

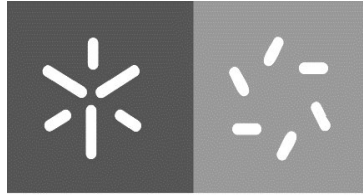


**Universidade do Minho**  
Escola de Ciências

Luis Manuel Coelho Oliveira

***DNA barcodes* de peixes marinhos da Europa:  
compilação de uma biblioteca de referência  
validada e investigação de padrões de  
divergência genética**





**Universidade do Minho**  
Escola de Ciências

Luís Manuel Coelho Oliveira

***DNA barcodes* de peixes marinhos da Europa:  
compilação de uma biblioteca de referência  
validada e investigação de padrões de  
divergência genética**

Dissertação de Mestrado

**Mestrado em Ecologia**

Trabalho efetuado sob a orientação do

**Professor Doutor Filipe José Oliveira Costa**

E sob a coorientação do

**Professor Doutor Pedro Soares**

## DECLARAÇÃO

**Nome:** Luís Manuel Coelho Oliveira

**Endereço eletrónico:** luismco12@gmail.com

**Telefone:** 918823832

**Cartão do Cidadão:** 13858442

**Título da dissertação:** *DNA barcodes* de peixes marinhos da Europa: compilação de uma biblioteca de referência validada e investigação de padrões de divergência genética

**Orientador:** Professor Doutor Filipe José Oliveira Costa

**Coorientador:** Professor Doutor Pedro Soares

**Ano de conclusão:** 2015

**Designação do Mestrado:** Mestrado em Ecologia

DE ACORDO COM A LEGISLAÇÃO EM VIGOR, NÃO É PERMITIDA A REPRODUÇÃO DE QUALQUER PARTE DESTA DISSERTAÇÃO

Universidade do Minho, \_\_\_\_/\_\_\_\_/\_\_\_\_

Assinatura: \_\_\_\_\_

## **AGRADECIMENTOS**

Um agradecimento em particular ao orientar, Doutor Filipe Costa, pela oportunidade de me deixar fazer um trabalho nesta área que me despertou interesse desde cedo no mestrado. Agradeço também pela paciência, até ao último minuto e pela disponibilidade e apoio que foi transmitindo ao longo deste trabalho.

Ao meu coorientado, Doutor Pedro Soares por ajudar com aquilo que tinha mais dificuldades e me sentia perdido e por ter apoiado as decisões tomadas neste trabalho. À Doutora Monica Landi, sempre disponível para ajudar, mesmo quando não lhe era possível. Um especial agradecimento ao Jorge Lobo e à Cláudia Hollatz, pelo companheirismo no laboratório e acima de tudo pela ajuda que prestaram, mesmo não estando envolvidos no trabalho.

Agradeço especialmente à minha mãe e ao Sérgio, por me ajudarem em tudo, dentro e fora deste trabalho, sem eles não poderia estar neste posição e ter acabado este ciclo de estudos. Ao resto da família, às tias, tios e avós, por me incentivarem a continuar a estudar e trabalhar sempre mais.

À Ilisa, em especial, por me aturar a todas as horas, dentro e fora do trabalho, sempre disponível para ajudar e perceber. Aos restantes colegas de laboratório (Ana Paula, Ana Sofia, Maria Luís) e do mestrado, em especial a Joana e a Mafalda, que fizeram com que estes dois anos passassem de uma melhor forma, sempre com companheirismo.

Este trabalho foi financiado por Fundos FEDER através do Programa Operacional de Factores de Competitividade - COMPETE e por Fundos Nacionais através da FCT "Fundação para a Ciência e a Tecnologia (FCT)" no âmbito dos projetos FCOMP-01-0124-FEDER-007381, FCOMP-01-0124-FEDER-010596 e PEst-C/BIA/UI4050/2011



# **DNA barcodes de peixes marinhos da Europa: compilação de uma biblioteca de referência validada e investigação de padrões de divergência genética**

## **RESUMO**

Nos últimos anos foram publicadas várias bibliotecas de referência de *DNA barcodes* com foco regional para peixes marinhos europeus, mas ainda existe a ausência de uma avaliação global do progresso da compilação de uma biblioteca de referência para a ictiofauna europeia. Neste estudo foi reunida pela primeira vez uma biblioteca de referência abrangente e de grande escala para este grupo de organismos, a partir de todos os *DNA barcodes* publicamente disponíveis, com o objetivo de examinar e anotar a consistência e a confiança das amostras obtidas independentemente por vários estudos de várias regiões geográficas distintas.

Foi compilado neste estudo um numeroso *dataset* compreendendo 4118 *DNA barcodes* de espécies de peixe marinhos amostrados na Europa, que foram reunidos a partir de 18 projetos na base de dados BOLD (*Barcode of Life Data System*), no total de 13 publicações e representando 358 espécies. Foi gerado um relatório de discordância de BINs (*Barcode Index Number*) para as sequências do *dataset*, de forma a atribuir unidades taxonómicas operacionais moleculares ao conjunto de sequências, produzindo 366 BINs dos quais 213 foram concordantes (1 BIN = 1 espécie), 141 discordantes e 12 *singletons* (1 BIN = 1 sequência). Inspeção pormenorizada da composição de cada BIN revelou a presença de potenciais artificios (sinónimos, más identificações), resultando num máximo de 73% de identificações taxonómicas concordantes. Um número considerável de espécies com importância económica foram encontradas nas 14% das espécies com identificação ambígua. Quinze espécies apresentaram distância intraespecíficas relativamente elevadas, atingindo 18,5% e foram atribuídas a 36 BINs (entre 2 a 4 BINs por espécie).

Este estudo demonstrou que apenas cruzando dados de múltiplas fontes, num processo de compilação de uma biblioteca de referência de escala continental, foi possível desvendar casos pertinentes de incerteza taxonómica e diversidade específica oculta que de outra forma permaneceria despercebida. Os casos de profunda estrutura intraespecífica encontrados constituem informação significativa que precisa de ser anotada e considerada no monitoramento das espécies. O aperfeiçoamento do conhecimento da diversidade, o controlo de pescas ilegais e a autentificação dos produtos pescados são algumas das aplicações mais relevantes desta biblioteca de referência.

Palavras-Chave: Peixes, Europa, *DNA barcodes*





# **DNA barcoding european marine fishes: assemble of a validated reference library and research into genetic diversity patterns**

## **ABSTRACT**

Over the last few years, several comprehensive reference libraries of DNA barcodes for marine fishes of Europe have been published with regional focus, but a global appraisal of the progress of the compilation of a reference library for European marine ichthyofauna is still missing. Here was assemble for the first time a large-scale comprehensive reference library for this ichthyofauna, based on all publicly available DNA barcodes, with the aim to examine and annotate consistency and reliability of records obtained independently from multiple regions and studies.

Was assembled here a large dataset comprising 4118 DNA barcodes from fish specimens collected in Europe, which were mined from 18 BOLD projects in a total of 13 publications and representing 358 species. A BIN (Barcode Index Number) Discordance Report was generated for the BOLD dataset, producing 366 BINs of which 213 were concordant (1 BIN= 1 species), 141 discordant and 12 were singletons (1 BIN = 1 sequence). Subsequent inspection of the BIN composition revealed potential artifacts (i.e. synonyms, misidentifications), resulting in a maximum of 73% concordant species IDs. A number of economically important species such as mackerel, salmonids and various sharks are included in the 14% of species with ambiguous IDs. Sixteen species displayed comparatively high intraspecific divergences, up to 18.5%, and were assigned to as much as 36 BINs (ranging from 2 to 4 BINs per species).

Here is shown that only by crossing data from multiple sources, in the process of assembling an European-scale library, it was possible to unravel pertinent cases of taxonomic uncertainties and hidden species diversity that otherwise would have remain unnoticed. The cases of deep within-species genetic structure detected within the European region constitute biologically meaningful information that needs to be annotated and considered in fish species monitoring. Improving the knowledge about the diversity, illegal fisheries control and authentication of fish products are some of the most important applications of this reference library.

Keywords: Fishes, Europe, DNA barcodes



# ÍNDICE

Agradecimentos.....	iii
Resumo.....	v
Abstract.....	vii
Lista de Figuras.....	xi
Lista de Tabelas.....	xiii
1. Introdução.....	1
1.1 Diversidade e relevância da ictiofauna marinha europeia.....	1
1.2 <i>DNA barcodes</i> na identificação taxonómica de espécies.....	3
1.3 Campanha FISH-BOL.....	5
1.4 Motivação e pertinência do estudo.....	6
1.5 Objetivos.....	7
2. Metodologia.....	9
2.1 Compilação das sequências de COI-5P.....	9
2.2 Análise das discordâncias entre as unidades taxonómicas operacionais moleculares (MOTUs) e a morfologia na identificação das espécies.....	10
2.3 Atribuição de classificação de fiabilidade taxonómica aos <i>DNA barcodes</i> .....	11
2.4 Tratamento e análise de dados.....	13
3. Resultados.....	15
3.1 Biblioteca de referência.....	15
3.2 Alocação de espécimes em BINs.....	17
3.3 Classificação de fiabilidade taxonómica aos <i>DNA barcodes</i> .....	18
3.4 Análise das distâncias genéticas.....	18
3.5 Casos de distância intraespecífica comparativamente elevada.....	19
3.6 Reconstrução da filogenia por inferência bayesiana.....	22
3.7 Análise da composição e variabilidade nucleotídica.....	23
4. Discussão.....	27
4.1 Fiabilidade da biblioteca de referência para peixes marinhos europeus.....	27
4.2 Divergência intraespecífica comparativamente elevada.....	28
4.3 Padrões de variabilidade genética.....	30

4.4	Importância e utilidade da biblioteca de referência.....	31
5.	Considerações finais.....	33
	Referências.....	35
	Anexo I.....	41

## LISTA DE FIGURAS

Figura 1 – Exemplo de página de um BIN (disponível em <a href="http://www.boldsystems.org/">http://www.boldsystems.org/</a> ).....	11
Figura 2 – Distribuição geográfica das amostras da biblioteca de referência que dispunham de coordenadas GPS (WGS84).....	16
Figura 3 – Representação gráfica da distribuição taxonómica dos espécimes da biblioteca de referência.....	17
Figura 4 – Partes da árvore NJ nucleotídica das espécies <i>Belone belone</i> , <i>Lepidorhombus whiffiagonis</i> , <i>Diplodus annularis</i> , <i>Dicentrarchus labrax</i> . A legenda de cada nóculo representa a distribuição geográfica, o número de espécimes.....	22
Figura 5 – Inferência bayesiana das sequências da biblioteca de referência. As cores interiores representam as três classes de organismos. A coloração exterior pretende representar as diferentes ordens dentro das respetivas classes.....	22
Figura 6 – Composição em GC de todas as posições dos codões das sequências da biblioteca de referência (A) e apenas da primeira (B), segunda (C) e terceira (D) posições dos codões das sequências nucleotídicas.....	23
Figura 7 – Representação gráfica dos valores de percentagem de conteúdo em GC nas duas classes de peixes mais abundantes da biblioteca de referência. Os espécimes estão representados em percentagem da sua representação na respetiva classe.....	25
Figura 8 – Composição em GC de todas as posições dos codões das sequências da classe Actinopterygii (A) e apenas da primeira (B), segunda (C) e terceira (D) posições dos codões das sequências nucleotídicas.....	26
Figura 9 – Composição em GC de todas as posições dos codões das sequências da classe Elasmobranchii (A) e apenas da primeira (B), segunda (C) e terceira (D) posições dos codões das sequências nucleotídicas.....	26



## LISTA DE TABELAS

Tabela 1 – Progresso do <i>DNA barcode</i> nas diversas classes de peixes (dados retirados de <a href="http://www.fishbol.org">http://www.fishbol.org</a> ) .....	5
Tabela 2 – Publicações com sequências públicas utilizadas na DS-EUROFISH, bem como o seu código da base de dados BOLD, a região e o número de sequências.....	15
Tabela 3 - Comparação da concordância taxonómica entre BINs sem revisão e após revisão .....	17
Tabela 4 - Atribuição de categorias às sequências das 345 espécies de peixes marinhos europeus....	18
Tabela 5 – Divergências médias, mínimas e máximas intraespecífica, congénérica e confamiliar de peixes marinhos europeus, calculadas segundo o modelo K2P. ....	19
Tabela 6 - Lista de espécies com distâncias intraespecíficas elevadas (excepto <i>Etmopterus annularis</i> ) e que foram alocadas em mais que um BIN. ....	19
Tabela 7 - Valores de percentagem de conteúdo em GC na totalidade da sequência e nas diferentes posições dos codões para todos os espécimes presentes na biblioteca de referência.....	24
Tabela 8 - Valores de percentagem de conteúdo em GC na totalidade da sequência e nas diferentes posições dos codões para todos os espécimes da classe Actinopterygii .....	24
Tabela 9 - Valores de percentagem de conteúdo em GC na totalidade da sequência e nas diferentes posições dos codões para todos os espécimes da classe Elasmobranchii .....	24
Tabela 10 - Número de acesso do GenBank de todas as sequências da biblioteca de referências para peixes marinhos europeus .....	41





# 1. INTRODUÇÃO

## 1.1 Diversidade e relevância da ictiofauna marinha europeia

O estudo da ictiofauna tem uma enorme importância, seja ela pelos impactos económicos ou pela sua importância na biodiversidade local e regional (Carneiro *et al.*, 2014). Em 2000 a pesca representava 15% da proteína animal consumida e é evidente a importância social das pescas, empregando cerca de 35 milhões de pessoas e gerando globalmente 81 mil milhões de dólares anualmente, sendo produtos de enorme importância para a alimentação humana (FAO, 2002) Esse facto é ainda mais evidente em países costeiros como Portugal em que a indústria emprega uma grande parte da população, chegando mesmo a ser, em 2011, o terceiro país da União Europeia com mais empregos no setor. Os recursos piscatórios fazem parte da dieta europeia e Portugal é o país que mais produtos das pescas e da aquacultura consumiu em 2010 (European Commission, 2014). Os peixes representam cerca de 50% de todas as espécies de vertebrados e estima-se que existam descritas até ao momento perto de 30 mil espécies, incluindo espécies marinhas e de água doce. São um grupo de organismos muito diverso e é representado desde espécies ancestrais de peixes sem mandíbula como a lampreia, peixes cartilagueiros como os tubarões e raias até peixes ósseos como os salmonídeos e o atum (Ward *et al.*, 2005). Apesar da abundância enorme de espécies de peixes, a maioria da pesca é focada numa percentagem reduzida das mesmas. Cerca de 75% do produto pescado globalmente consiste em 200 espécies, cerca de 1% da diversidade existente (Holmlund e Hammer, 1999). Para a Europa estão listadas entre 1220 (Nieto *et al.*, 2015) e 1349 (Costello *et al.*, 2006) espécies de peixes marinhos, sendo que 88% dessas espécies estão presentes na Zona Económica Exclusiva (ZEE) portuguesa, se considerarmos a zona proposta de extensão da plataforma continental. Portugal é deste modo um dos países da União Europeia mais ricos em ictiofauna, constituindo a costa portuguesa um importante ponto de encontro e mistura de espécies provenientes diversas áreas marinhas adjacentes. Considerando toda a área ocupada pela ZEE portuguesa, com um total de 1.727.408 km<sup>2</sup>, que engloba a área continental e dos arquipélagos da Madeira e dos Açores, Portugal é um dos países com o maior número de espécies de peixes marinhos na Europa, sendo que o inventário mais recente lista 1058 espécies (Carneiro *et al.*, 2014), apesar da base de dados FishBase listar apenas 844 espécies para o conjunto destas áreas marítimas (Froese e Pauly, 2010). No espaço marinho europeu é possível identificar algumas áreas diferenciadas como o nordeste Atlântico, o Mar Mediterrâneo, o Mar do Norte e o Mar Báltico. O Mar Mediterrâneo é particularmente importante no cenário europeu e representa 7% da ictiofauna marinha

global (Bianchi e Morri, 2000), sendo assim uma bacia de enorme importância na ictiofauna marinha europeia (Landi *et al.* 2014). As ligações com o oceano Atlântico pelo estreito de Gibraltar, ao oceano Índico pelo canal do Suez e com o Mar Negro fazem com que se caracterize por um espaço de confluência de um número grande de espécies, incluindo espécies endêmicas (Coll *et al.*, 2010; Azzurro *et al.*, 2011). O Mar do Norte é caracterizado por uma fauna mais específica, com um menor número de espécies, quando comparado com as restantes bacias europeias (Kneibelsberger *et al.*, 2014). No entanto é responsável por cerca de 5% da pesca global (MUMM, 2002).

Ao nível da relevância ecológica, os peixes são uma fonte de alimento essencial para os predadores de topo e desempenham um papel importante na regulação das cadeias tróficas (Costa *et al.*, 2012). Este grupo de organismo tem uma substancial importância na provisão de serviços de ecossistemas. Não só proporciona serviços de regulação, como a regulação das cadeias tróficas, reciclagem de nutrientes e redistribuição de substratos como proporciona serviços culturais como a produção de alimento e serviços de informação na avaliação da pressão e resiliência de um ecossistema (Holmlund e Hammer, 1999).

O papel destas espécies de peixes nos seus ecossistemas, tanto a nível das suas funções, manutenção ou recuperação só pode ser abordado com a correta identificação dessa mesma biodiversidade (Sutherland *et al.*, 2006). No entanto, até 2002 apenas 66% da totalidade do pescado capturado a nível mundial foi identificado corretamente até à espécie (Lleonart *et al.*, 2006). Estes números são facilmente explicados pela difícil identificação clássica dos espécimes, principalmente quando na presença de características morfológicas de difícil identificação ou na ausência de chaves taxonómicas atualizadas (Kochzius *et al.*, 2010). De acordo com um relatório de Daan (2001), sobre um programa de monitorização intensiva de espécies no Mar do Norte, podem mesmo existir discrepâncias no número de espécies identificadas quando a mesma é feita por diferentes países. Essa discrepância chegou mesmo a 30% e é consequência da utilização de chaves de identificação diferentes e na dificuldade na identificação de espécies morfológicamente similares, bem como estados juvenis (Kneibelsberger *et al.*, 2014). De mais difícil identificação são também a presença de espécies crípticas e a identificação de ovos e larvas (Costa e Carvalho, 2007). Todos estes impedimentos taxonómicos fazem com que as estatísticas do pescado possam ser falaciosas e que não traduzam a realidade dos mananciais pesqueiros, que podem posteriormente levar à errada gestão dos mesmos (Lleonart *et al.*, 2006). A conservação e a gestão da pesca das espécies fica comprometida com a falta de rigor na identificação das mesmas, particularmente quando não existem ferramentas para testar quão rigorosa são essas identificações. A acumulação de erros de identificação nas estatísticas dos produtos pescados

não é assim identificável e pode produzir estimativas de abundâncias e recrutamento erradas, conduzindo à ineficaz gestão e conservação das espécies marinhas capturadas (Costa *et al.*, 2012).

## **1.2 DNA barcodes na identificação taxonómica de espécies**

As ferramentas moleculares para a identificação e delimitação de espécies têm vindo a ser desenvolvidas e refinadas ao longo das últimas décadas, e são especialmente úteis quando a identificação tradicional por características morfológicas é ineficaz, ou a capacidade de identificar um grande número de amostras torna a tarefa logisticamente impraticável (Kneibelsberger *et al.*, 2014). O *DNA barcoding* (Hebert *et al.*, 2003a) é atualmente a metodologia molecular mais difundida e usada para a identificação de espécies. Esta metodologia baseia-se na comparação da sequência de *DNA* de uma região padronizada do genoma, para a identificação de espécies em grandes agrupamentos taxonómicos (e.g. *Animalia*, *Plantae*, *Fungi*). O *DNA barcode* padrão estabelecido para as espécies animais localiza-se na extremidade 5' da subunidade 1 do gene mitocondrial do citocromo c oxidase (COI) e tem um comprimento de 658 pares de bases (Hebert *et al.*, 2003b). A identificação de espécies animais a partir da sequência do *DNA barcode* é possível porque a variação da mesma entre indivíduos da mesma espécie é muito baixa quando comparada com a variação entre espécies. Desde 2003, o *DNA barcoding* é usado como marcador padrão para muitos grupos de organismos, especialmente espécies animais (Ward e Holmes, 2007).

