UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA ANIMAL

Evolutionary history and genetic diversity of native cyprinids from the Portuguese West region: implications for conservation management

Ana Rita Pereira de Almeida

Mestrado em Ecologia e Gestão Ambiental

Dissertação orientada por:
Doutora Carla Sousa Santos
Prof. Doutora Maria Filomena de Magalhães

2016

Agradecimentos

Abstract

Freshwater fish populations face a decrease in genetic diversity at a global scale due to multiple, often cumulative, pressures, and the Portuguese Cyprinidae populations are not an exception. This situation is particularly serious in the West region, where the extant river basins are under strong human pressure and there is still a considerable gap in knowledge about the genetic structure of the populations of native cyprinid species, most of which are imperilled. Thus, the main goal of this thesis was to provide genetic and biogeographic data which can be useful for a more efficient conservation management of native cyprinid species in this region. The mitochondrial cytochrome *b* gene was used to assess the genetic diversity and estimate the divergence time between populations from West region and those from the neighbour river basins of Mondego, Tejo and Sado. The results suggest that cyprinids colonized the West region during the Holocene via two independent routes: 1) *Squalius carolitertii*, *Achondrostoma oligolepis*, *Luciobarbus bocagei* and *Pseudochomdrostoma polylepis* followed a southward route from Mondego to the West region; 2) *Squalius pyrenaicus*, *Iberochondrostoma lusitanicum*, *Luciobarbus bocagei* and *Pseudochomdrostoma polylepis* followed a West-Northwestward route from Tejo to the West region. Thus, *L. bocagei* and *P. polylepis* colonized both ends of the West region by distinct routes. Cyprinids from the West region display low intra-population variability and moderate to low levels of molecular diversity, which are lower than those found in neighbour river basins. Intrinsic drivers related to the life-history traits of the species may explain the different genetic variations found between sympatric species. On the contrary, the distinct levels of intra-population variability detected for each species could be explained by the different effect of extrinsic drivers due to the different geomorphological characteristics of each river basin. Thus, the observed patterns suggest the presence of accumulated differences within each population, reflecting their unique evolutionary histories as primary freshwater fish, which are intimately related to the evolution of paleobasins and geomorphological rearrangements of the West region. As such, all populations must be viewed as independent conservation units. Conservation plans should target *A. occidentale*. *I. lusitanicum*, *S. pyrenaicus* and *A. oligolepis* populations, by this order of priority, and ideally should include species-specific measures. Future management plans concerning the West region should include global measures such as preventing the discharge of pollutants, limiting water abstraction during the summer, restoring fluvial connectivity, eradicating invasive fauna and flora and habitat rehabilitation, while avoiding translocations at all cost.

Key words: Conservation genetics; Cyprinids; Portuguese West region; Mitochondrial cytochrome *b* gene; Evolutionary Significant Units.

Resumo alargado

A ictiofauna dulçaquícola é um dos grupos taxonómicos mais ameaçados à escala global, enfrentando um acentuado decréscimo de diversidade genética particularmente durante as últimas décadas. Esta problemática encontra-se principalmente associada à pressão antropogénica crescente sobre os sistemas dulciaquícolas, nomeadamente através de pressões como a poluição, a degradação e fragmentação de habitats, a proliferação de espécies invasoras, a captação de água e a regulação de caudais. O mesmo acontece com as comunidades ictiofaunísticas dos rios da Península Ibérica, onde 70-80% das espécies foram classificadas como vulneráveis, em perigo ou criticamente em perigo de extinção.

Os ciprinídeos são a família de peixes de água doce com maior distribuição a nível global, sendo também o grupo mais abundante na Península Ibérica. Sendo peixes primários, a sua dispersão só poderá ter ocorrido através de contactos ancestrais entre bacias hidrográficas atualmente independentes, como as capturas fluviais nas zonas de cabeceiras e a confluência de fozes, estando a sua história evolutiva estreitamente relacionada com a evolução paleogeomorfológica de cada região. A Península Ibérica foi colonizada por linhagens ancestrais de ciprinídeos durante o Oligocénico que conseguiram ultrapassar as barreiras naturais existentes antes da elevação da Cordilheira dos Pirinéus estar completa. As bacias endorreicas existentes nesta época, através de vários rearranjos geomorfológicos, originaram a atual recente rede hidrográfica e permitiram a dispersão de ciprinídeos por toda a Península. Atualmente, a Península Ibérica alberga uma grande variedade de espécies de ciprinídeos com pequenas áreas de distribuição e um grande número de espécies endémicas.

A região Oeste de Portugal inclui várias bacias hidrográficas independentes e de pequenas dimensões, confinadas pelas cadeias montanhosas Sicó-Aire-Candeeiros, Montejunto e Sintra e as bacias hidrográficas dos rios Mondego e Tejo. Devido à elevada ocupação humana, à industrialização e à atividade agrícola, o contexto desta região pode ser considerado como preocupante para as populações das espécies nativas de ciprinídeos. Como tal, e tendo em conta que alguns aspetos genéticos destas populações são desconhecidos, a região Oeste constitui um caso de estudo interessante no âmbito da genética aplicada à conservação da biodiversidade.

O objetivo principal desta tese era fornecer informação genética e biogeográfica útil para uma gestão e conservação mais eficientes das espécies nativas de ciprinídeos que ocorrem na região Oeste de Portugal. Para tal, estimou-se a diversidade genética de cada população e os tempos de divergência entre populações, e reconstituiram-se as histórias evolutivas das populações, relacionando-as com a paleogeomorfologia da região. Com base nestes elementos, teceram-se considerações acerca do que podem ser consideradas unidades de conservação e unidades evolutivas significativas (ESUs) e sugeriram-se futuras linhas de investigação e medidas mitigadoras no âmbito da conservação de peixes de água doce.

O estudo incidiu sobre as 7 espécies de ciprinídeos nativos que ocorrem na região do Oeste, designadamente, *Iberochondrostoma lusitanicum* (Collares-Pereira, 1980), *Achondrostoma occidentale* (Robalo, Almada, Sousa-Santos, Moreira e Doadrio 2005), *Achondrostoma oligolepis* (Robalo, Doadrio, Almada e Kottelat, 2005), *Squalius pyrenaicus* (Günther, 1868), *Squalius caroliterii* (Doadrio, 1988), *Pseudochondrostoma polylepis* (Steindachner, 1864) e *Luciobarbus bocagei* (Steindachner, 1864). O trabalho foi desenvolvido em 14 bacias hidrográficas, nomeadamente Lis, São Pedro, Alcoa, Tornada, Real, Alcabrichel, Sizandro, Safarujo, Lizandro, Colares, Lage, Ossos e Jamor. Para atingir os objetivos propostos, usou-se o gene mitocondrial citocromo *b* e, em alguns casos particulares, o gene nuclear da beta-actina, ambos amplificados a partir de ADN extraído da barbatana dorsal de aproximadamente 20 indivíduos de cada espécie e de cada bacia. Para fins comparativos, e de modo a reconstituir a colonização da região Oeste, foram ainda analisadas amostras das bacias hidrográficas vizinhas, dos rios

Mondego, Tejo e Sado. As amostras de barbatana utilizadas foram recolhidas através de pesca elétrica, em operações realizadas no âmbito do projeto FISHATLAS e desta dissertação.

As análises genéticas baseadas nos tempos de divergência obtidos para o gene to citocromo *b* sugerem que os ciprinídeos nativos colonizaram a região Oeste muito recentemente, durante o Holocénico, posteriormente às últimas grandes glaciações, através de conexões com as bacias vizinhas do Mondego e Tejo. De acordo com os padrões observados, a explicação mais plausível para a colonização desta região integra duas rotas de colonização distintas: 1) *S. carolitertii*, *A. oligolepis*, *L. bocagei* e *P. polylepis* chegaram à região através do Mondego, utilizando uma rota no sentido norte-sul e 2) *S. pyrenaicus*, *I. lusitanicum*, *L. bocagei* e *P. polylepis* entraram na região a partir do Tejo, utilizando uma rota no sentido este-oeste/noroeste. De acordo com este cenário, *L. bocagei* e *P. polylepis* colonizaram os extremos opostos da região Oeste por duas rotas distintas, através do Mondego e do Tejo.

As populações de ciprinídeos da região Oeste exibem baixa a moderada variabilidade intrapopulacional, inferior à encontrada nas populações das grandes bacias vizinhas do Mondego, Tejo e Sado. Esta baixa diversidade parece corroborar a recente colonização mas pode também ser o resultado de sucessivos efeitos gargalo causados por transgressões marinhas ou por eventos recentes, como as descargas de poluentes e as secas estivais cíclicas, que tendencialmente podem reduzir o tamanho das populações de cada espécie.

As variações observadas entre populações da mesma espécie em diferentes bacias sugerem que a diversidade genética pode ser condicionada por fatores extrínsecos à espécie, nomeadamente pelas condições ambientais a que as populações estão sujeitas e que podem variar consideravelmente entre bacias. Por outro lado, espécies que ocorrem nas mesmas bacias apresentam diferenças no que diz respeito à diversidade genética, o que parece apontar para a eventual existência de condicionalismos intrínsecos à espécie, nomeadamente o tempo de geração, o tamanho corporal e a capacidade migratória.

No geral, e apesar da colonização muito recente, os padrões genéticos observados evidenciam a acumulação de diferenças em cada população, reflexo do isolamento a que ficaram sujeitas e, provavelmente, da adaptação a condições ambientais específicas de cada bacia. De facto, por se tratar de peixes primários e por isso intolerantes à água salgada, a evolução das espécies estudadas é similar à das espécies confinadas a ilhas: a partir do momento em que uma bacia se torna independente, os peixes que nela ocorrem deixam de trocar genes com outras populações e passam a acumular mutações próprias, iniciando uma história evolutiva exclusiva.

Assim, para propósitos de gestão e conservação, as populações de ciprinídeos da região Oeste devem ser consideradas como unidades evolutivas significativas e geridas independentemente como unidades de conservação distintas, com medidas específicas de acordo com a espécie e a bacia em questão. No entanto, não sendo possível englobar todas as populações em planos de conservação e gestão, algumas devem ser consideradas prioritárias. Assim, considerando as análises genéticas obtidas para as sete espécies estudadas, sugere-se que deve ser dada elevada prioridade à conservação das populações de *A. occidentale,* uma vez que a espécie tem uma área de distribuição muito reduzida e estas populações terão resultado de um processo de vicariância antigo. Numa segunda linha de prioridade, devem ser consideradas as populações de *I. lusitanicum* e *S. pyrenaicus* devido aos seus elevados estatutos de conservação e por serem as únicas populações destas espécies que não ocorrem em grandes bacias, representando assim as orlas das respetivas radiações. Em terceiro lugar, devem ainda ser consideradas as populações de *A. oligolepis* pela sua elevada divergência em relação à fonte de colonização.

