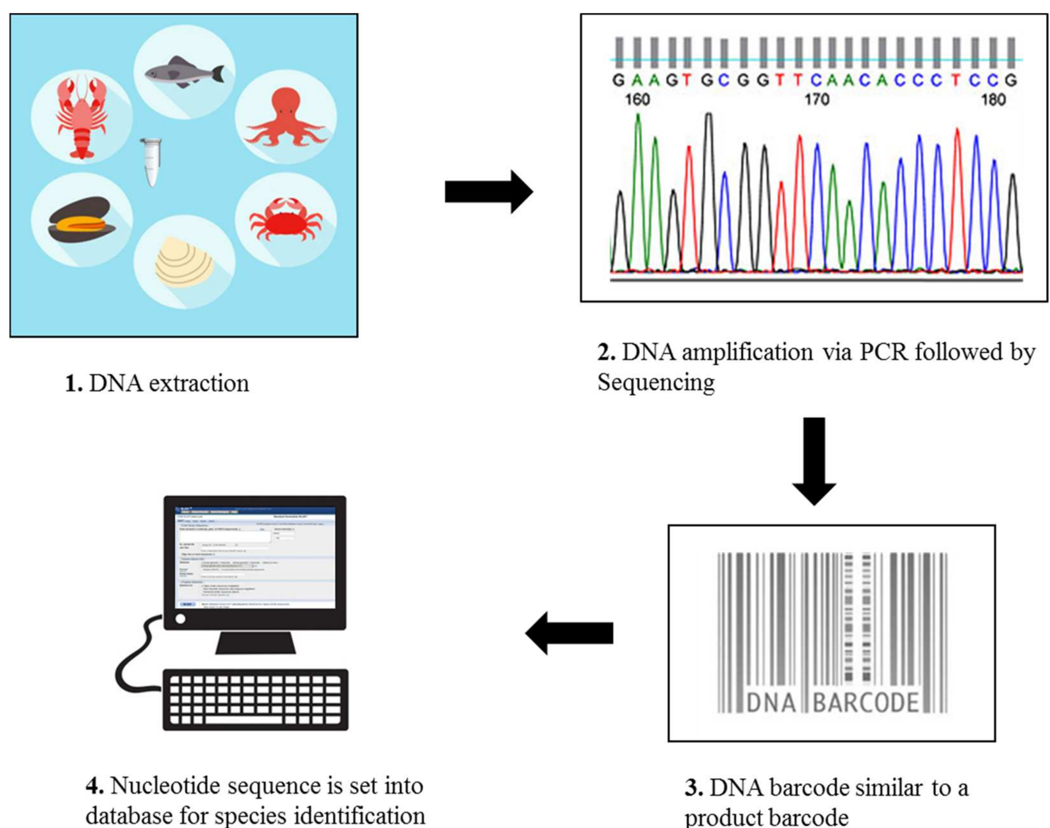


Figure 1. Squematic representation of DNA barcoding major steps for seafood species identification.

Moreover, the length of amplicon must be appropriate for the type of tested food because some of the processing and preservation methods used with seafood products can reduce DNA integrity, making the longer fragments difficult to recover from highly processed seafood. This can be a limitation when analysing full-length barcodes (i.e., ~ 650 bp) from severely processed foods. In these cases, the DNA is usually degraded into DNA fragments with sizes lower than 350 bp (Naaum & Hanner, 2016). Shorter regions of standardized gene targets (100-300 bp) are also used as barcodes and often named as DNA mini-barcodes, being particularly suitable to discriminate species in severely processed foods (Armani et al., 2015a; Fernandes et al., 2017a, 2017b, 2018; Günther et al., 2017; Lo & Shaw, 2018; Mitchell & Hellberg, 2016; Shokralla et al., 2015; Tomás et al., 2017). When shorter DNA regions are used, the requirements of the DNA barcode should be fulfilled and, at the same time identifying small nucleotide regions with high homology among different species for primer design.
Nonetheless, regardless the DNA target region or amplicon size, the DNA barcode should be selected based on the required level of resolution and on the target taxonomic groups to be discriminated.

Mitochondrial DNA

Mitochondrial DNA (mtDNA) has several advantages in comparison with nDNA for species identification. Firstly, mtDNA is inherited from the mother and is free of recombination, leading to minimum nucleotide variations within the same species (Hellberg et al., 2016; Scarano & Rao, 2014). Additionally, mitochondrial genome presents a high mutation rate and includes multiple copies, comparatively with the single copy of nDNA present in each cell. These facts are particularly relevant when analysing degraded and small samples (Angers et al., 2017; Clark, 2015; Hellberg et al., 2016; Paracchini et al., 2017). A large variety of mitochondrial genetic targets has been investigated for seafood species identification, being COI, cytochrome b (*cytb*) and 16S rRNA the most used barcode regions over the last years. Tables 1 and 2 summarise the relevant information of reports regarding the application of DNA barcoding to seafood authentication, from which it can be highlighted that COI is the top barcode used. The great number of available sequences for COI gene (~650 bp) in the available databases makes this region the preferential choice for seafood species differentiation. For instance, more than 100,000 COI barcodes have already been deposited on Fish Barcode of Life Initiative website (FISH-BOL, 2017), representing more than 11,000 species.

Table 1. Resumed information on the reports of the application of DNA barcoding based on sequencing analysis for seafood species authentication.

Target *locus/loci*	Amplicon(s) size	Sample collection	Identified species (scientific name)	Collection location	Reference
COI gene	nr	60 seafood species including fish, crustaceans and bivalves	e.g.: *Gadus macrocephalus, Merluccius capensis, Thunnus albacares, Salmo salar, Litopenaeus vannamei, Cancer pagarus, Mytilus chilensis.*	Portugal	Harris et al. (2016)
COI gene	313, 652, 658 bp	118 seafood products including fish and shrimp species	e.g.: *Salmo salar, Theragra chalchogramma, Oncorchynkus mykiss, Clupea harengus, Crangon crangon, Pandalus borealis.*	Germany	Günther et al. (2017)
COI gene	nr	11 reference Gadiforme fish species 70 dried salted cod fillets 40 battered cod chunks	*Gadus morhua, G. macrocephalus, Pollachius virens, Brosme brosme.*	Italy	Di Pinto et al. (2013)
COI gene	650, 190 bp	75 Sparidae fish species: 80 fresh fish specimens 200 ethanol-preserved reference tissues	e.g.: *Boops boops, Evynnis cardinalis, Pagellus bellottii, Sarpa salpa, Dentex angolensis, Pagrus auratus, Stenotomus caprinus.*	Europe	Armani et al. (2015a)
COI gene	nr	178 samples	*Salmo salar, Oncorhyncus keta, O. kisutch, O. tshawytscha, O. gorbuscha, O. nerka, O. masou.*	USA	Cline (2012)
COI gene	700, 150 bp	62 seafood products	e.g.: *Oncorhyncus gorbuscha, Gadus macrocephalus, Salmo salar, Crassostrea angulate, Priancanthus macranthus, Anguilla Anguilla.*	Malaysia	Chin et al. (2016)
COI gene	650 bp	364 sushi samples	*Thunnus* spp., *Seriola lalandi, Hippoglossus hippoglossus, H. stenolepis, Scomber* spp., *Scomberomorus* spp., *Salmo salar, Oncorhynkus* spp.	USA	Willette et al. (2017)
COI gene	479-692 bp	255 fish products	e.g.: *Genypterus blacodes, Thunnus albacares, Gadus morhua, Micropogonias furnieri, Pagrus pagrus, Orechromis niloticus.*	Brazil	Carvalho et al. (2017a)
COI gene	577 bp	137 fish fillets	*Macrodon ancyclodon, Lutjanus pupureus, L. campechanus, Plagioscion squamosissimus, P. auratus, Nebris microps, Cephalopolis fulva, Menticirrhus americanus, Cynoscion microlepidotus, C. jamaicensi.s*	Brazil	de Brito et al. (2015)
COI gene	604-625 bp	90 fish samples	*Lates niloticus, Pangasionodon hypophthalmus, Oreochromis niloticus.*	Egypt, Vietnam	Galal-Khallaf et al. (2014)

Table 1. (continued)

Target *locus/loci*	Amplicon(s) size	Sample collection	Identified species (scientific name)	Collection location	Reference
COI gene	>600 bp	63 catfish samples	*Brachyplatystoma platynemum, Genidens barbus, Pseudoplatystoma fasciatum, P. tigrinum, P. corruscans, Cynoscion jamaicensis, C. virescens.*	Brazil	Carvalho et al. (2011)
COI gene	≥530 bp	51 fish products	e.g.:*Thunnus albacares, Merluccius paradoxus, Lates calcifer, Genypterus blacodes, Salmo salar, Oncorhyncus mykiss, Mustelus antarcticus, Macruronus magellanicus.*	Tasmania	Lamendin et al. (2015)
COI gene	nr	120 hake food products	*Merluccius gayi, M. hubbsi, M. capensis, Theragra chalcogramma, M. productus, M. senegalensis, M. paradoxus, Pleuronectes platessa, Lepidopsetta polyxystra.*	Italy	Di Pinto et al. (2016)
COI gene	nr	200 samples of fish fillets	*Epinephelus marginatu, E. diacanthus, Xiphias gladius, Lates niloticus, Pangasius hypophthalmus, P. fluviatilis, P. sanitwongsei, Prionace glauca, Thunnus obesus, Isurus oxyrinchus.*	Italy	Di Pinto et al. (2015)
COI gene	650 bp	11 crab species	*Atergatis integerrimus, Scylla tranquebarica, Charybdis lucífera, Portunus sanguinolentus, Scylla serrata, Charybdis feriata, Portunus pelagicus, Barytelphusa cuniculareis, Varuna literata, Charybdis helleri, Portunus reticulatus.*	India	Vartak et al. (2015)
COI gene	698, 705, 190 bp	68 seafood processed products from chinese and bangladeshi food markets	e.g.: *Engraulis japonicus, E. australis, Neotropius khavalchor, Trichiurus japonicus, Ommastrephes bartramii, Upeneus japonicu.s*	Italia	Armani et al (2015b)
COI gene	669-693 bp	534 seafood species	e.g.: *Tinca tinca, Silurus asotus, Cololabis saira, Channa maculata, Jasus lalandii, Litopenaeus vannamei.*	China	Shen et al. (2016)
COI gene	660 bp	20 flatfish species 40 flatfish fillet samples	*Solea solea, Pangassius hypophthalmus, Pleuronectes platessa, Platichthys flesus, Limanda limanda, Arnoglossus laterna.*	Italy	Pappalardo & Ferrito (2015a)
COI gene	652, 127-314 bp	88 fish species 44 seafood products	e.g.:*Clupea harengus, Melanogrammus aeglefinus, Sardina pilchardus, Salmo salar, Scomber scombrus, Thunnus obesus.*	USA	Shokralla et al. (2015)
COI gene	nr	12 seafood dishes from 6 restaurants	*Litopenaeus vannamei, Uroteuthis edulis, Thunnus albacares, Farfantepenaeus aztecus, Dissostichus mawsoni, Lates calcifer, Dissostichus elefinoides, Lujtanus futtatus, L. synagris, Thunnus obesus.*	USA	Stern et al. (2017)

Table 1. (continued)

Target *locus/loci*	Amplicon(s) size	Sample collection	Identified species (scientific name)	Collection location	Reference
COI gene	650 bp	172 seafood samples form restaurants	e.g.: *Dissostichus eleginoides, Pagrus major, Salmo salar, Thunnus albacares,Gadus morhua, Octopus vulgaris, Litopenaeus vannamei.*	USA	Khaskar et al. (2015)
COI gene	650 bp	28 convenience seafood products	e.g.: *Theragra chalcogramma, Gadus macrocephalus, G. ogac, Merluccius hubbsi, Semipterus mesoprion.*	England	Huxley-Jones et al. (2012)
COI gene	650 bp	34 fish samples	e.g.: *Stolephorus holodon, Spratelloides gracillis, Istiompax indica, Priacanthus hamrur, Sardinella lemurus, Huso clauricus.*	Taiwan	Chang et al. (2016)
COI gene	234-645 bp	22 specimens of frozen fish and fillets	e.g.: *Glyptocephalus stelleri, Pleuronectes platessa, Limanda áspera, Gadus morhua, Torichthys helleri, Oreochromis niloticus.*	Russia	Nedunoori et al. (2017)
COI gene	nr	19 seafood samples	e.g.: *Sardinella tawilis, Thunnus tonggol, T. albacares, Pangasianodon hypophthalmus, Penaeus monodon, Litopenaeus vannamei.*	Philippines	Maralit et al. (2013)
COI gene	650 bp	14 fish samples	*Scomberomorus niphonius, Pomadasys hasta, Argyrosomus hololepidotus, Parastromateus niger, Nenipterus japonicus.*	Iran	Changizi et al. (2013)
COI gene	655, 208-226 bp	6 types of fish with different cooking processes	*Salmo salar, Oreochromis* spp., *Thunnus* spp., *Decapterus macarellus, Theragra chalcogramma, Pangasianodon hypophthalmus.*	USA	Pollack et al. (2018)
COI gene	699-720, 485-655 bp 201-207, 134-137 bp	52 codfish samples	e.g.: *Macrourus carinatus, Atheresthes evermanni, Dissostichus mawsoni, Albatrossia pectoralis, Pollachius virens, Theragra chalcogramma.*	China	Xiong et al. (2016)
16S rRNA gene	151, 362, 213 bp	6 Penaeid shrimp species	*Litopenaeus vannamei, L. stylirostris, Penaeus monodon, P. semisulcatus, Fenneropenaeus indicus, F. merguiensis.*	Spain	Pascoal et al. (2011)
COI gene 16S rRNA gene	542-638 bp	104 shellfish products	e.g.: *Fenneropenaeus indicus, F. merguiensis, Litopenaeus vannamei, Solenocera* spp., *Solenocera crassicornis, Pleoticus* spp.	Egypt, Spain	Galal-Khallaf et al. (2016)
COI gene 16S rRNA gene	526-658 bp 503-513 bp	95 traditional processed cephalopod products	e.g.: *Sepia pharaonic, S. esculenta, S. recurvirostra, Dosidicus gigas, Amphioctopus marginatus, Uroteuthis chinensis.*	China	Wen et al. (2017)

Table 1. (continued)

Target *locus/loci*	Amplicon(s) size	Sample collection	Identified species (scientific name)	Collection location	Reference
cytb gene COI gene	360 bp 655 bp	58 seafood samples	e.g.: *Sardina pilchardus, Engraulis encrasicolus, Micromesistius poutassou, Phycis blennoides, Merluccius merluccius, Lophius piscatorius.*	Italy	Cutarelli et al. (2014)
COI gene 16S rRNA PEPCK	192-710 bp 152, 630 bp 644 bp	277 seafood samples comprising fish, crustacean, cephalopods and molusks	e.g.:*Epinephelus bleekerii, Psettodes bennetti, Breogadus saída, Merluccius paradoxus, Mustelus punctulatus, Seriola quinqueradiata.*	Italy	Guardone et al. (2017)
COI gene 16S rDNA *cob* gene	700 bp 500 bp 850 bp	37 seafood samples	e.g.: *Perna canaliculus, Merluccius hubbi, Salmo salar, Amphioctopus marginatus, Solea solea.*	Italy	Nicolé et al. (2012)
5S rRNA gene	74-544 bp	10 fish species	*Dicentrarchus labrax, Lates niloticus, Pangasius hypohthalmus, Perca fluviatilis, Pleuronectes platessa, Salmo salar, S. truta, Solea vulgaris, Thunnus thynnus, Xiphias gladius.*	Italy	Tognoli et al. (2011)
16S rRNA gene 12S rRNA gene	560 bp 490 bp	53 ray-finned fish species	e.g.:*Clupea harengus, Sardinops sagax, Engraulis encrasicolus, Merluccius capensis, M. paradoxus.*	South Africa	Cawthorn et al. (2012)
COI gene Rhodopsin gene	619-663 bp nm	44 fish species	e.g.: *Lutjanus argentimaculatus, Pinjalo pinjalo, Plotosus canius, Priachantus tayenus, Epinephelus tauvina.*	Indonesia	Abdullah & Rehbein (2017)
CR gene ITS1 gene	236 bp 179 bp	53 canned tuna fish products	*Thunnus alalunga, T. thynnus, Katsuwonus pelamis, T. tongol, Sarda orientalis.*	USA	Mitchell & Hellberg (2016)
16S rRNA gene COI gene	nr nr	4 cutlassfish species (1277 individuals)	*Tentoriceps cristatus, Trichiurus japonicus, T. nanhaiensis, T. lepturus.*	Taiwan	Wang et al. (2017)
COI gene 16S rRNA gene	655 bp 335 bp	22 Amazonian fish species	e.g.: *Astronotus ocellatus, Schizon fasciatus, Curimata inornata, Arapaima gigas, Oxydoras niger.*	Brazil	Ardura et al. (2010)
COI gene 16S rDNA gene	453, 649 bp nr	29 oyster species	e.g.: *Crassostrea nippona, C. rhizophorae, C. sikamea, C. virgnica, Saccostrea commercialis.*	China	Liu et al. (2011)
cytb gene	nr	Cod, tuna and sole species (245 samples)	e.g.: *Solea solea, Synaptura lusitânica, Cynoglossus senegalensis, Pangasianodon hypophthalmus.*	Germany	Kappel & Schröder (2015)

Table 1. (continued)

Target *locus/loci*	Amplicon(s) size	Sample collection	Identified species (scientific name)	Collection location	Reference
28S rDNA COI gene ITS-1 gene	800 bp 700 bp 400-600 bp	9 oyster species (10 individuals per species)	e.g.: *Crassostrea gigas, C. sikamea, C. virginica, C. corteziensis, Saccostrea palmula.*	Mexico, California, Cuba, Chile	Mazón-Suástegui et al. (2016)
COI gene Rhodopsin gene	500-652 bp 460 bp	10 billfish species (n = 296)	e.g.: *Istiompax indica, Makaira nigricans, Istiophorus platypterus, Kajikia audax, Tetrapturus angustirostris.*	nr	Hanner et al. (2011)

nr – not reported.

Table 2. Resumed information on the reports of the application DNA barcoding based on several molecular approaches (other than Sanger sequencing) for sequencing analysis for seafood species authentication.

Target *locus / loci*	Amplicon(s) size	Methodology	Sample collection	Identified species (scientific name)	Collection location	Reference
COI gene	651 bp	PCR-RFLP	54 seafood products	*Theragra chalcogramma, Merluccius productus, M. merluccius, Merluccius paradoxus.*	Italy	Ferrito et al. (2016)
COI gene	439 bp	PCR-RFLP	106 fish species	e.g.: *Gadus morhua, G. macrocephalus, Oncorhyncus keta, O. mykiss, Salmo salar.*	USA, UK	Mueller et al. (2015)
COI gene	654 bp	PCR-RFLP	15 specimens of European anchovy 50 seafood products	*Sardinella aurita, Sardina pilchardus, Engraulis japonicus, E. encrasicolus.*	Italy	Pappalardo & Ferrito (2015b)
16S rRNA/tRNA^{Val} gene	530 bp	PCR-RFLP	4 Penaeidae species Frozen, cooked, canned, fried and raw tretaments.	*Litopenaeus vannamei, Penaeus monodon, P. semisulcatus, Fenneropenaeus indicus.*	India	Wilwet et al. (2018)
16S rRNA/tRNA^{Val} gene	530 bp	PCR-RFLP	41 shrimp and prawn food products, Penaeidae species	e.g.: *Litopenaeus vannamei, Fenneropenaeus indicus, Pleoticus muelleri, Penaeus semisulcatus, L. setiferus.*	Spain	Pacoal et al. (2008a)
COI /II gene 16S rDNA gene	nd 561-562 bp	PCR-RFLP	5 Penaeid shrimp species	*Penaeus semisulcatus, P. monodon, Fenneropenaeus merguiensis, Metapenaeus japonicus, Litopenaeus vannamei.*	Thailand	Khamnamtong et al. (2005)
16S rRNA/tRNA^{Val} gene	960 bp	PCR-RFLP	19 Penaeid shrimp species	e.g.: *Litopenaeus vannamei, Farfantepenaeus notialis, F. brasiliensis, Fenneropenaeus indicus, Penaeus monodon.*	Spain	Pascoal et al. (2008b)
Pantophysin gene *cytb* gene	201-202 bp 464 bp	PCR-RFLP	20 samples of frozen fish products	*Merluccius gayi, M. productus, M. hubbsi*	Czech Republic	Hubalková et al. (2009)

Table 2. (continued)

Target *locus / loci*	Amplicon(s) size	Methodology	Sample collection	Identified species (scientific name)	Collection location	Reference
cytb gene Parvalbumin gene	464 bp 227-670 bp	PCR-SSCP PCR-RFLP	27 authentic fish samples from Scombridae family	e.g.: *Thunnus alalunga, T. tonggol, T. albacares, T. maccoiyii, T. obesus.*	Indonesia	Abdullah & Rehbein (2016)
16S rRNA gene ND2 COI gene	30 bp 30 bp nd	Pyrosequencing	116 fish products	e.g.: *Engraulis encrasicolus, Sardinops sagax, Limanda limanda, Solea solea, S. senegalensis.*	Italy	De Battisti et al. (2014)
16S r RNA gene COI gene	203-288 bp 444 bp	Pyrosequencing	293 bivalve samples	e.g.: *Flexopecten glaber, Aeqyuoecten opercularis, Mytillus galloprovincialis, M. edulis, Crassostrea gigas.*	Italy	Abbadi et al. (2017)
cytb gene	126, 131 bp	Illumina MiSeq	9 standard mixtures of tuna species	*Thunnus alalunga, T. albacares, T. obesus, Katsuwonus pelamis.*	Germany	Kappel et al. (2017)
cytb gene COI gene	100-150 bp	Ion torrent PGM	22 codfish products	*Gadus macrocephalus, Theragra chalcogramma, Micromesistius australis, Pollachius virens, Gadus morhua, Oreochromis niloticus.*	Brazil	Carvalho et al. (2017b)
16S rRNA gene	250-260 bp 190-200 bp	Ion torrent PGM	16 surimi-based products	*Theragra chalcogramma, Merluccius merluccius, Todarodes pacificus, Melanogrammus aeglefinus, Architeuthis dux*	Italy, Spain	Giusti et al. (2017)
18S rDNA gene	220 bp	HRM analysis	6 bivalve species	*Crassostrea gigas, Ostrea edulis, Anomia ephippium, Modiolus barbatus, Mytilus edulis, M. galloprovincialis.*	France	Meistertzheim et al. (2017)
16S rRNA gene	83 bp	HRM analysis	95 oysters	*Crassostrea angulata, C. hongkongensis, C. ariakensis, C. gigas, C. sikamea.*	China	Jin et al. (2015)
COI gene	655, 100 bp	HRM analysis	33 codfish samples	*Gadus morhua, G. macrocephalus, Theragra chalcogramma.*	Portugal	Tomás et al. (2017)

Table 2. (continued)

Target *locus / loci*	Amplicon(s) size	Methodology	Sample collection	Identified species (scientific name)	Collection location	Reference
COI gene	106 , 316, 556 bp	HRM analysis	5 Penaeidae shrimp species 33 processed foodstuffs	*Litopenaeus vannamei, Penaeus monodon, Fenneropenaeus indicus, Metapenaeus affinis, M. kerathurus.*	Portugal	Fernandes et al. (2017a)
COI gene *cytb* gene	400, 157 bp 663, 134 bp	HRM analysis	4 Gadidae fish species 42 fish-containing foods	*Gadus morhua, G. macrocephalus, Theragra chalcogramma, Pollachius virens.*	Portugal	Fernandes et al. (2017b)
COI gene	400, 102 bp	HRM analysis	5 Merluccidae fish species 45 processed seafood products	*Merluccius merluccius, M. productus, M. hubbsi, M. capensis, M. paradoxus.*	Portugal	Fernandes et al. (2018)
COI gene	540, 217 bp	HRM analysis	4 Macrouridae fish species	*Macrourus carinatus, M. whitsoni, M. holotrachys, Macrourus* spp.	UK	Fitzcharles (2012)
PAPM	nr	HRM analysis	471 mussels	*Mytilus chilensis, M. edulis, M. galloprovincialis.*	Chile, Canada, Mexico, New Zealand	Jilberto et al. (2017)

nr – not reported.

Several authors have used COI sequences for multiple sample assays testing the compliance of seafood labelling. Harris et al. (2016) demonstrated that COI gene was useful to identify clearly the species of the great majority of the seafood samples (e.g. *Gadus macrocephalus*, *Merluccius capensis*, *Salmo salar*, *Cancer pagarus*). Almost 19% of the examined seafood samples were mislabelled, being this fraud more common among crustacean and bivalve foodstuffs (2/20 fish and 9/39 invertebrates) (Harris et al., 2016). Following the same trend, Günther et al. (2017), Cline (2012), Carvalho et al. (2017a), Armani et al. (2015b), Kashkar et al. (2015), Galal-Khallaf et al. (2016), Guardone et al. (2017) also demonstrated the great discriminatory power of COI gene in large numbers of seafood samples (Table 1).

The *cytb* gene has also proved to be a useful target for the identification of seafood species and for resolving species phylogenies (Imoto et al., 2013). Few works reported its use as a single barcode for seafood species discrimination (Kappel & Schröder, 2015; Kappel et al., 2017). Conversely, this barcode is often used in multi-*loci* approaches as a complementary target, as in the case of species differentiation in a restricted taxonomic seafood group included in a large collection of specimens. Cutarelli et al. (2014) performed a bi-directional sequence analysis of COI and *cytb* for commercial fraud detection, reporting a similar level of species identification in both barcodes (58/58 and 56/58, respectively), allowing the detection of *Gadus morhua* (Atlantic cod) replacement by *Pollachius virens* (saithe) in frozen cod fillets (Cutarelli et al., 2014). Fernandes et al. (2017b) reported that *cytb* mini-barcode was more efficient than COI in the differentiation of four Gadidae species (*G. morhua, G. macrocephalus, Theragra chalcogramma, Pollachius virens*) using HRM analysis (Fernandes et al., 2017b).

Although *cytb* and COI have demonstrated to be effective barcodes for seafood species identification, some limitations can still be pointed out. The lack of sufficient polymorphism to differentiate closely related species based on the genetic distance was referred by Hellberg et al. (2016). DNA barcodes based on the mitochondrial 12S and 16S rRNA genes are suitable alternatives within the mitochondrial genome to identify a wide range of fish and shellfish species in processed foods. These markers contain internal regions that are strongly conserved across taxa, suitable for universal primer designing, being alternated with short hypervariable regions that may be species-specific (Staats et al., 2016).

Pascoal et al. (2008 a, b) targeted the 16S rRNA gene for shrimp species differentiation using PCR-RFLP (restriction fragment length polymorphisms), both in fresh and frozen whole specimens and commercial food products, allowing the discrimination of 19 penaeid shrimp species. Jin et al. (2015) targeted the same gene, but using HRM analysis that showed its discriminatory potential for the identification of five commercially

relevant *Crassostrea* oysters (Table 2). Similarly to *cytb*, the 12S and 16S rRNA genes have been currently used in multi-target approaches as described by Galal-Khallaf et al. (2016), Wen et al. (2017), Guardone et al. (2017), Wang et al. (2017), Liu et al. (2011), among others (Tables 1 and 2).

Other mtDNA regions have also been targeted in seafood species differentiation, namely the control region (CR) (Mitchell & Hellberg, 2016), the apocytochrome b (*cob*) gene (Nicolé et al., 2012) and the NADH dehydrogenase subunit 2 (ND2) gene (De Battisti et al., 2014) (Table 1).

Nuclear DNA

Besides mtDNA, sequences of nDNA have been successfully targeted for the identification of seafood species (Paracchini et al., 2017; Tognoli et al., 2011) (Tables 1 and 2). Tognoli et al. (2011) were able to detect mislabelling cases in European fish markets, reporting the detection of 10 fish species targeting the 5S rRNA that proved to be suitable for easy and rapid seafood identification. This gene consists of a small conserved region (120 bp) coding for 5S rRNA and a variable sequence of noncoding DNA, designated as nontranscribed spacer (NTS) (Tognoli et al., 2011).

The intronless nuclear rhodopsin gene has also been suggested as an additional marker for fish species identification in combination with mtDNA targets. Abdullah and Rehbein (2014, 2017) and Hanner et al. (2011) demonstrated the usefulness of this gene as a molecular target for fish species differentiation, particularly tuna, Indonesian commercial fish and billfish species, respectively.

The gene for the integral membrane pantophysin is another studied barcode for fish species identification (Mariani & Bekkevold, 2014). Although its physiological function remains unknown, pantophysin has been used as a DNA marker in some Atlantic cod studies (Glover et al., 2011; Pampoulie et al., 2012). Genomic sequences of the partial pantophysin gene were also used as complementary markers to *cytb* gene using PCR-RFLP and sequencing to differentiation of hake species (*Merluccius gayi*, *M. productus*, *M. hubbsi*) (Hubálková et al., 2009).

Recently, Mazón-Suástegui et al. (2016) explored 28S rRNA as an alternative barcode for oyster species identification, presenting clear advantages in terms of simplicity and confidence when compared to COI and nuclear internal-transcribed spacer 1 (ITS-1) genes. Indeed, ITS-1 gene was also targeted by Micthell et al. (2016) along with mitochondrial CR gene for tuna species discrimination, as mentioned in the previous subsection. Other nDNA targets can be found in the literature included in barcoding methodologies, such as parvalbumin, polyphenolic adhesive protein of mussels (PAPM),

phosphoenolpyruvate carboxykinase (PEPCK) and 18S rDNA genes as reported by Abdullah and Rehbein (2016), Jilberto et al. (2017), Guardone et al. (2017), and Meistertzheim et al. (2017), respectively (Tables 1 and 2).

Reference databases

The strength of DNA barcode-based technologies relies on the availability of comprehensive and consistent databases, also referred as DNA sequence libraries (Galimberti et al., 2013; Hellberg et al., 2016). In order to assess the identity of a specimen or sample at the species level, the database must contain reference sequences representing that species, being particularly challenging in the case of seafood due to the worldwide huge number of species. Furthermore, the morphological similarities of many seafood species gives the rise to the possibility of misidentification and erroneous sequence information (Hellberg & Morissey, 2011).

A large amount of sequence data is available at the public database GenBank, which is included in the NCBI (National Center for Biotechnology Information) website (NCBI, 2017). NCBI is a very useful source of information, with multiple biological databases (sequence analysis, proteins, genes and expression, taxonomic, publications, DNA and RNA, etc.). A particularly advantageous tool accessible at the GenBank website is the software BLAST (Basic Local Alignment Search Tool) (Altschul et al., 1990), which enables the comparison of the experimentally obtained sequences with the available sequences at the database automatically by the alignment of algorithms, allowing the identification of unknown species (Lago et al., 2014). Nevertheless, GenBank is not specifically standardised for seafood species and it has been criticised as being susceptible to problems, such as unreliable terminology and missing information (Hellberg et al., 2016).

Barcode of Life Database (BOLD, 2017), coordinated by the International Barcode of Life project (iBOL, 2017), is a user-friendly search engine that supports the collection of DNA barcodes, with the main goal of creating a reference library for all living species. Many of the sequences deposited in BOLD are generated from specimens identified by expert taxonomists, reducing the likelihood of inaccurate identifications (Galimberti et al., 2013; Naaum & Hanner, 2016). Concerning reference databases specifically designed for seafood species, the most established and comprehensive resource is the Fish Barcode of Life Initiative (FISH-BOL, 2017). Conceived in 2004, FISH-BOL involves hundreds of researchers aiming at obtaining standardised reference sequences, focused on COI gene for all fish species and based on taxonomically verified voucher specimens. Sequences

collected by FISH-BOL are publicly archived in several other databases, including GenBank (Steinke & Hanner, 2011; Ward et al., 2009).

FishTrace (FishTrace, 2017) is a smaller European database that is currently limited to European fish species and does not use the COI gene as barcode. The main goal of FishTrace was to compile accessible information and material needed for the genetic discrimination of marine fish species from European waters and/or marketed in Europe (Clark, 2015).

DNA sequence reference databases are clearly advantageous computer science novelties due to the large quantity of sequence information. These data became available for the design of rapid DNA-based assays, being particularly useful for pre or post analysis of DNA barcodes. However, the existence of multiple databases may create a certain degree of complexity in species identification because each database includes variations on the information sources.

Methodologies

The methodologies based on DNA analysis are the most commonly used approaches in food authentication studies, mainly due to the well-known advantages associated with DNA molecules (Espiñeira & Santaclara, 2016). PCR amplification followed by DNA sequencing is the gold standard tool among genetic methods, being currently applied in numerous approaches for seafood species identification and authentication, many of them using DNA barcoding. In this section, the most widely used methodologies to exploit DNA barcoding will be described, from the classical Sanger's sequencing, to the most recent and emerging technologies such as NGS and HRM analysis.

Sanger's sequencing

DNA sequencing based on Sanger's method enables the incorporation of chain-terminating dideoxynucleotides (ddNTP) into sequences by means of DNA polymerase activity (Sanger et al., 1977). This technique remained the method of choice for more than two decades and was used for the sequencing of complete genomes of a high number of species. The automated Sanger's method of DNA sequencing is known as the "first generation technology", while innovative methods are referred as "next generation sequencing" (Kumar & Kocour, 2017). Despite the relevant advances in molecular biology technology, which improved the method effectiveness, Sanger's sequencing still follows these main steps: amplification of the targeted genetic region by PCR and clean-up step of the produced amplicons, followed by the addition of a dye-labelled ddNTP in the cycle sequencing reaction; afterwards, a sequencing clean-up step is performed, followed by

capillary electrophoresis to determine the identity of the nucleotides in the sequence; finally, the collection and revision of raw data in order to generate a consensus sequence (Hellberg et al., 2016).

Since COI is the most commonly targeted DNA barcode, Sanger's sequencing surely remains as the most used methodology for its analysis (Table 1). Despite the numerous reports and advances prompted by the classical Sanger's sequencing, some drawbacks have been pointed out, mainly concerning the low speed and high costs associated with the technique (Kumar & Kocour, 2017). Furthermore, as it can be verified in Table 1, first generation sequencing methods mostly use larger DNA fragments. These amplicons are often too long to be recovered from highly processed food products. Moreover, samples containing multiple species (e.g. canned products, surimi) cannot be identified with sequencing unless a time-consuming and expensive cloning step is added (Hellberg & Mourissey, 2011).

NGS

Recently, to overcome the limitations of first generation sequencing, new and improved technologies have emerged, which are designated as NGS and grouped into second, third and fourth generation sequencing approaches. Several terms have also been used regarding new sequencing technologies, such as high-throughput sequencing, massive parallel sequencing or clonal sequencing, being sometimes applied as synonyms or actually representing different tools. Nonetheless, NGS was one of the first terms used and is currently the most common name used for sequencing techniques other than Sanger's method (Kumar & Kocour, 2017).

Current NGS platforms are able to yield millions of DNA reads in a relatively short period of time, being the performance of the instrumentation improving every year. NGS combined with DNA barcoding has been termed as metabarcoding, which relies on the use of universal PCR primers to amplify massively one or more taxonomically informative targets. The prefix 'meta' refers to the collection of barcode sequences from different species. Generally, NGS consists of: (i) DNA extraction from food or other sample; (ii) amplification of a specific DNA barcode or other useful target region of taxonomic interest; (iii) sequencing the corresponding DNA amplicons using NGS technology; (iv) analysis the sequences using suitable bioinformatics tools; (5) species identification from the unknown samples (Staats et al., 2016). In the field of fishery research, metabarcoding has been widely applied to genomics, population and traceability studies (Kumar & Kocour, 2017), including environmental marine studies (Staats et al., 2016). Regarding seafood species

identification, the number of metabarcoding applications is still limited, and has relied on second generation sequencing technologies (Table 2).

The so-called second generation technologies are based on sequencing by synthesis or ligation, which may include pyrosequencing and reversible chain termination sequencing. These approaches both eliminate the *in vivo* bacterial cloning stage of the Sanger's methodology. Among the available NGS platforms, 454 pyrosequencing technology of Roche was the first commercially available NGS system and the one with the greatest potential in species identification. The main advantages of pyrosequencing rely on the speed and capacity to generate longer sequence reads, allowing a more informative set of DNA barcodes to be sequenced, along with high accuracy and automation potential (Kumar & Kocour, 2017). Generally, pyrosequencing is based on the real-time detection of pyrophosphate (PPi) molecules, which are released during the incorporation of nucleotides by DNA polymerase. Afterwards, PPi molecules initiate a series of enzymatic reactions leading to the production of light by the firefly enzyme luciferase (Figure 2) (Galimberti et al., 2013; Hellberg & Mourissey, 2011). Abbadi et al. (2017) showed the potential of pyrosequencing as a simple, rapid and cost-effective tool for the differentiation of bivalve species in fresh and processed seafood samples. De Battisti et al. (2014) were able to identify 25 fish species regarding Clupeiformes and Pleuronectiformes groups (Table 2).

Kappel et al. (2017) exploited metabarcoding using the Illumina MiSeq platform targeting two short *cytb* genes for tuna species identification in prepared mixtures of different fish species. This system is a reversible chain termination sequencing in which fluorescently labelled reversible dNTP terminators are coupled to a single-stranded DNA template. A unique wavelength and intensity is emitted by each nucleotide base as each dNTP is added to the growing DNA chain, allowing the detection of the fluorescent signal (Hellberg et al., 2016).

Ion Torrent technology is based on pyrosequencing with change in pH detection (Ion torrent Personal Genome Machine (PGM)) due to nucleotide incorporation into the complementary DNA strand, which causes hydrogen ion release (Hellberg et al., 2016). Recently, Carvalho et al. (2017b) explored metabarcoding with Ion torrent PGM methodology for the identification of fish species in highly processed codfish products, reporting a mislabelling rate of 41%. Giusti et al (2017a) applied same technology to short 16S rRNA barcode from a wide range of fish and cephalopod species to analyse surimi products. The authors verified that 37.5% of the products were mislabeled, from which 25% declared a species different from those identified and 25% did not label the presence of mollusks.

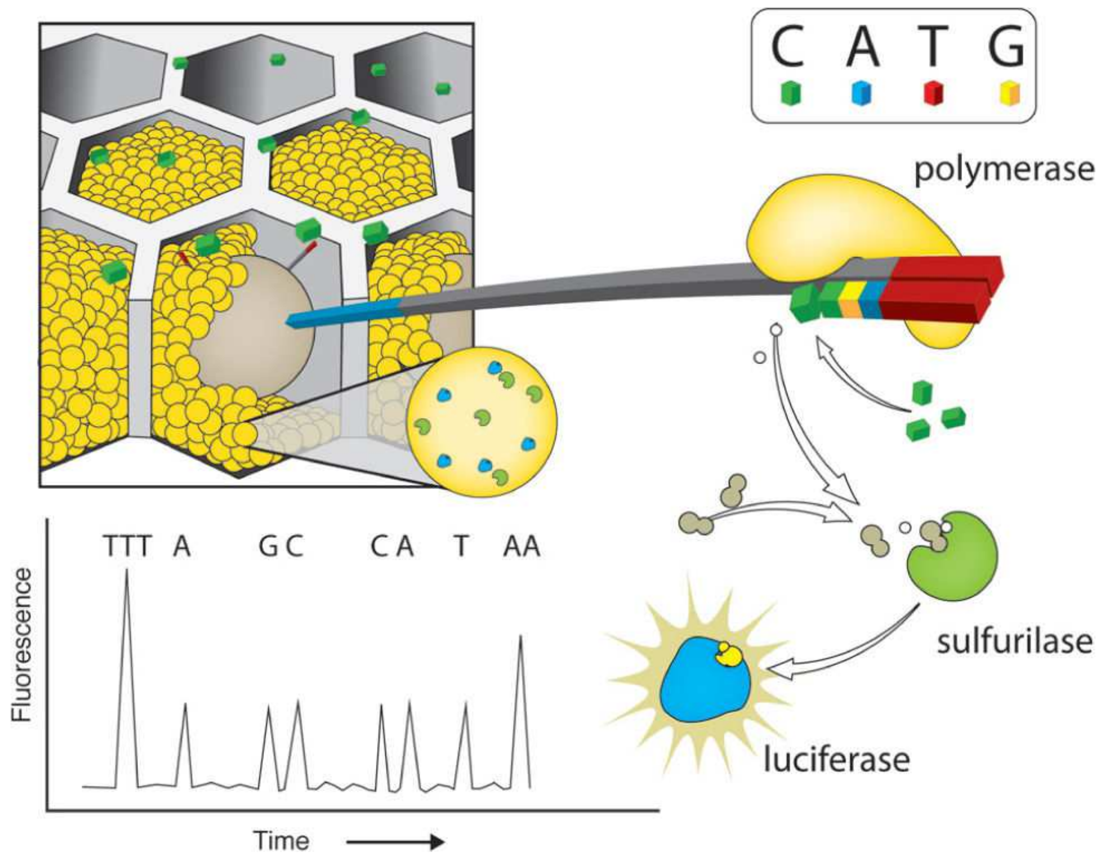


Figure 2. Pyrosequencing using Roche/454 Titanium platform. After loading the DNA-amplified beads (libraries) into individual Pico Titer Plate (PTP) wells, other type of beads, coupled with luciferase and sulphurilase, are added. The figure shows just one type of 2'-deoxyribonucleoside triphosphate (dNTP) - cytosine - that flows through the wells. Once the polymerase adds one nucleotide, a sulfurilase-luciferase reaction occurs, emitting light. A fibre optic slide is attached to a microfluidics camera allowing the reagents to reach the wells packed with beads. Underneath the fibre optic slide there is a direct connection to a high-resolution camera (charge coupled device or CCD), which allows detection of light emitted by each PTP when a pyrosequencing reaction occurs. Reprinted from Escalante et al. (2014). Copyright Elsevier 2017.

Sequencing technologies remain as the most reliable and comprehensive ways to obtain information from DNA barcodes. However, due to the inherent costs of the technique associated to instrumentation and time-consuming steps that are sometimes involved, its application is often unfeasible for routine use in many laboratories (Naaum & Hanner, 2016).

PCR-RFLP

PCR-RFLP involves the amplification of a preselected DNA fragment with universal primers, exploiting the polymorphisms in homologous DNA sequences by a digestion with restriction endonucleases, which recognize specific short sequences (four to six nucleotides) of the amplified fragment and cut them at those sites (Lago et al., 2014; Wilwet et al., 2018). Afterwards, the digested fragments can be separated and observed by gel electrophoresis. In order to establish a protocol for species identification using PCR-RFLP, sequence information for the DNA fragment of interest is required (DNA sequencing) to select appropriate restriction endonucleases that produce species-specific DNA restriction patterns after enzymatic digestion (Espiñeira & Santaclara, 2016; Hellberg & Morrisey, 2011).

PCR-RFLP is a well-established method and broadly accepted in seafood authentication because of its simplicity, speed and low-cost (Ferrito et al., 2016; Wilwet et al., 2018). Owing to the referred advantages, this technique has been successfully used in combination with DNA barcoding to investigate the authenticity of several seafood species and processed products (Table 2). Ferrito et al. (2016) investigated the application of PCR-RFLP targeting COI barcode in the identification of hake species in 20 convenience seafood products, detecting 35% of mislabelled products, mainly by the substitution of *M. merluccius* by other hake species with lower market level. Labelling accuracy in processed anchovy products was assessed by Pappalardo and Ferrito (2015b) using a similar approach, showing that 14% of the samples were the result of *Engraulis encrasicolus* fraudulent substitution, proving the usefulness of the method for routine analysis, mainly for large-scale sample screening. The application of PCR-RFLP targeting other mitochondrial barcode regions (16S rRNA/tRNA^{val}) was also effective in the differentiation of crustacean species, namely shrimps from Penaeidae family (Pascoal et al., 2008a, 2008b; Wilwet et al., 2018) (Table 2).

Although it is a simple, robust and easy to perform technique when compared with other PCR-based tools, PCR-RFLP is often less effective for species discrimination due to low reproducibility in the obtained DNA patterns. Another problem is the possibility of incomplete fragment digestion that could create additional restriction sites or even eliminate intraspecific variation (Lago et al., 2014).

HRM analysis

High-resolution melting analysis is a post-PCR method that detects small variations between sequences based on the melting properties of the double-stranded DNA (dsDNA). It has emerged with the advances of high resolution instrumentation and with

the enhanced fluorescent DNA-binding dyes (Reed et al., 2007). The method involves the gradual denaturation (melting) of PCR amplicons and detection of subsequent subtle fluorescent changes by the so-called new generation dyes present in the amplification reaction, such as EvaGreen (Biotium) and SYTO9 (Invitrogen). The new dyes can be used at higher concentrations than the classical SYBR Green I dye, generating superior fluorescent signals and improved sensitivity without causing PCR inhibition.

Generally, HRM consists of: (i) DNA extraction from food or other sample; (ii) amplification by real-time PCR with an enhanced fluorescent dye targeting a specific DNA barcode or other useful target region of taxonomic interest (100-200 bp); (iii) melting curve analysis; (iv) statistical data analysis using a specific HRM software. HRM relies on the amplification of the target DNA with subsequent melting of the amplicons performed in the presence of a saturation dye, which shows low levels of fluorescence when unbound, but is highly fluorescent in the bound state (Angers et al., 2017; Druml & Cichna-Markl, 2014; Meistertzheim et al., 2017). The melting temperature, which is the temperature at which half of the dsDNA is denatured, depends mainly on the number of nucleotides guanine and cytosine (GC content) and on the length of the sequence. Indeed, the melting profiles are monitored through the release of the DNA binding dye as the temperature increases. Even a single base variation between two dsDNA can be detected by a shift in melting temperature and a difference of the curve shape (conventional melting analysis). Afterwards, a specific software enables the statistical treatment of the melting curves to systematically and accurately assign samples into different clusters, allowing a suitable data manipulation (Lo & Shaw, 2018; Reed et al., 2007).

Recently, real-time PCR coupled to HRM analysis has proved to be a simple, cost-effective and rapid tool that can be applied in several molecular biology areas, such as DNA barcoding, fingerprinting or gene mapping (Angers et al., 2017). Particularly, DNA barcoding combined with HRM analysis has been designated as Bar-HRM and considered a powerful tool to screen and discriminate among closely related species (Meistertzheim et al., 2017), particularly in foods (Druml & Cichna-Markl, 2014). Recently, Fernandes et al. (2017a, 2017b, 2018) proposed HRM systems targeting COI mini-barcodes for the differentiation of shrimp, codfish and hake species (Table 2). The three systems were successfully applied to the analysis of processed seafood samples, demonstrating their usefulness in detecting mislabelled products. The HRM approach developed by Meistertzheim et al. (2017) targeted the nuclear 18S rRNA gene as a mini-barcode that enabled the fast and cost-effective identification of bivalve species from the Ostreida, Mytilida and Pectinida orders. Table 2 summarises the reported DNA barcoding combined with HRM applied to seafood species discrimination.

DNA barcoding combined with HRM analysis has demonstrated to be a high-through-put, reliable, cost-effective a simple tool to screen and discriminate closely related species. Since HRM is preceded of real-time PCR amplification, a quantitative analysis can be further performed, although without using melting curve data (Fernandes et al., 2018). However, in the case of mixed species, HRM cannot provide their accurate identification, but can elucidate differences among them (Angers et al., 2017). HRM analysis cannot determine the exact nucleotide difference(s) responsible for the distinct melting behaviour, but that can be complemented by Sanger's sequencing in the case of single species (Fernandes et al., 2017b, 2018). In the case of mixed species, Sanger's sequencing is not surely the adequate approach, which could only be given probably by NGS (Angers et al., 2017).

Conclusion

Currently, global efforts are required to trace and authenticate seafood to ensure fair competition among the fishery, industry sectors, avoiding fraudulent practices, with the final goal of protecting the consumers. Since seafood products are very diverse with regard to the range of fish and shellfish species considered as a food and various processing methods are presently used, their labelling compliance poses constant challenges. In this context, DNA barcoding has revealed itself as an extremely valuable tool. A DNA barcode should be easily amplified across a large range of related individuals, but at the same time should have low intraspecies variability, which has been accomplished by the COI gene, as the preferential targeted DNA region in animal species, including fish and shellfish. In spite of that, other mitochondrial genes (*cytb*, 12S rRNA, 16S rRNA) and nuclear genes (5S rRNA, 28S rRNA, rhodopsin, pantophysin) have also been targeted as barcodes for seafood species discrimination. DNA barcode sequencing based on Sanger's method is, undoubtedly, the most currently used approach for seafood authentication. However, its low speed and incapacity of analysing mixed samples are drawbacks that are being overcome with the advent of NGS. DNA metabarcoding provides fast and massive sequencing, being capable of multiple target analysis. Therefore, DNA metabarcoding has gained much attention in seafood authentication. The second generation NGS technologies are currently the metabarcoding approaches that have been applied to seafood authentication. As a well-established approach, PCR-RFLP has provided simple and cost-effective tools to differentiate fish and shellfish species. HRM analysis represent a promising approach in barcoding since it provides simple, cost-effective and high-throughput species discrimination in processed seafoods.

Acknowledgements

This work was supported by FCT (Fundação para a Ciência e Tecnologia) through project UID/QUI/50006/2013 – POCI/01/0145/FEDER/007265 with financial support from FCT/MEC through national funds and co-financed by FEDER, under the Partnership Agreement PT2020, by the project NORTE-01-0145-FEDER-000011 and FOODINTEGRITY (FP7-KBBE-2013-single-stage, No 613688). T. J. R. Fernandes is grateful to FCT grant (SFRH/BD/93711/2013) financed by POPH-QREN (subsidised by FSE and MCTES).

References

Abbadi, M., Marciano, S., Tosi, F., De Battisti, C., Panzarin, V., Arcangeli, G., & Cattoli, G. (2017). Species identification of bivalve molluscs by pyrosequencing. *Journal of the Science of Food and Agriculture, 97*, 512-519.

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*, 695-706.

Abdullah, A., & Rehbein, H. (2016). The differentiation of tuna (family: Scombridae) products through the PCR-based analysis of the cytochrome b gene and parvalbumin introns. *Journal of Science and Food Agriculture, 96*, 456-464.

Abdullah, A., & Rehbein, H. (2017). DNA barcoding for the species identification of commercially important fishery products in Indonesian markets. *International Journal of Food Science & Technology, 52*, 266-274.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology, 215*, 403-410.

Angers, A., Ballin, N. Z., Hofherr, J., Kagkli, D. M., Lievens, A., Maquet, A., Martinsohn, J. T., Paracchini., V., Petrillo, M., Puertas-Gallardo, A. (2017). Enhancing fish species identification using novel markers and emerging technologies. *JRC Technical Reports,* EUR 28505, 1-21.

Ardura, A., Pola, I. G., Ginuino, I., Gomes, V., & Garcia-Vazquez, E. (2010). Application of barcoding to Amazonian commercial fish labelling. *Food Research International, 43*, 1549-1552.

Ardura, A., Planes, S., & Garcia-Vazquez, E. (2013). Applications of DNA barcoding to fish landings: authentication and diversity assessment. *ZooKeys,* 365, 49-65.

Armani, A., Guardone, L., Castigliego, L., D'Amico, P., Messina, A., Malandra, R., Gianfaldoni, D., & Guidi, A. (2015a). DNA and Mini-DNA barcoding for the identification of Porgies species (family Sparidae) of commercial interest on the international market. *Food Control, 50*, 589-596.

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

Bartlett, S. E., & Davidson, W. S. (1992). FINS (forensically informative nucleotide sequencing): A procedure for identifying the animal origin of biological specimens. *Biotechniques, 13*, 518.

BOLD, Barcode of Life Database. Available at: http://www.boldsystems.org/ (Last accessed 21 September 2017).

Carvalho, D. C., Neto, D. A. P., Brasil, B. S. A. F., & Oliveira, D. A. A. (2011). DNA barcoding unveils a high rate of mislabeling in a commercial freshwater catfish from Brazil. *Mitochondrial DNA, 22*, 97-105.

Carvalho, D. C., Guedes, D., da Gloria Trindade, M., Coelho, R. M. S., & de Lima Araujo, P. H. (2017a). Nationwide Brazilian governmental forensic programme reveals seafood mislabelling trends and rates using DNA barcoding. *Fisheries Research, 191*, 30-35.

Carvalho, D. C., Palhares, R. M., Drummond, M. G., & Gadanho, M. (2017b). Food metagenomics: Next generation sequencing identifies species mixtures and mislabeling within highly processed cod products. *Food Control, 80*, 183-186.

Cawthorn, D.-M., Steinman, H. A., & Witthuhn, R. C. (2012). Evaluation of the 16S and 12S rRNA genes as universal markers for the identification of commercial fish species in South Africa. *Gene, 491*, 40-48.

Chang, C.-H., Lin, H.-Y., Ren, Q., Lin, Y.-S., & Shao, K.-T. (2016). DNA barcode identification of fish products in Taiwan: Government-commissioned authentication cases. *Food Control, 66*, 38-43.

Changizi R., Farahmad, H., Soltani, M., Asareh, R., & Ghiasvand, Z. (2013), Species identification reveals mislabeling of important fish products in Iran by DNA barcoding. *Iranian Journal of Fisheries Sciences, 12*, 783-791.

Chin, T. C., Adibah, A. B., Danial Hariz, Z. A., & Siti Azizah, M. N. (2016). Detection of mislabelled seafood products in Malaysia by DNA barcoding: Improving transparency in food market. *Food Control, 64*, 247-256.

Clark, L. F. (2015). The current status of DNA barcoding technology for species identification in fish value chains. *Food Policy, 54*, 85-94.

Cline, E. (2012). Marketplace substitution of Atlantic salmon for Pacific salmon in Washington State detected by DNA barcoding. *Food Research International, 45*, 388-393.

Cohen, N. J., Deeds, J. R., Wong, E. S., Hanner, R. H., Yancy, H. F., White, K. D., Thompson, T. M., Wahl, M., Pham, T. D., Guichard, F. M., Huh, I., Austin, C., Dizikes, G., & Gerber, S. I. (2009). Public health response to puffer fish (Tetrodotoxin) poisoning from mislabeled product. *Journal of Food Protection, 72*, 810-817.

Cutarelli, A., Amoroso, M. G., De Roma, A., Girardi, S., Galiero, G., Guarino, A., & Corrado, F. (2014). Italian market fish species identification and commercial frauds revealing by DNA sequencing. *Food Control, 37*, 46-50.

De Battisti, C., Marciano, S., Magnabosco, C., Busato, S., Arcangeli, G., & Cattoli, G. (2014). Pyrosequencing as a tool for rapid fish species identification and commercial fraud detection. *Journal of Agricultural and Food Chemistry, 62*, 198-205.

de Brito, M. A., Schneider, H., Sampaio, I., & Santos, S. (2015). DNA barcoding reveals high substitution rate and mislabeling in croaker fillets (Sciaenidae) marketed in Brazil: The case of “pescada branca” (Cynoscion leiarchus and Plagioscion squamosissimus). *Food Research International, 70*, 40-46.

Di Pinto, A., Di Pinto, P., Terio, V., Bozzo, G., Bonerba, E., Ceci, E., & Tantillo, G. (2013). DNA barcoding for detecting market substitution in salted cod fillets and battered cod chunks. *Food Chemistry, 141*, 1757-1762.

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.

Druml, B., & Cichna-Markl, M. (2014). High resolution melting (HRM) analysis of DNA – Its role and potential in food analysis. *Food Chemistry, 158*, 245-254.

EUMOFA (2016). Directorate-General for Maritime Affairs and Fisheries of the European Commission, European Market Observatory for Fisheries and Aquaculture Products (EUMOFA) (2016). EUMOFA The EU Fish Market.

Escalante, A. E., Jardón Barbolla, L., Ramírez-Barahona, S., & Eguiarte, L. E. (2014). The study of biodiversity in the era of massive sequencing. *Revista Mexicana de Biodiversidad, 85*, 1249-1264.

Espiñeira, M., & Santaclara, F. J. (2016). The use of molecular biology techniques in food traceability. In M. Espiñeira, & F. J. Santaclara (Eds.), *Advances in Food Traceability Techniques and Technologies*, (pp. 91-118). Cambridge: Woodhead Publishing.

Fernandes, T. J. R., Silva, C. R., Costa, J., Oliveira, M. B. P. P., & Mafra, I. (2017a). High resolution melting analysis of a COI mini-barcode as a new approach for Penaeidae shrimp species discrimination. *Food Control, 82*, 8-17.

Fernandes, T. J. R., Costa, J., Oliveira, M. B. P. P., & Mafra, I. (2017b). DNA barcoding coupled to HRM analysis as a new and simple tool for the authentication of Gadidae fish species. *Food Chemistry, 230*, 49-57.

Fernandes, T. J. R., Costa, J., Oliveira, M. B. P. P., & Mafra, I (2018). COI barcode-HRM as a novel approach for the discrimination of hake species. *Fisheries Research*, *197*, 50-59.

Ferrito, V., Bertolino, V., & Pappalardo, A. M. (2016). White fish authentication by COIBar-RFLP: Toward a common strategy for the rapid identification of species in convenience seafood. *Food Control, 70*, 130-137.

Fitzcharles, E. M. (2012). Rapid discrimination between four Antarctic fish species, genus *Macrourus*, using HRM analysis. *Fisheries Research, 127–128*, 166-170.

FISH-BOL, Fish Barcode of Life. Available at http://www.fishbol.org/ (Last accessed 21 September 2017).

FishTrace, genetic and taxonomic reference database for species identification of European commercial marine fish species. Available at https://fishtrace.jrc.ec.europa.eu/ (Last accessed 21 September 2017).

Galal-Khallaf, A., Ardura, A., Mohammed-Geba, K., Borrell, Y. J., & Garcia-Vazquez, E. (2014). DNA barcoding reveals a high level of mislabeling in Egyptian fish fillets. *Food Control, 46*, 441-445.

Galal-Khallaf, A., Ardura, A., Borrell, Y. J., & Garcia-Vazquez, E. (2016). PCR-based assessment of shellfish traceability and sustainability in international Mediterranean seafood markets. *Food Chemistry, 202*, 302-308.

Galimberti, A., De Mattia, F., Losa, A., Bruni, I., Federici, S., Casiraghi, M., Martellos, S., & Labra, M. (2013). DNA barcoding as a new tool for food traceability. *Food Research International, 50*, 55-63.

Giusti, A., Armani, A., & Sotelo, C. G. (2017a). Advances in the analysis of complex food matrices: Species identification in surimi-based products using Next Generation Sequencing technologies. *PLoS ONE, 12*, e0185586.

Glover, K. A., Dahle, G., & Jørstad, K. E. (2011). Genetic identification of farmed and wild Atlantic cod, *Gadus morhua*, in coastal Norway. *ICES Journal of Marine Science, 68*, 901-910.

Guardone, L., Tinacci, L., Costanzo, F., Azzarelli, D., D'Amico, P., Tasselli, G., Magni, A., Guidi, A., Nucera, D., & Armani, A. (2017). DNA barcoding as a tool for detecting mislabeling of fishery products imported from third countries: An official survey

conducted at the Border Inspection Post of Livorno-Pisa (Italy). *Food Control, 80*, 204-216.

Günther, B., Raupach, M. J., & Knebelsberger, T. (2017). Full-length and mini-length DNA barcoding for the identification of seafood commercially traded in Germany. *Food Control, 73*, 922-929.

Hanner, R., Floyd, R., Bernard, A., Collette, B. B., & Shivji, M. (2011). DNA barcoding of billfishes. *Mitochondrial DNA, 22*, 27-36.

Harris, D. J., Rosado, D., & Xavier, R. (2016). DNA barcoding reveals extensive mislabeling in seafood sold in Portuguese supermarkets. *Journal of Aquatic Food Product Technology, 25*, 1375-1380.

Hebert, P. D., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences, 270*, 313-321.

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S., & Francis, C. M. (2004). Identification of birds through DNA barcodes. *PLOS Biology, 2*, 1657-1663.

Hellberg, R. S., & Morrissey, M. T. (2011). Advances in DNA-based techniques for the detection of seafood species substitution on the commercial market. *Journal of Laboratory Automation, 16*, 308-321.

Hellberg, R. S., Pollack, S. J., & Hanner, R. H. (2016). Seafood species identification using DNA sequencing. In R. Hanner, & A. Naaum (Eds.), *Seafood Authenticity and Traceability*, (pp. 113-132). San Diego: Academic Press.

Hubálková, Z., Králík, P., Kasaloví, J., & Rencoví, E. (2009). Application of DNA-based techniques for intraspecies differentiation of hake fish. *Acta Veterinaria Brno, 78*, 673-678.

Huxley-Jones, E., Shaw, J. L. A., Fletcher, C., Parnell, J., & Watts, P. C. (2012). Use of DNA Barcoding to reveal species composition of convenience seafood. *Conservation Biology, 26*, 367-371.

iBOL, International Barcode of Life project, Ontario, Canada. Available at: http://ibol.org/ (Last accessed 21 September 2017).

Imoto, J. M., Saitoh, K., Sasaki, T., Yonezawa, T., Adachi, J., Kartavtsev, Y. P., Miya, M., Nishida, M., & Hanzawa, N. (2013). Phylogeny and biogeography of highly diverged freshwater fish species (Leuciscinae, Cyprinidae, Teleostei) inferred from mitochondrial genome analysis. *Gene, 514*, 112-124.

Jilberto, F., Araneda, C., & Larraín, M. A. (2017). High resolution melting analysis for identification of commercially-important *Mytilus* species. *Food Chemistry, 229*, 716-720.

Jin, Y., Li, Q., Kong, L., Yu, H., & Zhong, X. (2015). High-resolution melting (HRM) analysis: a highly sensitive alternative for the identification of commercially important *Crassostrea* oysters. *Journal of Molluscan Studies, 81*, 167-170.

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.

Kappel, K., & Schröder, U. (2015). Species identification of fishery products in Germany. *Journal für Verbraucherschutz und Lebensmittelsicherheit, 10*, 31-34.

Khamnamtong, B., Klinbunga, S., & Menasveta, P. (2005). Species identification of five penaeid shrimps using PCR-RFLP and SSCP analyses of 16S ribosomal DNA. *Journal of Biochemistry and Molecular Biology, 38*, 491-499.

Kumar, G., & Kocour, M. (2017). Applications of next-generation sequencing in fisheries research: A review. *Fisheries Research, 186*, 11-22.

Lago, F. C., Alonso, M., Vieites, J. M., & Espiñeira, M. (2014). Fish and Seafood Authenticity - Species Identification. In I. S. Boziaris (Ed), *Seafood Processing*, (pp. 419-452). New Jersey: John Wiley & Sons, Ltd.

Lamendin, R., Miller, K., & Ward, R. D. (2015). Labelling accuracy in Tasmanian seafood: An investigation using DNA barcoding. *Food Control, 47*, 436-443.

Larsen, R., Eilertsen, K. E., & Elvevoll, E. O. (2011). Health benefits of marine foods and ingredients. *Biotechnology Advances, 29*, 508-518.

Lo, Y.-T., & Shaw, P.-C. (2018). DNA-based techniques for authentication of processed food and food supplements. *Food Chemistry, 240*, 767-774.

Maralit, B. A., Aguila, R. D., Ventolero, M. F. H., Perez, S. K. L., Willette, D. A., & Santos, M. D. (2013). Detection of mislabeled commercial fishery by-products in the Philippines using DNA barcodes and its implications to food traceability and safety. *Food Control, 33*, 119-125.

Mariani, S., & Bekkevold, D. (2014). The Nuclear Genome: Neutral and Adaptive Markers in Fisheries Science. In *Stock Identification Methods (Second Edition)*, (pp. 297-327). San Diego: Academic Press.

Mazón-Suástegui, J. M., Fernández, N. T., Valencia, I. L., Cruz-Hernández, P., & Latisnere-Barragán, H. (2016). 28S rDNA as an alternative marker for commercially important oyster identification. *Food Control, 66*, 205-214.

Meistertzheim, A.-L., Héritier, L., & Lejart, M. (2017). High-Resolution Melting of 18S rDNA sequences (18S-HRM) for discrimination of bivalve’s species at early juvenile stage: application to a spat survey. *Marine Biology, 164*, 133.

Mitchell, J. K., & Hellberg, R. S. (2016). Use of the mitochondrial control region as a potential DNA mini-barcoding target for the identification of canned tuna species. *Food Analytical Methods, 9*, 2711-2720.

Mueller, S., Handy, S. M., Deeds, J. R., George, G. O., Broadhead, W. J., Pugh, S. E., & Garrett, S. D. (2015). Development of a COX1 based PCR-RFLP method for fish species identification. *Food Control, 55*, 39-42.

Muñoz-Colmenero, M., Klett-Mingo, M., Díaz, E., Blanco, O., Martínez, J. L., & Garcia-Vazquez, E. (2015). Evolution of hake mislabeling niches in commercial markets. *Food Control, 54*, 267-274.

Naaum, A. M., & Hanner, R. H. (2016). An introduction to DNA-based tools for seafood identification. In R. Hanner, & A. Naaum (Eds), *Seafood Authenticity and Traceability*, (pp. 99-111). San Diego: Academic Press.

NCBI, National Center for Biotechnology Information, Bethesda, USA. Available at: http://www.ncbi.nlm.nih.gov/ (Last accessed 21 September 2017).

Nedunoori, A., Turanov, S. V., & Kartavtsev, Y. P. (2017). Fish product mislabeling identified in the Russian far east using DNA barcoding. *Gene Reports, 8*, 144-149.

Nicolé, S., Negrisolo, E., Eccher, G., Mantovani, R., Patarnello, T., Erickson, D. L., Kress, W. J., & Barcaccia, G. (2012). DNA barcoding as a reliable method for the authentication of commercial seafood products. *Food Technology and Biotechnology*, 50, 387-398.

Pampoulie, C., Daníelsdóttir, A. K., Thorsteinsson, V., Hjörleifsson, E., Marteinsdóttir, G., & Ruzzante, D. E. (2012). The composition of adult overwintering and juvenile aggregations of Atlantic cod (*Gadus morhua*) around Iceland using neutral and functional markers: a statistical challenge. *Canadian Journal of Fisheries and Aquatic Sciences, 69*, 307-320.

Pappalardo, A. M., & Ferrito, V. (2015a). DNA barcoding species identification unveils mislabeling of processed flatfish products in southern Italy markets. *Fisheries Research, 164*, 153-158.