A partir do momento da criação da base de dados *Barcode of Life Data System* (BOLD, Ratnasingham e Hebert, 2007) foi disponibilizada uma plataforma de acesso universal que permite a identificação rápida de espécimes a partir de sequências da região padronizada do gene COI (doravante referida como COI-5P). Esta plataforma permitiu também o desenvolvimento de uma comunidade especializada na utilização desta metodologia e fomentou o estabelecimento de padrões de qualidade dos dados depositados, permitindo assim uma melhor organização e correção da informação disponibilizada (Hanner, 2005). A biblioteca foi crescendo e melhorando ao longo dos anos e já conta com mais de 5 milhões de sequências depositadas, num total de mais de 350 mil espécies formalmente descritas (disponível em <http://www.boldsystems.org>). Esta metodologia revela ser bastante útil, uma vez que não só facilita a identificação, como permite também a identificação de espécimes nas várias fases do ciclo de vida (e.g. ovos, larvas), ou de adultos mesmo a partir de fragmentos (Costa e Carvalho, 2007; Ward *et al.*, 2009). A utilidade do *DNA barcodes* estende-se bastante além da facilitação da identificação de espécies, de forma mais rápida e rigorosa, consistindo também numa abordagem ao

estudo dos organismos que complementa e auxilia a taxonomia, a filogenia molecular e a genética de populações (Hajibabaei *et al.*, 2007). A importância para a catalogação da biodiversidade é inegável, mas talvez mais importante que isso será a relevância na elucidação desse catálogo. Um grande entrave para a catalogação da biodiversidade é a clarificação da taxonomia, o que torna difícil a correta identificação e delimitação das espécies. A base de dados BOLD usa já ferramentas para combater estas dificuldades. A criação de uma taxonomia integrativa, aproximando a taxonomia e a filogenia, para a criação de um código identificativo para cada espécie é o grande objetivo. A delimitação de unidades taxonômicas operacionais a partir da análise de dados moleculares é prática corrente (Ratnasingham e Hebert, 2013). Estas abordagens permitem, por exemplo, sinalizar casos de potenciais espécies crípticas e de complexos de espécies. Além da identificação padronizada, simples e universal, o *DNA barcode* possibilita também a identificação de espécies crípticas, a clarificação de ambiguidades taxonômicas e a identificação de material histórico presente em museus (Costa e Carvalho, 2007). O poder discriminatório dos *DNA barcodes* é normalmente elevado. Estudos anteriores com peixes marinhos mostram que o baixo nível de variação intraespecífica permite a correta identificação de mais de 98% das espécies estudadas (Ward *et al.*, 2005; Costa *et al.*, 2012; Landi *et al.*, 2014). Casos particulares de espécies com elevadas divergências intraespecíficas podem ser evidência da presença de diversidade críptica, que se pode expressar no aumento das espécies conhecidas, bem como a diferenciação das populações de uma determinada espécie (Knebelsberger *et al.*, 2014).

Todos os genomas retêm características ancestrais e mostram uma história evolutiva da espécie. No entanto a sequenciação do genoma completo de todas as espécies pode não ser a melhor abordagem para pesquisar padrões de evolução molecular, seja pela quantidade de informação que é gerada e não consegue ser analisada, seja pela redundância de algumas partes do genoma. A utilização de *DNA barcodes* para a investigação de padrões de divergência molecular tem como vantagens o grande alcance a nível de espécies e a facilidade de sequenciação. Alguns estudos já investigaram padrões de divergência molecular recorrendo a *DNA barcodes* (Kerr *et al.*, 2009; McDevit e Saunders, 2010). No entanto existem alguns grupos taxonômicos onde as taxas de substituição nucleotídica do gene COI são muito baixas (Cárdenas *et al.*, 2005) ou muito altas (Johnson *et al.*, 2003), comprometendo o sucesso da análise dos *DNA barcodes*. As aplicações do *DNA barcoding* não passam apenas pela identificação das espécies. Esse é apenas o primeiro passo para a integração desta metodologia em estudos ecológicos. A gestão dos quotas de pescas, pela delimitação do manancial das espécies/populações e pelo acompanhamento do recrutamento através da análise de ovos e larvas, e a fiscalização de capturas ilegais e a análise de

dietas são algumas das aplicações das bibliotecas de referência de *DNA barcodes* (Cárdenas *et al.*, 2005; Costa *et al.*, 2012; Keskin e Atar, 2013).

### 1.3 Campanha FISH-BOL

Integrada na iniciativa BOLI (*Barcode of Life Initiative*), a *Fish Barcode of Life* – FISH-BOL (<http://www.fishbol.org>) é a campanha criada em 2005 para a criação de uma biblioteca de referência de *DNA barcodes* para todos os peixes. Até à data foram realizadas sequências de 11050 (Setembro de 2015) espécies de peixes. Outro dado interessante são os 2384 *clusters* de *DNA barcodes* sem identificação, que podem incluir espécies não descritas. Estas 11 mil espécies de peixes representam apenas 34% das espécies descritas até ao momento. O progresso das diferentes classes de peixes é diferente e é notório que o esforço de amostragem incide mais na classe mais abundante (Actinopterygii) que contem mais de 90% as espécies com *DNA barcodes* (Tabela 1).

**Tabela 1 – Progresso do *DNA barcode* nas diversas classes de peixes (dados retirados de <http://www.fishbol.org>)**

<b>Classe</b>	<b><i>Barcoded</i></b>	<b>Espécies</b>	<b>Progresso</b>
Actinopterygii	10329	30904	33%
Cephalaspidomorphi	31	43	72%
Elasmobranchii	638	1178	54%
Holocephali	33	50	66%
Myxini	17	74	23%
Sarcopterygii	2	8	25%
<b>Total</b>	<b>11050</b>	<b>32257</b>	<b>34%</b>

Mais do que catalogar e acompanhar o progresso das espécies com *DNA barcodes*, a FISH-BOL tem como objetivo e responsabilidade a organização dos esforços de vários grupos regionais, permitindo a partilha global da informação (Ward *et al.*, 2009). O número de publicações de *DNA barcoding* em peixes tem vindo a aumentar, muito devido ao poder discriminatório do mesmo na identificação de espécies, como demonstrado em muitos estudos realizados em várias partes do globo (Ward *et al.*, 2005; Costa *et al.*, 2012; Keskin e Atar, 2013; Landi *et al.*, 2014; Knebelsberger *et al.*, 2014).

## 1.4 Motivação e pertinência do estudo

Ao longo dos últimos 5 anos foi publicado um número considerável de estudos envolvendo a utilização de *DNA barcodes* em peixes marinhos europeus. Pela sua abordagem, estes estudos poderão ser divididos em dois tipos: uns de âmbito mais regional e sem especificidade taxonómica, como são exemplos o estudo realizado no Mar do Norte (Knebelsberger *et al.*, 2014), no Mar Mediterrâneo (Landi *et al.*, 2014; Keskin and Atar, 2013) e na costa Portuguesa (Costa *et al.*, 2012). O outro tipo de estudo tem um âmbito mais taxonómico, focando-se apenas numa família ou género, ou até mesmo numa espécie apenas. Foram também realizados vários estudos deste segundo tipo na Europa, como são exemplos o trabalho realizado sobre tubarões de águas profundas (Moura *et al.*, 2008), sobre góbios (Knebelsberger and Thiel, 2014) ou raias (Lynghammar *et al.*, 2014; Serra-Pereira *et al.*, 2011). Contudo, apesar do número considerável de estudos efetuados nos últimos 5 anos, em nenhum deles é realizada uma síntese global dos dados gerados, nomeadamente envolvendo a compilação e anotação de uma biblioteca de referência de *DNA barcodes* de peixes marinhos a nível europeu. Deste modo é da maior pertinência coligir e confrontar a informação dispersa em múltiplos estudos, por forma a iniciar a construção de uma biblioteca de referência de larga escala para peixes marinhos europeus, a partir de todos os códigos de barras de *DNA* publicamente disponíveis para a região. O objetivo é analisar e anotar a consistência e fiabilidade de *DNA barcodes* obtidos de forma independente por vários grupos de investigação. Apesar da distância geográfica entre os espécimes desta ampla amostragem, resultados de estudos anteriores mostram a capacidade do *DNA barcode* para o diagnóstico de espécies de peixes marinhos, independentemente da distância geográfica (Ward *et al.*, 2008; Zemlak *et al.*, 2009).

Devido à automação das análises das sequências nas bases de dados públicas como a BOLD e o GenBank, estas estão sujeitas a erros intrínsecos a alguma falha na cadeia de processamento dos *DNA barcodes* (Knebelsberger *et al.*, 2014). Estes erros podem ocorrer em qualquer etapa da metodologia, desde o pré ao pós *barcoding*. No pré *barcoding* a fonte da falha é a identificação morfológica do espécime e pela falta de dados associados ao mesmo. Na fase laboratorial do *barcoding* a contaminação e a incorreta etiquetagem de amostras podem ser a causa de falhas na metodologia. No pós *barcoding* os erros passam pela não verificação das sequências. É impreterível que sejam criadas medidas de controlo da metodologia para todas estas etapas. Para o pré *barcoding* passa por assegurar que as amostras são identificadas por taxonomistas certificados e que os dados requeridos pelo *Barcode Data Standards* (Hanner, 2005) são seguidos. O controlo das condições laboratoriais é essencial para uma maior confiança do resultado das amostras. No pós *barcoding* as ferramentas de anotação e validação das bibliotecas são então a chave para a garantia e manutenção da qualidade das mesmas. Um exemplo

é o sistema de classificação proposto por Costa e colaboradores (Costa *et al.*, 2012), que tem como objetivo atribuir categorias de fiabilidade taxonómica aos *DNA barcodes*. Estas ferramentas aliadas às análises automatizadas garantem a qualidade da biblioteca, servindo o propósito do utilizador final da biblioteca, seja ele especialista ou não. Desta forma o utilizador tem uma indicação do grau de fiabilidade da sua identificação, podendo usar a biblioteca com mais confiança (Knebelsberger *et al.*, 2014). Tal biblioteca representa assim um recurso válido e robusto para a identificação de sequências desconhecidas (Costa *et al.*, 2012). As suas aplicações vão além da identificação de espécimes nos seus diferentes estados de vida ou mesmo de partes de organismos. Uma biblioteca de referência de *DNA barcodes* constitui um suporte para as mais diversas aplicações como a autenticação de produtos piscatórios (Hanner *et al.*, 2011), desde peixe fresco, produtos enlatados, e mesmo peixe já confeccionado (Carvalho *et al.*, 2015). A biossegurança e a deteção do uso ilegal de espécies protegidas pode ser também uma das aplicações (Armstrong e Ball, 2005; Rasmussen e Morrissey, 2008). O controlo e a gestão das pescas é uma das aplicações mais importantes da biblioteca em questão (Costa e Carvalho, 2007). Tendo em conta o contexto integrado europeu, esta biblioteca pode ser importante no que toca à delimitação e gestão dos mananciais pesqueiros, principalmente pelo facto de muitos destes serem partilhados pelos estados membros (Landi *et al.*, 2014). A análise destes dados com alcance geográfico tão amplo pode revelar padrões de divergência e distribuição entre espécies ou mesmo dentro das mesmas, dando suporte para a clarificação de taxonomia, elucidar fatores de impacto demográficos e até mesmo identificar padrões de resposta a impactos antropogénicos (Knebelsberger *et al.*, 2014; Landi *et al.*, 2014). As bibliotecas de referência de *DNA barcodes*, devidamente anotadas e validadas, são essenciais para que seja possível a identificação rigorosa de espécimes através desta metodologia (Costa *et al.*, 2012).

## 1.5 Objetivos

Os objetivos desta tese passam pela compilação e anotação de uma biblioteca de referência de *DNA barcodes* para peixes marinhos europeus, bem como a análise da mesma:

- Compilação e anotação de uma biblioteca de referência de *DNA barcodes* para peixes marinhos da Europa, a partir das amostras analisadas pelo grupo de investigação e de sequências disponíveis em bases de dados públicas.
- Examinar e anotar a congruência entre as identificações baseadas na morfologia e a as unidades taxonómicas operacionais de raiz molecular ("*molecular operational taxonomic units*" - MOTUs)

após análise de sequências de *DNA barcodes* obtidas independentemente de múltiplas regiões e estudos.

- Categorizar a fiabilidade taxonómica dos registos da biblioteca de referência de forma a fornecer uma indicação do grau de confiança das identificações aos utilizadores.
- Comparação dos padrões de divergência molecular a diferentes níveis da hierarquia taxonómica, entre grupos taxonómicos (ex: género, família, ordem) e entre ictiofaunas de diferentes regiões.



## 2. METODOLOGIA

### 2.1 Compilação das sequências de COI-5P

O primeiro passo para a compilação desta biblioteca foi a criação de um *dataset* (DS-EUROFISH) na base de dados BOLD. Esta ferramenta permite a compilação de registos de espécimes, e respetivas sequências, de vários projetos num só local, permitindo subsequentes análises filogenéticas, de distribuição, composição e qualidade sem movimentar efetivamente os registos. Foram prontamente adicionadas as sequências previamente geradas pelo grupo de investigação, de projetos que englobavam amostras do Atlântico (Portugal, Reino Unido) e do Mar Mediterrâneo. Ainda numa primeira fase foram adicionadas as sequências provenientes de projetos de um grupo de investigação alemão que colaborou em estudos anteriores, sendo essas amostras provenientes do Mar do Norte e do Mar Báltico. O passo seguinte foi a procura, na base dados BOLD, de projetos que focassem o seu estudo em peixes marinhos na Europa. Esta procura foi realizada em duas etapas: primeiramente a procura foi focada em estudos de carácter geográfico, onde a amostragem não fosse limitada a nível taxonómico. De seguida a procura foi alargada a qualquer estudo que continha *DNA barcodes* de peixes marinhos europeus e que fossem de acesso livre. O último passo foi a adição de sequências de peixes marinhos europeus presentes no GenBank que estavam associados a publicações

A compilação de um número considerável de amostras sem qualquer tipo de filtragem poderia levar à introdução de dados que carecessem de qualidade. Por esse facto o próximo passo foi a validação dos *DNA barcodes*. Para isso foram seguidos os critérios de qualidade sugeridos para sequências de COI (adaptado de Hanner, 2005):

#### **Critérios obrigatórios:**

- a) Incluir um nome de espécie documentado;
- b) Incluir o país de origem;
- c) Incluir pelo menos 500 pares de bases contíguos e não ambíguos de sequenciação bidirecional;
- d) Não incluir mais que 1% de caracteres ambíguos na sequência;
- e) Incluir o nome da região genómica usada;

#### **Critérios recomendados:**

f) Incluir dados associados aos espécimes, registo fotográfico, localização GPS, identificador da amostra e detalhes sobre o ciclo de vida do organismo;

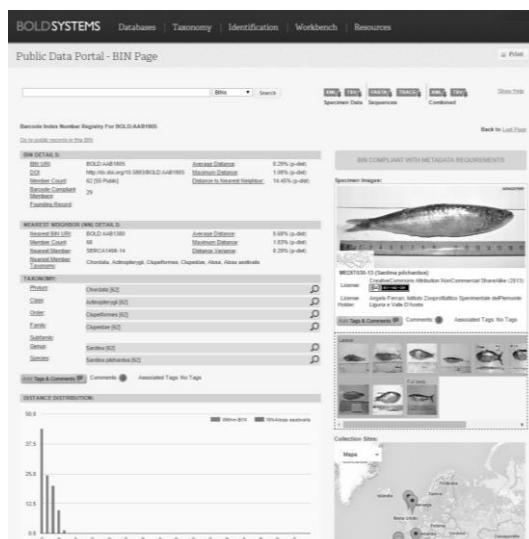
g) Incluir *trace files* submetidos ao NCBI *Trace Archive* ou ao *Ensembl Trace Server*;

h) Incluir sequências dos *primers*, *forward* e *reverse*, ou *cocktail* de *primers* usados.

Estes critérios foram aplicados a todas as sequências compiladas e as que não satisfizeram os mesmos foram eliminadas.

## **2.2 Análise das discordâncias entre as unidades taxonómicas operacionais moleculares (MOTUs) e a morfologia na identificação das espécies**

Para a atribuição de unidades taxonómicas operacionais moleculares (MOTUs) às amostras foi utilizado o “*Barcode Index Number*” (BIN), que é o sistema integrado na base de dados BOLD que agrupa sequências usando um algoritmo com base nas distâncias genéticas entre as mesmas. Cada BIN é único, corresponde a uma MOTU, contem sequências geneticamente idênticas, é registado e recebe uma página na base de dados BOLD (Figura 1). Esta página contem informação molecular e taxonómica, bem como dados como os registos fotográficos e localização geográfica das amostras (Ratnasingham e Hebert, 2013). Foi gerado um relatório de discordância BIN disponível na base de dados BOLD para o conjunto de sequências presentes na DS-EUROFISH. O relatório utiliza o algoritmo para separar as sequências em BINs e de seguida classifica cada BIN em concordante se o mesmo conter apenas sequências da mesma espécie (1 BIN = 1 espécie), discordante se forem atribuídas sequências de mais que uma espécie a um mesmo BIN ou *singleton* se um BIN for constituído por apenas uma sequência.



**Figura 1 – Exemplo de página de um BIN (disponível em <http://www.boldsystems.org/>)**

## 2.3 Atribuição de classificação de fiabilidade taxonómica aos *DNA barcodes*

A fim de examinar a fiabilidade taxonómica das amostras foi atribuído a cada espécie uma categoria de A (maior concordância) a E (menor concordância), consoante o nível de confiança na identificação (Costa *et al.*, 2012). As classificações são atribuídas com base na concordância das identificações obtidas e na quantidade e qualidade de informação existente na base de dados. O sistema de anotações é essencial para a criação de bibliotecas de referência de *DNA barcodes* (Costa *et al.*, 2012; Kneibelsberger *et al.*, 2014). As categorias atribuídas são as seguintes (adaptado de Costa *et al.*, 2012):

Categoria A – Concordância externa: Correspondência de BIN inequívoca, com amostras de outros projetos e sequências publicadas.

Categoria B – Concordância interna: Concordância de BIN com os dados do próprio projeto quando estão disponíveis, pelo menos, três espécimes. Sem sequências correspondentes na base de dados BOLD.

Categoria C – Concordância subótima (possível estrutura genética da espécie): pelo menos três espécimes da mesma morfoespécie disponíveis mas as mesmas dividem-se em mais que um BIN vizinho.

Categoria D – Dados insuficientes: baixo número de espécimes analisados (um ou dois indivíduos e sem correspondência na base de dados BOLD).

Categoria E – Correspondência discordante: sequências de uma determinada espécie não correspondem ao BIN ou BINs para essa mesma espécie na base de dados BOLD. Esses espécimes podem ter correspondência com um BIN de outra espécie ou em um BIN não vizinho.