A utilização de translocações de indivíduos como medida de conservação é frequentemente controversa. De acordo com os resultados deste trabalho, as translocações devem ser evitadas a todo o custo

principalmente por implicarem a mistura de unidades evolutivas significativas diferentes e a consequente perda da integridade genética das populações. Alternativamente, o reforço de populações com indivíduos produzidos em cativeiro a partir de reprodutores selvagens provenientes da bacia em questão poderá ser uma boa ferramenta para aumentar o tamanho efetivo das populações e assim reduzir o seu risco de extinção. No entanto, esta medida só será plenamente eficaz se as condições ambientais dos rios forem melhoradas, pelo que a adoção de medidas globais de melhoria dos habitats fluviais e de mitigação de ameaças devem ser uma prioridade a curto prazo.

Mais especificamente, futuros projetos de gestão e conservação de ciprinídeos na região Oeste devem incluir a implementação de medidas globais como a prevenção de descargas de poluentes; a limitação de captações de água, principalmente durante o verão; a erradicação de espécies invasivas de fauna e flora, como por exemplo a cana *Arundo donax* e o lagostim-vermelho *Procambarus clarkii;* o restauro da conectividade fluvial, eliminando barreiras transversais; a reabilitação de habitats através de restauro de galerias ripícolas e preservação de vegetação aquática; e o aumento da quantidade de habitats-refúgios disponíveis durante o verão.

Estudos futuros deverão focar-se nos agentes ecológicos, biológicos e ambientais que influenciam a diversidade genética das populações de ciprinídeos nativos para uma melhor compreensão dos padrões salientados neste estudo. Para além disso, adquirir um maior conhecimento sobre a disponibilidade e conectividade de habitats adequados para cada espécie e a forma como estes podem variar face a futuros cenários climáticos seria fundamental para o delineamento de planos de gestão e conservação mais eficazes para as bacias hidrográficas da região Oeste.

Palavras-chave: Genética da conservação; Ciprinídeos; Região Oeste Portuguesa; Gene mitocondrial citocromo *b*; Unidades Evolutivas Significativas.

Table of Contents

Index of figures

Index of tables

1. <u>INTRODUCTION</u>

1.1 Genetics as a conservation tool

Conservation genetics encompasses the use of genetic theory and techniques to manage populations and eventually contribute to reduce the risk of extinction, mainly in threatened species (e.g. DeSalle and Amato, 2004; Frankham *et al.,* 2010). Without conservation genetics, an unappropriated management and allocation of resources can occur (Frankham, 2003, 2005). However, its application in the management of threatened species is still far from being common (Frankham, 2010).

The application of genetics is crucial to resolve taxonomic uncertainties, detect hybridization and gene flow, identify distinct populations and determine if the populations are genetically healthy (e.g. DeSalle and Amato, 2004; Frankham, 2010; Allendorf *et al.*, 2010; Frankham *et al.,* 2010). These applications provide information that allow the identification of relevant populations for conservation purposes and the definition of evolutionary significant units (ESUs), which are indispensable due to the current high number of vulnerable species (e.g. DeSalle and Amato, 2004; Allendorf *et al.*, 2010; Frankham *et al.,* 2010). Evolutionary significant units (ESUs) are populations that have a high priority for separate conservation (Frankham *et al.,* 2010) and, according to Moritz (1994), should be defined based on the historical population structure and mtDNA phylogeny, to achieve their long-term conservation.

The genetic characterization of populations and the determination of conservation units are particularly important in management projects that imply conservation efforts such as translocation, captive breeding, and reintroductions (Moritz, 1999; Weeks *et al.*, 2011). They provide valuable information to the prevention of outbreeding depression risks and the preservation of specific population characteristics that could reflect adaptation to different environmental conditions (Frankham *et al.,* 2010; Weeks *et al.*, 2011).

Freshwater fish are a good example of the need to use genetic data for conservation purposes. Fish populations are often highly fragmented, mainly due to habitat loss and damming, and are expected to vary among basins reflecting adaptations to the local environment (Frankham *et al.,* 2010). Thus, although challenging, the management of imperiled freshwater fish species may largely benefit from the data retrieved from conservation genetics approaches.

1.2 European and Iberian colonizations by Cyprinid fishes

The Cyprinidae is one of the most successful and widespread families of freshwater fish, with more than 367 genera and 3006 species distributed throughout Eurasia, North America and Africa (Nelson *et al.,* 2016).

Cyprinids are primary fish (Fenolio *et al.,* 2013; Nelson *et al.,* 2016) intolerant to marine salinity that are, thus, restricted to river basins and incapable of migrate between disconnected ones or through the sea (Darlington, 1957; Sousa-Santos *et al.,* 2016). Thereby, the evolutionary history of cyprinids is highly related to connections and disconnections between river basins associated with geomorphological rearrangements, which were responsible, respectively, for gene flow and isolation of populations (e.g. Aboim *et al.,* 2013; Sousa-Santos *et al.,* 2014a and Sousa-Santos *et al.,* 2014b). This has already prompted several studies combining genetic and geological dating (e.g. Marchordom and Doadrio, 2001; Durand *et al.,* 2002; Kotlik *et al.,* 2008; Houston *et al.,* 2010; Sousa-Santos *et al.,* 2014a, 2014b; Perea

et al., 2016) and highlights the need to use data on the geomorphological evolution to understand the evolution of cyprinids.

Phylogenetic studies revealed that cyprinids originated in Asia, where there is the highest number of species, and that most of the North American, African and European species seem to have derived from Asiatic lineages (Robalo *et al.*, 2007; Levy *et al.*, 2009).

Europe was only colonized by cyprinids in the Oligocene (Briggs, 1995), when the drought of the Turgai Sea allowed the connection with Asia (Levy *et al.*, 2009). Two main hypotheses are recurrently considered to explain the colonization and radiation of cyprinids across European rivers. One hypothesis is that Europe was gradually colonized as a result of river captures from the Oligocene until the Pliocene (35.0-1.7Ma), with fish reaching the Iberian Peninsula before the formation of the Pyrenees and the intensification of the alpine orogeny, in the Miocene (reviewed by Levy *et al.,* 2009). The other hypothesis is that the dispersal of cyprinids occurred through the Mediterranean sea during the Lago Mare phase, in the Messinian (5.5Ma), in the end of the salinity crises (Krijgsman *et al.*, 1999) when Parathetys, a central European lake, drained to the Mediterranean basin that was dried because of the isolation from the Atlantic and the decrease of rainfall (Bianco *et al.,* 1990). This hypothesis is supported by evidence indicating that some cyprinids, as *Luciobarbus,* have more affinity with North African species than with central European species (e.g. Tsigenopoulos *et al.*, 2003; Gante *et al.,* 2009). However, besides the Lago Mare phase, the Iberian Peninsula and Africa were connected in other episodes, and river captures can have occurred between them, as well as coastal diffusion caused by a decrease in the sea level (Levy *et al.,* 2009, Doadrio *et al.,* 2011). Both hypotheses can be complementary in explaining the colonization of the Iberian Peninsula by cyprinids (reviewed by Levy *et al.*, 2009 and by Perea *et al.,* 2010).

Currently, two distinct groups of cyprinids are found in Europe: one in Central Europe, with a low number of widespread species, and the other in the Mediterranean Peninsulas, with a high number of species of few genera, with small geographic ranges and high levels of endemism. One possible explanation for these two patterns is that, in the last glaciation, the meridional peninsulas have constituted a refuge for freshwater fishes, from where colonizers of only a few number of species radiated to Central Europe after the end of the glaciation (Banarescu, 1992; Taberlet *et al.,* 1998; Hewitt, 1999; Nesbø *et al.,* 1999; Robalo *et al.*, 2007 and Doadrio *et al.*, 2011).

The Iberian Peninsula is surrounded by the Mediterranean Sea, the Atlantic Ocean and the Pyrenees and this isolation results in a somewhat insular character of the Iberian freshwater fish fauna. Moreover, the Iberian Peninsula has been colonized only by a small number of fish lineages that were able to surpass the natural barriers existent in the Oligocene, before the complete elevation of the Pyrenean Mountains (Doadrio *et al.*, 2011). In fact, the oldest cyprinid fossil known in the Iberian Peninsula is from *Rutilus antiquus*, dated from the Upper Oligocene, and was found in the Ebro basin, in the eastern margin of the Peninsula (reviewed by Levy *et al.,* 2009).

Since the beginning of the colonization of the Iberian Peninsula, several modifications have occurred in its hydrographic network (Andeweg, 2002). In the Miocene, the Iberian hydrographic network was formed by a large number of inland lakes that did not drained to the sea (Andeweg, 2002; Pais *et al.*, 2012; *Figure 1-1*). These endorheic basins were captured by the fluvial network between the Late Miocene and the Pliocene (Casas-Sainz and De Vicent, 2009), with the formation of the Iberian Central Massif during the Miocene, which resulted in the separation of Northern and Southern river basins - (Aboim *et al.*, 2013).

Figure 1-1: The evolution of the Miocene inland lakes (adapted from Sousa-Santos, 2007)

Nonetheless, the system only began to be exorheic in the Neogene, as a result of the combination of the alpine orogeny activity, with the Iberian Peninsula tilting towards the Atlantic (Sousa-Santos *et al.*, 2007), and climate constrains (Casas-Sainz and De Vicent, 2009). Regressive erosion could also have been linked to the endorheic-exorheic transition (Casas-Sainz and De Vicent, 2009) with the formation of large fluvial valleys and several endorheic basins captures (Cunha and Martins, 2004). Thus, the current Iberian fluvial network is very recent, of Plio-Pleistocene origin (Andeweg, 2002).