Pappalardo, A. M., & Ferrito, V. (2015b). A COIBar-RFLP strategy for the rapid detection of *Engraulis encrasicolus* in processed anchovy products. *Food Control, 57*, 385-392.

Paracchini, V., Petrillo, M., Lievens, A., Puertas Gallardo, A., Martinsohn, J. T., Hofherr, J., Maquet, A., Silva, A. P. B., Kagkli, D. M., Querci, M., Patak, A., & Angers-Loustau, A. (2017). Novel nuclear barcode regions for the identification of flatfish species. *Food Control, 79,* 297-308.

Pascoal, A., Barros-Velazquez, J., Cepeda, A., Gallardo, J. M., & Calo-Mata, P. (2008a). Survey of the authenticity of prawn and shrimp species in commercial food products by

PCR-RFLP analysis of a 16S rRNA/tRNA^{Val} mitochondrial region. *Food Chemistry, 109*, 638-646.

Pascoal, A., Barros-Velázquez, J., Cepeda, A., Gallardo, J. M., & Calo-Mata, P. (2008b). A polymerase chain reaction-restriction fragment length polymorphism method based on the analysis of a 16S rRNA/tRNA^{Val} mitochondrial region for species identification of commercial penaeid shrimps (Crustacea: Decapoda: Penaeoidea) of food interest. *Electrophoresis, 29*, 499-509.

Reed, G. H., Kent, J. O., Wittwer, C. T. (2007). High-resolution DNA melting analysis for simple and efficient molecular diagnostics. *Pharmacogenomics, 8,* 597-608.

Regulation (EU) No 1169/2011 of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No. 1924/2006 and (EC) No. 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004, *Official Journal of the European Union*, *L304*, 18-63.

Regulation (EC) No. 1379/2013 of 11 December 2013 on the common organisation of the markets in fishery and aquaculture products, amending Council Regulations (EC) No 1184/2006 and (EC) No 1224/2009 and repealing Council Regulation (EC) No 104/2000. 354, 1-21.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA, 74*, 5463-5467.

Scarano, D., & Rao, R. (2014). DNA markers for food products authentication. *Diversity, 6*, 579-596.

Shokralla, S., Hellberg, R. S., Handy, S. M., King, I., & Hajibabaei, M. (2015). A DNA mini-barcoding system for authentication of processed fish products. *5*, 15894.

Sicherer, S. H. (2011). Epidemiology of food allergy. *Journal of Allergy and Clinical Immunology*, 127, 594-602.

Staats, M., Arulandhu, A. J., Gravendeel, B., Holst-Jensen, A., Scholtens, I., Peelen, T., Prins, T. W., & Kok, E. (2016). Advances in DNA metabarcoding for food and wildlife forensic species identification. *Analytical and Bioanalytical Chemistry, 408*, 4615-4630.

Steinke, D., & Hanner, R. (2011). The FISH-BOL collaborators' protocol. *Mitochondrial DNA, 22*, 10-14.

Stern, D. B., Castro Nallar, E., Rathod, J., & Crandall, K. A. (2017). DNA Barcoding analysis of seafood accuracy in Washington, D.C. restaurants. *PeerJ, 5*, e3234.

Tognoli, C., Saroglia, M., Terova, G., Gornati, R., & Bernardini, G. (2011). Identification of fish species by 5S rRNA gene amplification. *Food Chemistry, 129*, 1860-1864.

Tomás, C., Ferreira, I. M. P. L. V. O., & Faria, M. A. (2017). Codfish authentication by a fast short amplicon high resolution melting analysis (SA-HRMA) method. *Food Control, 71*, 255-263.

Vartak, V. R., Narasimmalu, R., Annam, P. K., Singh, D. P., & Lakra, W. S. (2015). DNA barcoding detected improper labelling and supersession of crab food served by restaurants in India. *Journal of the Science of Food and Agriculture, 95*, 359-366.

Wang, D., & Hsieh, Y.-H. P. (2016). The use of imported pangasius fish in local restaurants. *Food Control, 65*, 136-142.

Wang, H.-Y., Dong, C. A., & Lin, H.-C. (2017). DNA barcoding of fisheries catch to reveal composition and distribution of cutlassfishes along the Taiwan coast. *Fisheries Research, 187*, 103-109.

Ward, R. D., Hanner, R., & Hebert, P. D. (2009). The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology, 74*, 329-356.

Wen, J., Tinacci, L., Acutis, P. L., Riina, M. V., Xu, Y., Zeng, L., Ying, X., Chen, Z., Guardone, L., Chen, D., Sun, Y., Zhao, J., Guidi, A., & Armani, A. (2017). An insight into the Chinese traditional seafood market: Species characterization of cephalopod products by DNA barcoding and phylogenetic analysis using COI and 16SrRNA genes. *Food Control, 82*, 333-342.

Willette, D. A., Simmonds, S. E., Cheng, S. H., Esteves, S., Kane, T. L., Nuetzel, H., Pilaud, N., Rachmawati, R., & Barber, P. H. (2017). Using DNA barcoding to track seafood mislabeling in Los Angeles restaurants. *Conservation Biology*, 0, 1-10.

Wilwet, L., Jeyasekaran, G., Shakila, R. J., Sivaraman, B., & Padmavathy, P. (2018). A single enzyme PCR-RFLP protocol targeting 16S rRNA/tRNA^{val} region to authenticate four commercially important shrimp species in India. *Food Chemistry, 239*, 369-376.

Xiong, X., Guardone, L., Giusti, A., Castigliego, L., Gianfaldoni, D., Guidi, A., & Andrea, A. (2016). DNA barcoding reveals chaotic labeling and misrepresentation of cod (鳕, Xue) products sold on the Chinese market. *Food Control, 60*, 519-532.

EXPERIMENTAL PART

DNA barcoding coupled to HRM analysis as a new and simple tool for the authentication of Gadidae fish species
Food Chemistry, **2017**, 230, 49-57.

High resolution melting analysis of a COI mini-barcode as a new approach for Penaeidae shrimp species discrimination.
Food Control, **2017**, 82, 8-17.

COI barcode-HRM as a novel approach for the discrimination of hake species.
Fisheries Research, **2018**, 197, 50-59.

Food Chemistry 230 (2017) 49–57

Contents lists available at ScienceDirect

Food Chemistry

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

DNA barcoding coupled to HRM analysis as a new and simple tool for the authentication of Gadidae fish species

Telmo J.R. Fernandes, Joana Costa, M. Beatriz P.P. Oliveira, Isabel Mafra*

REQUIMTE-LAQV, Faculdade de Farmácia, Universidade do Porto, Rua de Jorge Viterbo Ferreira, 228, 4050-313 Porto, Portugal

*Corresponding author: Tel: +351 220428640. Fax: +351 226093390. E-mail: isabel.mafra@ff.up.pt

Abstract

This work aimed to exploit the use of DNA mini-barcodes combined with high resolution melting (HRM) for the authentication of gadoid species: Atlantic cod (*Gadus morhua*), Pacific cod (*Gadus macrocephalus*), Alaska pollock (*Theragra chalcogramma*) and saithe (*Pollachius virens*). Two DNA barcode regions, namely cytochrome c oxidase subunit I (COI) and cytochrome b (*cytb*), were analysed *in silico* to identify genetic variability among the four species and used, subsequently, to develop a real-time PCR method coupled with HRM analysis. The *cytb* mini-barcode enabled best discrimination of the target species with a high level of confidence (99.3%). The approach was applied successfully to identify gadoid species in 30 fish-containing foods, 30% of which were not as declared on the label. Herein, a novel approach for rapid, simple and cost-effective discrimination/clustering, as a tool to authenticate Gadidae fish species, according to their genetic relationship, is proposed.

Keywords: DNA barcode; COI; *cytb*; high resolution melting analysis; real-time PCR; codfish; gadoid species.

Introduction

The nutritional value of seafood, as part of a healthy diet, has resulted in a significant increase in demand for fish over several decades. The collapse of some fish stocks, allied with price and the pursuit for increased profit, has led to high levVan, K. T., Shen, Y., & Williams-Hill, D. M., 2014; Mueller et al., 2015; Wetten, Wilson & Andersen, 2012; Sampels, 2015). In the European Union (EU), to control and reduce economic fraud, through substitution and/or mislabelling, fish products must be labelled with the commercial and scientific names, fishing and production methods, catch area and the fishing gear. For other processed foods, such as canned, composite products and breaded products, this information is voluntary (Armani, Guardone, La Castellana, Gianfaldoni, Guidi, & Castigliego, 2015; Di Pinto et al., 2013; Regulation (EC) No 206/2009; Regulation (EC) No 104/2000; Regulation (EU) No 1379/2013). Approximately 18% of the world's total catch is fish from the Gadiform order, which corresponded to almost 5.9 million tonnes in 2011 (FAO, 2014). The Gadidae family, in particular, represents an important marine resource and includes commercially important species that are common traditional dietary components for populations around the world. Atlantic cod (*Gadus morhua*), Pacific cod (*Gadus macrocephalus*), Alaska pollock (*Theragra chalcogramma*) and saithe (*Pollachius virens*) belong to this family, and are usually referred to as cod-like species or gadoids. Most of these species are similar in appearance, which makes their morphological identification very difficult or almost impossible. This can contribute to the misbranding of processed codfish products (Calo-Mata et al., 2003; Di Pinto et al., 2013; Heldberg et al., 2014; Wetten et al., 2012).

Food authenticity can be assessed using a broad variety of methods, such as those based on protein or DNA analysis. Within the sphere of DNA-based approaches, much attention has been devoted to DNA barcoding, which relies on sequence variation within a short and standardised region of the genome, designated as “barcode”, which provides accurate species identification (Hebert, Cywinska, Ball, & deWaard, 2003). Currently, the mitochondrial genes coding for cytochrome c oxidase subunit I (COI) and cytochrome b (*cytb*) are considered reliable DNA barcodes for the discrimination of animal species (Hebert et al., 2003; Hellberg et al., 2014; Mueller et al., 2015). For the identification of fish species, mitochondrial *loci* have been preferred to nuclear genes because fish genomes are haploid; they are present in high copy numbers (particularly in fish tissues) and their mutation rate is greater than that of nuclear genes (Cline, 2012; Rehbein et al., 2013). The number of DNA barcodes deposited in databases is growing continuously (http://www.fishbol.org/). So far, several studies have used COI or *cytb* mitochondrial DNA barcoding to identify seafood products and investigate broad patterns in fish mislabelling

(Cline, 2012; Di Pinto et al., 2013; Miller & Mariani, 2010; Rasmussen & Morissey, 2008; Wong & Hanner, 2008; Yancy et al., 2008). Despite being a good basis for species differentiation, DNA barcoding cannot be considered cost-effective since it depends on DNA sequencing. To overcome this, real-time PCR coupled to high resolution melting (HRM) analysis offers a rapid, reliable and more economic alternative to exploit DNA barcoding.

HRM analysis is regarded as an excellent tool for the identification and differentiation of closely related species or cultivars, identification of pathogens, screening of genetically modified organisms and detection of food allergens (Druml & Cichna-Markl, 2014; Madesis, Ganopoulos, Sakaridis, Argiriou, & Tsaftaris, 2013). This analysis has emerged from recent advances in high-resolution instrumentation using new generation fluorescent DNA-binding dyes (e.g. EvaGreen, LC Green and SYTO9). Owing to their ability to specifically bind double stranded DNA, these new dyes can be used at high concentrations and are less likely to cause non-specific amplification or PCR inhibition. This enables detection of subtle fluorescent changes during gradual melting of PCR fragments. HRM consists of a closed-tube post-PCR analysis based on the shape of melting transitions for real-time PCR products, allowing identification of small variations in DNA sequences, such as a single base change (deletion or addition). Recently, real-time PCR coupled to HRM analysis has provided simple, rapid, cost-effective and high-throughput approaches for food testing (Costa, Mafra, & Oliveira, 2012; Costa, 2013). So far, few works have reported the application of HRM analysis to differentiate fish species. McGlauflin et al. (2010) described the use of HRM analysis for the identification of 11 single-nucleotide polymorphisms (SNP) to distinguish rainbow trout (*Oncorhynchus mykiss*) and cutthroat trout (*Oncorhynchus clarkii*). Fitzcharles (2012) also reinforced the potential of HRM analysis as a rapid, robust and reliable technique for the discrimination of species among four Antarctic fish, even when analysing samples with poor DNA quality and quantity. However, none concerned the identification of codfish or cod-like species.

The aim of this work was to develop a robust and highly sensitive methodology for the rapid discrimination of four closely-related fish species from the Gadidae family. For this purpose, two DNA barcode regions, namely COI and the *cytb* genes of the selected species (*G. morhua*, *G. macrocephalus*, *T. chalcogramma* and *P. virens*) were analysed *in silico* to search for genetic variability among them. This variability was exploited using HRM analysis, targeting DNA mini-barcode regions to develop a specific, rapid and cost-effective approach for the identification of gadoid fish species. The proposed method was applied to processed fish-containing foods for species identification and verification of labelling compliance.

Materials and Methods

Sample collection and preparation

Authentic samples of Atlantic cod (*G. morhua*) (*n*=4), Pacific cod (*G. macrocephalus*) (*n*=3), Alaska pollock (*T. chalcogramma*) (*n*=2) and saithe (*P. virens*) (*n*=2) were kindly provided by Pascoal & Filhos SA. In order to evaluate the specificity of the proposed approach, a total of 34 samples, including different fish, crustacean and mollusc species (*n*=8) (Atlantic salmon - *Salmo salar*, gilt-head bream - *Sparus aurata*, ray - *Raja* spp., common sole - *Solea solea*, European pilchard - *Sardina pilchardus*, common shrimp - *Crangon crangon*, yellowfin tuna - *Thunus albacares*, squid - *Loligo* spp.), as well as other animal (*n*=9) (rabbit, deer, cow, chicken, turkey, pork, lamb, goat, ostrich) and plant species (*n*=17) (soybean, oat, rye, mint, wheat, lupine, maize, rice, pumpkin seeds, rapeseed, sunflower, tomato, peach, nectarine, apricot, strawberry, raspberry) acquired at local markets, were tested. A total of 42 fish-containing foods were also purchased at local markets, comprising a variety of frequently consumed traditional Portuguese products (codfish with cream, codfish cakes, "brás" style codfish, codfish "patanisca"), fish sticks, fish fillets, and patties, among others. The samples were milled and homogenised separately using a laboratory knife mill, Grindomix GM200 (Retsch, Haan, Germany), before being stored at -20 ºC for analysis. All containers and material used during this procedure were treated previously with a DNA decontamination solution (DNA-ExitusPlus™, AppliChem, Darmstadt, Germany).

In silico DNA barcode analysis and primer design

Sequences from *G. morhua*, *G. macrocephalus, T. chalcogramma* and *P. virens* were obtained from the NCBI database using the respective accession numbers for COI (KC015380.1, JQ354097.1, JQ354518.1 and KF930285.1, respectively) and *cytb* (EU492141.1, AB078152.1, AB078151.1 and EU492147.1, respectively) regions. Sequence alignment was performed with BioEdit v.7.2.5 (Ibis Biosciences, Carlsbad, USA) to examine variation in both COI (Figure 1) and *cytb* (Figure 2) genes among the selected species and their suitability for primer design. Accordingly, two sets of primers were designed for each DNA barcode region (COI and *cytb*), with the aim of using these sequences (Gad1COI-F/Gad2COI-R and Gad2CytB-F/Gad1Cytb-R) in the development of new real-time PCR system (Gad1COI-F/Gad1COI-R and Gad1CytB-F/Gad1Cytb-R) (Table 1). The nucleotide sequences were submitted to a basic local alignment search tool BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi), which identifies regions of local similarity among homologue sequences from different species and calculates the statistical significance of the matches. Primer specificity was assessed using the Primer-

BLAST tool that allows homologies in relation to all sequences available in the NCBI database to be revealed. Primer properties, and the absence of hairpins and self-hybridisation, were assessed using OligoCalc (http://www.basic.northwestern.edu./biotools/oligocalc.html). Primers were synthesised by STABVIDA (Lisbon, Portugal).

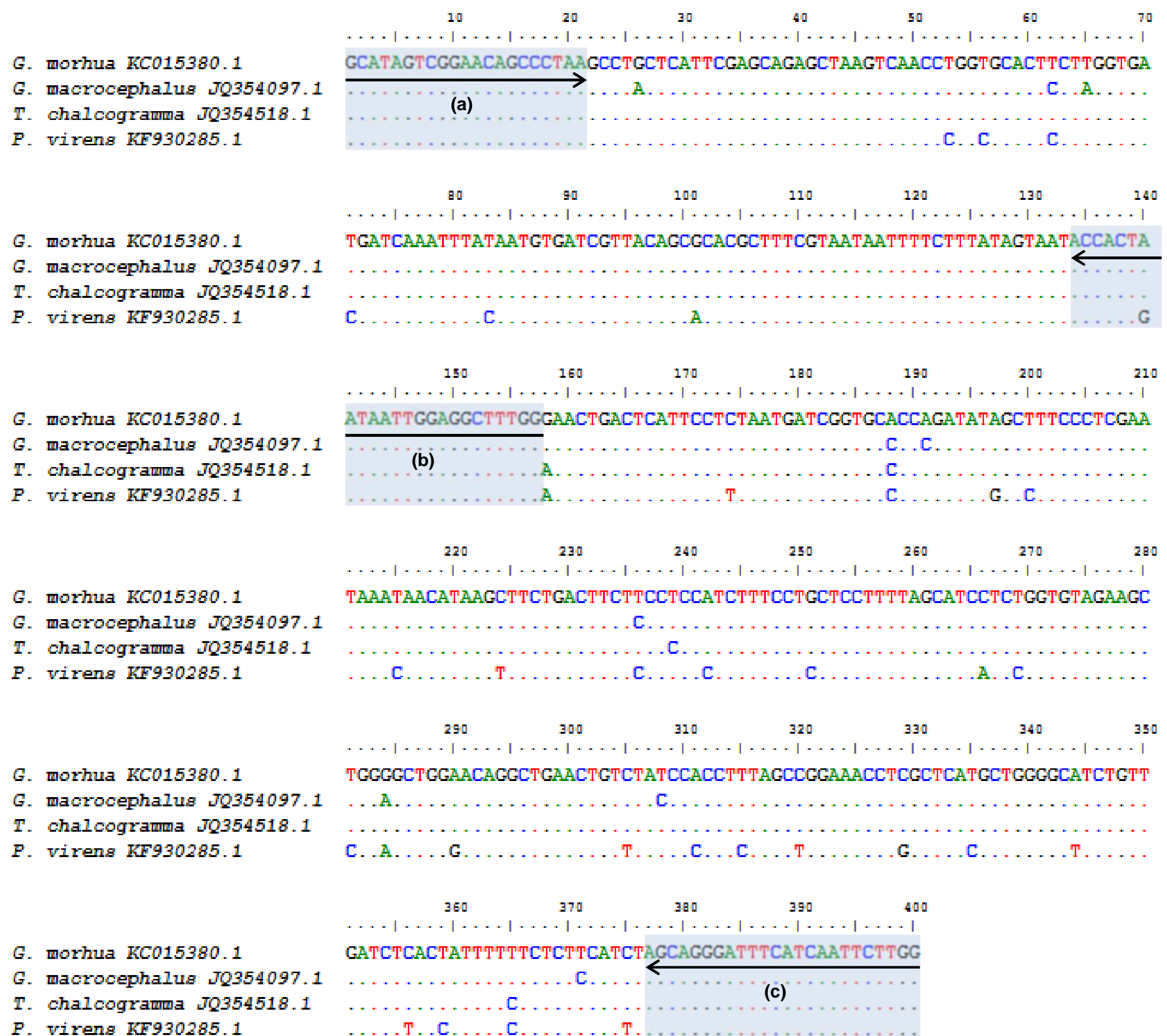


Figure 1. Sequence alignment of NCBI entries of COI locus showing nucleotide differences between species and primer design regions. Legend: a, Gad1COI-F; b, Gad1COI-R; and c, Gad2COI-R.

```
                                     10        20        30        40        50        60        70
G. morhua EU492141.1         GTAGGTGATGCCTTAGTTCAATGGATCTGAGGAGGTTTCTCAGTAGATAATGCTACCCTAACTCGGTTTT
G. macrocephalus AB078152.1  .......................A........G..C.................C...T.G..........
T. chalcogramma AB078151.1   .......................A........G..C.....................T.......A....
P. virens EU492147.1         .................C.....A.....G....................C..C.....G..G.....C.
                             (e) Gad1CytB-F: positions 1-23 →

                                     80        90       100       110       120       130       140
G. morhua EU492141.1         TTGCATTCCATTTCTTATTCCCCTTTGTTGTTGCTGCTTTTACAATACTCCACCTACTTTTTCTCCATGA
G. macrocephalus AB078152.1  .......................................................G........T.....
T. chalcogramma AB078151.1   .......T........................................................T.....
P. virens EU492147.1         ....................................................T..G..............

                                    150       160       170       180       190       200       210
G. morhua EU492141.1         AACAGGCTCAAATAATCCCACAGGAATCAATTCAAATGCAGACAAAATTCCATTCCACCCATATTTCACC
G. macrocephalus AB078152.1  ........................G....................G........................
T. chalcogramma AB078151.1   ...............C........T.....C....................C..T..............G
P. virens EU492147.1         ...............C........C.....C...........T.....CT....T.....T........T

                                    220       230       240       250       260       270       280
G. morhua EU492141.1         TACAAAGACCTACTTGGCTTTGCTGTGATGCTTCTGGGCTTAACCGCCCTCGCCCTCTTCGCACCTAATT
G. macrocephalus AB078152.1  ...........G..C...........A...........................................
T. chalcogramma AB078151.1   ...........G..............A..A..........................T.............
P. virens EU492147.1         ...........G.......................A..TC....TT....T........T.....C..CC

                                    290       300       310       320       330       340       350
G. morhua EU492141.1         TACTCGGAGATCCAGATAATTTCACCCCTGCTAACCCCATCGTTACCCCACCTCATGTTAAGCCCGAATG
G. macrocephalus AB078152.1  ....T...........................................................T.....
T. chalcogramma AB078151.1   .G..T.................T..G.......................T.................G..
P. virens EU492147.1         .G..A.....C...........T.....G....................T...........A.....G..

                                    360       370       380       390       400       410       420
G. morhua EU492141.1         ATATTTCTTGTTTGCCTATGCCATCTTACGCTCTATTCCAAATAAACTAGGTGGCGTACTTGCACTCCTA
G. macrocephalus AB078152.1  ..................C......C.............T......T.......................
T. chalcogramma AB078151.1   ..................C....................T...........C..T..............G
P. virens EU492147.1         ......TC................TC.............T..C........C.....C............

                                    430       440       450       460       470       480       490
G. morhua EU492141.1         TTCTCGATTCTAGTCCTCATGGTTGTACCCTTTCTCCATACGTCAAAACAACGAGGTTTAACATTCCGCC
G. macrocephalus AB078152.1  .....A...........T..A.....G.....C.....C..T...........G..C.............
T. chalcogramma AB078151.1   .....C...........T.......................T.......................T....
P. virens EU492147.1         .....C.....G..T...........C......T.A..C..T..............C.............

                                    500       510       520       530       540       550       560
G. morhua EU492141.1         CTCTTACCCAAATACTATTCTGAGTCCTCGTTGCAGATATACTAGTTCTTACATGAATTGGAGGCGTACC
G. macrocephalus AB078152.1  ...................T............................................A.....
T. chalcogramma AB078151.1   ....C.................G..T...........C................................
P. virens EU492147.1         .......T.....G.....T.....T...........C.....................C.....A..G..
                                                                    (d) Gad2CytB-F: positions 530-552 →

                                    570       580       590       600       610       620       630
G. morhua EU492141.1         TGTAGAGCACCCCTTCATTATCATCGGACAAGTGGCATCAGTACTATATTTCTCCCTCTTCCTAGTTTTA
G. macrocephalus AB078152.1  ......A................................T.................A............
T. chalcogramma AB078151.1   ......A..T................................G..............A............
P. virens EU492147.1         ......A........T..........................G........T.....A............

                                    640       650       660
G. morhua EU492141.1         TTCCCCCTTGCAGGAATAACTGAAAATAAGGCC
G. macrocephalus AB078152.1  .................................
T. chalcogramma AB078151.1   ..T..............................
P. virens EU492147.1         .................................
                                     ← (f) Gad1CytB-R: positions 639-663
```

Figure 2. Sequence alignment of NCBI entries *cytb* gene showing nucleotide differences between species and primer design regions. Legend: d, Gad2CytB-F; e, Gad1CytB-F; and f, Gad1CytB-R.

Table 1. Oligonucleotide primers for qualitative and real-time PCR.

Primer		Target	Sequence (5' →3')	Amplicon (bp)	
				HRM	Sequencing
Gad1COI-F	(a)	cytochrome c oxidase subunit I (COI)	GCATAGTCGGAACAGCCCTAA	157[a,b]	400 [a,c]
Gad1COI-R	(b)		CCAAAGCCTCCAATTATTAGTGGT		
Gad2COI-R	(c)		CCAAGAATTGATGAAATCCCTGCT		
Gad1CytB-F	(d)	cytochrome b (*cytb*)	TACTAGTTCTTACATGAATTGGAGG	134 [e,f]	663 [d,f]
Gad2CytB-F	(e)		GTAGGTGATGCCTTAGTTCAATG		
Gad1CytB-R	(f)		GGCCTTATTTTCAGTTATTCCTGCA		

DNA Extraction

DNA from all samples was extracted with a Nucleospin® Food kit (Macherey-Nagel, Düren, Germany), according to the manufacturer's instructions, using 200 mg of each sample. The extractions were performed in duplicate. All extracts were kept at -20 °C until further analysis.

DNA quality

Yield and purity of extracts were assessed using UV spectrophotometric DNA quantification on a Synergy HT multi-mode microplate reader (BioTek Instruments, Inc., Vermont, USA) fitted with a Take3 micro-volume plate accessory. The DNA content was determined using the nucleic acid quantification protocol with sample type defined as double-strand DNA in Gen5 data analysis software version 2.01 (BioTek Instruments, Inc., Vermont, USA). The quality of extracted DNA was also analysed by electrophoresis in a 1 % agarose gel containing 1× Gel Red (Biotium, Hayward, CA, USA) for staining and carried out in 1× STGB (GRISP, Porto, Portugal) for 25 min at 200 V. The agarose gel was visualised under a UV light tray Gel Doc™ EZ System (Bio-Rad Laboratories, Hercules, CA, USA) and a digital image was obtained with Image Lab software version 5.1 (Bio-Rad Laboratories, Hercules, CA, USA).

PCR amplification

For sequencing COI and *cytb* amplicons, PCR amplifications were carried out in duplicate with 50 μL of total reaction volume containing 4 μL of DNA extract (200 ng), 67 mM of Tris-HCl (pH 8.8), 16 mM of $(NH_4)_2SO_4$, 0.01% of Tween 20, 200 μM of each dNTP, 1.0 U of SuperHot Taq DNA Polymerase (Genaxxon Bioscience, Ulm, Germany), 3.0 mM of $MgCl_2$ and 200 nM of each primer (Gad1COI-F/Gad2COI-R or Gad2CytB-F/Gad1Cytb-R) (Table 1). The amplifications were performed in a MJ Mini™ Gradient

Thermal Cycler (Bio-Rad Laboratories, Hercules, CA, USA) according to the temperature programs: initial denaturation at 95 ºC for 5 min; 40 cycles at 95 ºC for 30 s, 60 ºC for 45 s, 72 ºC for 1 min or 1.5 min (for Gad1COI-F/Gad2COI-R or Gad2CytB-F/Gad1Cytb-R, respectively); and a final extension at 72 ºC for 5 min.

For the amplification of target fish species, and cross-reactivity testing with other animal and plant species, PCR components were the same but the total reaction volume was 25 µL containing 2 µL of DNA extract (100 ng) and 200 nM of each primer (Gad1COI-F/Gad1COI-R or Gad1CytB-F/Gad1Cytb-R) (Table 1). The temperature program was the same.

Amplified fragments were analysed by electrophoresis, in a 1.5% agarose gel containing 1× Gel Red (Biotium, CA, USA) for staining, in 1× SGTB buffer (GRISP, Porto, Portugal) for 20-25 min at 200 V. The agarose gel was visualised under UV (Gel Doc™ EZ System, Bio-Rad Laboratories, Hercules, USA) and a digital image was obtained with Image Lab software version 5.1 (BioRad Laboratories, Hercules, CA, USA).

Real-time PCR and HRM analysis

Real-time PCR assays were carried out in 20 µL of total reaction volume containing 2 µL of DNA (10 ng), 1× SsoFast™ Evagreen® Supermix (Bio-Rad Laboratories, Hercules, CA, USA) and 300 nM of each primer (Gad1COI-F/Gad1COI-R or Gad1CytB-F/Gad1Cytb-R) (Table 1). The real-time PCR runs were performed on a fluorometric thermal cycler CFX96 Real-time PCR Detection System (Bio-Rad Laboratories, Hercules, CA, USA) using the following conditions: 95 ºC for 5 min; 45 cycles at 95 ºC for 15 s and 60 ºC for 45 s, with collection of fluorescence signal at the end of each cycle. Data were processed using the software Bio-Rad CFX Manager 3.1 (Bio-Rad Laboratories, Hercules, CA, USA).

For HRM analysis, PCR products were denatured at 95 ºC for 1 min and held at 70 ºC for 5 min, in order to promote the correct annealing of the DNA duplexes. These steps were followed by melting from 70 ºC up to 95 ºC with temperature increments of 0.2 ºC every 10 s. Fluorescence data were acquired at the end of each melting phase and processed using Precision Melt Analysis Software 1.2 (Bio-Rad Laboratories, Hercules, CA, USA) to generate melting curves, as a function of temperature, and difference curves for easy visual identification of clusters. Melting curve shape sensitivity determines the stringency used to classify into different clusters, while temperature of melting (Tm) difference threshold determines the lowest Tm difference between samples. In both cases, cluster detection parameters were set to high sensitivity and threshold yields, with the aim of providing more heterozygote clusters (Bio-Rad, 2012). Thus, melting curve

shape sensitivity was adjusted to the default value of 50% and the Tm difference threshold was set as 0.20.

Precision Melt Analysis software determined a probability distribution for each cluster, based on the standard deviation of melt curves within the same cluster. Each sample was mapped on to each cluster probability distribution, based on similarity of the sample to the mean melt curve across each sample in the cluster. The calculated confidence value indicates the relative probability of the sample being included a certain cluster (Bio-Rad, 2012).

DNA extracts from gadoid fish species and commercial samples were analysed in triplicate (*n*=3) in two independent assays.

Sequencing of PCR products

PCR products of *G. morhua*, *G. macrocephalus, T. chalcogramma* and *P. virens* were purified using a Jetquick PCR purification kit (Genomed, Löhne, Germany) to remove any possible interfering components. The purified products were sent to a specialised research facility (STABVIDA, Lisbon, Portugal) for sequencing. Each target fragment was sequenced twice, performing the direct sequencing of both strands in opposite directions, which allowed production of four complementary sequences of high quality. The sequencing data were analysed using BioEdit v7.2.5 (Ibis Biosciences, Carlsbad, CA, USA) and FinchTV (Geospiza, Seattle, WA, USA).

Results and Discussion

Specificity

Prior to the specific amplification of COI and *cytb* genes, all DNA extracts were tested for amplifiability with universal eukaryotic primers 18SRG-F/18SRG-R, as described by Costa, Mafra and Oliveira (2013). All samples tested positively as producing the expecting 113 bp fragment, which confirmed the absence of any false negative results that might occur as result of potential PCR inhibition or ineffective DNA extraction.

Specificity assays were performed for each set of primers, using a total of 34 extracts that encompassed different untargeted species from plants (soybean, oat, rye, mint, wheat, lupine, maize, rice, pumpkin seeds, rapeseed, sunflower, tomato, peach, nectarine, apricot, strawberry, raspberry), animals (rabbit, deer, cow, chicken, turkey, pork, lamb, goat, ostrich) and other seafood (Atlantic salmon, gilt-head bream, ray, common sole, European pilchard; common shrimp, yellowfin tuna, squid). For each set of primers, no cross-reactivity was found (data not shown), confirming the specificity of the

primers for COI and *cytb loci* for the Gadidae family. The four species (*G. morhua*, *G. macrocephalus, T. chalcogramma* and *P. virens*) were amplified successfully using qualitative PCR targeting for both COI (157 pb) and *cytb* (134 pb) DNA mini-barcodes (data not shown).

Real-time PCR coupled with HRM analysis

COI mini-barcode

A real-time PCR assay using EvaGreen dye coupled with HRM analysis was developed targeting the COI region as a mini-barcode to differentiate four genetically related gadoid species. The results show that, based on the conventional melting analysis of COI amplicons (Figure 3A), two groups of melt peaks could be observed. Despite clear identification of a melt peak at 82 ºC, enabling differentiation of *P. virens*, a multitude of melt curve profiles that corresponded to the other three gadoid species was observed at approximately 80.5 ºC.

To increase the specificity of the assay, HRM analysis was applied, aiming to discriminate the four species clearly. The results allowed identification of three distinct clusters (Figure 3B), from which *P. virens* species was classified as a distinct group (cluster 3), confirming the conventional melting analysis (Figure 3A). From the multitude of melting profiles at 80.5 ºC, two distinct clusters were identified using HRM analysis with a level of confidence greater than 99.4% (Figure 3B). Cluster 1 comprised *G. morhua* and *T. chalcogramma* species, suggesting similar size and nucleotide composition of the amplicons. Previously included in the melt curves at 80.5 ºC, *G. macrocephalus* was grouped in a distinct cluster (cluster 2), allowing its differentiation from the other gadoid species (Figure 3B).

HRM analysis is based on minimal differences of amplicon size and composition, providing more detailed information, when compared to conventional melting analysis. In fact, a simple variation in one nucleotide could be enough to differentiate two otherwise identical sequences. PCR products are grouped in different clusters on the basis of similar size and composition (Costa et al., 2012; Druma et al., 2014; McGlauflin et al., 2010). Accorddingly, the results obtained with HRM analysis were in good agreement with the *in silico* analysis. Considering that PCR products are included in clusters based on to size and nucleotide composition, *G. morhua* and *T. chalcogramma* were classified in the same group since the amplicons did not present any nucleotide differences (Figure 1). *G. macrocephalus* was included in cluster 2 based on a three nucleotide difference with cluster 1. *P. virens* was also classified as cluster 3 owing to a seven-nucleotide difference with sequences from cluster 1.

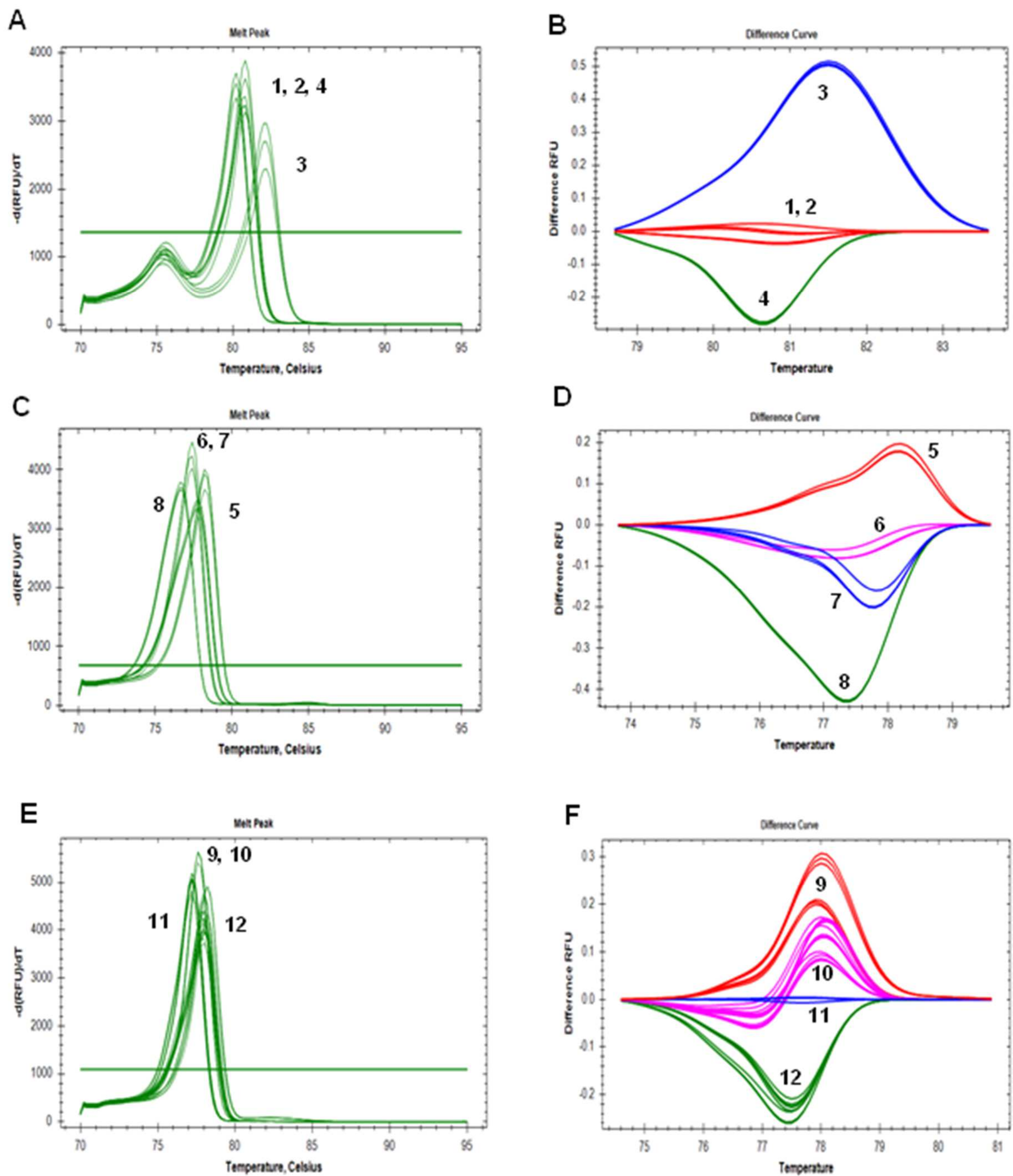


Figure 3. Conventional melting analysis (A, C, E) and HRM analysis (B, D, F) of real-time PCR products targeting COI (A, B) and *cytb* (C, D, E, F) regions. Legend: 1 and 5, *G. morhua*; 2 and 6, *T. chalcogramma*; 3 and 7, *P. virens*; 4 and 8, *G. macrocephalus*; 9, *G. morhua* and commercial sample (*n*=1, "Brás" style codfish); 10, *T. chalcogramma* and commercial samples (*n*=3, codfish cakes, fish fillet and shellfish cream); 11, *P. virens*; 12, *G. macrocephalus* and commercial samples (*n*=2, codfish pie and fresh codfish sticks).

Cytb mini-barcode

Similarly, a real-time PCR assay coupled with HMR analysis targeting the *cytb*, as a mini-barcode region was proposed to differentiate the four gadoid species. Using conventional melting analysis, *cytb* amplicons presented three groups of melt peaks at 76.5-78 ºC (Figure 3C). Despite exhibiting peaks that differentiated *G. morhua* and *G. macrocephalus*, conventional melting analysis did not allow *P. virens* and *T. chalcogramma* to be distinguished. With the HRM analysis, *cytb* products were classified in four clusters, which enabled differentiation of all species with a high level of confidence (>99.3%) (Figure 3D). As in the case of COI mini-barcode, the HRM analysis results were in good accordance with analysis *in silico* for the *cytb* gene. With *G. morhua* as the reference cluster (cluster 1), *G. macrocephalus* was grouped in cluster 2 based on a four-nucleotide difference between the sequences. *T. chalcogramma* and *P. virens* were classified in clusters 3 and 4 based on five- and eight-nucleotide differences with the reference sequence for *G. morhua* (Figure 2).

Sequencing data

The expected COI and *cytb* products from real-time PCR assays were designed to have lengths of 157 bp and 134 bp, which are relatively small for direct sequencing. To overcome this, larger fragments from both genes were amplified, allowing accurate sequencing of the target regions contained therein. Therefore, COI and *cytb loci* were amplified using specific primers (Table 1), which resulted in 400 and 663 bp amplicons, respectively (Figures 1 and 2). To confirm the identity of the specimens, the four species under study were sequenced and their alignment with sequences retrieved from NCBI are available in Figures S1 and S2. Table 2 summarises data from qualitative PCR, real-time PCR with HRM analysis, the number of nucleotide differences between sequencing results and *G. morhua* sequence, the number of nucleotide differences between sequencing data and NCBI sequences and BLAST data from sequencing. Sequencing results confirmed all base differences within the target region (157 bp and 134 bp, respectively) for the COI and *cytb* amplicons.

Sequencing results for COI from different species (400 bp) were in good agreement with COI sequences retrieved from the NCBI database (Figure S1, see supplementary material). The alignments showed almost no differences, except for 1 nucleotide in the cases of *G. morhua* and *G. macrocephalus* (Table 2). However, this small nucleotide difference was not located in the 157 bp region used for HRM analysis. Regarding *cytb* (663 bp), minor nucleotide differences could be found between sequencing and the relevant NCBI accessions (Figure S2, see supplementary material), namely 1, 6 and 3 nucleotide differences for *G. morhua*, *T. chalcogramma* and *P. virens*, respectively (Table 2), but again this was not located in the 134 bp target region.

Table 2. Summary of PCR results, HRM cluster analysis, sequence comparison and BLAST data targeting the COI and *cytb* genes for the tested species (*Gadus morhua, G. macrocephalus, Theragra chalcogramma* and *Pollachius virens*).

Species	Qualitative PCR		HRM analysis (level of confidence %)[a]		Differences between sequencing data and *G. morhua* sequence		Differences between sequencing data and NCBI sequences		NCBI sequences		BLAST (Identity %)	
	COI	*cytb*	COI	*cytb*	COI	*cytb*	COI	*cytb*	COI	*cytb*	COI	*cytb*
G. morhua	+	+	Cluster 1 (99.4±0.4%)	Cluster 1 (100.0±0.0%)	1	1	1	1	KC015380.1	EU492141.1	100	100
G. macrocephalus	+	+	Cluster 2 (100.0±0.1%)	Cluster 2 (99.7±0.2%)	10	31	1	0	JQ354097.1	AB078152.1	100	100
T. chalcogramma	+	+	Cluster 1 (99.4±0.4%)	Cluster 4 (99.7±0.2%)	4	41	0	6	JQ354518.1	AB078151.1	100	99
P. virens	+	+	Cluster 3 (99.9±0.1%)	Cluster 3 (99.3±0.4%)	31	66	0	3	KF930285.1	EU492147.1	100	99

[a] mean ± standard deviation of level of confidence (%) of *n*=6 replicates from two independent real-time PCR runs.

Analysis of commercial samples

The *cytb* barcode-HRM approach offered the best performance for differentiation of the four species and, thus, was applied to authenticate commercial samples. Table 3 lists the fish-products analysed, the results for qualitative PCR and HRM cluster analysis targeting the *cytb* gene, and a comparison with the label. Most of the samples were positive for qualitative PCR amplification, indicating the presence of gadoid species. The 12 negative samples were labelled as containing hake, fish and/or shrimp and not as containing gadoid fish species, so these results do not imply non-compliance with labelling statements. For the 30 positive samples, the *cytb* barcode-HRM approach developed was applied successfully to identify gadoid species (Figure 3C and 3D; Table 3). Species identification was achieved with levels of confidence ranging from 95.2 to 99.8 %, and a mean value of 99.0%, which was a very satisfactory outcome considering the complexity of the food matrices and degree of processing. *G. morhua* was identified in two of the samples, although it was declared in six. In contrast, *G. macrocephalus* was identified in seven samples, with only two declaring its presence, and *T. chalcogramma* was identified in 17 samples, but only 10 declared it on the label. Likewise, *P. virens*, declared in only one sample, was identified in four products. Of the nine samples labelled with “codfish”, most contained *G. macrocephalus* (6), two contained *T. chalcogramma* and one *G. morhua*. In summary, 17 of 30 samples (57%) containing gadoid species were in accordance with labelled information, while nine (30%) did not comply with declared species or “codfish” designations. *T. chalcogramma* was identified in the four remaining gadoid-containing samples (labelled with “fish”).

According to Regulation (EC) No. 1379/2013, inclusion of scientific names in all processed fish samples is voluntary (excluding salted and smoked fish). Moreover, fish species that constitute an ingredient of another food may be named as "fish", provided the name and presentation of such does not refer to a specific species. This might explain the variation in information displayed on the labels of processed fish products: inclusion of the scientific name of species, common fish name and the presence of fish on the list of ingredients. The Portuguese legislation (Decreto-Lei nº 25/2005) states that *G. morhua*, *G. macrocephalus* and *G. ogac* must be named as Atlantic, Pacific and Greenland cod (rarely sold at local markets), respectively, or under the commercial denomination of cod or codfish (“bacalhau”). Crossing then information from European and national regulations, it is not clear whether it is correct or not to name codfish to these three species in commercial processed foodstuffs. In this study, the samples declaring “codfish” on the list of ingredients were labelled correctly when *G. morhua* or *G. macrocephalus* were identified.

Table 3. Analysed processed fish-containing samples, relevant labelled information and results of qualitative PCR and HRM cluster analysis targeting *cytb* gene.

Code	Commercial samples	Labelled species	PCR	HRM cluster	Level of confidence (%)[a]	Labelling compliance (gadoid species)
Typical Portuguese dishes						
1	Codfish with cream	*Gadus morhua*	+	*G. macrocephalus*	96.8 ± 1.5	No
2	Codfish with cream	(codfish)	+	*G. macrocephalus*	99.0 ± 0.4	Yes
3	Codfish with cream	(codfish)	+	*T. chalcogramma*	99.7 ± 0.3	No
4	Codfish cakes	*G. morhua* *G. macrocephalus*	+	*T. chalcogramma*	98.8 ± 0.9	No
5	Codfish cakes	*G. morhua*	+	*T. chalcogramma*	99.3 ± 0.3	No
6	Codfish cakes	(codfish)	+	*G. macrocephalus*	99.5 ± 0.4	Yes
7	"Gomes de Sá" style codfish	(codfish)	+	*G. morhua*	99.0 ± 0.3	Yes
8	"Brás" style codfish	(codfish)	+	*G. macrocephalus*	99.8 ± 0.0	Yes
9	"Brás" style codfish	*G. morhua* *G. macrocephalus*	+	*G. morhua*	99.2 ± 0.6	Yes
10	Codfish "patanisca"	*G. morhua*	+	*T. chalcogramma*	99.6 ± 0.2	No
11	Codfish "alheira"	(codfish)	+	*G. macrocephalus*	99.7 ± 0.1	Yes
12	Codfish pie	(codfish)	+	*G. macrocephalus*	99.3 ± 0.4	Yes
Fish sticks						
13	Fresh codfish sticks	(codfish)	+	*G. macrocephalus*	99.0 ± 0.6	Yes
14	Fish sticks	*Theragra chalcogramma* *Merluccius productus* *G. morhua* *Pollachius virens*	+	*P. virens*	98.6 ± 0.9	Yes
15	Fish sticks	*P. virens*	+	*P. virens*	99.6 ± 0.3	Yes
16	Fish sticks	*T. chalcogramma*	+	*P. virens*	98.6 ± 0.9	No
17	Fish sticks	*M. hubbsi*	+	*T. chalcogramma*	98.8 ± 0.6	No
18	Fish sticks	(hake)	-			--
19	Fish sticks	(hake)	-			--
20	Fish sticks	(hake)	-			--
21	Fish sticks	(hake)	-			--
22	Fish sticks	*T. chalcogramma*	+	*T. chalcogramma*	98.7 ± 0.4	Yes
23	Fish sticks	(hake)	-			--
Fish fillets & crunchies						
24	Fish fillet	*T. chalcogramma*	+	*T. chalcogramma*	99.0 ± 0.3	Yes
25	Fish fillet	*T. chalcogramma*	+	*T. chalcogramma*	98.7 ± 0.4	Yes
26	Crunchy hake	(hake)	-			--
27	Crunchy fish (Spinach)	*T. chalcogramma*	+	*T. chalcogramma*	99.3 ± 0.4	Yes
28	Crunchy fish (Tomato & Mozzarella)	*T. chalcogramma*	+	*P. virens*	98.2 ± 0.2	No
29	Breaded hake	(hake)	-			--
Patties						
30	Hake & Shrimp patties	(hake, shrimp)	-			--
31	Hake pattys	(hake)	-			--
32	Hake pattys	(hake)	-			--

[a] mean level of confidence ± standard deviation (%) of *n*=6 replicates from two independent real-time PCR runs.

Table 3. (continuation)

Code	Commercial samples	Labelled species	PCR	HRM cluster	Level of confidence (%)[a]	Labelling compliance (gadoid species)
Crab sticks						
33	Crab sticks	*T. chalcogramma*	+	*T. chalcogramma*	99.6 ± 0.3	Yes
34	Crab sticks (to fry)	*T. chalcogramma*	+	*T. chalcogramma*	99.5 ± 0.3	Yes
35	Crab stick	(fish)	+	*T. chalcogramma*	99.6 ± 0.3	--
Others						
36	Seafood “risotto”	(hake, fish broth)	+	*T. chalcogramma*	99.5 ± 0.5	--
37	“Paella”	(fish)	-			--
38	Codfish lasagne	(codfish)	+	*T. chalcogramma*	98.5 ± 0.4	No
39	Shellfish cream	(fish)	-			--
40	Shellfish cream	(fish)	+	*T. chalcogramma*	95.2 ± 1.9	--
41	Seafood soup	(fish)	+	*T. chalcogramma*	98.9 ± 0.4	--
42	Fish nuggets	*T. chalcogramma*	+	*T. chalcogramma*	99.4 ± 0.0	Yes

[a] mean level of confidence ± standard deviation (%) of *n*=6 replicates from two independent real-time PCR runs.

Conclusion

In the present work, two DNA mini-barcodes (COI and *cytb*) were evaluated extensively for authentication of codfish and cod-like species. Both mini-barcodes allowed the development of two real-time PCR systems coupled with HRM analysis for discrimination of Gadidae species. The *cytb* barcode-HRM offered the best performance, enabling discrimination of the four fish species in different clusters. This approach was applied successfully to identify gadoid species in commercial fish-containing foods showing that, of 42 analysed samples, 30 contained Gadidae species, 17 of which were in accordance with the labelled species, nine (30%) of which did not comply with declared species or “codfish” designation, and four of which, labelled as “fish”, contained *T. chalcogramma*. Among the mislabelled cases, *G. morhua* seemed to have been replaced by other lower-cost species, namely *T. chalcogramma*.

In summary, a novel *cytb* barcode-HRM was proposed as a simple, rapid, specific and high-throughput tool for Gadidae species identification in foods. The proposed system overcame drawbacks associated with sequencing, enabling reduced costs per sample and time of analysis to 2-3 hours. To our knowledge, this is the first attempt to discriminate fish members from Gadidae family by means of real-time PCR combined with HRM analysis for authentication purposes.

Acknowledgements

This work was supported by FCT (Fundação para a Ciência e Tecnologia) through project UID/QUI/50006/2013 – POCI/01/0145/FEDER/007265 with financial support from

FCT/MEC through national funds and co-financed by FEDER, under the Partnership Agreement PT2020 and by the project NORTE-01-0145-FEDER-000011. T. J. R. Fernandes and J. Costa are grateful to FCT grants (SFRH/BD/93711/2013 and SFRH/BPD/102404/2014, respectively) financed by POPH-QREN (subsidised by FSE and MCTES). The gadoid species were kindly provided by Pascoal & Filhos SA.

References

Armani, A., Guardone, L., 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.

Bio-Rad (2012). Precision Melt Analysis software – Instruction Manual (http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_10014811.pdf).

Calo-Mata, P., Sotelo, C. G., Pérez-Martín, R. I., Rehbein, H., Hold, G. L., Russel, V. J., Pryde, S., Quinteiro, J., Rey-Méndez, M., Rosa, C., & Santos, A. T. (2003). Identification of gadoid fish species using DNA-based techniques. *European Food Research and Technology, 217*, 259-264.

Cline, E. (2012). Marketplace substitution of Atlantic salmon for Pacific salmon in Washington State detected by DNA barcoding. *Food Research International, 45*, 388-393.

Costa, J., Mafra, I., & Oliveira, M. B. P. P. (2012). High resolution melting analysis as a new approach to detect almond DNA encoding for Pru du 5 allergen in foods. *Food Chemistry, 133*, 1062-1069.

Costa, J. S. B. (2013). Development and application of molecular-based methods for the detection of tree nut allergens: the cases of almond, hazelnut and walnut, PhD Thesis, Faculty of Pharmacy, University of Porto, Porto.

Costa, J., Oliveira, M. B. P. P., & Mafra, I. (2013). Effect of thermal processing on the performance of the novel single-tube nested real-time PCR for the detection of walnut allergens in sponge cakes. *Food Research International*, *54*, 1722-1729.

Decreto-lei nº 25/2005. Ministério da Agricultura, Pescas e Florestas. *Diário da República*, *I série, nº 20*, 696-703.

Di Pinto, A., Di Pinto, P., Terio, V., Bozzo, G., Bonerba, E., Ceci, E., & Tantillo, G. (2013). DNA barcoding for detecting market substitution in salted cod fillets and battered cod chunks. *Food Chemistry, 141*, 1757-1762.

Druml, B., & Cichna-Markl, M. (2014). High resolution melting (HRM) analysis of DNA – Its role and potential in food analysis. *Food Chemistry, 158*, 245-254.

FAO (2014). Food and Agriculture Organization of the United Nations, Rome, Italy. Available at: http://www.fao.org/ (Last accessed 28 September 2015).

Fitzcharles, E. M. (2012). Rapid discrimination between four Antarctic fish species, genus Macrourus, using HRM analysis. *Fisheries Research, 127–128*, 166-170.

Hebert, P. D. N., Cywinska, A., Ball, S. L., & de Waard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 270*, 313-321.

Hellberg, R. S., Kawalek, M. D., Van, K. T., Shen, Y., & Williams-Hill, D. M. (2014). Comparison of DNA extraction and PCR setup methods for use in high-throughput DNA barcoding of fish species. *Food Analytical Methods*, 7, 1950-1959.

Madesis, P., Ganopoulos, I., Sakaridis, I., Argiriou, A., & Tsaftaris, A. (2014). Advances of DNA-based methods for tracing the botanical origin of food products. *Food Research International*, 60, 163-172.

McGlauflin, M. T., Smith, M. J., Wang, J. T., Young, S. F., Chen, N., Lee, Y. C., Pascal, C., Seeb, L. W., Stevens, J., & Seeb, J. E. (2010). High-resolution melting analysis for the discovery of novel single-nucleotide polymorphisms in rainbow and cutthroat trout for species identification. *Transactions of the American Fisheries Society, 139*, 676-684.

Miller, D. D., & Mariani, S. (2010). Smoke, mirrors, and mislabeled cod: poor transparency in the European seafood industry. *Frontiers in Ecology and the Environment, 8*, 517-521.

Mueller, S., Handy, S. M., Deeds, J. R., George, G. O., Broadhead, W. J., Pugh, S. E., & Garrett, S. D. (2015). Development of a COX1 based PCR-RFLP method for fish species identification. *Food Control, 55*, 39-42.

Rasmussen, R. S., & Morrissey, M. T. (2008). DNA-based methods for the identification of commercial fish and seafood species. *Comprehensive Reviews in Food Science and Food Safety, 7*, 280-295.

Regulation (EU) No 1169/2011 of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No 1924/2006 and (EC) No 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004. *Official Journal of the European Union*, *L304*, 18-63.

Regulation (EC) No 206/2009 of 5 March 2009 on the introduction into the Community of personal consignments of products of animal origin and amending Regulation (EC) No 136/2004. *Official Journal of the European Union*, *L77*, 1-19.

Regulation (EC) No 104/2000 of 17 December 1999 on the common organization of the markets in the fishery and aquaculture products. *Official Journal of the European Communities*, *L17*, 22-52.

Regulation (EU) No 1379/2013 of the European Parliament and of the Council of 11 December 2013 on the common organisation of the markets in fishery and aquaculture products, amending Council Regulations (EC) No 1184/2006 and (EC) No 1224/2009 and repealing Council Regulation (EC) No 104/2000. *Official Journal of the European Union*, *L 354*, 1-21.

Rehbein, H. (2013). Differentiation of fish species by PCR-based DNA analysis of nuclear genes. *European Food Research and Technology, 236*, 979-990.

Sampels, S. (2015). The effects of processing technologies and preparation on the final quality of fish products. *Trends in Food Science and Technology, 44*, 131-146.

Wetten, O. F., Wilson, R. C., & Andersen, Ø. (2012). High-resolution melting analysis of common and recombinant genotypes of the Atlantic cod (*Gadus morhua*) hemoglobin β1 gene in transatlantic populations. *Canadian Journal of Fisheries and Aquatic Sciences, 69*, 525-531.

Wong, E. H. K., & Hanner, R. H. (2008). DNA barcoding detects market substitution in North American seafood. *Food Research International, 41*, 828-837.

Yancy, H. F., Zemlak, T. S., Mason, J. A., Washington, J. D., Tenge, B. J., Nguyen, N.-L. T., Barnett, J. D., Savary, W. E., Hill, W. E., Moore, M. M., Fry, F. S., Randolph, S. C., Rogers, P. L., & Hebert, P. D. N. (2008). Potential Use of DNA Barcodes in Regulatory Science: Applications of the Regulatory Fish Encyclopedia. *Journal of Food Protection, 71*, 210-217.

Food Control 82 (2017) 8–17

Contents lists available at ScienceDirect

Food Control

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

High resolution melting analysis of a COI mini-barcode as a new approach for Penaeidae shrimp species discrimination

Telmo J.R. Fernandes, Catarina R. Silva, Joana Costa, M. Beatriz P.P. Oliveira, Isabel Mafra*

REQUIMTE-LAQV, Faculdade de Farmácia, Universidade do Porto, Rua de Jorge Viterbo Ferreira, 228, 4050-313 Porto, Portugal

*Corresponding author: Tel: +351 220428640. Fax: +351 226093390. E-mail: isabel.mafra@ff.up.pt

Abstract

Penaeidae family includes shrimp species of high commercial value, sharing noticeable morphological similarities, which makes them potential targets of adulteration. Therefore, mechanisms for authentication and certification of such crustaceans, frequently included in processed foods, constitute a benefit for the food industry. *Litopenaeus vannamei*, *Penaeus monodon, Fenneropenaeus indicus, Metapenaeus affinis* and *Melicertus kerathurus* are some of the most relevant penaeid shrimps, being their differentiation of high importance. This work intended to develop a new approach for the specific detection and differentiation of those five closely related shrimp species based on high resolution melting (HRM) analysis targeting a cytochrome oxidase subunit I (COI) mini-barcode. The method enabled the differentiation of the five species with high levels of confidence (>99%), being successfully applied to analyse processed seafood samples. *F. indicus* and *L. vannamei* were the main identified species in the commercial products. When verifying labelling compliance, four samples suggest adulterations based on the complete or partial substitution of declared species. The proposed method proved to be a potential tool for the rapid and cost-effective differentiation of penaied shrimp species.

Keywords: DNA barcoding; high resolution melting analysis; Penaeidae; seafood; crustaceans; authentication.

Introduction

Seafood is currently a significant market niche, being one of the most common protein sources consumed worldwide. It has also established itself as one of the top food categories associated with fraud, presenting nowadays a considerable rate of mislabelling (Galal-Khallaf, Ardura, Borrell, & Garcia-Vasquez, 2016; Rittenschober, Stadlmayr, Nowak, Du, & Charrondiere, 2016; Tagliavia et al., 2016).

In fact, food authentication is presently of utmost importance in order to avoid commercial fraud as well as to prevent hazards, such as allergies or intoxications from undeclared ingredients (Pascoal et al., 2011; Ortea et al., 2012). In the European Union, the recent regulation (Regulation (EC) No. 1379/2013) states that "*Crustaceans, whether in shell or not, live, fresh, chilled, frozen, dried, salted or in brine; crustaceans, in shell, cooked by steaming or by boiling in water, whether or not chilled, frozen, dried, salted or in brine; flours, meals and pellets of crustaceans, fit for human consumption*" should be labelled with the commercial designation and scientific name, fishing gear category, net weight, food operator, identification mark, production method, catch area and storage conditions. On the other hand, products such as canned, composite products, breaded products, among others, must be labelled with the name of the food, net quantity, food operator, identification mark, list of ingredients, "best before"/"use by date" and storage conditions, being no other information compulsory. In spite of the implemented legislation, seafood adulterations usually take place by unintentional or deliberate species substitutions, generally through the replacement of high-valued species by low-priced ones (Chin, Adibah, Hariz, & Azizah, 2016; Pascoal, Barros-Velásquez, Ortea, Cepeda, & Calo-Mata, 2008; Shen, Kang, Chen, & He, 2016).

Shrimps are one the most appreciated seafood groups all over the world, accounting for 7.8 million tons (aquaculture and captured) in 2013, which resulted in more than 50% of crustacean yield (FAO, 2016). Particularly, penaeid shrimps (belonging to Penaeidae family of the Decapoda order) represent more than 30% of the worldwide demand of crustaceans, comprising many commercially relevant species, such as whiteleg shrimp (*Litopenaeus vannamei*), giant tiger prawn (*Penaeus monodon*), Indian white prawn (*Fenneropenaeus indicus*), jinga shrimp (*Metapenaeus affinis*), and striped shrimp (*Melicertus kerathurus* also known as *Penaeus kerathurus*) (Pascoal et al., 2008). The phenotypic similarities of shrimps and the absence of external anatomical parts (e.g. carapace) after food processing make their morphological identification almost impossible, explaining why they are often regarded as adulteration targets.

Molecular tools relying on biomarkers targeting proteins or DNA have been proposed as suitable strategies for seafood species identification. Electrophoretic techniques, such

as sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and isoelectric focusing (IEF) have been used to differentiate shrimp species, but they frequently fail due to the lack of protein stability upon food processing. Immunochemical assays have also been applied, even in thermally treated samples, but are mostly generic for crustacean species and not useful for shrimp species differentiation (Ortea et al., 2012). More recently, approaches based on proteomics has been used for the characterisation of species-specific peptide markers that enabled the unequivocal identification of shrimp species (Ortea et al., 2012). On the other hand, the limitations of electrophoretic and immunochemical assays have been overcome with the introduction of methods based on DNA analysis, without requiring advanced mass spectrometry (MS) platforms as for proteomics. Thus, DNA-based techniques, in particular polymerase chain reaction (PCR), have been increasingly used for the identification and differentiation of closely related species, cultivars, genetically modified organisms and food allergens (Arroyave & Stiassny, 2014; Druml & Cichna-Markl, 2014; Madesis, Ganopoulos, Sakaridis, Argiriou, & Tsaftaris, 2012, Mafra, Ferreira, & Oliveira, 2008; Tagliavia et al., 2016). High resolution melting (HRM) analysis has been highlighted as a new alternative methodology. This post-PCR analysis, based on the shape of melting transitions of real-time PCR products, has been currently applied to food authentication, allowing the discrimination of little variations, such as a single base change on DNA sequence (Druml & Cichna-Markl, 2014).

In foods, HRM analysis has been applied in two main target regions: microsatellites or DNA barcodes (Costa, Mafra, & Oliveira, 2012; Druml & Cichna-Markl, 2014; Eischeid, Stadig, Handy, Fry, & Deeds, 2016; Fitzcharles, 2012; Jaakola, et al., 2010; Mackay, Wright, & Bonfiglioli, 2008; Sakaridis, Ganopoulos, Argiriou, & Tsaftaris, 2013). DNA barcodes are considered as some of the most reported species-specific markers, being the mitochondrial region of the cytochrome oxidase subunit I (COI) gene used in animal species differentiation including seafood (Galal-Khallaf et al., 2016, Haye, Segovia, Vera, Gallardo, Gallardo-Escárate, 2012). However, DNA barcoding itself cannot be regarded as a cost-effective tool since it has been intrinsically dependent on sequencing. To overcome this disadvantage, real-time PCR coupled to HRM analysis can be advanced as a faster, reliable and more economical alternative to exploit DNA barcoding.

The aim of this work was to develop a robust and specific new approach for the rapid detection and discrimination of five closely related shrimp species from the Penaeidae family. A DNA barcode from the COI region of the selected species (*L. vannamei, P. monodon, F. indicus, M. affinis* and *M. kerathurus*) was *in silico* analysed to identify genetic variability among them, thus enabling their differentiation. This variability was exploited by real-time PCR coupled to HRM analysis targeting a mini-barcode region to

propose a rapid and cost-effective methodology for the identification and differentiation of penaied shrimp species. Afterwards, the methodology was applied to test shrimp-containing foods in order to verify its feasibility for species identification and to verify the label statements.

Materials and Methods

Sample collection and preparation

Shrimp species (whiteleg shrimp, *L. vannamei* (*n*=4), giant tiger prawn, *P. monodon* (*n*=4), Indian white prawn, *F. indicus* (*n*=4), jinga shrimp, *M. affinis* (*n*=4), striped shrimp, *M. kerathurus* (*n*=4)) were kindly supplied by Marfresco (Loures, Portugal) and Brasmar Seafood Companies (Trofa, Portugal), whose identity was further confirmed by sequencing.

A total of 58 species were acquired at local market to evaluate the specificity of the proposed method including fish, crustaceans and molluscs (*n*=25) (Atlantic cod – *Gadus morhua,* Alaska pollock – *Theragra chalcogramma,* saithe - *Pollachius virens*, common ling – *Molva molva*, Herring hake – *Merluccius merluccius*, North Pacific hake – *Merluccius productus*, Argentine hake – *Merluccius hubbsi*, Whiting – *Merluccius capensis*, Atlantic salmon - *Salmo salar*, gilt-head bream - *Sparus aurata*, ray - *Raja* spp., common sole – *Solea solea*, European pilchard – *Sardina pilchardus*, yellowfin tuna – *Thunus albacares*, squid – *Loligo* spp., mussel – *Mytilus* spp., undulated surf clam – *Paphia undulata*, crab – *Portunus validus*, edible crab – *Cancer pagurus*, Caribbean spiny lobster – *Panulirus argus*, Norway lobster – *Nephrops norvegicus*, Argentine red shrimp – *Pleoticus muelleri*, Razor mud shrimp – *Solenocera melantho*, Udang merah – *Solenocera crassicornis*, Knife shrimp - *Haliporoides triarthrus*); other animals (*n*=9) (rabbit, deer, cow, chicken, turkey, pork, lamb, goat, ostrich) and plants (*n*=24) (soybean, oat, rye, mint, wheat, lupine, maize, rice, pumpkin seeds, rapeseed, sunflower, tomato, peach, nectarine, apricot, strawberry, raspberry, lemongrass, dandelion, fennel, honey, cashew, heather, lavender). Moreover, 33 frozen and/or processed seafood products were acquired at local markets, comprising typical Portuguese (shrimp with beans “feijoada”, shrimp “açorda”) and other pre-cooked (risotto, pizza, paella) dishes, seafood soups and sauces (powders), shrimp kernels, among others.