Esta classificação consistiu na análise detalhada de cada BIN gerado no ponto anterior. Primeiro foram separados os BINs com resultado concordante do discordante. Pela sua natureza concordante, cada um desses BINs corresponde a uma espécie. Foram então separadas essas espécies em duas categorias (A e B) dependendo da disponibilidade ou não de amostras de estudos independentes. De seguida as espécies que estavam presentes em mais que um BIN concordante foram classificadas com a categoria “C”. Todas aquelas espécies com poucas amostras (menos de 3 sequências) e em que não existia mais quaisquer dados na base de dados BOLD foram classificadas como “D”. O segundo passo foi analisar os BINs discordantes, em que já não existe uma correspondência BIN – espécie. O passo seguinte foi verificar quais as espécies que estavam em cada BIN e atribuir, ainda que temporariamente, a categoria “E”. Foi apenas atribuído temporariamente porque cada BIN foi um caso específico. Por causa da forma como a análise automática é feita um BIN pode ser discordante mas a fiabilidade da identificação da espécie não fica comprometida. Isto acontece porque por vezes algum passo da *barcode pipeline* de uma amostra tem um falha. Este falha pode ser a simples troca de etiquetas entre amostras de extratos de *DNA* no laboratório até à incorreta identificação do *voucher*. Foram também encontrados nomes inválidos e/ou sinónimos nas identificações taxonómicas de algumas amostras, fazendo com que o BIN contivesse mais do que um nome de espécie, tornando o BIN discordante quando na realidade ele era concordante. Noutros casos o próprio algoritmo foi incapaz de resolver a fronteira entre as espécies. Este último caso mais incomum ocorreu quando dentro de um BIN ocorriam mais do que uma espécies (normalmente do mesmo género) mas a correta identificação das espécies não era afetada. A todas estas espécies que estavam nestes BINs com algum tipo de artefacto foram atribuídas as categorias mais elevadas (A, B ou C) dependendo dos critérios anteriores. A todas as outras espécies foi atribuído, agora definitivamente a categoria “E”.

## 2.4 Tratamento e análise de dados

Depois de compilada e anotada, a biblioteca de referência foi submetida a algumas análises disponíveis na base de dados BOLD. Essas análises incluem um sumário das distâncias genéticas, análise da composição nucleotídica e a análise do *barcoding gap*. As distâncias genéticas intraespecífica, congênica e confamiliar foram calculadas usando o modelo de substituição nucleotídica Kimura de dois parâmetros (K2P) (Kimura, 1980). A análise de composição nucleotídica calcula a porcentagem de cada nucleótido e pares de nucleótidos de todas as sequências, bem como as diferentes porcentagens dos mesmos nas três posições de um códon. A opção para a análise do *barcoding gap* compara as distâncias genéticas intraespecíficas e congênicas na procura de uma clara separação das mesmas num mesmo eixo. Os resultados foram de seguida descarregados, analisados e editados.

Posteriormente foi também efetuada uma árvore filogenética, uma inferência bayesiana e uma análise ao conteúdo em GC. Para essas análises fora da base de dados BOLD foram descarregadas as sequências da biblioteca de referência e prontamente alinhadas no programa MAFFT, versão 7 (Katoh e Standley, 2013) com a opção da escolha automática de algoritmo. O programa faz a escolha do algoritmo tendo em conta o número e a dimensão das sequências, tal como a similaridade das mesmas em tamanho. Esse alinhamento foi então cuidadosamente verificado e traduzido em aminoácidos no programa MEGA, versão 6 (Tamura *et al.*, 2013) para detetar a possível presença de inserções, deleções ou códon stop. Este alinhamento foi usado então para as análises posteriores. A árvore filogenética foi construída no programa MEGA, versão 6 (Tamura *et al.*, 2013), utilizando o algoritmo Neighbour-Joining (NJ) (Saitou e Nei, 1987) e o teste *bootstrap* (Felsenstein, 1985) baseado em 500 replicações para determinação do grau de suporte dos nós. A inferência bayesiana para a estimação da filogenia foi calculada no programa MrBayes, versão 3.2 (Ronquist *et al.*, 2012) utilizando o modelo de substituição nucleotídica GTR+I+G, previamente calculado no programa IQ-TREE, versão 1.3.0 (Nguyen *et al.*, 2014). As configurações usadas para o cálculo da inferência foram as aconselhadas para alinhamentos com este tipo de substituição nucleotídica, segundo o manual do próprio programa (disponível em [http://mrbayes.sourceforge.net/mb3.2\\_manual.pdf](http://mrbayes.sourceforge.net/mb3.2_manual.pdf)). A inferência usa simulações Monte Carlo via cadeias de Markov para a aproximação das probabilidades posteriores das árvores (Hastings, 1970; Metropolis *et al.*, 1953) de forma a estimar a filogenia (Yang e Rannala, 1997). As árvores geradas anteriormente foram manipuladas no programa Archaeopteryx, versão 0.9901 (disponível em <https://sites.google.com/site/cmzmasek/home/software/archaeopteryx>) O conteúdo em GC (Guanina – Citosina) foi calculado na aplicação *geecee* do pacote de aplicações EMBOSS (Rice *et al.*, 2000) A aplicação calcula a porcentagem de GC presente em cada uma das sequências, sem ter em conta a

posição do nucleótido no codão. Apesar da semelhança com a análise de composição nucleotídica presente na BOLD, esta permite saber exatamente a percentagem de GC de cada uma das sequências, enquanto a anterior apresenta apenas valores médios. Com esta informação o passo seguinte foi separar os resultados por classes e respetivas ordens e comparar as mesmas, na busca de padrões que associem espécies com resultados extremos (baixo e alto conteúdo em GC) e a sua taxa de substituição nucleotídica no gene estudado.

Foi contruído também um mapa representativo da distribuição geográfica das amostras da biblioteca de referência. Aqui foi utilizado o programa QGIS, versão 2.10 (QGIS Development Team, 2015) tirando partido da informação GPS (*global positioning system*) das amostras da biblioteca.

### 3. RESULTADOS

#### 3.1 Biblioteca de referência

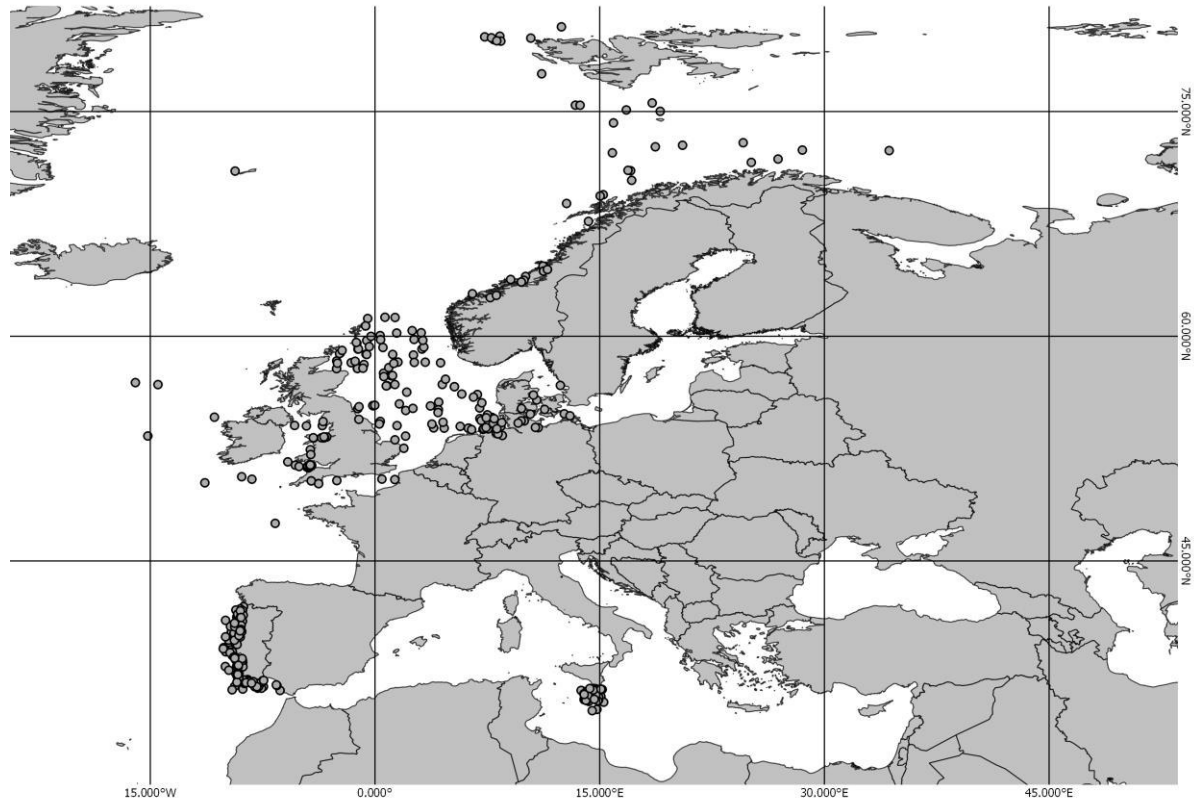
Os números da biblioteca compilada são os seguintes: 4118 *DNA barcodes* distribuídos por 358 espécies de 34 ordens. Os *DNA barcodes* estão repartidos em 18 projetos na base de dados BOLD e perfazendo um total de 13 publicações (Tabela 2). Quatro das publicações aqui compiladas só apresentavam sequências na base de dados GenBank e não tinham qualquer projeto associado à base de dados BOLD (Moura *et al.*, 2008; Straube *et al.*, 2010; Serra-Pereira *et al.*, 2011; Ardura *et al.*, 2013).

**Tabela 2 – Publicações com sequências públicas utilizadas na DS-EUROFISH, bem como o seu código da base de dados BOLD, a região e o número de sequências**

Código do projeto	Publicação	Região	Sequências
BNSFI			
FCFUK	Knebelsberger <i>et al.</i> (2014)	Mar do Norte	855
FCFBI			
FCFP			
FCFPI	Costa <i>et al.</i> (2012)	Costa portuguesa	624
FCFPS			
FCFPW			
MLFP			
FCFMT	Landi <i>et al.</i> (2014)	Mar Mediterrâneo	573
CSFOM			
NEAS	Lynghammar <i>et al.</i> (2014)	Atlântico Nordeste	80
DNATR	Keskin and Atar (2013)	Mar Mediterrâneo (Turquia)	1663
BGNBS	Knebelsberger and Thiel (2014)	Mar Báltico	73
MLFPZ	Martins <i>et al.</i> (2012)	Costa portuguesa	5
EFBC	Luchetti <i>et al.</i> (2011)	Atlântico Nordeste	8
DSNSF	Ward <i>et al.</i> (2008)	Atlântico Norte	40
MOURA	Moura <i>et al.</i> (2008)	Atlântico Nordeste	15
STRAUBE	Straube <i>et al.</i> (2010)	Atlântico Norte	59
SERRA.P	Serra-Pereira <i>et al.</i> (2011)	Atlântico Nordeste	47
ARDURA	Ardura <i>et al.</i> (2013)	Mar Mediterrâneo	33
MLFPI	Neste estudo	Costa portuguesa	27
FCFBS	Neste estudo	Mar Báltico	16
<b>TOTAL</b>			<b>4118</b>

A distribuição geográfica das amostras (Figura 2) revela que a maioria dos espécimes foram obtidos ao longo das costas dos países europeus, nas principais massas de água como o Mar do Norte,

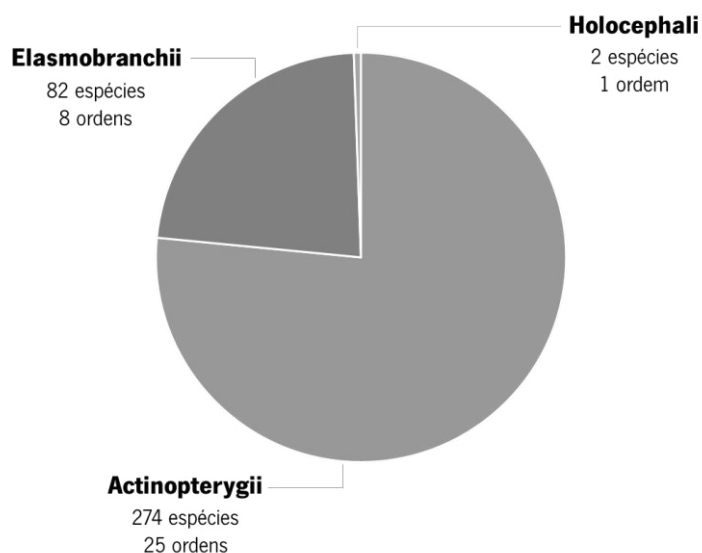
Oceano Atlântico, Mar Mediterrâneo e Mar Báltico. Algumas amostras não continham informação GPS, não constando no mapa de distribuição.



**Figura 2 – Distribuição geográfica das amostras da biblioteca de referência que dispunham de coordenadas GPS (WGS84)**

A DS-EUROFISH representa 3 classes de peixes, sendo que mais de  $\frac{3}{4}$  das espécies pertencem à classe Actinopterygii (peixes ósseos), seguido da classe Elasmobranchii (peixes cartilagineos) e por último a classe Holocephali, um grupo particular de peixes cartilagosos, com apenas duas espécies (Figura 3). A distribuição das amostras seguiu o mesmo padrão e a os peixes ósseos são representados por mais de 3 mil sequências, sendo as restantes maioritariamente da classe Elasmobranchii.





**Figura 3 – Representação gráfica da distribuição taxonómica dos espécimes da biblioteca de referência**

### 3.2 Alocação de espécimes em BINs

O relatório mostrou que as sequências das 345 espécies (13 não estavam presentes na base de dados BOLD) se agruparam em 366 BINs, dos quais 213 foram concordantes (1 BIN = 1 espécie), 141 discordantes (1 BIN = mais que 1 espécie) e 12 *singletons* (1 BIN = 1 sequência). Inspeção posterior da composição de cada BIN revelou potenciais artefactos como, por exemplo, sinónimos, erros de identificação e em alguns casos a incapacidade do algoritmo para resolver os limites das espécies. A tabela 3 mostra a comparação entre o resultado do relatório de discordância de BINs e o resultado da classificação de fiabilidade da biblioteca já anotada. Assumindo que na melhor hipótese possível um BIN corresponde a uma espécie, podemos ver o decréscimo de espécies com identificação discordante e o aumento das espécies com *DNA barcodes* de melhor fiabilidade taxonómica.

**Tabela 3 - Comparação da concordância taxonómica entre BINs sem revisão e após revisão**

Método	Concordante	Discordante	<i>Singleton</i>
BINs sem revisão	58%	39%	3%
Após revisão	82%	14%	4%

### 3.3 Classificação de fiabilidade taxonómica aos *DNA barcodes*

A classificação das espécies pela sua fiabilidade taxonómica mostrou que 82% das espécies pode ser identificada com confiança a partir de *DNA barcodes*. Este número engloba todas as espécies que ficaram alocadas num BIN concordante e que do ponto de vista do utilizador final oferece uma identificação sem ambiguidade. Essa percentagem corresponde a 281 espécies, sendo que 26 delas foram alocadas em BINs com sequências de um só estudo e 15 delas foram alocadas em dois ou mais BINs mas os mesmos são concordantes. Estas 15 últimas espécies apresentam então uma separação dos espécimes em mais que um BIN. Essa diferenciação esteve na maior parte dos casos associada à sua localização geográfica. Das espécies analisadas, 13 delas apresentaram baixa quantidade de sequências (menos de 3 amostras). Encontram-se ainda 51 espécies com ambiguidade taxonómica na biblioteca de referência. Todas essas espécies foram alocadas em BINs discordantes, que continham amostras de mais do que uma espécie. Em todos esses casos o BIN era partilhado por amostras de mais de que uma espécie mas todas do mesmo género. A tabela 4 discrimina a frequência de cada categoria bem como número de espécies que representa.

**Tabela 4 - Atribuição de categorias às sequências das 345 espécies de peixes marinhos europeus**

<b>Categoria</b>	<b>Frequência</b>	<b>N (Espécies)</b>
A	69%	239
B	8%	26
C	5%	16
D	4%	13
E	14%	51
	<b>100%</b>	<b>345</b>

### 3.4 Análise das distâncias genéticas

A distância média intraespecífica foi de 0,68%, com a distância mínima de 0% a distância máxima de 19,10%, registada na espécie *Sporpaena notata*. A distância congénica mais baixa foi de 0,39%, registada no género *Alosa* (entre as espécies *Alosa alosa* e *Alosa fallax*) e a máxima de 23,80% entre as espécies *M. boscanion* e *M. ocellatus* do género *Microchirus*. A distância máxima confamiliar é de 31,28%, que corresponde à família Macrouridae. A família Triglidae foi a que registou uma divergência confamiliar menor, apenas de 2,97% (Tabela 5).

**Tabela 5 – Divergências médias, mínimas e máximas intraespecífica, congênica e confamiliar de peixes marinhos europeus, calculadas segundo o modelo K2P.**

<b>Tipo de divergência</b>	<b>Nº de comparações</b>	<b>Distância Mínima</b>	<b>Valor médio</b>	<b>Distância Máxima</b>
Intraespecífica	42634	0%	0,68%	19,10%
Congênica	28615	0,39%	8,88%	23,80%
Confamiliar	174810	2,97%	16,51%	31,28%

### 3.5 Casos de distância intraespecífica comparativamente elevada

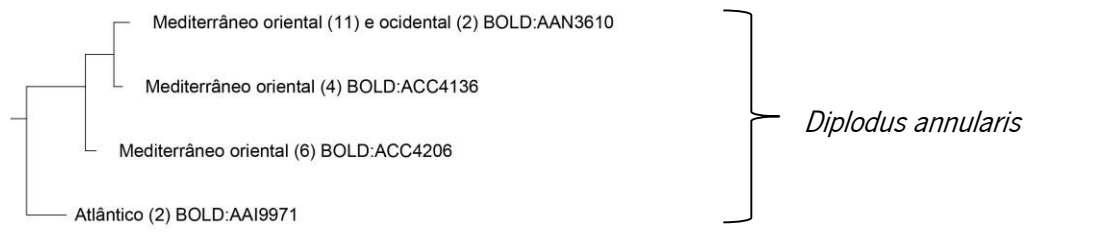
A análise das espécies que foram anteriormente classificadas como sendo da categoria C (16 espécies) mostra que a distância entre diferentes BINs é sempre superior a 2%. Nestas espécies figuram algumas comercialmente importantes como o robalo, o peixe-agulha, o sargo, o linguado e duas espécies de solha. Na maior parte dos casos as espécies foram divididas em dois BINs, separando as sequências de acordo com a sua origem geográfica. A espécie *Etmopterus annularis* não estava presente na base de dados BOLD e a mesma não consta na tabela 6.

**Tabela 6 - Lista de espécies com distâncias intraespecíficas elevadas (excepto *Etmopterus annularis*) e que foram alocadas em mais que um BIN.**

<b>Espécies</b>	<b>Nome comum</b>	<b>Distância entre BINs (K2P)</b>	<b>Região (nº de BINs)</b>
<i>Belone belone</i>	Peixe-agulha	1,3 – 4,3%	Atlântico (2) Mediterrâneo oriental (2)
<i>Chaunax pictus</i>	Sapo-mole	8,0%	Atlântico (2)
<i>Dicentrarchus labrax</i>	Robalo	6,0%	Atlântico (1) Mediterrâneo oriental (1)
<i>Diplodus annularis</i>	Sargo-alcorraz	1,9 – 5%	Atlântico (1), Mediterrâneo ocidental e Mediterrâneo oriental (1), Mediterrâneo oriental (1)
<i>Lepidorhombus whiffiagonis</i>	Solha da fundura	2,6%	Atlântico (1) Mar do Norte (1)
<i>Microchirus ocellatus</i>	Solha linguado	3,4%	Mediterrâneo ocidental (1) Mediterrâneo oriental (1)
<i>Molva molva</i>	Maruca	9,8%	Atlântico (1) Mar do Norte (1)
<i>Pegusa lascaris</i>	Linguado-da-areia	14,7%	Atlântico (2)
<i>Sarda sarda</i>	Sarrajão	2,0%	Atlântico (1) Mediterrâneo oriental (1)
<i>Scorpaena notata</i>	Rascasso-escorpião	18,5%	Atlântico e Mediterrâneo ocidental (2)
<i>Scorpaena scrofa</i>	Rascasso-vermelho	1,1 – 6,3%	Mediterrâneo ocidental (2) Mediterrâneo oriental (2)
<i>Spicara maena</i>	Trombeiro-choupa	7,4%	Atlântico (1) Mediterrâneo ocidental (1)
<i>Spondyliosoma cantharus</i>	Choupa	3,6%	Atlântico (1) Mediterrâneo oriental (1)
<i>Zeus faber</i>	Peixe Galo	6,8%	Atlântico e Mar do Norte (1) Mediterrâneo oriental (1)
<b>14 Espécies</b>			<b>34 BINS</b>

Estas espécies foram alocadas em 38 BINs (incluindo *Centroscyllium fabricii* e *Etmopterus annularis*). Das 14 espécies analisadas, 8 foram alocadas em BINs de regiões geográficas diferentes com maior prevalência na diferença entre o Atlântico e o Mediterrâneo oriental (4 espécies), seguida da diferença entre o Atlântico e o Mar do Norte (2 espécies) e entre o Mediterrâneo ocidental e oriental (2 espécies). Três das espécies foram divididas em 2 BINs da mesma região, todas elas do Atlântico. As restantes espécies foram divididas em um ou mais BINs geograficamente heterogêneos.