The last glacial maximum (LGM) culminated between 23.000 and 18.000 years ago (Jones and Mann, 2003), in the Late Pleistocene. At this time, the polar front reached the Northern Portugal latitude (Dias *et al.,* 2000), most of the Portuguese continental shelf was above sea level, and the confluence of some rivers was possible, especially between the Minho and Ave basins and in the vicinity of the Nazaré submarine canyon (Rodrigues and Dias, 1989). Moreover, several submarine canyons formed in the Pliocene (such as the Nazaré and Cascais submarine canyons), as a consequence of glacio-eustatic variations (Alves *et al.*, 2003), could have acted as isolation barriers during marine regressions and, thus, could have promoted the differentiation of populations separated by them, as already suggested for lampreys (Mateus *et al.* 2013). Between 13.000 and 11.000 years ago, the sea level rose rapidly and flooded the fluvial valleys (Cunha and Gouveia, 2015; *Figure 1-2*), which consequently dug the headwaters of the rivers. The current sea level was only reached approximately 3.500 years ago, after the Flandrian transgression (Teixeira and Gonçalves, 1980; Dias *et al.,* 2000; Cunha and Gouveia, 2015; *Figure 1-2*).
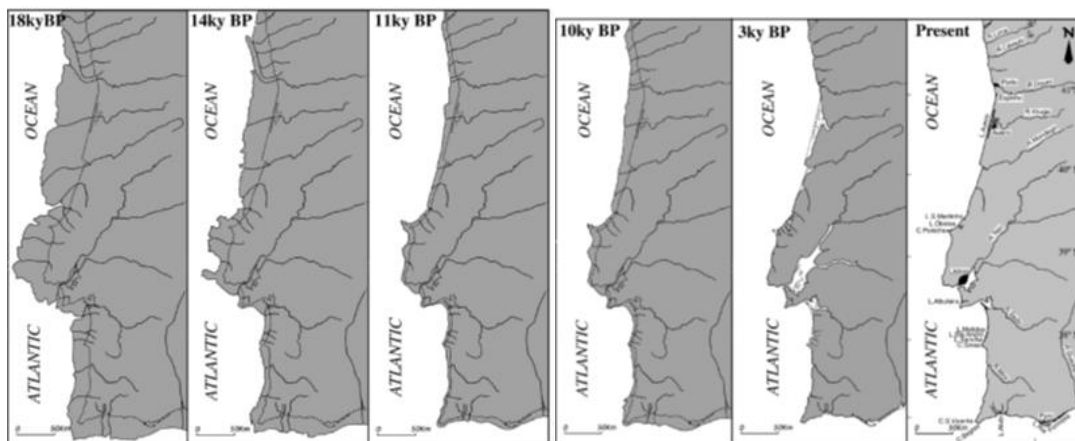
Figure 1-2: Portuguese coastal evolution since the LGM (Dias et al., 2000).

Thus, ancient freshwater fish lineages most likely colonized the different river basins in the Iberian Peninsula through connections between the inland lakes and, afterwards, through connections between the established river basins (e.g. Sousa-Santos *et al.,* 2007; Perea *et al.,* 2015). These connections could have resulted from fluvial captures of headwaters, confluence of river mouths during regressions and connections between previously existent coastal lagoons (*Figure 1-3*).

Figure 1-3: Coastal lagons of Pederneira, Alfeizerão and Óbidos at 2000 years B.C. (author: Joaquim Pereira da Silva; date:1982)

After each river basin became independent from the neighbour basins, primary fish populations also became isolated and evolved without further gene interchange with other populations.

1.3 Portuguese West region

The Portuguese West region is located between the Mondego and Tejo river basins and is considered the westernmost tip of the European continent. Nowadays, the West region is composed by approximately 20 small independent coastal river basins, some of them (from Lis to Real) previously connected by coastal lagoons (*Figure 1-3*). These river basins were colonized by cyprinids through connections with other river basins, following the processes above mentioned for the radiation of freshwater fish throughout the Iberian Peninsula.

The cyprinid species which occur in these region are relatively well known, including the phylogenetic relationships with close related species (for more detailed information see section 3). However, the species genetic structuration and their evolutionary history are still unclarified, namely concerning the colonization routes followed by the fish lineages and the eventual role of the three mountains ranges surrounding the Western rivers (Sicó-Aire-Candeeiros, Montejunto and Sintra) in the isolation and evolution of fish populations.

Considered the most populated region in the Portuguese Coast, the West region is highly industrialized. Thus, the environment is under severe anthropogenic pressure, particularly freshwater habitats.

Pollution and habitat degradation and fragmentation are the main consequences of this pressure, constituting major threats for the cyprinids and other freshwater fauna in the region.

Thus, since this region has several small river basins which are under high pressure and there is still a knowledge gap concerning some genetic aspects of the inhabitant cyprinid species, most of them imperiled, the West region configures an interesting geographical area to be studied from a conservation genetics perspective.

1.4 Objectives

The main goal of this thesis was to provide genetic and biogeographic data which can be useful for a more efficient conservation management of native cyprinid fish species from the Portuguese West region. To achieve this, the specific aims of this work were to:

1. Estimate the genetic diversity of cyprinid populations from the study area;
2. Estimate divergence times among cyprinid populations from the study area and between those populations and the ones from the neighbour river basins;
3. Reconstitute the evolutionary history of the cyprinid species from the study area and relate them with the paleogeomorphological evolution of the region;
4. Designate units of conservation and evolutionary significant units (ESUs) for native cyprinids and highlight future concerns for freshwater fish conservation.

2. STUDY AREA

The studied area is confined by two large river basins, the River Mondego in the north and the River Tejo in the south, and by three mountain ranges: Sicó-Aire-Candeeiros, Montejunto and Sintra (*Figure 2-1*). In administrative terms, this coastal area in central Portugal includes all municipalities between Leiria and Oeiras, from Leiria and Lisbon districts. To simplify, henceforth, the studied area will be referred as the West region.
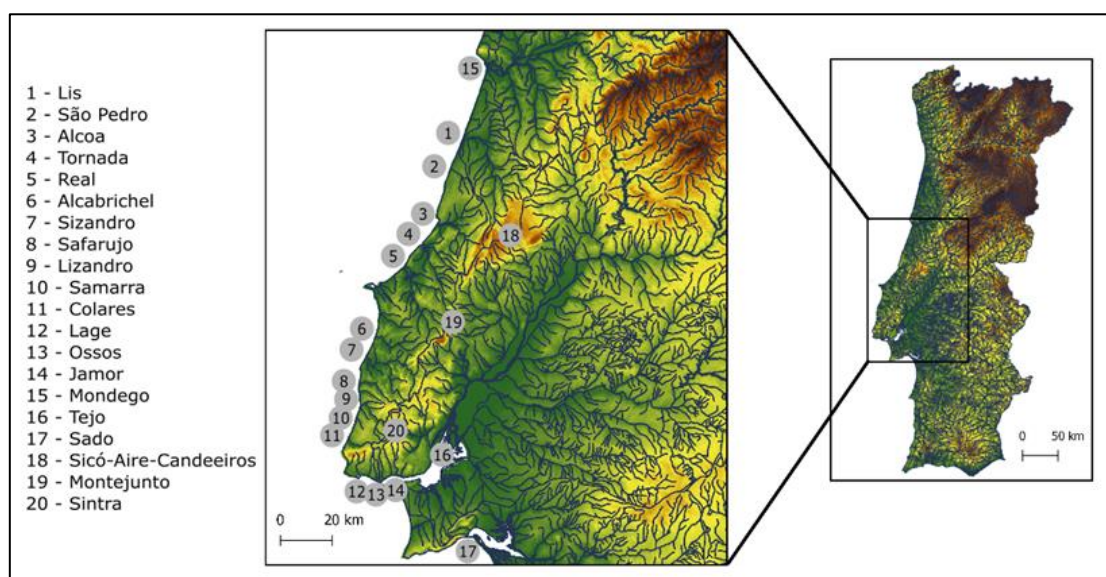
Figure 2-1: The Portuguese West region, including the river basins analyzed in this study (1-17) and the three mountain ranges of the region (18-20).

This is a region of extreme geological importance, mainly formed by limestone and with a complex network of subterranean karst caves (Cunha, 1990), which corresponds to the ancient Lusitanian basin formed in the beginning of the Alpine orogeny, during the Mesozoic (Kullberg *et al.*, 2013). Since that time, the region suffered many geological changes and the hydrographical network configuration changed accordingly (Pais *et al.*, 2012; for more details see section 1.2). Currently, the West region comprises approximately 20 small sized independent river basins, 14 of which still harbour native freshwater ichthyofauna (*Figure 2-1;* Ribeiro *et al.*, 2007).

The climate in the West region is Atlantic-humid (precipitation >2800 mm/yr; Feio *et al.*, 2009). However, in drier years, some of the streams are reduced to small disconnected pools that act as faunal refugia for fish (Magalhães *et al.*, 2002). The congregation of fish in these pools increase the predation, the competition for the limited space and food resources and mainly in the smallest pools, the fish are subject to hypoxia, hyperthermia and higher probability of being affected by infectious diseases (e.g. Magalhães *et al.*, 2002; Magoulick and Kobza, 2003; Dekar and Magoulick, 2007).

Currently, only the River São Pedro (*Figure 2-1*) is categorized as having good ecological quality, in the frame of the Water Framework Directive. The remainder of the rivers in this region have statuses inferior than good (Agência Portuguesa do Ambiente, I.P., 2016 available in: <http://sniamb.apambiente.pt/pgrh/>).

Freshwater fish assemblages in the West region include eight native species, including seven cyprinids (see section 3) and the cobitid *Cobitis paludica*. Besides that, at least five exotic species occur in the region: *Carassius auratus* in Lis, Tornada, Real and Lizandro; *Lepomis gibbosus* in Lis, S. Pedro and Lizandro; *Gambusia holbrooki* in Lis and Real; *Cyprinus carpio* in Lis, Lizandro and Colares and *Micropterus salmoides* in Lizandro (Ribeiro *et al.*, 2007).

3. STUDY SPECIES

This work focused in the seven cyprinid species occurring in the West region: *Iberochondrostoma lusitanicum* (Collares-Pereira, 1980), *Achondrostoma occidentale* (Robalo, Almada, Sousa-Santos, Moreira & Doadrio 2005), *Achondrostoma oligolepis* (Robalo, Doadrio, Almada & Kottelat, 2005), *Squalius pyrenaicus* (Günther, 1868), *Squalius carolitertii* (Doadrio, 1988), *Pseudochondrostoma polylepis* (Steindachner, 1864) and *Luciobarbus bocagei* (Steindachner, 1864). The distribution of each species across river basins in the West region is shown in *Table 3-1*.