The samples were ground and homogenised separately in a laboratory knife mill Grindomix GM200 (Retsch, Haan, Germany) and stored at -20 ºC until further analysis. Prior to grinding process, whenever it was possible, frozen/processed samples were divided in two portions: one including only the seafood itself (e.g. shrimp filling) and another including the entire foodstuff (e.g. shrimp patty). All containers and material used

during this procedure were previously treated with DNA decontamination solution (DNA-ExitusPlus™, AppliChem, Darmstadt, Germany).

In silico DNA barcode analysis and primer design

L. vannamei, P. monodon, F. indicus, M. affinis and *M. kerathurus* COI sequences were obtained from NCBI database with the respective accession numbers (KJ679916.1, KM528138.1, KF604889.1, KJ879298.1 and KC789253.1). Considering *M. affinis*, from the 7 available COI sequences on NCBI, only one (KJ879298.1) was able to be aligned with the corresponding sequences of the other 4 species. Even so, the first 38 nucleotides were absent in this sequence retrieved from NCBI (Fig. S1, supplementary material). BioEdit v.7.2.5 software (Ibis Biosciences, Canada) was used for sequence alignment, to identify differences among species and to determine appropriate regions for primer design. Three sets of primers were designed, aiming at developing a new method of real-time PCR coupled to HRM analysis (a. Crust2-F/ b. Crust2-R) and sequencing of the different crustacean species (a. Crust2-F/ b. Crust2-R, e. Crust1-F/ b. Crust2-R and c. Crust3-F/ d. Crust3-R) (Table 1). The nucleotide sequences were submitted to a basic local alignment search tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi), which identifies regions of local similarity among homologue sequences of different species and calculates the statistical significance of the matches. Primer specificity was assessed using the Primer-BLAST tool that enables revealing homologies in relation to all sequences available in Genbank database. Primer properties and the absence of hairpins and self-hybridisation were assessed using the software OligoCalc (http://www.basic.northwestern.edu./biotools/oligocalc.html). Primers were synthesised by STABVIDA (Lisbon, Portugal).

Table 1. Oligonucleotide primers for qualitative and real-time PCR.

Primer		Sequence (5' →3')	Amplicon (bp)	
			HRM	Sequencing
Crust2-F	a	ACAGGAATAACTATAGACCG	106[a,b]	106[a,b]
Crust2-R	b	GTATAGTAATAGCTCCTGCTA		
Crust3-F	c	TGGAGGATTTGGTAATTGAC		316[c, d]
Crust3-R	d	CGGTCTATAGTTATTCCTGT		
Crust1-F	e	GCTTGAGCTGGAATAGTAGG		556[e, b]

DNA Extraction

DNA from all tissues and food samples was extracted using Nucleospin® Food kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions with minor alterations. Two different protocols were tested using 200 mg of each sample, one of them by adding 2 µL of RNAse (2 mg/mL) for 5 min at room temperature immediately after the lysis step and the other one following the standard Nucleospin® Food kit protocol. The extractions were performed in duplicate for each sample. All the extracts were kept at -20 °C until further analysis.

DNA quality

Yield and purity of extracts were assessed by UV spectrophotometric DNA quantification on a Synergy HT multi-mode microplate reader (BioTek Instruments, Inc., Vermont, USA), using a Take3 micro-volume plate accessory. DNA content was determined using the nucleic acid quantification protocol with sample type defined for double-strand DNA in the Gen5 data analysis software version 2.01 (BioTek Instruments, Inc., Vermont, USA). The quality of extracted DNA was also analysed by electrophoresis in a 1.0% agarose gel containing 1× Gel Red (Biotium, Hayward, CA, USA) for staining and carried out in 1× STGB (GRISP, Porto, Portugal) for 20 min at 200 V. The agarose gel was visualised under a UV light tray Gel Doc™ EZ System (Bio-Rad Laboratories, Hercules, CA, USA) and a digital image was obtained with Image Lab software version 5.1 (Bio-Rad Laboratories, Hercules, CA, USA).

PCR amplification

For the amplification of shrimp species and frozen/processed seafood products, as well as cross-reactivity testing with other animal and plant species relevant in food, PCR amplifications were carried out in duplicate with 25 µL of total reaction volume containing 2 µL of DNA extract (100 ng), 67 mM of Tris-HCl (pH 8.8), 16 mM of $(NH_4)_2SO_4$, 0.01% of Tween 20, 200 µM of each dNTP, 1.0 U of SuperHot Taq DNA Polymerase (Genaxxon Bioscience, Ulm, Germany), 3.0 mM of $MgCl_2$ and 280 nM of each primer (Crust2-F/Crust2-R) (Table 1). The amplifications were performed in a MJ Mini™ Gradient Thermal Cycler (Bio-Rad Laboratories, Hercules, CA, USA) using different programs of temperature but with similar component proportions. For the set of primers producing a fragment of 556 bp (Crust1-F/Crust2-R) the following temperature program was used: initial denaturation at 95 ºC for 5 min, 40 cycles at 95 ºC for 45 s, 55 ºC for 45 s, 72 ºC for 75 s and a final extension at 72 ºC for 5 min. Considering Crust3-F/Crust3-R primer set

(316 bp) an initial denaturation at 95 ºC for 5 min was performed followed by, 40 cycles at 95 ºC for 30 s, 55 ºC for 45 s, 72 ºC for 1 min and a final extension at 72 ºC for 5 min.

The amplified fragments were analysed by electrophoresis in a 1.5% agarose gel containing 1× Gel Red (Biotium, CA, USA) for staining and carried out in 1× SGTB buffer (GRISP, Porto, Portugal) for about 20-25 min at 200 V. The agarose gel was visualised under a UV light tray Gel Doc™ EZ System (Bio-Rad Laboratories, Hercules, CA, USA) and a digital image was obtained with Image Lab software version 5.1 (BioRad Laboratories, Hercules, CA, USA).

Sequencing of PCR products

PCR products of *L. vannamei, P. monodon, F. indicus, M. affinis* and *M. kerathurus* species were purified with GRS PCR & Gel Band Purification Kit (GRISP, Porto, Portugal) to remove any possible interfering components. The purified products were sent to a specialised research facility for sequencing (GATC Biotech, Constance, Germany). Each target fragment was sequenced twice, performing the direct sequencing of both strands in opposite directions, which allowed the production of four complementary sequences of high quality. The sequencing data were analysed using the available software BioEdit v7.2.5 (Ibis Biosciences, Carlsbad, CA, USA) and FinchTV (Geospiza, Seattle, WA, USA).

Real-time PCR and HRM analysis

Real-time PCR assays were carried out in 20 µL of total reaction volume containing 2 µL of DNA (10 ng), 1× of SsoFast™ Evagreen® Supermix (Bio-Rad Laboratories, Hercules, CA, USA) and 300 nM of each primer (Crust2-F/Crust2-R) (Table 1). The real-time PCR runs were performed on a fluorometric thermal cycler CFX96 real-time PCR detection system (Bio-Rad Laboratories, Hercules, CA, USA) using the following conditions: 95 ºC for 5 min; 45 cycles at 95 ºC for 10 s, 55 ºC for 15 s and 72 ºC for 20 s, with collection of fluorescence signal at the end of each cycle. Data were processed using the software Bio-Rad CFX Manager 3.1 (Bio-Rad Laboratories, Hercules, CA, USA).

For HRM analysis, PCR products were denatured at 95 ºC for 1 min and then annealed at 60 ºC for 3 min, in order to allow the correct annealing of the DNA duplexes. These two steps were followed by melting curve ranging from 60 ºC up to 90 ºC with temperature increments of 0.2 ºC every 10 s. The fluorescence data were acquired at the end of each melting temperature and further processed using the Precision Melt Analysis Software 1.2 (Bio-Rad Laboratories, Hercules, CA, USA) to generate melting curves as a function of temperature and difference curves for easier visual identification of clusters. Melting curve shape sensitivity determines the stringency used to classify melting curves into different

clusters, while temperature of melting (Tm) difference threshold is a parameter that determines the lowest amount of Tm difference between samples. In both cases, cluster detection parameters were set to high sensitivity and threshold yields aiming at providing more heterozygote clusters. Therefore, melting curve shape sensitivity was adjusted as a default value of 50% and Tm difference threshold was set as 0.20. DNA extracts of shrimp species were analysed in replicates (*n*=3) in two independent assays.

Results and discussion

Sequencing data

To obtain reliable species identification and to further justify HRM results, all the reference specimens were sequenced. For that purpose, three sets of primers were designed for direct sequencing analysis. The primer pair Crust1-F/Crust2-R was proposed to amplify relatively large fragments (556 bp) containing the target sequence for HRM analysis (Table 1, Fig. S1, supplementary material). However, those primers only produced amplicons for *P. monodon* and *L. vannamei* species (Fig. S2A and S2B, supplementary material). As a result, other primers (Crust3-F/Crust3-R) (Table 1, Fig. S1, supplementary material) were designed to allow the amplification of smaller fragments (316 bp) from *F. indicus*, *M. affinis* and *M. kerathurus*, which were also sequenced (Fig. S2C, S2D and S2E, supplementary material). The obtained sequences from each specimen were aligned with the respective NCBI retrieved sequences to verify their identity (Fig. S2, supplementary material). Although some nucleotide differences between sequences could be noticed, the BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) results of the query region showed an identity above 99% for all the five species, thus confirming the authenticity of all reference crustaceans.

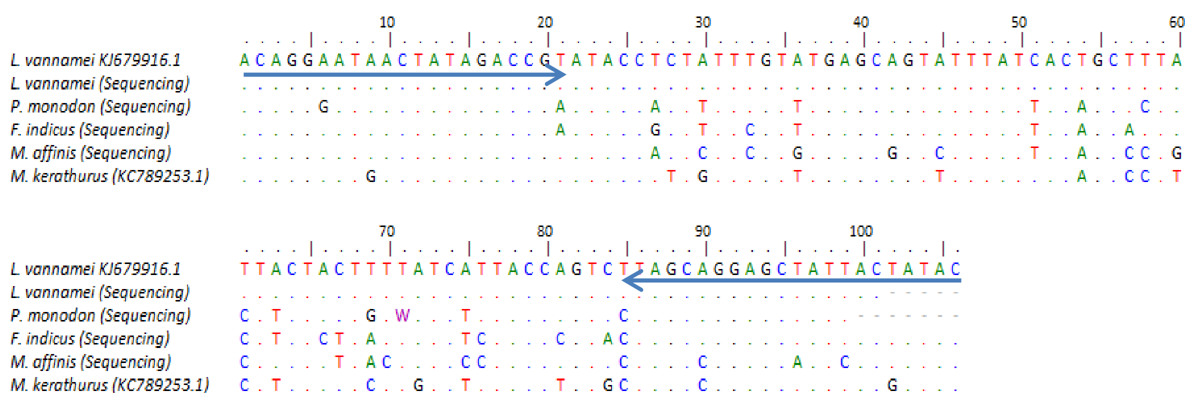


Fig. 1. Mini-barcode sequence alignment (109 bp) showing primer regions (Crust2-F/Crust2-R) and nucleotide differences among *L. vannamei*, *P. monodon*, *M. affinis*, *F. indicus* and *M. kerathurus*.

To complement the sequencing data of the 316 bp fragment (Fig. S2C and S2D, supplementary material), in the cases of *M. affinis* and *F. indicus*, the target HRM region (106 bp) was further sequenced with primers Crust2-F/Crust2-R. Figure 1 represents the alignment of sequencing data of the target HRM *locus* (mini-barcode), showing several nucleotide differences among the five penaeid species and the primer region.

Method optimisation

A real-time PCR assay using EvaGreen dye coupled to HRM analysis was developed and optimised targeting a COI region as a mini-barcode to differentiate five genetically related shrimp species. Different temperature programs were tested in order to achieve the best performance concerning the kinetics of the amplification curves, cycle threshold (Ct) values and the shape of the melting curve shapes. One of the program changes that most positively influenced the performance of the real-time PCR amplifications was splitting the annealing/extension step, commonly performed in one single step in this kind of assays, in two distinct temperature steps.

Besides optimising the program of temperatures, the effect of using RNAse during DNA extraction was also evaluated by real-time PCR with HRM analysis. Figure 2 shows the differences between the two DNA extraction protocols, with (Figs. 2A, 2B, 2C) and without the addition of RNAse (Fig. 2D, 2E, 2F), regarding real-time PCR amplification curves (Figs. 2A, 2D), conventional melting (Figs. 2B, 2E) and HRM analysis (Figs. 2C, 2F), respectively. In both extraction protocols, Ct values ranged from 19 to 32 cycles. It is possible to notice that Ct values for *P. monodon* (green)*, F. indicus* (blue)*, M. affinis* (violet) species were much closer to each other in the extraction with RNAse (Fig. 2A) than without it (Fig. 2D), suggesting that amplification kinetics were similar. The use of RNAse during extraction also allowed the enhancement of the melting profile of *M. kerathurus* species because without RNAse two melting peaks could be observed (Fig. 1B) and with its use a single melting peak was obtained (Fig. 2E). Considering the results of HRM analysis (Fig. 2C and 2F), in both cases with or without the use of RNAse, the five shrimp species were discriminated into distinct clusters with high levels of confidence (>99%) (Table 2), which highlight the robustness of the method.

Focusing on the most favourable conditions with the use of RNAse (Fig. 2A, 2B and 2C), the results of the conventional melting analysis of COI amplicons presented two main groups of melt peaks. In spite of the clear identification of a melt peak at 73.6 ºC, enabling the differentiation of *L. vannamei*, a multitude of melt curve profiles could be observed with very close melting temperatures (75-76 ºC) for the other four shrimp species. Therefore, the proximity of melting temperatures of *P. monodon, F. indicus, M. affinis* and

M. kerathurus disabled their differentiation based on conventional melting curve analysis. In opposition, HRM analysis allowed COI amplicons of the five species to be classified and included in five distinct clusters, thus enabling their correct differentiation with high level of confidence (>99%) (Table 2). According to conventional melting analysis, *L. vannamei* showed a distinct melt peak from the other four species. In HRM analysis, PCR products are classified in different clusters on the basis of similar size and composition and a simple variation in one nucleotide can be enough to differentiate two identical sequences (Druml & Cichna-Markl, 2014; McGlauflin et al., 2010). Based on this principle, the alienation of *L. vannamei* species can also be explained by the highest number of differences in nucleotide composition when aligned with the other four different species (Fig. 1). Its earlier amplification can be attributed to the higher affinity of the primers (Crust2-F/Crust2-R) with the target sequence (lowest number of nucleotide differences) (Fig. 1). Still based on the sequencing data of Figure 1, the identified nucleotide mismatches among sequences can justify the high level of species differentiation obtained by HRM analysis.

Within DNA-based methodologies, most of the scientific works rely frequently on PCR-RFLP (restriction fragment length polymorphism) and DNA barcoding followed by sequencing. Khamnamtong, Klinbunga and Menasveta (2005), Hisar, Aksakal, Hisar, Yanik and Mol (2008) and Pascoal et al. (2008) were able to identify shrimp species by means of PCR-RFLP. Khamnamtong et al. (2005) targeted a 16S ribosomal DNA region for the differentiation of *P. monodon*, *Penaeus semisulcatus*, *Fenneropenaeus merguiensis* and *L. vannamei*, which represent important shrimp species in Thailand, often prone to adulteration issues in that country. Hisar et al. (2008) applied a similar approach targeting cytochrome b gene that allowed the discrimination of *P. semisulcatus, M. kerathurus* (two high valued shrimp species in Turkey), *Penaeus longirostris* and *Metapenaeus monoceros*. Pascoal et al. (2008) performed the most complete PCR-RFLP approach since it allowed the differentiation of 19 penaeid shrimp species of food interest. On the other hand, Bremer, Ditty, Turner and Saxton (2010) performed a multiplex PCR assay followed by DNA sequencing to identify *Farfantepenaeus aztecus*, *Farfantepenaeus duorarum*, *Farfantepenaeus brasiliensis* and *Litopenaeus setiferus* species, all native to the Gulf of Mexico, while Pascoal et al. (2011) targeted the 16S rRNA mtDNA for the PCR amplification of *P. monodon*, *L. vannamei* and *F. indicus*, being three of the most commercially relevant species. Although some methods have already been reported for shrimp species identification, they relied on post-PCR analysis by RFLP or sequencing. In the present work, a HRM method was proposed for the first time to differentiate shrimp species, without requiring any further analysis.

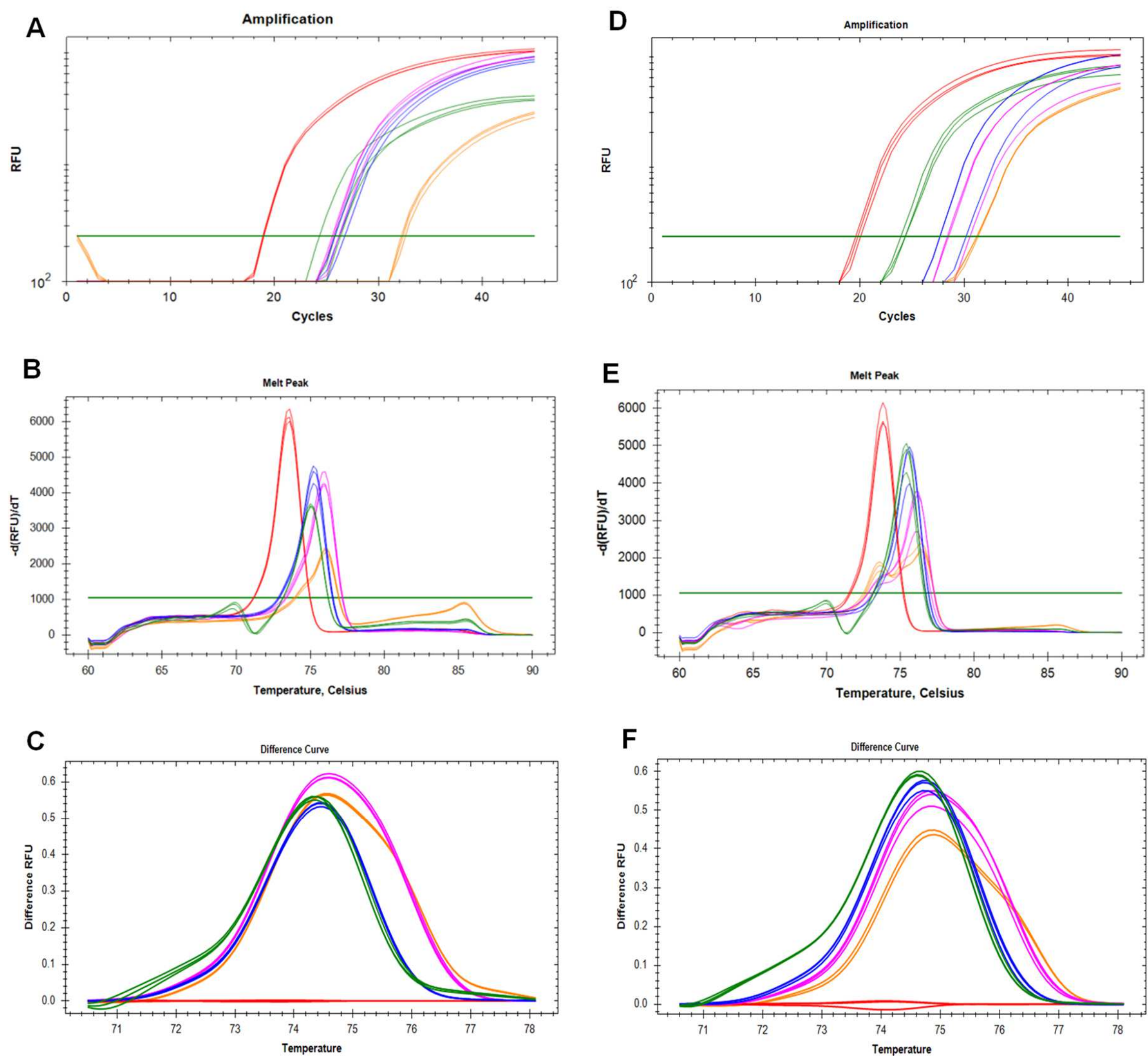


Fig. 2. Amplification curves (A, D), melting curves (B, E) and difference curves (C, F) obtained by real-time PCR with EvaGreen dye, conventional melting curve analysis and HRM analysis targeting the COI gene of Penaeidae species. Assays performed with DNA extracts using RNAse (A, B, C) and without RNAse (D, E, F). Legend: Red, *L. vannamei*; 2, Green, *P. monodon*; Violet, *M. affinis*; Blue, *F. indicus;* Orange, *M. kerathurus.*

Regarding protein-based techniques, some proteomic approaches have been proposed in the past years for shrimp/prawn identification/differentiation (Ortea et al., 2012). Ortea, Cañas, Calo-Mata, Barros-Velázquez and Gallardo (2010) used native IEF of water-soluble sarcoplasmic proteins to differentiate 14 shrimp species of the Decapoda order, but their identification was only achieved with tandem mass spectrometry (MS/MS). The same group of researchers (Ortea, Cañas, Calo-Mata, Barros-Velázquez, & Gallardo, 2009a; Ortea, Cañas, & Gallardo, 2009b) found interspecific variability within arginine kinase sequences of different shrimp species of Decapoda order. Using matrix-assisted

laser desorption/ionisation time-of-flight (MALDI-TOF) and liquid chromatography with electrospray ionisation-ion trap (LC-ESI-IT) MS/MS with target on arginine kinase, the authors were able to perform a selective differentiation between the superfamilies Penaeoidea and Pandaloidea and between the families Penaeidae, Solenoceridae, and Pandalidae, therefore allowing the unequivocal identification of several shrimp species (Ortea et al., 2009b). Pascoal et al. (2012) identified and characterised new *Pandalus borealis*-specific peptides that could be useful as potential markers of this species using the referred proteomic approaches. In spite of the great potential of MS platforms for unequivocal single or multiple species identification, the high cost associated with the equipment and the time of sample analysis are major drawbacks for their application in routine analysis.

Assay specificity

All DNA extracts were tested for their amplifiability with universal primers targeting an eukaryotic region (EG-F/EG-R), as described by Villa, Costa, Oliveira and Mafra (2017). All shrimp samples tested positively with the universal primers, producing the expected fragment of 109 bp (Table 2), which confirmed the absence of any false negative result that could occur due to PCR inhibition or to DNA extracts of poor quality. The five penaeid shrimp species (*L. vannamei, P. monodon, F. indicus, M. affinis and M. kerathurus*) were successfully amplified by qualitative PCR targeting COI mini-barcode (106 bp) (Table 2). No cross-reactivity was observed within non-target animal and plant species (Table 2). Nonetheless, the designed primers were capable of amplifying some crustaceans out of the Penaeidae family, namely some shrimp species from Solenoceridae family (*P. muelleri, S. melantho*, *S crassicornis* and *H. triarthrus*) and *N. norvegicus* (Table 2). In view of that, these non-target crustaceans were further analysed by real-time PCR coupled to HRM along with the target penaeid shrimps (Fig. 3). The results showed a complete discrimination of all species, which was accomplished with a percentage of confidence ranging from 97.8 to 99.6 % (Fig. 3B). Therefore, the specificity of the HRM approach for the five species of Penaeidae family was further demonstrated, being the method also capable of increasing the number of detectable crustacean species.

Analysis of food samples

To demonstrate the applicability of the proposed HRM method, a set of 33 shrimp-containing processed foods was analysed, being the resumed results presented in Table 3. It is possible to notice that only samples containing isolated shrimps or shrimp kernels

Table 2. Results of qualitative PCR and real-time PCR coupled with HRM analysis with universal and specific primers with target and non-target species.

Species	Qualitative PCR		Real-time PCR		
	Eukaryotic (EG-F/EG-R)	Crustaceans (Crust2-F/ Crust2-R)	Ct ± SD [a]	HRM cluster	Level of confidence ± SD [b] (%)
Target shrimp species					
Litopenaeus vannamei	+	+	19.3 ± 0.5	Cluster 1	100.0 ± 0.1
Penaeus monodon	+	+	24.7 ± 1.0	Cluster 2	99.8 ± 0.2
Fenneropenaeus indicus	+	+	26.7 ± 0.7	Cluster 3	99.0 ± 1.2
Metapenaeus affinis	+	+	26.1 ± 0.6	Cluster 4	99.2 ± 0.7
Melicertus kerathurus	+	+	31.4 ± 1.3	Cluster 5	99.5 ± 0.6
Non-target crustacean species					
Nephrops norvegicus	+	+	24.8 ± 0.2	Cluster 6	99.5 ± 0.5
Pleoticus muelleri	+	+	30.4 ± 0.2	Cluster 7	99.6 ± 0.0
Solenocera melantho	+	+	24.0 ± 0.0	Cluster 8	98.8 ± 0.0
Solenocera crassicornis	+	+	21.3 ± 0.1	Cluster 9	97.8 ± 0.5
Haliporoides thriarthus	+	+	25.9 ± 0.2	Cluster 10	98.7 ± 0.2
Portunus validus	+	-	NA[c]		
Cancer pagarus	+	-	NA		
Panulirus argus	+	-	NA		
Molluscs					
Lolligo spp.	+	-	NA		
Mytilus spp.	+	-	NA		
Paphia undulata	+	-	NA		
Fish					
Gadus morhua	+	-	NA		
Theragra chalcogramma	+	-	NA		
Pallachius virens	+	-	NA		
Molva molva	+	-	NA		
Merluccius merluccius	+	-	NA		
Merluccius productus	+	-	NA		
Merluccius hubbsi	+	-	NA		
Merluccius capensis	+	-	NA		
Salmo salar	+	-	NA		
Sparus aurata	+	-	NA		
Raja spp.	+	-	NA		
Solea solea	+	-	NA		
Sardina pilchardus	+	-	NA		
Thunus albacares	+	-	NA		
Other animals *(n=9* species*)*	+	-	NA		
Plants *(n=24* species*)*	+	-	NA		

[a] Mean cycle threshold values (Ct) ± standard deviation (SD) of *n*=6 replicates from two independent real-time PCR runs; [b] Mean level of confidence values ± standard deviation (SD); [c] NA – not applied.

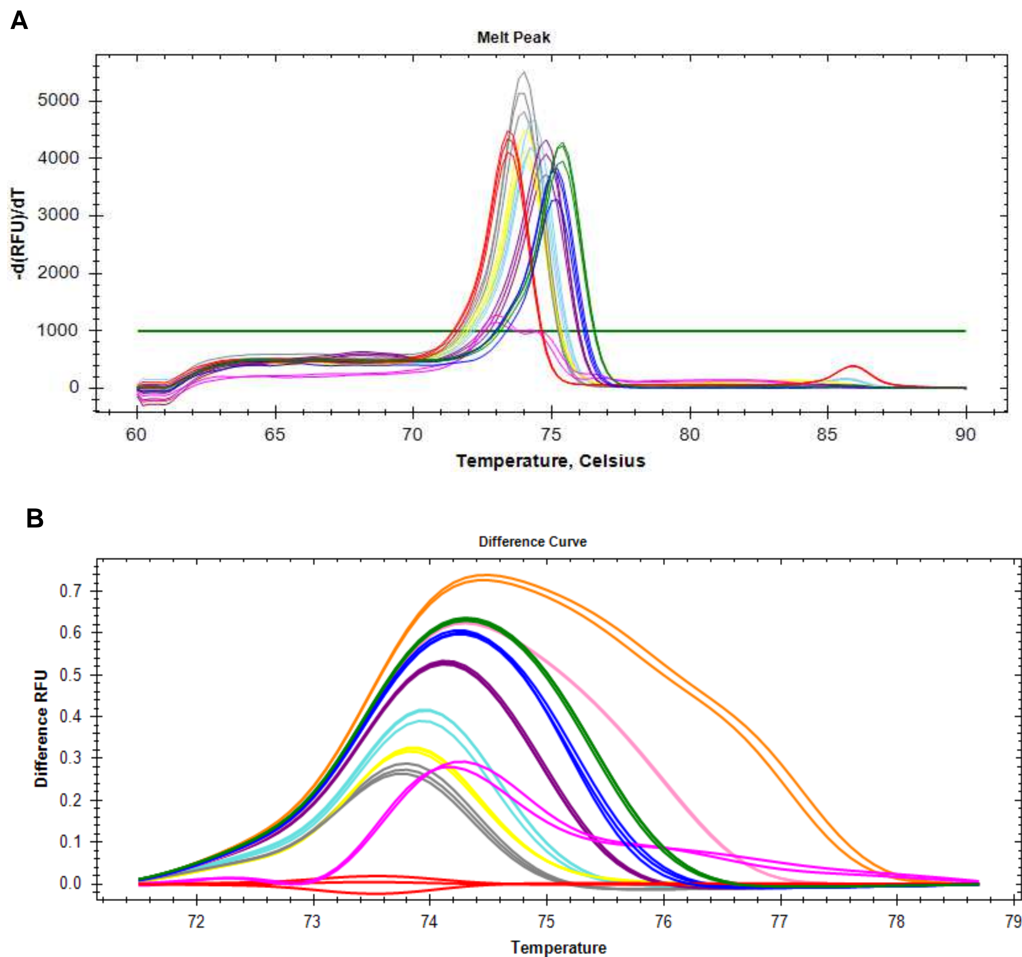


Fig. 3. Melting curves (A) and difference curves (B) obtained by real-time PCR with EvaGreen dye, conventional melting curve analysis and HRM analysis targeting the COI gene of crustacean species. Legend: Red, *L. vannamei*; 2, Green, *P. monodon*; Violet, *M. affinis*; Blue, *F. indicus;* Orange, *M. kerathurus;* Pink, *P. muelleri*; Purple, *S. melantho*; Grey, *S. crassicornis*; Yellow, *H. triarthrus*; Light blue, *N. norvegicus*.

display the scientific name of the species on their labels (samples #5 and #20-27). This fact may be explained by the compliance with the Regulation (EC) No. 1379/2013, which imposes the indication of the scientific name of the species for crustaceans with or without shell in several types of forms. However, the same regulation is not specific in the case of processed/complex foods, thus providing some space for the food industry to omit the scientific designation of shrimp species, as occurred for the majority of processed and shrimp-derived foodstuffs that only declare “shrimp”, “crustaceans” or “may contain crustaceans” in their labels (Table 3).

Concerning the qualitative PCR results, all the samples, except one (#28), were positive with universal primers, confirming the quality of the DNA extracts for PCR amplification. However, several samples were negative with the COI primers (Crust2-

F/Crust2-R), being the majority shrimp patties, seafood powders and fish-containing foodstuffs, the latter declaring “may contain crustaceans” on their label. DNA degradation as a result of food processing, along with the complete exclusion or the presence of minor amounts of shrimp species, as well as the possible substitution by other crustaceans may explain most of the negative PCR results for samples that declare shrimp/crustaceans. Foods labelled with “may contain crustaceans” (samples #6 to #9) must in fact declare such indication due to the common practice of precautionary labelling, considering that they are fish-containing products probably prepared in facilities that are also used to process crustacean-derived foods.

In Figure 4, the results of the conventional melting (A) and HRM (B) analysis of some of the commercial foods along with the five reference shrimp species are shown. Regarding the identified panaeid species in the food samples based on HRM analysis, the results showed that they were mostly distributed within *F. indicus* and *L. vannamei* clusters with levels of confidence that ranged from 97.4 to 99.7 % (mean value of 99.2%) (Table 3). This outcome was expected since these two species are the most commonly available shrimps on the Portuguese market. With the exception of samples #3 (shrimp “açorda”, entire foodstuff) and #10 (shrimp patty) that were discriminated into an unknown cluster out of the five penaeid species, all the positive samples were grouped within a penaeid reference cluster. The inclusion of samples #3 (entire foodstuff) and #10 in a distinct cluster may be explained by the presence of a different crustacean and/or a mixture of species not declared on the label. In sample #5 (a shellfish rice preparation), two portions were isolated, one containing shrimps with shell (labelled as *L. vannamei*) and another one with shrimp kernel (labelled as *P. indicus*). According to HRM analysis, *L. vannamei* species was correctly labelled, but *P. indicus* was not identified, suggesting its substitution by *M. affinis*. In the group of shrimp kernels, unexpectedly, three samples were negative with the proposed COI primers, suggesting the presence of a more distantly related crustacean species. From the 5 positive shrimp kernel samples for the target gene, the HRM results show that only one (sample #21) was assigned according with the labelled species (*P. monodon*). The other four did not comply with the declared species due to identification of *P. monodon* in two samples and *N. norvegicus* in another, suggesting a high level of mislabelling. The non-identified cluster in sample #20 was probably due to a mixture of species.

A crucial issue and novelty of the present work was the true application of the developed methodology to identity penaeid species in processed seafood samples, which was not performed in previous papers. Generally, the results highlight a low variability of the penaeid shrimp species in the products available in the Portuguese market since *F.*

Table 3. Analysed shrimp-containing commercial food samples, with relevant labelled information and results of qualitative PCR and real-time PCR coupled with HRM analysis.

Sample #	Commercial sample	Declared species or relevant information the on label	Qualitative PCR		Real-time PCR		
			Eukaryotic (EG-F/EG-R)	Crustaceans (Crust2-F/ Crust2-R)	Ct ± SD [a]	HRM cluster	Level of confidence ± SD [b] (%)
Typical Portuguese dishes							
1	Shrimp “feijoada”	shrimp, lobster extract	+	+	18.9 ± 0.2	*F. indicus*	99.5 ± 0.2
2	Shrimp “açorda”	shrimp					
	Only shrimp		+	+ [c]	> 38	-	-
	Entire foodstuff		+	-	NA[d]	NA	NA
3	Shrimp “açorda”	shrimp					
	Only shrimp		+	+	23.2 ± 0.2	*F. indicus*	99.7 ± 0.2
	Entire foodstuff		+	+	25.6 ± 0.5	NRC	99.5 ± 0.3
4	Shrimp “açorda”	shrimp	+	+	18.8 ± 0.9	*F. indicus*	99.7 ± 0.1
5	Shellfish rice preparation						
	Only shrimp (with shell)	*L. vannamei*	+	+	22.2 ± 0.6	*L. vannamei*	99.0 ± 0.8
	Only shrimp kernel	*P. indicus*	+	+	22.8 ± 0.2	*M. affinis*	99.6 ± 0.2
	Surimi	-	+	+	27.4 ± 0.5	*L. vannamei*	97.4 ± 1.4
6	Codfish “patanisca”	may contain crustaceans	+	-	NA	NA	NA
7	“Brás” style codfish	may contain crustaceans	+	-	NA	NA	NA
8	Codfish with cream	may contain crustaceans	+	-	NA	NA	NA
9	Codfish cookie	may contain crustaceans	+	-	NA	NA	NA
Patties							
10	Shrimp patties	shrimp	+	+	27.7 ± 1.1	NRC	99.7 ± 0.2
11	Shrimp patties	shrimp	+	-	NA	NA	NA
12	Shrimp patties	shrimp	+	-	NA	NA	NA
13	Shrimp patties	shrimp	+	+ [c]	> 38	-	-
14	Hake and shrimp patties	shrimp	+	+ [c]	> 38	-	-
15	Shrimp patties	shrimp	+	-	NA	NA	NA

[a] Mean cycle threshold values (Ct) ± standard deviation (SD); [b] mean level of confidence values ± standard deviation (SD); [c] faint band of qualitative PCR; [d] NA – not applied; [e] NRC – no reference cluster.

Table 3. (continuation)

Sample #	Commercial sample	Declared species or relevant information the on label	Qualitative PCR		Real-time PCR		
			Eukaryotic (EG-F/EG-R)	Crustaceans (Crust2-F/ Crust2-R)	Ct ± SD [a]	HRM cluster	Level of confidence ± SD [b] (%)
Seafood powders							
16	Seafood cream	shrimp	+	+	20.3 ± 0.4	*L. vannamei*	99.3 ± 0.7
17	Seafood cream	shrimp	+	-	NA	NA	NA
18	Seafood soup	shrimp	+	-	NA	NA	NA
19	Seafood broth	shrimp	+	-	NA	NA	NA
Shrimp kernels							
20	Shrimp kernel	*Penaeus* spp.	+	+	21.4 ± 0.2	NRC [e]	100.0 ± 0.1
21	Shrimp kernel	*P. monodon*	+	+	21.9 ± 0.2	*P. monodon*	100.0 ± 0.1
22	Shrimp kernel	*Parapenaeopsis* spp.	+	+	28.1 ± 0.2	*P. monodon*	99.9 ± 0.1
23	Shrimp kernel	*M. ensis*	+	+	22.4 ± 0.1	*P. monodon*	99.7 ± 0.2
24	Shrimp kernel	*Metapenaeus* spp	+	+	29.8 ± 0.3	*N. norvegicus*	99.6 ± 0.1
25	Shrimp kernel	*M. affinis/S. crassicornis*	+	-	NA	NA	NA
26	Shrimp kernel	*Solenocera* spp./*Metapenaeus* spp.	+	-	NA	NA	NA
27	Shrimp kernel	*Penaeus* spp.	+	-	NA	NA	NA
Other shrimp-containing foods							
28	Shrimp noodles	crustaceans	-	-	-	-	-
29	Fried shrimp	shrimp	+	+	20.0 ± 0.2	*L. vannamei*	99.3 ± 0.3
30	Crab stick (to fry)	crab	+	-	NA	NA	NA
31	Seafood pizza	shrimp	+	+	18.2 ± 0.1	*L. vannamei*	99.0 ± 0.4
32	Seafood “risotto”	shrimp	+	+ [c]	>38	-	-
33	“Paella”	shrimp	+	+ [c]	>38	-	-

[a] Mean cycle threshold values (Ct) ± standard deviation (SD); [b] mean level of confidence values ± standard deviation (SD); [c] faint band of qualitative PCR; [d] NA – not applied; [e] NRC – no reference cluster.

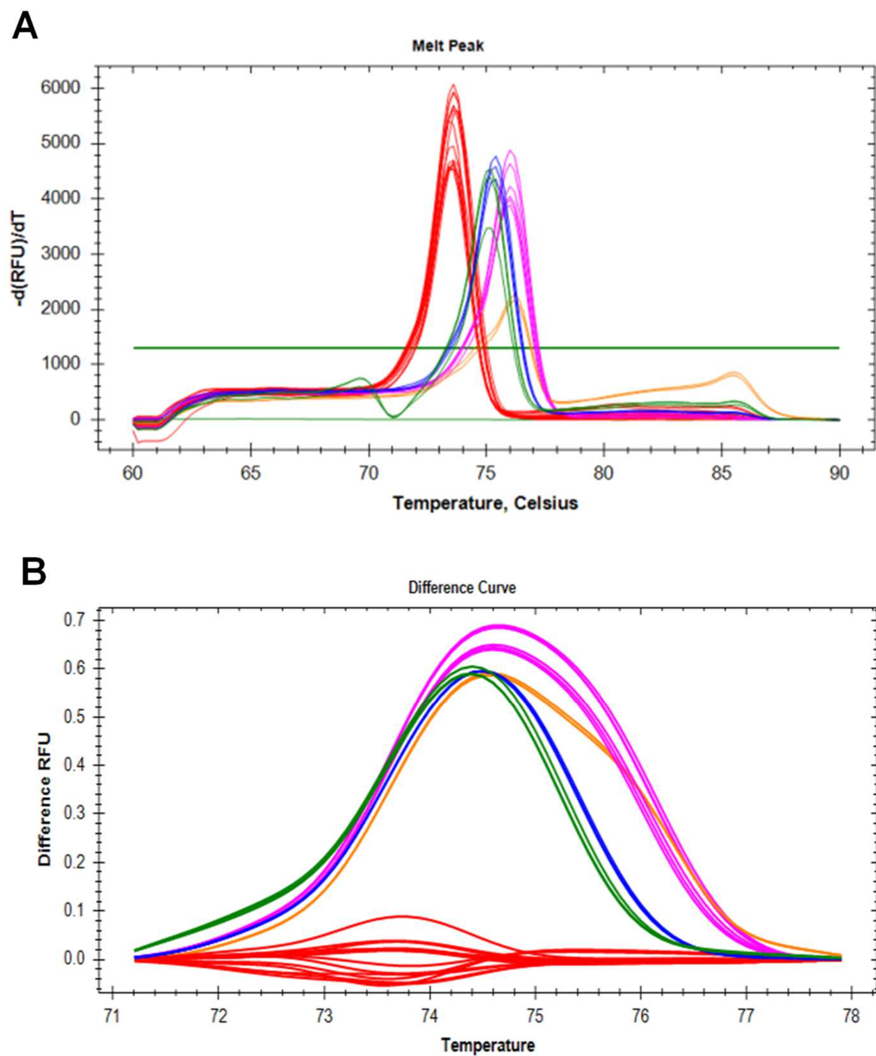


Fig. 4. Melting curves (A) and difference curves (B) obtained by real-time PCR with EvaGreen dye, conventional melting curve analysis and HRM analysis targeting the COI gene of shrimp species and commercial seafood samples. Legend: Red, *L. vannamei,* samples #5 (surimi), #16, #21, #23; Green, *P. monodon*; Violet, *M. affinis*, sample #5 (only shrimp kernel); Blue, *F. indicus;* Orange, *M. kerathurus.*

indicus and *L. vannamei* were the main identified species. When verifying labelling compliance, four samples suggest adulterations owing to complete or partial substitution of shrimp species, highlighting the need to control the authenticity of these products.

Conclusion

In the present work, a novel method based on real-time PCR coupled to HRM analysis is proposed to target a COI mini-barcode as a robust tool for the detection and discrimination of penaeid shrimp species, evidencing clear advantages in terms of speed, cost and simplicity, comparing with the available reported methods. This approach can be

very useful for the food industry and control laboratories, allowing the identification of penaeid shrimp species in commercialised raw and processed crustacean products and protecting fair competing and consumers' from fraudulent practices.

Supplementary Material

Detailed data regarding sequence analysis and respective alignments for primer design targeting the five penaied shrimps (*L. vannamei*, *P. monodon*, *F. indicus*, *M. affinis* and *M. kerathurus*) (Fig. S1) and sequencing results of each shrimp species (Fig. S2) were provided as supplementary material.

Acknowledgements

This work was supported by FCT (Fundação para a Ciência e Tecnologia) through project UID/QUI/50006/2013 – POCI/01/0145/FEDER/007265 with financial support from FCT/MEC through national funds and co-financed by FEDER, under the Partnership Agreement PT2020 and by the project NORTE-01-0145-FEDER-000011. Telmo J. R. Fernandes and Joana Costa are grateful to FCT grants (SFRH/BD/93711/2013 and SFRH/BPD/102404/2014, respectively) financed by POPH-QREN (subsidised by FSE and MCTES). The authors acknowledge the kind supply of shrimp species by Marfresco (Loures, Portugal) and Brasmar Seafood Companies (Trofa, Portugal).

References

Arroyave, J., & Stiassny, M. L. J. (2014). DNA barcoding reveals novel insights into pterygophagy and prey selection in distichodontid fishes (Characiformes: Distichodontidae). *Ecology and Evolution, 4*, 4534-4542.

Bremer, J. R. A., Ditty, J. G., Turner, J. S., & Saxton, B. L., 2010. Molecular species identification of commercially important penaeid shrimp from the Gulf of Mexico using a multiplex haplotype-specific PCR assay. *Biochemical Systematics and Ecology, 38*, 715-721.

Chin, T. C., Adibah, A. B., Hariz, Z. A. D., & Azizah, M. N. S. (2016). Detection of mislabelled seafood products in Malaysia by DNA barcoding: Improving transparency in food market. *Food Control, 64*, 247-256.

Costa, J., Mafra, I., & Oliveira, M. B. P. P. (2012). High resolution melting analysis as a new approach to detect almond DNA encoding for Pru du 5 allergen in foods. *Food Chemistry, 133*, 1062-1069.

Druml, B., & Cichna-Markl, M. (2014). High resolution melting (HRM) analysis of DNA – Its role and potential in food analysis. *Food Chemistry, 158*, 245-254.

Eischeid, A. C., Stadig, S. R., Handy, S. M., Fry, F. S., & Deeds, J. (2016). Optimization and evaluation of a method for the generation of DNA barcodes for the identification of crustaceans. *LWT – Food Science and Technology*, *73*, 357-367.

FAO (2016). Food and Agriculture Organization of the United Nations, Rome, Italy. Available at: http://www.fao.org/ (Last accessed 31 August 2016).

Fitzcharles, E. M. (2012). Rapid discrimination between four Antarctic fish species, genus *Macrourus*, using HRM analysis. *Fisheries Research, 127–128*, 166-170.

Galal-Khallaf, A., Ardura, A., Borrell, Y. J., & Garcia-Vazquez, E. (2016). PCR-based assessment of shellfish traceability and sustainability in international Mediterranean seafood markets. *Food Chemistry, 202*, 302-308.

Haye, P. A., Segovia, N. I., Vera, R., Gallardo, M. Á., & Gallardo-Escárate, C. (2012). Authentication of commercialized crab-meat in Chile using DNA Barcoding. *Food Control, 25*, 239-244.

Hisar, S. A., Aksakal, E., Hisar, O., Yanik, T., & Mol, S. (2008). Discrimination of penaeid shrimps with PCR-RFLP analysis. *Journal of Shellfish Research, 27*, 917-920.

Jaakola, L., Suokas, M., & Häggman, H. (2010). Novel approaches based on DNA barcoding and high-resolution melting of amplicons for authenticity analyses of berry species. *Food Chemistry*, *123*, 494-500.

Khamnamtong, B., Klinbunga, S., & Menasveta, P. (2005). Species identification of five penaeid shrimps using PCR-RFLP and SSCP analyses of 16S ribosomal DNA. *Journal of Biochemistry and Molecular Biology*, *38*, 491-499.

Mackay, J., Wright, C., & Bonfiglioli, R. (2008). A new approach to varietal identification in plants by microsatellite high resolution melting analysis: application to the verification of grapevine and olive cultivars. *Plant Methods*, *4*, 8.

Madesis, P., Ganopoulos, I., Sakaridis, I., Argiriou, A., & Tsaftaris, A. (2014). Advances of DNA-based methods for tracing the botanical origin of food products. *Food Research Internacional*, *60*, 163-172.

Mafra, I., Ferreira, I. M. P. L. V. O., & Oliveira, M. B. P. P. (2008). Food authentication by PCR-based methods. *European Food Research and Technology*, *227*, 649-665.

McGlauflin, M. T., Smith, M. J., Wang, J. T., Young, S. F., Chen, N., Lee, Y. C., Pascal, C., Seeb, L. W., Stevens, J., & Seeb, J. E. (2010). High-resolution melting analysis for the discovery of novel single-nucleotide polymorphisms in rainbow and cutthroat trout for species identification. *Transactions of the American Fisheries Society, 139*, 676-684.

Ortea, I., Cañas, B., Calo-Mata, P., Barros-Velázquez, J., & Gallardo, J. M. (2009a). Arginine kinase peptide mass fingerprinting as a proteomic approach for species

identification and taxonomic analysis of commercially relevant shrimp species. *Journal of Agricultural and Food Chemistry, 57*, 5665-5672.

Ortea, I., Cañas, B., Calo-Mata, P., Barros-Velázquez, J., & Gallardo, J. M. (2010). Identification of commercial prawn and shrimp species of food interest by native isoelectric focusing. *Food Chemistry*, *121*, 569-574.

Ortea, I., Cañas, B., & Gallardo, J. M. (2009b). Mass spectrometry characterization of species-specific peptides from arginine kinase for the identification of commercially relevant shrimp species. *Journal of Proteome Research, 8*, 5356-5362.

Ortea, I., Pascoal, A., Cañas, B., Gallardo, J. M., Barros-Velázquez, J., & Calo-Mata, P. (2012). Food authentication of commercially-relevant shrimp and prawn species: From classical methods to Foodomics. *Electrophoresis, 33*, 2201-2211.

Pascoal, A., Barros-Velazquez, J., Cepeda, A., Gallardo, J. M., & Calo-Mata, P. (2008). Survey of the authenticity of prawn and shrimp species in commercial food products by PCR-RFLP analysis of a 16S rRNA/tRNAVal mitochondrial region. *Food Chemistry, 109*, 638-646.

Pascoal, A., Barros-Velázquez, J., Ortea, I., Cepeda, A., Gallardo, J. M., & Calo-Mata, P. (2011). Molecular identification of the black tiger shrimp (*Penaeus monodon*), the white leg shrimp (*Litopenaeus vannamei*) and the Indian white shrimp (*Fenneropenaeus indicus*) by PCR targeted to the 16S rRNA mtDNA. *Food Chemistry, 125*, 1457-1461.

Pascoal, A., Ortea, I., Gallardo, J. M., Cañas, B., Barros-Velázquez, J., & Calo-Mata, P. (2012). Species identification of the Northern shrimp (*Pandalus borealis*) by polymerase chain reaction–restriction fragment length polymorphism and proteomic analysis. *Analytical Biochemistry, 421*, 56-67.

Regulation (EC) No. 1379/2013 of 11 December 2013 on the common organisation of the markets in fishery and aquaculture products, amending Council Regulations (EC) No 1184/2006 and (EC) No 1224/2009 and repealing Council Regulation (EC) No 104/2000. *Official Journal of the European Union*, *L 354*, 1-21.

Rittenschober, D., Stadlmayr, B., Nowak, V., Du, J., & Charrondiere, U. R. (2016). Report on the development of the FAO/INFOODS user database for fish and shellfish (uFiSh) – Challenges and possible solutions. *Food Chemistry, 193*, 112-120.

Sakaridis, I., Ganopoulos, I., Argiriou, A., & Tsaftaris, A. (2013). High resolution melting analysis for quantitative detection of bovine milk in pure water buffalo mozzarella and other buffalo dairy products. *International Dairy Journal*, *28*, 32-35.

Shen, Y., Kang, J., Chen, W., & He, S. (2016). DNA barcoding for the identification of common economic aquatic products in Central China and its application for the supervision of the market trade. *Food Control, 61*, 79-91.

Tagliavia, M., Nicosia, A., Salamone, M., Biondo, G., Bennici, C. D., Mazzola, S., & Cuttitta, A. (2016). Development of a fast DNA extraction method for sea food and marine species identification. *Food Chemistry, 203*, 375-378.

Villa C., Costa, J., Oliveira, M.B.P.P., Mafra, I. (2017). Novel quantitative real-time PCR approach to determine safflower (*Carthamus tinctorius*) adulteration in saffron (*Crocus sativus*). *Food Chemistry, 229, 680-687.*

Fisheries Research 197 (2018) 50–59

Contents lists available at ScienceDirect

Fisheries Research

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

COI barcode-HRM as a novel approach for the discrimination of hake species

Telmo J.R. Fernandes, Joana Costa, M. Beatriz P.P. Oliveira, Isabel Mafra*

REQUIMTE-LAQV, Faculdade de Farmácia, Universidade do Porto, Rua de Jorge Viterbo Ferreira, 228, 4050-313 Porto, Portugal

*Corresponding author: Tel: +351 220428640. Fax: +351 226093390. E-mail: isabel.mafra@ff.up.pt

Abstract

Hake species of *Merluccius* genus represent an important group of fish commonly sold all over the world. Therefore, they are highly prone to be adulterated, particularly the substitution of *M. merluccius* by other species with lower market value. The present work intended the development of a highly sensitive methodology for the rapid detection and differentiation of hake species based on mini-barcoding of cytochrome c oxidase subunit I (COI) gene combined with high resolution melting (HRM) analysis. The method allowed the full discrimination of *M. merluccius*, *M. productus*, *M. hubbsi*, *M. capensis* and *M. paradoxus* with high levels of confidence. Real-time PCR assay targeting COI mini-barcode provided a high sensitive tool to detect hake species down to 0.2-20 pg of DNA with adequate performance parameters. The application of the COI-HRM approach to 45 fish-containing foods showed that two samples did not comply with the declared species, suggesting mislabelling or species substitution. These findings highlight the need of controlling processed fish-containing foods and the feasibility of the proposed tool for their authentication at trace levels.

Keywords: High resolution melting; *Merluccius*; DNA barcoding; species differentiation; authenticity.

Introduction

The fish sector is significantly growing over the years due to the global increase of fish consumption worldwide. Merluccidae family encompasses some of the most commercially relevant fish species that are heavily captured nowadays. *Merluccius merluccius* (European hake) is the most appreciated species and its market value is often higher than other hakes within the *Merluccius* genus. Nonetheless, there are important and commonly sold Merluccidae species from other geographical locations, such as *M. productus* (North Pacific hake), *M. hubbsi* (Argentine hake), *M. capensis* (whiting) and *M. paradoxus* (deepwater hake) (Hubalkova et al., 2009; Machado-Schiaffino et al., 2008; Muñoz-Colmenero et al., 2015; Sánchez et al., 2009). Similarly to other fish families, the identification based on visual analysis is very difficult or even impossible due to the phenotypic resemblances of *Merluccius* spp. or when the morphological characteristics are lost (e.g. fillets, tails, loins, compressed blocks, fish sticks), making them prone to mislabelling and fraudulent practices. Species substitution has been widely reported both in the media and in the scientific community, being usually through the deliberate replacement of high-valued species by lower cost ones without the respective labelling correction (Abdullah and Rehbein, 2016; Castigliego et al., 2015; Di Pinto et al., 2013; Hubalkova et al., 2008). According to the European Union (EU) legislation, unprocessed and some processed seafood products (e.g. salted, smoked products) must be labelled with the commercial and the scientific name of the species, while in other products, such as, canned, composite products and breaded products, the inclusion of this type of information is voluntary (Regulation (EU) No 1379/2013). Additionally, Regulation (EU) No 1169/2011 demands the obligatory labelling of a list of allergenic foods, from which fish is included, that should be highlighted from the listed ingredients, regardless of their quantity. Therefore, clear and reliable mechanisms for species certification constitute a benefit for consumers and producers, not only for authentication purposes, but also for assessing the presence of potentially allergenic foods that might be inadvertently undeclared (Hubalkova et al., 2009; Sánchez et al., 2009).

DNA-based methods have proved to be suitable alternatives to the classical protein-based methods and became the methods of preference for the analysis of a wide range of seafood products. Species-specific polymerase chain reaction (PCR), multiplex PCR, PCR followed by analysis of restriction fragment length polymorphisms (PCR-RFLP), forensically informative nucleotide sequencing (PCR-FINS) and single-stranded conformational polymorphism (PCR-SSCP), targeting mostly mitochondrial (mt) DNA markers, are some examples of methods developed for the identification and discrimination of fish species, such as flatfish, gadoids, scombroids, salmonids and

percoids (Mafra et al., 2008; Rasmussen and Morissey, 2008; 2009). More recently, real-time PCR methods have been widely used for fish species identification in foods due to the advantages of high specificity and sensitivity, combined with simplicity and rapidity (Armani et al., 2012; Hird et al., 2012; Sánchez et al., 2009; Sánchez et al., 2013; Taboada et al., 2017). Another recent approach applied to fish species discrimination is based on DNA barcoding, for which the cytochrome c oxidase subunit I (COI) gene is the target most commonly used (Armani et al., 2015; Deli Antoni et al., 2015; Di Pinto et al., 2013; Fernandes et al., 2017; Ferrito et al., 2016). DNA barcoding relies on the sequence variation within a short and standardised region of the genome, designated as a "barcode", to enable species identification. In particular, COI fragments of ~650 base pairs have shown to discriminate animal species reliably, including fish (Carvalho et al., 2017; Costa and Carvalho, 2007; Handy et al., 2011; Hebert et al., 2003; Pappalardo and Ferrito, 2015; Pappalardo et al., 2015; Ward et al., 2005). The use of even smaller DNA barcodes (mini-barcodes) has also been used and further recommended when analysing highly processed foods containing degraded DNA (Fernandes et al., 2017; Fields et al., 2015; Little, 2014; Mitchell and Hellberg, 2016; Villa et al., 2016). Besides, mini-barcodes can be combined with high resolution melting (HRM) analysis, which can be used as a cost-effective, specific and high-throughput tool for the discrimination of fragments with small nucleotide differences, therefore avoiding the need of sequencing (Fernandes et al., 2017; Villa et al., 2016; Xanthopoulou et al., 2016). However, only few studies report the approach of mini-barcoding combined with HRM analysis for fish species differentiation (Fernandes et al., 2017; Fitzcharles, 2012).

The main goal of the present work was to develop a highly sensitive methodology for the rapid detection and differentiation of hake species based on mini-barcoding combined with HRM analysis. A COI mini-barcode was *in silico* analysed regarding five *Merluccius* species (*M. merluccius*, *M. products*, *M. paradoxus*, *M. hubbsi* and *M. capensis*), taking into account their genetic variability for primer design and HRM analysis. The method was validated through sequencing and its applicability was further demonstrated by analysing processed fish-containing foods to identify Merluccidae species and to verify labelling compliance.

Materials and Methods

Sampling and preparation

Twenty specimens of five *Merluccius* species were acquired at different local markets and further authenticated by sequencing: European hake – *M. merluccius* (*n*=4); North Pacific hake – *M. productus* (*n*=4); Argentine hake - *M. hubsi* (*n*=4); whiting – *M. capensis*

(*n*=4) and deepwater hake – *M. paradoxus* (*n*=4). A total of 49 specimens, one of each species, comprising other fish, crustaceans and molluscs (*n*=21) (Atlantic cod – *Gadus morhua*; Alaska pollock - *Theragra chalcogramma;* saithe - *Pollachius virens*; Atlantic salmon - *Salmo salar*; gilt-head bream - *Sparus aurata*; ray - *Raja* spp.; common sole - *Solea solea*; European pilchard - *Sardina pilchardus*; yellowfin tuna - *Thunnus albacares*; squid - *Loligo* spp.; mussel – *Mytilus spp.*; undulated surf clam – *Paphia undulata*; crab – *Portunus validus*; edible crab – *Cancer pagurus*; Caribbean spiny lobster – *Panulirus argus*; Norway lobster – *Nephrops norvegicus*; Whiteleg shrimp - *Litopenaeus vannamei*; Giant tiger prawn - *Penaeus monodon*; Indian white prawn - *Fenneropenaeus indicus*; Jinga shrimp - *Metapenaeus affinis* and Striped shrimp, *Melicertus kerathurus*), meat species (*n*=9) (cow, chicken, rabbit, deer, turkey, pork, lamb, goat, ostrich) and plant species (*n*=19) (soybean, oat, rye, mint, wheat, lupine, maize, rice, pumpkin seeds, rapeseed, sunflower, tomato, peach, nectarine, apricot, strawberry, raspberry, honey, cashew), commonly used as food, were tested to assess the assay specificity. Several processed seafood products (fish sticks, hake/fish fillets, breaded hake/fish fillets, hake patties, hake medallion, fish pies, surimi, pre-cooked dishes, among others, *n*=45) were also acquired at local markets and analysed to verify assay applicability.

All the specimens and samples were individually ground and homogenised in a laboratory knife mill Grindomix GM200 (Retsch, Haan, Germany) and stored at -20 ºC until additional analysis. All containers and material used during this procedure were previously treated with DNA decontamination solution (DNA-ExitusPlus™, AppliChem, Darmstadt, Germany).

In silico DNA barcode analysis and primer design

COI sequences of *Merluccius* spp. were obtained from NCBI database with the respective accession numbers (*M. merluccius* - KJ679916.1*; M. productus* - KM528138.1; *M. hubbsi* - KF604889.1; *M. campensis* - KJ879298.1*; M. paradoxus* - KC789253.1). The sequences were aligned with BioEdit v.7.2.5 software (Ibis Biosciences, Canada) and further analysed to check the regions with high sequence homology for primer design, but also including several nucleotide variations to allow amplicon differentiation. Accordingly, two sets of primers were designed, namely MER1COI-F (TCACGGCACACGCCTTCGTAA)/ MER1COI-R (TGTCGGGGGCTCCGATCATTA) to produce a fragment of 102 bp for the development of a real-time PCR assay combined with HRM analysis and MER2COI-F (GCATAGTCGGAACAGCCCTAA)/MER2COI-R (CCCAGAATTGATGAAACGCC) targeting a region of 400 bp for sequencing purposes. To check the primer properties, namely GC content, estimated annealing temperature, the

absence of hairpins and self-hybridisation, as well as to evaluate their *in silico* specificity towards the available database nucleotide sequences (GenBank), the OligoCalc (http://www.basic.northwestern.edu./biotools/oligocalc.html) and the Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) softwares, respectively, were used. Primers were synthesised by STABVIDA (Lisbon, Portugal).

DNA Extraction and its quality assessment

For DNA extraction, the Nucleospin® Food kit (Macherey-Nagel, Düren, Germany) was used according to the manufacturer's instructions with minor alterations. Briefly, 200 mg of each sample were used and 2 µL of RNAse (2 mg/mL) were added for 5 min at room temperature immediately after the lysis step (1 h incubation at 65 ºC). The extractions were performed in duplicate for each sample and all the extracts were kept at -20 ºC until further analysis.

For DNA yield and purity determinations, the UV spectrophotometric DNA quantification on a Synergy HT multi-mode microplate reader (BioTek Instruments, Inc., Vermont, USA) and a Take3 micro-volume plate accessory were used, following the defined nucleic acid quantification protocol (double-strand DNA) in the Gen5 data analysis software version 2.01 (BioTek Instruments, Inc., Vermont, USA). Electrophoresis in a 1.0% agarose gel containing Gel Red 1x (Biotium, Hayward, CA, USA) for staining and carried out in STGB 1x (GRISP, Porto, Portugal) for 25 min at 200 V was performed to check DNA integrity of the extracts. The agarose gel was visualised under a UV light tray Gel Doc™ EZ System (Bio-Rad Laboratories, Hercules, CA, USA) and a digital image was obtained with Image Lab software version 5.1 (Bio-Rad Laboratories, Hercules, CA, USA).

PCR amplification

Reactional mixtures of 25 µL of total volume containing 2 µL of DNA extract (100 ng), 67 mM of Tris-HCl (pH 8.8), 16 mM of $(NH_4)_2SO_4$, 0.01% of Tween 20, 200 µM of each dNTP, 1.0 U of SuperHot Taq DNA Polymerase (Genaxxon Bioscience, Ulm, Germany), 3.0 mM of $MgCl_2$ and 200 nM of each primer (MER1COI-F/MER1COI-R or MER2COI-F/MER2COI-R). The amplifications were performed in a MJ Mini™ Gradient Thermal Cycler (Bio-Rad Laboratories, Hercules, CA, USA) using the following temperature program: initial denaturation at 95 ºC for 5 min; 40 cycles at 95 ºC for 30 s, 60 ºC for 30 s (or 45 s with primers MER2COI-F/MER2COI-R) and 72 ºC for 30 s (or 60 s with primers MER2COI-F/MER2COI-R); and a final extension at 72 ºC for 5 min.

PCR amplicons were run in a 1.5% agarose gel containing 1× Gel Red (Biotium, CA, USA) for staining and carried out in 1× SGTB buffer (GRISP, Porto, Portugal) for about 20-25 min at 200 V. A digital image was obtained with Image Lab software version 5.1 (BioRad Laboratories, Hercules, CA, USA) using a UV light tray Gel Doc™ EZ System (Bio-Rad Laboratories, Hercules, CA, USA).

Real-time PCR and HRM analysis

Real-time PCR amplifications were carried out in 20 µL of total reactional mixture contained 2 µL of DNA (10 ng), 1× of SsoFast™ Evagreen® Supermix (Bio-Rad Laboratories, Hercules, CA, USA) and 300 nM of each primer (MER1COI-F/MER1COI-R). The runs were performed on a fluorometric thermal cycler CFX96 Real-time PCR Detection System (Bio-Rad Laboratories, Hercules, CA, USA) according to the program: 95 ºC for 5 min; 45 cycles at 95 ºC for 10 s, 60 ºC for 15 s and 72 ºC for 20 s, with collection of fluorescence signal at the end of each cycle. Data were processed using the software Bio-Rad CFX Manager 3.1 (Bio-Rad Laboratories, Hercules, CA, USA).

Fig. 1. Alignment of sequencing data, highlighting the region used in HRM analysis (shaded blue).

For HRM analysis, PCR products were denatured at 95 ºC for 1 min and then annealed at 65 ºC for 3 min, in order to allow the correct annealing of the DNA duplexes. Afterwards, there was an increase of 0.2 ºC every 10 s, from 65 ºC up to 95 ºC. The fluorescence data were acquired at the end of each melting cycle and were further processed using the Precision Melt Analysis Software 1.2 (Bio-Rad Laboratories, Hercules, CA, USA). DNA extracts were analysed in replicates (*n*=3), at least in two independent assays.

Sequencing of PCR products

PCR products of *M. merluccius, M. productus, M. hubsi, M. capensis,* and *M. paradoxus* samples were purified with GRS PCR & Gel Band Purification Kit (GRISP, Porto, Portugal) and sent to a specialised research facility (GATC Biotech, Constance, Germany) for sequencing. Each target fragment was sequenced twice, performing the direct sequencing of both strands in opposite directions, which allowed the production of four complementary sequences of high quality. The sequencing data were aligned using the available software BioEdit v7.2.5 (Ibis Biosciences, Carlsbad, CA, USA) and the respective electropherograms analysed with FinchTV (Geospiza, Seattle, WA, USA).

Results

Specificity

After satisfactory *in silico* analysis, the assay specificity with the designed new primers MER1COI-F/MER1COI-R was further tested for any eventual amplification with extracts of several (*n*=49) food-related plant and animal species, listed in the previous section. No cross-reactivity was obtained with the untargeted species, but the expected COI fragment of 102 bp was observed in all tested *Merluccius* species. (*M. merluccius, M. productus*, *M. hubbsi, M. capensis* and *M. paradoxus*) (Figure S1, see Supplementary Material). The absence of any false negative result was assured based on the positive PCR amplification of a fragment targeting an universal eukaryotic gene (Costa et al., 2013).