A espécie *Belone belone* é dividida em quatro BINs. Na árvore filogenética (Figura 4) é possível verificar a separação mais evidente das amostras de Turquia do restante das amostras e de seguida a diferenciação das amostras do Mar do Norte e da costa portuguesa. As amostras da Turquia foram de facto divididas em dois BINs com distância entre os mesmos de 1,28% (K2P). As sequências do Mar do Norte e uma da costa sul de Portugal agruparam num BIN e a outra amostra da costa portuguesa ficou alocada num BIN com apenas a sua sequência. A distância entre estes dois BINs é de 2,25% e a distância máxima dentro desta espécie é de 4,3%. O robalo (*Dicentrarchus labrax*) apresenta dois grupos bem definidos na árvore filogenética (Figura 4). Esses dois grupos correspondem aos dois BINs atribuídos à espécie e a distância entre os dois é de 6% (K2P) em média e a distância dentro destes não ultrapassa de 0,5% (K2P). Cada BIN está associado a uma região geográfica distinta, ocorrendo a separação das amostras da Turquia e as do Reino Unido. As amostras de sargo-acorraz (*Diplodus annularis*) foram alocadas em 4 BINs (Figura 4). Neste caso a separação não seguiu uma lógica geográfica. A distância máxima entre sequências desta espécie é de 6,6% (K2P) entre o BIN com amostras de Sicília e Turquia e o BIN que contém amostras da costa sul de Portugal. Os outros dois BINs contêm apenas sequências da Turquia e a distância entre os mesmos é de 1,93% (K2P). A espécie *Lepidorhombus whiffiagonis* ou solha-da-fundura foi alocada em 2 BINs. Registou-se de novo uma separação de amostras por região geográfica (Figura 4). Um dos BINs contém amostra do Mar do Norte enquanto o outro BIN contém uma amostra do sul de Portugal e uma do Mar Mediterrâneo (Malta). A distância genética média entre estes dois grupos é de 2,57%, já as distâncias dentro dos grupos não ultrapassa os 0,5%, nos dois casos.

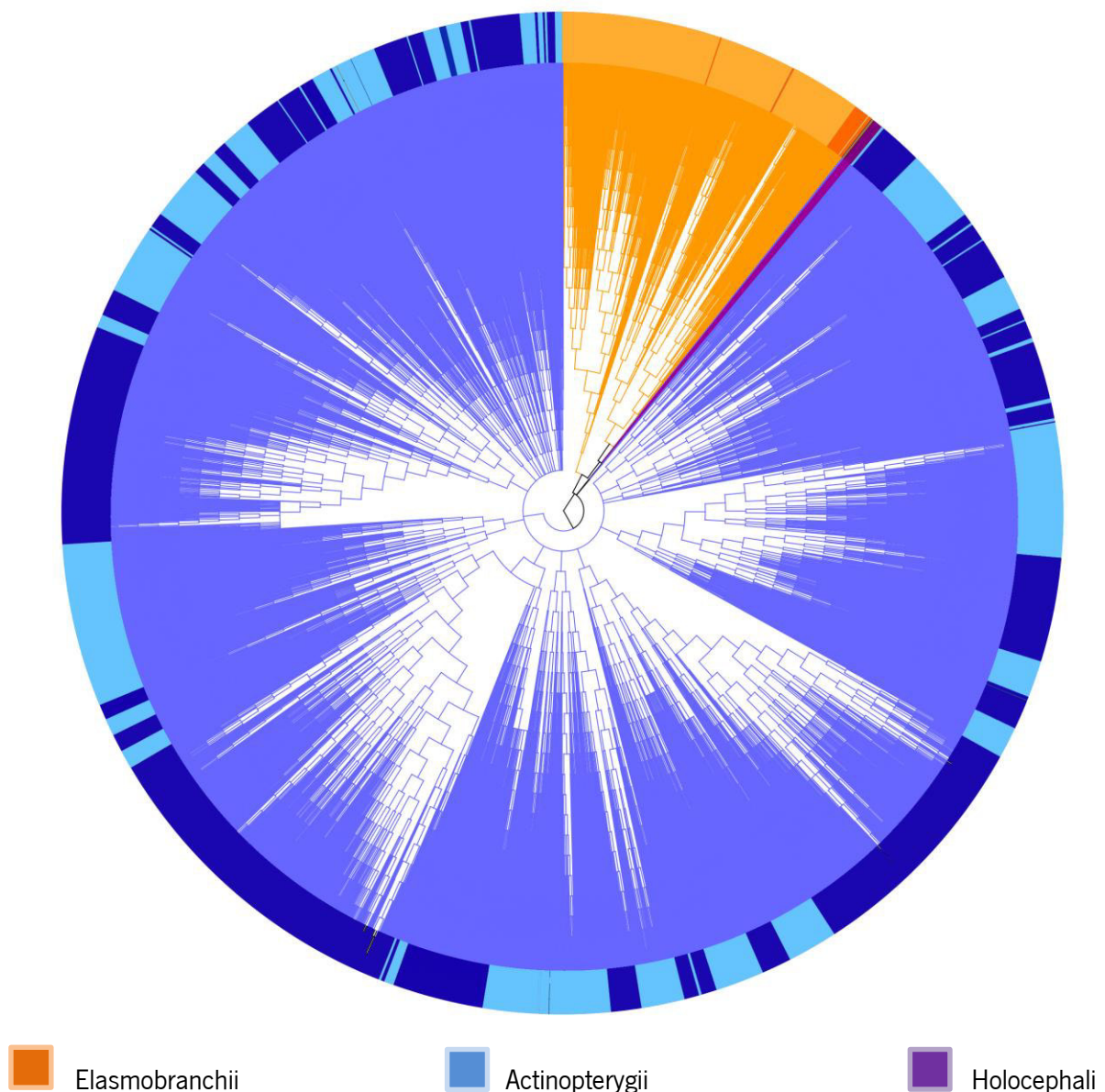


0.01

**Figura 4 – Partes da árvore NJ nucleotídica das espécies *Belone belone*, *Lepidorhombus whiffiagonis*, *Diplodus annularis*, *Dicentrarchus labrax*. A legenda de cada nóculo representa a distribuição geográfica, o número de espécimes**

### 3.6 Reconstrução da filogenia por inferência bayesiana

A reconstrução da filogenia por inferência bayesiana (Figura 5) da biblioteca mostra a clara separação das 3 classes de peixes representadas.

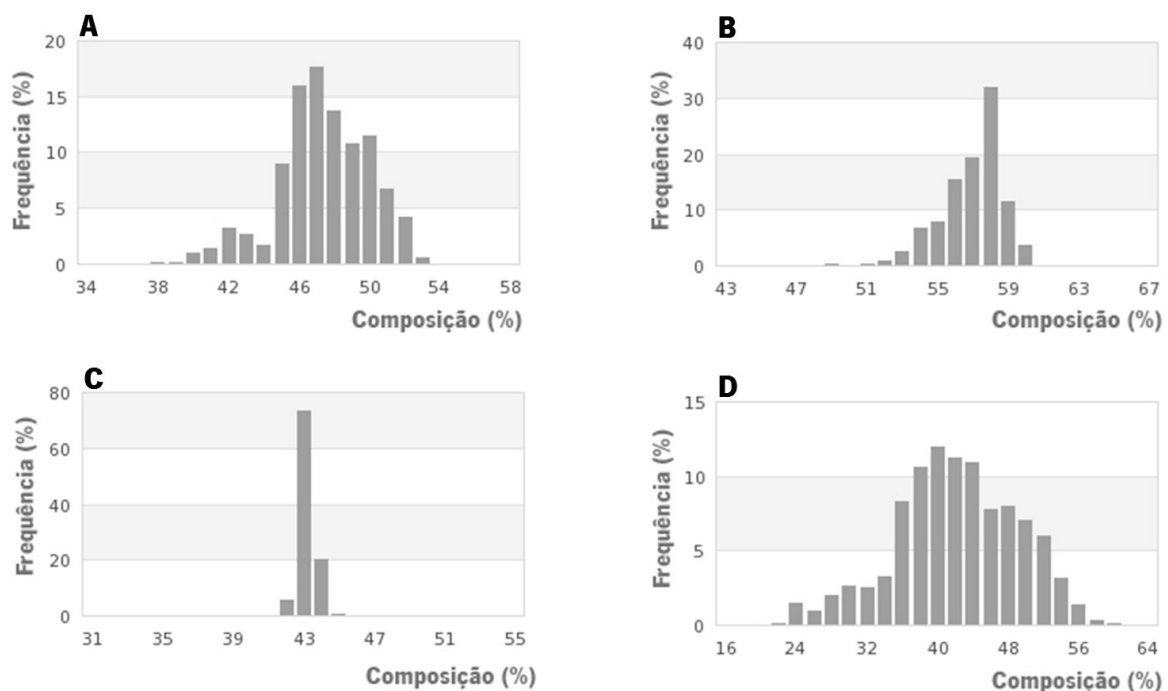


**Figura 5 – Inferência bayesiana das sequências da biblioteca de referência. As cores interiores representam as três classes de organismos. A coloração exterior pretende representar as diferentes ordens dentro das respectivas classes**

A separação das ordens dentro das classes já não é tão evidente, pelo menos na classe Actinopterygii, onde algumas ordens com apenas uma ou duas espécies aparecerem entre ordens muito melhor representadas, tanto a nível de espécies como de sequências. Na classe Elasmobranchii a separação das ordens é perfeita, no entanto o número de ordens nesta classe (8) é inferior comparado com a classe Actinopterygii (25).

### 3.7 Análise da composição e variabilidade nucleotídica

Na tabela 7 é possível verificar que as sequências apresentam diferentes padrões no que toca ao conteúdo em GC (Guanina-Citosina) presente nas amostras. A média geral é de 46,86%, no entanto as diferentes posições do codão apresentam valores diferentes. A primeira posição regista uma média de 56,51%, a segunda posição 42,75% e a terceira posição 41,11%. É possível também verificar pela figura 6 que a variabilidade é bastante diferente se tivermos em conta a posição no codão. A segunda posição apresenta muito baixa variabilidade entre as amostras, com valores de conteúdo em GC mínimos de 39,9% e máximos de 45,10%. Comparativamente, a terceira posição do codão apresenta a maior variabilidade com os valores de conteúdo em GC a oscilar entre 17,55% e 59,25%. Por sua vez a primeira posição do codão apresenta um valor intermédio de variabilidade com valores máximos de conteúdo em GC de 60,74% e mínimos de 46,49%



**Figura 6 – Composição em GC de todas as posições dos codões das sequências da biblioteca de referência (A) e apenas da primeira (B), segunda (C) e terceira (D) posições dos codões das sequências nucleotídicas**

**Tabela 7 - Valores de percentagem de conteúdo em GC na totalidade da sequência e nas diferentes posições dos codões para todos os espécimes presentes na biblioteca de referência.**

	<b>Mínimo</b>	<b>Valor Médio</b>	<b>Máximo</b>	<b>Erro Padrão</b>
<b>GC %</b>	36.48	46.86	53.07	0.04
<b>GC % Posição 1</b>	46.49	56.51	60.74	0.03
<b>GC % Posição 2</b>	39.9	42.75	45.10	0.01
<b>GC % Posição 3</b>	17.55	41.11	59.25	0.11

A mesma análise realizada separadamente para as duas classes mais abundantes revelou diferenças nos conteúdos em GC (Tabela 8 e 9). A média geral para a classe Actinopterygii foi de 47,16% e para a classe Elasmobranchii de 43,36%. A variabilidade foi semelhante para os dois grupos mas a classe Elasmobranchii apresenta uma distribuição com duas modas, enquanto a outra classe apresenta uma distribuição normal (Figura 7).

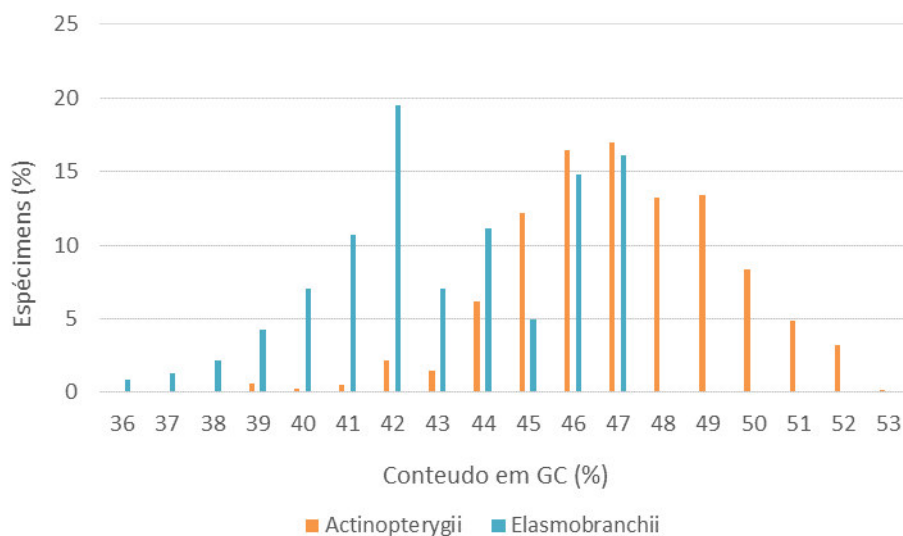
**Tabela 8 - Valores de percentagem de conteúdo em GC na totalidade da sequência e nas diferentes posições dos codões para todos os espécimes da classe Actinopterygii**

	<b>Mínimo</b>	<b>Valor Médio</b>	<b>Máximo</b>	<b>Erro Padrão</b>
<b>GC %</b>	39.14	47.19	53.07	0.04
<b>GC % Posição 1</b>	46.49	56.81	60.74	0.02
<b>GC % Posição 2</b>	39.9	42.74	45.1	0.01
<b>GC % Posição 3</b>	21.63	41.82	59.25	0.11

**Tabela 9 - Valores de percentagem de conteúdo em GC na totalidade da sequência e nas diferentes posições dos codões para todos os espécimes da classe Elasmobranchii**

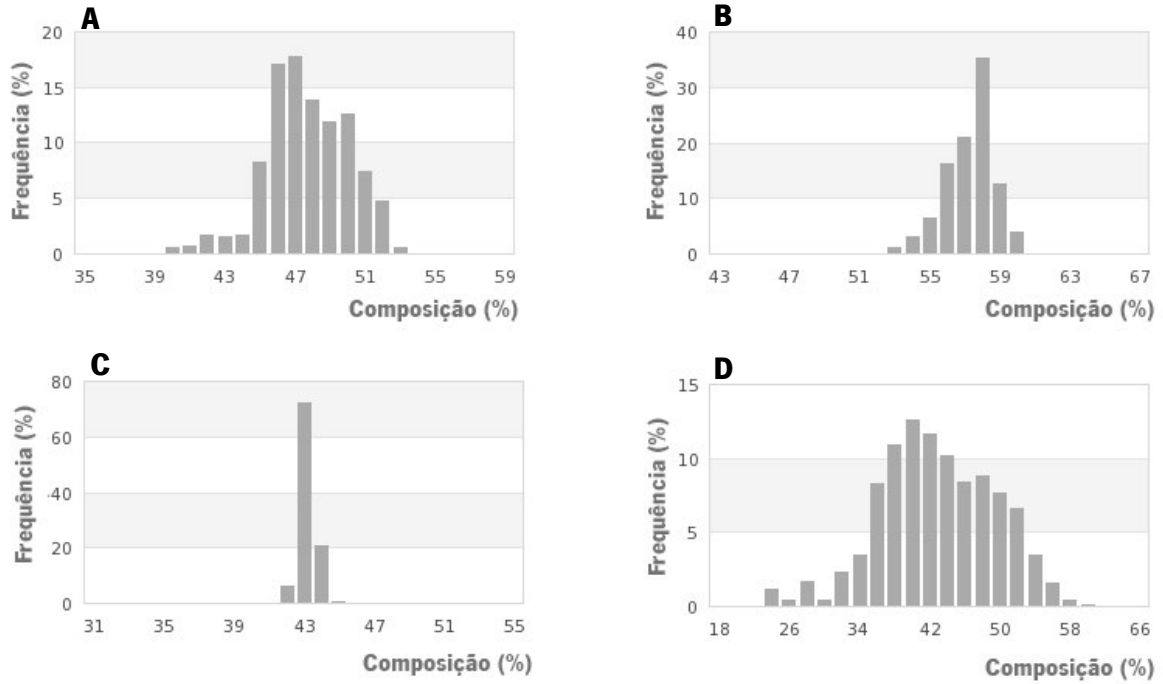
	<b>Mínimo</b>	<b>Valor Médio</b>	<b>Máximo</b>	<b>Erro Padrão</b>
<b>GC %</b>	37.48	43.59	47.49	0.14
<b>GC % Posição 1</b>	47.8	53.53	56.51	0.07
<b>GC % Posição 2</b>	41.56	42.84	44.58	0.02
<b>GC % Posição 3</b>	17.55	34.05	46.08	0.38



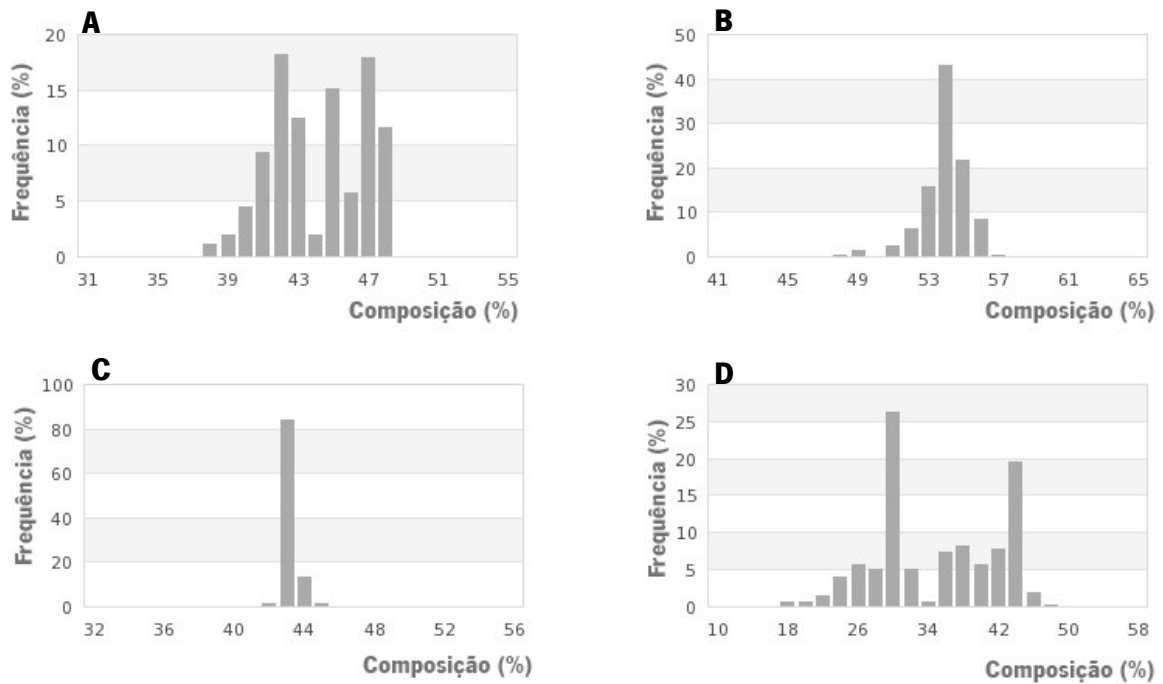


**Figura 7 – Representação gráfica dos valores de porcentagem de conteúdo em GC nas duas classes de peixes mais abundantes da biblioteca de referência. Os espécimes estão representados em porcentagem da sua representação na respetiva classe.**

As diferenças entre as duas classes expressou-se mais na primeira posição e terceira posição do codão. Na primeira posição a diferença entre as médias foi de 3,28%, sendo que o conteúdo em GC a variação dos valores foi maior para a classe Actinopterygii (valores entre 46,49% e 60,74%). Na terceira posição registou-se a maior diferença entre as duas classes, tanto na média dos valores como da dispersão dos mesmos (Figura 8 e 9). A classe Actinopterygii apresentou valores médios de conteúdo em GC de 41,82% e uma distribuição normal enquanto a classe Elasmobranchii apresentou valores médios de 34,05% e uma distribuição com duas modas (Figura 9 “D”). Já na segunda posição a média foi semelhante, 42,74% para a classe Actinopterygii (com os valores a oscilar entre 39,9% e 45,1%) e 42,84% para a classe Elasmobranchii (com os valores a oscilar entre 41,56% e 44,58%).