Table 3-1: Cyprinid species occurring in the West region river basins.

	L. bocagei	*S. carolitertii*	*S. pyrenaicus*	*A. Oligolepis*	*A. occidentale*	*P. polylepis*	*I. lusitanicum*
Lis	X		X	X		X	
São Pedro				X			
Alcoa	X	X		X		X	
Tornada				X			
Real				X			
Alcabrichel					X		

Sizandro					X	
Safarujo					X	
Lizandro	X		X			X
Samarra			X			X
Colares	X		X		X	X
Lage			X			X
Ossos						X
Jamor			X			X

Three of the studied species - *I. lusitanicum*, *A. occidentale* and *A.oligolepis* - are Portuguese endemisms with very reduced distribution ranges, whereas he remaining species, *S. pyrenaicus*, *S. carolitertii*, *P. polylepis* and *L. bocagei,* are Iberian endemisms (Ribeiro *et al.,* 2007) and present larger distribution ranges (*Figure 3-2*).
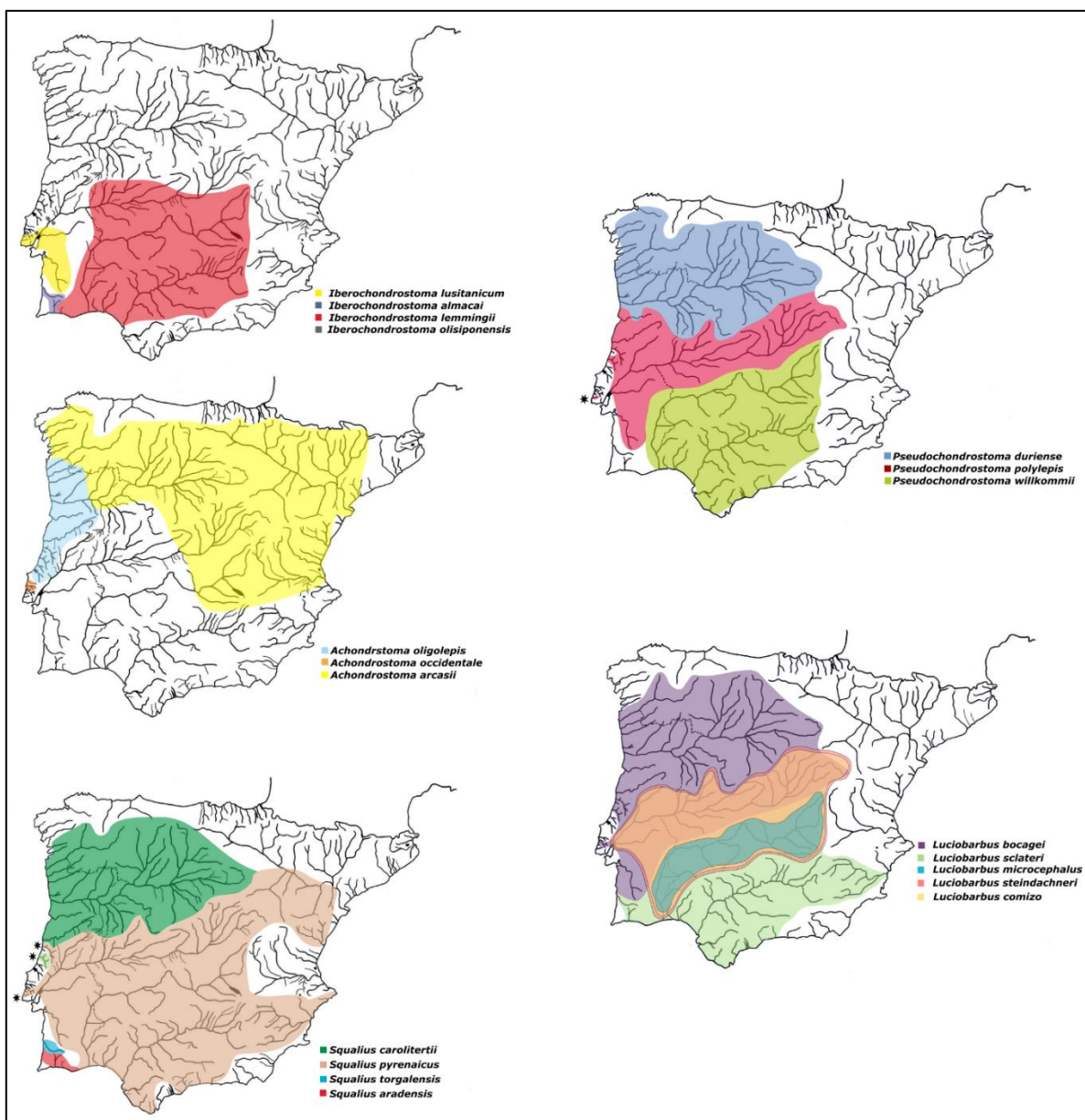
Figure 3-1: Iberian distribution of Portuguese Iberochondrostoma species (adapted from Robalo, 2007 and Gante et al., 2010); Achondrostoma species (adapted from Robalo et al., 2007); Squalius species (with the exception of the hybridogenetic complex Squalius alburnoides; adapted from Waap et al., 2011); Pseudochondrostoma species (adapted from Aboim et al., 2013) and Luciobarbus species (adapted from Gante et al., 2015). * occurrence locations newly found during this thesis.

Three out of these seven species are of conservation concern, with *I. lusitanicum* and *S. pyrenaicus* being listed, respectively, as Critically Endangered (CR) and Endangered (EN) in the Portuguese Red Data Book (Cabral *et al.* 2005), and *A. occidentale* listed as Endangered (EN) by the IUCN (Freyhof and Kottelat, 2008).

The study species vary considerably in body size and longevity. Maximum body length and longevity is found for *L. bocagei* (100cm and 11 years; Ribeiro *et al.,* 2007). In contrast, the Portuguese endemisms *A. occidentale*, *A. oligolepis* and *I. lusitanicum* show the smallest body lenghts (<15 cm) and the minimum longevity was described for *I. lusitanicum* (4 years; Ribeiro *et al.,* 2007). *P. polylepis* and both *Squalius* species are considered as medium size cyprinids (<40 and <26 cm, respectively).

All species spawn from April to June (Carmona and Doadrio, 2000; Robalo *et al.*, 2008) and at least *I. lusitanicum*, *A. occidentale* and *P. polylepis* spawn in groups, laying adhesive eggs (e.g. Doadrio *et al.*, 2011). *P. polylepis* and *L. bocagei* are potamodromous species that migrate in the beginning of the reproductive season to spawn upstream (Kottelat and Freyhof, 2007; Mateus *et al.*, 2008).

As mentioned above, the phylogenetic relationships of the species occurring in the West region with their congeners are known, as well as the estimated times of divergence between species (summarized in *Figure 3-3).*
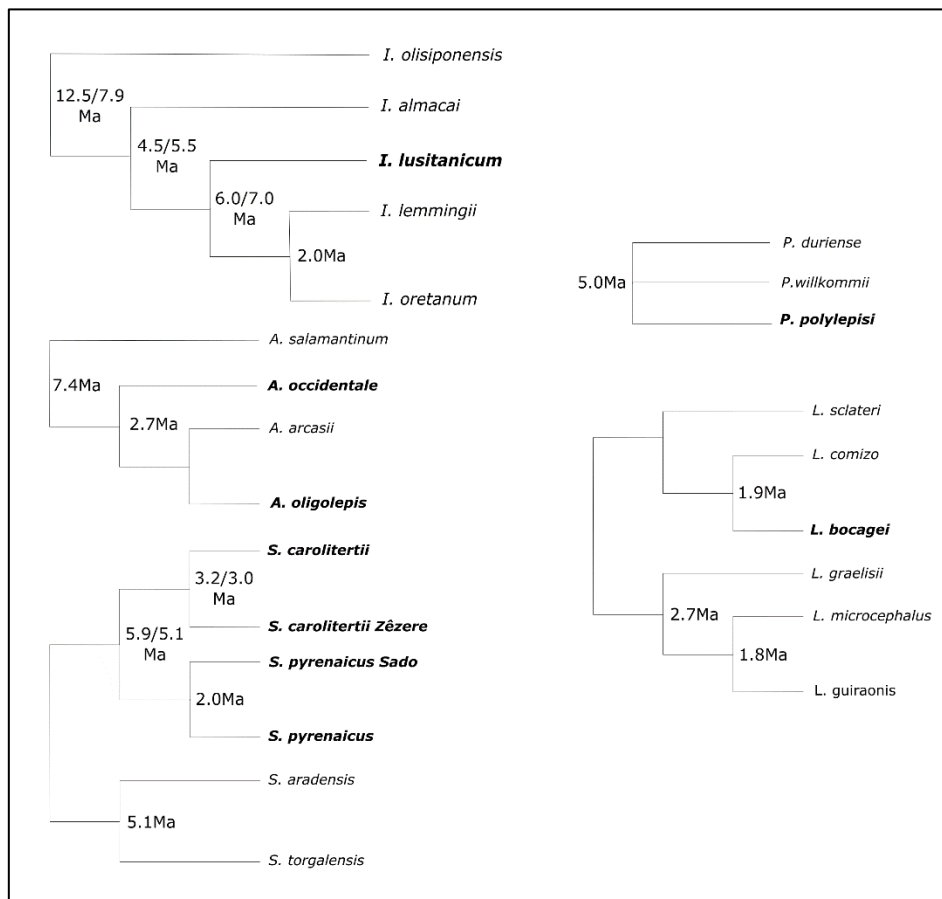
Figure 3-2: Diagrams summarizing the phylogenies of Iberian Iberochondrostoma species (adapted from Robalo, 2007 and Gante et al., 2010); Achondrostoma species (adapted from Robalo et al., 2006a); Squalius (adapted from Almada and Sousa-Santos, 2010); Pseudochondrostoma species (adapted from Doadrio and Carmona, 2004 and Aboim et al., 2013) and Luciobarbus species (adapted from Gante et al., 2009).