Sequencing data

The second set of designed new primers (MER2COI-F/MER2COI-R) produced amplicons of 400 bp, encompassing the target COI mini-barcode (102 bp) of *Merluccius* spp. to enable reliable sequencing of the query region. Sequencing results (Figure 1) were in good agreement with the available COI sequences retrieved from the NCBI database

since BLAST ID values ranged from 99 to 100% (Table 1) for each species, which confirmed the authenticity of the hake specimens.

Table 1. Summary of PCR results, HRM cluster analysis and sequence comparison for the five *Merluccius* species under study.

Species	Qualitative PCR	HRM cluster	Level of confidence (%) [a]	Sequencing data BLAST ID (%)	NCBI sequences
M. merluccius	+	1	99.1 ± 0.3	100	KP975783.1
M. productus	+	2	99.8 ± 0.2	99	JQ354224.1
M. hubbsi	+	3	99.5 ± 0.3	99	GU702480.1
M. capensis	+	4	99.5 ± 0.4	100	JF493883.1
M. paradoxus	+	5	99.5 ± 0.5	100	KP975790.1

[a] Mean values of level of confidence ± standard deviation (%) of n=6 replicates from two independent real-time PCR runs.

Real-time PCR and HRM analysis

A real-time PCR assay (Figure 2) using EvaGreen dye was successfully developed and further combined with HRM analysis targeting the COI mini-barcode sequence of 102 bp to specifically detect *Merluccius* spp. and further differentiate the five species under study. The conventional melting analysis of COI amplicons (Figure 2A) shows groups of curves with close melt peaks around 78-79 ºC, without being able to differentiate the five *Merluccius* species under study. The application of HRM analysis allowed the discrimination of the five hake species with levels of confidence above 99% (Figure 2B, 2C, Table 1). The differences among all melting curves presented in the normalized plot (Figure 2B), are evidenced in the temperature-shifted difference curve chart, thus allowing the identification of five distinct clusters (Figure 2C). These data are in good agreement with the sequencing results and alignment within the region of the analysed amplicon HRM. Considering *M. merluccius* NCBI sequence as the target used for primer design, it is possible to verify that, as expected, the obtained sequencing data for this species were completely aligned with the query region. However, the other four species showed all different nucleotide mismatches, thus corroborating with the HRM results (Figure 1). Being *M. merluccius* set as the reference cluster 1, *M. productus* was included in cluster 2, *M hubbsi* in the cluster 3, both with 5 nucleotide differences; *M. capensis* was set as cluster 4 and *M. paradoxus* as cluster 5, both with 4 mismatches (Figure 1, Table 1). It is also relevant to emphasise that HRM analysis was able to discriminate *M. productus* from *M. hubbsi and M. capensis* from *M. paradoxus* since only 2 nucleotide differences are noted within each pair.

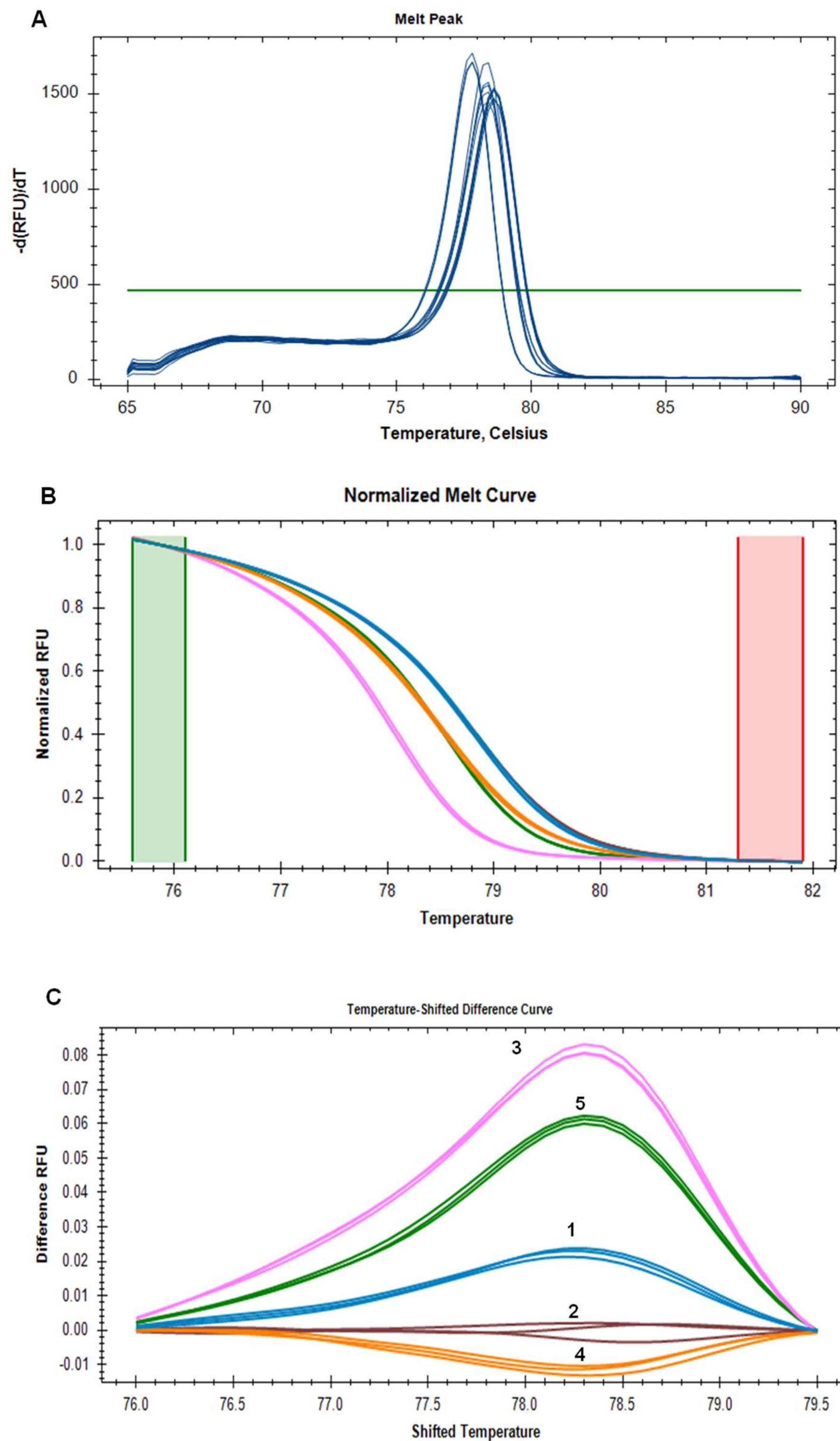


Figure 2. Conventional melting curves (A) and HRM data, namely, normalised melting curves (B) and temperature-shifted difference curves (C) of PCR products targeting the COI region of *Merluccius* spp. Legend: 1, *M. merluccius* cluster; 2, *M. productus* cluster; 3, *M. hubbsi* cluster; 4, *M. capensis* cluster*; 5, M. paradoxus* cluster.

Analytical performance

For method development, the performance criteria established for real-time PCR assays of Bustin et al. (2009) and European network of GMO laboratories (ENGL, 2015) were carefully considered. To assess method sensitivity, 10-fold serially diluted DNA extracts (20 ng to 0.2 pg) from each of the five hake species were amplified by real-time PCR to cover at least 4 orders of magnitude (Figure 3). *M. merluccius* presented the lowest Ct values of the five species, comparing at each dilution level, and, consequently the highest sensitivity, with a limit of detection (LOD) of 0.2 pg since all the tested replicates amplified at this level. The limit of quantification (LOQ) was also established as 0.2 pg as it was within the linear dynamic range of the calibration curve (Figures 3A, 3B). *M. hubbsi* (Figures 3E, 3F), *M. capensis* (Figures 3G, 3H) and *M. paradoxus* (Figures 3I, 3J) displayed both LOD and LOQ of 2 pg and dynamic ranges of 5 orders of magnitude. *M. productus* was the species revealing the worst sensitivity (LOD and LOQ of 20 pg) and, therefore, a dynamic range of 4 orders of magnitude (Figure 3C, 3D).

To evaluate the performance of the real-time PCR methods, other parameters have to comply with the acceptance criteria established for this type of assay, namely PCR efficiency that should be between 90-110%, the slope within -3.6 and -3.1 and the correlation coefficient (R^2) above 0.98 (ENGL, 2015). Thus, the real-time PCR assays evidenced high performance, presenting values of PCR efficiency in the range of 100.5-109.6%, slopes from -3.311 to -3.111 and R^2 between 0.986 and 0.998, for the calibration curves obtained with the five *Merluccius* species (Figure 3).

Accordingly, the sensitivity of the method enables the correct identification of the hake species down to 0.2-20 pg of DNA target, suggesting that this technique is highly adequate for the differentiation of hake species, even when present at trace levels.

Analysis of commercial foodstuffs

Table 2 presents the 45 analysed seafood samples, with respective relevant labelling information and summarised results of PCR and HRM analysis. It is possible to observe that almost half of the products (44%) did not indicate the scientific name of the species as they only presented the common name of the fish (e.g. hake) or the indication of “may contain fish”. In fact, the inclusion of the scientific name of the species is facultative for processed fish products. According to the Regulation (EC) No 1379/2013, the inclusion of the scientific names of the species in all the processed fish samples (excluding salted and smoked fish) is voluntary. Fish species that constitute an ingredient of another food maybe termed as "fish", provided that the name and presentation of such food does not refer to a specific species.

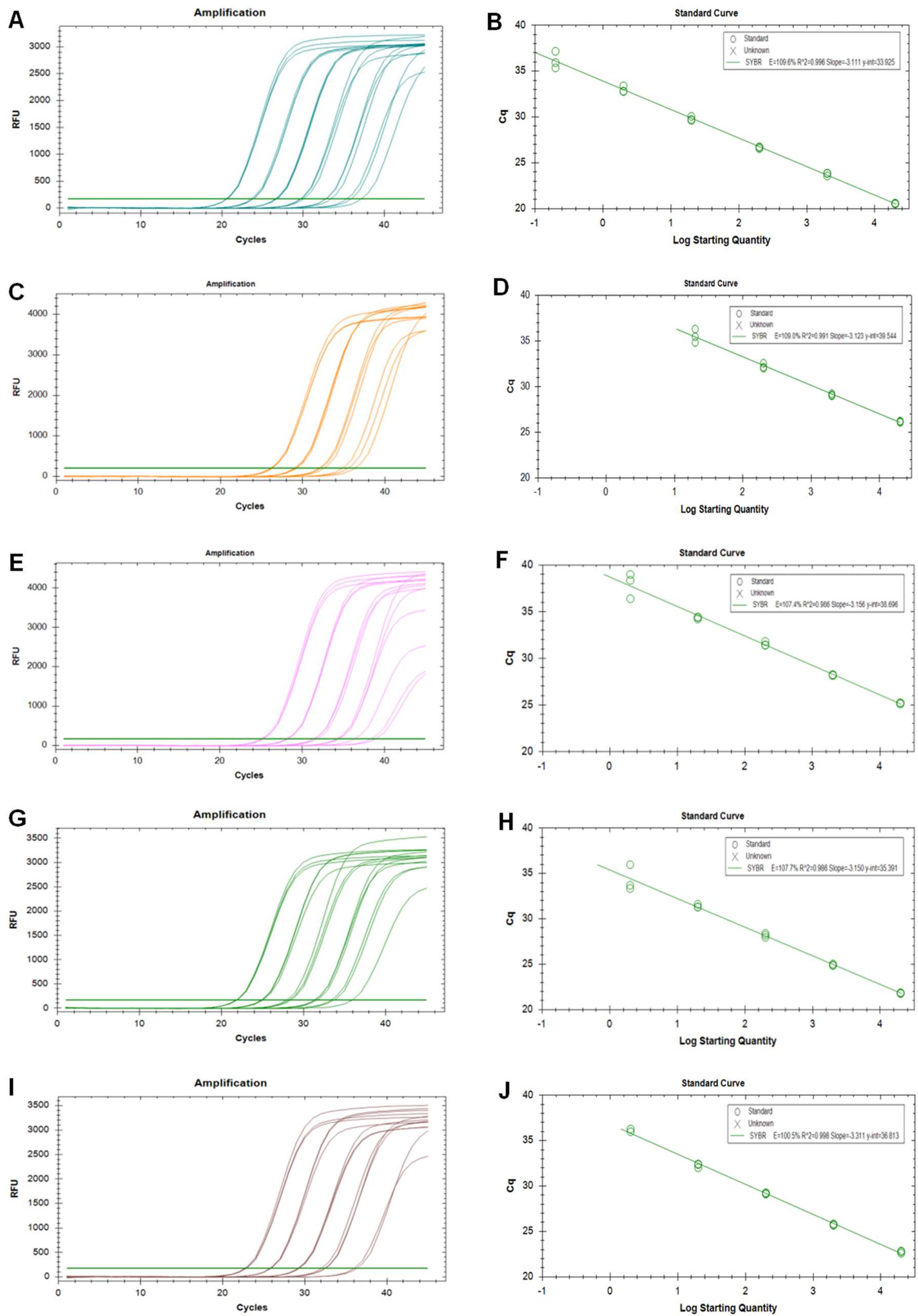


Figure 3. Real-time PCR amplification and respective calibration curves of 10-fold dilutions of target DNA (20 ng to 0.2 pg) for hake species. (A, B) *M. merluccius*; (C, D) *M. productus*; (E, F) *M. hubbsi*; (G, H) *M. capensis*; (I, J) *M. paradoxus*.

Table 2. Resumed results of qualitative PCR, real-time PCR and HRM analysis of seafood samples

Sample #	Commercial sample	Declared species or relevant information on the label	Qualitative PCR [a]		Real-time PCR		
			Eukaryotic (18SRG-F/ 18SRG-R)	Hake (MER1COI-F/ MER1COI-R)	Ct ± SD [b]	HRM cluster	Level of confidence ± SD [c] (%)
Frozen fish sticks							
1	Fish sticks	*Merluccius hubbsi*	+	+	23.9 ± 0.2	*M. hubbsi*	99.6 ± 0.3
2	Fish sticks	Hake	+	+	22.5 ± 0.1	*M. productus*	99.4 ± 0.2
3	Fish sticks	Hake	+	+	21.6 ± 0.1	*M. productus*	99.0 ± 1.1
4	Fish sticks	Hake	+	+	20.5 ± 0.1	*M. productus*	99.4 ± 0.7
5	Fish sticks	Hake	+	+	22.6 ± 0.5	*M. productus*	99.5 ± 0.1
6	Fish sticks	Hake	+	+	21.6 ± 0.1	*M. productus*	99.5 ± 0.3
7	Fish sticks	Hake	+	+	22.8 ± 0.4	*M. productus*	99.5 ± 0.4
8	Fish sticks	*Theragra chalcogramma*	+	-	NA [d]	NA	NA
9	Fish sticks	*T. chalcogramma, M. productus, Gadus morhua, Pollachius virens*	+	-	NA	NA	NA
10	Fish sticks	*P. virens*	+	-	NA	NA	NA
11	Fish sticks	*T. chalcogramma*	+	-	NA	NA	NA
12	Fish sticks	*P. virens*	+	+/-	> 38	NA	NA
13	Codfish sticks	Codfish	+	-	NA	NA	NA
14	Fish nuggets	*T. chalcogramma*	+	-	NA	NA	NA
15	Breaded hake	Hake	+	+	22.5 ± 0.1	*M. productus*	99.5 ± 0.3
16	Breaded hake	Hake	+	+	20.7 ± 0.2	*M. productus*	99.5 ± 0.3
17	Breaded fish	*T. chalcogramma*	+	-	NA	NA	NA
18	Breaded fish	*T. chalcogramma*	+	-	NA	NA	NA
19	Breaded fish	*T. chalcogramma*	+	-	NA	NA	NA
20	Breaded fish	*T. chalcogramma*	+	+/-	> 38	NA	NA

[a] (-) no amplification, (+) positive amplification with strong/moderate bands, (+/-) faint bands of qualitative PCR; [b] mean cycle threshold values (Ct) ± standard deviation (SD); [c] mean level of confidence values ± SD; [d] NA – not applied; [e] NRC – no reference cluster.

Table 2. (continuation)

Sample #	Commercial sample	Declared species or relevant information on the label	Qualitative PCR [a]		Real-time PCR		
			Eukaryotic (18SRG-F/ 18SRG-R)	Hake (MER1COI-F/ MER1COI-R)	Ct ± SD [b]	HRM cluster	Level of confidence ± SD [c] (%)
Frozen fish patties/pies/cakes							
21	Hake patties	Hake	+	+	28.2 ± 0.1	*M. capensis*	99.4 ± 0.3
22	Hake and shrimp patties	Hake	+	+	22.3 ± 0.4	*M. hubbsi*	99.9 ± 0.1
23	Codfish pie	Codfish	+	-	NA	NA	NA
24	Codfish cakes	*G. morhua*	+	-	NA	NA	NA
25	Codfish cakes	*G. morhua, G. macrocephalus*	+	-	NA	NA	NA
26	Codfish “patanisca”	*G. morhua*	+	+/-	> 38	NA	NA
27	Codfish “alheira”	Codfish	+	-	NA	NA	NA
Frozen pre-cooked fish meals							
28	Seafood risotto	Hake	+	+	27.2 ± 0.3	*M. hubbsi*	99.7 ± 0.1
29	“Brás” style Codfish	*G. microcephalus, G. morhua*	+	+	33.8 ± 0.6	*M. hubbsi*	98.9 ± 1.7
30	“Brás” style Codfish	Codfish	+	-	NA	NA	NA
31	“Gomes de Sá” style codfish	Codfish	+	-	NA	NA	NA
32	Codfish with cream	*G. morhua*	+	-	NA	NA	NA
33	Codfish with cream	Codfish	+	+/-	> 38	NA	NA
34	Paella	May contain fish	+	-	NA	NA	NA
35	Codfish lasagne	Codfish	+	+/-	> 38	NA	NA
Surimi							
36	Surimi	Fish	+	+	26.4 ± 0.4	NRC [e]	99.0 ± 0.3
37	Surimi	*T. chalcogramma*	+	-	NA	NA	NA
38	Surimi	*T. chalcogramma*	+	-	NA	NA	NA

[a] (-) no amplification, (+) positive amplification with strong/moderate bands, (+/-) faint bands of qualitative PCR; [b] mean cycle threshold values (Ct) ± standard deviation (SD); [c] mean level of confidence values ± SD; [d] NA – not applied; [e] NRC – no reference cluster

Table 2. (continuation)

Sample #	Commercial sample	Declared species or relevant information on the label	Qualitative PCR [a]		Real-time PCR		
			Eukaryotic (18SRG-F/ 18SRG-R)	Hake (MER1COI-F/ MER1COI-R)	Ct ± SD [b]	HRM cluster	Level of confidence ± SD [c] (%)
Fresh and frozen fish fillets							
39	Hake medallion	*M. capensis, M. paradoxus*	+	+	21.6 ± 0.1	*M. capensis*	99.3 ± 0.2
40	Hake medallion	*M. campensis, M. productus*	+	+	20.4 ± 0.4	*M. capensis*	99.5 ± 0.2
41	Fresh hake	*M. merluccius*	+	+	20.1 ± 0.1	*M. merluccius*	99.0 ± 0.6
42	Hake fillet	*M. productus*	+	+	18.5 ± 0.1	*M. productus*	99.9 ± 0.1
43	Hake fillet	*M. hubbsi*	+	+	19.7 ± 0.1	*M. hubbsi*	99.4 ± 0.2
44	Hake fillet	*M. capensis*	+	+	20.1 ± 0.1	*M. paradoxus*	99.1 ± 0.4
45	Hake fillet	*M. capensis, M. paradoxus*	+	+	23.1 ± 0.1	*M. paradoxus*	99.7 ± 0.1

[a] (-) no amplification, (+) positive amplification with strong/moderate bands, (+/-) faint bands of qualitative PCR; [b] mean cycle threshold values (Ct) ± standard deviation (SD); [c] mean level of confidence values ± SD; [d] NA – not applied; [e] NRC – no reference cluster.

As it is possible to notice (Table 2), all the samples provided extracts with amplifiable DNA since they were all positive with the universal eukaryotic primers (18SRG-F/18SRG-R) (Costa et al., 2013). The results of PCR amplification with primers MER1COI-F/MER1COI-R showed that almost half of the samples (47%) were positive for *Merluccius* species, which was confirmed by real-time PCR with Ct values ranging from 18.5 to 33.8 (Table 2). The conventional melting curve analysis shows amplified fragments with single melt peaks of about 78 ºC, suggesting that they were all specific products of *Merluccius* spp. (Figure 4A). The barcode-HRM approach was successfully applied to identify hake species (Figures 4B and 4C, Table 2), which was accomplished based on cluster classification with high levels of confidence (≥99%). It is important to highlight the high levels of confidence even for several samples that were pre-cooked meals and others subjected to extensive degree of processing. Regarding the clustering classification, 9 samples were included in the *M. productus* cluster, 5 in the *M. hubbsi* cluster, 3 in the *M. capensis*, 2 in the *M. paradoxus*, 1 in *M. merluccius* and 1 with no assigned reference cluster (sample #36, surimi) (Table 2). This later result might be explained by the possible mixture of different hake species in this type of fish paste or by the presence of other *Merluccius* spp. rather than the five reference species. When comparing the results with the labelling information, two samples (#28 and #44), out of 20 with identified *Merluccius* spp., were not according with the declared species, suggesting mislabelling or adulteration practices. It should be noted that sample #28 is a traditional codfish dish declaring two codfish species on the label, in which a hake species (*M. hubbsi*) was detected. This finding suggests the partial or complete substitution of codfish with hake species. As expected, *M. merluccius* was the least frequent hake species found in the analysed foods, probably due to its inherent high market value, being *M. productus* and *M. hubbsi* the predominant species found in Portuguese hake-containing products.

From the analysed samples, 25 declaring mainly codfish species on their labels were also included in this work since they could be substituted by undeclared hake species. This was the case of sample #28, while the other 24 were negative for *Merluccius* spp. The 5 samples producing faint bands and late real-time PCR amplification (Ct > 38) suggest possible cross-contamination with hake species during production/processing.

Discussion

Currently, few works have reported methods based on the combination of HRM analysis targeting DNA mini-barcodes for the differentiation of fish species. McGlauflin et al. (2010) described the use of HRM analysis for the identification of 11 single-nucleotide polymorphisms (SNP) to distinguish rainbow trout (*Oncorhynchus mykiss*) and cutthroat

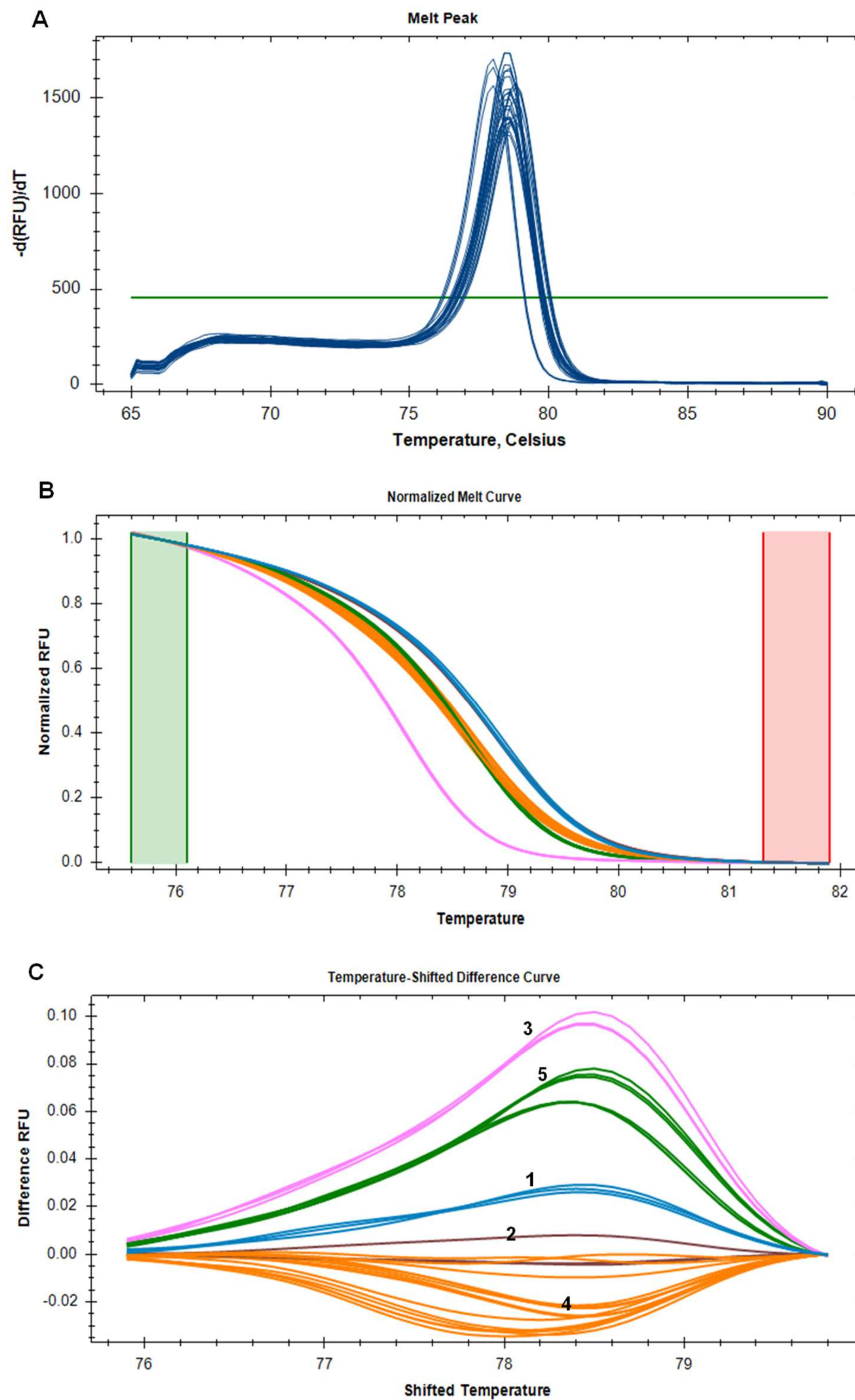


Figure 4. Conventional melting curves (A) and HRM data, namely, normalised melting curves (B) and temperature-shifted difference curves (C) of PCR products targeting the COI region of *Merluccius* spp. and analysed seafoods. Legend: 1, *M. merluccius* cluster; 2, *M. paradoxus* cluster; 3, *M. hubbsi* cluster; *4, M. productus* cluster, fish sticks (samples #2 to #5) and breaded hake (sample #15); 5, *M. capensis* cluster, hake medallion (sample #40).

trout (*Oncorhynchus clarkii*). Fitzcharles (2012) demonstrated the potential of HRM analysis as a fast and reliable technique for the differentiation Antarctic fish species, genus *Macrourus*. Recently, Fernandes et al. (2017) exploited the use of two DNA barcode regions, namely COI and *cytb*, combined with HRM analysis to differentiate four important gadoid species: Atlantic cod (*Gadus morhua*), Pacific cod (*Gadus macrocephalus*), Alaska pollock (*Theragra chalcogramma*) and saithe (*Pollachius virens*). Both barcode-HRM approaches were successfully developed, though the *cytb* barcode-HRM assay offered the best performance, enabling discriminating the four cod-like fish species in different clusters.

In what regards Merluccidae species, to our knowledge, this is the first barcode-HRM approach proposed for their differentiation. Relating to DNA barcoding combined with sequencing, there are some reports involving Merluccidae species differentiation. Nicolé et al. (2012) presented a multi-locus approach (apocytochrome b encoding gene (*cob*), COI and 16S rDNA) to genetically identify crustaceans, molluscs and fish (comprising *M. productus*, *M. hubbsi* and *M. paradoxus*) in seafood samples. Deli Antoni et al. (2015) were able to discriminate *M. hubbsi* and *M. australis* from *M. patagonicus*, exploiting COI gene. More recently, Ferrito et al. (2016) combined COI barcode amplification with the analysis of restriction fragment length polymorphisms (RFLP), being able to detect *Theragra chalcogramma*, *M. merluccius*, *M. productus* and *M. paradoxus* species in seafood products.

Concerning other DNA-based approaches, Garcia-Vazquez et al. (2011) explored 5S rDNA *loci* to differentiate *Merluccius* spp. by qualitative PCR, while Sanchéz et al. (2009) used real-time PCR to identify *M. merluccius* in fish samples by Ct value comparison. Machado-Schiaffino et al. (2008) applied a mitochondrial SNP approach using several sets of primers and capillary electrophoresis, which allowed a 20% of mislabelling detection within hake species. Chapela et al. (2007) developed a screening method based on PCR targeting the *cytb* gene and SSCP analysis, which enabled them to differentiate several species of *Merluccius*. Hubalkova et al. (2009) tested different approaches for the differentiation of hakes from other gadoid species or intraspecies identification of hake species based on PCR-RFLP and PCR-sequencing of two genes (*pantophysin I* and *cytb*), which enabled them to successfully identify three *Merluccius* species in 20 commercial samples. Comparing with the above-mentioned reports, the proposed method has the advantage of not requiring any post-PCR analysis, being able to identify the five *Merluccius* species based on the barcoding approach without the need of further sequencing.

Conclusion

In the present article, a COI mini-barcode combined with HRM analysis was successfully proposed for the authentication of hake species. The method allowed the full discrimination of five *Merluccius* species commonly used in processed foods (*M. merluccius*, *M. productus*, *M. hubbsi*, *M. capensis* and *M. paradoxus*) with high levels of confidence. The novel approach evidenced clear advantages in terms of speed, cost, simplicity and specificity when compared with most DNA barcoding approaches that rely on sequencing. The proposed COI mini-barcode, as a target for real-time PCR amplification, provided a high sensitive tool for the detection of each of the five hake species down to 0.2-20 pg of hake DNA, with adequate performance parameters of PCR efficiency and linearity. The application of the COI-HRM approach to 45 commercial fish-containing foods showed that from 20 samples declaring hake species, one did not complied with the declared *Merluccius* spp., while the others were in good agreement. The remaining samples declaring codfish/fish were mostly negative for hake species, suggesting possible cross-contamination in five samples, and the addition/substitution of undeclared hake to a codfish food product. These findings suggest mislabelling or adulteration by species substitution and the feasibility of the proposed tool to authenticate processed fish-containing foods at trace levels.

Acknowledgements

This work was supported by FCT (Fundação para a Ciência e Tecnologia) through project UID/QUI/50006/2013 – POCI/01/0145/FEDER/007265 with financial support from FCT/MEC through national funds and co-financed by FEDER, under the Partnership Agreement PT2020 and by the project NORTE-01-0145-FEDER-000011. T. J. R. Fernandes and J. Costa are grateful to FCT grants (SFRH/BD/93711/2013 and SFRH/BPD/102404/2014, respectively) financed by POPH-QREN (subsidised by FSE and MCTES).

References

Abdullah, A., Rehbein, H., 2016. The differentiation of tuna (family: Scombridae) products through the PCR-based analysis of the cytochrome b gene and parvalbumin introns. J. Sci. Food Agric. 96, 456-464.

Armani, A., Castigliego, L., Tinacci, L., Gianfaldoni, D., Guidi, A., 2012. Multiplex conventional and real-time PCR for fish species identification of Bianchetto (juvenile form of *Sardina pilchardus*), Rossetto (*Aphia minuta*), and Icefish in fresh, marinated and cooked products. Food Chem. 133, 184-192.

Armani, A., Guardone, L., 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.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clin. Chem. 55, 611-622.

Castigliego, L., Armani, A., Tinacci, L., Gianfaldoni, D., Guidi, A., 2015. Two alternative multiplex PCRs for the identification of the seven species of anglerfish (*Lophius* spp.) using an end-point or a melting curve analysis real-time protocol. Food Chem. 166, 1-9.

Chapela, M.J., Sanchez, A., Suarez, M.I., Perez-Martin, R.I., Sotelo, C.G., 2007. A rapid methodology for screening hake species (*Merluccius* spp.) by single-stranded conformation polymorphism analysis. J. Agric. Food Chem. 55, 6903-6909.

Carvalho, D.C., Guedes, D., Trindade, M.G., Coelho, R.M.S., Araujo, P.H.L., 2017. Nationwide Brazilian governmental forensic programme reveals seafood mislabelling trends and rates using DNA barcoding. Fish. Res. 191, 30-35.

Costa, F.O., Carvalho, G.R., 2007. The barcode of life initiative: synopsis and prospective societal impacts of DNA barcoding of fish. Genom. Soc. Policy. 3, 29-40.

Costa, J., Oliveira, M.B.P.P., Mafra, I., 2013. Effect of thermal processing on the performance of the novel single-tube nested real-time PCR for the detection of walnut allergens in sponge cakes. Food Res. Int. 54, 1722-1729.

Deli Antoni, M.Y., Gonzalez-Castro, M., Diaz de Astarloa, J.M., 2015. New tools (DNA barcoding), old hypothesis: the case of the taxonomic identity of the Argentine hakes (Actinopterygii: *Merluccius*). J. Fish Biol. 87, 783-793.

Di Pinto, A., Di Pinto, P., Terio, V., Bozzo, G., Bonerba, E., Ceci, E., Tantillo, G., 2013. DNA barcoding for detecting market substitution in salted cod fillets and battered cod chunks. Food Chem. 141, 1757-1762.

ENGL, 2015. Definition of minimum performance requirements for analytical methods of GMO testing. European network of GMO laboratories, Joint Research Center, EURL Available online at http://gmo-crl.jrc.ec.europa.eu/doc/MPR%20Report%20Application%2020_10_2015.pdf (Website last accessed April 6, 2017).

Fernandes, T.J.R., Costa, J., Oliveira, M.B.P.P., Mafra, I., 2017. DNA barcoding coupled to HRM analysis as a new and simple tool for the authentication of Gadidae fish species. Food Chem. 230, 49-57.

Ferrito, V., Bertolino, V., Pappalardo, A.M., 2016. White fish authentication by COIBar-RFLP: Toward a common strategy for the rapid identification of species in convenience seafood. Food Control. 70, 130-137.

Fields, A.T., Abercrombie, D.L., Eng, R., Feldheim, K., Chapman, D.D., 2015. A novel mini-DNA barcoding assay to identify processed fins from internationally protected shark species. PLoS One. 10, 1-10.

Fitzcharles, E.M., 2012. Rapid discrimination between four Antarctic fish species, genus Macrourus, using HRM analysis. Fish. Res. 127–128, 166-170.

Garcia-Vazquez, E., Perez, J., Martinez, J.L., Pardiñas, A.F., Lopez, B., Karaiskou, N., Casa, M.F., MacHado-Schiaffino, G., Triantafyllidis, A., 2011. High level of mislabeling in Spanish and Greek hake markets suggests the fraudulent introduction of African species. J. Agric. Food Chem. 59, 475-480.

Handy, S.M., Deeds, J.R., Ivanova, N.V., Hebert, P.D.N., Hanner, R.H., Ormos, A., Weigt, L.A., Moore, M.M., Yancy, H.F., 2011. A single-laboratory validated method for the generation of DNA barcodes for the identification of fish for regulatory compliance. J. AOAC Int. 94, 201-210.

Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. P. Roy. Soc. Lon. B Bio. 270, 313-321.

Hird, H., Chisholm, J., Kaye, J., Colyer, A., Hold, G., Conyers, C., Núñez, J., Macarthur, R., 2012. Development of real-time PCR assays for the detection of Atlantic cod (*Gadus morhua*), Atlantic salmon (*Salmo salar*) and European plaice (*Pleuronectes platessa*) in complex food samples. Eur. Food Res. Technol. 234, 127-136.

Hubalkova, Z., Kralik, P., Kasalova, J., Rencova, E., 2008. Identification of gadoid species in fish meat by polymerase chain reaction (PCR) on genomic DNA. J. Agric. Food Chem. 56, 3454-3459.

Hubalkova, Z., Kralík, P., Kasaloví, J., Rencova, E., 2009. Application of DNA-based techniques for intraspecies differentiation of hake fish. Acta Vet. Brno 78, 673-678.

Little, D.P., 2014. A DNA mini-barcode for land plants. Mol. Ecol. Resour. 14, 437-446.

Machado-Schiaffino, G., Martinez, J.L., Garcia-Vazquez, E., 2008. Detection of mislabeling in hake seafood employing mtSNPs-based methodology with identification of eleven hake species of the genus *Merluccius*. J. Agric. Food Chem. 56, 5091-5095.

Mafra, I., Ferreira, I.M.P.L.V.O., Oliveira, M.B.P.P., 2008. Food authentication by PCR-based methods. Eur. Food Res. Technol. 227, 649-665.

McGlauflin, M.T., Smith, M.J., Wang, J.T., Young, S.F., Chen, N., Lee, Y.C., Pascal, C., Seeb, L.W., Stevens, J., Seeb, J.E., 2010. High-resolution melting analysis for the discovery of novel single-nucleotide polymorphisms in rainbow and cutthroat trout for species identification. T. Am. Fish. Soc. 139, 676-684.

Mitchell, J.K., Hellberg, R.S., 2016. Use of the mitochondrial control region as a potential DNA mini-barcoding target for the identification of canned tuna species. Food Anal. Method. 9, 2711-2720.

Muñoz-Colmenero, M., Klett-Mingo, M., Díaz, E., Blanco, O., Martínez, J.L., Garcia-Vazquez, E., 2015. Evolution of hake mislabeling niches in commercial markets. Food Control 54, 267-274.

Nicolé, S., Negrisolo, E., Eccher, G., Mantovani, R., Patarnello, T., Erickson, D.L., Kress, W.J., Barcaccia, G., 2012. DNA barcoding as a reliable method for the authentication of commercial seafood products. Food Technol. Biotechnol. 50, 387-398.

Pappalardo, A., Ferrito, V., 2015. DNA barcoding species identification unveils mislabeling of processed flatfish products in southern Italy markets. Fish. Res., 164, 153-158.

Pappalardo, A.M., Cuttitta, A., Sardella, A., Musco, M., Maggio, T., Patti, B., Mazzola, S., Ferrito, V., 2015. DNA barcoding and COI sequence variation in Mediterranean lanternfishes larvae. Hydrobiol., 749, 155-167.

Rasmussen, R.S., Morrissey, M.T., 2008. DNA-Based Methods for the Identification of Commercial Fish and Seafood Species. Compr. Rev. Food Sci. F. 7, 280-295.

Rasmussen, R.S., Morrissey, M.T., 2009. Application of DNA-Based Methods to Identify Fish and Seafood Substitution on the Commercial Market. Compr. Rev. Food Sci. F. 8, 118-154.

Regulation (EU) No 1169/2011 of the European Parliament and of the Council of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No. 1924/2006 and (EC) No. 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004, Off. J. Eur. Union L304, 18-63.

Regulation (EU) No 1379/2013 of the European Parliament and of the Council of 11 December 2013 on the common organisation of the markets in fishery and aquaculture products, amending Council Regulations (EC) No 1184/2006 and (EC) No 1224/2009 and repealing Council Regulation (EC) No 104/2000. Off. J. Eur. Union L354, 1-21.

Sánchez, A., Quinteiro, J., Rey-Mendez, M., Perez-Martín, R.I., Sotelo, C.G., 2009. Identification of European hake species (*Merluccius merluccius*) using real-time PCR. J. Agric. Food Chem. 57, 3397-3403.

Sánchez, A., Vázquez, J.A., Quinteiro, J., Sotelo, C.G., 2013. Modeling real-time PCR kinetics: Richards reparametrized equation for quantitative estimation of European hake (*Merluccius merluccius*). J. Agric. Food Chem. 61, 3488-3493.

Taboada, L., Sánchez, A., Sotelo, C.G., 2017. A new real-time PCR method for rapid and specific detection of ling (*Molva molva*). Food Chem. 228, 469-475.

Villa, C., Costa, J., Meira, L., Oliveira, M.B.P.P., Mafra, I., 2016. Exploiting DNA mini-barcodes as molecular markers to authenticate saffron (*Crocus sativus* L.). Food Control. 65, 21-31.

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., Hebert, P.D.N., 2005. DNA barcoding Australia's fish species. Philos. Trans. R. Soc. B-Biol. Sci. 360, 1847-1857.

Xanthopoulou, A., Ganopoulos, I., Kalivas, A., Osathanunkul, M., Chatzopoulou, P., Tsaftaris, A., Madesis, P., 2016. Multiplex HRM analysis as a tool for rapid molecular authentication of nine herbal teas. Food Control. 60, 113-116.

CHAPTER 2. SEAFOOD ALLERGEN DETECTION

STATE-OF-THE-ART

An overview of fish and shellfish allergens and current methods of detection.
Food and Agricultural Immunology, **2015**, 26, 848-869.

Advances on the molecular characterisation, clinical relevance and detection methods of Gadiform parvalbumin allergens.
Critical Reviews in Food Science and Nutrition, **2017**, 15, 3281-3296.

EXPERIMENTAL PART

Exploiting 16S rRNA gene for the detection and quantification of fish as a potential allergenic food: A comparison of two real-time PCR approaches.
Food Chemistry, **2018**, 245, 1034-1041.

A new real-time PCR quantitative approach for the detection of shrimp crustaceans as potential allergens.
Journal of Food Composition and Analysis, **2017**, submitted.

STATE-OF-THE-ART

An overview of fish and shellfish allergens and current methods of detection.
Food and Agricultural Immunology, **2015**, 26, 848-869.

Advances on the molecular characterisation, clinical relevance and detection methods of Gadiform parvalbumin allergens.
Critical Reviews in Food Science and Nutrition, **2017**, 15, 3281-3296.

Food and Agricultural Immunology, 2015
Vol. 26, No. 6, 848–869, http://dx.doi.org/10.1080/09540105.2015.1039497

Taylor & Francis
Taylor & Francis Group

An overview on fish and shellfish allergens and current methods of detection

Telmo J.R. Fernandes, Joana Costa, M. Beatriz P.P. Oliveira, Isabel Mafra*

REQUIMTE-LAQV, Faculdade de Farmácia, Universidade do Porto, Rua de Jorge Viterbo Ferreira, 228, 4050-313 Porto, Portugal

*Corresponding author: Tel: +351 220428640. Fax: +351 226093390. E-mail: isabel.mafra@ff.up.pt

Abstract

Food induced allergies are considered an important problem of public health with special impact in the quality of life of the sensitised/allergic individuals. As highly consumed foods, fish and shellfish represent a valuable source of proteins for the general population. In spite of their economical and nutritional importance, these foods are known to induce hypersensitivity reactions in sensitised/allergic individuals. So far, parvalbumins (fish) and tropomyosins (crustaceans and molluscs) have been considered major allergens in seafood allergy, being responsible for most of the reported cases of adverse immunological responses. More recently, other proteins such as arginine kinases, myosin light chains, troponins and sarcoplasmic calcium-binding proteins have been regarded as relevant allergens in fish, crustaceans and molluscs. This review focuses on seafood allergens, reporting an updated and compiled list of allergens from fish, crustaceans and mollusc species, with an overview on the most representative analytical methods for their detection.

Keywords: allergens, seafood, parvalbumins, tropomyosins, detection methods.

Introduction

Seafood plays an important role in human nutrition and health, which is considered an excellent source of highly assimilated proteins, vitamins and polyunsaturated fatty acids, such as docosahexaenoic acid (22:6) and eicosapentaenoic acid (20:5) from the *n*-3 series (omega-3) (Perez-Gordo et al., 2011; Sharp & Lopata, 2014). There is a solid scientific background reporting the wide range of health benefits associated with the ingestion of omega-3 fatty acids, such as the prevention of cardiovascular diseases and cancer or the improvement of glycaemic control (Larsen, Eilertsen, & Elvevoll, 2011; Mozzafarian, Bryson, Lemaitre, Siscovick, & Burke, 2005; Sirot, Oseredczuk, Bemrah-Aouachria, Volatier, & Burke, 2008; Van Do, Elsayed, Florvaag, Hordvik, & Endresen, 2005).

Due to the referred well-established health benefits, the consumption of fish and shellfish has been continuously increasing worldwide. For the majority of the world's population, the growing interest in seafood intake can be considered a nutritional advantage. However, for a small but rather significant part of food-allergic individuals, the consumption of products containing undeclared seafood can pose severe health problems (e.g. systemic immunological reactions, anaphylaxis) as result of accidental exposure to the offending food (Madsen et al., 2012). In the recent years, more cases of fish and shellfish allergies have been frequently reported, being currently viewed as an emergent issue of public health. The clinical diagnosis of specific food allergies, such as seafood allergy, is based on self-reported symptoms (clinical history), specific immunoglobulin E (sIgE) blood tests or skin prick test (SPT) sensitisation, rather than open food challenges (OFC) or double blind placebo controlled food challenges (DBPCFC) (Burks et al., 2012), making the true prevalence of seafood allergy difficult to establish. In spite of this, recent data seem to suggest that 0.1-0.4% of general population is affected by fish allergy, while over 2% suffer from shellfish allergy. Nevertheless, the referred prevalence can vary with specific geographical and cultural eating habits and/or with the type of food processing (Chen et al., 2013a; Kamath et al., 2014; Kuehn et al., 2013). In the USA, the available reports suggest that 0.4% of the population suffers from fish allergy and 0.2% is affected by both fish and shellfish allergies. In Europe, fish allergy was estimated to have an overall incidence of 0.2%, while the allergy to shrimp, a major contributor in shellfish allergy, presented a prevalence of 5.4% (Burney et al., 2010; Sicherer, Muñoz-Furlong, & Sampson, 2004). Coastal countries like Portugal and Finland, where the consumption of seafood is very high, were not included in the referred European prevalence study. Therefore, the prevalence of seafood allergy in Europe is probably underestimated

(Lopata, O'Hehir, & Lehrer, 2010; Perez-Gordo et al., 2011; Sharp & Lopata, 2014; Tsabouri et al., 2012).

Fish and crustaceans are known to induce hypersensitivity reactions mediated by the IgE in sensitised/allergic individuals (Kuehn, Scheuermann, Hilger, & Hentges, 2010; Lee & Taylor, 2011), being two of the eight groups responsible for almost 90% of worldwide reported food allergies (CODEX, 1985). The routes of exposure for seafood allergy are ingestion, direct contact (skin) or inhalation of their odours or fumes created during preparation/cooking of derived foods (Kuehn et al., 2010; Lee & Taylor, 2011). For allergic individuals, the only effective means of preventing an adverse reaction is the total avoidance of seafood or the use of a therapeutic treatment (e.g. antihistaminic, corticosteroids, epinephrine) in the case of an accidental exposure to the allergenic food (van Hengel, 2011). Consequently, it became imperative to improve consumer's protection through an accurate food labelling system, in order to prevent potential life-threatening risks for sensitised/allergic individuals (Costa, Carrapatoso, Oliveira, & Mafra, 2014; Rencova, Kostelnikova, & Tremlova, 2013). According to the recent European Union (EU) regulations, food producers are obligated to declare the presence of fourteen groups of foods that are recognised as potentially allergenic, namely fish, crustaceans, molluscs, celery, mustard, sesame seed, gluten, tree nuts, peanuts, milk, eggs, soybeans, lupine and sulphites, as well as highlighting them from the rest of the list of ingredients (Directive 2007/68/EC; Directive 2000/13/EC; Regulation No 1169/2011).

This report intends to provide a general and updated overview on seafood allergens, focusing on fish and shellfish (crustaceans and molluscs), the main consumed seafood groups, including a brief description of the most representative analytical methods for their detection.

Fish allergens

So far, some families of proteins such as enolases, aldolases and parvalbumins have been classified as allergens in fish, although the most representative one correspond to parvalbumins. Included in the calcium-binding proteins, which comprise the second most important family of animal food allergens, parvalbumins are currently reported as responsible for more than 95% of food allergies induced by fish. Generally, the symptoms occur within 30 min after the contact with the offending food and can result in skin, respiratory and gastrointestinal symptoms, including less frequent fatal systemic responses such as anaphylaxis (Kuehn et al., 2010; Lee & Taylor, 2011; Weber & Paschke, 2010).

Parvalbumins

Parvalbumins are small (10-13 kDa), acidic and water-soluble proteins, presenting remarkable resistance to high temperatures, denaturing agents and proteolytic activity (Griesmeier et al., 2010; Lee & Taylor, 2011; Weber & Paschke, 2010). They are usually divided in two evolutionary lineages of isoforms: the α-parvalbumins, which are generally classified as non-allergenic; and the β-parvalbumins, wherein the majority of IgE reactive parvalbumins are included (Jenkins, Breiteneder, & Mills, 2007; Weber & Paschke, 2010; Wang et al., 2014). Parvalbumins are abundant proteins in the white muscle of many fish species, performing an important role in the relaxation of muscle fibres through the binding of free intracellular calcium. They are composed by two functional domains, each binding a calcium ion, and a third silent domain protecting the hydrophobic core of the protein. In these proteins, the binding of calcium is thought to be of critical relevance to the integrity of the conformation of the IgE epitopes. Calcium depletion is known to induce structural alterations in these proteins, decreasing the allergenic capacity of parvalbumins (Arif, Jabeen & Hasnain, 2007; Bugajska-Schretter et al., 1998; Bugajska-Schretter et al., 2000; Capony & Pechère, 1973). These sarcoplasmatic proteins are present at high proportion in the bottom dwelling fish species, such as cod, whiff or flounder. Therefore, these species are expected to have higher allergenicity than active fishes (rich in dark muscle) such as tuna, mackerel, swordfish and skipjack (Jenkins et al., 2007; Rencova et al., 2013; Tsabouri et al., 2012).

Tolerance to a certain fish species can vary greatly among allergic individuals, so there is about 50% of chance for a sensitised patient to be cross-reactive to more than one fish species. This happens because the secondary and tertiary structures of parvalbumins are highly conserved, though their amino acid sequences (primary structures) can differ substantially among fish species (Sharp & Lopata, 2014). In spite of evidencing very distinct IgE-binding epitopes, limited data on epitope alignment of four parvalbumins from different fish species, namely salmon, cod, mackerel and carp, seem to indicate the presence of a highly antigenic region (region IV) (Bugajska-Schretter et al., 1998; Perez-Gordo et al., 2012; Sharp & Lopata, 2014), which might be responsible for the polysensitisation to multiple fish species in allergic individuals (Griesmeier et al., 2010). Current studies on cross-reactivity phenomena highlight the need of sensitised/allergic individuals to eliminate from diet any kind of fish, even before the performance of allergy diagnosis by SPT, serum-specific IgE blood tests, or OFC (Carrapatoso, 2004; Lieberman & Sicherer, 2011; Muraro et al., 2014; Perez-Gordo et al, 2011; Sharp & Lopata, 2014; Tsabouri et al., 2012).

Another major aspect of fish allergy concerns the fate of allergenic proteins during food processing since parvalbumins are considered highly stable proteins, being most frequently resistant to common physical and chemical processes. So far, it is still uncertain how food processing may affect parvalbumins from distinct fish species. According to literature, heat treatments do not affect IgE-binding capacity since these proteins are able to return to their original conformation after cooling (Mills, Sancho, Rigby, Jenkins, & Mackie, 2009). IgE-binding capacity of parvalbumins can be reduced by chemical processes. Proteolysis, often combined with pH alterations, is another efficient way of decreasing allergenicity, however it may also contribute to expose pre-existing epitopes or create new epitopes by aggregation (Sletten, Van Do, Lindvik, Egaas, & Florvaag, 2010; Thomas et al., 2007).

Codfish allergy is presently the best well studied since most fish-allergic patients do not tolerate cod. The major allergen designated as Gad c 1, which is isolated from Baltic cod (*Gadus callarias*), is often used as a reference molecule for the study of parvalbumins. Other homologous allergens have been isolated from worldwide highly appreciated commercial fishes, namely other cod species (*Gadus morhua*), common carp (*Cyprinus carpio*), Atlantic salmon (*Salmo salar*), Japanese jack mackerel (*Trachurus japonicas*), bigeye tuna (*Thunnus obesus*), and European hake (*Merluccius merluccius*) (Perez-Gordo, 2011; Tsabouri et al., 2012; Kuehn et al., 2010).

In recent years, the number of identified allergenic proteins available at databases has increased, improving the establishment of evolutionary and structural relationships among allergens from distinct origins (Radaurer, Bublin, Wagner, Mari, & Breiteneder, 2008). Allergen platforms such as the Official List of Allergens issued by the International Union of Immunological Societies (IUIS) Allergen Nomenclature Sub-committee and the ALLERGOME database have become excellent tools for allergen classification since they report molecular, biochemical and clinical data about allergenic proteins. Tables 1 and 2 summarise all fish allergens that were already characterised, being currently available at databases (ALLERGEN, 2014; ALLERGOME, 2014). Most of the identified fish allergens correspond to parvalbumins, which account for more than 200 entries (Table 1), although different proteins, namely enolases and aldolases (Table 2) are also defined as IgE reactive in fish species.

Enolases and Aldolases

With respect to other families of allergenic proteins in fish (Table 2), 50 kDa enolases and 40kDa aldolases have been recently described as important allergens in highly consumed

Table 1. List of fish allergens (parvalbumins) and corresponding fish species.

Biochemical classification	Allergen	Fish species	Fish common name
Parvalbumin	Aca so 1	*Acanthocybium solandri*	Wahoo
	Aet ro 1	*Aethaloperca rogaa*	Redmouth grouper
	Ana l 1	*Anarhichas lupus*	Striped wolfish
(Calcium-binding protein)	Ang a 1	*Anguilla anguilla*	European eel
	Ang ja 1	*Anguilla japonica*	Japanese eel
	Ano fi 1	*Anoplopoma fimbria*	Sablefish
	Arc ja 1	*Arctoscopus japonicus*	Sailfin sandfish
	Arc pr 1	*Archosargus probatocephalus*	Sheepshead
	Ari fe 1	*Ariopsis felis*	Hardhead catfish
	Bag ma 1	*Bagre marinus*	Gafftopsail sea catfish
	Bal ca 1	*Balistes capriscus*	Gray triggerfish
	Ber sp 1	*Beryx splendens*	Splendid alfonsino
	Bor sa 1	*Boreogadus saida*	Arctic cod
	Bra br 1	*Brama brama*	Ray's bream
	Bra du 1	*Brama dussumieri*	Bream
	Bro br 1	*Brosme brosme*	Tusk
	Cal le 1	*Calamus leucosteus*	Whitebone porgy
	Can ma 1	*Canthidermis maculata*	Ocean triggerfish
	Car au 1	*Carassius auratus*	Goldfish
	Car cr 1	*Caranx crysos*	Hardtail
	Cau ch 1	*Caulolatilus chrysops*	Atlantic goldeye tilefish
	Cen s 1	*Centropristis striata*	Black sea bass
	Cep so 1	*Cephalopholis sonnerati*	Tomato hind
	Cha fa 1	*Chaetodipterus faber*	Atlantic spadefish
	Che sp 1	*Chelidonichthys spinosus*	Bluefin gurnard
	Clu h 1	*Clupea harengus*	Atlantic herring
	Clu pa 1	*Clupea pallasii*	Pacific herring
	Col sa 1	*Cololabis saira*	Pacific saury
	Con my 1	*Conger myriaster*	Whitespotted conger
	Cor hi 1	*Coryphaena hippurus*	Mahi-mahi
	Cte id 1	*Ctenopharyngodon idella*	Grass carp
	Cyn ar 1	*Cynoscion arenarius*	Sand seatrout
	Cyn ne 1	*Cynoscion nebulosus*	Spotted seatrout
	Cyp c 1	*Cyprinus carpio*	Common carp
	Dan re 1	*Danio rerio*	Zebrafish
	Das ak 1	*Dasyatis akajei*	Red stingray
	Das am 1	*Dasyatis americana*	Southern stingray
	Das sa 1	*Dasyatis sabina*	Atlantic stingray
	Dec ru 1	*Decapterus russelli*	Indian scad
	Dic la 1	*Dicentrarchus labrax*	Sea bass
	Dip ho 1	*Diplodus holbrookii*	Spottail pinfish
	Dip sa 1	*Diplodus sargus*	White seabream
	Dit te 1	*Ditrema temminckii*	Temminck's surfperch
	Elo sa 1	*Elops saurus*	Ladyfish
	Eng e 1	*Engraulis encrasicolus*	Anchovy
	Epi bl 1	*Epinephelus bleekeri*	Bleekeri's grouper
	Eng e 1	*Engraulis encrasicolus*	Anchovy
	Epi bl 1	*Epinephelus bleekeri*	Bleekeri's grouper
	Epi br 1	*Epinephelus bruneus*	Tawny grouper
	Epi co 1	*Epinephelus coioides*	Orange-spotted grouper
	Epi fl 1	*Epinephelus flavolimbatus*	Yellowfinned grouper
	Epi mc 1	*Epinephelus maculatus*	Highfin grouper

Table 1. (*Continued*)

Biochemical classification	Allergen	Fish species	Fish common name
Parvalbumin (Calcium-binding protein)	Epi mo 1	*Epinephelus morio*	Red grouper
	Epi po 1	*Epinephelus polyphekadion*	Camouflage grouper
	Epi un 1	*Epinephelus undulosus*	Wavyline grouper
	Eso lu 1	*Esox lucius*	Northern pike
	Evy j 1	*Evynnis japonica*	Crimson seabream
	Fun gr 1	*Fundulus grandis*	Gulf Killifish
	Fun he 1	*Fundulus heteroclitus*	Astronaut Fish
	Fun si 1	*Fundulus similis*	Longnose killifish
	Gad c 1	*Gadus callarias*	Baltic cod
	Gad m 1	*Gadus morhua*	Cod
	Gad ma 1	*Gadus macrocephalus*	Pacific cod
	Gen bl 1	*Genypterus blacodes*	Pink cusk-eel
	Gil mi 1	*Gillichthys mirabilis*	Long-jawed mudsucker
	Hem le 1	*Hemanthias leptus*	Longtail bass
	Hex ot 1	*Hexagrammos otakii*	Greenling
	Hip h 1	*Hippoglossus hippoglossus*	Atlantic halibut
	Hip pl 1	*Hippoglossoides platessoides*	American plaice
	Hip st 1	*Hippoglossus stenolepis*	Pacific halibut
	Hop a 1	*Hoplostethus atlanticus*	Rosy soldier fish
	Hus hu 1	*Huso huso*	Russian sturgeon
	Hym st 1	*Hymenocephalus striatissimus*	Rattail
	Hyp by 1	*Hyperoglyphe bythites*	Black driftfish
	Hyp mo 1	*Hypophthalmichthys molitrix*	Silver carp
	Hyp no 1	*Hypophthalmichthys nobilis*	Bighead carp
	Ict fu 1	*Ictalurus furcatus*	Blue catfish
	Ict pu 1	*Ictalurus punctatus*	Channel catfish
	Kat p 1	*Katsuwonus pelamis*	Skipjack tuna
	Kyp se 1	*Kyphosus sectator*	Bermuda chub
	Lac ma 1	*Labrus maximus*	Hogfish
	Lag la 1	*Lagocephalus laevigatus*	Smooth puffer
	Lar cr 1	*Larimichthys crocea*	Large yellow croaker
	Lar po 1	*Larimichthys polyactis*	Yellow croaker
	Lat c 1	*Lates calcarifer*	Giant seaperch
	Lat ja 1	*Lateolabrax japonicus*	Japanese sea perch
	Lei xa 1	*Leiostomus xanthurus*	Spot croaker
	Lep bo 1	*Lepidorhombus boscii*	Four-spot megrim
	Lep ca 1	*Lepidopus caudatus*	Silver scabbardfish
	Lep gi 1	*Lepomis gibbosus*	Pumpkinseed
	Lep i 1	*Leptomelanosoma indicum*	Indian threadfin
	Lep ma 1	*Lepomis macrochirus*	Bluegill sunfish
	Lep mi 1	*Lepidotrigla microptera*	Redwing searobin
	Lep mo 1	*Lepidopsetta mochigarei*	Dusky sole
	Lep w 1	*Lepidorhombus whiffiagonis*	Whiff
	Leu ce 1	*Leuciscus cephalus*	Chub
	Lim fe 1	*Limanda ferruginea*	Yellowtail flounder
	Lob su 1	*Lobotes surinamensis*	Tripletail
	Lop pi 1	*Lophius piscatorius*	Monkfish
	Lut a 1	*Lutjanus argentimaculatus*	Mangrove red snapper
	Lut c 1	*Lutjanus campechanus*	Red snapper
	Lut cy 1	*Lutjanus cyanopterus*	Cubera snapper
	Lut gr 1	*Lutjanus griseus*	Gray snapper
	Lut gu 1	*Lutjanus guttatus*	Spotted rose snapper

Table 1. (*Continued*)

Biochemical classification	Allergen	Fish species	Fish common name
Parvalbumin (Calcium-binding protein)	Lut jo 1	*Lutjanus jocu*	Dog teeth snapper
	Lut pu 1	*Lutjanus purpureus*	Southern red snapper
	Lut sy 1	*Lutjanus synagris*	Lane snapper
	Mac ma 1	*Macruronus magellanicus*	Patagonian grenadier
	Mac n 1	*Macruronus novaezelandiae*	Tailed hake
	Mal vi 1	*Mallotus villosus*	Capelin
	Meg sp 1	*Megalops spp*	Tarpons
	Mel ae 1	*Melanogrammus aeglefinus*	Haddock
	Men am 1	*Menticirrhus americanus*	Southern kingfish
	Mer ap 1	*Merluccius australis polylepis*	Patagonian hake
	Mer au 1	*Merluccius australis australis*	Southern hake
	Mer bi 1	*Merluccius bilinearis*	Silver hake
	Mer ca 1	*Merluccius capensis*	Stockfish
	Mer ga 1	*Merluccius gayi*	English hake
	Mer hu 1	*Merluccius hubbsi*	Argentine hake
	Mer me 1	*Merlangius merlangus*	Whiting
	Mer mr 1	*Merluccius merluccius*	European hake
	Mer pa 1	*Merluccius paradoxus*	Deepwater hake
	Mer po 1	*Merluccius polli*	Benguela hake
	Mer pr 1	*Merluccius productus*	North Pacific hake
	Mer se 1	*Merluccius senegalensis*	Black hake
	Mic po 1	*Micromesistius poutassou*	Blue whiting
	Mic sa 1	*Micropterus salmoides*	Largemouth bass
	Mic un 1	*Micropogonias undulatus*	Atlantic croaker
	Mor am 1	*Morone americana*	Sea perch
	Mor sa 1	*Morone saxatilis*	Striped bass
	Mor sc 1	*Morone saxatilis x chrysops*	White Bass x Striped Bass
	Mug c 1	*Mugil cephalus*	Striped mullet
	Mur mi 1	*Muraenolepis microps*	Smalleye moray cod
	Myc bo 1	*Mycteroperca bonaci*	Black grouper
	Myc mi 1	*Mycteroperca microlepis*	Gag
	Myc ph 1	*Mycteroperca phenax*	Scamp
	Nem vi 1	*Nemipterus virgatus*	Golden threadfin bream
	Ocy ch 1	*Ocyurus chrysurus*	Yellowtail snapper
	Onc k 1	*Oncorhynchus keta*	Chum salmon
	Onc ki 1	*Oncorhynchus kisutch*	Coho salmon
	Onc m 1	*Oncorhynchus mykiss*	Rainbow Trout and Steelhead
	Onc ma 1	*Oncorhynchus masou*	Masu salmon
	Onc n 1	*Oncorhynchus nerka*	Sockeye salmon
	Onc ts 1	*Oncorhynchus tshawytscha*	Winter Salmon
	Ore a 1	*Oreochromis aureus*	Blue tilapia
	Ore mo 1	*Oreochromis mossambicus*	Mozambique tilapia
	Ore ni 1	*Oreochromis nilonica*	Nile tilapia
	Ory la 1	*Oryzias latipes*	Japanese rice fish
	Pag bo 1	*Pagellus bogaraveo*	Blackspot seabream
	Pag ma 1	*Pagrus major*	Red seabream
	Pag pa 1	*Pagrus pagrus*	Couch's sea-bream
	Pam ar 1	*Pampus argenteus*	White pomfret
	Pam c 1	*Pampus chinensis*	Chinese pomfret
	Pan bo 1	*Pangasius bocourti*	Bocourtã¢âs catfish
	Pan hy 1	*Pangasianodon hypophthalmus*	Sutchi catfish
	Par a 1	*Paralichthys albigutta*	Gulf flounder

Table 1. (*Continued*)

Biochemical classification	Allergen	Fish species	Fish common name
Parvalbumin (Calcium-binding protein)	Par as 1	*Parasilurus asotus*	Amur catfish
	Par le 1	*Paralichthys lethostigma*	Southern flounder
	Par ol 1	*Paralichthys olivaceus*	Olive flounder
	Pen ar 1	*Pennahia argentata*	Silver white croaker
	Pla fl 1	*Platichthys flesus*	European flounder
	Ple ar 1	*Plectropomus areolatus*	Squaretail coralgrouper
	Ple le 1	*Plectropomus leopardus*	Coral trout
	Pol vi 1	*Pollachius virens*	Sillock
	Pom sa 1	*Pomatomus saltatrix*	Bluefish
	Rac ca 1	*Rachycentron canadum*	Cobia
	Rho au 1	*Rhomboplites aurorubens*	Vermilion snapper
	Riv ma 1	*Rivulus marmoratus*	Ocellated rivulus
	Sal al 1	*Salvelinus alpinus*	Charr
	Sal f 1	*Salvelinus fontinalis*	Brook trout
	Sal s 1	*Salmo salar*	Atlantic salmon
	San lu 1	*Sander lucioperca*	Pike-perch
	Sar m 1	*Sardinops melanostictus*	Japanese pilchard
	Sar p 1	*Sardina pilchardus*	Pilchard
	Sar sa 1	*Sardinops sagax*	Pacific sardine
	Sci oc 1	*Sciaenops ocellatus*	Red drum
	Sco a 1	*Scomber australasicus*	Japanese mackerel
	Sco ca 1	*Scomberomorus cavalla*	King mackeral
	Sco g 1	*Scomberomorus guttatus*	Indo-Pacific king mackerel
	Sco j 1	*Scomber japonicus*	Pacific chub mackerel
	Sco ma 1	*Scomberomorus maculatus*	Atlantic Spanish mackerel
	Sco s 1	*Scomber scombrus*	Atlantic mackerel
	Seb fa 1	*Sebastes fasciatus*	Acadian redfish
	Seb in 1	*Sebastes inermis*	Darkband rockfish
	Seb m 1	*Sebastes marinus*	Rose fish
	Ser d 1	*Seriola dumerili*	Greater amberjack
	Ser la 1	*Seriola lalandi*	Yellowtail amberjack
	Ser q 1	*Seriola quinqueradiata*	Yellowtail
	Ser ri 1	*Seriola rivoliana*	Almaco jack
	Sil ja 1	*Sillago japonica*	Japanese whiting
	Sin ch 1	*Siniperca chuatsi*	Chinese perch
	Sol so 1	*Solea solea*	Slip
	Spa a 1	*Sparus aurata*	Gilthead seabream
	Sol so 1	*Solea solea*	Slip
	Spa a 1	*Sparus aurata*	Gilthead seabream
	Sph ti 1	*Sphyrna tiburo*	Bonnet hammerhead
	Ste ci 1	*Stephanolepis cirrhifer*	Threadsail filefish
	Sto i 1	*Stolephorus indicus*	Indian anchovy
	Tak ob 1	*Takifugu obscurus*	Obscure pufferfish
	Tak ru 1	*Takifugu rubripes*	Torafugu
	Tet ni 1	*Tetraodon nigroviridis*	Spotted green pufferfish
	The ch 1	*Theragra chalcogramma*	Walleye pollock
	Thu a 1	*Thunnus albacares*	Yellowfin tuna
	Thu al 1	*Thunnus alalunga*	Albacore tuna
	Thu at 1	*Thunnus atlanticus*	Blackfin tuna
	Thu o 1	*Thunnus obesus*	Big eye tuna

Table 1. (*Continued*)

Biochemical classification	Allergen	Fish species	Fish common name
Parvalbumin (Calcium-binding protein)	Thu t 1	*Thunnus thynnus*	Atlantic bluefin tuna
	Thu to 1	*Thunnus tonggol*	Oriental bonito
	Thy at 1	*Thyrsites atun*	Cape Snoek
	Tra ca 1	*Trachinotus carolinus*	Pompano
	Tra j 1	*Trachurus japonicus*	Japanese jack mackerel
	Tra tr 1	*Trachurus trachurus*	Crake-herring
	Tri le 1	*Trichiurus lepturus*	Atlantic cutlassfish
	Uro te 1	*Urophycis tenuis*	Boston ling
	Xip g 1	*Xiphias gladius*	Swordfish
	Zeu fa 1	*Zeus faber*	John dory

ALLERGEN (2014), ALLERGOME (2014), Encyclopedia of Life (2014).

fishes such as cod, salmon and tuna (Kuehn et al., 2013, Kuehn et al., 2014). Both enzymes have biological functions in metabolic glycolysis, being involved in the sugar degradation for the production of energy (Kuehn et al., 2014). The biochemical characterisation of enolases seems to indicate that they are dimeric proteins, while aldolases present oligomeric profiles. Additionally, no post-translational modifications such as glycosylation and/or phosphorylation have been found in fish enolases and aldolases (Kuehn et al., 2013). Preliminary results revealed limited inter-species cross-reactivity, with fish enolases being more cross-reactive than aldolases. Both enzymes have been described as heat-labile since thermal treatments above 90ºC for 1-5 minutes seem to destroy their tri-dimensional structures. Despite eliminating some conformational allergenic epitopes, new linear epitopic regions can be created during food processing, which might contribute to increase their potential allergenicity (Kuehn et al., 2013).