**Figura 8 – Composição em GC de todas as posições dos códons das seqüências da classe Actinopterygii (A) e apenas da primeira (B), segunda (C) e terceira (D) posições dos códons das seqüências nucleotídicas**



**Figura 9 – Composição em GC de todas as posições dos códons das seqüências da classe Elasmobranchii (A) e apenas da primeira (B), segunda (C) e terceira (D) posições dos códons das seqüências nucleotídicas**

## 4. DISCUSSÃO

### 4.1 Fiabilidade da biblioteca de referência para peixes marinhos europeus

A biblioteca compilada neste estudo confirma a utilidade do *DNA barcoding*, mais especificamente da região mitocondrial COI-5P, na identificação de espécies de peixes marinhos. A identificação morfológica correspondeu à identificação por unidades taxonómicas operacionais moleculares (MOTUs) em mais de  $\frac{3}{4}$  das espécies amostradas. Estes resultados vão de encontro a estudos anteriormente realizados para o mesmo tipo de organismos, na Europa (Landi *et al.*, 2014; Knebelsberger *et al.*, 2014; Costa *et al.*, 2012) ou até mesmo noutros continentes (Ward *et al.*, 2005). Apesar da quantidade de espécies presentes nesta biblioteca com elevada fiabilidade taxonómica ser satisfatória quando comparada com estudos semelhantes, a percentagem de espécies aqui representadas corresponde a menos de 30% das espécies de peixes marinhos reportadas para a Europa. Note-se que, devido ao facto dos espécimes terem sido recolhidos ao longo das costas dos países, a maioria das espécies de águas profundas e dos Arquipélagos atlânticos estão em falta neste estudo e consequentemente nas bases de dados públicas.

As análises automatizadas disponibilizadas na base de dados BOLD constituem um primeiro passo importante para curar estas bibliotecas. Facultam um primeiro diagnóstico sobre o conjunto de dados, especialmente o relatório de discordância de BINs, que mostra a potencialidade da identificação de espécimes através de *DNA barcodes* ao agrupar as sequências em unidades taxonómicas operacionais moleculares (Ratnasingham e Hebert, 2013). No entanto, o relatório gerado neste estudo revelou um número elevado de BINs discordantes, com quase 40% dos BINs a conterem sequências de mais do que uma espécie. Este número notoriamente elevado, e que a confirmar-se comprometeria a fiabilidade da biblioteca de referência na sua globalidade, desceu consideravelmente após revisão individual de cada BIN. Nesse processo evidenciaram-se vários artifícios que podem provocar uma sobrestimação das discordâncias dos BINs, que não são passíveis de deteção pelo algoritmo que realiza o agrupamento das sequências, como por exemplo erros de sintaxe nos nomes das espécies, ou falhas operacionais como a contaminação de amostras. Estas insuficiências dos métodos automatizados de atribuição de sequências a espécies, confirmam a absoluta necessidade de intervenção humana na revisão e anotação das bibliotecas de referência (Costa *et al.*, 2012). O método BIN também poderá em certos casos falhar não por razões externas, mas por insuficiência da capacidade de resolução do próprio algoritmo. Um caso particular desta incapacidade de resolução foi observado com o género *Trachurus*, onde um BIN

continha três espécies desse mesmo género. No entanto, uma análise mais aprofundada da árvore filogenética mostrou que as espécies se agrupam em ramos monofiléticos bem delimitados, não existindo em consequência ambiguidade na sua diferenciação através dos respetivos *DNA barcodes*. Nestes casos de distância congénica comparativamente baixa, o algoritmo não foi capaz de discriminar as três espécies em três BINs diferentes (Ratnasingham e Hebert, 2013). Outro artefacto encontrado na revisão dos BINs foi de espécies que foram agrupadas em dois ou mais BINs internamente concordantes, mas com distâncias genéticas entre os mesmos de 1,1% até 18,5% (K2P). Isto ocorre porque o critério de concordância do relatório apenas requer que um BIN só contenha sequências de uma espécie (1 BIN = 1 espécie). No entanto, em 15 ocasiões uma espécie continha sequências em mais que um BIN (1 BIN = 1 espécie mas 1 espécie = 2 ou mais BINS). Esta inspeção pormenorizada permitiu revelar estes casos de diversidade anteriormente oculta no resultado do relatório de discordância de BINs. Foram encontrados também erros de identificação, que são de fácil identificação mas podem facilmente induzir em erro o usuário final da base de dados. A presença de sinónimos, más classificações e/ou identificações também afetam o resultado do relatório. Nestes casos, os BINs são classificados discordantes no relatório de discordância, devido à presença de sequências de mais de uma espécie, mas uma análise detalhada da taxonomia mostrou que em muitos desses casos a simples correção de sintaxe ou deteção de sinónimos e nomes não aceites permite a retificação da classificação dada ao BIN. Estes erros ocorrem principalmente na fase pré *barcode* e para evitar estas situações é essencial a criação de ferramentas de correção taxonómica dentro das próprias bases de dados, de forma a auxiliar o aperfeiçoamento das mesmas. Apesar de mais de  $\frac{3}{4}$  das espécies desta biblioteca apresentarem fiabilidade taxonómica elevada subsistiram ainda 51 espécies com ambiguidade na sua identificação. Estas incertezas taxonómicas podem ser eficazmente endereçadas com a contínua revisão e anotação, incluindo a atribuição de graus de fiabilidade taxonómica (Costa *et al.*, 2012).

## **4.2 Divergência intraespecífica comparativamente elevada**

Neste estudo ficou patente a importância deste tipo de compilação pois, apesar de não ter gerado qualquer contributo direto em novos *DNA barcodes*, só através do cruzamento de dados de origem múltipla foi possível detetar diversos casos pertinentes de ambiguidade nas identificações de espécies de peixes, ou no seu estatuto taxonómico. Neste âmbito merecem particular apreço as 15 espécies que apresentam distâncias intraespecíficas comparativamente elevadas, tendo por referência o padrão típico de variabilidade determinado a partir do enorme volume de sequências de COI-5P disponíveis para

animais, incluindo numerosos peixes marinhos (Ward e Holmes, 2007; Costa e Carvalho, 2010). De facto, ficou evidenciado que os diferentes grupos de investigação de diferentes países tipicamente produzem bibliotecas de referência para uma determinada região geográfica e interpretam isoladamente os dados gerados localmente (Moura *et al.*, 2008; Straube *et al.*, 2010; Luchetti *et al.*, 2011; Serra-Pereira *et al.*, 2011; Martins *et al.*, 2012; Ardura *et al.*, 2013; Keskin and Atar, 2013; Lynghammar *et al.*, 2014). Ao compilar e anotar dados gerados a nível regional foi possível identificar espécies com distâncias intraespecíficas atípicas e que previamente se julgavam dentro dos padrões de variabilidade habituais. Um destes exemplos é a espécie *Lepidorhombus whiffiagonis* que no estudo referente aos peixes marinhos do Mar do Norte (Kneibelsberger *et al.*, 2014) e da costa portuguesa (Costa *et al.*, 2012) apresentam distâncias intraespecíficas menores que 0,5% entre os seus respetivos espécimes. No entanto quando os dados foram analisados em conjunto foi possível verificar distâncias até 3,0% entre os espécimes das duas regiões.

Podemos identificar três tipos de padrões de separação em diferentes BINs das linhagens intraespecíficas: BINs de regiões geográficas diferentes; BINs da mesma região; e BINs geograficamente heterogéneos. Das 14 espécies analisadas, 8 foram divididas em dois BINs e em metade desses casos a divisão geográfica ocorreu entre o oceano Atlântico e o Mediterrâneo oriental. A existência de uma barreira filogeográfica entre o Mar Mediterrâneo e o oceano Atlântico, foi revista num estudo de Patarnello *et al.* (2007) em que, a partir de vários marcadores moleculares, se avalia a existência de divergências genéticas pronunciadas entre populações de espécies amostradas nos dois lados do estreito de Gibraltar. O trabalho aqui realizado reforça o papel desta barreira filogeográfica na estruturação genética de um número considerável de espécies de peixes.

Os padrões de diferenciação genética entre regiões são em alguns casos muito distintos. As sequências da espécie *Zeus faber* foram distribuídas por dois BINs, um com amostras do oceano Atlântico e do Mar do Norte e outro com as amostras do Mediterrâneo oriental. Numa análise pormenorizada foi possível ainda verificar que as amostras do Mediterrâneo oriental estão presentes num BIN composto por amostras dos oceanos Índico e Pacífico, contendo amostras desde a África do Sul, até à Austrália, China e Japão. Isto sugere que estes espécimes de *Zeus faber* capturados no Mediterrâneo oriental (Turquia) são resultado da migração de populações provenientes do oceano Índico através do canal do Suez, como já foi sugerido em estudos anteriores (Castro, 2013; Landi *et al.*, 2014). No seu conjunto, as espécies com diversidade intraespecífica comparativamente elevada, cujos espécimes foram divididos em dois ou mais BINs, constituem casos de possíveis espécies crípticas ou, no mínimo, unidades evolutivas significantes (i.e. ESU - Evolutionarily Significant Units) que exigem um tratamento

diferenciado num quadro de gestão de mananciais pesqueiros ou de conservação da biodiversidade. Devem por esse motivo ser alvo de estudos pormenorizados para que a sua taxonomia, ecologia e história evolutiva seja compreendida.

### 4.3 Padrões de variabilidade genética

A inferência bayesiana permitiu demonstrar a capacidade de discriminação do *DNA barcoding*, não só ao nível da espécie mas também ao nível de *taxa* mais elevados, como ordens e classes. A estimação da filogenia com base em dados moleculares mostrou ser capaz de discriminar os espécimes por espécie e compreender a sua estrutura e complexidade. Esta estrutura intraespecífica é mais evidente quanto maior o número e a dispersão geográfica das amostras e esta biblioteca de referência é a situação ideal para encontrar essas evidências. A análise das distâncias intraespecífica, congénérica e confamiliar vão de encontro ao que se tem verificado em estudos anteriores (Landi *et al.*, 2014; Kneibelsberger *et al.*, 2014; Costa *et al.*, 2012; Steinke *et al.*, 2009; Ward *et al.*, 2005). Algumas espécies desviam-se da maioria, seja pela enorme divergência intraespecífica ou baixa divergência congénérica. Os casos das espécies com distância intraespecíficas elevadas, onde se verifica a separação notória de duas populações, evidência a presença de espécies crípticas (Hubert *et al.*, 2012). Alguns géneros com baixa divergência entre as suas espécies, como é exemplo o género *Alosa* e o género *Trachurus*, demonstram duas características genéticas e populacionais que podem fazer com que o *DNA barcoding* falhe na discriminação correta das espécies. O género *Alosa* é representado nesta biblioteca por duas espécies (*Alosa alosa* e *Alosa fallax*) que hibridam frequentemente (Alexandrino 1996; Marques, 2011). Essa característica, aliada à transmissão unilateral do genoma mitocondrial faz com que seja possível a identificação morfológica de um espécime de *Alosa alosa* com genoma mitocondrial de *Alosa fallax*. No caso da biblioteca aqui compilada foi possível verificar a presença de espécimes das duas espécies no mesmo BIN, confirmando a possibilidade anterior. As espécies do género *Trachurus* neste estudo representadas (*Trachurus trachurus*, *Trachurus mediterraneus* e *Trachurus picturatus*) apresentaram entre si distâncias interespecíficas inferiores a 1,5%. No entanto as três espécies apresentam clados monofiléticos na filogenia estimada por inferência bayesiana e a identificação de espécimes a partir de *DNA barcodes* não ficou comprometida. Este caso evidencia a presença de um género com taxas de substituição nucleotídica comparativamente baixa, uma vez que a hipótese de ser um género com história evolutiva recente foi rejeitada anteriormente (Cárdenas *et al.*, 2005). Estas evidências são de inegável

importância uma vez que estas espécies de carapau (*Trachurus*) apresentam um enorme interesse económico.

Uma das vantagens da compilação de dados com grande amplitude geográfica e taxonómica é a possibilidade de descobrir padrões que não ocorrem quando os mesmos são analisados apenas a nível regional. A amplitude taxonómica permitiu observar que o conteúdo em GC revelou padrões diferentes nas duas classes mais abundantes da biblioteca. Os resultados obtidos nesta dissertação acompanham os que foram obtidos no estudo de *DNA barcoding* dos peixes da Austrália (Ward *et al.*, 2005), onde estão representadas 207 espécies de um trabalho de amplitude continental. Foi consistente a observação de um erro padrão maior na terceira posição dos codões, independentemente da classe. Isto reflete o facto das mutações sinónimas ocorrerem mais nessa posição, seguida da primeira posição do codão. Na segunda posição não ocorrem mutações sinónimas, e o erro padrão e média do conteúdo em GC foi idêntico nas duas classes, mostrando a conservação nucleotídica ao longo dos vários grupos taxonómicos.

#### **4.4 Importância e utilidade da biblioteca de referência**

A biblioteca compilada neste estudo constitui um recurso valioso para a comunidade científica e utilizadores finais, nas mais diversas aplicações. O propósito mais relevante será no auxílio ao conhecimento da diversidade da ictiofauna marinha europeia. A informação aqui compilada permite a deteção de diversidade que de outra forma estaria oculta. Esta compilação é uma oportunidade para padrões que de outra forma não seriam perceptíveis, como é o exemplo das variações do conteúdo em GC e a diferenciação geográfica de 15 espécies. Os dados da biblioteca evidenciam que essas espécies apresentam uma diversidade acima dos valores habituais para este grupo taxonómico extensamente estudado (Ward e Holmes, 2007). Essa diversidade nem sempre estará associada ao aumento do número de espécies conhecidas mas pode também identificar populações diferenciadas de uma determinada espécie. O conhecimento dessas dinâmicas das espécies torna-se pertinente na gestão e conservação das mesmas. A facilidade em identificar larvas ou mesmo ovos pode facilitar a identificação de zonas e épocas de postura das espécies, que pode ser uma informação importante na gestão conservação das mesmas. Os resultados sugerem que espécies com valores de divergência elevados, que posteriormente se refletem na distribuição das mesmas, podem ter populações com pouca ou nenhuma interação. Uma compilação com esta amplitude geográfica pode também auxiliar a apurar

áreas de distribuição das espécies e mesmo barreiras de dispersão, particularmente quando são encontradas espécies com grande divergência entre populações geograficamente próximas. Estes resultados sugerem que, pelo menos para algumas espécies, possam existir barreiras geográficas significativas ao fluxo genético, previamente insuspeitas. Nos casos de espécies comercialmente importantes a informação sobre a diferenciação populacional torna-se ainda mais importante e deve ter sido em conta quando se estabelecem unidades de gestão de mananciais pesqueiros e respetivas quotas de pescas numa comunidade integrada como a europeia. A ineficiente gestão das diferentes populações de uma destas espécies pode levar ao aumento de pressão ou mesmo perda de uma delas (FAO, 2012). Isso teria implicações na conservação das mesmas pois significaria a possível perda de haplotipos do *pool* genético da espécie. Mesmo quando essa gestão diferenciada das populações de uma espécie é efetuada de forma eficiente, a falta de controlo das capturas pode anular o esforço anterior (Cawthorn *et al.*, 2014). Esta biblioteca pode também ser usada nesse controlo de pescas, na correta identificação dos produtos pescados, podendo também ser importante no controlo de pescas ilegais. São vários os exemplos do uso de métodos moleculares e especialmente do *DNA barcoding* no controlo de pescas ilegais e não reportadas em que os resultados mostram a utilidade dessas ferramentas para o auxílio na identificação e autenticação de produtos de pesca (Nicolè *et al.*, 2012; Torres *et al.*, 2013; Helyar *et al.*, 2014). Para além do controlo do que é pescado, a biblioteca pode auxiliar a controlar o que é vendido provando a autenticidade dos produtos, tanto nos mercados como na restauração, sendo vários os exemplos do uso do *DNA barcoding* para este fim. O estudo sobre a substituição de espécies em mercados de Itália (Filonzi *et al.*, 2010) conclui que mais de 30% da etiquetagem dos peixes é incorreta e que na maioria dos casos ocorria fraude económica e nutricional. Apenas com a informação molecular foi possível identificar algumas das substituições, já que o estudo também se focou em filetes de peixe já processados. Este tópico é emergente e faz com que seja essencial encontrar soluções para resolver as fraudes que ocorrem não só na Europa (Mendes e Silva, 2015; Kappel e Schröder, no prelo) como no Canadá (Hanner *et al.*, 2011) e no Brasil (Carvalho *et al.*, 2015). A publicação de Carvalho e colaboradores (2015) utiliza o *DNA barcoding* como ferramenta legal de avaliação da autenticidade de produtos de pesca. Apenas uma biblioteca devidamente anotada e validada pode servir todos estes usos de uma forma confiável.



## 5. CONSIDERAÇÕES FINAIS

Apenas através da integração destes dados dispersos, obtidos por diferentes grupos de investigação de vários países, e pela revisão e anotação desta biblioteca de escala continental, foi possível desvendar casos pertinentes de ampla divergência intraespecífica e incertezas taxonómicas em mais de 60 espécies de peixes marinhos europeus. São essas as que precisam de uma atenção redobrada em estudos futuros, de forma a conseguir aprimorar estas bibliotecas de referência de grande escala. É também importante o contínuo processo de revisão e anotação destas bibliotecas. Note-se que a biblioteca é um trabalho em progresso, sempre em mudança com a introdução de novas espécies e de novos espécimes. Só com essa continuidade é possível que esta seja usada com confiança pelo utilizador final. Este estudo permitiu também mostrar que, entre as espécies de peixes marinhos europeus aqui examinadas, mais de  $\frac{3}{4}$  destas podem ser identificadas com elevada fiabilidade através dos respectivos *DNA barcodes* existentes na biblioteca de referência. Contudo, o volume de dados gerados pelos estudos já realizados cobre apenas cerca de 30% das espécies reportadas para a Europa, facto que deixa bem patente o enorme esforço de preenchimento da biblioteca de referência que ainda deverá ser envidado..