4. <u>METHODS</u>

4.1 Fish Sampling

Fish were sampled in the 14 river basins of the Western region harbouring cyprinids and in the three neighbour river basins (Mondego, Tejo and Sado) for comparison purposes. To simplify, and since all river basins are independent, throughout the text, the term "population" will be used to designate the individuals of the same species from the same river basin.

Most of the samples were previously obtained under the scope of the FISHATLAS project (FCT funded; PTDC/AAC-CLI/103110/2008) and were available at the MARE-ISPA research unit. However, in order to obtain a complete dataset of 20 individuals for each population (*Table 5-1,* section 5), additional sampling was conducted between September and December of 2015. In all cases, fish were sampled by electrofishing, using a SAMUS® portable device (500V impulses, with frequency and duration selected according to the water conductivity). Fish were maintained in aerated buckets and, after the removal of a dorsal fin clip (non-destructive sampling), they were safely returned to the river course.

4.2 DNA extraction, amplification and sequencing

DNA was extracted from fin clips (one per individual), previously preserved in ethanol, using the REDExtract-N-Amp Tissue PCR kit by Sigma-Aldrich®. Each fin clip was incubated with extraction and preparation solutions (100 µl and 25 µl, respectively) at 55°C for 10 minutes and subsequently at 95°C for 3 minutes. In the end of the cycle, a neutralization buffer (100 µl) was added.

One mitochondrial (cytochrome *b*) and one nuclear gene (beta-actin) were amplified by PCR (Polymerase Chain Reaction) using the primers LCB1-ACTTGAAGAACCACCGTTG (Sousa-Santos *et al.*, 2016) and HA–CAACGATCTCCGGTTTACAAGAC (Schmidt and Gold, 1993) for cyt*b* and BactFor-ATGGATGATGAAATTGCCGC and BactRev-AGGATCTTCATGAGGTAGTC (Robalo *et al.*, 2006b) for beta-actin. Tissue extracts (4 µl) were mixed with ultrapure water (4.4 µl), REDExtract-N-Amp PCR Reaction Mix (10 µl, Sigma-Aldrich®) and forward and reverse primers (0.8 µl each). PCR conditions were the following: 94ºC for 3 min + 35x[94ºC 1 min + 50ºC 1 min 30 s + 72ºC 1 min 30 s] + 72ºC 10 min for cyt*b*, and 94ºC for 3 min + 35x[94ºC 1 min + 55ºC 1min30 s + 72ºC 1 min 30 s] + 72ºC 10 min for beta-actin.

PCR products were purified and sequenced at GATC (Konstanz, Germany) using the forward primers LCB1 and BactFor. DNA sequencing was only conducted for the species which showed variation at the mitochondrial DNA level and, thus, for which the data obtained with the slower paced beta-actin gene could be informative.

4.3 DNA analysis

Obtained sequences were aligned and edited with CodonCode Aligner v.4.0.4 (Codoncode Corp., USA) and trimmed at the 3' and 5' ends so they had the same length. Sequences from heterozygous individuals for the beta-actin gene were manually phased using the procedures described by Sousa-Santos *et al.* (2005). Edited sequences were collapsed into haplotypes using DNA-colapser (FaBox v.1.41; available online at <<u>http://www.birc.au.dk/software/fabox</u>>).

The haplotype networks were computed using statistical parsimony based on a 95% confidence interval (TCS 1.21 software; Clement *et al.*, 2000). Haplotype classification, allowing the visualization of population correspondences and the relative haplotype frequencies as pie-chart like graphs, and layout

improvements were conducted using tcsBU (Santos *et al.*, 2015) and Inkscape v.0.48.4 softwares, respectively.

The ARLEQUIN software package v.3.5 (Excoffier and Lischer, 2010), henceforth designated as ARLEQUIN, was used to estimate molecular diversity indices (*h,* gene diversity; π, nucleotide diversity and *k*, mean number of pairwise differences), corrected population pairwise differences, and the number of migrants (M=Nm) between populations.

Genetic divergence between populations was calculated with the fixation index Φ_{ST} (Tamura-Nei model, 2000 permutations), implemented in ARLEQUIN. Estimates of the divergence time based on the cyt*b* gene were calculated using an evolutionary rate of 1.05% sequence divergence per million years (MY), as suggested by Dowling *et al.* (2002) for North American cyprinids. Molecular variance analyses (AMOVA) and mismatch distributions (Rogers and Harpending, 1992) were also computed in ARLEQUIN.

The parameters θ_0 and θ_1, M and τ, and and their confidence intervals were obtained by a parametric bootstrap approach using 10.000 replicates. Fu's Fs (Fu, 1997) and Tajima's D (Tajima, 1989) tests (based in 10.000 replicates) were performed using ARLEQUIN. For populations presenting deviations from mutation-drift equilibrium (i.e. showing negative and significant Tajima's D at $p < 0.05$; negative and significant Fu's Fs at $p < 0.02$; and non-significant Harpending's Raggedness p-values), it was determined the effective population sizes before (N_0) and after (N_1) a sudden expansion event, the effective population size prior to a spatial expansion event (N), the time since the expansion (t; in years), and the migration rate (M=2Nm), using the equations $\tau=2\mu t$ and $\theta=2N\mu$. To this end, we calculated the total mutation rate per sequence per generation (μ) considering the 1.05%/My divergence rate (0.0053×10^{-6} substitutions/lineage/year) of the cyt*b* gene estimated by Dowling *et al.* (2002), as before. According to the generation time known for each species, distinct mutations rates were calculated: 2.27×10^{-5} for *L. bocagei* ($5.25 \times 10^{-9} \times 720$ nucleotides x generation time of 6 years; Doadrio *et al.*, 2011); 7.56×10^{-6} for *A. oligolepis* and *A. occidentale* ($5.25 \times 10^{-9} \times 720$ nucleotides x generation time of 2 years; it was assumed that they have the same generation time as their congener *A. arcasii,* Doadrio *et al.*, 2011), 7.56×10^{-6} for *I. lusitanicum* ($5.25 \times 10^{-9} \times 720$ nucleotides x generation time of 2 years; it was assumed that it has the same generation as its congener *I. lemmingii*, Ribeiro *et al.*, 2007), 1.32×10^{-5} for *S. pyrenaicus* and *S. carolitertii* ($5.25 \times 10^{-9} \times 720$ nucleotides x generation time of 3.5 years; Rodrigues, 1999), and 1.32×10^{-5} for *P. polylepis* ($5.25 \times 10^{-9} \times 720$ nucleotides generation time of 3.5 years; Doadrio *et al.*, 2011). To evaluate the goodness-of-fit between the observed and the model frequency distributions, the sum of squared deviations (SSD) of the observed data related to the model and the Harpending's Raggedness indexes (and their respective *p* values based on 10.000 replicates) were calculated using ARLEQUIN.

A slower paced molecular marker (beta-actin gene) was used to search for shared haplotypes and infer more ancient genealogical relationships in *A. oligolepis, S. pyrenicus* and *S. carolitertii* species. Beta-actin sequences were not considered for *A. occidentale* populations since its phylogenetically closest relative, *Achondrostoma arcasii* (Robalo *et al.,* 2006a), is absent in the geographical vicinity of the study area, preventing the determination of the colonization routes followed towards the West region. Also, since *S. pyrenaicus* and *S. carolitertii* are not distinguishable species at the beta-actin gene level (Almada and Sousa-Santos, 2010), *Squalius* populations from the West region were compared with congeners from the river basins of Mondego, Tejo, and Sado.

5. RESULTS

The species composition of most of the river basins prospected during this thesis was already known and published (Ribeiro *et al.,* 2007, available in: <http://www.cartapiscicola.org/> and Sousa-Santos *et al.,* 2013, available in: <http://www.fishatlas.net/>). However, the occurrences of *S. pyrenaicus* in the Lis and Lage basins and *P. polylepis* in the Colares basin are described here for the first time (*Table 5-1*).

The complete dataset consisted of 2418 sequences, including 2043 sequences (720 bp) of the cyt*b* gene and 375 sequences of the beta-actin gene (935 and 905 bp for *Squalius* and *Achondrostoma oligolepis,* respectively) (*Table 5-1*). This included a set of sequences for the populations from the West region (561 cyt*b* and 135 beta-actin sequences) and another set for the populations from the Neighbour river basins (1482 cyt*b* and 240 beta-actin sequences). The total number of sequences used *per* population, for each molecular marker, is presented in *Table 5-1*.

Table 5-1: Number of DNA sequences by gene fragment and population.

	L. bocagei	*S. carolitertii*		*S. pyrenaicus*		*A. oligolepis*		*A. occidentale*	*P. polylepis*	*I. lusitanicum*	*Total*	
	Cyt*b*	Cyt*b*	Bact	Cyt*b*	Bact	Cyt*b*	Bact	Cyt*b*	Cyt*b*	Cyt*b*	Cyt*b*	Bact
Western basins												
Lis	20			20	6	20	12		19		79	18
São Pedro						20	9				20	9
Alcoa	20	20	13			20	6		20		80	19
Tornada						13	12				13	12
Real						20	4				20	4
Alcabrichel								23			23	
Sizandro								27			27	
Safarujo								21			21	
Lizandro	20			20	17					20	60	17
Samarra				20	20					19	39	20
Colares	19			21	7				16	19	75	7
Lage				20	10					23	43	10
Barcarena										21	21	
Jamor				20	19					20	40	19
Neighbour Basins												
Mondego	110	93	90			112	19		108		423	109
Tejo	200			150	87	21	8		192	158	721	95
Sado	158			83	36				20	77	338	36
Total	547	113	103	354	202	226	70	71	375	357	2043	375

For easy of comprehension, results will be presented independently for each species. The analyses will focus on the populations from the West region and will include comparisons with populations from Neighbour river basins. For simplicity, throughout the text, the terms "Western populations" and "Neighbour populations" will be used to designate, respectively, the populations from river basins in the West region and the populations from Mondego, Tejo and Sado basins.

5.1 Iberochondrostoma lusitanicum

The molecular diversity indices calculated for the Western populations of *Iberochondrostoma lusitanicum* and for the Neighbour populations from the Tejo and Sado basins are presented in *Table 5-2*. The mitochondrial cyt*b* dataset for *I. lusitanicum* Western populations only comprised one haplotype and, consequently, the molecular diversity indices were null. Contrastingly, in the Neighbour populations of the Tejo and Sado basins high values of molecular diversity and higher number of haplotypes were found.