In addition to the parvalbumins, enolases and aldolases, collagen has also been pointed out as an allergen in fish. However, despite the existence of some positive *in vitro* allergy tests, the potential risk of fish gelatine (collagen) for triggering an adverse immunological reaction among fish-allergic individuals remains unclear (Lee & Taylor, 2011; Weber & Paschke, 2010).

Shellfish allergens

The term shellfish is a non-taxonomic designation usually used in the context of seafood consumption. This group comprises crustaceans and molluscs, representing a significant market niche of marine species with commercial interest. Crustaceans areclassified as arthropods and include over 50,000 living species (shrimp, prawns, lobster, crayfish, and barnacles). A large number of crustacean species is consumed either raw or cooked/processed. Molluscs are subdivided into classes of bivalves, gastropods and cephalopods, comprising almost 100,000 different species (mussels,

oysters, abalone and squids). The nutritional value and intrinsic organoleptic characteristics of molluscs make them highly appreciated and consumed foods all over the world, especially in coastal regions (Lopata et al., 2010; Mao et al., 2013; Kamath et al., 2014).

Table 2. List of fish allergens (enolases and aldolases) and corresponding fish species.

Biochemical classification	Allergen	Fish species	Fish common name
Enolase (Glycolytic enzyme)	Ano fi 2	*Anoplopoma fimbria*	Sablefish
	Dan re 2	*Danio rerio*	Zebrafish
	Dic la 2	*Dicentrarchus labrax*	Sea bass
	Epi br 2	*Epinephelus bruneus*	Tawny grouper
	Epi co 2	*Epinephelus coioides*	Orange-spotted grouper
	Gad m 2	*Gadus morhua*	Cod
	Gas ac 2	*Gasterosteus aculeatus*	Three-spined stickleback
	Gil mi 2	*Gillichthys mirabilis*	Long-jawed mudsucker
	Ict fu 2	*Ictalurus furcatus*	Blue catfish
	Ict pu 2	*Ictalurus punctatus*	Channel catfish
	Lat ni 2	*Lates niloticus*	Nile perch
	Meg am 2	*Megalobrama amblycephala*	Wu-Chang fish
	Ore mo 2	*Oreochromis mossambicus*	Mozambique tilapia
	Ore ni 2	*Oreochromis nilonica*	Nile tilapia
	Ory la 2	*Oryzias latipes*	Japanese rice fish
	Sal s 2	*Salmo salar*	Atlantic salmon
	Sal tr 2	*Salmo trutta*	Brown trout
	Tak ru 2	*Takifugu rubripes*	Torafugu
	Tet ni 2	*Tetraodon nigroviridis*	Spotted green pufferfish
	Thu a 2	*Thunnus albacares*	Yellowfin tuna
Aldolase (Aldol reaction enzyme)	Gad m 3	*Gadus morhua*	Cod
	Lat ni 3	*Lates niloticus*	Nile perch
	Ore mo 3	*Oreochromis mossambicus*	Mozambique tilapia
	Sal s 3	*Salmo salar*	Atlantic salmon
	Thu a 3	*Thunnus albacares*	Yellowfin tuna

ALLERGEN (2014), ALLERGOME (2014), Encyclopedia of Life (2014).

Hypersensitivity reactions to seafood are normally immediate (approximately 30 minutes) up to 2 hours after its ingestion. Late-phase immunological responses are also possible to occur, when symptoms are developed up to 8 hours (Wang & Sampson, 2007). Clinical manifestations of shellfish allergy are very similar to fish allergy, resulting, not only from the ingestion of the offending food, but also from manipulating or inhaling the cooking vapours during food processing. Commonly, symptoms begin within minutes and may include oral allergy syndrome (OAS); cutaneous (urticaria, angioedema), gastrointestinal (vomiting, abdominal pains) and/or respiratory symptoms. Although less frequent, severe and systemic responses such as anaphylactic shocks may also occur upon shellfish consumption (Carrapatoso, 2004; Lopata et al., 2010; Yu et al., 2011).

Tropomyosins

In crustaceans and molluscs (Tables 3-5), different proteins are known to trigger observable clinical symptoms, although the majority of them is attributed to a family of proteins designated as tropomyosins, which are closely related alpha helical coiled-coil secondary structure proteins. Tropomyosins are present in muscle and non-muscle cells, and together with actin and myosin, they intervene in the regulatory process of muscle contraction (Breiteneder & Mills, 2005; Leung et al., 2014). So far, a great number of allergenic tropomyosins have already been described among crustaceans (Tables 3), namely in shrimp, crab, lobster and prawn, whose cross-reactivity is related to the high similarity in amino acid sequences among distinct species (Leung et al., 2014). The homology among tropomyosins is so high that most shellfish allergic individuals cross-react upon the ingestion of other crustacean or mollusc species. In fact, it is estimated that 75% of the individuals that present allergy to some type of shellfish is at risk of cross-reacting to a second species due to the high structural similarity among tropomyosins (Emoto, Ishizaki, & Shiomi, 2009; Lee & Taylor, 2011; Tsabouri et al., 2012; Weber and Paschke, 2010). With molecular weights ranging from 34-38 kDa, tropomyosins are considered heat-stable proteins. Upon food processing, these allergens can unfold at a limited extent during heating process, being able to return to their conformational structure after cooling. Chemical processes such as Maillard modification may potentiate tropomyosin allergenicity (Mills et al., 2009).

Arginine kinase

In recent years, other proteins such as arginine kinases have been reported as new allergens in crustaceans and molluscs (Tables 4 and 5). This enzyme is a 40-kDa water-soluble protein present in myosinogen, which is involved in cell metabolism of invertebrates (Chen et al., 2013b). Preliminary reports on identified allergenic arginine kinases from crustaceans seem to indicate that these enzymes are unstable at temperatures between 40-80ºC, being partially unfolded and revealing novel hidden epitopes that may be responsible for increasing IgE-reactivity. Above 80ºC, arginine kinase is thought to fully unfold and subsequently decrease its immunogenicity (Chen et al., 2013b, Giuffrida, Villalta, Mistrello, Amato, & Asero, 2014). Similarly to arginine kinase from crustaceans, the allergenic arginine kinase from molluscs (e.g. Oct f 2) is also unstable above 40ºC, though literature suggests that its allergenic properties are reduced at lower temperatures (>48ºC) (Shen et al., 2012).

Table 3. List of crustacean allergens (tropomyosin) and corresponding species.

Biochemical classification	Allergen	Crustacean species	Crustacean common name
Tropomyosin (Muscle contraction protein)	Ace ja 1	*Acetes japonicus*	Akiami paste shrimp
	Art fr 1	*Artemia franciscana*	Artemia sanfranciscana
	Cal cl 1	*Caligus clemensi*	Sea louse
	Cal fi 1	*Calanus finmarchicus*	Calanus
	Can p 1	*Cancer pagurus*	Edible crab
	Cap e 1	*Caprella equilibra*	Skeleton shrimp
	Cha f 1	*Charybdis feriatus*	Crucifix crab
	Che de 1	*Cherax destructor*	Yabbie
	Chi o 1	*Chionoecetes opilio*	Snow crab
	Cra c 1	*Crangon crangon*	Sand shrimp
	Eri i 1	*Erimacrus isenbeckii*	Hair crab
	Eri s 1	*Eriocheir sinensis*	Chinese freshwater edible crab
	Eup p 1	*Euphausia pacifica*	North Pacific krill
	Eup s 1	*Euphausia superba*	Antarctic krill
	Fen c 1	*Fenneropenaeus chinensis*	Fleshy prawn
	Fen me 1	*Fenneropenaeus merguiensis*	Banana prawn
	Geo de 1	*Geothelphusa dehaani*	Japanese Freshwater Crab
	Hom a 1	*Homarus americanus*	Northern lobster
	Hom g 1	*Homarus gammarus*	Lobster
	Jas ed 1	*Jasus edwardsii*	Cape spiny lobster
	Jas la 1	*Jasus lalandii*	Cape rock lobster
	Lep sa 1	*Lepeophtheirus salmonis*	Salmon louse
	Lim p 1	*Limulus polyphemus*	Horseshoe crab
	Lit se 1	*Litopenaeus setiferus*	White shrimp
	Lit v 1	*Litopenaeus vannamei*	Pacific white shrimp
	Mac r 1	*Macrobrachium rosenbergii*	Giant freshwater prawn
	Mar j 1	*Marsupenaeus japonicus*	Kuruma prawn
	Mel l 1	*Melicertus latisulcatus*	Western king prawn
	Met ba 1	*Metapenaeopsis barbata*	Red rice prawn
	Met e I	*Metapenaeus ensis*	Greasyback shrimp
	Met j 1	*Metapenaeus joyneri*	Shiba shrimp
	Met ja 1	*Metanephrops japonicus*	Japanese lobster
	Met la 1	*Metapenaeopsis lata*	Broad velvet shrimp
	Nep n 1	*Nephrops norvegicus*	Langoustine
	Ora o 1	*Oratosquilla oratoria*	Japanese Mantis Shrimp
	Ova au 1	*Ovalipes australiensis*	Ocean Sand Crab
	Pan b 1	*Pandalus borealis*	Northern shrimp
	Pan e 1	*Pandalus eous*	Alaskan pink shrimp
	Pan h 1	*Panulirus homarus*	Scalloped spiny lobster
	Pan j 1	*Panulirus japonicus*	Japanese crayfish
	Pan s 1	*Panulirus stimpsoni*	Green lobster
	Par c 1	*Paralithodes camtschaticus*	Red king crab
	Par f 1	*Parapenaeus fissurus*	Neptune rose shrimp
	Pen a	*Penaeus aztecus*	Northern brown shrimp
	Pen i 1	*Penaeus indicus*	Indian white shrimp
	Pen m 1	*Penaeus monodon*	Tiger prawn
	Pen se 1	*Penaeus semisulcatus*	Green tiger prawn
	Pen o 1	*Penaeus orientalis*	Fleshy prawn
	Pol po 1	*Pollicipes pollicipes*	Goose neck barnacle
	Por p 1	*Portunus pelagicus*	Blue swimming crab
	Por s 1	*Portunus sanguinolentus*	Three spot swimming crab
	Por tr 1	*Portunus trituberculatus*	Swimming crab

Table 3. (*Continued*)

Biochemical classification	Allergen	Crustacean species	Crustacean common name
Tropomyosin (Muscle contraction protein)	Pro cl 1	*Procambarus clarkii*	Red swamp crawfish
	Ran ra 1	*Ranina ranina*	Spanner crab
	Scy o 1	*Scylla olivacea*	Orange mud crab
	Scy pa 1	*Scylla paramamosain*	Green mud crab
	Scy s 1	*Scylla serrata*	Giant mud crab
	Ser lu 1	*Sergia lucens*	Sakura shrimp
	Sol me 1	*Solenocera melantho*	China Red Shrimp
	Squ ac 1	*Squilla aculeata*	No data
	Squ o 1	*Squilla oratoria*	Japanese Mantis Shrimp
	The or 1	*Thenus orientalis*	Flathead locust lobster
	Tra c 1	*Trachysalambria curvirostris*	Southern rough shrimp

ALLERGEN (2014), ALLERGOME (2014), Encyclopedia of Life (2014).

Sarcoplasmic calcium-binding and troponin C proteins

The sarcoplasmic calcium-binding and the troponin C proteins belong to the superfamily of EF-hand proteins. The sarcoplasmic calcium-binding proteins are present in the invertebrates, being considered the equivalent of the vertebrate parvalbumins that contribute to maintain the intracellular calcium (Gao, Gillen, & Wheatly, 2006). Due to this function, they are designated as calcium buffers, being acidic cytosolic proteins (20–22 kDa) with four potential EF-hand calcium-binding sites, from which two or three are functional (Hermann & Cox, 1995). Troponin C is a calcium sensor/modulator protein that regulates downstream target proteins in a calcium-dependent manner. So far, five troponin C and twenty-six sarcoplasmic calcium-binding proteins have been identified as allergens in crustaceans (Table 4), although their study is still at a very preliminary stage.

Myosin light chain proteins

Myosins are part of a complex that involves other proteins such as actin, tropomyosin and troponin, being all fundamental to muscle contraction. Myosin is composed of two heavy chains and four light chains. Each myosin heavy chain is surrounded by two light chains with 20 kDa each (Ayuso et al., 2008). Until now, six myosin light chains have been classified as allergens among crustaceans (Table 4). Myosin light chain allergens seem to present high IgE-reactivity with sera from shellfish allergic patients, being potentially considered as major allergens similarly to tropomyosins. They are also resistant to heat since they tend to maintain IgE-binding capacity upon processing (100ºC for 5 minutes) (Ayuso et al., 2008).

Table 4. List of crustacean allergens (other non-tropomyosin) and corresponding species.

Biochemical classification	Allergen	Crustacean species	Crustacean common name
Arginine kinase (Phosphagen kinase)	Art fr 2	*Artemia franciscana*	Artemia sanfranciscana
	Cal s 2	*Callinectes sapidus*	Blue crab
	Can mg 2	*Cancer magister*	Dungeness crab
	Car ma 2	*Carcinus maenas*	Green crab
	Cha f 2	*Charybdis feriatus*	Crucifix crab
	Chi o 2	*Chionoecetes opilio*	Snow crab
	Cra c 2	*Crangon crangon*	Sand shrimp
	Eri s 2	*Eriocheir sinensis*	Chinese freshwater edible crab
	Fen c 2	*Fenneropenaeus chinensis*	Fleshy prawn
	Fen me 2	*Fenneropenaeus merguiensis*	Banana prawn
	Hom g 2	*Homarus gammarus*	Lobster
	Lim p 2	*Limulus polyphemus*	Horseshoe crab
	Lit v 2	*Litopenaeus vannamei*	Pacific white shrimp
	Mac r 2	*Macrobrachium rosenbergii*	Giant freshwater prawn
	Mar j 2	*Marsupenaeus japonicus*	Kuruma prawn
	Met e 2	*Metapenaeus ensis*	Greasyback shrimp
	Met j 2	*Metapenaeus joyneri*	Shiba shrimp
	Met t 2	*Metanephrops thomsoni*	Red-banded lobster
	Pac ma 2	*Pachygrapsus marmoratus*	Marbled Rock Crab
	Pan b 2	*Pandalus borealis*	Northern shrimp
	Pen m 2	*Penaeus monodon*	Tiger prawn
	Por p 2	*Portunus pelagicus*	Blue swimming crab
	Por tr 2	*Portunus trituberculatus*	Swimming crab
	Pro cl 2	*Procambarus clarkii*	Red swamp crawfish
	Scy o 2	*Scylla olivacea*	Orange mud crab
	Scy pa 2	*Scylla paramamosain*	Green mud crab
	Scy s 2	*Scylla serrata*	Giant mud crab
Myosin light chain (ATP-dependent motor protein)	Art fr 5	*Artemia franciscana*	Artemia sanfranciscana
	Cra c 5	*Crangon crangon*	Sand shrimp
	Hom a 3	*Homarus americanus*	Northern lobster
	Lit v 3	*Litopenaeus vannamei*	Pacific white shrimp
	Pan b 3	*Pandalus borealis*	Northern shrimp
	Pen m 3	*Penaeus monodon*	Tiger prawn
Sarcoplasmic calcium-binding protein	Cha f 4	*Charybdis feriatus*	Crucifix crab
	Chi o 4	*Chionoecetes opilio*	Snow crab
	Cra c 4	*Crangon crangon*	Sand shrimp
	Eri s 4	*Eriocheir sinensis*	Chinese freshwater edible crab
	Far be 4	*Farfantepenaeus brevirostris*	Pink shrimp
	Far no 4	*Farfantepenaeus notialis*	Southern pink shrimp
	Fen me 4	*Fenneropenaeus merguiensis*	Banana prawn
	Hom a 4	*Homarus americanus*	Northern lobster
	Lit v 4	*Litopenaeus vannamei*	Pacific white shrimp
	Mac r 4	*Macrobrachium rosenbergii*	Giant freshwater prawn
	Mar j 4	*Marsupenaeus japonicus*	Kuruma prawn
	Mel l 4	*Melicertus latisulcatus*	Western king prawn
	Met e 4	*Metapenaeus ensis*	Greasyback shrimp
	Ora o 4	*Oratosquilla oratoria*	Japanese Mantis Shrimp
	Pan b 4	*Pandalus borealis*	Northern shrimp
	Pan e 4	*Pandalus eous*	Alaskan pink shrimp
	Par lo 4	*Parapenaeus longirostris*	Deep-water rose shrimp

Table 4. (*Continued*)

Biochemical classification	Allergen	Crustacean species	Crustacean common name
Sarcoplasmic calcium-binding protein	Pen a 4 Pen i 4 Pen m 4 Pen se 4 Pon l 4 Por p 4 Pro cl 4 Scy pa 4 Sol ag 4	*Penaeus aztecus* *Penaeus indicus* *Penaeus monodon* *Penaeus semisulcatus* *Pontastacus leptodactylus* *Portunus pelagicus* *Procambarus clarkii* *Scylla paramamosain* *Solenocera agassizii*	Northern brown shrimp Indian white shrimp Tiger prawn Green tiger prawn Galician crayfish Blue swimming crab Red swamp crawfish Green mud crab Kolobri shrimp
Troponin C (Calcium-binding protein)	Chi o 6 Cra c 6 Hom a 6 Pan b 6 Pen m 6	*Chionoecetes opilio* *Crangon crangon* *Homarus americanus* *Pandalus borealis* *Penaeus monodon*	Snow crab Sand shrimp Northern lobster Northern shrimp Tiger prawn
Triosephosphate isomerase	Arc s 8	*Archaeopotamobius sibiriensis*	Shrimp
(Glycolytic pathway enzyme)	Cra c 8	*Crangon crangon*	Sand shrimp

ALLERGEN (2014), ALLERGOME (2014), Encyclopedia of Life (2014).

Detection of seafood allergens

The presence of trace amounts of undeclared allergenic ingredients in seafood products poses a significant health risk to sensitised/allergic individuals of suffering abnormal immune episodes as consequence of accidental exposure (van Hengel, 2011, Lee & Taylor, 2011). Considering the severity and the frequency of cases involving seafood allergy, in addition with the low levels (3-32 mg of allergenic proteins from fish or shellfish, respectively) responsible for eliciting observable symptoms upon ingestion, the sensitised/allergic individuals are obliged to completely avoid products susceptible of containing seafood as ingredient (Reese et al., 2005; Untersmayr et al., 2007). Therefore, in order to protect these patients against the presence of hidden allergens as a result of cross-contamination or mislabelling, reliable and highly sensitive analytical methods are vital to verify labelling compliance and to help the industrial management of fish and shellfish allergens (Costa, Mafra, Carrapatoso, & Oliveira, 2012; Herrero, Vieites, & Espiñeira, 2014; Lee & Taylor, 2011). It is generally accepted that the ideal limit of detection (LOD) for allergens in food products should range between 1 and 100 mg/kg (Poms, Klein, & Anklam, 2004).

So far, several molecular tools targeting either proteins or DNA have been published for food allergen analysis, namely enzyme-linked immunosorbent assays (ELISA), lateral

Table 5. List of mollusc allergens and corresponding species.

Biochemical classification	Allergen	Mollusc species	Mollusc common name
Tropomyosin (Muscle contraction protein)	Ana br 1	*Anadara broughtonii*	Blood clam
	Arg i 1	*Argopecten irradians*	Bay scallop
	Bab ja 1	*Babylonia japonica*	Japanese Babylon
	Bal r 1	*Balanus rostratus*	Acorn Barnacle
	Buc mi 1	*Buccinum middendorffi*	Middendorf's whelk
	Cap m 1	*Capitulum mitella*	Japanese goose barnacle
	Chl n 1	*Chlamys nipponensis*	Japanese scallop
	Cra g 1	*Crassostrea gigas*	Pacific cupped oyster
	Ent d 1	*Enteroctopus dofleini*	North Pacific giant octopus
	Eup sc 1	*Euprymna scolopes*	Bobtail squid
	Ful mu 1	*Fulvia mutica*	Japanese Cockle
	Hal a 1	*Haliotis asinina*	Ass's abalone
	Hal d 1	*Haliotis diversicolor*	Many-colored abalone
	Hal di 1	*Haliotis discus*	Disk abalone
	Hal r 1	*Haliotis rubra*	Blacklip abalone
	Hal ru 1	*Haliotis rufescens*	Red abalone
	Hem t 1	*Hemifusus ternatanus*	Whelk
	Lol b 1	*Loligo bleekeri*	Inshore squid
	Lut p 1	*Lutraria philippinarum*	No data
	Mer ly 1	*Meretrix lyrata*	Lyrate asiatic hard clam
	Mer mt 1	*Meretrix meretrix*	Asiatic hard clam
	Mim n 1	*Mimachlamys nobilis*	Noble scallop
	Myt e 1	*Mytilus edulis*	Mussel
	Myt g 1	*Mytilus galloprovincialis*	Mediterranean mussel
	Nas ob 1	*Nassarius obsoletus*	Eastern mudsnail
	Nep po 1	*Neptunea polycostata*	Wrinkled Neptune
	Oct f 1	*Octopus fangsiao*	Gold-spot octopus
	Oct v 1	*Octopus vulgaris*	Common octopus
	Omm b 1	*Ommastrephes bartramii*	Webbed squid
	Pat y 1	*Patinopecten yessoensis*	Ezo giant scallop
	Pec fu 1	*Pecten fumatus*	Tasmanian Scallop
	Per v 1	*Perna viridis*	Asian green mussel
	Pin a 1	*Pinna atropurpurea*	Pen Shell
	Sep e 1	*Sepia esculenta*	Golden cuttlefish
	Sep l 1	*Sepioteuthis lessoniana*	Bigfin reef squid
	Sep m 1	*Sepia madokai*	Madokai's cuttlefish
	Sep of 1	*Sepia officinalis*	Margade
	Sin c 1	*Sinonovacula constricta*	Chinese razor clam
	Sol st 1	*Solen strictus*	Gould's razor shell
	Spi sa 1	*Spisula sachalinensis*	Surf-clam
	Teg gr 1	*Tegillarca granosa*	Blood Cockle
	Tod p 1	*Todarodes pacificus*	Japanese flying squid
	Tre ke 1	*Tresus keenae*	Horse clam
	Tur c 1	*Turbo cornutus*	Horned Turban
	Uro ed 1	*Uroteuthis edulis*	Swordtip squid
	Ven ph 1	*Venerupis philippinarum*	Japanese Cockle
Arginine kinase (Phosphagen kinase)	Oct f 2	*Octopus fangsiao*	Gold-Spot Octopus
	Oct v 2	*Octopus vulgaris*	Common octopus
	Sep in 2	*Sepiella inermis*	Cuttlefish
	Sep l 2	*Sepioteuthis lessoniana*	Bigfin reef squid

ALLERGEN (2014), ALLERGOME (2014), Encyclopedia of Life (2014).

flow devices (LFD), liquid chromatography (LC) coupled with mass spectrometry (MS), polymerase chain reaction (PCR), real-time PCR, microarrays or biosensors. However, the potential number of available methods within the sphere of seafood allergen detection is still limited (Pascoal et al., 2011; Rencova et al., 2013).

Protein-based methods

Protein-based methods, such as ELISA or LFD, test for the presence of allergens or specific food marker proteins by using specific mono- or polyclonal antibodies that are usually raised in animals (Schubert-Ullrich et al., 2009). These immunochemical assays are particularly useful in food industry since they offer high specificity and sensitivity (antigen/antibody interaction) for fast allergen detection, without requiring extensive sample preparation, expensive equipment or experienced personnel. In the specific case of fish allergens, the immunochemical assays targeting parvalbumins are among the most widely used methods for their detection in foods. As previously mentioned, important variations in parvalbumin levels can be found in some of the most consumed fish species (carp, cod, hake, salmon and tuna) as highlighted in several reports (ALLERGEN, 2014; ALLERGOME, 2014; Perez-Gordo et al.; 2011; Griesmeier et al., 2010; Tsabouri et al., 2012; Van Do et al., 2005). Due to the high prevalence of shellfish allergy, namely to crustacean species, the development of rapid methods for the detection of trace amounts of shellfish have become imperative. Despite some lack of available techniques to detect mollusc allergens, immunochemical methods have been the first choice for qualitative and quantitative analysis of shellfish allergens, such as the tropomyosins Cra c 1, Hom a 1, Pen m 1, Tod p 1 and Scy pa 1 (ALLERGEN, 2014, ALLERGOME, 2014; Emoto et al., 2009; Kamath et al., 2014; Yu et al., 2011). In spite of the simplicity and utility of the immunoassays, they also present some major drawbacks and, consequently, the results from their application should be carefully analysed. Both ELISA and LFD are prone to cross-reactivity phenomena and they are also highly affected by conformational changes in proteins upon food processing, which can lead to the possibility of false negative or positive results (Costa et al., 2014; Lee & Taylor, 2011).

Mass spectrometry methodologies have also found application in allergenomics, allowing protein identification, characterisation and quantification. The main advantages of these methods rely on their high accuracy, sensitivity, specificity and reproducibility, though the high cost of analysis represents an important drawback (Picariello, Mamone, Addeo, & Ferranti, 2011). Additionally, MS methods can overcome the problems of cross-reactivity phenomena often linked to immunoassays, allowing the unequivocal identification of the tested allergens/peptides. The analysis of allergens by MS

methodology can be performed by one of two approaches, either targeting intact proteins (analyte and reference standards) or peptides obtained from protein digestion using proteolytic enzymes (Picariello et al., 2011). This methodology was already applied for the direct detection of parvalbumin peptide biomarkers using a set of 16 species of fish that were analysed by LC-MS/MS approach (Carrera, Cañas, & Gallardo, 2012).

Currently, there are several ELISA kits commercially available that enable the detection of fish and shellfish in foods. For example, the "Fish Protein ELISA kit" (Elution Technologies, Vermont, USA) described as capable of directly detecting parvalbumin in different food matrices has a limit of quantification (LOQ) of 1 mg/kg. The "AgraQuant® Fish" (Romer Labs Division Holding GmbH, Austria) refers a LOQ of 4-100 mg/kg, though without clear information regarding the target class of fish proteins. The "Crustacean Protein ELISA kit" (Elution Technologies, Vermont, USA) and "AgraQuant® ELISA Crustacea" (Romer Labs Division Holding GmbH, Austria) claim LOQ of 2 mg/kg and 20-400 mg/kg, respectively. "Mollusk ELISA kit" (Elution Technologies, Vermont, USA) allows the detection of mollusk proteins from 1 mg/kg.

The limits of detection (LOD) of the most relevant reports from literature on protein-based techniques applied to seafood detection range from 0.046 mg/kg to 18.7 mg/kg (Carrera, Cañas, & Gallardo; Faeste & Plassen, 2008; Kuehn et al., 2010; Weber, Steinhart, & Paschke, 2009).

DNA-based methods

Technologies based on DNA analysis, namely PCR, present some advantages over the methodologies targeting proteins. DNA molecules are more stable and resistant to thermal treatments, pH alterations and partial hydrolysis than proteins, being less affected by processes that normally alter the integrity of proteins. In fact, DNA-based methods are notably helpful on the analysis of highly processed foodstuffs (Eischeid, Kim, & Kasko, 2013; Herrero et al., 2014; Hildebrandt & Garber, 2010; Lee, Nordlee, Koppelman, Baumert, & Taylor, 2012). PCR, preceded by DNA extraction, provides a sensitive tool for the specific detection of genomic sequences encoding allergenic proteins or species-specific markers. However, the amplification of a DNA sequence by PCR does not necessarily indicate the presence of an allergenic protein in the food matrix, being therefore considered an indirect method of allergen detection (Costa et al., 2014; Mafra, Ferreira, & Oliveira, 2008). There are already two commercial kits available for fish and mollusc DNA detection: "SureFood® ALLERGEN ID Fish" and "SureFood® ALLERGEN ID Molluscs" (R-Biopharm AG Darmstadt, Germany) (LOD ≤ 0.4 mg/kg).To our knowledge, the majority of studies in the literature reporting the application of PCR-based

methodologies for fish analysis have mainly been focused on species authentication purposes. However, there are some studies regarding the use of PCR and real-time PCR for the specific detection of parvalbumins, as the cases of the Atlantic and Pacific herrings (Clu h 1 and Clu pa 1) or the Pacific mackerel (Sco j 1) (Lee & Taylor, 2011; Rencova et al., 2013). Sun, Liang, Gao, Lin, and Deng (2009) developed a real-time PCR assay for the specific detection of parvalbumin gene in fish, allowing a sensitivity of 5 pg. Concerning the detection and quantification of shellfish allergens, despite the preponderance of immunochemical methods, real-time PCR has only been applied to tropomyosin analysis in blue crab (Cal s 2) and tiger prawn (Pen m 1) (LOD, 0.1-1 mg/kg) (Eischeid et al., 2013).

Several factors should be considered when choosing a technique for the detection of allergens in foods, such as the availability of expensive equipment and experienced personnel, the time consumed per analysis, the cost of analysis, among others. For instance, MS platforms present reliable results, but require expensive equipment, specialised personnel and are not very suited for routine analyses. In the case of ELISA, the time per sample analysis is relatively short, the need for specialised personnel and cost are low/moderate, but the reliability of results can be compromised by cross-reactivity phenomena and food processing. DNA-based methods using quantitative technology such as real-time PCR require specialised personnel and equipment, moderate time of analysis, present high specificity and sensitivity, but can only provide indirect information regarding the detection of allergens in foods. In summary, all the methods present major advantages and also some drawbacks, so the choice of the method should be critically analysed according to the food matrix, type of processing and the available equipment (Costa, Ansari, Mafra, Oliveira, & Baumgartner, 2014). Whenever possible, the combination between protein- and DNA-based methods is highly recommended for confirmation and identification purposes.

Final Remarks

Recent data suggest that there is a clear increase in the number of reported cases of allergy to seafood proteins, along with the growing consumption of fish and shellfish at a global scale The major allergens in fish and shellfish are parvalbumins and tropomyosins, respectively, with several occurrences as IgE-reactive proteins already reported, which have been currently characterised and publically made available in allergen databases. The establishment and development of novel methodologies for the detection and quantification of allergens in fish and shellfish are also of crucial importance for a better assessment and management of these proteins in processed foods. Protein-based assays

such as ELISA, LFD and MS-platforms are the most well-known used techniques for the detection/quantification of parvalbumins and tropomyosins in seafood products. Additionally, the use of DNA-based methods has already allowed the development of some commercial kits that detect different allergens by means of real-time PCR technology. Despite the available protein- and DNA-based methods, novel techniques allowing the detection and quantification of fish and shellfish in foods at trace amounts (with limits of detection as low as 1 mg/kg) are still much needed. Forthcoming advances should become available upon the development of testing/reference materials, the establishment of official methods for seafood allergen detection, as well as novel information regarding allergen threshold levels.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Fundação para a Ciência e a Tecnologia (FCT) [grant number. PEst-C/EQB/LA0006/2013; Telmo J.R. Fernandes and J. Costa are grateful to PhD [grant number SFRH/BD/93711/2013]; post-doctoral [grant number SFRH/BPD/102404/2014] from FCT financed by POPH-QREN (subsidised by Fundo Social Europeu [FSE] and Ministério da Ciência, Tecnologogia e Ensino Superior [MCTES]).

References

ALLERGEN, 2014, official site for the systematic allergen nomenclature. Retrieved from: http://www.allergen.org

ALLERGOME, 2014, allergome database, the platform for allergen knowledge, Latina, Italy. Retrieved from: http://www.allergome.org

Arif S. H., Jabeen, M., &Hasnain, A. U. (2007). Biochemical Characterization and thermostable capacity of parvalbumins: the major fish-food allergens. *Journal of Food Biochemistry*, 31, 121-137. doi: 10.1111/j.1745-4514.2007.00104.x

Ayuso, R., Grishina, G., Bardina, L., Carrillo, T., Blanco, C., Ibáñez, M. D., Sampson, H. A., & Beyer, K. (2008). Myosin light chain is a novel shrimp allergen, Lit v 3. *Journal of Allergy and Clinical Immunology*, 122, 795-802. doi: 10.1016/j.jaci.2008.07.023

Breiteneder, H., & Mills, E. N. C. (2005). Molecular properties of food allergens. *Journal of Allergy and Clinical Immunology*, *115*, 14-23. doi: 10.1016/j.jaci.2004.10.022

Bugajska-Schretter, A., Elfman, L., Fuchs, T., Kaplotis, S., Rumpold, H., Valenta, R., & Spitzauer, S. (1998). Parvalbumin, a cross-reactive fish allergen, contains IgE-binding epitopes sensitive to periodate treatment and Ca^{2+} depletion. *Journal of Allergy and Clinical Immunology*, *101*, 67-74.

Bugajska-Schretter, A., Grote, M., Vangelista, L., Valent, P., Sperr, W. R., Rumpold, H., Pastore, A., Reichelt, R., Valenta, R., & Spitzauer, S. (2000). Purification, biochemical, and immunological characterisation of a major food allergen: different immunoglobulin E recognition of the apo- and calcium-bound forms of carp parvalbumin. *Gut*, *46*, 661-669. doi:10.1136/gut.46.5.661

Burks, A. W., Tang, M., Sicherer, S., Muraro, A., Eigenmann, P. A., Ebisawa, M., Fiocchi, A., Chiang, W., Beyer, K., Wood, R., Hourihane, J., Jones, S. M., Lack, G., & Sampson, H. A. (2012). ICON: Food allergy. *Journal of Allergy and Clinical Immunology*, *129*, 906-920. doi: 10.1016/j.jaci.2012.02.001.

Burney, P., Summers, C., Chinn, S., Hooper, R., Van Ree, R., & Lidholm, J. (2010). Prevalence and distribution of sensitization to foods in the European Community Respiratory Health Survey: a EuroPrevall analysis. *Allergy*, *65*, 1182-1188. doi: 10.1111/j.1398-9995.2010.02346.x

Carrapatoso, I. (2004). Grupos de alimentos com maior reactividade cruzada: artigo de revisão. *Revista Portuguesa de Imunoalergologia*, *10*, 103-113. Retrieved from: http://hdl.handle.net/10400.4/1231

Carrera, M., Cañas, B., Gallardo, J. M. (2012). Rapid direct detection of the major fish allergen, parvalbumin, by selected MS/MS ion monitoring mass spectrometry. *Journal of Proteomics*, *75*, 3211-3220. doi: 10.1016/j.jprot.2012.03.030

Chen, H. L., Cao, M. J., Cai, Q. F., Su, W. J., Mao, H. Y., & Liu, G. M. (2013a). Purification and characterisation of sarcoplasmic calcium-binding protein, a novel allergen of red swamp crayfish (*Procambarus clarkii*). *Food Chemistry*, *139*, 213-223. doi: 10.1016/j.foodchem.2013.01.119

Chen, H.-L., Mao, H.-Y., Cao, M.-J., Cai, Q.-F., Su, W.-J., Zhang, Y.-X., & Liu, G.-M. (2013b). Purification, physicochemical and immunological characterization of arginine kinase, an allergen of crayfish (*Procambarus clarkii*). *Food and Chemical Toxicology*, *62*, 475-484. doi:10.1016/j.fct.2013.09.014

CODEX STAN 1-1985. Amended in 1991, 1999, 2001, 2003, 2005, 2008 and 2010 regarding the general standard for the labelling of pre-packaged foods. FAO/WHO Standards, Off Codex Stand 2010; Rome, Italy. Retrieved from: http://www.codexalimentarius.org/standards/list-of-standards

Costa, J., Ansari, P., Mafra, I., Oliveira, M. B. P. P., & Baumgartner, S. (2014). Assessing hazelnut allergens by protein- and DNA-based approaches: LC-MS/MS, ELISA and

real-time PCR. *Analytical and Bioanalytical Chemistry*, *406*, 2581-2590. doi: 10.1007/s00216-014-7679-x

Costa, J., Carrapatoso, I., Oliveira, M. B. P. P., & Mafra, I. (2014). Walnut allergens: Molecular characterization, detection and clinical relevance. *Clinical and Experimental Allergy, 44*, 319-341. doi: 10.1111/cea.12267

Costa, J., Mafra, I., Carrapatoso, I., & Oliveira, M. B. P. P. (2012). Almond allergens: molecular characterization, detection and clinical relevance. *Journal of Agricultural and Food Chemistry*, *60*, 1337-1349. doi: 10.1021/jf2044923

Directive 2000/13/EC of 20 March 2000 relating to the labelling, presentation and advertising of foodstuffs. *Official Journal of the European Communities*, *L109*, 29-42.

Directive 2007/68/EC of 27 November 2007 amending Annex IIIa to Directive 2000/13/EC regarding certain food ingredients. *Official Journal of the European Union*; *L310*, 11-14.

Eischeid, A. C., Kim, B. H., & Kasko, S. M. (2013). Two quantitative real-time PCR assays for the detection of penaeid shrimp and blue crab, crustacean shellfish allergens. *Journal of Agricultural and Food Chemistry*, *61*, 5669-5674. doi: 10.1021/jf3031524

Emoto, A., Ishizaki, S., & Shiomi, K. (2009). Tropomyosins in gastropods and bivalves: Identification as major allergens and amino acid sequence features. *Food Chemistry*, *114*, 634-641. doi: 10.1016/j.foodchem.2008.09.100

Encyclopedia of Life, official site for information and pictures of all species known to science. Retrieved from: http://www.eol.org

Giuffrida, M. G., Villalta, D., Mistrello, G., Amato, S., & Asero, R. (2014). Shrimp allergy beyond tropomyosin in Italy: clinical relevance of arginine kinase, sarcoplasmic calcium binding protein and hemocyanin. *European Annals of Allergy and Clinical Immunology*, *46*, 172-177.

Griesmeier, U., Vázquez-Cortés, S., Bublin, M., Radauer, C., Ma, Y., Briza, P., Fernández-Rivas, M., & Breiteneder, H. (2010). Expression levels of parvalbumins determine allergenicity of fish species. *Allergy: European Journal of Allergy and Clinical Immunology*, *65*, 191-198. doi: 10.1111/j.1398-9995.2009.02162.x

Faeste C. K., & Plassen C (2008). Quantitative sandwich ELISA for the determination of fish in foods. *Journal of Immunology Methodologies*, 329, 45-55. doi:10.1016/j.jim.2007.09.007

Hermann, A., & Cox, J. A. (1995). Sarcoplasmic calcium-binding protein. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *111*, 337-345. doi:10.1016/0305-0491(94)00218-J

Herrero, B., Vieites, J. M., & Espiñeira, M. (2014). Development of an in-house fast real-time PCR method for detection of fish allergen in foods and comparison with a commercial kit. *Food Chemistry*, *151*, 415-420. doi: 10.1016/j.foodchem.2013.11.042

Hildebrandt, S., & Garber, E. A. E. (2010). Effects of processing on detection and quantification of the parvalbumin gene in Atlantic salmon (*Salmo salar*). *Food Chemistry*, *119*, 75-80. doi: 10.1016/j.foodchem.2009.05.074

Jenkins, J. A., Breiteneder, H., & Mills, E. N. C. (2007). Evolutionary distance from human homologs reflects allergenicity of animal food proteins. *Journal of Allergy and Clinical Immunology, 120*, 1399-1405. doi: 10.1016/j.jaci.2007.08.019

Kamath, S. D., Rahman, A. M. A., Voskamp, A., Komoda, T., Rolland, J. M., O'Hehir, R. E., & Lopata, A. L. (2014). Effect of heat processing on antibody reactivity to allergen variants and fragments of black tiger prawn: A comprehensive allergenomic approach. *Molecular Nutrition & Food Research*, *58*, 1144-1155. doi: 10.1002/mnfr.201300584

Kuehn, A., Hilger, C., Lehners-Weber, C., Codreanu-Morel, F., Morisset, M., Metz-Favre, C., Pauli, G., de Blay, F., Revets, D., Muller, C. P., Vogel, L., Vieths, S., & Hentges, F. (2013). Identification of enolases and aldolases as important fish allergens in cod, salmon and tuna: Component resolved diagnosis using parvalbumin and the new allergens. *Clinical and Experimental Allergy*, *43*, 811-822. doi: 10.1111/cea.12117

Kuehn, A., Scheuermann, T., Hilger, C., & Hentges, F. (2010). Important variations in parvalbumin content in common fish species: A factor possibly contributing to variable allergenicity. *International Archives of Allergy and Immunology, 153*, 359-366. doi: 10.1159/000316346

Kuehn, A., Swoboda, I., Arumugam, K., Hilger, C., & Hentges, F. (2014). Fish allergens at a glance: variable allergenicity of parvalbumins, the major fish allergens. *Frontiers in Immunology*, *5*, 179. doi: 10.3389/fimmu.2014.00179

Larsen, R., Eilertsen, K.-E., & Elvevoll, E. O (2011). Health benefits of marine foods and ingredients. *Biotechnology Advances*, *29*, 508–518. doi: 10.1016/j.biotechadv.2011.05.017

Lieberman, J. A., & Sicherer, S. H. (2011). Diagnosis of food allergy: Epicutaneous skin tests, in vitro tests, and oral food challenge. *Current Allergy Asthma Reports*, *11*, 58-64. doi: 10.1007/s11882-010-0149-4

Lee, P-W., Nordlee, J. A., Koppelman, S. J., Baumert, J. L., & Taylor, S. L. (2012). Measuring parvalbumin levels in fish muscle tissue: Relevance of muscle locations and storage conditions. *Food Chemistry*, *135*, 502-507. doi: 10.1016/j.foodchem.2012.05.030

Lee, P-W, & Taylor, S. L.. (2011). Fish, Crustaceans, and Mollusks. In A. J. Hengel, and L. M. L. Nollet (Eds). *Food Allergens – Analysis Instrumentation and Methods*. Taylor and Francis Group, LLC, United States of America, 117-205.

Leung, N. Y. H., Wai, C. Y. Y., Shu, S. A., Wang, J., Kenny, T. P., Chu, K. H., & Leung, P. S. C. (2014). Current immunological and molecular biological perspectives on seafood

allergy: a comprehensive review. *Clinical Reviews in Allergy & Immunology*, *46*, 180-197. doi: 10.1007/s12016-012-8336-9

Lopata, A. L., O'Hehir, R. E., & Lehrer, S. B. (2010). Shellfish allergy. *Clinical and Experimental Allergy*, *40*, 850-858. doi: 10.1111/j.1365-2222.2010.03513.x

Madsen, C. B., Hattersley, S., Allen, K. J., Beyer, K., Chan, C. H., Godefroy, S. B., Hodgson, R., Mills, E. N. C., Muñoz-Furlong, A., Schnadt, S., Ward, R., Wickman, M., & Crevel, R. (2012). Can we define a tolerable level of risk in food allergy? Report from a EuroPrevall/UK Food Standards Agency workshop. *Clinical & Experimental Allergy*, *42*, 30-37. doi: 10.1111/j.1365-2222.2011.03868.x

Mafra, I., Ferreira, I. M. P. L. V. O., & Oliveira, M. B. P. P. (2008). Food authentication by PCR-based methods. *European Food Research and Technology*, *227*, 649-665. doi: 10.1007/s00217-007-0782-x

Mao, H. Y., Cao, M. J., Maleki, S. J., Cai, Q. F., Su, W. J., Yang, Y., & Liu, G. M. (2013). Structural characterization and IgE epitope analysis of arginine kinase from *Scylla paramamosain*. *Molecular Immunology*, *56*, 463-470, 2013. doi: 10.1016/j.molimm.2013.04.016.

Mills, E. N. C., Sancho, A. I., Rigby, N. M., Jenkins, J. A., & Mackie, A. R. (2009). Impact of food processing on the structural and allergenic properties of food allergens. *Molecular Nutrition & Food Research*, *53*, 963-969. doi: 10.1002/mnfr.200800236

Mozaffarian, D., Bryson, C. L., Lemaitre, R. N., Siscovick, D. S., & Burke, G. L. (2005). Fish intake and risk of incident heart failure. *Journal of the American College of Cardiology*, *45*, 2015-2021. doi: 10.1016/j.jacc.2005.03.038

Muraro, A., Werfel, T., Hoffmann-Sommergruber, K., Roberts, G., Beyer, K., Bindslev-Jensen, C., Cardona, V., Dubois, A., duToit, G., Eigenmann, P., Fernandez Rivas, M., Halken, S., Hickstein, L., Høst, A., Knol, E., Lack, G., Marchisotto, M. J., Niggemann, B., Nwaru, B. I., Papadopoulos, N. G., Poulsen, L. K., Santos, A. F., Skypala, I., Schoepfer, A., Van Ree, R., Venter, C., Worm, M., Vlieg–Boerstra, B., Panesar, S., de Silva, D., Soares-Weiser, K., Sheikh, A., Ballmer-Weber, B. K., Nilsson, C., de Jong, N. W., Akdis, C. A., the, E. F. A., & Anaphylaxis Guidelines, G. (2014). EAACI Food Allergy and Anaphylaxis Guidelines: diagnosis and management of food allergy. *Allergy*, *69*, 1008-1025. doi:10.1111/all.12429

Nakamura, A., Watanabe, K., Ojima, T., Ahn, D.-H., & Saeki, H. (2005). Effect of maillard reaction on allergenicity of scallop tropomyosin. *Journal of Agricultural and Food Chemistry*, *53*, 7559-7564. doi: 10.1021/jf0502045

Pascoal, A., Barros-Velázquez, J., Ortea, I., Cepeda,, A., Gallardo, J. M., & Calo-Mata, P. (2011). Molecular identification of the black tiger shrimp (*Penaeus monodon*), the white leg shrimp (*Litopenaeus vannamei*) and the Indian white shrimp (*Fenneropenaeus*

indicus) by PCR targeted to the 16S rRNA mtDNA. *Food Chemistry*, *125*, 1457-1461. doi; 10.1016/j.foodchem.2010.10.053

Perez-Gordo, M., Cuesta-Herranz, J., Maroto, A. S., Cases, B., Ibáñez, M. D., Vivanco, F., & Pastor-Vargas, C. (2011). Identification of sole parvalbumin as a major allergen: Study of cross-reactivity between parvalbumins in a Spanish fish-allergic population. *Clinical and Experimental Allergy*, *41*, 750-758.

Perez-Gordo, M., Lin, J., Bardina, L., Pastor-Vargas, C., Cases, B., Vivanco, F., Cuesta-Herranz, J., & Sampson, H. A. (2012). Epitope mapping of Atlantic salmon major allergen by peptide microarray immunoassay. *International Archives of Allergy and Immunology*, *157*, 31-40. doi: 10.1159/000324677

Picariello, G., Mamone, G., Addeo, F., & Ferranti, P. (2011). The frontiers of mass spectrometry-based techniques in food allergenomics. *Journal of Chromatography A*, *1218*, 7386-7398. doi: 10.1016/j.chroma.2011.06.033

Poms, R. E., Klein, C. L., & Anklam, E. (2004). Methods for allergen analysis in food: a review. *Food Additives and Contaminants*, *21*, 1-31. doi: 10.1080/02652030310001620423

Radauer, C., Bublin, M., Wagner, S., Mari, A., & Breiteneder, H. (2008). Allergens are distributed into few protein families and possess a restricted number of biochemical functions. *Journal of Allergy and Clinical Immunology*, *121*, 847-852. doi: 10.1016/j.jaci.2008.01.025

Reese, G., Viebranz, J., Leong-Kee, S. M., Plante, M., Lauer, I., Randow, S., Moncin, M. S.-M., Ayuso, R., Lehrer, S. B., & Vieths, S. (2005). Reduced allergenic potency of VR9-1, a mutant of the major shrimp allergen Pen a 1 (tropomyosin). *The Journal of Immunology*, *175*, 8354-8364. doi: 10.4049/jimmunol.175.12.8354

Regulation (EU) No 1169/2011 of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No 1924/2006 and (EC) No 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004. *Official Journal of the European Union*, *L304*, 18-63.

Rencova, E., Kostelnikova, D., & Tremlova, B. (2013). Detection of allergenic parvalbumin of Atlantic and Pacific herrings in fish products by PCR. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment*, *30*, 1679-1683. doi: 10.1080/19440049.2013.817024

Schubert-Ullrich, P., Rudolf, J., Ansari, P., Galler, B., Führer, M., Molinelli, A., & Baumgartner, S. (2009). Commercialized rapid immunoanalytical tests for

determination of allergenic food proteins: an overview. *Analytical and Bioanalytical Chemistry*, *395*, 69-81. doi: 10.1007/s00216-009-2715-y

Sharp, M. F., & Lopata, A. L. (2014). Fish Allergy: In Review. *Clinical Reviews in Allergy and Immunology*, *46*, 258-271. doi: 10.1007/s12016-013-8363-1

Shen, H.-W., Cao, M.-J., Cai, Q.-F., Ruan, M.-M., Mao, H.-Y., Su, W.-J., & Liu, G.-M. (2012). Purification, cloning, and immunological characterization of arginine kinase, a novel allergen of *Octopus fangsiao*. *Journal of Agricultural and Food Chemistry*, *60*, 2190-2199. doi: 10.1021/jf203779w

Sicherer, S. H., Muñoz-Furlong, A., & Sampson, H. A. (2004). Prevalence of seafood allergy in the United States determined by a random telephone survey. *Journal of Allergy and Clinical Immunology*, *114*, 159-165. doi: 10.1016/j.jaci.2004.04.018

Sirot, V., Oseredczuk, M., Bemrah-Aouachria, N., Volatier, J-L., & Leblanc, J-C (2008). Lipid and fatty acid composition of fish and seafood consumed in France: CALIPSO study. *Journal of Food Composition and Analysis*, *21*, 8-16. doi: 10.1016/j.jfca.2007.05.006

Sletten, G., Van Do, T., Lindvik, H., Egaas, E., & Florvaag, E. (2010). Effects of industrial processing on the immunogenicity of commonly ingested fish species. *International Archives of Allergy and Immunology*, *151*, 223-236. doi: 10.1159/000242360

Sun M., Liang, C., Gao, H., Lin, C., Deng, M. (2009). Detection of parvalbumin, a common fish allergen gene in food, by real-time polymerase chain reaction. *Journal of AOAC International*, 92, 234-240.

Thomas, K., Herouet-Guicheney, C., Ladics, G., Bannon, G., Cockburn, A., Crevel, R., Fitzpatrick, J., Mills, C., Privalle, L., & Vieths, S. (2007). Evaluating the effect of food processing on the potential human allergenicity of novel proteins: International workshop report. *Food and Chemical Toxicology*, *45*, 1116-1122. doi: 10.1016/j.fct.2006.12.016

Tsabouri, S., Triga, M., Makris, M., Kalogeromitros, D., Church, M. K., & Priftis, K. N. (2012). Fish and shellfish allergy in children: Review of a persistent food allergy. *Pediatric Allergy and Immunology, 23*, 608-615. doi: 10.1111/j.1399-3038.2012.01275.x

Untersmayr, E., Vestergaard, H., Malling, H.-J., Jensen, L. B., Platzer, M. H., Boltz-Nitulescu, G., Scheiner, O., Skov, P. S., Jensen-Jarolim, E., & Poulsen, L. K. (2007). Incomplete digestion of codfish represents a risk factor for anaphylaxis in patients with allergy. *Journal of Allergy and Clinical Immunology*, *119*, 711-717. doi: 10.1016/j.jaci.2006.10.039

Van Do, T., Elsayed, S., Florvaag, E., Hordvik, I., & Endresen, C. (2005). Allergy to fish parvalbumins: Studies on the cross-reactivity of allergens from 9 commonly consumed

fish. *Journal of Allergy and Clinical Immunology*, *116*, 1314-1320. doi: 10.1016/j.jaci.2005.07.033

van Hengel, A. J. (2011). Introduction. In L. M. L. Nollet, and A. J. van Hengel, (Eds). *Food allergens: analysis instrumentation and methods*, CRC Press, Taylor & Francis Group, Boca Raton, 1–11.

Wang, J., & Sampson, H. A. (2007). Food anaphylaxis. *Clinical & Experimental Allergy*, *37*, 651-660. doi: 10.1111/j.1365-2222.2007.02682.x

Wang, J., Zhou, C., Tian, J., Sun, N., Lu, J., Wang, C., & Che, H. (2014). The analysis of specific allergenicity of food allergens families. *Food and Agricultural Immunology*, *25*, 375-386. doi: 10.1080/09540105.2013.805733

Weber P., Steinhart H., Paschke, A. (2009). Competitive indirect ELISA for the determination of parvalbumins from various fish species in food grade fishgelatins and isinglass with PARV-19 anti-parvalbumin antibodies. *Journal of Agricultural and Food Chemistry*, 57, 11328-11334. doi: 10.1021/jf902470e

Weber, P., & Paschke, A. (2010). Fish Allergens. In L Jedrychowski, and H. J. Wichers (Eds). *Chemical and Biological Properties of Food Allergens*. Taylor and Francis Group, LLC, United States of America, 223-232.

Yu, H-L., Cao, M-J., Cai, Q-F., Weng, W-Y., Su, W-J., & Liu, G-M. (2011). Effects of different processing methods on digestibility of *Scylla paramamosain* allergen (tropomyosin). *Food and Chemical Toxicology, 49*, 791-798. doi: 10.1016/j.fct.2010.11.046

CRITICAL REVIEWS IN FOOD SCIENCE AND NUTRITION
2017, VOL. 57, NO. 15, 3281–3296
https://doi.org/10.1080/10408398.2015.1113157

Taylor & Francis
Taylor & Francis Group

Advances on the molecular characterisation, clinical relevance and detection methods of Gadiform parvalbumin allergens

Telmo J.R. Fernandes, Joana Costa, Isabel Carrapatoso, M. Beatriz P.P. Oliveira, Isabel Mafra*

[a] REQUIMTE-LAQV, Faculdade de Farmácia, Universidade do Porto, Rua de Jorge Viterbo Ferreira, 228, 4050-313 Porto, Portugal

[b] Serviço de Imunoalergologia, Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal

*Corresponding author: Tel: +351 220428640. Fax: +351 226093390. E-mail: isabel.mafra@ff.up.pt

Abstract

Gadiform order includes several fish families, from which Gadidae and Merlucciidae are part of, comprising the most commercially important and highly appreciated fish species, such as cod, pollock, haddock and hake. Parvalbumins, classified as calcium-binding proteins, are considered the main components involved in the majority of fish allergies. Nine and thirteen parvalbumins were identified in different fish species from Gadidae and Merlucciidae families, respectively. This review intends to describe their molecular characterisation and the clinical relevance, as well as the prevalence of fish allergy. Additionally, the main protein- and DNA-based methods to detect fish allergens are fully reviewed owing to their importance in the safeguard of sensitised/allergic individuals.

Keywords: food allergen; fish allergy; Gadidae; Merlucciidae; molecular characterisation; prevalence; parvalbumin detection; protein-based methods; DNA-based methods.

Introduction

The taxonomic order of Gadiformes is ubiquitous in every ocean of the world and includes several fish families of great commercial interest. In this order, the most commercially important and highly appreciated species are included in two main families: the Gadidae and the Merlucciidae. The Gadidae family (Gadoids) contains a well-known group of fishes such as Atlantic cod (*Gadus morhua*), pollock (*Pollachius pollachius*) and haddock (*Melanogrammus aeglefinus*), while Merlucciidae family encompasses species such as hakes (*Merluccius merluccius*) (Cohen et al., 1990; Di Finizio et al., 2007). Approximately 18% of the global marine catch regards Gadiform species, corresponding to almost 6.5 million tons in 2011 (FAO, 2014). Like other commonly consumed fish families, the species belonging to Gadidae and Merlucciidae are traditional dietary components in several worldwide populations, being used in many of the existing manufactured fish products (fresh or frozen fillets, smoked, salted, among others) (Calo-Mata et al., 2003).

Beyond the incontestable nutritional value of their flesh as a source of highly assimilated proteins, fish plays a vital role in human diet. Several health benefits have been related to fish consumption, such as prevention of cardiovascular diseases and cancer or glycaemic control, mainly associated with the ingestion of omega-3 polyunsaturated fatty acids (PUFA) (Larsen et al., 2011; Mozaffarian, et al., 2005; Sirot et al., 2008). Along with the increased awareness of the benefits of fish intake and consequent consumption, the number of fish allergies has also increased (Van Do et al., 2005a).

Fish is one of the eight groups responsible for almost 90% of food allergies reported at global scale (CODEX, 2010). For the sensitised/allergic patients, the prevention of an allergic reaction relies mostly on the total avoidance of the offending food. Therefore, accurate and reliable food labelling systems have become imperative to improve consumer's protection and life quality (Costa et al., 2014; Fæste et al., 2011; Monaci, and Visconti, 2010; Prado et al., 2015; Rencova et al., 2013; Taylor, and Baumert, 2015). According to the recent European Union (EU) regulations, food producers are obligated to declare and highlight from the rest of the list of ingredients, the presence of fourteen groups of potentially allergenic foods, which include fish, crustaceans, molluscs, celery, mustard, sesame seed, gluten, tree nuts, peanuts, milk, eggs, soybeans, lupine and sulphites (Directive 2006/142/EC, Directive 2007/68/EC, EFSA, 2014; Regulation (EU) No 1169/2011).

In the present work, a concise and updated overview on fish allergens, focusing on Gadiform order will be described. Issues such as the prevalence of fish allergy, the

molecular characterisation of Gadiform parvalbumins and their clinical relevance will be especially addressed. The analytical techniques used for the detection and quantification of fish allergens will be also particularly reviewed.

Prevalence of fish allergy

Despite the general perception of the increasing frequency of food allergies, estimates of their actual prevalence and incidence are still uncertain (Nwaru et al., 2014). Most of the diagnoses is based on self-reporting symptoms, specific immunoglobulin E (IgE) sensitisation assays or skin prick tests (SPT) to common food allergens, so the real frequency of food allergies could be probably overestimated (Burks et al., 2012; Nwaru et al., 2014). Although objective assessments such as SPT are considered reliable evaluations (Zuidmeer et al., 2012), relatively few epidemiological studies have been performed using open food challenges (OFC) or double-blind placebo-controlled food challenge (DBPCFC) as the gold standard diagnosis in defining food allergy (Nwaru et al., 2014).

Presently, overall data suggest that 1-10% of the general population suffers from food allergies (Chafen, et al., 2010), with a higher frequency among children than in adults (Sicherer, and Sampson, 2009; Sicherer, and Sampson, 2010). During the first years of life, the prevalence of food allergy reaches 6-8%, mainly due to the consumption of milk, egg, peanut, fish, and shellfish (Berin, and Sampson, 2013; Wang, and Sampson, 2011). Children often acquire natural tolerance to some foods such as egg, milk, wheat or soybean during childhood (mainly in the first decade), in opposition to fish, shellfish, peanut and tree nuts allergies that are more likely to be life-persisting (Priftis et al., 2008; Wang, and Sampson, 2011).

Fish allergy is estimated to affect 0.1-0.4% of the general population (Codex, 2010; Kuehn et al., 2013), although these numbers can vary among different countries/regions. In a population-based study performed by random call survey in Canada, the prevalence of perceived and further confirmed fish allergy was 0.51% and 0.10%, respectively. This discrepancy suggests that among Canadian population, the prevalence of fish allergy is probably overestimated, highlighting the importance of using confirmatory tests for food allergy diagnosis (Ben-Shoshan et al., 2010). In the USA, another random call survey estimates that 0.4% of the population suffers from fish allergy and 0.2% from both shellfish and fish allergies. In the same study, the prevalence of fish allergy was found to be higher among adults and among women, respectively, in terms of age and gender (Sicherer et al., 2004). In a different report from the USA targeting adult population, the incidence of self-reported and clinically diagnosed fish allergy was 0.7% and 0.6%,

respectively (Vierk et al., 2007). Contrarily to the previous report (Sicherer et al., 2004), no significant differences were found in the prevalence of fish allergy between gender or race/ethnic groups (Vierk et al., 2007). In Europe, the latest reported overall incidence of fish allergy was approximately 0.2%, with Germany presenting the highest prevalence (0.9%), followed by Spain with 0.5% (Burney et al., 2010; Burney et al., 2014). Although the percentage in Europe was estimated on the basis of a study that included eleven countries and other partners (Australia and USA), only six of them reported data regarding fish allergy. Similarly to other food allergies, there is a proportional relationship between the prevalence of fish allergy and the total consumption of fish in each country. In fact, fish is one of the most common food allergens in the European coastal countries like Norway, Finland, Portugal or Spain (Kamath et al., 2013; Perez-Gordo et al., 2011; Sharp, and Lopata, 2013; Tsabouri et al., 2012). Considering that countries such as Portugal or Finland were not included in the study described by Burney et al. (2010), the prevalence of fish allergy in Europe is most likely underestimated. The scarce data from Asia and South Africa suggest a high incidence of fish allergy among the allergic population. In Singapore, fish allergy presented an estimated prevalence of 4.1% (Thong et al., 2007).

It is also important to highlight that although food intolerance does not result from a specific immune response because it is mainly due to digestion, absorption or metabolic disorders, the clinical symptoms are extremely similar to the ones presented by allergic individuals. This fact may also contribute to frequent errors on the estimation of the prevalence of the perceived and confirmed food allergies (Prado et al., 2015).

Molecular characterisation of Gadiform parvalbumins

Over the last years, the number of identified allergen sequences has been exponentially increasing, which emphasises the need for a systematic classification. Together with the recent advances of bioinformatics, researchers are now more capable of establishing evolutionary and structural relationships among allergens from distinct origins (Radauer, and Breiteneder, 2007; Radauer et al., 2008).

As a result, allergen platforms such as the ALLERGOME, the Resource Program Allergen, the Official List of Allergens issued by the International Union of Immunological Societies (IUIS) Allergen Nomenclature Sub-committee and the InFormAll databases that report molecular, biochemical and clinical data about allergenic proteins have become excellent tools for allergen classification. Particularly, ALLERGOME database has been designed to provide relevant information on allergenic molecules selected from international renowned scientific journals, reporting significant IgE-mediated reactions. Even though most of the allergenic molecules are officially named and recognised by the IUIS Allergen Nomenclature Sub-committee, ALLERGOME database expands data

available from the scientific literature, following careful immunological and allergological criteria based on structural relationships with known allergens and IgE-binding capacity. Thereby, a wide number of allergenic molecule entries that are not listed in the IUIS allergen nomenclature website are included in the ALLERGOME database (Mari et al., 2009).

Although the number of food allergens has been continuously rising, literature seems to suggest that they are limited to some protein groups, corresponding to about 5% of all structural protein families (Radauer et al., 2008). So far, plant food allergens have been considered a diversified group of proteins that belong to few specific families, evidencing their distinct biological and biochemical roles (Breiteneder, and Radauer, 2004; Breiteneder, and Mills, 2005a; Breiteneder, and Mills, 2005b). In the case of animal food allergens, the number of families is even smaller. Contrasting with plant food or pollen allergens, practically all animal food allergens have homologs in the human proteome, which is thought to affect the recognition of such proteins by the human immune system (Jenkins et al., 2007).