O desenvolvimento de novas ferramentas bioinformáticas e o aperfeiçoamento das bases de dados públicas podem permitir que esta biblioteca seja considerada um produto que possa ser usado por qualquer pessoa, evocando as bases com que o *DNA barcoding* foi idealizado e concebido por Hebert e colaboradores, em 2003. Estas bibliotecas já mostraram ser capazes de auxiliar no conhecimento da diversidade e ecologia das espécies, principalmente em grupos bem estudados como é o caso dos peixes. A correta identificação da biodiversidade, a deteção de zonas de postura e acompanhamento do recrutamento e reconhecimento de barreiras filogeográficas são as principais valências da metodologia. Pelo facto deste grupo de organismos ter uma importância social e económica elevada as aplicações do *DNA barcoding* são ainda mais extensas e mostram ser pertinentes no controlo e gestão das pescas e a tudo a isso associado, desde deteção de pesca ilegal, autenticação de produtos pescadas pela avaliação da etiquetagem e a gestão dos mananciais pesqueiros pela correta identificação das espécies e respetivas populações. Apenas uma biblioteca devidamente anotada, validada e atualizada pode servir todos estes usos de uma forma confiável e segura.



## REFERÊNCIAS

- Alexandrino, P. J. (1996). *Estudo de Populações de sável (Alosa alosa L.) e savelha (Alosa fallax Lacépède). Análise de diferenciação interespecífica, subestruturação e hibridação.* (Tese de Doutoramento, Universidade do Porto, Porto, Portugal).
- Ardura, A., Planes, S., Garcia-Vazquez, E. (2013). Applications of *DNA barcoding* to fish landings: Authentication and diversity assessment. *ZooKeys*, 365, 49–65. <http://doi.org/10.3897/zookeys.365.6409>
- Armstrong, K. F., Ball, S. L. (2005). *DNA barcodes* for biosecurity: invasive species identification. *Philos T Roy Soc B* 360, 1813–1823.
- Azzurro, E., Moschella, P., Maynou, F. (2011). Tracking Signals of Change in Mediterranean Fish Diversity Based on Local Ecological Knowledge. *PLoS ONE* 6, e24885.
- Bianchi, C. N., Morri, C. (2000). Marine biodiversity of the Mediterranean Sea: situation, problems and prospects for future research. *Mar. Pollut. Bull.* 40, 367–376.
- Cárdenas, L., Hernández, C. E., Poulin, E., Magoulas, A., Kornfield, I., Ojeda, F. P. (2005). Origin, diversification, and historical biogeography of the genus *Trachurus* (Perciformes: Carangidae). *Molecular Phylogenetics and Evolution*, 35, 496–507. <http://doi.org/10.1016/j.ympev.2005.01.011>
- Carneiro, M., Martins, R., Landi, M., Costa, F. O. (2014). Updated checklist of marine fishes (Chordata: Craniata) from Portugal and the proposed extension of the Portuguese continental shelf. *European Journal of Taxonomy*, 73, 1–73. <http://doi.org/http://dx.doi.org/10.5852/ejt.2014.73>
- Carvalho, D. C., Palhares, R. M., Drummond, M. G., Frigo, T. B. (2015). DNA barcoding identification of commercialized seafood in South Brazil: A governmental regulatory forensic program. *Food Control*, 50, 784–788. <http://doi.org/10.1016/j.foodcont.2014.10.025>
- Castro, S. L. C. (2013). *Compilação de uma biblioteca de referência de DNA barcodes para peixes marinhos de Portugal e estudo filogeográfico da espécie Zeus faber.* (Dissertação de Mestrado, Universidade do Minho, Braga, Portugal).
- Cawthorn, D-M., Steinman, H. A., Witthuhn, R. C. (2012) DNA barcoding reveals a high incidence of fish species misrepresentation and substitution on the South African market. *Food Research International*, 46, 30-40. doi:10.1016/j.foodres.2011.11.011
- Coll, M., Piroddi, C., Steenbeek, J., Kaschner, K., Ben RaisLasram, F., Aguzzi, J., Ballesteros, E., Bianchi, C. N., Corbera, J., Dailianis, T., Danovaro, R., Estrada, M., Froglia, C., Galil, B. S., Gasol, J. M., Gertwagen, R., Gil, J., Guilhaumon, F., Kesner-Reyes, K., Kitsos, M-S., Koukouras, A., Lampadariou, N., Laxamana, E., de la Cuadra, C. M. L-F., Lotze, H. K., Martin, D., Mouillot, D., Oro, D., Raicevich, S., Rius-Barile, J., Saiz-Salinas, J. I., San Vicente, C., Somot, S., Templado, J., Turon, X., Vafidis, D., Villanueva, R., Voultsiadou, E. (2010). The biodiversity of the Mediterranean Sea: Estimates, Patterns, and Threats. *PLoS ONE* 5, e11842.
- Costa, F. O., Carvalho, G. R. (2007). The Barcode of Life Initiative : synopsis and prospective societal impacts of DNA barcoding of Fish. *Genomics Society and Policy*, 3, 29–40.

- <http://doi.org/10.1186/1746-5354-3-2-29>
- Costa, F. O., Carvalho, G. R. (2010). New insights into molecular evolution: Prospects from the barcode of life initiative (BOLI). *Theory in Biosciences*, 129, 149–157. <http://doi.org/10.1007/s12064-010-0091-y>
- Costa, F. O., Landi, M., Martins, R., Costa, M. H., Costa, M. E., Carneiro, M., Alves, M. J., Steinke, D., Carvalho, G. R. (2012). A Ranking System for Reference Libraries of DNA barcodes: Application to Marine Fish Species from Portugal. *PLoS ONE*, 7, e35858. <http://doi.org/10.1371/journal.pone.0035858>
- Costello, M. J., Bouchet, P., Emblow, C. S., Legakis, A. (2006). European marine biodiversity inventory and taxonomic resources: State of the art and gaps in knowledge. *Marine Ecology Progress Series*, 316, 257–268. <http://doi.org/10.3354/meps316257>
- European Commission (2014). *Facts and figures on the Common Fisheries Policy –Basic statistical data*. Luxemburgo: Publications Office of the European Union. <http://doi.org/10.2771/35745>
- FAO (2012) *The State of the World's Fisheries and Aquaculture*. Food and Agriculture Organization of the United Nations, Roma, Itália.
- FAO (2002) *The State of the World's Fisheries and Aquaculture*. Food and Agriculture Organization of the United Nations, Roma, Itália.
- Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39, 783-791.
- Filonzi, L., Chiesa, S., Vaghi, M., Marzano, F. N. (2010) Molecular barcoding reveals mislabelling of commercial fish products in Italy. *Food Research International*, 43, 1383-1388. doi:10.1016/j.foodres.2010.04.016
- Froese, R., Pauly, D. (2012). Global Information System on Fishes - FISHBASE (version Dec. 2012). Disponível em: <http://www.fishbase.org>
- Hajibabaei, M., Singer, G. A. C., Hebert, P. D. N., Hickey, D. A. (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, 23, 167–172. <http://doi.org/10.1016/j.tig.2007.02.001>
- Hanner, R. (2005). Proposed standards for BARCODE records in INSDC (BRIs). Disponível em em [http://barcoding.si.edu/pdf/dwg\\_data\\_standards-final.pdf](http://barcoding.si.edu/pdf/dwg_data_standards-final.pdf)
- Hanner, R., Becker, S., Ivanova, N. V., Steinke, D. (2011). FISH-BOL and seafood identification: Geographically dispersed case studies reveal systemic market substitution across Canada. *Mitochondrial DNA*, 22, 106-122
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57, 97–109. Retrieved from <http://biomet.oxfordjournals.org/content/57/1/97.short>
- Hebert, P. D. N., Cywinska, A., Ball, S. L., deWaard, J. R. (2003a). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, 270, 313–321. <http://doi.org/10.1098/rspb.2002.2218>
- Hebert, P. D. N., Ratnasingham, S., Jeremy, R., DeWaard, J. R., de Waard, J. R. (2003b). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270, S96.

<http://doi.org/10.1098/rsbl.2003.0025>

- Helyar, S. J., Lloyd, H. ap D., de Bruyn, M., Leake, J., Bennett, N., & Carvalho, G. R. (2014). Fish Product Mislabelling: Failings of Traceability in the Production Chain and Implications for Illegal, Unreported and Unregulated (IUU) Fishing. *PLoS ONE*, *9*, e98691. <http://doi.org/10.1371/journal.pone.0098691>
- Holmlund, C. M., Hammer, M. (1999) Ecosystem services generated by fish populations. *Ecological Economics*, *29*, 253-268
- Hubert, N., Meyer, C. P., Bruggemann, H. J., Guérin, F., Komeno, R. J. L., Espiau, B., Causse, R., Williams, J. T., Planes, S. (2012). Cryptic diversity in Indo-Pacific coral-reef fishes revealed by DNA-barcoding provides new support to the centre-of-overlap hypothesis. *PLoS One*, *7*, e28987. <http://doi.org/10.1371/journal.pone.0028987>
- Kappel, K., Schröder, U. (no prelo). Substitution of high-priced fish with low-priced species: Adulteration of common sole in German restaurants. *Food Control*, *59*, 478–486. <http://doi.org/10.1016/j.foodcont.2015.06.024>
- Katoh, K., Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, *30*, 772–780. <http://doi.org/10.1093/molbev/mst010>
- Kerr, K. C. R., Lijtmaer, D. A., Barreira, A. S., Hebert, P. D. N., Tubaro, P. L. (2009). Probing Evolutionary Patterns in Neotropical Birds through DNA barcodes. *PLoS ONE*, *4*, e4379. <http://doi.org/10.1371/journal.pone.0004379>
- Keskin, E., Atar, H. H. (2013). DNA barcoding commercially important fish species of Turkey. *Molecular Ecology Resources*, *13*, 788–797. <http://doi.org/10.1111/1755-0998.12120>
- Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* *16*, 111-120.
- Knebelsberger, T., Landi, M., Neumann, H., Kloppmann, M., Sell, A. F., Campbell, P. D., Laakmann, S., Raupach, M. J., Carvalho, G. R., Costa, F. O. (2014). A reliable DNA barcode reference library for the identification of the North European shelf fish fauna. *Molecular Ecology Resources*, *14*, 1060–1071. <http://doi.org/10.1111/1755-0998.12238>
- Knebelsberger, T., Thiel, R. (2014). Identification of gobies (Teleostei: Perciformes: Gobiidae) from the North and Baltic Seas combining morphological analysis and DNA barcoding. *Zoological Journal of the Linnean Society*, *172*, 831–845. <http://doi.org/10.1111/zoj.12189>
- Kochzius, M., Seidel, C., Antoniou, A., Botla, S. K., Campo, D., Cariani, A., Vasquez, E. G., Hauschild, J., Herve, C., Hjörleifsdottir, S., Hreggvidsson, G., Kappel, K., Landi, M., Magoulas, A., Marteinson, V., Nölte, M., Planes, S., Tinti, F., Turan, C., Venugopal, M. N., Weber, H., Blohm, D. (2010). Identifying fishes through DNA barcodes and microarrays. *PLoS ONE*, *5*, 1–15. <http://doi.org/10.1371/journal.pone.0012620>
- Landi, M., Dimech, M., Arculeo, M., Biondo, G., Martins, R., Carneiro, M., Carvalho, G. R., Brutto, S. L., Costa, F. O. (2014). DNA barcoding for species assignment: the case of Mediterranean marine fishes. *PLoS One*, *9*, e106135. <http://doi.org/10.1371/journal.pone.0106135>
- Leonart, J., Taconet, M., Lamboeuf, M. (2006). Integrating information on marine species identification

- for fishery purposes. *Marine Ecology Progress Series*, 316, 231–238. <http://doi.org/10.3354/meps316231>
- Luchetti, E. a, Iglésias, S. P., Sellos, D. Y. (2011). *Chimaera opalescens* n. sp., a new chimaeroid (Chondrichthyes: Holocephali) from the north-eastern Atlantic Ocean. *Journal of Fish Biology*, 79, 399–417. <http://doi.org/10.1111/j.1095-8649.2011.03027.x>
- Lynghammar, A., Christiansen, J. S., Griffiths, A. M., Fevolden, S. E., Hop, H., Bakken, T. (2014). DNA barcoding of the northern Northeast Atlantic skates (Chondrichthyes, Rajiformes), with remarks on the widely distributed starry ray. *Zoologica Scripta*, 43, 485–495. <http://doi.org/10.1111/zsc.12064>
- Marques, T. N. (2011). *Desenvolvimento Larvar da Savelha (Alosa fallax) – Descrição Morfológica e Influência da Temperatura e da Salinidade no Crescimento*. (Dissertação de Mestrado, Universidade de Lisboa, Lisboa, Portugal)
- Martins, R., Costa, F. O., Murta, A. G., Carneiro, M., Landi, M. (2012). First record of *Zenion hololepis* (Zenionidae) in Portuguese continental waters: the northernmost occurrence in the eastern Atlantic. *Marine Biodiversity Records*, 5, e30. <http://doi.org/10.1017/S1755267211000522>
- McDevit, D. C., Saunders, G. W. (2010). A DNA barcode examination of the Laminariaceae (Phaeophyceae) in Canada reveals novel biogeographical and evolutionary insights. *Phycologia* 49, 235-48.
- Mendes, R., Silva, H. (2015) Control of seafood labelling in Portugal. *Relatórios Científicos e Técnicos do IPMA*, 4. Disponível em <https://www.ipma.pt/export/sites/ipma/bin/docs/publicacoes/ipma/Reln4-IPMA.pdf>
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21, 1087–1092. <http://doi.org/doi:10.1063/1.1699114>
- Moura, T., Silva, M. C., Figueiredo, I., Neves, A., Muñoz, P. D., Coelho, M. M., Gordo, L. S. (2008). Molecular barcoding of north-east Atlantic deep-water sharks: species identification and application to fisheries management and conservation. *Marine and Freshwater Research* 59, 214–223.
- Nguyen, L. T., Schmidt, H. A., von Haeseler, A., Minh, B. Q. (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.*, 32, 268-274
- Nicolè, 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.
- Nieto, A., Ralph, G.M., Comeros-Raynal, M.T., Kemp, J., García Criado, M., Allen, D.J., Dulvy, N.K., Walls, R.H.L., Russel, B., Pollard, D., Garcia, S., Craig, M., Collette, B.B., Pollom, R., Biscoito, M., Chao, N.L., Abella, A., Afonso, P., Álvarez, H., Carpenter, K.E., Clò, S., Cook, R., José Costa, M., Delgado, J., Dureuil, M., Ellis, J.R., Farrell, E.D., Fernandes, P., Florin, A.-B., Fordham, S., Fowler, S., Gil de Sola, L., Herrera, J.G., Goodpaster, A., Harvey, M., Heessen, H., Herler, J., Jung, A., Karmovskaya, E., Keskin, C., Knudsen, S.W., Kobylansky, S., Kovacic, M., Lawson, J.M., Lorange, P., Phillips, S.M., Munroe, T., Nedreaas, K., Nielsen, J., Papaconstantinou, C., Polidoro, B., Pollock, C.M., Rijnsdorp, A.D., Sayer, C., Scott, J., Serena, F., Smith-Vaniz, W.F., Soldo, A., Stump, E., Williams,

- J.T. (2015). *European Red List of marine fishes*. Luxemburgo: Publication Office of the European Union. <http://doi.org/10.2779/082723>
- Patarnello, T., Volckaert, F. A. M. J., Castilho, R. (2007) Pillars of Hercules: is the Atlantic–Mediterranean transition a phylogeographical break? *Molecular Ecology*, *16*, 4426–4444. doi: 10.1111/j.1365-294X.2007.03477.x
- QGIS Development Team, 2015. QGIS Geographic Information System. Open Source Geospatial Foundation Project. Disponível em: <http://qgis.osgeo.org>
- Rasmussen, R. S., Morrissey, M. T. (2008). DNA-Based Methods for the Identification of Commercial Fish and Seafood Species. *Compr Rev Food Sci Food Saf* *7*, 280–295.
- Ratnasingham, S., Hebert, P. D. N. (2007). bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular Ecology Notes*, *7*, 355–364. <http://doi.org/10.1111/j.1471-8286.2007.01678.x>
- Ratnasingham, S., Hebert, P. D. N. (2013). A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. *PLoS ONE*, *8*, e66213. <http://doi.org/10.1371/journal.pone.0066213>
- Rice, P., Longden, I., Bleasby, A. (2000). EMBOSS: the European molecular biology open software suite. *Trends Genet.* *16*, 276–277
- Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. (2012). MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic Biology*, *61*(3), 539–542. <http://doi.org/10.1093/sysbio/sys029>
- Saitou, N., Nei, M. (1987). The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Molecular Biology and Evolution*, *4*, 406–425. Retrieved from <http://mbe.oxfordjournals.org/content/4/4/406.short>
- Serra-Pereira, B., Moura, T., Griffiths, A. M., Serrano Gordo, L., Figueiredo, I. (2011). Molecular barcoding of skates (Chondrichthyes: Rajidae) from the southern Northeast Atlantic. *Zoologica Scripta*, *40*, 76–84. <http://doi.org/10.1111/j.1463-6409.2010.00461.x>
- Steinke, D., Zemlak T. S., Hebert P. D. N. (2009). Barcoding Nemo: DNA-Based Identifications for the Ornamental Fish Trade. *PLoS ONE*, *4*, e6300. doi:10.1371/journal.pone.0006300
- Straube, N., Iglésias, S. P., Sellos, D. Y., Kriwet, J., Schlieven, U. K. (2010). Molecular phylogeny and node time estimation of bioluminescent Lantern Sharks (Elasmobranchii: Etmopteridae). *Molecular Phylogenetics and Evolution*, *56*, 905–917. <http://doi.org/10.1016/j.ympev.2010.04.042>
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, *30*, 2725–2729. <http://doi.org/10.1093/molbev/mst197>
- Torres, R. A., Feitosa, R. B., Carvalho, D. C., Freitas, M. O., Hostim-Silva, M., Ferreira, B. P. (2013) DNA barcoding approaches for fishing authentication of exploited grouper species including the endangered and legally protected goliath grouper *Epinephelus itajara*. *Sci. Mar.*, *77*, 409-418
- Ward, R., Costa, F., Holmes, B., Steinke, D. (2008). DNA barcoding of shared fish species from the North Atlantic and Australasia: minimal divergence for most taxa, but *Zeus faber* and *Lepidopus caudatus*

- each probably constitute two species. *Aquatic Biology*, 3, 71–78.  
<http://doi.org/10.3354/ab00068>
- Ward, R. D., Hanner, R., Hebert, P. D. N. (2009). The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology*, 74(2), 329–356. <http://doi.org/10.1111/j.1095-8649.2008.02080.x>
- Ward, R. D., Holmes, B. H. (2007). An analysis of nucleotide and amino acid variability in the barcode region of cytochrome c oxidase I (cox1) in fishes. *Molecular Ecology Notes*, 7, 899–907. <http://doi.org/10.1111/j.1471-8286.2007.01886.x>
- Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360, 1847–1857. <http://doi.org/10.1098/rstb.2005.1716>
- Yang, Z., Rannala, B. (1997). Bayesian phylogenetic Monte Carlo method inference using DNA sequences: a Markov Chain Monte Carlo method. *Molecular Biology and Evolution*, 14, 717–724.
- Zemlak, T. S., Ward, R. D., Connell, A. D., Holmes, B. H., Hebert, P. D. N. (2009). DNA barcoding reveals overlooked marine fishes. *Molecular Ecology Resources*, 9, 237–242. <http://doi.org/10.1111/j.1755-0998.2009.02649.x>



## ANEXO I

**Tabela 10 – Número de acesso do GenBank de todas as sequências da biblioteca de referências para peixes marinhos europeus. \*BOLD-ID**