Table 5-2: Molecular diversity indices for I. lusitanicum populations from the West region and Neighbour river basins. Sample size (N); Number of haplotypes (NH); Number (and percentage) of haplotypes shared with other populations from the West Region (SH); Gene diversity (h); Nucleotide diversity (π) and Mean number of pairwise differences (k) ± standard deviation (sd). The blank cells correspond to null values.

Population	N	NH	SH (%)	$h \pm sd$	$\pi \pm sd$	$k \pm sd$
Lizandro	20	1	1 (100%)			
Samarra	19	1	1 (100%)			
Colares	19	1	1 (100%)			
Lage	23	1	1 (100%)			
Ossos	21	1	1 (100%)			
Jamor	20	1	1 (100%)			
Tejo	158	12	2 (17%)	0.854±0.011	0.0031±0.0019	2.267±1.252
Sado	77	14	1 (7%)	0.675±0.051	0.0025±0.0016	1.777±1.040

The network of all *I. lusitanicum* haplotypes showed a reticulated pattern centred in the ancestral haplotype IL16 (*Figure 5-1*). The haplotype IL1, found in Western populations and distinct from the ancestral by three mutational steps, was the most frequent one, representing 40% of the individuals sampled, and was shared with individuals from the Tejo population. Contrastingly, most of the haplotypes found in the Sado population were distinct from the haplotypes found in the Western and Tejo populations by more than ten mutations.
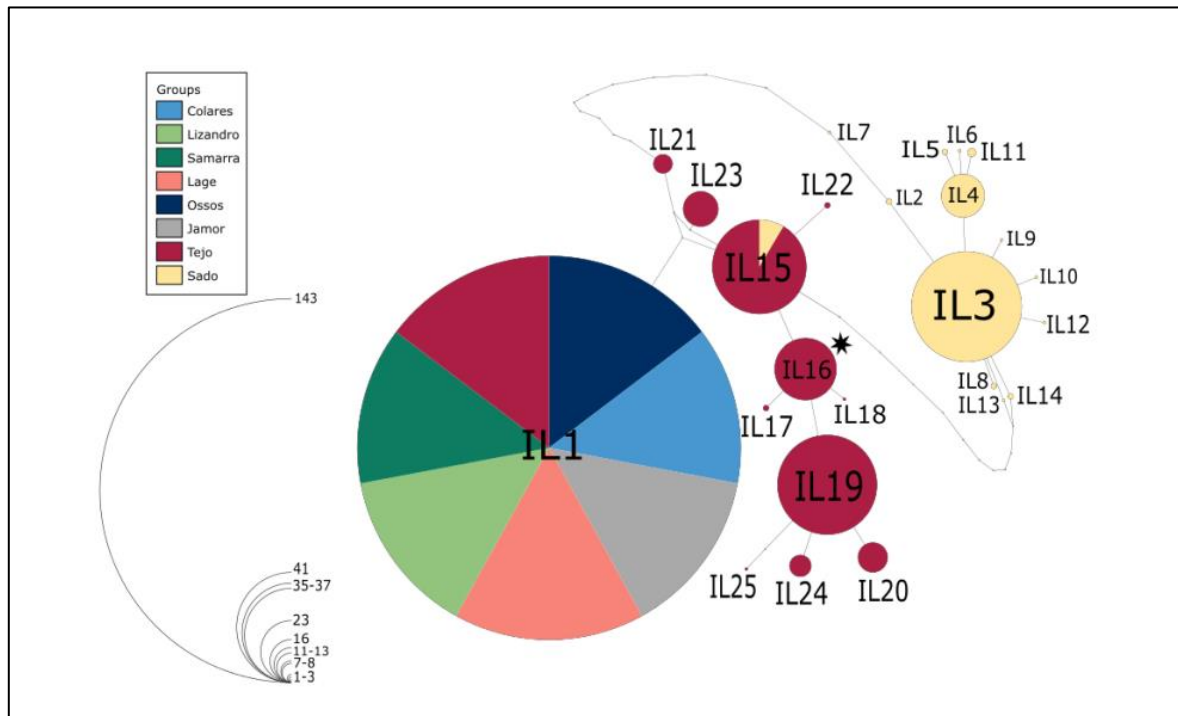
*Figure 5-1: Haplotype network of I. lusitanicum populations from the West region and Neighbour rivers basins of Tejo and Sado. *ancestral haplotype (outgroup weight=0.163)*

Considering Tejo, Sado and Western populations as three distinct populations, AMOVA results showed higher variation among populations (84.02%; Va=3.646; $p<0.001$) than within populations (15.98%; Vb=0.693; $p<0.001$). The overall fixation index was high and significant (FST=0.840, $p<0.001$), indicating that the *I. lusitanicum* populations were highly structured. This was also corroborated by the significant, corrected pairwise differences between Western populations and both Neighbour populations ($p<0.001$) (*Table 5-3*).

The percentages of divergence and estimated times of divergence between Western and Neighbour populations of the Tejo and Sado basins were 0.30% and 1.80%, and 2461 and 17274 years, respectively (*Table 5-3*). As expected, the lowest divergence times were obtained between the Western and Tejo populations, which are geographically closer. However, low gene flow (M<1) was detected between Western and Neighbour populations (*Table 9-16 section 9*).

Table 5-3: Relationships between I. lusitanicum populations from the West region and from the Neighbour populations of Tejo and Sado based on the average corrected pairwise distances between haplotypes. Above diagonal: corrected average pairwise differences; diagonal: average number of pairwise differences within populations; below diagonal: inter-population percentage of divergence between populations and respective estimated time of divergence based on the divergence rates calculated by Dowling et al. (2002) for the cytb gene. The blank cells correspond to null values.

	Lizandro	Samarra	Colares	Lage	Ossos	Jamor	Tejo	Sado
Lizandro							**1.860** **(p<0.001)**	**13.059** **(p<0.001)**
Samarra							**1.860** **(p<0.001)**	**13.059** **(p<0.001)**
Colares							**1.860** **(p<0.001)**	**13.059** **(p<0.001)**
Lage							**1.860** **(p<0.001)**	**13.059** **(p<0.001)**
Ossos							**1.860** **(p<0.001)**	**13.059** **(p<0.001)**
Jamor							**1.860** **(p<0.001)**	**13.059** **(p<0.001)**
Tejo	0.30% (2461y)	0.30% (2461y)	0.30% (2461y)	0.30% (2461y)	0.3% (2461y)	0.30% (2461y)		**11.445** **(p<0.001)**
Sado	1.80% (17274y)	1.80% (17274y)	1.80% (17274y)	1.80% (17274y)	1.80% (17274y)	1.80% (17274y)	1.60% (15138y)	

5.2 *Achondrostoma occidentale*

The mitochondrial cyt*b* dataset for Western populations of *Achondrostoma occidentale* comprised eight distinct haplotypes and showed moderate levels of genetic diversity: h=0.719±0.037, π=0.0023±0.0015 and k=1.640±0.979. The molecular diversity indices obtained for each population are presented in *Table 5-4*. The highest number of haplotypes (seven) and the highest values for the molecular diversity indices were found in the Alcabrichel population. By contrast, just one haplotype was found in the Sizandro population and, consequently, values of molecular diversity were null for all indices.

Population	N	NH	SH (%)	h±sd	π±sd	k±sd
Alcabrichel	23	7	1 (14%)	0.814±0.050	0.0020±0.0014	1.423±0.902
Safarujo	21	2	1 (50%)	0.095±0.084	0.0004±0.0005	0.286±0.316
Sizandro	27	1	1 (100%)			

The haplotype network of *A. occidentale* (*Figure 5-2*) showed a star-like pattern centred in the ancestral and the most frequent haplotype AOC2, was shared by individuals from the three Western populations. The remaining haplotypes, six private to the Alcabrichel population and one private to the Safarujo population, derived from the ancestor by one to five mutational steps.
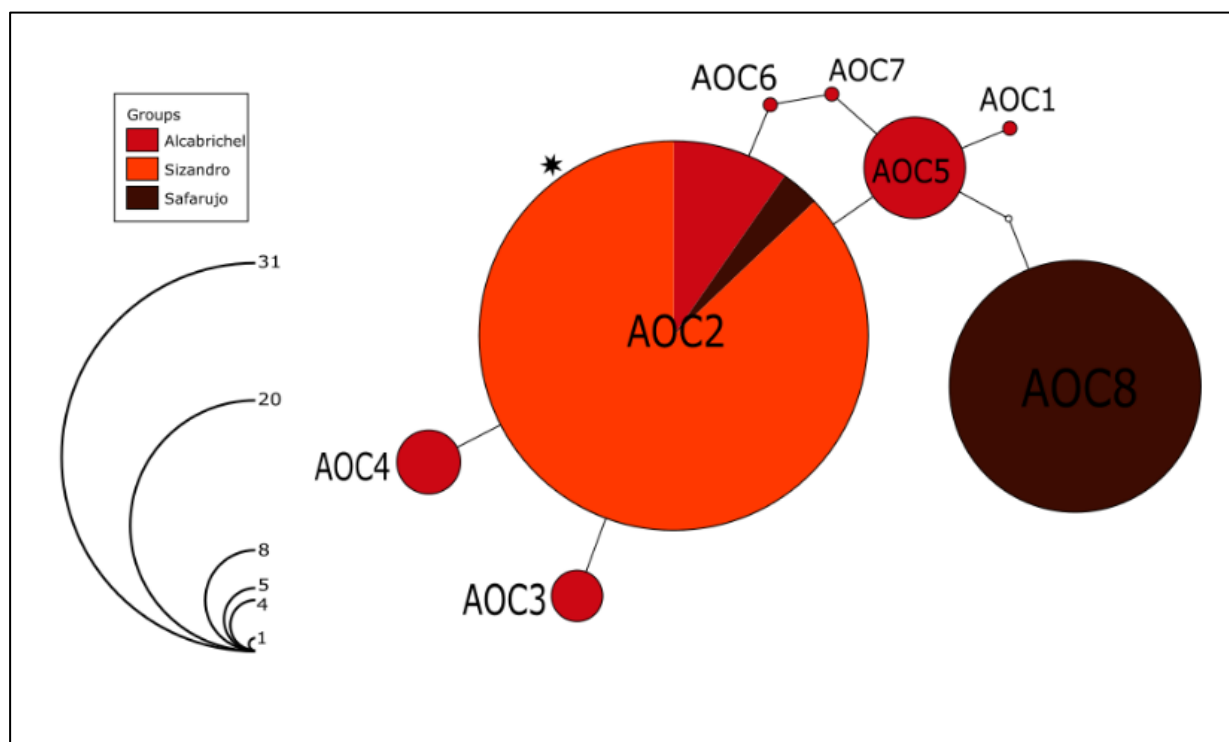
*Figure 5-2: Haplotype network of A. occidentale population from the West region. *ancestral haplotype (outgroup weight=0.331)*

AMOVA results showed that the majority of the variation (74.96%) occurred among populations (Va=0.815; p<0.001) and only 34.15%, within populations (Vb=0.272; p<0.001). The overall fixation index was high and significant (FST=0.750, p-value<0.001), indicating that *A. occidentale* populations were highly structured. This structuring was also corroborated by the significant, corrected pairwise differences (p<0.001) between all pairs of populations (*Table 5-5*).