Parvalbumins belong to the second largest family of animal food allergens, the calcium binding proteins (Jenkins et al., 2007; Breiteneder, and Radauer, 2005a). They are vertebrate-specific, primarily cytosolic, small in length (composed by 106-113 aa), with a globular tri-dimensional structure and acidic isoelectric points (3.9-5.5), being expressed in fast-twitch muscles, specific neurons, certain kidney and endocrine gland cells, Corti's cells and others (Permyakov et al., 2008). Biochemically, parvalbumins are characterised by the high proportions of glutamic and aspartic acids and phenylalanine, with little or no histidine, proline, cysteine, methionine, tyrosine, or tryptophan (Heizmann, 1984). Owing to their unbalanced amino acid composition (high content in phenylalanine), parvalbumins exhibit a characteristic ultraviolet absorption spectrum with low absorbance at 280 nm (Arif, 2009). They are well-known for possessing one or more EF-hand motifs that correspond to a 12-residue loop flanked on both sides by a 12-residue alpha-helical domain (Grabarek, 2006). Parvalbumins contain, specifically, three EF-hand motifs, in which two domains bind a calcium ion each and a third silent domain forms a cap covering the hydrophobic surface of the two functional domains, stabilising the two EF-hand motifs (Grabarek, 2006; Ikura, 1996). They are abundant proteins (up to 5 mg per g of fresh weight) in the white muscle of many fish species, with estimated biological functions related to the relaxation of muscle fibres through the binding of free intracellular calcium. The binding of calcium ion to the protein is thought to be determinant for the conformational integrity of the IgE epitopes, affecting their allergenic capacity. Parvalbumins have remarkable resistance to high temperatures, denaturing agents and

proteolytic activity (Arif et al., 2007; Bugajska-Schretter et al., 1998; Bugajska-Schretter et al., 2000; Capony, and Pechère, 1973).

Parvalbumins can be divided into two lineages – alpha-parvalbumins and beta-parvalbumins – that present very similar overall folding. These two groups can be distinguished by their isoelectric points (pI) (alpha: pI ≥5; beta: pI ≤4.5), sequence characteristics, affinities for calcium and magnesium, cell-type specific expression and physiologic functions (Wopfner et al., 2007). The alpha-parvalbumins are abundantly present in the muscle of fish and amphibians, but in general they are not allergenic. The beta-parvalbumins seem to present allergenic properties, which are evidenced by the great number of beta-parvalbumins identified as allergens in fish species (Jenkins et al., 2007). Beta-parvalbumins are considered important allergenic proteins with well conserved sequences, exhibiting at least 53% of sequence identity among homologues from distantly related or even unrelated fish species (Radauer et al., 2008). Beta-parvalbumins of fish represent one of the largest animal families of food allergens, reaching the second place just behind tropomyosin family in crustaceans and molluscs (Jenkins et al., 2007). Cross-reactions between beta-parvalbumins from cod and other species are frequently the main cause of observed polysensitisation to multiple fish species in allergic individuals (Griesmeier et al., 2010).

So far, the Gadidae and the Merlucciidae families of fish allergens are considered the most important in terms of clinical relevance. Few members of other families of Gadiform order, namely Phycidae and Muraenolepididae (Cohen et al., 1990), have been reported as allergenic fishes, though with lower consumption rate and, subsequently, with lower exposure. Among them, the Gadidae is, by far, the most representative family of allergens. The Merlucciidae is the family of cod-like fish with several of its members having commercial interest, which include 5 different genera of fish. The most well-known species of this family belong to the *Macruronus* and *Merluccius* genera (Lloris et al., 2005).

At the time, only Gad c 1 and Gad m 1 parvalbumins have been included in the WHO/IUIS official nomenclature list of allergens. Due to the high number of reported immunoreactivity of other fish parvalbumins, several of them have been proposed as allergens in the ALLERGOME database, which are expected that after careful evaluation by the WHO/IUIS Executive Committee are also included in the official list of allergens, justifying their inclusion in the present review.

Gadidae parvalbumins

Gad c 1 (former Allergen M)

This protein corresponds to the parvalbumin allergen present in Baltic cod (*Gadus callarias*) and it was formerly designated as allergen M. Its identification and characterisation was first reported almost four decades ago (Elsayedet al., 1974; Elsayed, and Bennich, 1975) and since then, Gad c 1 has been the source of extensive study.

Gad c 1 has a primary structure composed of 113 aa with an estimated molecular mass of 12.1 kDa and a pI of 4.75 (Table 1). Its sequence includes a monomer of glucose and residues of tyrosine, tryptophan and arginine (one of each), with arginine being thought to play an essential role in the conformational structure of the mature protein (Elsayed, et al., 1974). The trypic cleavage of the arginyl peptide bond of allergen M allowed the identification of two allergenically active fragments (TM1 and TM2), evidencing that their linear sequences are most directly related to cod allergy than conformational structure (Elsayed et al., 1974; Elsayed, and Bennich, 1975). Gad c 1 was the first allergenic protein to undergo epitope mapping, thus peptides have been generated by tryptic hydrolysis and analysed for immunoreactivity (Elsayed et al., 1974; Elsayed et al., 1981; Elsayed, and Apold, 1983). Based on Gad c 1.0101 isoform, four hexapeptides, namely DEDGFY, IADEDK, KGILSN and SNADIK were *in silico* analysed by Minkiewicz et al. (2012), revealing a wide distribution of Gad c 1 epitopic hexapeptides in the universal proteome. In the reported study, the four hexapeptides were found, not only across plant and animal kingdoms, but also in several microorganisms, which might explain potential cross-reactivity with unrelated food species (Minkiewicz et al., 2012).

In general, parvalbumins are thought to be well conserved among different fish species (Radauer et al., 2008), indicating a high potential for cross-reactivity with distinct parvalbumins in fish-allergic patients. The analysis of purified extracts of fish allergens by SDS-PAGE and immunoblotting, using sera of fish-allergic individuals, allowed Van Do et al. (2005a) to compare the IgE-binding proteins of nine fish species. In the reported study, an intensive coloured band at 12 kDa could be observed in cod, salmon, pollock and wolfish, confirming the presence of parvalbumins in these fish species. The use of rabbit polyclonal IgG raised against recombinant Gad c 1 by means of immunoblotting assay, also enabled the identification of strong reactivity of cod with salmon, pollock, herring, wolfish and flounder, while reactivity for other fish species (tuna, halibut and mackerel) was considered weak, confirming the immunoblotting results with IgE from sera of fish-allergic patients (Van Do et al., 2005a). Primary protein sequence of Gad c 1 showed high identity indices with the parvalbumins Sal s 1 (68%) from salmon and The c 1 (62%) from

Table 1. General data on fish parvalbumin isoallergens from the Gadiform order.

Allergen	Species	Common name	MW (kDa)	Isoallergens	Protein (UniProt)
Gad c 1	*Gadus callaris*	Baltic cod	12.1 (113 aa)	Gad c 1.01	P02622
Gad m 1	*Gadus morhua*	Atlantic cod	11.45 (109 aa)	Gad m 1.01	Q90YL0
			11.47 (109 aa)		A5I873
			11.55 (109 aa)	Gad m 1.02	Q90YK9
			11.54 (109 aa)		A5I874
Gad ma 1	*Gadus macrocephalus*	Pacific cod	No data available	No data available	No data available
Bor sa 1	*Boreogadus saida*	Arctic cod	11.57 (109 aa)	No data available	C0LEL4
Mic po 1	*Micromesistius poutassou*	Blue whiting	12	No data available	No data available
Mel ea 1	*Melanogrammus aeglefinus*	Haddock	12	No data available	No data available
Pol vi 1	*Pollachius virens*	Pollock	10	No data available	No data available
The ch 1	*Theragra chalcogramma*	Alaska pollock	11.50 (109 aa)	Parvalbumin beta 1 (predicted classification The ch 1.01)	Q90YK8
			11.51 (109 aa)	Parvalbumin beta 2 (predicted classification The ch 1.02)	Q90YK7
Mer me 1	*Merlangius merlangus*	Whiting	11.33 (108 aa)	Parvalbumin beta	P02621
Mer ap 1	*Merluccius australis polylepis*	Patagonian hake	11.27 (108 aa)	Parvalbumin beta 1 (predicted classification Mer ap 1.01)	P86749
			11.32 (108 aa)	Parvalbumin beta 2 (predicted classification Mer ap 1.02)	P86750
			8.22 (76 aa)	Parvalbumin beta 3 (predicted classification Mer ap 1.03)	P86751
Mer au 1	*Merluccius australis australis*	Southern hake	11.32 (108 aa)	Parvalbumin beta 1 (predicted classification Mer au 1.01)	P86745
			11.53 (108 aa)	Parvalbumin beta 2 (predicted classification Mer au 1.01)	P86747
			11.30 (108 aa)	Parvalbumin beta 3 (predicted classification Mer au 1.02)	P86748
			11.33 (69 aa)	Parvalbumin beta 4 (predicted classification Mer au 1.03)	P86746
Mer bi 1	*Merluccius bilinearis*	Silver hake	11.27 (108 aa)	Parvalbumin beta 1 (predicted classification Mer bi 1.01)	P56503
			11.24 (108 aa)	Parvalbumin beta 3 (predicted classification Mer bi 1.01)	P86753
			11.32 (108 aa)	Parvalbumin beta 2 (predicted classification Mer bi 1.02)	P86752
			11.27 (94 aa)	Parvalbumin beta 4 (predicted classification Mer bi 1.03)	P86754
Mer ca 1	*Merluccius capensis*	Stockfish	11.30 (108 aa)	Parvalbumin beta 1 (predicted classification Mer ca 1.01)	P86756
			11.38 (108 aa)	Parvalbumin beta 2 (predicted classification Mer ca 1.02)	P86757
			11.39 (58 aa)	Parvalbumin beta 3	P86755
Mer ga 1	*Merluccius gayi*	English hake	11.33 (108 aa)	Parvalbumin beta 1 (predicted classification Mer ga 1.01)	P86761
			11.35 (108 aa)	Parvalbumin beta 2 (predicted classification Mer ga 1.02)	P86759
			11.20 (75 aa)	Parvalbumin beta 3	P86758
			11.33 (91 aa)	Parvalbumin beta 4	P86760
Mer hu 1	*Merluccius hubbsi*	Argentine hake	11.33 (108 aa)	Parvalbumin beta 1 (predicted classification Mer hu 1.01)	P86764
			11.35 (108 aa)	Parvalbumin beta 2 (predicted classification Mer hu 1.02)	P86762
			11.35 (86 aa)	Parvalbumin beta 3	P86763

Table 1. (continuation)

Allergen	Species	Common name	MW (kDa)	Isoallergens	Protein (UniProt)
Mer mr 1	*Merluccius merluccius*	European hake	11.33 (108 aa)	Parvalbumin beta 1 (predicted classification Mer mr 1.01)	P02620
			11.30 (108 aa)	Parvalbumin beta 2(predicted classification Mer mr 1.01)	P86765
			11.38 (108 aa)	Parvalbumin beta 3(predicted classification Mer mr 1.02)	P86766
			11.39 (69 aa)	Parvalbumin beta 4(predicted classification Mer mr 1.03)	P86767
Mer pa 1	*Merluccius paradoxus*	Deepwater hake	11.36 (108 aa)	Parvalbumin beta 1 (predicted classification Mer pa 1.01)	P86768
			11.37 (108 aa)	Parvalbumin beta 2(predicted classification Mer pa 1.02)	P86769
			11.35 (95 aa)	Parvalbumin beta 3 (predicted classification Mer pa 1.03)	P86770
Mer po 1	*Merluccius polli*	Benguela hake	11.33 (108 aa)	Parvalbumin beta 1(predicted classification Mer po 1.01)	P86773
			11.35 (108 aa)	Parvalbumin beta 2 (predicted classification Mer po 1.02)	P86771
			11.35 (69 aa)	Parvalbumin beta 3 (predicted classification Mer po 1.03)	P86772
Mer pr 1	*Merluccius productus*	North Pacific hake	11.37 (108 aa)	Parvalbumin beta 1 (predicted classification Mer pr 1.01)	P86774
			11.31 (108 aa)	Parvalbumin beta 2 (predicted classification Mer pr 1.02)	P86775
			11.35 (88 aa)	Parvalbumin beta 3 (predicted classification Mer pr 1.03)	P86776
Mer se 1	*Merluccius senegalensis*	Black hake	11.33 (108 aa)	Parvalbumin beta 1 (predicted classification Mer se 1.01)	P86778
			11.38 (108 aa)	Parvalbumin beta 2 (predicted classification Mer se 1.02)	P86779
			11.37 (58 aa)	Parvalbumin beta 3(predicted classification Mer se 1.03)	P86777
Mac ma 1	*Macruronus magellanicus*	Patagonian grenadier	11.25 (98 aa)	Parvalbumin beta 1 (predicted classification Mac ma 1.01)	P86739
			11.34 (108 aa)	Parvalbumin beta 2 (predicted classification Mac ma 1.02)	P86741
			11.35 (74 aa)	Parvalbumin beta 3 (predicted classification Mac ma 1.03)	P86740
Mac n 1	*Macruronus novaezelandiae*	Blue grenadier	11.25 (98 aa)	Parvalbumin beta 1 (predicted classification Mac n 1.01)	P86739
			11.35 (108 aa)	Parvalbumin beta 2 (predicted classification Mac n 1.02)	P86741
			11.35 (83 aa)	Parvalbumin beta 3 (predicted classification Mac n 1.03)	P86740
Mur mi 1	*Muraenolepis microps*	Smalleye Moray Cod	No data available	No data available	No data available
Uro te 1	*Urophycis tenuis*	Boston ling	No data available	No data available	No data available

pollock, supporting the strong cross-reactivity among cod and these species. This cross-reactivity is of clinical relevance since individuals with a positive DBPCFC to cod also reacted to other fish, such as herring, plaice and mackerel.

Gad m 1

From *Gadus morhua* species, the Gad m 1 is another allergenic protein of major importance. So far, two isoallergens – Gad m 1.01 and Gad m 1.02 – and four respective isoforms – Gad m 1.0101/Gad m 1.0102 and Gad m 1.0201/Gad m 1.0202 – have been identified in Atlantic cod (ALLERGEN, 2014). The four isoforms of Gad m 1 present a primary structure of 109 aa (NCBI, 2014; UniProt, 2014) containing the residue alanine N-acetylated, although their encoded nucleotide sequences vary between 546 base pair (bp) and 797 bp (Table 1) (NCBI, 2014). The amino acid sequence and 3-D representation of Gad m 1 molecule is presented in Figure 1.

Gad m 1 was first reported by Das Dores et al. (2002) as an allergen from Atlantic cod, presenting a sequence identity of 62.3% with Gad c 1 (*Gadus callarias*) and 75% with Sal s 1 (*Salmo salar*) allergens. This finding suggested that, in cod and salmon species, the Gad m 1 and the Sal s 1 are probably encoded by the same gene, while Gad m 1 and Gad c 1 are most likely to be coded by two distinct genes (Das Dores et al., 2002).

Van Do et al., (2003) reported the identification and characterisation of two distinct parvalbumin transcripts (T1 and T2) in cod by means of reverse transcriptase-polymerase chain reaction (PCR) using the RNA of cod muscle. Both parvalbumin mRNA sequences

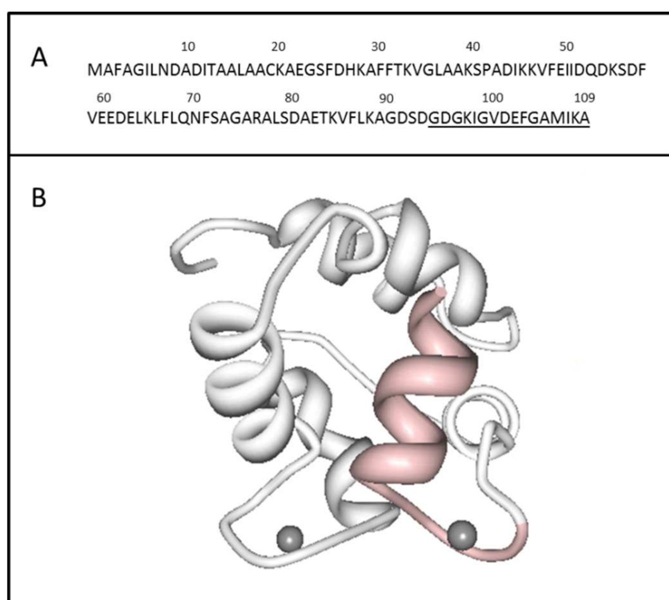


Fig. 1. (A) Aminoacid sequence of Gadus morhua (Gad m 1) parvalbumin – 2MBX (RCSB-PDB, http://www.rcsb.org). The main epitopes are underlined. (B) Ribbon representation of Gad m 1 (main epitopes highlighted) and calcium cations represented by spheres.

encoded similar polypeptide chains of 108 aa (including the Methionine start codon), and show unambiguously strong relationship to the beta-parvalbumins. The amino acid sequence of T1 had identity indices with beta-type parvalbumins of teleost fish spanning from 73.1% (Northern pike) to 92.6% (whiting). In relation to T2, the identity indices were from 71.3% (whiting) to 86.0% (European hake) and with alpha-type parvalbumins, identity ranged from 53.9% (human) to 59.8% (Northern pike). The identification of two distinct cDNA in Atlantic cod seems to suggest that isotypic variants are normally present in fish. The T1 isotype evidenced high similarity with beta-parvalbumin of whiting (92.6%), whereas the T2 isotype presented less identity (71.3%) with the same sequence (Van Do et al., 2003).

More recently, Ma et al. (2008) cloned and expressed two full-length nucleotide sequences (Gad m 1.01 and Gad m 1.02). The coding regions of each of the two parvalbumin cDNA encompass approximately 330 bp for proteins with 109 aa, both with a theoretical pI of 4.58. The deduced sequences of the two cod beta-parvalbumin isoforms were 71% identical, justifying their classification as isoallergens of Gad m 1. Both isoallergens also evidenced 80-81% sequence identity with Cyp c 1.01 from *Cyprinus carpio* (common carp). After purifying native Gad m 1 from muscle protein extract of cod, nGad m 1 and rGad m 1.02 were evaluated for their immunoreactivity to sera obtained from fish-allergic patients (n=26). All sera were reactive to both nGad m 1 and rGad m 1.02, and 25 out of 26 tested sera were also reactive to Cyp c 1.01, confirming that cod and carp parvalbumins share at least some B-cell epitopes. In the same study, supported by mass spectrometry (MS), authors speculate that the majority of nGad m 1 molecules undergo post-translational modification by acetylation. Therefore, sequences of Gad m 1 isoforms that were included in the available databases (ALLERGEN, 2014; NCBI, 2014; UniProt, 2014) present the acetylated N-terminal alanine (Ma et al., 2008).

Gad m 1 was the first beta-parvalbumin studied by NMR spectroscopy (Ma et al., 2008), although no structure is yet available in the Protein Data Bank (PDB). The protein revealed a secondary structure composed of six alpha-helices (7-10 aa each) and two small beta-sheets with residues 56-57 and 97-99. These regions are thought to correspond to the putative calcium-binding sites that are present in parvalbumin family members as an EF-hand motif (Moraes et al., 2013). Since Gad m 1 is a calcium-binding protein, in the absence of calcium, its structure becomes less ordered, losing integrity at a pressure of 200 MPa. According to infrared results, above 50ºC the polarity in protein core increases, indicating a loss of the tertiary structure that leaded to the appearance of a molten globule conformation (Somkuti et al., 2012). In the same study, the authors established that the complete unfolding of Gad m 1 was possible combining high temperatures and pressures (e.g. 40ºC at 1.14 GPa or 50ºC at 890 MPa). Besides the

normally observable native, unfolded and aggregated states, the authors also reported a molten globule and a partially unfolded state possessing different calcium-binding capacities. However, the immunoreactivity of those proteins seemed unaffected by any of the tested treatments since IgE-binding capacity was neither reduced nor enhanced in the presence of these proteins (Somkuti et al., 2012).

Mic po 1

Mic po 1 allergen was identified in the *Micromesistius poutassou* species, frequently known as blue whiting. Piñeiro et al. (1998) have described the identification of several proteins from *Micromesistius poutassou* species ranging from high molecular weight (~67 kDa) down to approximately 14 kDa and with high isoelectric points (pI>5.5), suggesting their classification as parvalbumins. However, the set of proteins with molecular mass below 14 kDa and pI in the interval of 3.9-4.6 were defined as potential allergenic parvalbumins in blue whiting. More recently, Sletten et al. (2010) reported the identification of a 12 kDa parvalbumin in blue whiting using two antibodies: a monoclonal anti-frog parvalbumin antibody PARV-19 and a polyclonal rabbit anti-cod parvalbumin antibody K991. In the referred study, data suggested that blue whiting resembles cod and haddock, with a prominent 12 kDa parvalbumin band and other additional bands between 17 and 62 kDa.

Mel ae 1

The *Melanogrammus aeglefinus* (haddock) has the parvalbumin Mel ae 1 identified as an allergen. Like other parvalbumins, it is present in the white muscle tissue of haddock and its route of sensitisation is ingestion (ALLERGOME, 2014). Mel ae 1 was identified by means of ELISA with the monoclonal PARV-19 mouse anti-frog parvalbumin antibody. In the same research, parvalbumins with molecular masses ranging from 8-11 kDa were detected in different species, namely cod, pollock, hake, haddock, salmon, sturgeon, and tilapia (Weber et al., 2009). Using polyclonal antibodies raised against Atlantic cod, Koppelman et al. (2012) were also able to identify a haddock parvalbumin that was 100% reactive with the developed antibodies, suggesting high homology between parvalbumins from cod and haddock. Supporting this fact, Sun et al. (2009) were able to amplify parvalbumin sequences from twenty-eight fish species using universal real-time PCR.

Sera from a total of nineteen fish-allergic patients were incubated with haddock parvalbumin, presenting 100% immunoreactivity with fresh haddock. With respect to processed haddock, the immunoreactivity seems to be increased in most of the sera tested with parvalbumin from the smoked fish (Sletten et al., 2010). The identification of

this parvalbumin in haddock (*Melanogrammus aeglefinus*) was also confirmed using a monoclonal anti-frog parvalbumin antibody PARV-19 and a polyclonal rabbit anti-cod parvalbumin antibody K991 (Sletten et al., 2010).

The ch 1

This protein has been identified as the allergenic parvalbumin in *Theragra chalcogramma* species, which is commonly known as Alaska pollock or Walleye pollock (ALLERGOME, 2014). The ch 1 presents a primary sequence of 109 aa, with an estimated molecular mass of 11.5 kDa that is encoded by the *Theragra chalcogramma* clone 2 parvalbumin mRNA with 679 bp (Table 1). It was first described by Van Do et al. (2005b) as the allergenic parvalbumin from pollock, exhibiting similar patterns of reactivity with polyclonal rabbit anti-cod and anti-pollock antibodies. The immunoreactivity of The ch 1 was also confirmed by immunoblotting using a panel of six sera of fish-allergic patients, which exhibited 100% reactivity to The ch 1 (Alaska pollock) as well as to Gad c 1 (Baltic cod).

Two isoforms (P1 and P2) were obtained by means of a reverse transcriptase-PCR, evidencing properties from the beta-lineage of parvalbumins. Both cDNA encoded proteins with 109 aa had theoretical molecular weights of approximately 11.5 kDa and pI of 4.39 and 4.60 for P1 and P2 isoforms, respectively. The alignment of the referred isoforms with cod and salmon parvalbumins revealed different sequence identities. Pollock P1 showed a 67% of sequence identity with P2, which suggested their classification as isoallergens. Regarding the other parvalbumins, isoform P1 presented a sequence identity that ranged from 59% with Baltic cod up to 75% with salmon. With respect to isoform P2, identity indices varied between 62% with Baltic cod and 77% with cod T2 (Van Do et al., 2003; Van Do et al., 2005b). Recombinant The ch 1 (P1 and P2 isoforms) were expressed in transformed *Escherichia coli*, but small IgE-reactivity was verified for each isoform. Contrarily to the expressed recombinant isoforms, native proteins may undergo post-translational modifications in calcium binding sites, increasing their capacity for IgE recognition (Van Do et al., 2005b). The differences observed by Sun et al. (2009) in the real-time PCR amplification of parvalbumin genes from Atlantic cod and Alaska pollock suggest a more distant genetic relationship between these two species (Sun et al., 2009). Van Do et al. (2005a), using SPT in twelve fish-allergic patients and respective sera, revealed high reactivity to nGad c 1 (cod), nThe ch 1 (Alaska pollock) and nSal s 1 (salmon).

Mer me 1

The species *Merlangius merlangus*, which was previously named *Gadus merlangus*, is popularly known as whiting. In this species, a major parvalbumin was first described by Joassin, and Gerday (1977) with a primary sequence containing 108 aa and an acetylated terminal aa group (Table 1). This protein has no disulphide bridges and the alignment of the hydrophobic amino acids suggests that the structures of the two calcium-binding sites, as well as its hydrophobic core are well preserved. The comparison of the amino acid composition of Mer me 1 with other parvalbumins is indicative of belonging to the beta-lineage (Joassin, and Gerday 1977). Additionally, data suggest a strong interaction and a positive cooperation between the two calcium-binding sites. The structure of whiting parvalbumin appears to be very stable at temperatures up to 70ºC, demonstrating some conformational alterations when heated between 70-90ºC, with expected complete denaturation at temperatures above 100ºC. Whiting parvalbumin is also stable in a wide range of pH (5.5-9.5), but out of this interval seems to lose integrity, lowering its affinity to calcium (Permyakov et al., 1980). Its tertiary structure was already determined by X-ray crystallography and is publically available in PDB (accession no. 1A75).

Gad ma 1, Bor sa 1 and Pol vi 1

Gad ma 1 and Bor sa 1 parvalbumins have been recently identified in *Gadus macrocephalus* and *Boreogadus saida*, also commonly known as Pacific cod and Polar cod/Artic cod, respectively (Table 1). Pol vi 1 is the allergenic protein of *Pollachius virens*, known as Pollock, which was described as presenting the same parvalbumin profile as cod and haddock, with molecular mass <14 kDa determined by immunoblotting with the monoclonal anti-frog parvalbumin antibody PARV-19 (Weber et al., 2009). Koppelman et al. (2012) identified pollock parvalbumins with polyclonal antibodies raised against Atlantic cod, revealing a reactivity of at least 92%. Like in the case of haddock, cod and pollock seem to share high sequence homology with allergenic parvalbumins. Gad ma 1, Bor sa 1 and Pol vi 1are present in the white muscle tissue and their route of exposure seems to be ingestion (ALLERGOME, 2014). In terms of molecular characterisation, no information is yet available.

Merlucciidae parvalbumins

Mer ap 1

Belonging to the species *Merluccius australis*, the allergen Mer ap 1 was identified in the subspecies *Merluccius australis polylepis*, more commonly known as Southern hake or Patagonian hake (UniProt, 2014). So far, three sequences defined as parvalbumins of

beta-lineage have been identified in this species (Table 1) (Carrera et al., 2010). The parvalbumin beta 1 has a primary structure of 109 aa, with a pI of 4.30 and a molecular mass of 11.30 kDa that was determined by two-dimensional (2D) gel electrophoresis and confirmed by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) (Carrera et al., 2006; Carrera et al., 2010). With the same number of residues, the parvalbumin beta 2 has a pI of 4.14 and an estimated molecular mass of 11.33 kDa (MALDI-TOF-MS determination). The third sequence defined as parvalbumin beta 3 presents identical molecular mass (~11.33 kDa) and a pI of 3.98. According to its molecular mass, the primary structure should present 109 aa, though only about 76 aa have been experimentally identified (Carrera et al., 2010). Parvalbumins beta 1 and beta 2 exhibited a sequence identity of 74%, suggesting their classification as isoallergens. Regarding the isoallergens beta 1 and beta 2, high sequence homology (>90%) was verified with different members of the *Merluccius* genus, being designated as isoform sequences (Carrera et al., 2010). Like other parvalbumins, the isoallergens from *Merluccius australis polylepis* present characteristic residues such as cysteine and arginine in positions 18 and 75, respectively.

Mer au 1

Mer au 1 was classified as an allergen of the subspecies *Merluccius australis australis*, frequently denominated as Austral hake or Southern hake (UniProt, 2014). In this subspecies, four protein sequences have been identified, exhibiting primary structures with 108 aa, with the exception of the parvalbumin beta 4 that revealed a structure with 69 aa. All four peptides are suggested to present pI ranges of 3.98-4.51 and molecular mass of approximately 11.3 kDa (2D-gel and MALDI-TOF-MS), even parvalbumin beta 4 with residues partially unidentified (Carrera et al., 2010). In Blastp analysis, parvalbumin beta 1 showed a high sequence homology with parvalbumin beta 2 (~99%), indicating that they are probably variants (isoforms) of the same isoallergen (Chapman et al., 2007; Chapman, 2008). Regarding parvalbumin beta 3, its sequence identity with parvalbumin beta 1 was of 74%, suggesting their forthcoming classification of isoallergens. Since parvalbumin beta 4 was not yet fully sequenced, its classification as an isoallergen of Mer ap 1 is still not yet possible, according to current nomenclature criteria. Like the isoallergens from Mer ap 1, high sequence homology (>90%) was verified among Mer au 1 and other parvalbumins from the *Merluccius* genus (Carrera et al., 2010). In addition to high homology, similar molecular masses and isoelectric points, the isoallergens from *Merluccius australis australis* also possess other characteristic traits of beta-parvalbumins such as a cysteine in position 18 and an arginine in position 75.

Mer bi 1

The *Merluccius bilinearis* species is frequently nominated as silver hake and is also identified as an allergenic fish. In this species, four proteins sequentially designated as parvalbumin beta 1 to parvalbumin beta 4 have been characterised and classified as isoallergens or isoforms of Mer bi 1 (Table 1). The Mer bi 1 includes three sequences that exhibit a primary structure composed of 108 aa and a fourth sequence with 94 aa. All sequences presented an estimated molecular size of approximately 11.3 kDa and pI ranging from 3.98-4.23 (UniProt, 2014). From Blastp search, the sequence parvalbumin beta revealed 95% of sequence identity with parvalbumin beta 3, suggesting that they are possibly two variants of the same isoallergen (NCBI, 2014; UniProt, 2014). In addition, parvalbumin beta 2 and parvalbumin beta 4 presented sequence identity of 76% and 68%, respectively with parvalbumin beta, confirming that they are isoallergens (NCBI, 2014; UniProt, 2014). When compared to other species from the same or even different families, sequence homology with Mer bi 1 is, in general, very high (Carrera et al., 2010).

The isoallergen parvalbumin beta, also referred as parvalbumin isoform B, was firstly described by Revett et al. (1997) and its tri-dimensional structure was determined by X-ray diffraction (PDB accession no. 1BU3) (PDB, 2014; Richardson et al., 2000). According to its conformation, parvalbumin isoform B exhibits the same structural features as other parvalbumins, consisting of six alpha-helices (A-F) arranged in a globular shape. Chains A and B are folded into antiparallel conformation, while helices C/D and E/F pair up and each adopt a perpendicular arrangement. Like in other parvalbumins, the loops between helices CD and EF form the two predominant metal ion (calcium) binding sites. Other invariant features of parvalbumins such as the Ile58-Ile97 beta-sheet, the Arg75-Glu81 salt bridge and the characteristic hydrophobic core composed by Phe, Leu, Ile and Val, are also present in this protein (Richardson et al., 2000).

Mer ca 1

In the *Merluccius capensis* (stockfish or cape hake as common names) species, three beta-parvalbumins have been identified, from which two presented 108 aa and the third one 58 aa. Exhibiting molecular masses of approximately 11.3 kDa and pI of 3.95-4.55, parvalbumin beta 1 presents 75% of sequence identity with parvalbumin beta 2, suggesting that both sequences are isoallergens (NCBI, 2014; UniProt, 2014). Regarding the parvalbumin beta 3, further information about its primary sequence is needed (Chapman et al., 2007; Chapman, 2008).

In a study described by Weber et al. (2009), the parvalbumin from *M. capensis* exhibited a molecular size of approximately 11 kDa as determined by SDS-PAGE with

silver staining and by immunostaining with monoclonal PARV-19 mouse anti-frog parvalbumin antibody. Using a competitive indirect ELISA with the referred antibody, the same authors reported 80% of cross-reactivity with cod (*Gadus* spp.).

Mer ga 1

The allergen identified in the *Merluccius gayi* species (known as South Pacific hake or English hake) is Mer ga 1. It was reported for the first time by Piñeiro et al. (1998), presenting four parvalbumins with molecular weights of 12-13 kDa and pI of 3.6-4.5 (Table 1). More recent data on the molecular size/pI and the primary structures of each parvalbumin were completely/partially identified, confirming previous findings (Carrera et al., 2010). The two sequences with 108 aa in length revealed a 74% identity, making them isoallergens of Mer ga 1 (Table 1). Parvalbumins beta 3 and 4 evidenced lower sequence identity (61-62%) with parvalbumin beta 1, probably due to their incomplete sequence (Uniprot, 2014).

Mer hu 1

Mer hu 1 is a parvalbumin from *Merluccius hubbsi* species, commonly known as Argentine hake. Like in *Merluccius gayi*, four parvalbumins were identified in *M. hubbsi* with pI of 3.9-4.5 (Piñeiro et al., 1998). MALDI-TOF-MS revealed two sequences with a primary structure composed of 108 aa and a third one with 86 aa, all presenting molecular masses of approximately 11.3 kDa and pI of 4.09-4.57 (Carrera et al., 2010), which confirmed the previous data (Piñeiro et al., 1998). However, the fourth parvalbumin identified by Piñeiro et al. (1998) with a pI of 3.9, common to other three *Merluccius* species (*M. capensis*, *M. australis* and *M. merluccius*), was not reported in the study of Carrera et al. (2010) or in ALLERGOME database.

Mer mr 1

The isolation of a major parvalbumin with 108 aa (11.47 kDa) from the white muscle of *Merluccius merluccius* species was firstly described by Pechère et al. (1971) and further confirmed by Capony, and Pechère (1973). Thereafter, other authors identified a set of three aa sequences containing 108 residues in length, molecular masses around 11.3 kDa and pI ranging from 3.9-4.5 (Carrera et al., 2006; Carrera et al., 2010; Piñeiro et al., 1998; UniProt, 2014). Together with these three sequences, a fourth structure displaying 69 aa was also classified as a parvalbumin from *M. merluccius*, sharing similar biochemical properties to other parvalbumins (UniProt, 2014). A Blast search on the available sequences of parvalbumins evidenced high homology (>99%) between

parvalbumin beta and beta 2, suggesting their classification as isoforms or variants of Mer mr 1 (Table 1). The primary structure of parvalbumin beta 4 was not yet fully sequenced, presenting only 61% of sequence identity with the remaining three parvalbumins. In spite of this fact, the four parvalbumins are considered isoallergens of Mer mr 1 (Carrera et al., 2010).

Mer pr 1

Mer pr 1 belongs to *Merluccius productus*, commonly named as North Pacific hake. It comprises two sequences with 108 aa and one with 88 aa, exhibiting more than 57% of homology among sequences. The characteristic traits of beta-parvalbumins such as a molecular weight of approximately 11.3 kDa and pI below 5 are also shared by Mer pr 1 (Wopfner et al., 2007).

So far, no nucleotide sequences encoding these proteins have become available in the GenBank database, although in the study of Sun et al. (2009), the amplification of an universal nucleotide sequence evidenced a more close genetic relation among *M. productus* and other species such as *Hymenocephalus striatissimus* (hoki), *Sciaenops ocellatus* (red fish) or *Theragra chalcogramma* (Alaska pollock) than with *Gadus macrocephalus* (Pacific cod) or *Aristichthys nobilis* (Bighead carp).

Mer se 1, Mer po 1 and Mer pa 1

Mer se 1, Mer po 1 and Mer pa 1 belong to *Merluccius senegalensis*, *M. polli* and *M. paradoxus* species, frequently known as Senegalese hake or Black hake, Benguela hake and deepwater hake/deepwater Cape hake, respectively. Each allergen is composed of three isoallergens and besides presenting the same properties of the beta-parvalbumins already described (Table 1), limited information is yet available regarding these molecules.

Mac ma 1 and Mac n 1

The *Macruronus magellanicus* or *Macruronus novaezelandiae magellanicus*, frequently designated as Patagonian grenadier, is also considered an allergenic fish. Recently, the isoallergens Mac ma 1 have been completely/partially sequenced (Table 1). Presenting primary structures with 98 aa, 108 aa and 74 aa for parvalbumins beta 1, beta 2 and beta 3, respectively, all isoallergens evidenced similar molecular weights (~11.3 kDa) and pI (<5). Besides the common features associated with beta-parvalbumins, the isoallergens Mac ma 1 also exhibit two EF-hand repeats and conserved amino acid regions, such as an aspartic acid in position 61, which is common to several parvalbumin sequences (Mer

ap 1, Mer au 1, Mer bi 1, Mer ca 1; Mer ga 1, Mer hu 1, Mer mr 1, Mer pa 1, Mer po 1, Mer pr 1 and Mer se 1) (Carrera et al., 2006; Carrera et al., 2010). Additionally, for all the referred sequences, as well as for Mac ma 1, a cysteine in position 18 and an arginine residue in position 75 were always conserved, which are considered characteristic traits of parvalbumins of beta-lineage. In relation to the parvalbumins that were completely sequenced, homology among different species from *Merluccius* or *Macruronus* genera is frequently higher than 80%, explaining the high incidence of cross-reactivity among them.

The *Macruronus novaezelandiae* or *Macruronus novaezelandiae novaezelandiae*, commonly known as Blue grenadier, has been described as an allergenic fish. Like the Mac ma 1, the Mac n 1 is composed by three sequences with 98 aa, 108 aa and 83 aa, which were firstly reported by Carrera et al. (2006). Mac n 1 isoallergens present all the biochemical traits of beta-parvalbumins, namely molecular weight of approximately 11.3 kDa, pI<5 and a set of well-conserved amino acid regions (a cysteine in position 18, an aspartic acid in position 61 and an arginine in position 75) (Carrera et al., 2006; Carrera et al., 2010).

Other Gadiform parvalbumins

The Mur mi 1 is the allergenic parvalbumin identified in the *Muraenolepis microps* species, commonly known as Smalleye Moray Cod and included in the Muraenolepididae family. It was firstly reported by Sun et al. (2009), presenting close sequence amplification with other members of Gadidae family (e.g. *Micromesistius poutassou*, *Gadus macrocephalus*), suggesting their high sequence homology.

The Uro te 1 allergen was identified in *Urophycis tenuis* species, commonly known as Boston ling or White hake, which is a member of Phycidae family. It was detected using the monoclonal anticarp antibody, suggesting high structural homology between respective sequences (Koppelman et al., 2012; Lee et al., 2011). In the same sense, parvalbumin Uro te 1 also evidenced 100% of cross-reactivity with the polyclonal anticod antibody described by Koppleman et al. (2012), which reinforces that most parvalbumins share similar traits. Regarding this allergen, no relevant biochemical information could be found in literature.

Clinical relevance of Gadiform parvalbumins

From a clinical point of view, parvalbumins are major allergens that are present in a broad spectrum of fish species (Beale et al., 2009; Lindstrom et al., 1996; Perez-Gordo et al., 2011; Van Do et al., 2005a). During an allergic episode, fish-allergic patients can exhibit individual symptoms or combined manifestations, depending on the target

organ/systems affected. The observable clinical presentations of fish allergy can be classified as mild, such as oral allergy syndrome (OAS) and general erythema, moderate (urticaria, vomiting) or severe (angioedema, asthma, anaphylaxis) (Bock et al., 2001; Helbling et al., 1999; Sicherer et al., 2000). In the specific case of fish allergy, patients with severe and systemic symptoms are often common. A great part of the described symptoms occur after ingestion, though urticaria and eczema may also appear upon skin contact with fish/fish products, as well as upper and lower airway syndromes as consequence of the inhalation of odours during fish preparing/processing (Jeebhay et al., 2008). Most fish sensitised/allergic patients are positively reactive to multiple species from taxonomically distinct classes (genetically close or distantly related species). Moreover, serological studies and skin testing seem to suggest high cross-sensitisation (70-90%) (Bernhisel-Broadbent et al., 1992), though clinical cross-reactivity is only confirmed in 50% to 70% of the cases (De Martino et al., 1990; Helbling et al., 1999; Sicherer, and Sampson, 2010). Reported differences in the allergenicity among different fish species are mainly related to variable allergen content and to the effects of food processing (Kuehn et al., 2010).

Several parvalbumins have been identified in members of the Gadidae family and comprised in the ALLERGOME database, but only few molecules have been included in the IUIS official list of allergens. In spite of their relevance, clinical information could only be found for Gad c 1, Gad m 1, Mic po 1 and The ch 1 parvalbumins, which belong to the most consumed fish species (Bugajska-Schretter et al., 1998; Griesmeier et al., 2010; Ma et al., 2008; Rancé et al., 1999; Sletten et al., 2010; Van Do et al., 2005a). With respect to Gad c 1, it has been classified as a major allergen in cod (Bugajska-Schretter et al., 1998), considering that more than 50% of the sera of fish-allergic patients were reactive to this allergen (Chapman, 2008). In a study described by Bugajska-Schretter et al. (1998), using a test population of 30 patients with clear history of at least one clinical symptom (dermatitis, urticaria, angioedema, diarrhoea, asthma and/or anaphylactic reaction) after ingestion, inhalation or skin contact with fish proteins, the sera of all individuals were IgE-reactive to Gad c 1. In the same study, Gad c 1 was submitted to different chemical treatments to evaluate its immunoreactivity. Similarly to Gad c 1, high immunoreactivity to Gad m 1 was also observed in patients' sera with clear history of well-defined symptoms that range from mild (skin rash or OAS) to potentially life-threatening (anaphylaxis) (Griesmeier et al., 2010; Ma et al., 2008). From a total of 19 fish-allergic patients presenting a wide variety of clinical symptoms (atopic dermatitis, asthma, urticaria, throat swelling, vomiting, breathlessness, abdominal pain and/or anaphylaxis), the sera of 18 individuals were IgE-reactive to blue whiting parvalbumin (Mic po 1) (Sletten et al., 2010).

For members of Merlucciidae family, only two studies reporting clinical symptoms were found in the literature (Beale et al., 2009; González-de-Olano et al., 2012), both regarding the Mer mr 1. Beale et al. (2009) reported that, using sera of ten patients with clinical and serological history of self-reported allergy after fish ingestion, they were able to identify four individuals IgE-reactive to *Merluccius merluccius*. The reported allergic reactions varied from moderate to severe, in general with systemic clinical symptoms of urticaria, dermatitis, vomiting, asthma, rhinoconjunctivitis and/or angioedema. Additionally, a rare clinical case evidencing the cross-reactivity between chicken and fish parvalbumins was recently described by González-de-Olano et al. (2012). In the referred study, a 23-year-old patient with a documented severe fish-allergic episode at age of nine (chest tightness, wheezing and facial angioedema) presented a very similar clinical condition (chest tightness and wheezing), within minutes upon the ingestion of chicken. Using this patient's serum, strong IgE-reactivity towards hake and chicken extracts could be observed.

Detection of Gadiform parvalbumins

The risk of allergic individuals of suffering from an abnormal immune episode remains a reality, even when following a daily diet with total avoidance of the offending foods. Despite the EU regulations requesting the compulsory labelling of food products that may contain food allergens (Directive 2007/68/EC; EFSA, 2014; Regulation (EU) No. 1169/2011), incorrect labelling or cross-contamination cases may occur upon food processing, resulting in the inadvertent presence of hidden allergens. This fact emphasises the need for accurate, fast and highly sensitive methods in order to detect trace levels of allergens in foods and protect fish-allergic consumers, while ensuring the accurate labelling of products (Carrera et al., 2012; Costa et al., 2014; Herrero et al., 2014; Lee, and Taylor, 2011; Rencova et al., 2013).

Despite the lack of available testing/reference materials and the absence of official methods, there are currently several reports describing molecular tools to detect food allergens. However, within the sphere of fish allergen detection, the number of available methods is still limited.

In recent years, methods targeting proteins and DNA have been proposed for the detection of fish allergens (mostly parvalbumins). Protein-based methods are the most widely used and they include two main groups of techniques: immunoassays and mass spectrometry methods. DNA-based analysis has allowed the development of indirect allergen detection methods, predominantly by means of PCR amplification (Carrera et al., 2012; Rencova et al., 2013; Zheng et al., 2012).

Protein-based methods

Protein-based methods are particularly helpful in the food industry since they offer high specific and sensitive tools for fast screening of allergens within a complex food matrix, without needing an extensive sample preparation step (Lee, & Taylor, 2011). Considering the detection of Gadiform allergens, the available commercial kits and literature reports mostly rely on immunochemical techniques based on allergen/antibody interaction (Tables 2 and 3).

Lateral flow devices

Lateral flow devices (LFD) are simple and rapid tests that provide qualitative or semi-quantitative information without the need for specialised equipment. Presently, only one LFD is commercially available for the analysis of general fish proteins (Table 2), with a sensitivity of 2 mg/kg of fish in foodstuffs and a test performance time of about 10 minutes.

To detect *Gadus macrocephalus* parvalbumin (Table 2), a superparamagnetic nanoparticle-based lateral flow immunoassay (SPMN-LFIA) was recently developed by means of monoclonal antibodies raised against fish parvalbumin, allowing its rapid detection in foods (~20 min) and a limit of detection (LOD) of 0.046 mg/kg (Zheng et al., 2012).

The simplicity of LFD with the possibility of visually reading the result makes these assays suitable for wide applications in the control of allergens at industrial level, though they are highly prone to false negative results in part caused by the complexity of the matrix (Diaz-Amigo, 2010).

ELISA

As noticed in Tables 2 and 3, the enzyme-linked immunosorbent assay (ELISA) is by far the most used method to trace Gadiform allergens, as in the case of other allergenic foods (e.g. tree nuts, peanut and soybean). ELISA allows the direct detection and quantification of allergens or other marker proteins, usually providing a rapid and low-cost method with no special requirements for expertise knowledge (Costa et al., 2014). Currently, there are two ELISA kits commercially available to enable the detection of fish in foods: the “Fish Protein ELISA kit” targeting parvalbumins with a limit of quantification (LOQ) of 1 mg/kg; and the “AgraQuant® Fish” with a LOQ of 4-100 mg/kg, though without clear information regarding the target class of fish proteins (Table 2). Due to the structural homology of parvalbumins, the possibility of cross-reactivity is referred in the Fish Protein ELISA kit namely for cod, haddock or mackerel species (Table 2).

Table 2. Commercial LFD, ELISA and real-time PCR kits available for the detection and quantification of fish allergens.

Commercial kit/brand	Assay type	Cross-reactivity	LOD	LOQ	Estimated time to perform assay
Fish Protein Rapid Test (Elution Technologies, Vermont, USA)	LFD	No cross-reactivity observed	2 mg/kg	No information available	~10 min (sample preparation)
Fish Protein ELISA Kit (Elution Technologies, Vermont, USA)	Sandwich ELISA	Salmon, Cod, Trout, Sole, Tilapia, Haddock, Bass, Sardine, Mackerel, Halibut.	No information available	1 mg/kg	~50 min
AgraQuant® Fish (Romer Labs Division Holding GmbH, Austria)	Quantitative – Sandwich ELISA	No information available	1.4 mg/kg	4-100 mg/kg	~40 min (sample extraction) + ~60 min (incubation time)
SureFood Allergen Fish (R-Biopharm AG Darmstadt, Germany)	Real-time PCR (qualitative)	No cross-reactivity observed	≤ 5 DNA copies, ≤ 0.4 mg/kg	No information available	~60 min (Real-time reaction)

Three ELISA, each one in the formats of sandwich-type, competitive indirect and quantitative, have been proposed in literature for the detection and quantification of parvalbumins in fish (e.g. *Gadus morhua* or *Pollachius virens*) and in derived-fish foodstuffs and gelatines (Table 3). Recently, Fæste and Plassen (2008) also developed a sandwich ELISA targeting a fish muscle protein (34 kDa) rather than parvalbumin. In general, these assays present high sensitivity and no cross-reactivity with plant or animal-derived food (fish not included), excluding the positive result for European squid (Fæste, and Plassen, 2008).

Despite the predominance of ELISA, both as commercial kits or as reported methods, their results should be carefully analysed. These assays are likely to present false positive results since the target proteins might be highly affected by conformational changes upon processing (thermal treatments, pH, fermentation and partial hydrolysis). However, the capacity of multiple epitope recognition by polyclonal antibodies leads to a better detection of modified proteins (Costa et al., 2014; Lee, & Taylor, 2011).

MS platforms

In recent years, the enormous progress in MS technology has given a boost to the field of proteomics, including the identification. characterisation and determination of food allergens (Fæste et al., 2011; Monaci, and Visconti, 2010). The application of proteomic methods for the analysis of allergenic proteins has been termed allergenomics. The many different allergenic proteins are included in the allergenome, whose information is

Table 3. Most relevant protein-based methods for the detection of Gadiform fish allergens available in the literature.

Method	Antibody	Cross-reactivity	Sensitivity level	Food matrices	References
Sandwich ELISA	Rabbit antibodies (raised against cod parvalbumin, *Gadus morhua*)	1/24 plant and animal-derived food (fish not included). Cross-reactivity with European squid (0.6 mg/kg parvalbumin)	LOD = 0.01 mg/kg LOQ = 0.02 mg/kg	Fish (32 species, including *Gadus morhua* and *Pollachius virens*) and food/ingredient (24) samples	Faeste, and Plassen (2008)
Competitive indirect ELISA	Rabbit antibodies (raised against frog muscle parvalbumin)	4/5 fish species. Strong cross-reactivity related to cod (80-179%)	LOD = 18.7 mg/kg fish gelatin	Fish (8 species, including *Gadus spp.*, *Pollachius virens*, *Merluccius capensis*, *Melanogrammus aeglefinus*) Fish gelatins Isinglass	Weber et al. (2009)
Sandwich ELISA	Polyclonal rabbit antibodies (raised against a fish muscle protein, 36 kDa)	0/16 animal and plant-derived food (fish not included)	LOD = 0.1 mg/kg	Fish (63 species, including *Gadus morhua* and *Pollachius pollachius*) Fish in crab meat mixtures	Chen, and Hsieh (2014)
Quantitative ELISA	Polyclonal rabbit antibodies (raised against purified parvalbumin from fish muscle)	Not verified	LOD = 2-6 ng parvalbumin/ml	Fish (8 species, including *Gadus morhua*)	Kuehn et al. (2010)
SPMNP-LFIA	Monoclonal antibody (raised against fish parvalbumin)	Not verified	LOD = 0.046 mg/kg	Fish (17 species, including *Gadus macrocephalus*) Plant and animal food matrices	Zheng et al. (2012)
LC-MS/MS (multi-target approach)	Not applicable (peptide biomarkers for parvalbumin detection)	Not verified	No information available	Fish (16 species, including *Gadus morhua*) Non-fish (6 species) Commercial sea-foodstuffs	Carrera et al. (2012)

continuously being accumulated in ALLERGOME database (Mari et al., 2009). For further reading, Fæste et al. (2011) provide a compendium of studies on proteomic analysis, marker peptides, liquid chromatography (LC) methods and quantitative assays for the 14 main food allergens.

The protein analysis by LC and MS has greatly advanced in recent years, allowing allergen identification, characterisation and quantification (Fæste et al., 2011). There are several advantages that make this technique suitable for allergen detection, such as high accuracy, sensitivity, specificity and reproducibility, while presenting less problems regarding cross-reactivity, frequently associated with immunoassays (Costa et al., 2014).

Carrera et al. (2012) proposed a MS-driven detection, as an alternative method for the direct identification of parvalbumin in any food product, in less than 2 h (Table 3). This method was tested in 16 fish species (including *Gadus morhua)* and involved reverse phase LC for the separation of peptides and selected MS/MS ion monitoring for multi-target analysis of beta-parvalbumin biomarkers.

MS strategies have been suggested as having high potential for unequivocal identification of food allergens. Nevertheless, the need for specialised personnel and the high cost of the equipment are some disadvantages that hamper the wide application of this technology.

DNA-based methods

Lately, methodologies based on DNA analysis have been referred as adequate alternatives to proteins and considered as methods of choice for the differentiation and identification of distinct components in foods (Mafra et al., 2008). DNA molecules can be found in most biological tissues, presenting higher stability/resistance to adverse conditions than proteins. Therefore, DNA is less prone to be affected by thermal treatments, pH alterations or partial hydrolysis than proteins, which normally suffer structural changes that, consequently, affect their detection. As a result, DNA-based methods are especially useful to analyse highly processed foodstuffs, emphasising their important role for the management of allergens in food industry (Eischeid et al., 2013; Herrero et al., 2014). In spite of the growing number of DNA-based methods for the detection and quantification of food allergens, only few studies report fish allergens, particularly considering the Gadiformes.

Table 4. Most relevant DNA-based methods for the detection of Gadiform fish allergens available in the literature.

Method	Target gene	Fragment size (bp)	Cross-reactivity	Sensitivity level	Food matrices	References
Conventional PCR	Atlantic herring parvalbumin partial sequence	189	1/26 fish species (Strong cross-reactivity with *Clupea Pallasii* (Pacific herring)	10 pg/μL	Fish (26 species, including *Gadus morhua*, *Gadus merlangus*, *Merluccius merluccius*, *Theragra chalcogramma*, *Pollachius virens*)	Rencova et al. (2013)
Real-time PCR for fish allergen detection	18S RNA gene sequence	Not available	0/47 plant and animal-derived food (fish not included)	50 pg	Fish (54 species, including *Gadus macrocephalus*, *Gadus morhua*, *Merluccius australis* and *Merluccius merluccius*)	Herrero et al. (2014)
Real-time PCR for fish parvalbumin detection	Parvalbumin gene	Not available	0/13 plant and animal-derived food (fish not included)	5 pg	Fish (30 species, including 8 Gadiform varieties)	Sun et al. (2009)
Real-time PCR for multiple fish species detection	Rhodopsin gene partial sequence	Not available	Not available	0.05 ng	Animal feeds with fish as an ingredient (targeting 22 fish species, including 12 Gadiform species)	Prado et al. (2012)
Real-time PCR for multiple fish species detection	12S rRNA gene sequence	87 bp	Not available	0.1 pg	Animal feeds with fish as an ingredient (targeting 39 fish species, including 3 Gadiform species)	Pegels et al. (2013)

PCR

PCR provides a fast, simple and sensitive tool for the specific detection of sequences encoding allergenic proteins or species-specific markers. Since it does not target directly the offending proteins, it is considered an indirect method of detecting food allergens (Costa et al., 2014; Hildebrandt, 2010; Mafra et al., 2008).

Up to date, the majority of the reports describing the application of PCR-based methods for fish analysis have mainly been focused on species-specific identification for authentication purposes. Concerning commercial applications, there is one kit available for fish DNA detection (Table 2). The five reported fish-specific PCR methods are resumed in Table 4. The specific detection of the parvalbumin gene of fish (including species from the Gadiform order) by a real-time PCR assay was solely proposed by Sun et al. (2009), allowing a sensitivity of 5 pg.

Generally, these methods provide high specificity by minimising or avoiding cross-reactivity with other non-target food species. Therefore, the application of DNA-based methods has high potential for the detection and quantification of parvalbumin sequences.

Biosensors

Biosensors represent probably one of the most promising ways of simple, fast, reproducible, and cheap multi-analyte detection alternatives. A biosensor is an integrated receptor-transducer device that converts the biological-recognition event into a measurable chemical physical signal, which is proportional to the target concentration. The receptor can be an antibody raised against an allergen, a single-stranded DNA molecule capable of hybridising with an allergen-specific DNA fragment, or an aptamer selected to recognise the target allergen directly (Pilolli et al., 2013). In the case of food allergens (Alves et al., 2015; Pilolli et al., 2013; Prado et al., 2015), as in the case of other food analysis, such as the detection of genetically modified organisms (Plácido et al., 2016), three major groups of biosensors, depending of the type of transducer can be found: electrochemical, optical (surface plasmon resonance, SPR) and piezoelectric (quartz crystal microbalance, QCM). Concerning the type of recognition element, immunosensors are the most commonly applied to food allergen analysis, followed by genosensors (DNA-sensors) and, more recently, aptasensors (Alves et al., 2015).

In the specific case of biosensing approaches applied to Gadiform parvalbumin detection, to our knowledge, no reports have become available. To detect fish parvalbumins, Lu et al. (2004) developed a SPR immunosensor based on a monoclonal antibody against tuna and carp parvalbumins that was applied to processed seafoods (Table 3). More recently, Wang et al. (2011) developed a highly efficient assay to

simultaneously identify eight food allergens (soybean, wheat, peanut, cashew, shrimp, fish, beef, and chicken) consisting of a silicon-based optical thin-film biosensor chip. The assay was based on two tetraplex PCR systems that were developed and validated, followed by an enzyme-labelled indirect detection. The enzymatic reaction products precipitated on the thin-film surface and modified the interference pattern of light on the biosensor surface, producing a significant colour change on the surface. The main advantage of this approach was the possibility to perceive the colour change with the naked eye, without any extra analysis/equipment. Fish detection was achieved after PCR amplification of a sequence of the mitrocondrial 16S rRNA gene of *Plecoglossus altivelis*, only reporting sensitivity for cashew nut (Wang et al., 2011).

Final remarks

The consumption of fish is growing as a result of the continuous recommendations for its inclusion in the context of a healthy diet. Nevertheless, the number of reported cases of fish allergy has increased in recent years, representing nowadays an important issue of food safety.

Recently, relevant information on the biochemical classification of fish allergens has become available, which has improved its correlation with the elicited clinical symptoms. Currently, the parvalbumins are considered the most important class of fish-allergens, being responsible for inducing moderate to severe and systemic adverse immunological reactions in fish-sensitised/allergic individuals. So far, cod has been frequently described as the model case regarding the study of fish allergy, being also the most well studied. Although the patterns of sensitisation to fish parvalbumins have been related to ingestion, direct contact and/or inhalation, they are yet not fully understood, suggesting the need for additional studies. From a clinical point of view, it is expected that the increasing consumption of fish could lead to a rising number of fish-allergic individuals, which emphasises the importance of this subject. Although there are reports stating the relevance of parvalbumins, more clinical data should be made available regarding other less studied allergens. Furthermore, it becomes clear that more reliable and extensive epidemiological studies are essential to support the existing data on the prevalence of fish allergy.

Until now, the avoidance of offending foods is the only effective means of preventing adverse immunological reactions. Therefore, regardless of being protein- or DNA-based techniques, the development of new methodologies for the detection and quantification of parvalbumins in fish is markedly significant for the better management of allergens by the food industry. Food technologies can induce changes due to processes, such as

mechanical stress, high temperature, pH variations, enzymatic activities or fermentations may result in significant degradation of both proteins and DNA, which can affect the sensitivity of the analysis, particularly the limits of detection and quantification (Fernandes et al., 2015; Prado et al., 2015).

Regarding Gadiform allergens, immunochemical assays, namely ELISA, are the most used methods for parvalbumin testing, in terms of both commercial and reported assays. Recently, PCR technologies have emerged as excellent alternatives to identify parvalbumin or other gene sequences and assess the effect of food processing on its detection/allergenicity, although much effort is also required for the establishment of novel assays as effective tools. New methodologies based on biosensor devices have also been developed for the detection of allergenic foods, as well as the use of nanomaterials for analytical methods improvement (Pilolli et al. 2013; Prado et al., 2015). However, in the particular case of fish allergen detection, the biosensing developments are still very scarce, highlighting the need for future research in such promising approaches. Moreover, forthcoming advances should become available upon the development of testing/reference materials, as well as further improvements to establish threshold levels.

Conflict of interests

The authors declare no conflict of interests.

Funding

This work has been supported by Fundação para a Ciência e a Tecnologia (FCT) through grant no. UID/QUI/50006/2013 and NORTE-07-0124-FEDER-000069-Food Science. Telmo J. R. Fernandes is grateful to FCT PhD grant (SFRH/BD/93711/2013) and Joana Costa to FCT Post-doc grant (SFRH/BPD/102404/2014), both financed by POPH-QREN (subsidised by FSE and MCTES).

References

ALLERGEN, official site for the systematic allergen nomenclature. Available at: http://www.allergen.org/ (Last accessed 8 December 2014).

ALLERGOME Database, the platform for allergen knowledge, Latina, Italy. Available at: http://www.allergome.org (Last accessed 8 December 2014).

Alves, R. C., Barroso, M. F., Gonzállez-García, M. B., Oliveira, M. B. P. P., Delerue-Matos, M. C. (2015). New trends in food allergens detection: towards biosensing stratagies. *Crit. Rev. Food Sci. Nutr.* doi:10.1080/10408398.2013.831026

Arif, S. H. (2009). A Ca^{2+} binding protein with numerous roles and uses: parvalbumin in molecular biology and physiology. *Bioessays*. **31**: 410-421.

Arif, S. H., Jabeen, M., and Hasnain, A. U. (2007). Biochemical Characterization and thermostable capacity of parvalbumins: The major fish-food allergens. *J Food Biochem.* **31**: 121-137.

Beale, J. E., Jeebhay, M. F., and Lopata, A. L. (2009). Characterisation of purified parvalbumin from five fish species and nucleotide sequencing of this major allergen from Pacific pilchard, *Sardinops sagax*. *Mol Immunol.* **46**: 2985-2993.

Ben-Shoshan, M., Harrington, D. W., Soller, L., Fragapane, J., Joseph, L., St Pierre, Y., Godefroy, S. B., Elliot, S. J., and Clarke, A. E. (2010). A population-based study on peanut, tree nut, fish, shellfish, and sesame allergy prevalence in Canada. *J Allergy Clin Immunol.* **125***:* 1327-1335.

Berin, M. C., and Sampson, H. A. (2013). Food allergy: an enigmatic epidemic. *Trends Immunol.* **34**: 390-397.

Bernhisel-Broadbent, J., Scanlon, S. M., and Sampson, H. A. (1992). Fish hypersensitivity. I. In vitro and oral challenge results in fish- allergic patients. *J Allergy Clin Immunol.* **89**: 730-737.

Bock, S. A., Muoz-Furlong, A., and Sampson, H. A. (2001). Fatalities due to anaphylactic reactions to foods. *J Allergy Clin Immunol.* **107**: 191-193.

Breiteneder, H., and Radauer, C. (2004). A classification of plant food allergens. *J Allergy Clin Immunol.* **113***:* 821-830.

Breiteneder, H., and Mills, E. N. C. (2005a). Molecular properties of food allergens. *J Allergy Clin Immunol.* **115**: 14-23.

Breiteneder, H., and Mills, E. N. C. (2005b). Plant food allergens - structural and functional aspects of allergenicity. *Biotechnol Adv.* **23**: 395-399.

Bugajska-Schretter, A., Elfman, L., Fuchs, T., Kaplotis, S., Rumpold, H., Valenta, R., and Spitzauer, S. (1998). Parvalbumin, a cross-reactive fish allergen, contains IgE-binding epitopes sensitive to periodate treatment and Ca^{2+} depletion. *J Allergy Clin Immunol.* **101**: 67-74.

Bugajska-Schretter, A., Grote, M., Vangelista, L., Valent, P., Sperr, W. R., Rumpold, H., Pastore, A., Reichelt, R., Valenta, R., and Spitzauer, S. (2000). Purification, biochemical, and immunological characterisation of a major food allergen: different immunoglobulin E recognition of the apo- and calcium-bound forms of carp parvalbumin. *Gut.* **46**: 661-669.

Burks, A. W., Tang, M., Sicherer, S., Muraro, A., Eigenmann, P. A., Ebisawa, M., Fiocchi, A., Chiang, W., Beyer, K., Wood, R., Hourihane, J., Jones, S. M., Lack, G., and Sampson, H. A. (2012). ICON: Food allergy. *J Allergy Clin Immunol.* **129**: 906-920.

Burney, P. G. J., Potts, J., Kummeling, I., Mills, E. N. C., Clausen, M., Dubakiene, R., Barreales, L., Fernandez-Perez, C., Fernandez-Rivas, M., Le, T. M., Knulst, A. C., Kowalski, M. L., Lidholm, J., Ballmer-Weber, B. K., Braun-Fahlander, C., Mustakov, T., Kralimarkova, T., Popov, T., Sakellariou, A., Papadopoulos, N. G., Versteeg, S. A., Zuidmeer, L., Akkerdaas, J. H., Hoffmann-Sommergruber, K., and van Ree, R. (2014). The prevalence and distribution of food sensitization in European adults. *Allergy*. **69**: 365-371.

Burney, P., Summers, C., Chinn, S., Hooper, R., Van Ree, R., and Lidholm, J. (2010). Prevalence and distribution of sensitization to foods in the European Community Respiratory Health Survey: a EuroPrevall analysis. *Allergy.* **65**: 1182-1188.

Chen, Y. T., and Hsieh, Y. H. P. (2014). A sandwich ELISA for the detection of fish and fish products. *Food Control.* **40**: 265-273.

Calo-Mata, P., Sotelo, C. G., Pérez-Martín, R. I., Rehbein, H., Hold, G. L., Russel, V. J., Pryde, S., Quinteiro, J., Rey-Méndez, M., Rosa, C., and Santos, A. T. (2003). Identification of gadoid fish species using DNA-based techniques. *Eur Food Res Technol.* **217**: 259-264.

Capony, J. P., and Pechère, J. F. (1973). The primary structure of the major parvalbumin from hake muscle. Tryptic peptides derived from the S-sulfo and the performic-acid-oxidized proteins. *Eur J Biochem.* **32**: 88-96.

Carrera, M., Cañas, B., and Gallardo, J. M. (2012). Rapid direct detection of the major fish allergen, parvalbumin, by selected MS/MS ion monitoring mass spectrometry. *J Proteomics.* **75**: 3211-3220.

Carrera, M., Cañas, B., Vázquez, J., and Gallardo, J. M. (2010). Extensive de Novo sequencing of new parvalbumin isoforms using a novel combination of bottom-up proteomics, accurate molecular mass measurement by FTICR-MS, and selected MS/MS ion monitoring. *J Proteome Res.* **9**: 4393-4406.

Carrera, M., Cañas, B., Piñeiro, C., Vázquez, J., and Gallardo, J. M. (2006). Identification of commercial hake and grenadier species by proteomic analysis of the parvalbumin fraction. *Proteomics.* **6**: 5278-5287.

Chafen, J. J. S., Newberry, S. J., Riedl, M. A., Bravata, D. M., Maglione, M., Suttorp, M. J., Sundaram, V., Paige, N. M., Towfigh, A.,, Hulley, B. J., and Shekelle, P. G. (2010). Diagnosing and managing common food allergies: A systematic review. *J Amer Med Assoc.* **303**: 1848-1856.