<b>Identificação</b>	<b>Nº de acesso GenBank</b>	<b>Identificação</b>	<b>Nº de acesso GenBank</b>
<i>Acantholabrus palloni</i>	JQ774957 - JQ774958 KJ768197 - KJ768199	<i>Anthias anthias</i>	JQ774769 JQ774959
<i>Acipenser gueldenstaedtii</i>	JQ623904 KC500088 - KC500102		KJ709462 KJ709463
<i>Acipenser nudiiventris</i>	JQ623905 KC500103 - KC500118	<i>Antigonia capros</i>	KJ768204 KJ768205
<i>Acipenser stellatus</i>	JQ623906 KC500105 KC500119 - KC500132	<i>Aphanius fasciatus</i>	KJ709464 - KJ709474
<i>Aculeola nigra</i>	GU130678.1 GU130703.1 GU130704.1	<i>Aphia minuta</i>	KJ204698 - KJ204702 KM077806 - KM077815
<i>Agonus cataphractus</i>	KJ204628 - KJ204645 KJ205264 - KJ205266 KJ205339	<i>Apristurus longicephalus</i>	GU130672.1
<i>Alburnus tarichi</i>	JQ623908 KC500153 - KC500172	<i>Argentina silus</i>	KJ204703 - KJ204704 JQ774570 - JQ774571 JQ774770 - JQ774773 JQ774960 - JQ774964
<i>Alepocephalus rostratus</i>	KJ768201	<i>Argentina sphyraena</i>	KJ204705 - KJ204721 KJ205267 - KJ205268 KJ205344
<i>Alosa alosa</i>	KC500173 - KC500192		KJ709476
<i>Alosa fallax</i>	KJ204646 - KJ204654 KJ768202		KJ709687 - KJ709693 KJ768206
<i>Amblyraja hyperborea</i>	KF604118 - KF604127	<i>Argyrosomus regius</i>	JQ623911 KC500213 - KC500232
<i>Amblyraja radiata</i>	JN312484 KF604140 KF604162 KF604165 KF604168 KF604175 KF604184 KF604199 KF604204 KF604207 KJ204655 - KJ204669 KJ205340 - KJ205343	<i>Arnoglossus imperialis</i>	JQ774774 - JQ774775 JN312467 - JN312470 JQ774776 - JQ774780 JQ774965 - JQ774968 KJ204722 - KJ204728
<i>Ammodytes marinus</i>	KJ204670 - KJ204693	<i>Arnoglossus laterna</i>	KJ205269 - KJ205272 KJ709694 - KJ709697
<i>Ammodytes tobianus</i>	FCFBS032-07*	<i>Arnoglossus rueppelii</i>	JQ774781 - JQ774785 JQ774969 - JQ774973
<i>Anarhichas lupus</i>	KJ204694 - KJ204697	<i>Arnoglossus thori</i>	KJ709698 - KJ709702
<i>Anguilla anguilla</i>	JQ623910 KC500193 - KC500212 KJ768203	<i>Atherina boyeri</i>	KJ709477
		<i>Atherina hepsetus</i>	JQ623913 - KC500272
		<i>Aulopus filamentosus</i>	KJ709481 - KJ709482 KJ768210
		<i>Auxis rochei rochei</i>	KJ709483 KJ768211 MLFPI355-14*

**Tabela 10 (continuação) – Número de acesso do GenBank de todas as sequências da biblioteca de referências para peixes marinhos europeus. \*BOLD-ID**

<b>Identificação</b>	<b>Nº de acesso GenBank</b>	<b>Identificação</b>	<b>Nº de acesso GenBank</b>
<i>Auxis thazard</i>	JQ623915 KC500273 - KC500292	<i>Capoeta capoeta</i>	JQ623921 KC500353 - KC500372
<i>Balistes capriscus</i>	JQ623916 KC500293 - KC500312	<i>Capros aper</i>	JQ774583 - JQ774802 JQ774989 - JQ774993 KJ204774
<i>Balistes capriscus</i>	KJ709484 - KJ709486 KJ768212		KJ709493 - KJ709494 KJ709713 - KJ709720
<i>Bathyraja spinicauda</i>	KF604208 - KF604217	<i>Centracanthus cirrus</i>	MLFPI088-09*
<i>Bathysolea profundicola</i>	KJ768213 - KJ768215	<i>Centrolophus niger</i>	KJ709721 EU003893.1 - EU003894.1
<i>Belone belone</i>	JQ623917 JQ774572 KC500313 - KC500332 KJ204729 - KJ204731 KJ768216	<i>Centrophorus granulosus</i>	JQ774505 EU003895.1 - EU003897.1
<i>Benthodesmus simonyi</i>	JQ774573 - JQ774574	<i>Centrophorus squamosus</i>	EU003896.1 GU130701.1 GU130705.1 - GU130708.1
<i>Beryx decadactylus</i>	KJ768217		EU003883.1 - EU003885.1 EU003886.1 - GU130695.1
<i>Blennius ocellaris</i>	JQ774786 - JQ774790 JQ774983 KJ205273 KJ205345 KJ709487 - KJ709489 KJ709703 - KJ709707 KJ768218	<i>Centrosyllium fabricii</i>	GU130696.1 GU130694.1
<i>Boops boops</i>	JQ623918 JQ774575 - JQ774579 JQ774791 - JQ774793 JQ774984 - JQ774987 KC500333 - KC500352 KJ709490 - KJ709491 KJ709708 - KJ709712	<i>Centrosymnus coelolepis</i>	EU003886.1 - EU003888.1 EU003889.1
		<i>Centrosymnus owstoni</i>	GU130695.1
<i>Bothus podas</i>	KJ709492	<i>Centrosymnus plunketi</i>	GU130696.1
<i>Brama brama</i>	KJ204732 KJ768224 - KJ768225	<i>Centroselachus crepidater</i>	GU130694.1
<i>Brosme brosme</i>	KJ204733	<i>Cepola macrophthalma</i>	JQ774589 - JQ774997 KJ205352 - KJ205358 KJ709722 - KJ709726
<i>Buenia jeffreysii</i>	KM077816 - KM077819		KJ709495 KJ709727
<i>Buglossidium luteum</i>	JN312471 - JN312474 KJ204734 - KJ205348	<i>Chauliodus sloani</i>	KJ709495 KJ709727
<i>Callionymus lyra</i>	JQ774580 - JQ774582 JQ774794 - JQ774796 JQ774988 KJ204742 - KJ204756 KJ205349 - KJ205351	<i>Chaunax pictus</i>	JQ774506 - JQ774507 MLFPI077-09*
		<i>Cheilopogon heterurus</i>	MLFPI062-09*
<i>Callionymus maculatus</i>	JQ774797	<i>Chelidonichthys cuculus</i>	JQ623912 JQ774974 - JQ774982 KC500233 - KC500252 KJ204775 - KJ204780 KJ205359 KJ709496 KJ709728 - KJ709732
		<i>Chelidonichthys lucerna</i>	JN312483 JQ623923 JQ774803 - JQ774807 JQ774998 - JQ775001

**Tabela 10 (continuação) – Número de acesso do GenBank de todas as sequências da biblioteca de referências para peixes marinhos europeus. \*BOLD-ID**

<b>Identificação</b>	<b>Nº de acesso GenBank</b>	<b>Identificação</b>	<b>Nº de acesso GenBank</b>
	JQ774999	<i>Ctenolabrus rupestris</i>	KJ204821
	JQ775000	<i>Cyclopterus lumpus</i>	JN311799
	KC500373 - KC500392		JN311800 - JN311803
	KJ204781 - KJ709498		JN311801
<i>Chelidonichthys obscurus</i>	JQ774508		JN311802
	JQ774594 - JQ774595		KJ204822 - KJ204830
	JQ774808 - JQ774812	<i>Cyttopsis rosea</i>	JQ774518 - JQ774524
	JQ775002 - JQ775003	<i>Dactylopterus volitans</i>	KJ709511 - KJ709512
<i>Chelon labrosus</i>	KJ768226 - KJ768227	<i>Dalatias licha</i>	GU130676.1
<i>Chimaera monstrosa</i>	JQ774509	<i>Deania calcea</i>	GU130699.1
	JQ774596 - JQ774599	<i>Deania profundorum</i>	JQ774525 - JQ774528
	EFBC005-09*	<i>Dentex dentex</i>	JQ623927
	EFBC006-09*		KC500453 - KC500472
	EFBC007-09*	<i>Dentex macrophthalmus</i>	JQ623928
	EFBC008-09*		KC500473 - KC500492
<i>Chimaera opalescens</i>	EFBC001-09*		KJ709513
	EFBC002-09*		KJ709745 - KJ709748
	EFBC003-09*	<i>Diaphus holti</i>	KJ709514
	EFBC004-09*	<i>Diaphus metopoclampus</i>	KJ709515
<i>Chlorophthalmus agassizi</i>	JQ774510 - JQ774514		KJ709749 - KJ709750
	KF714757	<i>Dicentrarchus labrax</i>	KC500493 - KC500512
	KJ709499		KJ205274 - KJ205275
<i>Ciliata mustela</i>	KJ204791 - KJ204807	<i>Dicologlossa cuneata</i>	JQ775008
<i>Citharus linguatula</i>	JQ623924	<i>Diplodus annularis</i>	JQ623930
	JQ774600 - JQ774602		JQ774608 - JQ774609
	JQ774813 - JQ774816		KC500513 - KC500532
	KC500393 - KC500412		KJ709516 - KJ709517
	KJ709500 - KJ709501	<i>Diplodus bellottii</i>	MLFPI128-10*
	KJ709734 - KJ709738		MLFPI336-13*
<i>Clarias gariepinus</i>	JQ623925	<i>Diplodus cervinus</i>	JQ623931
	KC500413 - KC500432		KC500533 - KC500552
<i>Clupea harengus</i>	KJ204808 - KJ204816	<i>Diplodus puntazzo</i>	JQ623932
	KJ205361 - KJ205365		KC500553 - KC500572
<i>Coelorinchus caelorhincus</i>	JQ774515 - JQ774516	<i>Diplodus sargus</i>	JQ623933
	KJ709502 - KJ709503		JQ775009
<i>Conger conger</i>	JQ774603 - JQ774607		KC500573 - KC500592
	JQ774817 - JQ774821	<i>Diplodus vulgaris</i>	JQ623934
	JQ775004 - JQ775007		KC500593 - KC500612
	KJ709504		KJ709519 - KJ709521
	KJ709739 - KJ709742		MLFPI196-10*
<i>Coris julis</i>	KJ709505 - KJ709510		MLFPI197-10*
<i>Coryphaena hippurus</i>	KJ709744	<i>Dipturus batis</i>	JQ774529
<i>Crystalllogobius linearis</i>	KJ204817 - KM077825		

**Tabela 10 (continuação) – Número de acesso do GenBank de todas as sequências da biblioteca de referências para peixes marinhos europeus. \*BOLD-ID**

<b>Identificação</b>	<b>Nº de acesso GenBank</b>	<b>Identificação</b>	<b>Nº de acesso GenBank</b>
	KJ204831	<i>Etmopterus dislineatus</i>	GU130681.1
<i>Dipturus linteus</i>	KF604231	<i>Etmopterus fusus</i>	GU130687.1
<i>Dipturus nidarosiensis</i>	KF604234	<i>Etmopterus gracilispinis</i>	GU130724.1
	KF604239 - KF604243	<i>Etmopterus granulosus</i>	GU130679.1
<i>Dipturus oxyrinchus</i>	HM043215.1 - HM043218.1		GU130736.1 - GU130738.1
	JQ774530		GU130682.1 -
	KF604244 - KJ709754	<i>Etmopterus lucifer</i>	GU130683.1
<i>Dussumieria elopsoides</i>	JQ623935		GU130711.1 - GU130714.1
	KC500613 - KC500632	<i>Etmopterus molleri</i>	GU130710.1
<i>Echelus myrus</i>	KJ709755		GU130715.1
<i>Echiichthys vipera</i>	JQ775010 - JQ775014	<i>Etmopterus polli</i>	GU130742.1
	KJ204832 - KJ204839	<i>Etmopterus princeps</i>	GU130727.1
	KJ205276 - KJ205279	<i>Etmopterus pseudosqualiolus</i>	GU130686.1
	KJ205366	<i>Etmopterus pusillus</i>	EU869807 - EU869810
<i>Electrona risso</i>	KJ709523		GU130688.1 - GU130689.1
<i>Enchelyopus cimbrius</i>	KJ204840 - KJ204849		GU130721.1 - GU130722.1
	KJ205367 - KJ205371		JQ774531
<i>Engraulis encrasicolus</i>	HM480814.1 - HM480816.1	<i>Etmopterus schultzi</i>	GU130719.1
	JN007762.1 - JN007768.1	<i>Etmopterus sentosus</i>	GU130720.1
	JQ623936	<i>Etmopterus sheikoi</i>	GU130680.1
	JQ774610 - JQ774611		GU130709.1
	JQ774823	<i>Etmopterus spinax</i>	GU130691.1
	JQ774826		GU130725.1 - GU130726.1
	JQ775015 - JQ775020		JQ774532
	KC500633 - KC500652		JQ774612 - JQ774616
	KJ204850 - KJ204859		KJ709757 - KJ709759
	KJ709524		GU130739.1 - GU130741.1
<i>Entelurus aequoreus</i>	KJ205372 - KJ205375	<i>Etmopterus unicolor</i>	GU130743.1
<i>Epigonus telescopus</i>	KJ709756	<i>Euthynnus alletteratus</i>	JQ623941
<i>Epinephelus aeneus</i>	JQ623937		KC500733 - KC500752
	KC500653 - KC500672		KJ709529
	KJ709526	<i>Eutrigla gurnardus</i>	JN312463 - JN312466
<i>Epinephelus caninus</i>	KJ709527		JN312479 - JN312482
<i>Epinephelus costae</i>	KJ709528		JQ775021 - JQ775023
<i>Epinephelus marginatus</i>	JQ623938		KJ204860 - KJ204863
	KC500673 - KC500692		KJ205377
<i>Etmopterus baxteri</i>	GU130737.1		KJ709530
<i>Etmopterus bigelowi</i>	GU130690.1		KJ709760
	GU130723.1		
	GU130716.1 -		
<i>Etmopterus brachyurus</i>	GU130718.1		
<i>Etmopterus dianthus</i>	GU130693.1		

**Tabela 10 (continuação) – Número de acesso do GenBank de todas as sequências da biblioteca de referências para peixes marinhos europeus. \*BOLD-ID**

<b>Identificação</b>	<b>Nº de acesso GenBank</b>	<b>Identificação</b>	<b>Nº de acesso GenBank</b>
<i>Facciolella oxyrhyncha</i>	JQ774533 - JQ774535 JQ774617		KJ709777 - KJ709780 KJ768238 - KJ768239
<i>Gadiculus argenteus</i>	JQ774618 - JQ775028 KJ204864 - KJ204873 KJ709531 - KJ709532 KJ709761	<i>Hygophum benoiti</i>	KJ709536
<i>Gadomus longifilis</i>	KJ768230 - KJ768232	<i>Hymenocephalus italicus</i>	KJ709781 - KJ709783
<i>Gadus morhua</i>	KJ204874 - KJ204885 KJ205378 - KJ205380	<i>Hyperoplus immaculatus</i>	KJ204926 - KJ204933
<i>Gaidropsarus mediterraneus</i>	JQ774626 KJ709762 - KJ709764	<i>Hyperoplus lanceolatus</i>	JQ775039 - JQ775043
<i>Gaidropsarus vulgaris</i>	KJ204886 - KJ204889	<i>Hyporthodus haifensis</i>	KJ709537 - KJ709539
<i>Galeorhinus galeus</i>	KJ204890 - KJ204892	<i>Isurus oxyrinchus</i>	EU869822
<i>Galeus atlanticus</i>	JQ774627 - JQ774837	<i>Labrus merula</i>	KJ709540 - KJ709541
<i>Galeus melastomus</i>	JQ774633 - JQ774634 JQ774838 - JQ774841 KJ709765 - KJ709769	<i>Labrus mixtus</i>	JQ775044 KJ768240 - KJ768243
<i>Gasterosteus aculeatus</i>	KJ204893 - KJ204898	<i>Lagocephalus lagocephalus</i>	MLFPI331-13*
<i>Glyptocephalus cynoglossus</i>	KJ204899 - KJ204915 KJ205381 - KJ205382	<i>Lampanyctus crocodilus</i>	KJ709784 - KJ709788
<i>Gobiosoma bosc</i>	KM077826 - KM077829	<i>Lepadogaster lepadogaster</i>	KJ768244 - KJ768246
<i>Gobius niger</i>	JQ623942 KC500753 - KC500772 KJ204916 KJ205280 - KJ205281 KJ768236 KM077836 - KM077839	<i>Lepidocybium flavobrunneum</i>	KJ768247
<i>Gobiusculus flavescens</i>	KM077830 - KM077835	<i>Lepidopus caudatus</i>	EU869827 - EU869832 KJ709542 - KJ768251
<i>Grammicolepis brachiusculus</i>	KJ768237		MLFPI367-14*
<i>Halobatrachus didactylus</i>	JQ774635	<i>Lepidorhombus boscii</i>	JQ774842 - JQ775048 KJ709793
<i>Helicolenus dactylopterus</i>	JQ774636 - JQ775038 KJ204917 - KJ204918 KJ709533 KJ709770 - KJ709774	<i>Lepidorhombus whiffiagonis</i>	JQ774536 KJ204934 - KJ204942 KJ709794
<i>Heptranchias perlo</i>	EU869819 KJ709775 KJ709776	<i>Lepidotrigla cavillone</i>	JQ774639 - JQ774643 JQ774848 - JQ774849 KJ709795 - KJ709800
<i>Hippoglossoides platessoides</i>	JN312184 - JN312191 KJ204919 - KJ204921 KJ205383	<i>Lepidotrigla dieuzeidei</i>	JQ774644 - JQ775054 KJ709801
<i>Hippoglossus hippoglossus</i>	KJ204922 - KJ204925		KJ768252 - KJ768253 HM043205.1 - HM043206.1
<i>Hoplostethus mediterraneus</i>	EU869820 - EU869821	<i>Leucoraja circularis</i>	JQ774537 KF604254 KF604256
		<i>Leucoraja fullonica</i>	KF604257 - KF604259 HM043211.1 - HM043214.1
		<i>Leucoraja naevus</i>	JQ774853 KF604260 - KF604263 KJ204943 - KJ204956 KJ205282 - KJ205384

**Tabela 10 (continuação) – Número de acesso do GenBank de todas as sequências da biblioteca de referências para peixes marinhos europeus. \*BOLD-ID**