Regarding divergence between populations, the Safarujo-Sizandro and Alcabrichel-Safarujo pairs showed, respectively, 0.40% and 0.30% of divergence, corresponding to divergence times of 3590 and 2819 years respectively (*Table 5-5*). The Sizandro and Safarujo were the closest populations, displaying the lowest percentage of divergence (0.03%) and the most recent estimated time of divergence (324 years) (*Table 5-5*).

Table 5-5: Relationships between A. occidentale populations from the West region based on the average corrected pairwise distances between haplotypes. Above diagonal: corrected average pairwise differences; diagonal: average number of pairwise differences within populations; below diagonal: inter-population percentage of divergence between populations and respective estimated time of divergence based on the divergence rates calculated by Dowling et al. (2002) for the cytb gene.

	Alcabrichel	Safarujo	Sizandro
Alcabrichel	1.423	**2.131** **(p<0.001)**	**0.245** **(p<0.001)**
Safarujo	0.30% (2819y)	0.286	**2.714** **(p<0.001)**
Sizandro	0.03% (324y)	0.40% (3590y)	0.000

No signature of expansion was detected for *A. occidentale* populations (Tajima's D $p>0.050$, Fu *Fs* $p>0.020$ and SSD $p<0.050$; *Table 9-10 section 9)*. The mismatch distributions were narrow and bimodal for Alcabrichel and unimodal for Safarujo (*Figure 9-1 section 9*), corroborating the pattern obtained in the haplotype network (*Figure 5-2*).

5.3 Achondrostoma oligolepis

The mitochondrial cyt*b* dataset obtained for the Western populations of *Achondrostoma oligolepis* comprised 17 distinct haplotypes and showed moderate levels of genetic diversity: $h=0.729\pm0.036$, $\pi= 0.0025\pm0.0016$ and $k=1.810\pm1.053$. The molecular diversity indices obtained for each population are presented in *Table 5-6*. Populations from the Alcoa and Tornada basins showed the lowest values for all molecular diversity indices, while the populations of the Lis, Real and São Pedro basins presented the highest ones Western populations showed three to seven haplotypes, most of them private, with the exception of the São Pedro population, with 100% of shared haplotypes.

When compared to the Western populations, the Neighbour population of the Mondego basin showed an higher number of haplotypes and an higher genetic diversity. Conversely, the population of the Tejo basin showed similar number of haplotypes and gene diversity to the Western population of Tornada basin, but π and k values were higher than those showed by any Western population (*Table 5-6*).

Table 5-6: Molecular diversity indices for A. oligolepis populations from the West region and Neighbour river basins. Sample size (N); Number of haplotypes (NH); Number (and percentage) of haplotypes shared with other populations from the West region (SH); Gene diversity (h); Nucleotide diversity (π) and Mean number of pairwise differences (k) ± standard deviation (sd).

Population	N	NH	SH (%)	*h±sd*	*π±sd*	*k±sd*
Lis	20	5	2 (40%)	0.679±0.080	0.0012±0.0010	0.832±0.619
São Pedro	20	3	3 (100%)	0.468±0.104	0.0014±0.0011	0.995±0.700
Alcoa	20	3	1 (33%)	0.195±0.114	0.0003±0.0004	0.200±0.259
Tornada	13	4	1 (25%)	0.423±0.164	0.0006±0.0007	0.461±0.432
Real	20	7	2 (29%)	0.521±0.135	0.0013±0.0011	0.968±0.687
Mondego	112	14	2 (14%)	0.781±0.056	0.0037±0.0023	2.676±1.484
Tejo	21	5	4 (80%)	0.385±0.068	0.0036±0.0021	1.556±0.937

The network of *A. oligolepis* haplotypes (N=33) showed a reticulated pattern centred in the ancestral haplotype, AOL14, which was shared by most individuals from Mondego, Tejo and São Pedro (*Figure 5-3*) basins. The most frequent haplotypes were AOL14 and AOL1, which were shared by individuals

from Alcoa, Real and Tornada basins, and represented 22% (N=49) and 19% (N=42) of the individuals sampled, respectively (*Figure 5-3*). From the ancestral haplotype (AOL14), 16 other haplotypes diverged by one to three mutational steps, most of them within the Mondego basin , with 12 private haplotypes and one shared with the Tejo population, and the remaining three within the Tejo population (AOL31-33; *Figure 5-3*). Besides this more ancestral clade, the haplotype network showed two other clades: 1) a star-like group centred in the second most common haplotype (AOL1), found in the Western populations of Tornada, Alcoa and Real basins, from which 10 haplotypes, nine of them private to a single Western population, diverged by one to three mutations; and 2) a terminal clade derived from the previously described one, which included five haplotypes, three private to the Lis population and the remaining two shared between Lis, São Pedro and Tejo populations. Thus, globally, these results showed that more ancestral haplotypes were found in some individuals outside of the Mondego basin, namely in the Tejo and São Pedro populations), and that a considerable diversification occurred in the West region, particularly in the Lis and São Pedro populations. Although showing some private haplotypes, the Tejo population harboured haplotypes shared with the Mondego and the Western populations of the Lis, São Pedro and Real basins.
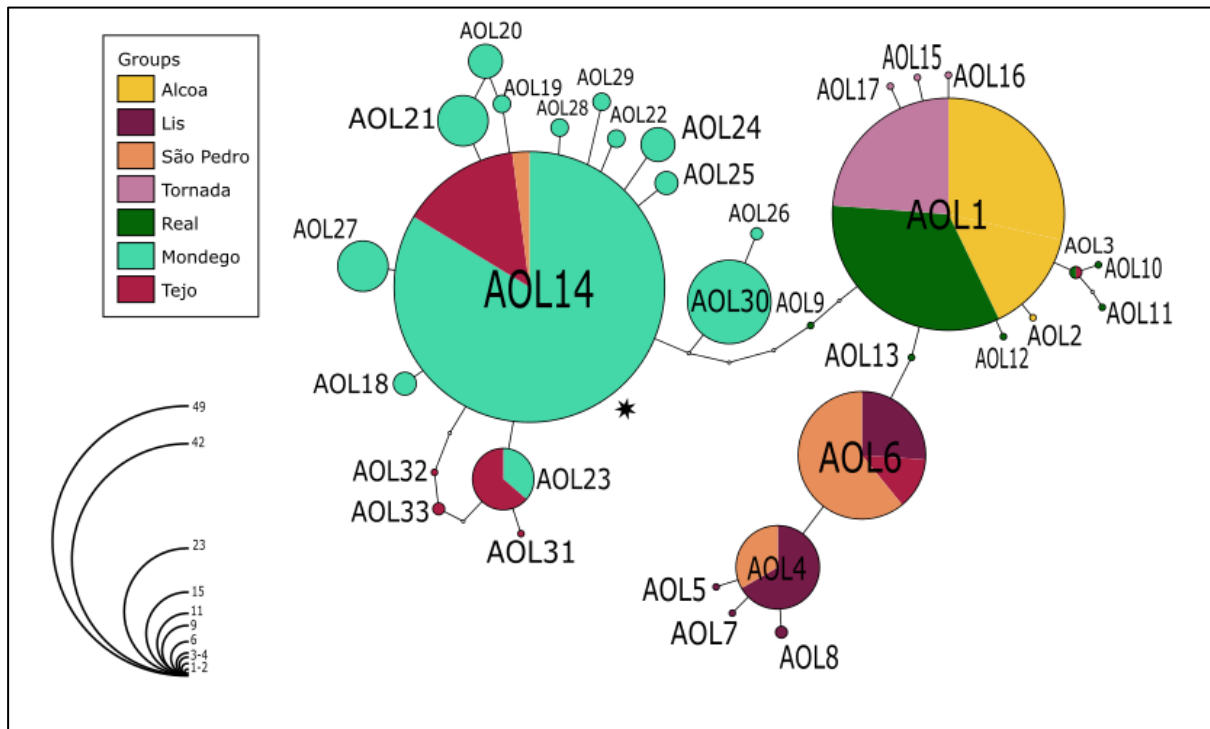
*Figure 5-3: Haplotype network of A. oligolepis populations from the West region and the Neighbour rivers basins of Mondego and Tejo. *ancestral haplotype (outgroup weight=0.175)*

AMOVA results for the Western populations showed that 65.85% of the variation occurred among populations (Va=0.684; p<0.001) and 34.15% within populations (Vb=0.355; p<0.001). The overall fixation index was high and significant (FST=0.658, p-value<0.001), indicating that populations of *A. oligolepis* from the West region were highly structured.

When assuming the Western, Tejo and Sado populations as distinct, AMOVA showed that 55.87% of the variation was explained among groups (FCT=0.559; *p*=0.048±0.002), 22.21% among populations within groups (FSC=0.503; *p*<0.001±0.000) and 21.92% within groups (FST=0.781; *p*<0.001±0.000), indicating that these populations were highly structured. This structuring was corroborated by the significant, corrected pairwise differences between most population pairs, with the exception of

populations from Alcoa, Real and Tornada basins ($p>0.050$). The percentage of divergence between Western populations was extremely low, ranging from 0% to 0.30%, which corresponds to recent estimated times of divergence (<3300 years ago; *Table 5-7*).

Concerning the divergence from the Neighbour populations, the Western and Tejo populations were closely related. Indeed, the divergence from the Tejo population occurred between 5493 and 6744 years ago (0.60%-0.70% of divergence) whereas the estimated time of divergence from the population in the Mondego basin was at 6868 to 8209 years ago (0.70-90% of divergence; *Table 5-7*).