Chapman, M. D., Pomés, A., Breiteneder, H., and Ferreira, F. (2007). Nomenclature and structural biology of allergens. *J Allergy Clin Immunol.* **119**: 414-420.

Chapman, M. D. (2008). Allergen Nomenclature. **In**: Allergens and Allergen Immunotherapy, pp. 47-58. Lockey R. F., & Ledford, D. K., Eds., Informa Healthcare, New York.

CODEX STAN 1-1985. (2010). Amended in 1991, 1999, 2001, 2003, 2005, 2008 and 2010 regarding the general standard for the labelling of pre-packaged foods. FAO/WHO Standards, Off Codex Stand. Rome, Italy. Available at: http://www.codexalimentarius.org/standards/list-of-standards. (Last accessed 4 December 2014).

Cohen D. M., Inada, T., Iwamoto, T., and Scialabba, N. (1990). FAO species catalogue. Gadiform fishes of the world (Order Gadiformes). An annotated and illustrated catalogue of cods, hakes, grenadiers and other gadiform fishes known to date. **In**: FAO Fisheries Synopsis, pp 442, FAO, Rome.

Costa, J., Carrapatoso, I., Oliveira, M. B. P. P., and Mafra, I. (2014). Walnut allergens: Molecular characterization, detection and clinical relevance. *Clin Experim Allergy.* **44**: 319-341.

Das Dores, S., Chopin, C., Villaume, C., Fleurence, J., and Guéant, J. L. (2002). A new oligomeric parvalbumin allergen of Atlantic cod (Gad mI) encoded by a gene distinct from that of Gad cI. *Allergy: Eur J Allergy Clin Immunol, Supp.* **57**: 79-83.

De Martino, M., Novembre, E., Galli, L., De Marco, A., Botarelli, P., Marano, E., and Vierucci, A. (1990). Allergy to different fish species in cod-allergic children: In vivo and in vitro studies. *J Allergy Clin Immunol.* **86**: 909-914.

Diaz-Amigo, C. (2010). Antibody-based detection methods: from theory to practice. **In**: Molecular biological and immunological techniques and applications for food chemists, pp. 223-245. Popping, B., Diaz-Amigo, C., Hoenicke, K., Eds., John Wiley & Sons Inc., New Jersey.

Di Finizio, A., Guerriero, G., Russo, G. L., and Ciarcia, G. (2007). Identification of gadoid species (Pisces, Gadidae) by sequencing and PCR-RFLP analysis of mitochondrial 12S and 16S rRNA gene fragments. *Eur Food Res Technol.* **227**: 337-344.

Directive 2006/142/EC of 22 December 2006 amending Annex IIIa of Directive 2000/13/EC of the European Parliament and of the Council listing the ingredients which must under all circumstances appear on the labelling of foodstuffs. *Official Journal of the European Union*, *L368*, 110-111.

Directive 2007/68/EC of 27 November 2007 amending Annex IIIa to Directive 2000/13/EC regarding certain food ingredients. *Official Journal of the European Union*, *L310*, 11-4.

EFSA (2014). Scientific Opinion on the evaluation of allergenic foods and food ingredients for labelling purposes. *EFSA J.* **12.** doi:10.2903/j.efsa.2014.3894

Eischeid, A. C., Kim, B. H., and Kasko, S. M. (2013). Two quantitative real-time PCR assays for the detection of penaeid shrimp and blue crab, crustacean shellfish allergens. *J Agri Food Chem.* **61**: 5669-5674.

Elsayed, S., and Apold, J. (1983). Immunochemical analysis of cod fish Allergen M: Locations of the immunoglobulin binding sites as demonstrated by the native and synthetic peptides. *Allergy: Eur J Allergy Clin Immunol.* **38**: 449-459.

Elsayed, S., & Bennich, H. (1975). The Primary Structure of Allergen M from Cod. *Scand J Immunol.* **4**: 203-208.

Elsayed, S., Ragnarsson, U., Apold, J., Florvaag, E., and Vik, H. (1981). Allergenic synthetic peptide corresponding to the second calcium-binding loop of cod Allergen M. *Scand J Immunol.* **14**: 207-211.

Elsayed, S., von Bahr Lindström, H., and Bennich, H. (1974). The Primary Structure of Fragment TM2 of Allergen M from Cod. *Scand J Immunol.* **3**: 683-686.

Fernandes, T. J. R. Costa, J., Plácido A., Villa, C., Grazina, L., Meira, L., Oliveira, M. B. P .P., and Mafra, I. (2015). GMO analysis as affected by DNA degradation. In Genetically Modified Organisms in Foods, pp. 111-118. Watson, R., and Preedy, V. R., Eds, Elsevier, USA.

Fæste, C. K., and Plassen, C. (2008). Quantitative sandwich ELISA for the determination of fish in foods. *J Immunol Meth.* **329**: 45-55.

Fæste, C. K., Rønning, H. T., Christians, U., and Granum, P. E. (2011). Liquid chromatography and mass spectrometry in food allergen detection. *J Food Protection.* **74**: 316-345.

FAO (2014). Food and Agriculture Organization of the United Nations, Rome, Italy. Available at: http://www.fao.org/ (Last accessed 3 December 2014).

González-de-Olano, D., Bartolomé, B., Maroto, A. S., Vivanco, F., and Pastor-Vargas, C. (2012). Asthma after chicken consumption due to cross-reactivity between fish and chicken parvalbumin. *J Invest Allergol Clin Immunol.* **22**: 227-228.

Grabarek, Z. (2006). Structural basis for diversity of the EF-hand calcium-binding proteins. *J Mol Biol.* **359**: 509-525.

Griesmeier, U., Vázquez-Cortés, S., Bublin, M., Radauer, C., Ma, Y., Briza, P., Fernández-Rivas, M., and Breiteneder, H. (2010). Expression levels of parvalbumins determine allergenicity of fish species. *Allergy: Eur J Allergy Clin Immunol.* **65**: 191-198.

Heizmann, C. W. (1984). Parvalbumin, an intracellular calcium-binding protein; distribution, properties and possible roles in mammalian cells. *Experientia.* **40**: 910-921.

Helbling, A., Haydel Jr, R., McCants, M. L., Musmand, J. J., El-Dahr, J., and Lehrer, S. B. (1999). Fish allergy: Is cross-reactivity among fish species relevant? Double- blind placebo-controlled food challenge studies of fish allergic adults. *Ann Allergy, Asthma, Immunol.* **83**: 517-523.

Herrero, B., Vieites, J. M., and Espiñeira, M. (2014). Development of an in-house fast real-time PCR method for detection of fish allergen in foods and comparison with a commercial kit. *Food Chem.* **151**: 415-420.

Hildebrandt, S. (2010). Multiplexed identification of different fish species by detection of parvalbumin, a common fish allergen gene: A DNA application of multi-analyte profiling (xMAP™) technology. *Anal Bioanal Chem.* **397**: 1787-1796.

InFormAll Database, communicating about Food Allergies. Available at: http://www.inflammation-repair.manchester.ac.uk/informAll/ (Last accessed 9 December 2014).

Ikura, M. (1996). Calcium binding and conformational response in EF-hand proteins. *Trends Biochem Sci.* **21**: 14-17.

Jeebhay, M. F., Robins, T. G., Miller, M. E., Bateman, E., Smuts, M., Baatjies, R., and Lopata, A. L. (2008). Occupational allergy and asthma among salt water fish processing workers. *Amer J Indust Med.* **51**: 899-910.

Jenkins, J. A., Breiteneder, H., and Mills, E. N. C. (2007). Evolutionary distance from human homologs reflects allergenicity of animal food proteins. *J Allergy Clin Immunol.* **120**: 1399-1405.

Joassin, L., and Gerday, C. (1977). The amino acid sequence of the major parvalbumin of the whiting (*Gadus merlangus*). *Comp Biochem Physiol B.* **57**: 159-161.

Kamath, S. D., Abdel Rahman, A. M., Komoda, T., and Lopata, A. L. (2013). Impact of heat processing on the detection of the major shellfish allergen tropomyosin in crustaceans and molluscs using specific monoclonal antibodies. *Food Chem.* **141***:* 4031-4039.

Koppelman, S. J., Nordlee, J. A., Lee, P. W., Happe, R. P., Hessing, M., Norland, R., Manning, T., Deschene, R., de Jong, G. A. H., and Taylor, S. L. (2012). Parvalbumin in fish skin-derived gelatin: Is there a risk for fish allergic consumers? *Food Addit Contam Part A Chem Anal Control Expo Risk Assess.* **29**: 1347-1355.

Kuehn, A., Hilger, C., Lehners-Weber, C., Codreanu-Morel, F., Morisset, M., Metz-Favre, C., Pauli, G., de Blay, F., Revets, D., Muller, C.P., Vogel, L., Vieths, S., and Hentges, F. (2013). Identification of enolases and aldolases as important fish allergens in cod, salmon and tuna: Component resolved diagnosis using parvalbumin and the new allergens. *Clin Exp Allergy*. **43**: 811-822.

Kuehn, A., Scheuermann, T., Hilger, C., and Hentges, F. (2010). Important variations in parvalbumin content in common fish species: A factor possibly contributing to variable allergenicity. *Int Arch Allergy Immunol.* **153**: 359-366.

Larsen, R., Eilertsen, K-E., and Elvevoll, E. O. (2011). Health benefits of marine foods and ingredients. *Biotechnol Adv.* **29**. 508–518.

Lee, P. W., Nordlee, J. A., Koppelman, S. J., Baumert, J. L., and Taylor, S. L. (2011). Evaluation and comparison of the species-specificity of 3 antiparvalbumin IgG antibodies. *J Agri Food Chem.* **59**: 12309-12316.

Lee, P. W, and Taylor SL. (2011). Fish, Crustaceans, and Mollusks. **In**: Food Allergens – Analysis Instrumentation and Methods, pp. 117-205. Hengel, A. J., Nollet, L. M. L., Eds., CRC Press, Taylor & Francis Group, Boca Raton.

Lindstrom, C. D. V., Van Do, T., Hordvik, I., Endresen, C., and Elsayed, S. (1996). Cloning of two distinct cDNAs encoding parvalbumin, the major allergen of Atlantic salmon (*Salmo salar*). *Scand J Immunol.* **44**: 335-344.

Lloris, D., Matallanas, J., and Oliver, P. (2005). Hakes of the world (Family Merlucciidae). An annotated and illustrated catalogue of hake species known to date. **In**: Species Catalogue for Fishery Purposes, pp 1-57. FAO, Rome.

Lu, Y., Ohshima, T., and Ushio, H. (20004). Rapid detection of fish major allergen parvalbumin by surface plasmon resonance biosensor. J Food Sci. **69**: C652-C658.

Ma., Y, Griesmeier, U., Susani, M., Radauer, C., Briza, P., Erler, A., Bublin, M., Alessandri, S., Himly, M., Vàzquez-Cortés, S., de Arellano, I. R., Vassilopoulou, E., Saxoni-Papageorgiou, P., Knulst, A. C., Fernández-Rivas, M., Hoffmann-Sommergruber, K., and Breiteneder, H. (2008). Comparison of natural and recombinant forms of the major fish allergen parvalbumin from cod and carp. *Mol Nutr Food Res.* **52**: 196-207.

Mafra, I., Ferreira, I. M. P. L. V. O., and Oliveira, M. B. P. P. (2008). Food authentication by PCR-based methods. *Eur Food Res Technol.* **227**: 649-665.

Mari, A., Rasi, C., Palazzo, P., and Scala, E. (2009). Allergen Databases: Current Status and Perspectives. *Curr Allergy Asthma Rep.* **9**: 376-383.

Minkiewicz, P., Bucholska, J., Darewicz, M., and Borawska, J. (2012). Epitopic hexapeptide sequences from Baltic cod parvalbumin beta (allergen Gad c 1) are common in the universal proteome. *Peptides.* **38**: 105-109.

Monaci, L., and Visconti, A. (2010). Immunochemical and DNA-based methods in food allergen analysis and quality assurance perspectives. *Trends Food Sci Tech*, **21**: 272–283.

Moraes, A. H., Ackerbauer, D., Kostadinova, M., Bublin, M., Ferreira, F., Almeida, F. C. L., Breiteneder, H., and Valente, A. P. (2013). 1H, 13C and 15N resonance assignments

and second structure information of Gad m 1: a β-parvalbumin allergen from Atlantic cod (*Gadus morhua*). *Biomolecular NMR Assignments.* **7**: 133-136.

Mozaffarian, D., Bryson, C. L., Lemaitre, R. N., Siscovick, D. S., and Burke, G. L. (2005). Fish intake and risk of incident heart failure. *J Amer Col Cardiol.* **45**: 2015-2021.

NCBI, National Center for Biotechnology Information, Bethesda, USA. Available at: http://www.ncbi.nlm.nih.gov/ (Last accessed 12 November 2014).

Nwaru, B. I., Hickstein, L., Panesar, S. S., Muraro, A., Werfel, T., Cardona, V., Dubois, A. E. J., Halken, S., Hoffmann-Sommergruber, K., Poulsen, L. K., Roberts, G., Van Ree, R., Vlieg-Boerstra, B. J., and Sheikh, A. (2014). The epidemiology of food allergy in Europe: A systematic review and meta-analysis. *Allergy: Eur J Allergy Clin Immunol.* **69:** 62-75.

Pechère, J. F., Capony, J. P., and Ryden, L. (1971). The primary structure of the major parvalbumin from hake muscle. Isolation and general properties of the protein. *Eur J Biochem.* **23**: 421-428.

Pegels, N., González, I., López-Calleja, I., García, T., Martín, R. (2013). Detection of Fish-Derived Ingredients in Animal Feeds by a TaqMan Real-Time PCR Assay. *Food Anal Methods.* **6:** 1040-1048.

Perez-Gordo, M., Cuesta-Herranz, J., Maroto, A. S., Cases, B., Ibáñez, M. D., Vivanco, F., and Pastor-Vargas, C. (2011). Identification of sole parvalbumin as a major allergen: Study of cross-reactivity between parvalbumins in a Spanish fish-allergic population. *Clin Exp Allergy.* **41**: 750-758.

Permyakov, S. E., Bakunts, A. G., Denesyuk, A. I., Knyazeva, E. L., Uversky, V. N., and Permyakov, E. A. (2008). Apo-parvalbumin as an intrinsically disordered protein. *Proteins: Struct, Funct, Genet.* **72**: 822-836.

Permyakov, E. A., Yarmolenko, V. V., Emelyanenko, V. I., Burstein, E. A., Closset, J., and Gerday, C. (1980). Fluorescence studies of the calcium binding to whiting (Gadus merlangus) parvalbumin. *Eur J Biochem.* **109**: 307-315.

Pilolli, R., Monaci, L., and Visconti, A. (2013). Advances in biosensor development based on integrating nanotechnology and applied to food-allergen management. *Trends Anal. Chem.* **47**: 12-26.

Piñeiro, C., Barros-Velázquez, J., Sotelo, C. G., Pérez-Martín, R. I., and Gallardo, J. M. (1998). Two-Dimensional Electrophoretic Study of the Water-Soluble Protein Fraction in White Muscle of Gadoid Fish Species. *J Agri Food Chem.* **46**: 3991-3997.

Plácido, A., Amaral, J. S., Costa, J., Fernandes, T. J. R., Oliveira, M. B. P. P., Delerue-Matos, C., Mafra, I. (2015). Novel Strategies for Genetically Modified Organism Detection. In Genetically Modified Organisms in Foods, pp. 119-131. Watson, R., and Preedy, V. R., Eds, Elsevier, USA.

Prado, M., Boix, A., von Holst, C. (2012). Novel approach for the simultaneous detection of DNA from different fish species based on a nuclear target: quantification potential. *Anal Bioanal Chem*. **403**: 3041–3050.

Prado, M., Ortea, I., Vial, S., Rivas, J., Calo-Mata, P., Barros-Velázquez, J. (2015). Advanced DNA- and protein-based methods for the detection and investigation of food allergens. *Crit. Rev. Food Sci. Nutr*. doi:10.1080/10408398.2013.873767

Priftis, K. N., Mermiri, D., Papadopoulou, A., Papadopoulos, M., Fretzayas, A., and Lagona, E. (2008). Asthma symptoms and bronchial reactivity in school children sensitized to food allergens in infancy. *J Asthma.* **45***:* 590-595.

Radauer, C., and Breiteneder, H. (2007). Evolutionary biology of plant food allergens. *J Allergy Clin Immunol.* **120**: 518-525.

Radauer, C., Bublin, M., Wagner, S., Mari, A., and Breiteneder, H. (2008). Allergens are distributed into few protein families and possess a restricted number of biochemical functions. *J Allergy Clin Immunol.* **121**: 847-852.e847.

Rancé, F., Kanny, G., Dutau, G., and Moneret-Vautrin, D. A. (1999). Food hypersensitivity in children: Clinical aspects and distribution of allergens. *Pediatr Allergy Immunol.* **10**: 33-38.

Regulation (EU) No 1169/2011 of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No 1924/2006 and (EC) No 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004. *Official Journal of the European Union*, *L304*, 18-63.

Rencova, E., Kostelnikova, D., and Tremlova, B. (2013). Detection of allergenic parvalbumin of Atlantic and Pacific herrings in fish products by PCR. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess.* **30**: 1679-1683.

Resource Program Allergen Database, Home of the FARRP allergen protein database. Available at http://www.allergenonline.com (Last accessed 9 December 2014).

Revett, S. P., King, G., Shabanowitz, J., Hunt, D. F., Hartman, K. L., Laue, T. M., and Nelson, D. J. (1997). Characterization of a helix-loop-helix (EF hand) motif of silver hake parvalbumin isoform B. *Protein Sci.* **6**: 2397-2408.

Richardson, R. C., King, N. M., Harrington, D. J., Sun, H., Royer, W. E., and Nelson, D. J. (2000). X-ray crystal structure and molecular dynamics simulations of silver hake parvalbumin (Isoform B). *Protein Sci.* **9**:73-82.

RCSB-PDB Research Collaboratory for Structural Bioinformatics-Protein Data Bank, an information portal to biological macromolecular structures. Available at http://www.rcsb.org (Last accessed 12 November 2014).

Sharp, M. F., and Lopata, A. L. (2013). Fish Allergy: In Review. *Clin Rev Allergy Immunol.* 1-14.

Sicherer, S. H., Morrow, E. H., and Sampson, H. A. (2000). Dose-response in double-blind, placebo-controlled oral food challenges in children with atopic dermatitis. *J Allergy Clin Immunol.* **105**: 582-586.

Sicherer, S. H., Muñoz-Furlong, A., and Sampson, H. A. (2004). Prevalence of seafood allergy in the United States determined by a random telephone survey. *J Allergy Clin Immunol.* **114***:* 159-165.

Sicherer, S. H., and Sampson, H. A. (2009). Food Allergy: Recent Advances in Pathophysiology and Treatment. *Ann Rev Med.* **60**: 261-277.

Sicherer, S. H., and Sampson, H. A. (2010). Food allergy. *J Allergy Clin Immunol.* **125**: 116-125.

Sirot, V., Oseredczuk, M., Bemrah-Aouachria, N., Volatier, J-L., and Leblanc, J-C. (2008). Lipid and fatty acid composition of fish and seafood consumed in France: CALIPSO study. *J Food Compos Anal.* **21**: 8-16.

Somkuti, J., Bublin, M., Breiteneder, H., and Smeller, L. (2012). Pressure-temperature stability, Ca^{2+} binding, and pressure-temperature phase diagram of cod parvalbumin: Gad m 1. *Biochem.* **51**: 5903-5911.

Sletten, G., Van Do, T., Lindvik, H., Egaas, E., and Florvaag, E. (2010). Effects of industrial processing on the immunogenicity of commonly ingested fish species. *Int Arch Allergy Immunol.* **151**: 223-236.

Sun, M., Liang, C., Gao, H., Lin, C., and Deng, M. (2009). Detection of parvalbumin, a common fish allergen gene in food, by real-time polymerase chain reaction. *J AOAC Int.* **92**: 234-240.

Taylor, S. L., & Baumert, J. L. (2015). Worlwide food allergy labeling and detection of allergens in processed foods. **In**: Food Allergy: Molecular Basis and Clinical Practice. Chemical Immunology and Allergy, vol 101, pp 227-234. Ebisawa M, Ballmer-Weber B. K., Vieths S., Wood R. A., Eds., Basel, Karger,.

Thong, B. Y. H., Cheng, Y. K., Leong, K. P., Tang, C. Y., and Chng, H. H. (2007). Immediate food hypersensitivity among adults attending a clinical immunology/allergy centre in Singapore. *Singapore Med J.* **48***:* 236-240.

Tsabouri, S., Triga, M., Makris, M., Kalogeromitros, D., Church, M. K., and Priftis, K. N. (2012). Fish and shellfish allergy in children: Review of a persistent food allergy. *Pediatr Allergy Immunol.* **23***:* 608-615.

UniProt, Protein knowledgebase, UniProt Consortium. Available at: http://www.uniprot.org/ (Last accessed 12 November 2014).

Van Do, T., Hordvik, I., Endresen, C., and Elsayed, S. (2003). The major allergen (parvalbumin) of codfish is encoded by at least two isotypic genes: cDNA cloning, expression and antibody binding of the recombinant allergens. *Mol Immunol.* **39**: 595-602.

Van Do, T., Elsayed, S., Florvaag, E., Hordvik, I., and Endresen, C. (2005a). Allergy to fish parvalbumins: Studies on the cross-reactivity of allergens from 9 commonly consumed fish. *J Allergy Clin Immunol.* **116**: 1314-1320.

Van Do, T., Hordvik, I., Endresen, C., and Elsayed, S. (2005b). Characterization of parvalbumin, the major allergen in Alaska pollack, and comparison with codfish Allergen M. *Mol Immunol.* **42**: 345-353.

Vierk, K. A., Koehler, K. M., Fein, S. B., and Street, D. A. (2007). Prevalence of self-reported food allergy in American adults and use of food labels. *J Allergy Clin Immunol.* **119**. 1504-1510.

Wang, J., and Sampson, H. A. (2011). Food allergy. *J Clin Invest.* **121**: 827-835.

Wang, W., Han, J., Wu, Y., Yuan, F., Chen, Y., and Ge, Y. (2011). Simultaneous detection of eight food allergens using optical thin-film biosensor chips. *J Agric Food Chem*. **59**: 6889-6894.

Weber, P., Steinhart, H., and Paschke, A. (2009). Competitive indirect ELISA for the determination of parvalbumins from various fish species in food grade fishgelatins and isinglass with PARV-19 anti-parvalbumin antibodies. *J Agri Food Chem.* **57**: 11328-11334.

Wopfner, N., Dissertori, O., Ferreira, F., and Lackner, P. (2007). Calcium-Binding Proteins and Their Role in Allergic Diseases. *Immunol Allergy Clin North Amer.* **27**: 29-44.

Zheng, C., Wang, X., Lu, Y., and Liu, Y. (2012). Rapid detection of fish major allergen parvalbumin using superparamagnetic nanoparticle-based lateral flow immunoassay. *Food Control.* **26**: 446-452.

Zuidmeer, L., Goldhahn, K., Rona, R. J., Gislason, D., Madsen, C., Summers, C., Sodergren, E., Dahlstrom, J., Lindner, T., Sigurdardottir, S. T., McBride, D., and Keil, T. (2008). The prevalence of plant food allergies: A systematic review. *J Allergy Clin Immunol.* **121**: 1210-1218.e1214.

EXPERIMENTAL PART

Exploiting 16S rRNA gene for the detection and quantification of fish as a potential allergenic food: A comparison of two real-time PCR approaches.
Food Chemistry, **2018**, 245, 1034-1041.

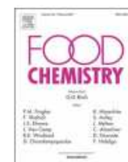
A new real-time PCR quantitative approach for the detection of shrimp crustaceans as potential allergens.
Journal of Food Composition and Analysis, **2017**, submitted.

Food Chemistry 245 (2018) 1034–1041

Contents lists available at ScienceDirect

Food Chemistry

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



Exploiting 16S rRNA gene for the detection and quantification of fish as a potential allergenic food: A comparison of two real-time PCR approaches

Telmo J.R. Fernandes, Joana Costa, M. Beatriz P.P. Oliveira, Isabel Mafra*

REQUIMTE-LAQV, Faculdade de Farmácia, Universidade do Porto, Rua de Jorge Viterbo Ferreira, 228, 4050-313 Porto, Portugal

*Corresponding author: Tel: +351 220428640. Fax: +351 226093390. E-mail: isabel.mafra@ff.up.pt

Abstract

Fish is one of the most common allergenic foods that should be accurately labelled to protect the health of allergic consumers. In this work, two real-time PCR systems based on the EvaGreen dye and a TaqMan probe are proposed and compared. New primers were designed to target the 16S rRNA gene, as a universal marker for fish detection, with fully demonstrated specificity for a wide range of fish species. Both systems showed similar absolute sensitivities, down to 0.01 pg of fish DNA, and adequate real-time PCR performance parameters. The probe system showed higher relative sensitivity and dynamic range (0.0001-50%) than the EvaGreen (0.05-50%). They were both precise, but trueness was compromised at the highest tested level with the EvaGreen assay. Therefore, both systems were successful, although the probe one exhibited the best performance. Its application to verify labelling compliance of foodstuffs suggested a high level of mislabelling and/or fraudulent practices.

Keywords: Fish detection; quantitative real-time PCR; TaqMan probe; EvaGreen dye; mitochondrial DNA, fish allergen.

Introduction

In geographical areas where fish has an important role in socioeconomic and cultural behaviours, such as Portugal, Spain and the Scandinavian countries, the prevalence of fish allergy also tends to be higher than in countries without a traditional strong fish industry (Calderon-Rodriguez, Pineda, Perez, & Muñoz, 2016;. Mourad & Bahna, 2015). Fish allergy is typically IgE-mediated with clinical features that can range from mild to potentially life-threatening (Mourad & Bahna, 2015; Sharp & Lopata, 2014; Thalayasingam & Lee, 2015). Common clinical symptoms of fish allergy are gastrointestinal (vomiting, diarrhoea and abdominal pain), cutaneous (urticaria or angioedema) and respiratory (rhinitis, conjunctivitis, cough and wheezing), although more severe symptoms, such as laryngeal oedema or anaphylaxis, might also occur (Mourad & Bahna, 2015; Sharp & Lopata, 2014; Thalayasingam & Lee, 2015). The main fish allergen is parvalbumin, a small calcium-binding muscle protein, although other proteins, such as enolases and aldolases are also responsible for fish allergy (Fernandes, Costa, Oliveira, & Mafra, 2015; Mourad & Bahna, 2015). The content in parvalbumin differs considerably among fish species. Cartilaginous fish muscle (dark tissues) contains mainly alpha-parvalbumin (non-allergenic protein), whereas bony fish muscle (white tissues) expresses its beta-homolog (allergenic protein), which explains their distinct potential of allergenicity (Kobayashi et al., 2006; Kuehn, Scheuermann, Hilger, & Hentges, 2010; Mourad & Bahna, 2015). Owing to differences in the amount of expressed allergenic parvalbumin in white or dark fish muscle, some sensitised/allergic patients tolerate one or more species of fish with less parvalbumin. Amongst the white fish species that are globally consumed, hakes (Merluccidae family) are one group of the most allergenic, representing a high proportion of food allergies, namely in children (Calderon-Rodriguez et al., 2016; Pascual, Esteban, & Crespo, 1992; Triantafyllidis et al., 2010). Due to cross-reactivity phenomena among different species, parvalbumins are responsible for the polysensitisation of allergic individuals to several fish species, who are advised avoiding their consumption (Fernandes et al., 2015; Triantafyllidis et al., 2010).

Despite the compulsory labelling of food products that may contain allergenic ingredients (Regulation (EU) No. 1169/2011), the allergic individuals can inadvertently consume the offending foods due to mislabelling or cross-contamination during processing. This circumstance emphasises the need of verifying labelling compliance, which should rely on highly sensitive, accurate, specific and fast methods to enable the detection of trace levels of allergens in foods. Currently, several methods targeting protein and DNA markers have been suggested to detect fish allergens. The most widely used approaches are the enzyme linked immunosorbent assay (ELISA) (Chen & Hsieh, 2014;

Faeste & Plassen, 2008; Weber, Steinhart, & Paschke, 2009) and the DNA amplification by real-time polymerase chain reaction (PCR) (Herrero, Vieites, & Espiñeira, 2014; Houhoula, Dimitriou, Mengjezi, Kyrana, & Lougovois, 2015; Sun, Liang, Gao, Lin, & Deng, 2009; Tetzlaff & Mäde, 2017). ELISA allows the direct detection and quantification of allergens, mainly parvalbumins in the case of fish, usually providing a rapid and simple method. PCR-based methods are able to detect marker sequences of the allergenic species, which can be allergen encoding genes, providing an indirect detection of the target (Fernandes, Costa, Carrapatoso, Oliveira, & Mafra, 2017a). However, immunochemical methods, such as ELISA, are particularly prone to cross-reactivity phenomena and to be affected by structural changes in proteins after food processing, often conducting to false positive or negative results. On the other hand, PCR methods take advantage of the higher stability of DNA markers compared to proteins, being adequate to detect allergens in processed and complex food matrices with high specificity and sensitivity (Fernandes et al., 2017a; Prado et al., 2016). A real-time PCR assay for the specific detection of parvalbumin gene in fish was developed by Sun et al. (2009), allowing a sensitivity of 5 pg. Herrero et al. (2014) proposed a real-time PCR method targeting a universal DNA marker of fish (18S rRNA gene), which detected fish DNA down to 50 pg and was successfully applied to commercial food samples. More recently, Tetzlaff and Mäde (2017) reported a real-time PCR assay targeting the *Hoxc13* gene to detect the Teleostei subclass of fish, reaching sensitivities of 10-100 mg/kg, depending on the matrix, and 2.5 DNA copies. Those real-time PCR reports are all based on the use of TaqMan probes targeting mostly nuclear genes. Mitochondrial DNA markers, such as the encoding genes of 16S rRNA, 12S rRNA, cytochrome c oxidase subunit I (COI) and cytochrome b (cyt b) have been frequently targeted in fish species differentiation for authentication purposes due to their high copy numbers, allowing for the universal primer design (Abdullah & Rehbein, 2016; Feng, Wu, Xie, Dai, & Liu, 2017; Fernandes, Costa, Oliveira, & Mafra, 2017b; Giusti et al., 2017; Pegels, González, López-Calleja, García, & Martín, 2013).

In the present work, the development of two real-time PCR approaches is proposed targeting the 16S rRNA mitochondrial region, as a universal marker for fish detection in foods at trace levels. The two approaches based of the universal EvaGreen dye and TaqMan probes are compared in terms of assay performance, validation and applicability to model mixtures with spiked fish. Finally, several commercial samples are analysed and the results compared with the label statements.

Materials and Methods

Sample collection and preparation

Several specimens of commercially relevant fish species available on the Portuguese market (*n*=26) were acquired: Atlantic cod (*Gadus morhua*), Pacific cod (*Gadus macrocephalus*), Alaska pollock (*Theragra chalcogramma*), saithe (*Pollachius virens*) and ling (*Molva molva*) were kindly provided by Pascoal & Filhos S.A; European hake (*Merluccius merluccius*), North Pacific hake (*Merluccius productus*), Argentine hake (*Merluccius hubsi*), whiting (*Merluccius capensis*), deepwater hake (*Merluccius paradoxus*) were acquired at local markets and authenticated by sequencing (Fernandes, Costa, Oliveira, & Mafra, 2018); European pilchard (*Sardina pilchardus*), ray (*Raja* spp.), Atlantic horse mackerel (*Trachurus trachrurus*), gilt-head bream (*Sparus aurata*), common sole (*Solea solea*) and Pacific mackerel (*Scomber japonicus*) were acquired at local markets as fresh whole specimens, allowing morphological identification; Atlantic salmon (*Salmo salar*), yellowfin tuna (*Thunus albacares*), European seabass (*Dicentrarchus labrax*), whiting-pout (*Trisopterus luscus*), tadpole codling (*Saliota australis*), rose fish (*Sebastes marinus*), rock ling (*Genypterus blacodes*), black scabbardfish *(Aphanopus carbo*), Nile perch (*Lates niloticus*) and sutchi catfish (*Pangasius hipophthalmus*) were acquired at local markets as frozen specimens. Other non-fish species were also acquired for reactivity testing of the assay (*n*=40), namely crustaceans and molluscs (squid, mussel, undulated surf clam, crab, edible crab, Caribbean spiny lobster, Norway lobster, whiteleg shrimp, giant tiger prawn, Indian white prawn, jinga shrimp and striped shrimp), meat species (cow, chicken, rabbit, deer, turkey, pork, lamb, goat, ostrich) and plant species (soybean, potato, oat, rye, wheat, lupine, maize, rice, pumpkin seeds, rapeseed, sunflower, tomato, peach, nectarine, cassava, strawberry, raspberry, honey, cashew).

Several processed seafood products (fish sticks (*n*=12), breaded fish fillets (*n*=2), powder soups (*n*=3), fish patties/pies (*n*=6), surimi (*n*=2), pre-cooked dishes (*n*=9), among others) were also acquired at local markets and analysed to verify assay applicability.

All the specimens and samples were individually ground and homogenised in a laboratory knife mill Grindomix GM200 (Retsch, Haan, Germany) and stored at -20 ºC until additional analysis. All containers and material used during this procedure were previously treated with DNA decontamination solution (DNA-ExitusPlus™, AppliChem, Darmstadt, Germany).

Preparation of model mixtures and samples

The preparation of binary model mixtures was attempted to simulate the stuffing/filling of processed foods, such as patties and pies. Hake fillets were boiled in water during 15

min, drained and ground in a laboratory knife mill Grindomix GM200 (Retsch, Haan, Germany). The first mixture containing 50% (w/w) of ground-cooked fish in béchamel sauce was prepared by the addition of 200 g of fish to the same amount of sauce. Successive additions of béchamel sauce to the 50% homogenised mixture enabled the preparation of the following binary mixtures: 10, 5, 1, 0.5, 0.1, 0.05, 0.01, 0.005, 0.001, 0.0005 and 0.0001% (w/w) of fish in béchamel. For method validation, blind mixtures were similarly prepared, in order to contain 8.0, 4.0, 2.5 and 0.25% (w/w) of fish in béchamel.

All the binary mixtures were also ground and homogenised in a laboratory knife mill and stored at -20 ºC until further analysis. All containers and material were previously treated with DNA decontamination solution (DNA-ExitusPlus™, AppliChem, Darmstadt, Germany).

Oligonucleotide design

Sequences of the 16S rRNA gene were retrieved from NCBI database from a set of 24 different fish species (Figure S1, supplementary material). The majority of these species match with the ones most frequently consumed as food. However, it was not always possible to find available sequences on the web for all fish species commonly marketed, and vice-versa. Sequence alignments were performed with BioEdit v.7.2.5 software (Ibis Biosciences, Canada) and examined for suitable regions for primer and probe design. Accordingly, primers (16SFish1-F – 5'-AACGGCCGCGGTATTTTAAC-3'/16SFish1-R – 5'-TGCACGGGGAGGTCAATTT-3') and probe (16SFish-P – FAM-5'-CTTTTAAATGAAGACCTGTATGAATGGCAT-3'-BHQ1) were designed and submitted to a basic local alignment search tool BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) for the in silico analysis of primer specificity. Primer properties, self-hybridisation and the absence of hairpins were evaluated using the software OligoCalc (http://www.basic.northwestern.edu./biotools/oligocalc.html) and, finally, the oligonucleotides were synthesised by STABVIDA (Lisbon, Portugal).

DNA Extraction

The DNA was extracted using the Nucleospin® Food kit (Macherey-Nagel, Düren, Germany) with minor modifications. All DNA extractions were done in duplicate assays using 200 mg of each sample and 2 µL of RNAse (2 mg/mL) that were added for 5 min at room temperature immediately after the lysis step. All the extracts were kept at -20 °C until further analysis.

The concentration and purity of the extracts were assessed by UV spectrophotometric DNA quantification on a Synergy HT multi-mode microplate reader (BioTek Instruments,

Inc., Vermont, USA), using a Take3 micro-volume plate accessory. DNA content was determined using the nucleic acid quantification protocol with sample type defined for double-strand DNA in the Gen5 data analysis software version 2.01 (BioTek Instruments, Inc., Vermont, USA). The quality of extracted DNA was also analysed by electrophoresis in a 1.0% agarose gel containing Gel Red 1× (Biotium, Hayward, CA, USA) for staining and carried out in STGB 1× (GRISP, Porto, Portugal) for 25 min at 200 V. The agarose gel was visualised under a UV light tray Gel Doc™ EZ System (Bio-Rad Laboratories, Hercules, CA, USA) and a digital image was obtained with Image Lab software version 5.1 (Bio-Rad Laboratories, Hercules, CA, USA).

Qualitative PCR

Amplifications by qualitative PCR were carried out in 25 µL of total reaction volume containing 2 µL of DNA extract (100 ng), 67 mM of Tris-HCl (pH 8.8), 16 mM of $(NH_4)_2SO_4$, 0.01% of Tween 20, 3.0 mM of $MgCl_2$, 1.0 U of SuperHot Taq DNA Polymerase (Genaxxon Bioscience, Ulm, Germany), 200 µM of each dNTP (Bioron, Ludwigshafen, Germany) and 200 nM of each primer (16SFish1-F/16SFish1-R). The PCR runs were performed in a MJ Mini™ Gradient Thermal Cycler (Bio-Rad Laboratories, Hercules, CA, USA) using the following cycling program: initial denaturation at 95 ºC for 5 min; 40 cycles at 95 ºC for 30 s, 57 ºC for 30 s and 72 ºC for 30 s; and a final extension at 72ºC for 5 min. The amplicons were then analysed by electrophoresis in a 1.5% agarose gel containing 1× Gel Red (Biotium, Hayward, CA, USA) for staining and carried out in 1× SGTB buffer (GRISP, Porto, Portugal) for about 20-25 min at 200 V. The agarose gel was visualised under a UV light tray Gel Doc™ EZ System (Bio-Rad Laboratories, Hercules, CA, USA) and a digital image was obtained with Image Lab software version 5.1 (Bio-Rad Laboratories, Hercules, CA, USA).

Real-time PCR

Two different real-time PCR approaches were assayed using EvaGreen dye and melting curve analysis and a TaqMan probe. Real-time PCR assays with Evagreen were carried out in 20 µL of total reaction volume containing 2 µL of DNA, 1× SsoFast™ Evagreen® Supermix (Bio-Rad Laboratories, Hercules, CA, USA) and 300 nM of each primer (16SFish1-F/16SFish1-R). The real-time PCR runs were performed on a fluorometric thermal cycler CFX96 Real-time PCR Detection System (Bio-Rad Laboratories, Hercules, CA, USA) using the following conditions: 95 ºC for 5 min; 45 cycles at 95 ºC for 10 s, 57 ºC for 15 s and 72 ºC for 20 s, with collection of fluorescence signal at the end of each cycle. Data were processed using the software Bio-Rad CFX

Manager 3.1 (Bio-Rad Laboratories, Hercules, CA, USA). For melting curve analysis, PCR products were denatured at 95 ºC for 1 min and then annealed at 60 ºC for 3 min, for the correct formation of DNA duplexes. These two steps were followed by melting curve ranging from 60 ºC up to 90 ºC with temperature increments of 0.2 ºC per 10 s. The fluorescence data were acquired at the end of each melting temperature. For the real-time PCR amplifications with a TaqMan probe, the 20 µL of the reaction mixture included 2 µL of DNA extract, 1× SsoFast Probes Supermix (Bio-Rad Laboratories, Hercules, CA, USA), 300 nM of each primer (16SFish1-F/16SFish1-R) and 200 nM of probe (16SFish1-P). The following temperature conditions were used: 95 ºC for 5 min, 50 cycles at 95 ºC for 15 s and 60 ºC for 45 s, with collection of fluorescence signal at the end of each cycle. Bio-Rad CFX Manager 3.1 software was used to process and analyse the fluorescence data. Cycle of quantification (Cq), also known as cycle threshold (Ct), values were calculated using the software at automatic threshold settings.

Results and Discussion

Specificity

In the present work, a sequence of the 16S rRNA gene was selected as a marker for the detection of a wide range of fish species, including those most commonly consumed. For that purpose, an exhaustive *in silico* study was performed to identify homologue regions and enable the design of universal primers for fish (Figure S1, supplementary material). To confirm, experimentally, the specificity of the primers, several relevant fish species were tested by qualitative PCR with the newly designed primers (16SFish1-F/16SFish1-R). Table 1 shows that all the 26 fish species were amplified with the proposed set of primers, producing the expected fragment of 131 bp (Figure S2A, supplementary material). The amplification of DNA from the binary mixtures showed that the optimised fish-specific PCR assay allow reaching a sensitivity down to 0.0001% (w/w) of cooked fish in béchamel (Figure S3, supplementary material). The specificity of primers was further demonstrated using the DNA from several non-target species of crustaceans, molluscs, mammals, avian and plants, all relevant as foods. No cross-reactivity was obtained with the non-target species, thus confirming the specificity of the designed primers for fish detection (Figure S2B, supplementary material). Prior to the specific amplification, all DNA extracts were tested for their amplifiability with universal eukaryotic primers (18SRG-F/18SRG-R), as described by Costa, Oliveira and Mafra (2013). All the fish (Table 1), the other animal and the plant species produced the expected PCR fragment (113 bp) with the universal primers, confirming the absence of any false negative result.

Table 1. Qualitative PCR results for eukaryotic and fish testing and cycle threshold values for several fish species with real-time PCR amplification with a TaqMan probe.

Species	Scientific name	Qualitative PCR[a]		Real-time PCR
		Eukaryotic (18S rRNA gene)	Fish (16S rRNA gene)	Ct ± SD[b]
Atlantic cod	*Gadus morhua*	+	+	18.81 ± 0.09
Pacific cod	*Gadus macrocephalus*	+	+	20.59 ± 0.26
Alaska pollock	*Theragra chalcogramma*	+	+	20.67 ± 0.58
Saithe	*Pollachius virens*	+	+	18.94 ± 0.08
Ling	*Molva molva*	+	+	18.86 ± 0.16
European hake	*Merluccius merluccius*	+	+	18.31 ± 0.03
Argentine hake	*Merluccius hubbsi*	+	+	23.63 ± 0.09
North Pacific hake	*Merluccius productus*	+	+	23.75 ± 0.27
Deepwater hake	*Merluccius paradoxus*	+	+	22.30 ± 0.01
Whiting	*Merluccius capensis*	+	+	18.62 ± 0.13
Atlantic salmon	*Salmo Salar*	+	+	21.19 ± 0.12
European pilchard	*Sardina pilchardus*	+	+	29.43 ± 0.86
Yellowfin tuna	*Thunnus albacares*	+	+	18.72 ± 0.21
Gilt-head bream	*Sparus aurata*	+	+	21.04 ± 0.03
European seabass	*Dicentrarchus labrax*	+	+	19.56 ± 0.16
Whiting-pout	*Trisopterus luscus*	+	+	21.30 ± 0.14
Pacific mackerel	*Scomber japonicus*	+	+	20.89 ± 0.37
Atlantic horse mackerel	*Trachurus trachurus*	+	+	22.44 ± 0.12
Rose fish	*Sebastes marinus*	+	+	19.79 ± 0.10
Black scabbardfish	*Aphanopus carbo*	+	+	23.32 ± 0.45
Nile perch	*Lates niloticus*	+	+	20.30 ± 0.11
Common sole	*Solea solea*	+	+	21.78 ± 0.11
Ray	*Raja* spp.	+	+	21.87 ± 0.42[c]
Tadpole codling	*Salilota australis*	+	+	19.28 ± 0.19[c]
Rock ling	*Genypterus blacodes*	+	+	22.46 ± 0.30[c]
Sutchi catfish	*Pangasius hipophthalmus*	+	+	19.24 ± 0.08[c]

[a] (+) positive qualitative PCR amplification; [b] Mean values of cycle threshold (Ct) ± standard deviation (SD) of replicate assays (*n*=4); [c]These values were obtained by real-time PCR with EvaGreen dye.

Real-time PCR with EvaGreen dye

DNA binding dyes allow the acquisition of high fluorescence upon intercalation into double-stranded DNA (dsDNA) or by binding to the minor grooves of dsDNA. This type of binding chemistry enables the quantitative estimation of the DNA target due to the increase of the total fluorescence signal in direct proportion to the DNA quantity in the sample (Salihah, Hossain, Lubis, & Ahmed, 2016). In this work, the enhanced fluorescent EvaGreen dye was used for method development and further compared with a TaqMan probe system.

The performance of the real-time PCR system with EvaGreen dye was analysed using a serially diluted fish-DNA extract and the binary model mixtures with fish to estimate both absolute and relative sensitivities, respectively. The guidelines of "Minimum Information for Publication of Quantitative Real-Time PCR Experiments" (Bustin et al., 2009) and the requirements set by the European Network of GMO Laboratories (ENGL) document "Definition of minimum performance requirements for analytical methods of GMO testing" (ENGL, 2015) were carefully considered for real-time PCR method development.

Table 2 presents the average results of two independent assays for each set of standards, while Figure 1 shows an example assay with the amplification, calibration and melting curves. For the determination of the absolute sensitivity, a 10-fold serially diluted fish (hake) DNA extract (100,000 pg to 0.001 pg) was used to establish the dynamic range, which allowed covering the amplification of 8 orders of magnitude of the target analyte (Table 2, Figure 1A, 1C). The limit of detection (LOD) was considered as the lowest concentration level with amplification at least 95% of the times. Accordingly, the LOD was established as 0.01 pg of fish DNA, which was also considered as the limit of quantification (LOQ) since it was within the linear range of the calibration curve. The average parameters of PCR efficiency (110.6%), slope (-3.092) and R^2 (0.998) were all within the acceptance criteria, suggesting a high performance of the assay (Table 2) (Bustin et al., 2009, ENGL, 2015). The presence of melt peaks with the same temperature (81.0 ºC) is indicative of one single group of PCR products and the absence of primer dimer or other non-specific fragments (Figure 1C).

The use of binary mixtures as reference standards with known amounts of spiked fish enabled to establish a distinct dynamic range (0.05-50%) and the relative LOD. In this case, a calibration curve with 4 orders of magnitude was obtained, with LOD and LOQ down to 0.05% (500 mg/kg) (Table 2, Figure 1B, 1D) and adequate analytical performance parameters (PCR efficiency of 96.3%, slope of -3.415 and R^2 of 0.999). According to the melting curve analysis, the amplification of fragments with similar melt peaks (81.0 ºC) was obtained until the level of 0.05%. After which, other groups of melt peaks at 79.8-80.0 ºC appeared, suggesting the presence of non-specific products (Figure 1F, blue and pink lines).

Real-time PCR with TaqMan probe

In real-time PCR, the use of specific fluorescent probes, such as the hydrolysis TaqMan probes, has the advantage of conferring increased specificity to the assay, without needing melting curve analysis. For comparative purposes, the performance of the

Table 2. Real-time PCR results with EvaGreen dye and TaqMan probe for the absolute and relative detection of fish targeting the 16S rRNA gene

	Cycle threshold ± SD[a]	
	EvaGreen dye	TaqMan probe
Absolute amount (pg fish DNA)		
100 000	12.93 ± 0.31 (16/16)	14.78 ± 0.19 (16/16)
10 000	15.79 ± 0.35 (16/16)	18.08 ± 0.12 (16/16)
1 000	18.83 ± 0.32 (16/16)	21.09 ± 0.14 (16/16)
100	22.06 ± 0.28 (16/16)	24.33 ± 0.27 (16/16)
10	25.17 ± 0.28 (16/16)	27.36 ± 0.17 (16/16)
1	28.36 ± 0.29 (16/16)	30.73 ± 0.22 (16/16)
0.1	31.56 ± 0.35 (16/16)	33.56 ± 0.36 (16/16)
0.01	34.15 ± 0.81 (16/16)	37.23 ± 0.91 (16/16)
0.001	43.25 ± 0.29 (2/16)	nd[b]
PCR efficiency (%)	110.6 ± 1.9	107.4 ± 1.3
Correlation coefficient (R^2)	0.998 ± 0.002	0.999 ± 0.001
Slope	-3.092 ± 0.04	-3.156 ± 0.028
Relative amount (%, w/w of fish in béchamel)		
50	16.84 ± 0.11 (8/8)	17.84 ± 0.40 (8/8)
10	19.01 ± 0.07 (8/8)	19.53 ± 0.77 (8/8)
5	20.21 ± 0.09 (8/8)	21.21 ± 0.13 (8/8)
1	22.64 ± 0.11 (8/8)	23.57 ± 0.17 (8/8)
0.5	23.81 ± 0.90 (8/8)	24.22 ± 0.17 (8/8)
0.1	26.02 ± 0.37 (8/8)	26.53 ± 0.64 (8/8)
0.05	27.11 ± 0.10 (8/8)	27.80 ± 0.18 (8/8)
0.01	ns[c]	31.29 ± 0.54 (8/8)
0.005	ns	32.26 ± 0.34 (8/8)
0.001	ns	34.61 ± 0.55 (8/8)
0.0005	ns	35.88 ± 0.57 (8/8)
0.0001	ns	36.67 ± 0.99 (8/8)
0	ns	nd
PCR efficiency (%)	96.3 ±1.6	90.3 ± 0.6
Correlation coefficient (R^2)	0.999 ± 0.001	0.987 ± 0.004
Slope	-3.415 ± 0.042	-3.578 ± 0.018

[a] Mean values ± standard deviation (SD) of replicates (in brackets are the positive amplifications/total replicates) obtained in two independent assays; [b] nd, not detected; [c] ns, non-specific amplification.

real-time PCR system with a TaqMan probe was analysed using a serially diluted fish-DNA extract and the binary model mixtures with fish to estimate both absolute and relative sensitivities, respectively, as made in the previous system. The performance results of the probe system are presented in Table 2 and Figure 2. Regarding the amplification of the serially diluted DNA, the results are very similar to the obtained in the previous EvaGreen system. Despite a delayed amplification of about 2 cycles in the probe system, the parameters of dynamic range (8 orders of magnitude) and sensitivity were identical (LOD=LOQ=0.01 pg) (Figure 2A, 2C). The other parameters suggested again a high

performance of the assay (PCR efficiency of 107.4%, slope of -3.156 and R^2 of 0.999) (Table 2).

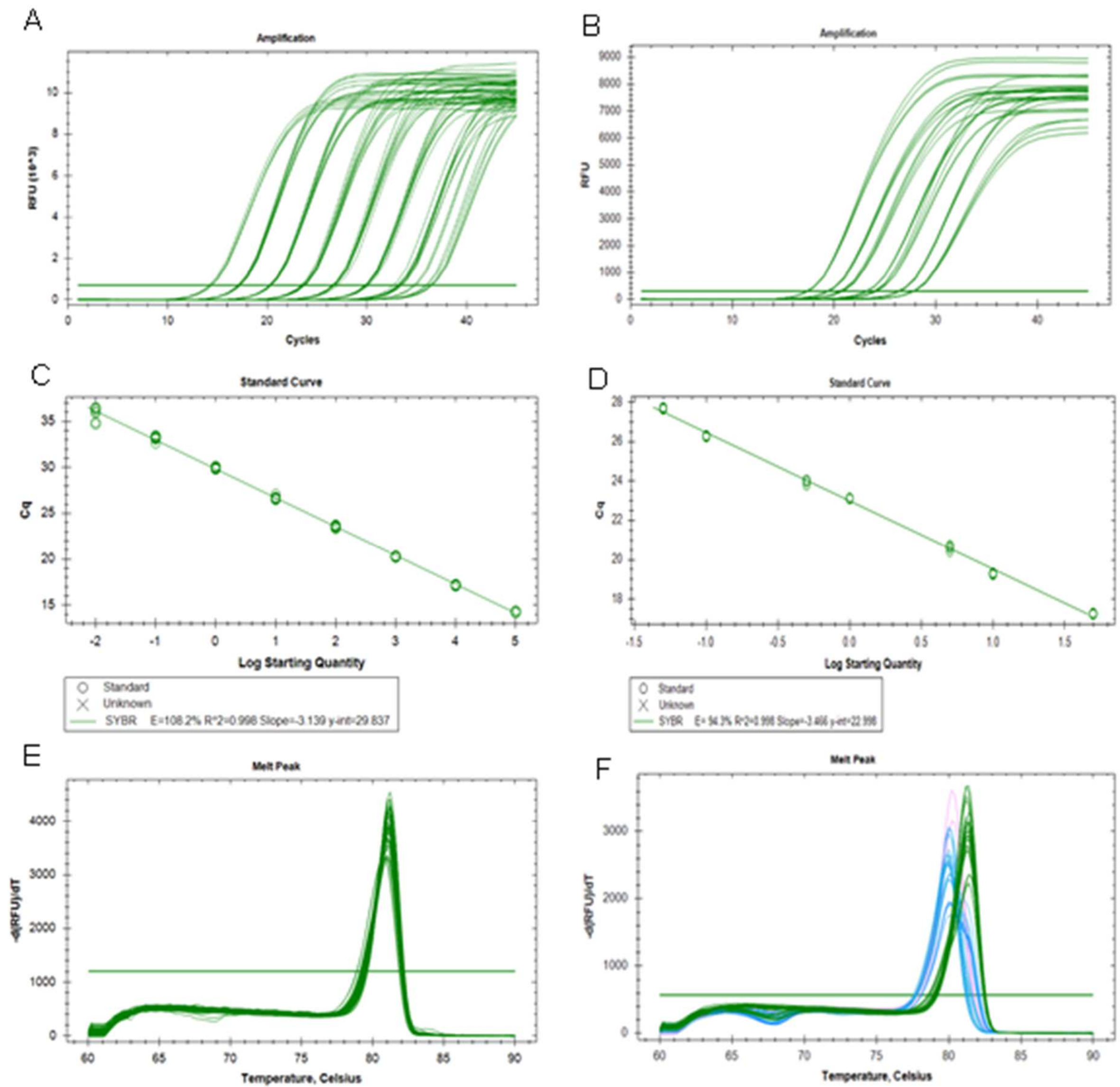


Figure 1. Real-time PCR amplification (A, B), calibration (C, D) and melting (E, F) curves with EvaGreen dye targeting the 16S rRNA of fish using a serially diluted fish-DNA extract (100,000-0.001 pg) (A, C, E) and DNA extracts from binary model mixtures of cooked fish in béchamel sauce (50-0.001% (w/w)) (B, D, F). Non-specific melt curves are presented in blue and pink lines (F) from 0.01% to 0.0001% and 0%, respectively.

In opposition, the results for the relative sensitivity were quite improved with the probe system since both LOD and LOQ were reduced to 0.0001% (w/w) of fish in béchamel sauce and the dynamic range was increased to 6 orders of magnitude with high performance assay parameters (PCR efficiency of 90.3%, slope of -3.578 and R^2 of 0.987)

(Table 2, Figure 2B, 2D). This sensitivity (1 mg of fish per kg of food) can be considered an adequate level regarding the threshold dose (5 to 6000 mg of fish) for eliciting an allergic response to fish (Taylor et al., 2002).

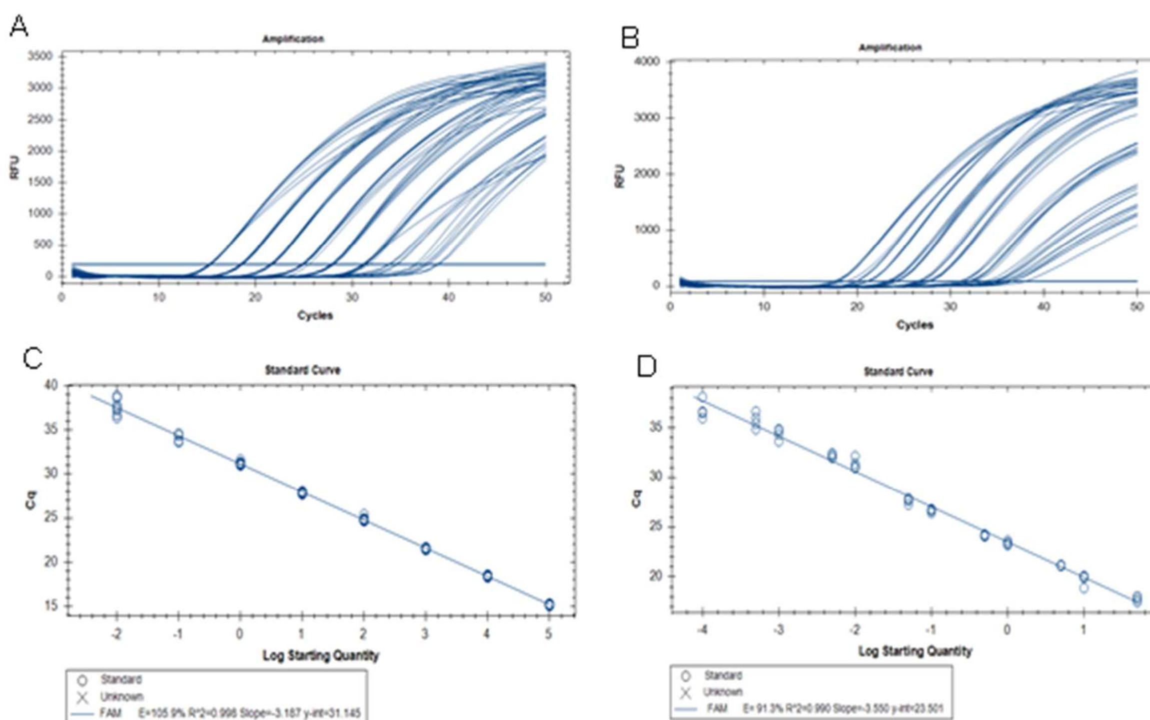


Figure 2. Real-time PCR amplification (A, B) and calibration (C, D) curves with a TaqMan probe targeting the 16S rRNA of fish using a serially diluted fish-DNA extract (100,000-0.001 pg) (A, C) and DNA extracts from binary model mixtures of cooked fish in béchamel sauce (50-0.001% (w/w)) (B, D).

In summary, the proposed real-time PCR with the probe enabled absolute and relative LOD of 0.01 pg of DNA and 0.0001% (1 mg/kg) of fish in béchamel sauce, which were considerably lower than the values reported in literature. Benedetto, Abete and Squadrone (2011) and Pegels et al. (2013), who also targeted mitochondrial DNA using real-time PCR with probe systems, reached sensitivities of 0.2 pg and 0.1 pg of fish DNA, respectively. Comparing the performance of both proposed systems, it can be inferred that the probe system was the best. Despite the same absolute sensitivities (0.01 pg) obtained by both approaches with adequate real-time PCR performance parameters, the probe system showed higher relative sensitivity and dynamic range (0.0001-50%) than the EvaGreen (0.05-50%).

Method validation

Considering that the specificity of the primers was previously demonstrated by qualitative PCR, it is also important to verify the specificity of the probe. Accordingly, all fish species were amplified by real-time PCR with the TaqMan probe, confirming most results with Ct values ranging from 18 to 23 (Table 1). The amplification with TaqMan probe was negative in four fish species (*Raja* spp*., Salilota australis, Genypterus blacodes, Pangasius hypophthalmus*). This might be due to nucleotide differences in the probe region, which can be observed in the cases of *G. blacodes* and *P. hypophthalmus* sequences (Figure S1, supplementary material). To overcome this finding, the four species were further successfully amplified using the EvaGreen assay (Table 1).

For the validation of the quantitative assays, both EvaGreen and probe systems were tested for their aptitude to estimate the fish content using blind mixtures with fish contents of 8%, 4%, 2.5% and 0.25% (w/w) (Table 3). The comparison between the estimations and the actual values suggests that both real-time PCR systems demonstrate precision over the tested dynamic range (⩽25%) since the coefficients of variation varied between 3.31% and 15.2% (ENGL, 2015). The measured trueness, expressed as bias, was generally higher for the estimates with the EvaGreen than for those with the probe system, but still within ±25% of the actual value (ENGL, 2015) for all the tested levels, except for the highest one (8%) determined by the EvaGreen system that exhibited 45.5% of bias. This means that the probe system can provide accurate determinations in all tested range (0.25-8%), while in the EvaGreen system the range is lower (0.25-4%). Therefore, it can be inferred that both real-time PCR approaches were successfully developed and validated, although the TaqMan probe exhibited better performance than the EvaGreen in terms of sensitivity, dynamic range and trueness.

Table 3. Results for the validation of the quantitative real-time PCR assays with EvaGreen dye and TaqMan probe using blind samples

Real-time PCR assay	Fish content (%, w/w)		CV (%)[b]	Bias (%)[c]
	Actual	Estimated (mean ± SD [a])		
EvaGreen dye	8.0	11.64 ± 1.43	12.3	45.5
	4.0	3.27 ± 0.24	7.28	-9.14
	2.5	2.80 ± 0.09	3.31	3.75
	0.25	0.23 ± 0.01	6.31	-0.29
TaqMan probe	8.0	9.52 ± 1.45	15.2	19.0
	4.0	3.60 ± 0.46	12.7	-5.02
	2.5	2.64 ± 0.20	7.52	1.80
	0.25	0.26 ± 0.02	8.33	0.10

[a] Mean values ± standard deviation (SD) of replicates (*n* = 8) obtained in two independent real-time PCR assays; [b] CV, coefficient of variation; [c] Bias = ((mean estimated value-true value)/true value x 100).

Analysis of processed food samples

Based on the validation results and on the previous data of assay performance, the real-time PCR system with the probe was considered the most adequate for fish allergen detection and quantification due to its higher sensitivity and dynamic range for quantification compared with the EvaGreen dye assay. For that reason, the real-time PCR assay with probe was selected to be further applied in the analysis of processed foods.

Concerning the application of the method to processed foods, the results of the 34 analysed samples are presented in Table 4. It can be noticed that most foods indicated the percentage of fish as a food component on the label. With the exception of sample #6 (Paella, "it contains fish"), all the samples were positive with the 16S rRNA primers in qualitative PCR and real-time PCR, which supports the fact that the primers are universal markers for the presence of most fish species found in processed foods. The Ct values ranged from 20.72 to 32.66, which allowed estimating the fish content based on the developed calibration curve using the binary mixtures (Figure 2D). The estimated contents varied from 0.02 to 48.5%, but they were, in general, much lower than the labelled values, which questions the veracity of the claimed information to consumers. In samples #9 (Surimi), #19, #20 (soups) and #28, #29 (fish sticks), the estimated fish contents are in close relation with the labelled information. In the samples #6 and #17, fish was not detected or found at trace levels, which is in good agreement with the type of foods that are not supposed to contain fish, whose labels should state the precautionary labelling "May contain fish" instead of "It contains fish". In opposition, in sample #23, with an estimated fish content of 7.07% and labelled as "May contain fish", fish should only be present at trace levels and not as an ingredient. These findings highlight inaccurate labelling of samples #6 and #17, and adulteration practices by partial or total replacement of shrimp by fish in sample #23. The remaining samples, with estimated contents much lower than the labelled, also suggest adulteration practices by partial substitution of fish by other ingredients.

The application of other real-time PCR methods in the detection of fish in food samples has been reported. The method of Herrero et al. (2014) was applied to processed food analysis, but only at the level of detection without quantification. The methods of Benedetto et al. (2011) and Pegels et al. (2013), although not focusing on fish as a potential allergen, but on feed authentication, were effectively applied to quantitative analysis. Benedetto et al. (2011) attempted a normalised method with binary mixtures of fish in plant material, but they did not reach an accurate quantification. The authors highlighted the need for a reliable endogenous reference for the accurate quantitative determination of fish. Pegels et al. (2013) successfully applied the assay to the authentication of industrial farm animal feeds and commercial pet foods, and exploited the

quantitative potential with reproducible results. However, they were not able to obtain accurate estimates. The matrix, the DNA extraction method and the level of processing can greatly affect the DNA yield, integrity and purity, which are crucial factors that influence the quantitative real-time PCR performance.

Table 4. Analysed processed food samples with relevant labelled information and results of qualitative PCR and real-time PCR targeting the 16S rRNA gene of fish species

Sample		Relevant label information	PCR[a]	Real-time PCR[b]	
				Ct ± SD	Estimated fish content (%, w/w)
1	Codfish “patanisca”	Codfish 18%	+	22.94 ± 0.13	6.74 ± 0.57
2	“Brás style” codfish	Codfish 18%	+	22.56 ± 0.16	8.64 ± 0.91
3	“Gomes de Sá” style codfish	Codfish 20%	+	22.44 ± 0.02	13.43 ± 0.18
4	Codfish “alheira”	Codfish 34%	+	24.67 ± 0.29	3.47 ± 0.63
5	Codfish lasagne	Codfish 47%	+	24.64 ± 0.04	3.83 ± 0.09
6	Paella	It contains fish	-	-	
7	Codfish with cream	Codfish 11%	+	24.42 ± 0.17	4.38 ± 0.44
8	Codfish with cream	Codfish 18%	+	24.72 ± 0.06	2.14 ± 0.08
9	Surimi	Alaska pollock 49.5%	+	20.36 ± 0.02	48.51 ± 1.56
10	Surimi	Alaska pollock 49%	+	22.00 ± 0.14	19.21 ± 1.66
11	Seafood risotto	Hake 8%. Fish sauce 1%	+	26.02 ± 0.13	0.93 ± 0.08
12	Codfish pies	Codfish 23%	+	24.13 ± 0.07	5.21 ± 0.21
13	Hake/shrimp patties	Filling 42.5% (containing hake)	+	25.21 ± 0.11	2.71 ± 0.18
14	Hake patties	Hake 11%	+	23.95 ± 0.03	3.51 ± 0.07
15	Codfish patties	Codfish 26%	+	24.47 ± 0.02	2.51 ± 0.03
16	Codfish patties	Codfish 30%	+	25.04 ± 0.19	2.75 ± 0.33
17	Shrimp patties	It contains fish	+	32.66 ± 0.04	0.02 ± 0.00
18	Seafood soup	Fish 4.1%	+	30.07 ± 0.11	0.14 ± 0.01
19	Seafood soup	Fish 4.1%	+	23.91 ± 0.05	5.46 ± 0.14
20	Seafood soup	Fish 5.2%	+	24.09 ± 0.05	4.90 ± 0.14
21	Breaded fish	Alaska pollock 56%	+	21.40 ± 0.06	18.23 ± 0.69
22	Breaded fish	Alaska pollock 70%	+	20.81 ± 0.04	29.35 ± 0.79
23	Shrimp stick	May contain fish	+	23.63 ± 0.05	7.07 ± 0.21
24	Fish sticks	Fish 65%	+	22.01 ± 0.09	19.09 ± 1.05
25	Fish sticks	White fish 65%	+	21.76 ± 0.13	22.27 ± 1.80
26	Fish sticks	White fish 48%	+	24.57 ± 0.07	2.35 ± 0.11
27	Fish sticks	Hake 50%	+	21.39 ± 0.06	18.39 ± 0.61
28	Fish sticks	Hake 55%	+	20.72 ± 0.13	31.11 ± 2.41
29	Fish sticks	Hake 55%	+	20.94 ± 1.08	39.66 ± 0.87
30	Fish sticks	Hake 54%	+	22.15 ± 0.05	13.07 ± 0.43
31	Fish sticks	Hake 60%	+	21.20 ± 0.02	23.24 ± 0.31
32	Codfish sticks	Codfish 58%	+	22.33 ± 0.10	11.72 ± 0.68
33	Fish sticks	Alaska pollock 65%	+	25.34 ± 0.03	1.86 ± 0.04
34	Fish sticks	Hake 55%	+	22.20 ± 0.06	14.08 ± 0.50

[a] (+) positive qualitative PCR amplification, (-) non-detected amplification; [b] Mean values ± standard deviation (SD) of replicates (n = 8) obtained in two independent real-time PCR assays.