<b>Identificação</b>	<b>Nº de acesso GenBank</b>	<b>Identificação</b>	<b>Nº de acesso GenBank</b>
<i>Lichia amia</i>	JQ623944 KC500773 - KC500792		KC500912 - KJ205026 KJ205393 - KJ205396
<i>Limanda limanda</i>	JN312159 - JN312162 JN312180 - JN312183 KJ204957 - KJ204963 KJ205385 - KJ205387	<i>Merluccius merluccius</i>	JQ623955 JQ774659 - JQ774668 JQ774859 - JQ774863 JQ775071 - JQ775075 KC500913 - KC500932
<i>Liparis liparis</i>	KJ204964 - KJ204977		KJ205027 - KJ205034 KJ709559 - KJ709560 KJ709813 - KJ709816
<i>Lithognathus mormyrus</i>	JQ623945 KC500793 - KC500812 KJ709548 - KJ709549		JQ774669 - JQ774672 JQ774673 - JQ774677 KJ709561 KJ768258
<i>Liza aurata</i>	JQ623946 KC500813 - KC500832	<i>Microchirus azevia</i> <i>Microchirus boscanion</i> <i>Microchirus ocellatus</i>	JQ774669 - JQ774672 JQ774673 - JQ774677 KJ709561 KJ768258
<i>Liza carinata</i>	JQ623947  KC500833 - KC500852		MLFPI212-11*
<i>Liza ramada</i>	JQ775055 - JQ775059	<i>Microchirus variegatus</i>	JQ774678 - JQ774681 JQ774864 - JQ774866 JQ775076 - JQ775080 KJ205035 - KJ205042 KJ205291 KJ205397 - KJ205400 KJ768259
<i>Lobianchia doffeini</i>	KJ709550		
<i>Lobianchia gemellarii</i>	KJ709802		
<i>Lobotes surinamensis</i>	JQ623948 KC500853 - KC500872		
<i>Lophius budegassa</i>	JQ774651 - JQ774652 KJ204978 - KJ204981 KJ709551 - KJ709552 KJ709803 - KJ709806	<i>Micromesistius poutassou</i>	JQ774682 - JQ774686 JQ774868 - JQ774871 JQ775081 - JQ775086 KJ205043 - KJ205045 KJ709817 - KJ709821
<i>Lophius piscatorius</i>	JQ623950 KC500873 - KC500892 KJ204982 - KJ204994 KJ205284 - KJ205287 KJ709553 KJ709807	<i>Microstomus kitt</i>	JN312167 - JN312170 KJ205046 - KJ205058 KJ205401 - KJ205405
<i>Lumpenus lampretaeformis</i>	KJ204995 - KJ205000 KJ205388 - KJ205390 KJ205389	<i>Mola mola</i> <i>Molva molva</i>	JQ775087 - JQ775088 JQ774540 KJ205059 - KJ205066 KJ768260
<i>Macroramphosus scolopax</i>	JQ774653 - JQ775070 KJ709554 - KJ709556 KJ709808 - KJ709812	<i>Mugil cephalus</i>	JQ623956 KC500933 - KC500952
<i>Malacocephalus laevis</i>	JQ774539 KJ768255 - KJ768256	<i>Mullus barbatus</i>	JQ623957 JQ774687
<i>Maurollicus muelleri</i>	KJ709557 - KJ709558		KC500953 - KC500972 KJ709562 - KJ709568
<i>Melanogrammus aeglefinus</i>	KJ205001 - KJ205013 KJ205288 - KJ205391		
<i>Merlangius merlangus</i>	JQ623954 KC500893 - KC500911		

**Tabela 10 (continuação) – Número de acesso do GenBank de todas as sequências da biblioteca de referências para peixes marinhos europeus. \*BOLD-ID**

<b>Identificação</b>	<b>Nº de acesso GenBank</b>	<b>Identificação</b>	<b>Nº de acesso GenBank</b>
	KJ709822 - KJ709827	<i>Ophisurus serpens</i>	MLFPI366-14*
	KJ768261 - KJ768263	<i>Osmerus eperlanus</i>	KJ205104 - KJ205109
<i>Mullus surmuletus</i>	JQ623958	<i>Oxynotus paradoxus</i>	GU130674.1
	JQ774872 - JQ774876	<i>Pagellus acarne</i>	JQ774690 - JQ774691
	KC500973 - KC500992		JQ774877 - JQ774881
	KJ205067 - KJ205295		JQ775089 - JQ775093
<i>Muraena helena</i>	KJ709828		KJ709573 - KJ709574
	KJ768264	<i>Pagellus bogaraveo</i>	KJ768270 - KJ768272
<i>Mustelus asterias</i>	KJ205082 - KJ205091	<i>Pagellus erythrinus</i>	JQ623965
	KJ205296		JQ774882 - JQ774886
	KJ205407		KC501033 - KC501052
	KJ709829 - KJ709832		MLFPI274-11*
<i>Mustelus mustelus</i>	KJ709833 - KJ709836		MLFPI275-11*
	KJ768265 - KJ768266		MLFPI279-11*
<i>Mustelus punctulatus</i>	KJ709837		MLFPI280-11*
<i>Mycteroperca rubra</i>	KJ709569	<i>Pagrus auriga</i>	JQ623966
<i>Myctophum punctatum</i>	KJ709570		KC501053 - KC501072
			KJ768273
	KJ709838 - KJ709839	<i>Pagrus pagrus</i>	JQ623967
<i>Myliobatis aquila</i>	KJ709571 - KJ709572		
<i>Myoxocephalus scorpioides</i>	FCFBS033-07*		KC501073 - KC501092
	FCFBS034-07*		KJ709847 - KJ709850
	FCFBS035-07*		KJ768274 - KJ768275
	FCFBS036-07*	<i>Pegusa impar</i>	KJ709575
<i>Myoxocephalus scorpius</i>	JN312475 - JN312478	<i>Pegusa lascaris</i>	KJ205297 - KJ205298
	KJ205092 - KJ205103		KJ205411
	KJ205408 - KJ205410		KJ768276
<i>Neogobius melanostomus</i>	JQ623961	<i>Peristedion cataphractum</i>	JQ774692 - JQ774693
	KC500993 - KC501012		KJ709577 - KJ709581
	KM077840 - KM077846		KJ709851 - KJ709855
	FCFBS010-07*		KJ768277
	FCFBS011-07*	<i>Petromyzon marinus</i>	MLFPI230-11*
<i>Neoraja iberica</i>	HM043186.1 - HM043191.1	<i>Pholis gunnellus</i>	KJ205110 - KJ205118
<i>Nettastoma melanurum</i>	KJ709840 - KJ709841		KJ205299
	KJ768267	<i>Phrynorhombus norvegicus</i>	KJ205119 - KJ205121
<i>Nezumia sclerorhynchus</i>	JQ774541 - JQ774544	<i>Phycis blennoides</i>	JQ774545 - KJ709860
	KJ709842 - KJ709843	<i>Phycis phycis</i>	KJ768278 - KJ768279
<i>Notacanthus bonaparte</i>	KJ709844 - KJ709846	<i>Platichthys flesus</i>	JQ623970
	KJ768268 - KJ768269		JQ775094 - JQ775096
<i>Oblada melanura</i>	JQ623963		KC501093 - KC501112
	KC501013 - KC501032		KJ205122 - KJ205132
<i>Odontaspis ferox</i>	GU130673.1		KJ205300 - KJ205302

**Tabela 10 (continuação) – Número de acesso do GenBank de todas as sequências da biblioteca de referências para peixes marinhos europeus. \*BOLD-ID**

<b>Identificação</b>	<b>Nº de acesso GenBank</b>	<b>Identificação</b>	<b>Nº de acesso GenBank</b>
	KC501073 - KC501092		FCFBS017-07*
	KJ709847 - KJ709850		
	KJ768274 - KJ768275		FCFBS018-07*
<i>Pegusa impar</i>	KJ709575		FCFBS019-07*
<i>Pegusa lascaris</i>	KJ205297 - KJ205298		FCFBS020-07*
	KJ205411		FCFBS021-07*
	KJ768276		FCFBS022-07*
<i>Peristedion cataphractum</i>	JQ774692 - JQ774693	<i>Pomatoschistus norvegicus</i>	KM077867 - KM077874
	KJ709577 - KJ709581	<i>Pomatoschistus pictus</i>	KM077875 - KM077878
	KJ709851 - KJ709855	<i>Pomatoschistus tortonesei</i>	KJ709585 - KJ709586
	KJ768277	<i>Ponticola kessleri</i>	JQ623973
<i>Petromyzon marinus</i>	MLFPI230-11*		KC501133 - KC501152
<i>Pholis gunnellus</i>	KJ205110 - KJ205118	<i>Pontinus kuhlii</i>	JQ774695 - JQ774696
	KJ205299	<i>Prionace glauca</i>	EU869837
<i>Phrynorhombus norvegicus</i>	KJ205119 - KJ205121		MLFPI321-12*
<i>Phycis blennoides</i>	JQ774545 - KJ709860		MLFPI322-12*
<i>Phycis phycis</i>	KJ768278 - KJ768279		MLFPI324-12*
<i>Platichthys flesus</i>	JQ623970	<i>Pterois miles</i>	KJ709587 - KJ709588
	JQ775094 - JQ775096		HM043201.1 -
	KC501093 - KC501112	<i>Raja brachyura</i>	HM043203.1
	KJ205122 - KJ205132		JQ774887 - JQ775098
			KF604264 - KF604267
	KJ205300 - KJ205302		KJ205303
	KJ768280		KJ768288 - KJ768290
<i>Pleuronectes platessa</i>	JN312163 - JN312179	<i>Raja clavata</i>	HM043193.1 -
	KJ205133 - KJ205134		HM043196.1
	KJ205412 - KJ205413		JQ774697 - JQ775101
<i>Pollachius pollachius</i>	KJ205135 - KJ205137		KF604268 - KF604277
<i>Pollachius virens</i>	KJ205138 - KJ205149	<i>Raja maderensis</i>	KJ205150 - KJ709867
<i>Polymetme corythaeola</i>	JQ774548 - JQ774549		HM043185.1
<i>Polyprion americanus</i>	KJ768281	<i>Raja microocellata</i>	HM043197.1 -
<i>Pomadasyus incisus</i>	KJ768282 - KJ768284		HM043200.1
<i>Pomatomus saltatrix</i>	JQ623971	<i>Raja miraletus</i>	KJ205305 - KJ205308
	KC501113 - KC501132		HM043182.1 -
<i>Pomatoschistus lozanoi</i>	JQ775029 - JQ775033		HM043184.1
	KM077847 - KM077849		JQ774701 - JQ774704
<i>Pomatoschistus marmoratus</i>	KJ709583 - KJ709584		JQ774892 - JQ774894
<i>Pomatoschistus microps</i>	KJ768285 - KJ768287	<i>Raja montagui</i>	KJ709590 - KJ709592
	KM077850 - KM077856		KJ709868 - KJ709872
	FCFBS012-07*		HM043204.1
<i>Pomatoschistus minutus</i>	KM077857 - KM077866		HM043207.1 -
	FCFBS014-07*		HM043210.1
	FCFBS015-07*		JQ774705 - JQ774707
			JQ774895



**Tabela 10 (continuação) – Número de acesso do GenBank de todas as sequências da biblioteca de referências para peixes marinhos europeus. \*BOLD-ID**

<b>Identificação</b>	<b>Nº de acesso GenBank</b>	<b>Identificação</b>	<b>Nº de acesso GenBank</b>
	KF604279 - KF604282		JQ623985
	KJ205152 - KJ205155		JQ774720 - JQ774725
	KJ205416		JQ774901 - JQ774905
<i>Raja polystigma</i>	KJ709593 - KJ709594		JQ775115 - JQ775118
	HM043219.1		KC501353 - KC501372
<i>Raja undulata</i>	HM043222.1		KJ205159 - KJ205163
<i>Rajella fyllae</i>	KF604283 - KF604292		KJ205417 - KJ205420
			KJ709609
<i>Ranzania laevis</i>	KJ768291		KJ768300 - KJ768302
<i>Rostroraja alba</i>	HM043192.1		
<i>Salmo ohridanus</i>	KJ709597	<i>Scomberesox saurus</i>	KJ768299
<i>Salmo trutta trutta</i>	JQ623976	<i>Scomberomorus commerson</i>	JQ623986
	KC501153 - KC501172		KC501333 - KC501352
	KJ709598 - KJ709600	<i>Scombrolabrax heterolepis</i>	KJ768303
<i>Sarda sarda</i>	JQ623978	<i>Scophthalmus maximus</i>	JQ623987
	KC501193 - KC501212		KC501373 - KC501392
	KJ709601		KJ205164 - KJ205173
	KJ768292 - KJ768295		KJ205421 - KJ205427
<i>Sardina pilchardus</i>	JQ623979	<i>Scophthalmus rhombus</i>	KJ205174 - KJ205179
	JQ774708 - JQ774712		KJ205428
	JQ774897 - JQ775107		KJ768304
	KC501213 - KC501232	<i>Scorpaena notata</i>	JQ774726 - JQ774727
	KJ205156 - KJ205158		JQ774906 - JQ774910
	KJ768296 - KJ768297		KJ709610 - KJ709612
<i>Sargocentron rubrum</i>	JQ623980		KJ709878
	KC501233 - KC501252		KJ768305 - KJ768308
<i>Sarpa salpa</i>	JQ623981	<i>Scorpaena porcus</i>	JQ623988
	KC501253 - KC501272		KC501393 - KC501412
	KJ709602 - KJ709604	<i>Scorpaena scrofa</i>	JQ623989
<i>Saurida undosquamis</i>	JQ623982		KC501413 - KC501432
	KC501273 - KC501292		KJ709879 - KJ709890
<i>Schedophilus medusophagus</i>	KJ709605	<i>Scorpaenodes arenai</i>	KJ709613
<i>Scomber colias</i>	JQ774713 - JQ774719	<i>Scyliorhinus canicula</i>	JQ774728 - JQ774732
	JQ775108 - JQ775114		JQ774911 - JQ775126
	KJ709606 - KJ709608		KJ205180 - KJ205190
	KJ709873 - KJ709877		KJ205309 - KJ205313
	KJ768298		KJ205429
<i>Scomber japonicus</i>	JQ623984		KJ709614 - KJ709620
	KC501313 - KC501332		KJ709891 - KJ709898
	HM480797.1 -		
<i>Scomber scombrus</i>	HM480799.1	<i>Scyliorhinus stellaris</i>	KJ205314
	HM480819.1		KJ205430
	JN007745.1 - JN007777.1		KJ709621 - KJ709622

**Tabela 10 (continuação) – Número de acesso do GenBank de todas as sequências da biblioteca de referências para peixes marinhos europeus. \*BOLD-ID**

<b>Identificação</b>	<b>Nº de acesso GenBank</b>	<b>Identificação</b>	<b>Nº de acesso GenBank</b>
	KJ709899 - KJ709900		KJ709921 - KJ709922
<i>Scymnodon ringens</i>	GU130697.1	<i>Spondyllosoma cantharus</i>	JQ624002
<i>Seriola dumerili</i>	JQ623993		JQ774743 - JQ774747
	KC501433 - KC501452		JQ774923 - JQ775133
<i>Serranus cabrilla</i>	JQ623994		KC501613 - KC501632
	KC501453 - KC501472		KJ205433
	KJ709623 - KJ709630	<i>Sprattus sprattus</i>	JQ624003
	KJ709901 - KJ768309		KC501633 - KC501652
<i>Serranus hepatus</i>	JQ774733 - JQ774737		KJ205197 - KJ205205
	JQ774916 - JQ774917		KJ205434
	KJ709631	<i>Squaliolus aliae</i>	GU130675.1
		<i>Squalus acanthias</i>	KJ205206 - KJ205210
	KJ709909 - KJ709913		
<i>Serranus scriba</i>	JQ623995	<i>Squalus blainville</i>	KJ709641
	KC501473 - KC501492		KJ709923 - KJ709926
	KJ709632	<i>Squalus megalops</i>	GU130698.1
<i>Solea aegyptiaca</i>	KJ709633	<i>Squatina squatina</i>	JQ624004
<i>Solea senegalensis</i>	JQ775127		KC501653 - KC501672
	KJ768310	<i>Stomias boa</i>	KJ709927 - KJ709928
<i>Solea solea</i>	JQ623997	<i>Stromateus fiatola</i>	MLFPI020-09*
	JQ774918 - JQ774922	<i>Symbolophorus veranyi</i>	KJ709643
	KC501513 - KC501532	<i>Symphodus bailloni</i>	MLFPI091-09*
	KJ205191 - KJ205196	<i>Symphodus tinca</i>	KJ709644
	KJ205431 - KJ205432	<i>Symphurus nigrescens</i>	JQ774554 - JQ774555
<i>Somniosus microcephalus</i>	GU130677.1	<i>Synaphobranchus kaupii</i>	JQ774748
<i>Sparisoma cretense</i>	JQ623998	<i>Synchiropus phaeton</i>	JQ774556
	KC501533 - KC501552		JQ774749 - JQ774751
<i>Sparus aurata</i>	JQ623999	<i>Syngnathus acus</i>	KJ709646 - KJ709650
	KC501553 - KC501572	<i>Syngnathus rostellatus</i>	KJ205211 - KJ205221
	KJ709635	<i>Taurulus bubalis</i>	KJ205222
<i>Sphoeroides pachygaster</i>	EU869841 - EU869843		KJ205315
	KJ709636	<i>Tetrapturus belone</i>	KJ709929
	KJ709914 - KJ709918	<i>Thunnus alalunga</i>	JN007752.1 - JN007761.1
	KJ768311		JQ624006
<i>Sphyraena sphyraena</i>	JQ624000		KC501673 - KC501692
	KC501573 - KC501592		KJ709651
	KJ709637 - KJ709638	<i>Thunnus thynnus</i>	JQ624007
<i>Spicara maena</i>	JQ774738 - JQ774742		KC501693 - KC501712
	KJ709919		KJ709930
	KJ768312	<i>Torpedo marmorata</i>	JQ774752 - JQ774753
<i>Spicara smaris</i>	JQ624001		KJ709652
	KC501593 - KC501612		KJ768313
	KJ709639 - KJ709640		

**Tabela 10 (continuação) – Número de acesso do GenBank de todas as sequências da biblioteca de referências para peixes marinhos europeus. \*BOLD-ID**

<b>Identificação</b>	<b>Nº de acesso GenBank</b>	<b>Identificação</b>	<b>Nº de acesso GenBank</b>
<i>Torpedo nobiliana</i>	JQ774557		KJ205440 - KJ205443
<i>Torpedo torpedo</i>	MLFPI341-13*	<i>Trisopterus minutus</i>	JQ775159 - JQ775163
<i>Trachinotus ovatus</i>	JQ624009		KJ205245 - KJ205257
	KC501733 - KC501752		KJ205316 - KJ205321
	KJ768314		KJ205444
<i>Trachinus draco</i>	JQ774754	<i>Umbrina cirrosa</i>	JQ624013
	JQ774928 - JQ774932		KC501813 - KC501832
	KJ205435	<i>Upeneus moluccensis</i>	JQ624014
	KJ709931 - KJ709936		KC501833 - KC501852
<i>Trachinus radiatus</i>	KJ709937 - KJ709939	<i>Uranoscopus scaber</i>	KJ709673
<i>Trachurus mediterraneus</i>	JQ624010		KJ709950
	KC501753 - KC501772		KJ768320
	KJ709654 - KJ709657	<i>Vinciguerria attenuata</i>	KJ709674
	KJ768315 - KJ768316	<i>Vinciguerria poweriae</i>	KJ709675
		<i>Xenodermichthys copei</i>	JQ774563 - JQ774568
<i>Trachurus picturatus</i>	JQ774755 - JQ774759		
	JQ774933 - JQ775146	<i>Xiphias gladius</i>	JQ624016
	KJ709658 - KJ709659		KC501873 - KC501892
<i>Trachurus trachurus</i>	JQ624011		KJ709676
	JQ774760 - JQ774766		KJ709951
	JQ774938 - JQ775151	<i>Xyrichtys novacula</i>	KJ709677 - KJ709679
	KC501773 - KC501792	<i>Zenion hololepis</i>	JF718831 - JF718835
	KJ205223 - KJ709944		KJ768321
	KJ768317 - KJ768319	<i>Zenopsis conchifer</i>	KJ768322
<i>Trachyrincus scabrus</i>	JQ774558 - JQ774559	<i>Zeus faber</i>	EU869849 - EU869870
<i>Trachyrincus scabrus</i>	JQ774560 - JQ774562		JQ624017
<i>Trigla lyra</i>	JQ624012		KC501893 - KC501912
	JQ774767 - JQ774768		KJ205258 - KJ205262
	JQ774943 - JQ774946		KJ709680
	JQ775152 - JQ775153		KJ709952
	KC501793 - KC501812		KJ768323 - KJ768324
	KJ709661 - KJ709662		MLFPI359-14*
	KJ709945 - KJ709947		MLFPI360-14*
<i>Trigloporus lastoviza</i>	JQ774948 - JQ774951		MLFPI361-14*
	KJ205439		MLFPI362-14*
	KJ709664 - KJ709665	<i>Zoarces viviparus</i>	KJ205263
<i>Trigonognathus kabeyai</i>	GU130702.1		
<i>Trisopterus capelanus</i>	KJ709666 - KJ709672		
	KJ709948 - KJ709949		
<i>Trisopterus esmarkii</i>	KJ205233 - KJ205240		
<i>Trisopterus luscus</i>	JQ774952 - JQ774956		
	JQ775154 - JQ775158		
	KJ205241 - KJ205244		