In the present work, the proposed method was effectively applied to verify labelling compliance of processed fish-containing foods, being able to detect fish at trace levels, required as a food allergen, and to obtain estimates that could be successfully validated with blind samples. To our knowledge, this was performed for the first time.

Conclusion

Fish-allergic individuals may react to the ingestion of a wide range of fish species at trace levels, thus relying on the labelled information of processed foods. Therefore, the development of reliable and sensitive methods is of utmost importance to verify labelling compliance and protect the health of allergic consumers. In view of that, after an exhaustive DNA sequence search of the 16S rRNA gene, new primers for the universal detection of fish were designed whose specificity was fully demonstrated, both *in silico* and experimentally. Two real-time PCR systems based on the universal EvaGreen dye and on a TaqMan probe were successfully developed and further evaluated. Both systems were able to reach a high absolute sensitivity (0.01 pg) with adequate real-time PCR performance parameters. However, the probe system enabled a relative LOD down to 0.0001% (1 mg/kg) of cooked fish in béchamel sauce, while the EvaGreen one revealed specific amplification only until a level of 0.05%, with consequent distinct dynamic ranges of 0.0001-50% and 0.05-50%, respectively. The validation results with blind samples showed that both systems were precise in the tested range, but trueness, expressed as bias, was within ±25% for all tested levels (0.25-8%) with the probe system, while with EvaGreen the error was too high for the highest tested level (8%). Therefore, the real-time PCR system with the probe exhibited the best performance in terms of relative sensitivity, dynamic range and trueness.

The application of the real-time PCR system with the probe proved to be useful to detect trace amounts of fish in processed foods. On the other hand, the application of real-time PCR system with EvaGreen dye extended the number of fish species detected that, due to nucleotide probe mismatches, could not be amplified with the probe system. Moreover, the verification of labelling compliance of foodstuffs suggested a high level of mislabelling and/or fraudulent practices. Therefore, the proposed approaches can provide useful tools to contribute for the safety assessment of foods, particularly protecting the health of fish allergic consumers.

Acknowledgements

This work was supported by FCT (Fundação para a Ciência e Tecnologia) through project UID/QUI/50006/2013 – POCI/01/0145/FEDER/007265 with financial support from

FCT/MEC through national funds and co-financed by FEDER, under the Partnership Agreement PT2020 and by the project NORTE-01-0145-FEDER-000011. T. J. R. Fernandes and J. Costa are grateful to FCT grants (SFRH/BD/93711/2013 and SFRH/BPD/102404/2014, respectively) financed by POPH-QREN (subsidised by FSE and MCTES). The codfish species were kindly provided by Pascoal & Filhos SA.

References

Abdullah, A., & Rehbein, H. (2016). The differentiation of tuna (family: Scombridae) products through the PCR-based analysis of the cytochrome b gene and parvalbumin introns. *Journal of the Science of Food and Agriculture, 96*, 456-464.

Benedetto, A., Abete, M. C., & Squadrone, S. (2011). Towards a quantitative application of real-time PCR technique for fish DNA detection in feedstuffs. *Food Chemistry, 126*(3), 1436-1442.

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., & Wittwer, C. T. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, *55*, 611-622.

Calderon-Rodriguez, S., Pineda, F., Perez, R., & Muñoz, C. (2016). Tolerability to dogfish in children with fish allergy. *Allergologia et Immunopathologia, 44*, 167-169.

Chen, Y. T., & Hsieh, Y. H. P. (2014). A sandwich ELISA for the detection of fish and fish products. *Food Control, 40*, 265-273.

Costa, J., Oliveira, M. B. P. P., & Mafra, I. (2013). Effect of thermal processing on the performance of the novel single-tube nested real-time PCR for the detection of walnut allergens in sponge cakes. *Food Research International*, *54*, 1722-1729.

ENGL (2015). Definition of minimum performance requirements for analytical methods of GMO testing. European network of GMO laboratories, Joint Research Center, EURL Available online at <http://gmo-crl.jrc.ec.europa.eu/doc/MPR%20Report%20Application%2020_10_2015.pdf>. Last accessed on March 3, 2017.

Faeste, C. K., & Plassen, C. (2008). Quantitative sandwich ELISA for the determination of fish in foods. *Journal of Immunology Methodologies*, 329, 45-55.

Feng, J., Wu, Z., Xie, X., Dai, Z., & Liu, S. (2017). A real-time polymerase chain reaction method for the identification of four commercially important salmon and trout species. *Mitochondrial DNA Part A, 28*, 104-111.

Fernandes, T. J. R., Costa, J., Oliveira, M. B. P. P., & Mafra, I. (2015). An overview on fish and shellfish allergens and current methods of detection. *Food and Agricultural Immunology, 26*, 848-869.

Fernandes, T. J. R., Costa, J., Carrapatoso, I., Oliveira M. B. P. P., Mafra, I. (2017a). Advances on the molecular characterisation, clinical relevance and detection methods of Gadiform parvalbumin allergens, *Critical Reviews in Food Science and Nutrition*, *57*, 3281-3296.

Fernandes, T. J. R., Costa, J, Oliveira M. B. P. P., Mafra, I. (2017b). DNA barcoding coupled to HRM analysis as a new and simple tool for the authentication of Gadidae fish species, *Food Chemistry*, *230*, 49-57.

Fernandes, T. J. R., Costa, J, Oliveira M. B. P. P., Mafra, I. (2018). COI barcode-HRM as a novel approach for the discrimination of hake species. *Fisheries Research, 197*, 50-59.

Giusti, A., Tinacci, L., Sotelo, C. G., Marchetti, M., Guidi, A., Zheng, W., & Armani, A. (2017). Seafood identification in multispecies products: assessment of 16S rRNA, 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.

Herrero, B., Vieites, J. M., & Espiñeira, M. (2014). Development of an in-house fast real-time PCR method for detection of fish allergen in foods and comparison with a commercial kit. *Food Chemistry, 151*, 415-420.

Houhoula, D., Dimitriou, P., Mengjezi, G., Kyrana, V., & Lougovois, V. (2015). Quantification of parvalbumin in commercially important mediterranean seafood species using real time PCR. *Czech Journal of Food Science, 33*, 143-147.

Kobayashi, A., Tanaka, H., Hamada, Y., Ishizaki, S., Nagashima, Y., & Shiomi, K. (2006). Comparison of allergenicity and allergens between fish white and dark muscles. *Allergy, 61*, 357-363.

Kuehn, A., Scheuermann, T., Hilger, C., & Hentges, F. (2010). Important variations in parvalbumin content in common fish species: A factor possibly contributing to variable allergenicity. *International Archives of Allergy and Immunology, 153*, 359-366.

Mourad, A. A., & Bahna, S. L. (2015). Fish-allergic patients may be able to eat fish. *Expert Review of Clinical Immunology, 11*, 419-430.

Pascual, C., Esteban, M. M., & Crespo, J. F. (1992). Fish allergy: Evaluation of the importance of cross-reactivity. *The Journal of Pediatrics, 121*, S29-S34.

Pegels, N., González, I., López-Calleja, I., García, T., & Martín, R. (2013). Detection of fish-derived ingredients in animal feeds by a TaqMan real-time PCR Assay. *Food Analytical Methods, 6*, 1040-1048.

Prado, M., Ortea, I., Vial, S., Rivas, J., Calo-Mata, P., & Barros-Velazquez, J. (2016). Advanced DNA- and protein-based methods for the detection and investigation of food allergens. *Critical Reviews in Food Science and Nutrition, 56*, 2511-2542.

Regulation (EU) No 1169/2011 of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No. 1924/2006 and (EC) No. 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004. Official Journal of the European Union, L304, 18–63.

Salihah, N. T., Hossain, M. M., Lubis, H., & Ahmed, M. U. (2016). Trends and advances in food analysis by real-time polymerase chain reaction. *Journal of Food Science and Technology, 53*, 2196-2209.

Sharp, M. F., & Lopata, A. L. (2014). Fish allergy: in review. *Clinical Reviews in Allergy & Immunology, 46*, 258-271.

Sun, M., Liang, C., Gao, H., Lin, C., & Deng, M. (2009). Detection of parvalbumin, a common fish allergen gene in food, by real-time polymerase chain reaction. *Journal of AOAC International, 92*, 234-240.

Taylor, S. L., Hefle, S. L., Bindslev-Jensen, C., Bock, S. A., Burks, A. W., Christie, L., Hill, D. J., Host, A., Hourihane, J. O. B., Lack, G., Metcalfe, D. D., Moneret-Vautrin, D. A., Vadas, P. A., Rance, F., Skrypec, D. J., Trautman, T. A., Yman, I. M., & Zeiger, R. S. (2002). Factors affecting the determination of threshold doses for allergenic foods: How much is too much? *Journal of Allergy and Clinical Immunology, 109*, 24-30.

Tetzlaff, C., & Mäde, D. (2017). Development of a real-time PCR system for the detection of the potential allergen fish in food. *European Food Research and Technology*, *243*, 849-857.

Thalayasingam, M., & Lee, B. -W. (2015). Fish and shellfish allergy. M. Ebisawa, B. K. Ballmer-Weber, S. Vieths, R. A. Wood (Eds.), *Food allergy: molecular basis and clinical practice. Chem Immunol Allergy* (pp. 152-161). Basel: Karger.

Triantafyllidis, A., Karaiskou, N., Perez, J., Martinez, J. L., Roca, A., Lopez, B., & Garcia-Vazquez, E. (2010). Fish allergy risk derived from ambiguous vernacular fish names: Forensic DNA-based detection in Greek markets. *Food Research International, 43*, 2214-2216.

Weber, P., Steinhart, H., & Paschke, A. (2009). Competitive indirect ELISA for the determination of parvalbumins from various fish species in food grade fish gelatins and isinglass with PARV-19 anti-parvalbumin antibodies. *Journal of Agricultural and Food Chemistry*, 57, 11328-11334.

Journal of Food Composition and Analysis (submitted)

Contents lists available at ScienceDirect

Journal of Food Composition and Analysis

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

A new real-time PCR quantitative approach for the detection of shrimp crustaceans as potential allergens

Telmo J.R. Fernandes, Joana Costa, M. Beatriz P.P. Oliveira, Isabel Mafra*

REQUIMTE-LAQV, Faculdade de Farmácia, Universidade do Porto, Rua de Jorge Viterbo Ferreira, 228, 4050-313 Porto, Portugal

*Corresponding author: Tel: +351 220428640. Fax: +351 226093390. E-mail: isabel.mafra@ff.up.pt

Abstract

Allergy to crustaceans is an important food safety issue nowadays. To protect people from experiencing adverse allergic reactions, reliable methodologies are necessary to verify the labelling of processed seafoods. In the present work, two new DNA-based approaches targeting the 16S rRNA mitochondrial gene are proposed to detect crustaceans in foods using a qualitative PCR assay specific for crustaceans (shrimps, lobsters and crabs) and a quantitative real-time PCR assay specific for shrimp crustaceans. The real-time PCR system allowed the detection and quantification down to 0.1 pg and 0.0001% (w/w) of shrimp DNA and shrimp in model mixtures, respectively. The method exhibited high performance for quantitative analysis in the range of 0.0001% to 50% as inferred by the calibration curve parameters, being effectively validated with blind mixtures. The qualitative PCR assay can provide a simple, fast and high throughput tool for screening the presence of crustaceans in processed foods, while the proposed real-time PCR method proved to be a useful tool for the accurate detection and quantification of shrimp in foods at trace levels.

Keywords: Crustaceans; allergen detection; real-time PCR; 16S rRNA; ELISA; quantification.

Introduction

The international trade of seafood species and products has been growing over the years mainly due to the high nutritive value and popularity of these foods across many countries. Nonetheless, along with the increased production and consumption of seafood, food allergic reactions among consumers have been more frequently reported (Daul et al., 1993; Lopata et al., 2010; Günther et al., 2017). Crustaceans such as shrimps, crabs or lobsters represent a major food resource with high commercial value that may elicit allergic reactions (Pedrosa et al., 2015; Thalayasingam et al., 2015; Khora et al., 2016). About 60% of all crustaceans (13.9 million tonnes), on the global capture and production, belong to shrimps/prawns, from which *Litopenaeus vannamei* is the main crustacean species with a production of 3.9 million tonnes in 2015 (FAO, 2017).

Since the allergic individuals can have a very strong adverse immunologic reaction, even to the ingestion of minute quantities of an allergen, they should rely on the labelling information to avoid the consumption of any offending food. Besides, cross-contamination occurrences during production, storage or transport may lead to the inadvertent presence of allergens in foods (Herrero et al., 2012; Fernandes et al., 2017). Hence, in order to protect sensitised persons, the labelling of fourteen allergenic ingredients (including crustaceans) is demanded by the European Union legislation (Directive 2007/68/EC, Regulation (EU) No 1169/2011). To comply with it, several producers use frequently the precautionary labelling “may contain traces of…”, which often causes some indecision or rejection by food allergic consumers at a global scale (Fernandes et al., 2017). Therefore, sensitive analytical methods aiming at detecting trace levels of the allergenic ingredients in complex matrices are needed to verify the labelling of crustaceans.

According to the Regulation (EU) No 1169/2011, the detection of the allergenic food is demanded, but the target analyte does not necessarily have to be the allergenic protein itself. In this sense, both protein- and DNA-based methods have been widely used for allergen detection (Zhenxing et al., 2010; Eischeid et al., 2013; Prado et al., 2016). For protein analysis, the enzyme-linked immunosorbent assay (ELISA) is by far the most widely used immunochemical technique to identify and quantify allergens in food (Lopata et al., 2010; Zhenxing et al., 2010; Wang et al., 2011; Gomaa & Boye, 2015). In spite of the major advantages of this methodology associated with simplicity, fastness, high specificity and sensitivity inherent to the antigen/antibody interaction, it also presents some drawbacks. Cross-reactivity phenomena can lead to the occurrence of false positive results, while conformational changes of proteins caused by food processing can conduct to false negative results (Costa et al., 2014; Prado et al., 2016). On the other hand, DNA-based methods meet some advantages in comparison with immunochemical assays. DNA

molecules are almost ubiquitous in any organic matter and more thermostable than proteins, which allows their analysis from difficult matrices, such as processed and complex foods (Herrero et al., 2012; Eischeid et al., 2013). For allergen analysis, real-time PCR-based methods have been widely applied (Mafra et al., 2008; Prado et al., 2016; Costa et al., 2017). Their specificity has relied mainly on primers and probes targeting sequences of genes encoding for allergenic proteins, but mitochondrial genes have also demonstrated to be useful targets (Prado et al., 2016; Costa et al., 2017). In the case of crustacean detection and differentiation, the mitochondrial genes, such as 16S rRNA (Cao et al., 2011; Pascoal et al., 2011; Herrero et al., 2012; Eischeid et al., 2013; Mäde & Rohmberger, 2017; Zagon et al., 2017; Wilwet et al., 2018), 12S rRNA (Eischeid et al., 2013; Eischeid, 2016; Eischeid & Stadig, 2018) and cytochrome oxidase subunit I (COI) (Eischeid et al., 2013; Fernandes et al., 2017) have been used as specific markers, taking advantage of their high copy number that usually results in more sensitive assays. From the available reports on real-time PCR, several of them allow high sensitivity levels of detecting particular groups of crustacean species (Cao et al., 2011; Herrero et al., 2012; Eischeid et al., 2013; Eischeid, 2016; Mäde & Rohmberger, 2017; Zagon et al., 2017; Eischeid & Stadig, 2018). However, only few reports are based on quantitative real-time PCR methods using model mixtures simulating low levels of spiked crustaceans (Eischeid et al., 2013; Eischeid, 2016; Eischeid & Stadig, 2018). Eischeid et al. (2013) describes two real-time PCR systems able to quantify DNA from penaeid shrimps and crabs. Eischeid (2016) reports a real-time PCR quantitative system specific for lobsters based on model mixture as calibrants, followed by a similar approach for crabs (Eischeid & Stadig, 2018), being both methods effective for allergen quantification.

Considering the scarce methods on crustacean allergen quantification by DNA-based methods, the development of a quantitative real-time PCR system specific of a wide range of shrimp species is proposed, validated and applied to verify labelling compliance of foods. Additionally, the performance of quantitative analysis of the new method is compared with ELISA.

Materials and Methods

Sample collection and preparation

Several crustacean specimens from different species (*n*=20) (whiteleg shrimp - *Litopenaeus vannamei;* speckled prawn - *Metapenaeus monoceros*; Indian white prawn - *Fenneropenaeus indicus*; giant tiger prawn - *Penaeus monodon;* jinga shrimp - *Metapenaeus affinis;* green tiger prawn – *Penaeus semisulcatus*; *Penaeus* spp; Argentine red shrimp - *Pleoticus muelleri;* razor mud shrimp - *Solenocera melantho;* Udang merah -

Solenocera crassicornis; knife shrimp - *Haliporoides triarthrus; Solenocera* spp.; common shrimp – *Palaemon serratus*; Scarlet shrimp - *Aristaeopsis edwardsiana,* Norway lobster - *Nephrops norvegicus;* European lobster – *Hommarus grammarus*; Caribbean spiny lobster - *Panalirus argus;* crab - *Portunus validus;* velvet swimming crab – *Necora puber;* edible crab - *Cancer pagarus*) were purchased from local markets or provided by Marfresco (Loures, Portugal) and Brasmar Seafood Companies (Trofa, Portugal). The selection of species was made regarding the relevance in terms of consumption and availability in different commercial areas of the Portuguese market.

A total of 62 non-crustacean species were also acquired to evaluate the assay specificity: seafood species (*n*=28) (Atlantic cod, Pacific cod, Alaska Pollock, saithe, ling, Atlantic salmon, gilt-head bream, common sole, European pilchard, yellowfin tuna, European hake, North Pacific hake, Argentine hake, whiting, deepwater hake, ray, European seabass, whiting-pout, Atlantic horse mackerel, tadpole codling, rose fish, rock ling, Pacific mackerel, black scabbardfish, Nile perch, squid, mussel, undulated surf clam), meat species (*n*=15) (boar, duck, partridge, hare, quail, pheasant, deer, rabbit, chicken, turkey, lamb, ostrich, cow, horse, pig) and plant species (*n*=19) (onion, garlic, parsley, pepper, bay leaf, sweet chili, tomato, maize, soybean, potato, manioc, lupine, chestnut, walnut, broad bean, rye, wheat, rice, pumpkin).

Several seafood products (*n*=21), including shrimp patties, seafood broth/soups, surimi and pre-cooked dishes, were also acquired at local markets for assay applicability.

The specimens and food samples were ground and homogenised separately in a laboratory knife mill Grindomix GM200 (Retsch, Haan, Germany) and stored at -20 ºC until further analysis. All containers and material used during this procedure were previously treated with DNA decontamination solution (DNA-ExitusPlus™, AppliChem, Darmstadt, Germany).

Preparation of model mixtures

The preparation of binary model mixtures intended to simulate a processed shrimp stuffing/filling (e.g. shrimp patty fill). Shrimp kernels were boiled in water during 5 min, drained and minced in a laboratory knife mill Grindomix GM200 (Retsch, Haan, Germany). A mixture containing 50% (w/w) of ground-cooked shrimp kernel in béchamel sauce was prepared by the addition of 200 g of shrimp to the same amount of sauce. Afterwards, the following binary proportions: 10, 5, 1, 0.5, 0.1, 0.05, 0.01, 0.005, 0.001, 0.0005 and 0.0001% (w/w) of shrimp in béchamel. Identically, for method validation, blind mixtures were prepared to contain 8.0, 4.0, 2.5 and 0.25% (w/w) of shrimp in béchamel.

The binary mixtures were ground and homogenised in a laboratory knife mill and stored at -20 ºC until further analysis. All containers and material were previously treated with DNA decontamination solution (DNA-ExitusPlus™, AppliChem, Darmstadt, Germany).

Primer and probe design

Sequences of the 16S rRNA gene were selected and uploaded from NCBI database from a set of 17 different crustacean species (Figure 1). The majority of the crustacean species match with the ones most frequently consumed as food. However, it was not always possible to find available sequences on the available databases for all species commonly marketed, and vice-versa. Sequence alignment was performed with BioEdit v.7.2.5 software (Ibis Biosciences, Canada) and examined for suitable regions for primer and probe design. Primers (16SCrust2-F: TAA AGT CTG GCC TGC CCA CTG; 16SCrust1-R: AGC TTT ATA GGG TCT TAT CGT C) and probe (16SCrust1-P: FAM-TTA ATT GAA GGC TTG TAT GAA TGG TTG GAC-BHQ1) were then submitted to a basic local alignment search tool BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) in order to identify regions of local similarity among homologue sequences of different species and to calculate the statistical significance of the matches. Primer properties, self-hybridisation and the absence of hairpins were evaluated using the software OligoCalc (http://www.basic.northwestern.edu./biotools/oligocalc.html). The oligonucleotides were synthesised by STABVIDA (Lisbon, Portugal).

DNA Extraction

The DNA was extracted using the SureFood® Prep advanced kit (CONGEN Biotechnologie GmbH, Berlin, Germany). All DNA extractions were done in duplicates and extracts were kept at -20 ºC until further analysis. The concentration and purity of the extracts were assessed by UV spectrophotometric DNA quantification on a Synergy HT multi-mode microplate reader (BioTek Instruments, Inc., Vermont, USA), using a Take3 micro-volume plate accessory. DNA content was determined using the nucleic acid quantification protocol with sample type defined for double-strand DNA in the Gen5 data analysis software version 2.01 (BioTek Instruments, Inc., Vermont, USA).

The quality of extracted DNA was also analysed by electrophoresis in a 1.0% agarose gel containing Gel Red 1x (Biotium, Hayward, CA, USA) for staining and carried out in STGB 1x (GRISP, Porto, Portugal) for 25 min at 200 V. The agarose gel was visualised under a UV light tray Gel Doc™ EZ System (Bio-Rad Laboratories, Hercules, CA, USA) and a digital image was obtained with Image Lab software version 5.1 (Bio-Rad Laboratories, Hercules, CA, USA).

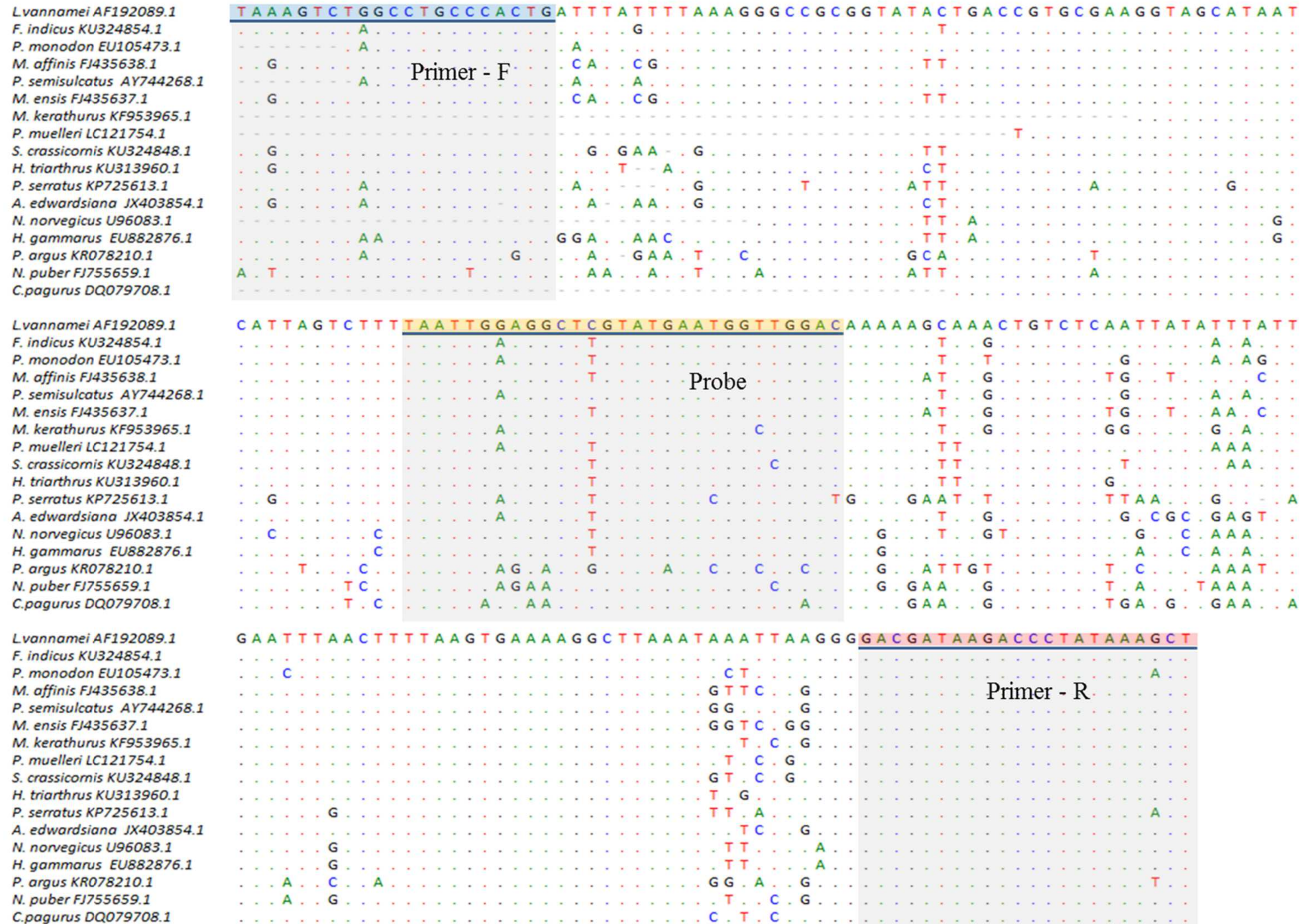


Figure 1. Sequence alignment of NCBI 16S rRNA entries showing primer (16SCrust2-F/16SCrust1-R) and probe (16SCrust1-P) design regions.

Qualitative PCR amplification

The following PCR protocol was used for DNA amplification: 25 µL of total reaction volume containing 2 µL of DNA extract (100 ng), 67 mM of Tris-HCl (pH 8.8), 16 mM of $(NH_4)_2SO_4$, 0.01% of Tween 20, 200 µM of each dNTP (Bioron, Ludwigshafen, Germany), 1.0 U of SuperHot Taq DNA Polymerase (Genaxxon Bioscience, Ulm, Germany), 3.0 mM of $MgCl_2$ and 200 nM of each primer (16SCrust2-F/16SCrust1-R). The amplifications were performed in a SimpliAmp Thermal Cycler (Applied Biosystems, California, USA) using the following cycling program: initial denaturation at 95 ºC for 5 min; 40 cycles at 95 ºC for 30 s, 54 ºC for 30 s and 72 ºC for 30 s; and a final extension at 72 ºC for 5 min.

The amplicons were then analysed by electrophoresis in a 1.5% agarose gel containing 1× Gel Red (Biotium, Hayward, CA, USA) for staining and carried out in 1× SGTB buffer (GRISP, Porto, Portugal) for about 20-25 min at 200 V. The agarose gel was visualised under a UV light tray Gel Doc™ EZ System (Bio-Rad Laboratories, Hercules, CA, USA) and a digital image was obtained with Image Lab software version 5.1 (Bio-Rad Laboratories, Hercules, CA, USA).

Real-time PCR amplification

For the real-time PCR amplifications, 20 µL of total reaction mixture included 1× of SsoFast Probes Supermix (Bio-Rad, Hercules, CA, USA), 300 nM of each primer (16SCrust2-F/16SCrust1-R), 200 nM of probe (16SCrust1-P) and 2 µL of DNA extract. The following temperature conditions performed on a fluorometric thermal cycler CFX96 real-time PCR detection system (Bio-Rad Laboratories, Hercules, CA, USA) were used: 95 ºC for 5 min, 50 cycles at 95 ºC for 15 s and 60 ºC for 45 s, with collection of fluorescence signal at the end of each cycle. Fluorescence data were and analysed using the software Bio-Rad CFX Manager 3.1. Cycle of quantification (Cq), also known as cycle threshold (Ct), values were calculated using the software at automatic threshold settings.

ELISA

Two different ELISA kits for the detection of crustaceans in foods were used for comparative purposes with the proposed real-time PCR assay. The MonoTrace™ Crustacea ELISA Kit (BioFront Technologies, Florida, USA) is a monoclonal antibody (Mab)-based assay used for the quantification of crustacean (1-40 mg/kg) presence/contamination in several food matrices, while RidaScreen® Fast Crustacean (R-Biopharm AG, Darmstadt, Germany) claims the detection of crustacean proteins (mainly tropomyosin) in a range of 20-160 mg/kg.

The extraction of proteins from model mixtures and food samples used in both ELISA kits was carried out according to the manufacturers' description. Briefly, a pre-heated extraction buffer (9 mL or 20 mL for MonoTrace or RidaScreen kits, respectively) was added to 1 g (double extraction for each sample) of previously ground sample and incubated for 10 min (42 ºC or 60 ºC for MonoTrace or RidaScreen kits, respectively) with continuous mixing. The sample was centrifuged (2,000×*g* or 2,500×*g* for MonoTrace or RidaScreen kits, respectively) and the aqueous phase was transferred to new tubes for subsequent analysis.

Afterwards, similar protocols were performed in both ELISA kits. Microplates coated with specific antibodies were incubated for 10 min at room temperature with 100 µL or 200 µL of standards and samples for MonoTrace or RidaScreen kits, respectively. Each microplate was rinsed 3-times with 200 µL or 250 µL of washing buffer solution for MonoTrace or RidaScreen kits, respectively, followed by incubation in the dark for 10 min at room temperature with 100 µL of antibody-conjugate (peroxidase). After a new washing step, the addition of 100 µL of substrate allowed the conversion of the complex to a coloured product. Finally, the addition of a stop solution (100 µL) lead to a new colour change, allowing the absorbance measurement at 450 nm using a Synergy HT multi-mode microplate reader (BioTek Instruments, Inc., Vermont, USA). The results were evaluated with Gen5 data analysis software version 2.01 (BioTek Instruments, Inc., Vermont, USA). The absorbance values were plotted against the logarithmic concentration of the crustacean protein standard solutions. A non-linear regression function was carried using a sigmoid four parametric logistic function:

$$B \equiv Y = \frac{A - D}{1 + \left(\frac{X}{C}\right)^{b}} + D$$

being Y the optical density (absorbance), A the maximum absorbance, b the slope of the calibration curve in linear range, C the 50% inhibition-concentration – IC_{50} (µg/L), D the minimum absorbance, X the analyte concentration (µg/L) (Figure S1, supplementary material). Each sample was analysed in triplicate.

Results and Discussion

16S rRNA primer specificity

A sequence of the mitochondrial of 16S rRNA was used as a group specific target for the crustacean detection. This evidence was firstly demonstrated by *in silico* analysis, which enables identifying homologue regions of crustaceans among the available sequences at NCBI (Figure 1). The PCR amplification with the proposed new primers (16SCrust2-F/16SCrust1-R) showed the expected 203 bp fragment for the all the tested

crustacean species (Table 1, Figure S2, supplementary material). The specificity of primers was further assessed using the DNA of 62 food-related species (seafood, other animal species and plants), previously described (section 2.1). Since no cross-reactivity was observed with any of non-target species, the specificity of the assay was confirmed for crustacean species.

To guarantee the absence of any false negative results, the amplification capacity of all the DNA extracts were previously amplified by PCR with universal eukaryotic primers (18SRG-F/18SRG-R), as described by Costa et al. (2013).

Real-time PCR method development

A real-time PCR probe system targeting the 16S rRNA gene was proposed for the detection and quantification of crustacean species, focusing on the group of shrimps. Considering the results of qualitative PCR (section 3.1), the specificity of the real-time PCR system was verified for all crustacean species, being the cycle threshold (Ct) values presented in Table 1. As it can be observed, Ct values ranged from 20.20 (*Litopenaeus vannamei*) to 23.88 (*Penaeus semisulcatus*) and five crustacean species were not amplified by real-time PCR, namely *Palaemon serratus*, *Panalirus argus, Portunus validus*, *Necora puber* and *Cancer pagarus* (Table 1). The negative amplifications are supported by the 16S rRNA sequence alignment (Figure 1), particularly when the probe region is considered. Indeed, *C. pagarus* and *P. serratus* presented 4 nucleotide differences comparing with the consensus sequence, while *N. puber* and *P. argus* showed 5 and 8 nucleotide mismatches, respectively. Generally, the real-time PCR system was effective in detecting all the species from the Penaeoidea superfamily, in which are included the most commonly consumed shrimp species. Only one shrimp species (*P. serratus*) was not amplified with the probe system because it belongs to a more distantly related taxonomic group at the level of suborder (Pleoyemata). However, this is a coastal seasonal species, with small size, being consumed as a snack rather than processed in foods. On the other hand, in the same suborder, the species of Nephropidae family were amplified, while crab species from Palinuridae, Portunidae and Cancridae families were not. The cross-reactivity of species of Nephropidae family also occurred in the method of Eischeid and Stadig (2018) specific for crabs, probably because they share common features to both crabs and shrimps. Therefore, the proposed real-time PCR method can be considered as a specific approach for the detection and quantification of the most common shrimp species, particularly those included in the superfamily of Penaeoidea (Penaeidae, Solenoceridae and Aristeidae).

Table 1. Qualitative PCR results for eukaryotic and crustacean-specific amplifications and real-time PCR cycle threshold values for several crustacean species.

Order	Suborder	Infraorder	Superfamily	Family	Species	Scientific name	Qualitative PCR [a]		Real-time PCR
							Eukaryotic (18S rRNA)	Crustacean (16S rRNA)	Ct ± SD [b]
Decapoda	Dendrobranchiata		Penaeoidea	Penaeidae	Whiteleg shrimp	*Litopenaeus vannamei*	+	+	20.20 ± 0.30
					Speckled prawn	*Metapenaeus monoceros*	+	+	23.53 ± 0.36
					Indian white prawn	*Fenneropenaeus indicus*	+	+	23.40 ± 0.21
					Giant tiger prawn	*Penaeus monodon*	+	+	22.17 ± 0.09
					Jinga shrimp	*Metapenaeus affinis*	+	+	20.22 ± 0.09
					Green tiger prawn	*Penaeus semisulcatus*	+	+	23.88 ± 0.16
				Solenoceridae	Argentine red shrimp	*Pleoticus muelleri*	+	+	21.79 ± 0.21
					Razor mud shrimp	*Solenocera melantho*	+	+	22.01 ± 0.07
					Coastal mud shrimp	*Solenocera crassicornis*	+	+	20.60 ± 0.38
					Knife shrimp	*Haliporoides triarthrus*	+	+	20.53 ± 0.28
				Aristeidae	Scarlet shrimp	*Aristaeopsis edwardsiana*	+	+	20.47 ± 0.11
	Pleocyemata	Caridea	Palaemonoidea	Palaemonidae	Common prawn	*Palaemon serratus*	+	+	-
		Astacidea	Nephropoidea	Nephropidae	Norway lobster	*Nephrops norvegicus*	+	+	20.28 ± 0.06
					European lobster	*Homarus grammarus*	+	+	21.14 ± 0.12
		Achelata		Palinuridae	Caribbean spiny lobster	*Panulirus argus*	+	+	-
		Brachyura	Portunoidea	Portunidae	Smoth swimcrab	*Portunus validus*	+	+	-
					Velvet swimming crab	*Necora puber*	+	+	-
			Cancroidea	Cancridae	Edible crab	*Cancer pagurus*	+	+	-

[a] (+) positive qualitative PCR amplification; [b] Mean values of cycle threshold (Ct) ± standard deviation (SD) of replicate assays (n=4); (-) non-detected amplification.

For the development of the real-time PCR quantification approach, the requirements set by the European Network of GMO Laboratories (ENGL) document "Definition of minimum performance requirements for analytical methods of GMO testing" (ENGL, 2015) and the guidelines of "Minimum Information for Publication of Quantitative Real-Time PCR Experiments" (Bustin et al., 2009) were taken into consideration. The real-time PCR system was applied to a 10-fold serially diluted shrimp DNA extract (100 000 pg to 0.01 pg) that allowed establishing a dynamic range of 7 orders of magnitude and an absolute limit of detection (LOD) of 0.1 pg (Table 2, Figure 2). The LOD could also be considered as the limit of quantification (LOQ) since it was within the linear range of the calibration curve (Figure 2). The calibration curve parameters, namely PCR efficiency (97.4%), R^2 (0.997) and slope (-3.371) were all within the acceptance criteria, suggesting a high performance of the assay (Table 2) (Bustin et al., 2009; ENGL, 2015). These results are in good agreement with a reported method that enabled the detection of penaeid shrimp DNA down to 0.1 pg (Eischeid et al., 2013).

Table 2. Real-time PCR results targeting the 16S rRNA gene of crustaceans for the absolute detection of shrimp DNA using 10-fold serially diluted extracts and for the relative determination of shrimp meat using binary model mixtures.

Absolute amount (pg shrimp)	Ct ± SD[a]	Relative amount (% w/w of shrimp in béchamel)	Ct ± SD[a]
100 000	17.74 ± 0.33 (16/16)	50	18.24 ± 0.28 (12/12)
10 000	21.21 ± 0.21 (16/16)	10	20.22 ± 0.48 (12/12)
1 000	24.79 ± 0.26 (16/16)	5	22.05 ± 0.29 (12/12)
100	28.21 ± 0.24 (16/16)	1	22.96 ± 0.45 (12/12)
10	31.56 ± 0.22 (16/16)	0.5	25.53 ± 0.28 (12/12)
1	34.74 ± 0.50 (16/16)	0.1	27.44 ± 0.45 (12/12)
0.1	38.06 ± 0.57 (16/16)	0.05	29.27 ± 0.38 (12/12)
0.01	nd[b]	0.01	30.93 ± 0.20 (12/12)
		0.005	31.68 ± 0.22 (12/12)
		0.001	34.49 ± 0.37 (12/12)
		0.0005	36.47 ± 0.99 (12/12)
		0.0001	37.28 ± 0.49 (12/12)
PCR efficiency (%)	97.4 ± 0.9	PCR efficiency (%)	97.7 ± 2.3
Correlation coefficient (R^2)	0.997 ± 0.000	Correlation coefficient (R^2)	0.987 ± 0.003
Slope	-3.371 ± 0.023	Slope	-3.380 ± 0.056

[a] Mean cycle threshold (Ct) values ± standard deviation (SD) of replicate assays, in brackets (positive amplifications/total replicates);[b] nd, not detected.

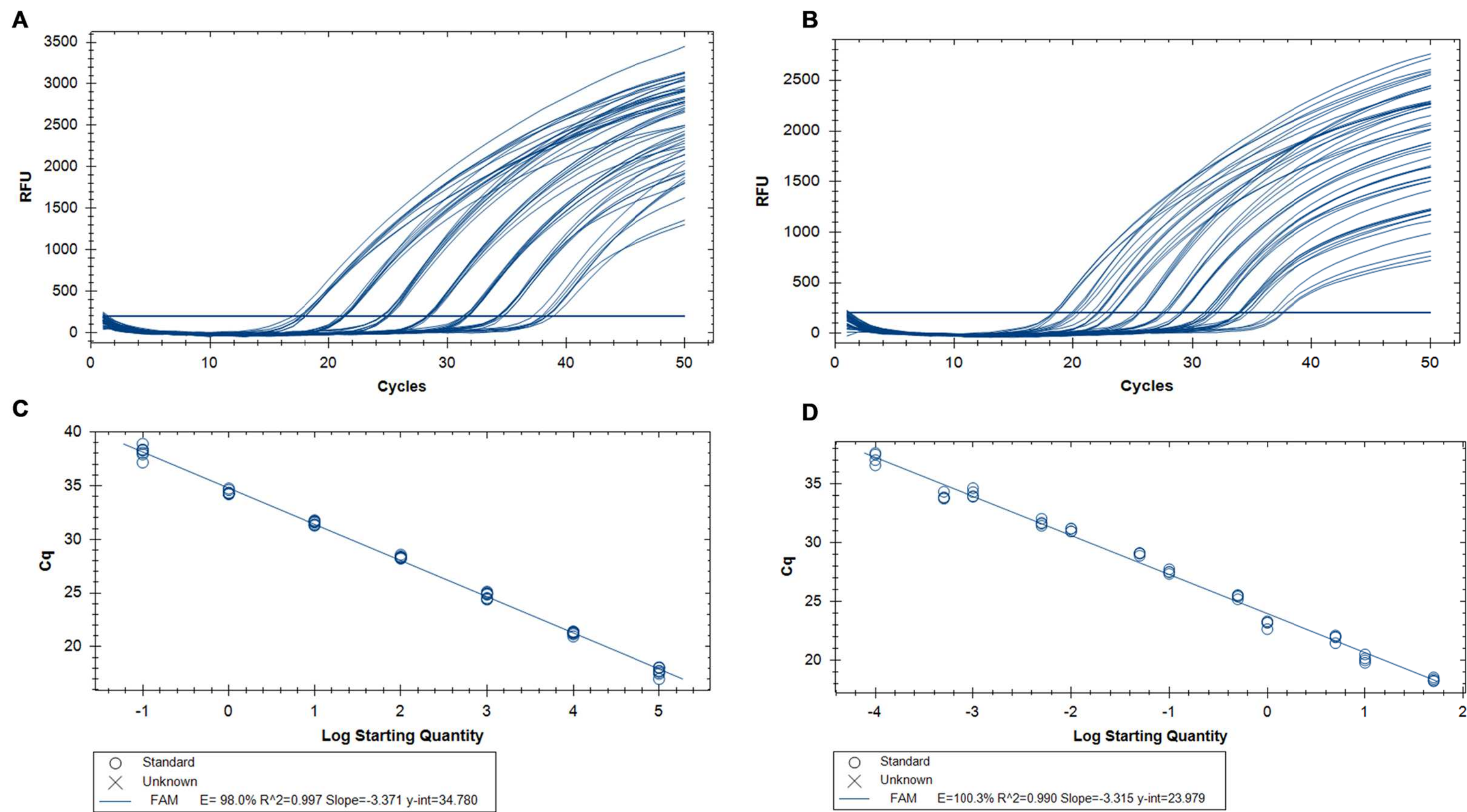


Figure 2. Real-time PCR amplification (A, B) and calibration (C, D) curves with a TaqMan probe targeting the 16S rRNA gene using a serially diluted shrimp-DNA extract (100,000-0.1 pg) (A, C) and DNA extracts from binary model mixtures of cooked shrimp in béchamel sauce (50-0.0001% (w/w)) (B, D).

To propose a system that enables estimating the amount of shrimp crustaceans in foods, the use of binary mixtures as reference standards with known amounts of spiked shrimp is crucial. Their use allowed establishing a new calibration curve with a dynamic range of 6 orders of magnitude, covering all the tested concentration range of 50-0.0001% (w/w) (Table 2, Figure 2). Accordingly, both LOD and LOQ were identical, being down to 0.0001% (1 mg/kg). The calibration curve parameters (PCR efficiency of 97.7%, slope of -3.380 and R^2 of 0.987) suggested again a high performance of the assay (Table 2) (Bustin et al., 2009, ENGL 2015). These findings are in good agreement with the levels of quantification obtained by other authors for crabs (Eischeid & Stadig, 2018) and lobsters (Eischeid, 2016), which ranged between 0.00001% and 10% (w/w). Regarding the relative quantification of shrimp meat using model mixtures as calibrants, no data was found in the literature.

For assay validation, the developed real-time PCR system was evaluated for its ability to estimate the shrimp content using blind mixtures with 8%, 4%, 2.5% and 0.25% (w/w) of shrimp. The coefficient of variation expressing the relative standard deviation of the estimated values ranged between 13.0% and 19.9%, which indicates that the method is precise (⩽25%) (Table 3). Comparing the actual values with the estimated ones, it is evident a high proximity between them. The measured trueness, expressed as bias (0.83-11.3%), is within ±25% of the actual value over the tested dynamic range, demonstrating the high accuracy of the method (ENGL, 2015). Therefore, in the present work, a real-time PCR method able to quantify shrimp crustacean meat in food is proposed and validated for the first time.

Table 3. Results for the validation of the quantitative real-time PCR assays using blind samples

Blind samples	Crustacean content % (w/w)		CV (%)[b]	Bias (%)[c]
	Actual Value	Estimated value (Mean ± SD) [a]		
A	8.0	7.13 ± 0.93	13.0	-10.9
B	4.0	4.51 ± 0.59	13.0	12.6
C	2.5	2.52 ± 0.42	16.5	0.83
D	0.25	0.28 ± 0.06	19.9	11.3

[a] Mean values ± standard deviation (SD) of replicate assays (n = 12) from three independent real-time PCR runs; [b] CV, coefficient of variation; [c] Bias = ((mean estimated value - true value)/true value x 100).

Analysis of processed foods by real-time PCR and ELISA

To verify the applicability of the proposed PCR methods, 19 food samples containing crustaceans or labelled as susceptible of containing them were analysed (Table 4). Concerning the qualitative PCR results, as expected, most samples were positive

producing the 203 bp fragment of the 16S rRNA gene. The exceptions were the sample #17 ("Brás" style codfish) with a faint band, and samples #12 (noodles) and #18 (codfish "patanisca") that were negative. This is complying with the labelled information since none of the samples declare shrimp as ingredient (samples #17 and #18 may contain crustaceans and sample #12 only contains shrimp flavour). The real-time PCR results confirmed these findings, with Ct values ranging from 21 to 36, allowing estimating the shrimp content in all samples (Table 4). The foods containing the highest amounts of shrimp were the #15 (shrimp "feijoada") and #16 (shrimp "açorda"), which is in good agreement with the composition of these traditional dishes. The different samples of tested patties (#1-#5) reveal variable shrimp contents, which might be related to the addition of shrimp extracts/flavours instead of shrimp meat. The same finding can be attributed to the samples of shellfish soups (#6 and #7) and surimi (#10 and #11), while samples #9, #13 and #14 evidence a very low proportion of shrimps among other seafood species.

For comparative purposes, two ELISA kits for crustacean detection, namely MonoTrace and RidaScreen, were also applied to verify their feasibility to estimate crustaceans. The calibration curve of MonoTrace ELISA Kit was determined according to manufacturer instructions with the provided ready-to-use crustacean standards (0, 1, 4, 10, 20 and 40 mg/kg) with a R^2 of 0.999 (Figure S1A, supplementary material). Similarly, the calibration curve of RidaScreen ELISA kit was built using crustacean standards (0, 20, 40, 80 and 160 mg/kg) with a R^2 of 1.00, which is presented in Figure S1B (supplementary material). From the tested samples, 9 were selected to be analysed by both kits in n=3 replicates (Table 4). All the analysed samples were positive for the presence of crustaceans. When attempting estimating the crustacean content, most samples needed to be diluted (from 10- to 50-fold) and even so, some were out of the quantitative range using the MonoTrace kit (#5, #7, #9 and #14). The non-quantified or high estimated contents for samples #2, #5, #7, #9 and #14 are not really well correlated with the real-time PCR determinations. It is important to refer that these assays are optimised to quantify trace levels of crustaceans, rather than crustaceans as ingredients in foods. Moreover, these kits are optimised to target tropomyosins from crustaceans, based on their strong resistance to food processing and high abundance. However, tropomyosins are also present in molluscs, which are ingredients in samples #9 and #14, suggesting potential cross-reactivity in these cases. The presence of spices, such as curcuma, and other condiments are also prone to develop cross-reactivity with ELISA kits, justifying the

Table 4. Results of the analysed processed food samples by qualitative PCR, real-time PCR and ELISA for the detection and quantification of crustaceans.

Sample	Type of food	Relevant label information	Qualitative PCR		Real-time PCR[a]		ELISA	
			Eukaryotic (18S rRNA)	Crustacean (16S rRNA)	Ct ± SD	Estimated amount ± SD (mg/kg)	MonoTrace Estimated amount ± SD (mg/kg)[b]	RidaScreen Estimated amount ± SD (mg/kg)[b]
1	Shrimp patties	shrimp, shrimp powder	+	+	27.04 ± 0.34	892 ± 22	nt[e]	nt
2	Shrimp patties	shrimp, shrimp powder	+	+	27.37 ± 0.08	710 ± 37	1392 ± 44*	3071 ± 55*
3	Shrimp patties	shrimp, shrimp powder	+	+	25.34 ± 0.49	2570 ± 360	nt	nt
4	Shrimp patties	shrimp, shrimp powder	+	+	25.09 ± 0.32	3110 ± 680	nt	nt
5	Shrimp and hake patties	Shrimp, fish	+	+	28.76 ± 0.39	168 ± 17	>LQ	3618 ± 841*
6	Shellfish soup	shrimp, lobster extract	+	+	26.66 ± 0.34	770 ± 116	nt	nt
7	Shellfish soup	shrimp, lobster extract	+	+	27.58 ± 0.87	307 ± 54	>LQ	5061 ± 881*
8	Shellfish broth	shrimp, shrimp and lobster extract	+	+	24.63 ± 0.62	4510 ± 560	nt	nt
9	Shellfish rice preparation	crustaceans	+	+	30.37 ± 0.40	65.7 ± 17.9	>LQ	828.3 ± 31.1*
10	Surimi	crab (extract and flavour)	+	+	36.03 ± 0.88	1.16 ± 0.17	1.24 ± 0.07	79.8 ± 3.8
11	Surimi	crab (extract and flavour)	+	+	33.99 ± 0.75	4.71 ± 0.54	3.60 ± 0.08	<LQ (4.94 ± 0.44)
12	Shrimp noodles	crustaceans, shrimp flavour	+	-	-		nt	nt
13	Seafood risotto	shrimp, squid, mussels, fish, crustacean extract	+	+	27.03 ± 0.63	483 ± 72	nt	nt
14	"Paella"	shrimp, squid, mussels, chicken	+	+	29.65 ± 0.63	147 ± 43	>LQ	4245 ± 365*
15	Shrimp "feijoada"	shrimp, lobster extract	+	+	22.54 ± 0.74	11400 ± 900	nt	nt
16	Shrimp "açorda"	shrimp	+	+	21.05 ± 0.20	43200 ± 5500	nt	nt
17	"Brás" style codfish	may contain crustaceans	+	+/-	-		<LQ	<LQ
18	Codfish "patanisca"	may contain crustaceans	+	-	-		28.4 ± 7.0	82.6 ± 8.4

[a] Mean values ± standard deviation (SD) of replicate assays (*n* = 8); [b] Mean values ± standard deviation (SD) of replicate assays (*n* = 3). [c] +/-, faint band. [d] LQ, limit of quantification. [e] nt- not tested. * Estimated amounts were obtained from previously diluted protein extracts.

high estimates of samples #2, #5, #7 and #14. The low estimated amounts for surimi samples (#10 and #11) with both kits are according to the real-time PCR results, highlighting again the use of flavours instead of crustacean meat. The results for the samples declaring the precautionary labelling for crustaceans (#17 and #18) were also at trace or non-quantifiable levels with both kits, agreeing with the PCR results.

Conclusion

In the present work, two new DNA-based approaches targeting the 16S rRNA mitochondrial gene are proposed to detect crustaceans in foods: a qualitative PCR assay specific for crustaceans (shrimps, lobsters and crabs) and a quantitative real-time PCR assay specific for shrimp crustaceans. The qualitative PCR assay can provide a simple, fast and high throughput tool for screening the presence of crustaceans in processed foods. The proposed real-time PCR system enabled the detection and quantification down to 0.1 pg and 0.0001% (w/w) of shrimp DNA and shrimp in model mixtures, respectively. The method exhibited high performance for quantitative analysis in the range of 0.0001% to 50% as inferred by the calibration curve parameters. The validation results with blind mixtures showed that the system was accurate and precise in the tested range.

Both qualitative PCR and real-time PCR methods were successfully applied to analyse commercial food samples and verify labelling compliance. Generally, the quantitative results suggest the use of shrimp flavour/extracts instead of shrimp meat. The comparison of shrimp estimates with two ELISA kits highlights that, at trace amounts of crustaceans, real-time PCR and ELISA are well correlated. However, for the higher contents (crustaceans or shrimps as ingredients), both ELISA kits display much higher estimates than those of real-time PCR, suggesting potential cross-reactivity of immunoassays with other food components, such as molluscs and spices.

Acknowledgements

This work was supported by FCT (Fundação para a Ciência e Tecnologia) through project UID/QUI/50006/2013 – POCI/01/0145/FEDER/007265 with financial support from FCT/MEC through national funds and co-financed by FEDER, under the Partnership Agreement PT2020 and by the project NORTE-01-0145-FEDER-000011. T. J. R. Fernandes and J. Costa are grateful to FCT grants (SFRH/BD/93711/2013 and SFRH/BPD/102404/2014, respectively) financed by POPH-QREN (subsidised by FSE and MCTES). The authors acknowledge the kind supply of shrimp species by Marfresco (Loures, Portugal) and Brasmar Seafood Companies (Trofa, Portugal).

References

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., & Wittwer, C. T. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, *55*, 611-622.

Cao, J., Yu, B., Ma, L., Zheng, Q., Zhao, X., & Xu, J. (2011). Detection of Shrimp-Derived Components in Food by Real-Time Fluorescent PCR. *Journal of Food Protection, 74*, 1776-1781.

Costa, J., Oliveira, M. B. P. P., & Mafra, I. (2013). Effect of thermal processing on the performance of the novel single-tube nested real-time PCR for the detection of walnut allergens in sponge cakes. *Food Research International*, *54*, 1722-1729.

Costa, J., Ansari, P., Mafra, I., Oliveira, M. B. P. P., & Baumgartner, S. (2014). Assessing hazelnut allergens by protein- and DNA-based approaches: LC-MS/MS, ELISA and real-time PCR. *Analytical and Bioanalytical Chemistry, 406*, 2581-2590.

Costa, J., Fernandes, T. J. R., Villa, C., Oliveira, M. B. P. P., & Mafra I. (2017). Advances in food allergen analysis. In U.G. Spizzirri, & G. Cirillo (Eds.), *Food safety: innovative analytical tools for safety assessment* (pp. 305-360). Beverly, MA, USA: Scrivener Publishing LLC (Wiley).

Daul, C. B., Morgan, J. E., & Lehrer, S. B. (1993). Hypersensitivity reactions to crustacea and mollusks. *Clinical Reviews in Allergy, 11*, 201-222.

Directive 2007/68/EC of 27 November 2007 amending Annex IIIa to Directive 2000/13/EC regarding certain food ingredients. Official Journal of the European Union, L310, 11–14.

Eischeid, A. C., Kim, B.-h., & Kasko, S. M. (2013). Two quantitative real-time PCR assays for the detection of Penaeid shrimp and blue crab, crustacean shellfish allergens. *Journal of Agricultural and Food Chemistry, 61*, 5669-5674.

Eischeid, A. C. (2016). Development and evaluation of a real-time PCR assay for detection of lobster, a crustacean shellfish allergen. *Food Control, 59*, 393-399.

Eischeid, A. C., & Stadig, S. R. (2018). A group-specific, quantitative real-time PCR assay for detection of crab, a crustacean shellfish allergen, in complex food matrices. *Food Chemistry, 244*, 224-231.

ENGL (2015). Definition of minimum performance requirements for analytical methods of GMO testing. European network of GMO laboratories, Joint Research Center, EURL Available online at <http://gmo-crl.jrc.ec.europa.eu/doc/MPR%20Report%20Application%2020_10_2015.pdf>. Last accessed on July 28, 2017.

FAO (2017). Fisheries and aquaculture statistics 2015, FAO Yearbook. Food and Agriculture Organization of the United Nations, Rome, Italy. Available at: http://www.fao.org/3/a-i7989t.pdf (Last accessed 11 December 2017).

Fernandes, T. J. R., Costa, J., Oliveira, M. B. P. P., & Mafra, I. (2015). An overview on fish and shellfish allergens and current methods of detection. *Food and Agricultural Immunology, 26*, 848-869.

Fernandes, T. J. R., Silva, C. R., Costa, J., Oliveira, M. B. P. P., & Mafra, I. (2017). High resolution melting analysis of a COI mini-barcode as a new approach for Penaeidae shrimp species discrimination. *Food Control, 82*, 8-17.

Gomaa, A., & Boye, J. (2015). Simultaneous detection of multi-allergens in an incurred food matrix using ELISA, multiplex flow cytometry and liquid chromatography mass spectrometry (LC–MS). *Food Chemistry, 175*, 585-592.

Günther, B., Raupach, M. J., & Knebelsberger, T. (2017). Full-length and mini-length DNA barcoding for the identification of seafood commercially traded in Germany. *Food Control, 73*, 922-929.

Herrero, B., Vieites, J. M., & Espiñeira, M. (2012). Fast real-time PCR for the detection of crustacean allergen in foods. *Journal of Agricultural and Food Chemistry, 60*, 1893-1897.

Khora, S. S. (2016). Seafood-associated shellfish allergy: a comprehensive review. *Immunological Investigations, 45*, 504-530.

Lopata, A. L., O'Hehir, R. E., & Lehrer, S. B. (2010). Shellfish allergy. *Clinical and Experimental Allergy, 40*, 850-858.

Mäde, D., & Rohmberger, D. (2017). Development of sensitive and specific real-time PCR systems for the detection of crustaceans in food. *European Food Research and Technology, 243*, 2105-2113.

Mafra, I., Ferreira, I. M. P. L. V. O., & Oliveira, M. B. P. P. (2008). Food authentication by PCR-based methods. *European Food Research and Technology, 227*, 649-665.

Pascoal, A., Barros-Velázquez, J., Ortea, I., Cepeda, A., Gallardo, J. M., & Calo-Mata, P. (2011). Molecular identification of the black tiger shrimp (*Penaeus monodon*), the white leg shrimp (*Litopenaeus vannamei*) and the Indian white shrimp (*Fenneropenaeus indicus*) by PCR targeted to the 16S rRNA mtDNA. *Food Chemistry, 125*, 1457-1461.

Pedrosa, M., Boyano-Martínez, T., García-Ara, C., & Quirce, S. (2015). Shellfish allergy: a comprehensive review. *Clinical Reviews in Allergy & Immunology, 49*, 203-216.

Prado, M., Ortea, I., Vial, S., Rivas, J., Calo-Mata, P., & Barros-Velazquez, J. (2016). Advanced DNA- and protein-based methods for the detection and investigation of food allergens. *Critical Reviews in Food Science and Nutrition, 56*, 2511-2542.

Regulation (EU) No 1169/2011 of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No. 1924/2006 and (EC) No. 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004. Official Journal of the European Union, L304, 18–63.

Thalayasingam, M., & Lee, B. W. (2015). Fish and shellfish allergy. *Chemical Immunology and Allergy*, *101*, 152-161.

Wang, W., Han, J., Wu, Y., Yuan, F., Chen, Y., & Ge, Y. (2011). Simultaneous detection of eight food allergens using optical thin-film biosensor chips. *Journal of Agricultural and Food Chemistry, 59*, 6889-6894.

Wilwet, L., Jeyasekaran, G., Shakila, R. J., Sivaraman, B., & Padmavathy, P. (2018). A single enzyme PCR-RFLP protocol targeting 16S rRNA/tRNAval region to authenticate four commercially important shrimp species in India. *Food Chemistry, 239*, 369-376.

Zagon, J., Schmidt, J., Schmidt, A. S., Broll, H., Lampen, A., Seidler, T., & Braeuning, A. (2017). A novel screening approach based on six real-time PCR systems for the detection of crustacean species in food. *Food Control, 79*, 27-34.

Zhenxing, L., Yiqun, Z., Hong, L., Haider, S., & Jie, J. (2010). Quantitative analysis of shrimp allergen in food matrices using a protein chip based on sandwich immunoassay. *European Food Research and Technology, 231*, 47-54.

FINAL REMARKS

The potential for seafood fraud and the actual cases of seafood mislabeling have increased in recent years, as a result of fishery industry expansion. Likewise, as a consequence of the current high rates of seafood consumption, allergic reactions after seafood ingestion have been widely reported. In this sense, efficient analytical methodologies must ensure the control of seafood species authenticity and the assessment of potential seafood allergen presence. The main advantages of using DNA-based tools for seafood analysis rely mostly on its ability to circumvent the problems associated with morphological identification and food processing. The current research project aimed at providing highly sensitive and specific PCR-based methods for the reliable detection, discrimination and quantification of relevant fish and crustacean species, covering both authenticity assessment and allergen analysis purposes. It is important to highlight that many efforts were effectively done and were not fully described in this thesis, namely in choosing the most adequate method and conditions for DNA extraction and evaluating the most suitable DNA region (nuclear and mitochondrial) for primer and/or probe designing.

Almost 20% of the world's fish catch regards the Gadiform order, which include Gadidae and Merluccidae families, two of the most relevant fish groups in terms of market value and traditional nutrition habits worldwide. *G. morhua* and *M. merluccius* are the most commercially important Gadidae and Merluccidae species, respectively, and they are often mislabeled due to species substitution phenomena. For the specific detection and discrimination of gadoid species, two DNA mini-barcodes (COI and *cytb*) were targeted using real-time PCR coupled with HRM. The *cytb* system enabled the best performance in discriminating/clustering the four gadoid species and was successfully applied to commercial fish products. 30% of the positive samples for gadoid species did not comply with the declared species, namely caused by the substitution of *G. morhua* for *T. chalcogramma*. Furthermore, a COI barcode real-time PCR system combined with HRM analysis was effectively developed for the differentiation of hake species providing a sensitive method to detect Merluccidae species down to 0.2-20 pg of hake DNA. From the 45 fish-containing foodstuffs that were analyzed, two samples did not comply with the declared species in the label.

Crustaceans are one of the most appreciated seafood groups and Penaeidae family comprise more than 30% of the worldwide demand of crustaceans. *L. vannamei*, *P. monodon* and *F. indicus, M. affinis* and *M. kerathurus* are some of the most important penaeid shrimps and HRM targeting a COI mini-barcode also showed its discriminatory power in this seafood group with high levels of confidence (>99%). Regarding labelling compliance assessment, four samples out of 33 suggest adulterations based on the complete or partial replacement of the species declared on the label. These approaches

proved to be important advances in mislabeling and fish species substitution evaluation. Moreover, the proposed systems overcome some drawbacks associated with DNA sequencing, enabling the reduction of costs and time of analysis.

The growing number of mislabeled cases in seafood industry has also been accompanied by an increase in the number of reported cases of allergy to seafood. The ingestion or contact with fish and crustaceans may trigger IgE antibody-mediated reactions in sensitized individuals and, in order to help them to avoid certain seafood groups, the control and detection of trace levels of fish and crustaceans is of utmost importance. PCR-based methods may either target the allergenic protein, or indirectly detect a universal DNA region for the allergen presence, which is much more useful.

In this context, two real-time PCR systems based on EvaGreen dye and TaqMan probe were proposed and compared for detection and quantification of fish as a potential allergen in foods. Both systems were able to detect 0.01 pg of fish DNA with adequate real-time PCR performance parameters. Nonetheless, the probe system allowed a relative LOD down to 0.0001% (1 mg/kg) of cooked fish in béchamel sauce, while the EvaGreen approach showed a specific amplification only until a level of 0.05. The trueness of the obtained data, expressed as bias, was within ±25% for all tested levels (0.25-8%) with the probe system, while the error was too high for the highest tested level (8%) for EvaGreen system. In spite of the best performance of the probe system, the application of real-time PCR system with EvaGreen dye allowed to extend the number of detected fish species that could not be amplified with the former methodology. Additionaly, two new DNA-based approaches (qualitative and quantitative real-time PCR) were also proposed to detect crustaceans in foods: The qualitative PCR assay provide a simple and fast tool for screening the presence of crustaceans (shrimps, lobsters, crabs) in processed foods. The real-time PCR system allowed the detection and quantification down to 0.1 pg and 0.0001% (w/w) of shrimp DNA and shrimp in model mixtures, respectively. Similarly to the real-time PCR approach for fish detection, the validation results with blind mixtures showed that the system was accurate and precise in the tested range. Both qualitative PCR and real-time PCR methods were successfully applied to analyse commercial food samples and verify labelling compliance. Generally, the quantitative results suggest the use of shrimp flavour/extracts instead of shrimp meat. The comparison of shrimp estimates with two ELISA kits highlights that, at trace amounts of crustaceans, real-time PCR and ELISA are well correlated. However, for the higher contents (crustaceans or shrimps as ingredients), both ELISA kits display much higher estimates than those of real-time PCR, suggesting potential cross-reactivity of immunoassays with other food components, such as molluscs and spices. The present works proposed and validated novel PCR tools to quantify fish and shrimp DNA in processed foods at trace levels,

targeting the 16S rRNA mitochondrial region and can be considered as advances in the field of seafood allergen management approaches.

In conclusion, much research is still required aiming seafood species authentication in processed foodstuffs and the implementation of reliable analytical methods for the detection of minute quantities, which is relevant when considering seafood as potential allergens. Large collections of authenticated specimens are of utmost importance for methodology development and more data should be addressed concerning mollusk species detection and quantification, both by authentication tools as well as allergen detection approaches. As future perspectives, next-generation sequencing (NGS) appears to represent the most accurate response in the case of complex samples, for instance when species mixtures are present.