Table 5-7: Relationships between A. oligolepis populations from the West region and the Neighbour river basins of Mondego and Tejo based on the average corrected pairwise distances between haplotypes. Above diagonal: corrected average pairwise differences; diagonal: average number of pairwise differences within populations; below diagonal: inter-population percentage of divergence between populations and respective estimated time of divergence based on the divergence rates calculated by Dowling et al. (2002) for the cytb gene.

	Lis	São Pedro	Alcoa	Tornada	Real	Mondego	Tejo
Lis	0.832	**0.187** **(*p*=0.010)**	**2.484** **(*p*<0.001)**	**2.484** **(*p*<0.001)**	**2.400** **(*p*<0.001)**	**6.206** **(*p*<0.001)**	**5.098** **(*p*<0.001)**
São Pedro	0.03% (247y)	0.995	**1.953** **(*p*<0.001)**	**1.953** **(*p*<0.001)**	**1.858** **(*p*<0.001)**	**5.192** **(*p*<0.001)**	**4.153** **(*p*<0.001)**
Alcoa	0.30% (3286y)	0.30% (2583y)	0.200	0.000 (*p*=1.000)	0.001 (*p*=0.530)	**5.776** **(*p*<0.001)**	**4.900** **(*p*<0.001)**
Tornada	0.30% (3286y)	0.30% (2583y)	0.00%	0.461	0.016 (*p*=0.270)	**5.771** **(*p*<0.001)**	**4.900** **(*p*<0.001)**
Real	0.30% (3175y)	0.30% (2458y)	0.00%	0.00%	0.968	**5.491** **(*p*<0.001)**	**4.637** **(*p*<0.001)**
Mondego	0.90% (8209y)	0.70% (6868y)	0.80% (7640y)	0.80% (7634y)	0.80% (7263y)	1.556	**0.323** **(*p*<0.001)**
Tejo	0.70% (6744y)	0.60% (5493y)	0.70% (6481y)	0.70% (6481y)	0.60% (6133y)	0.04% (427y)	2.678

Signatures of expansion were only detected for the *A. oligolepis* populations from Alcoa and Tornada basins (Tajima's D *p*<0.050, Fu *Fs* *p*<0.020 and SSD *p*>0.050; *Table 9-11 section 9*). The effective population sizes and time since the expansion estimated for Alcoa were lower than those of Tornada (0; 9767 years and 930; 36916 years respectively; *Table 9-21 section 9*). On the contrary, the migration rate was lower in the Tornada population (*Table 9-21 section 9*).

The mismatch distributions were narrow and unimodal for Lis, Alcoa, Tornada and Real populations and bimodal and wider only for the São Pedro population (*Figure 9-2 section 9*).

High gene flow was detected between the two Neighbour populations of the Tejo-Mondego basins (M>1; *Table 9-17 section 9*).

5.3.1 Achondrostoma oligolepis Nuclear DNA data

The network of *A. oligolepis* beta-actin haplotypes showed some diversification within the populations at the nuclear level, particularly within the Western populations, showing some divergence from the Neighbour populations. Indeed, from the seven private haplotypes found three are private to Western populations (two in Lis and one in Tornada) and from the remaining five haplotypes, four are shared

between Western populations and a single haplotype (AOL_B1) is shared by Neighbour and a Western population (Tornada; *Figure 5-4*). The topology of this network showed a star-like pattern centred in the ancestral haplotype AOL_B2, shared by individuals of the Alcoa, Lis, Tornada and Real populations. The most frequent haplotypes were the ancestral haplotype, which represented 44% (N=31) of the sampled individuals, and the haplotype AOL_B1, shared by individuals from Tornada, Mondego and Tejo basins, which represented 57% (N=40) of the individuals sampled. The three private haplotypes were found in the Western populations: AOL_B6 and AOL_B12 in Lis (derived from the ancestral haplotype AOL_B2 by, respectively, three and one mutational steps) and AOL_B7 in Tornada (also derived from the ancestral haplotype by 12 mutational steps).
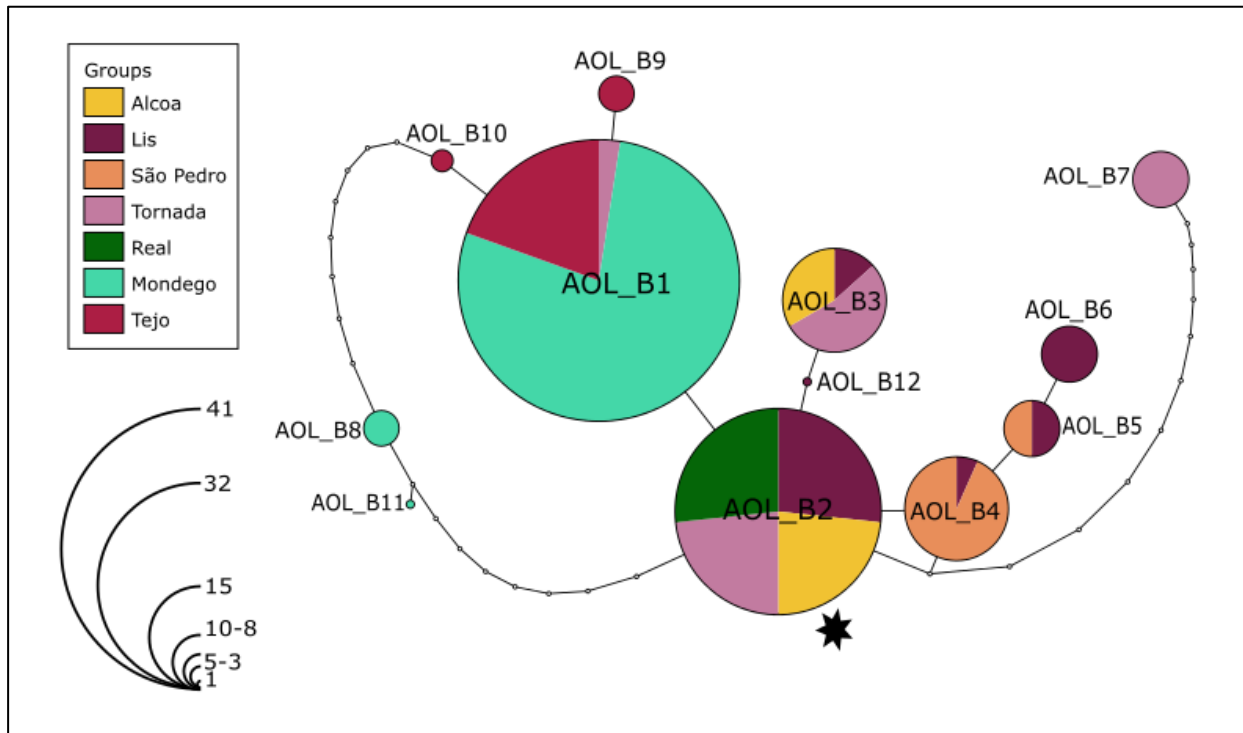
Figure 5-4: Haplotype network of A. oligolepis populations from the West region and Neighbour rivers basins of Mondego and Tejo. *ancestral haplotype (outgroup weight=0.239)

5.4 Squalius pyrenaicus

The mitochondrial cyt*b* dataset obtained for *Squalius pyrenaicus* populations from the West region comprised 15 distinct haplotypes and shows moderate levels of genetic diversity: h=0.676±0.039, π=0.0016±0.0011 and k=1.134±0.742. The molecular diversity indices obtained for each Western and Neighbour (Tejo and Sado) population are presented in *Table 5-8*. Between the Western populations, the highest number of haplotypes and also the highest values for all the molecular diversity indices were found in the Lizandro population (*Table 5-8*), while the lowest values were shown by the Samarra, Lage and Jamor populations (*Table 5-8*).

All haplotypes found in the Lage and Jamor populations were shared with other Western populations (*Table 5-8*). Contrastingly, the haplotypes found in the Samarra population were private (*Table 5-8*).

Both the genetic diversity indices and the number of haplotypes were generally lower in the Western than Neighbour populations (*Table 5-8*).

It is worth mentioning that it was expected that the *Squalius* individuals sampled in the Lis River were *S. carolitertii,* similar to what occurs in the geographically close Mondego and Alcoa basins. However, the obtained mtDNA sequences were from *S. pyrenaicus*, and this was further confirmed with the sequencing of the nuclear beta-actin gene (see section 5.7).

Table 5-8: Molecular diversity indices for S. pyrenaicus populations from the West region and the Neighbour river basins of Tejo and Sado. Sample size (N); Number of haplotypes (NH); Number (and percentage) of haplotypes shared with other populations from the West region (SH); Gene diversity (h); Nucleotide diversity (π) and Mean number of pairwise differences (k) ± standard deviation (sd).

Population	N	NH	SH (%)	h±sd	π±sd	k±sd
Lis	20	3	2 (67%)	0.195±0.114	0.0010±0.0008	0.689±0.546
Lizandro	20	7	3 (43%)	0.642±0.118	0.0022±0.0015	1.574±0.978
Samarra	20	2	0 (0%)	0.100±0.088	0.0001±0.0003	0.100±0.178
Colares	21	4	1 (25%)	0.471±0.116	0.0007±0.0007	0.514±0.451
Lage	20	2	2 (100%)	0.189±0.108	0.0003±0.0004	0.189±0.251
Jamor	20	2	2 (100%)	0.100±0.088	0.0008±0.0008	0.600±0.499
Tejo	150	41	4 (10%)	0.943±0.009	0.0056±0.0031	4.028±2.023
Sado	83	11	3 (27%)	0.385±0.068	0.0036±0.0021	2.585±1.400

The network of haplotypes showed two distinct clades: one smaller clade grouping most of the haplotypes found in the Sado basin and the other, larger and more diversified, grouping the remaining haplotypes (*Figure 5-5*). Regarding the later clade, the network was highly branched and showed a reticulate pattern, centred in the ancestral and most common haplotype (SP1), which was shared by the Colares, Jamor, Lizandro, Lage, Sado and Tejo populations. From the ancestral haplotype, 54 haplotypes diverge: three shared and 51 private From the 15 haplotypes found in Western populations, only four were shared with the Neighbour populations (SP1, SP6, SP7 and SP9) while ten were private to a given population and one (SP5) was shared by individuals from Lizandro and Jamor populations.
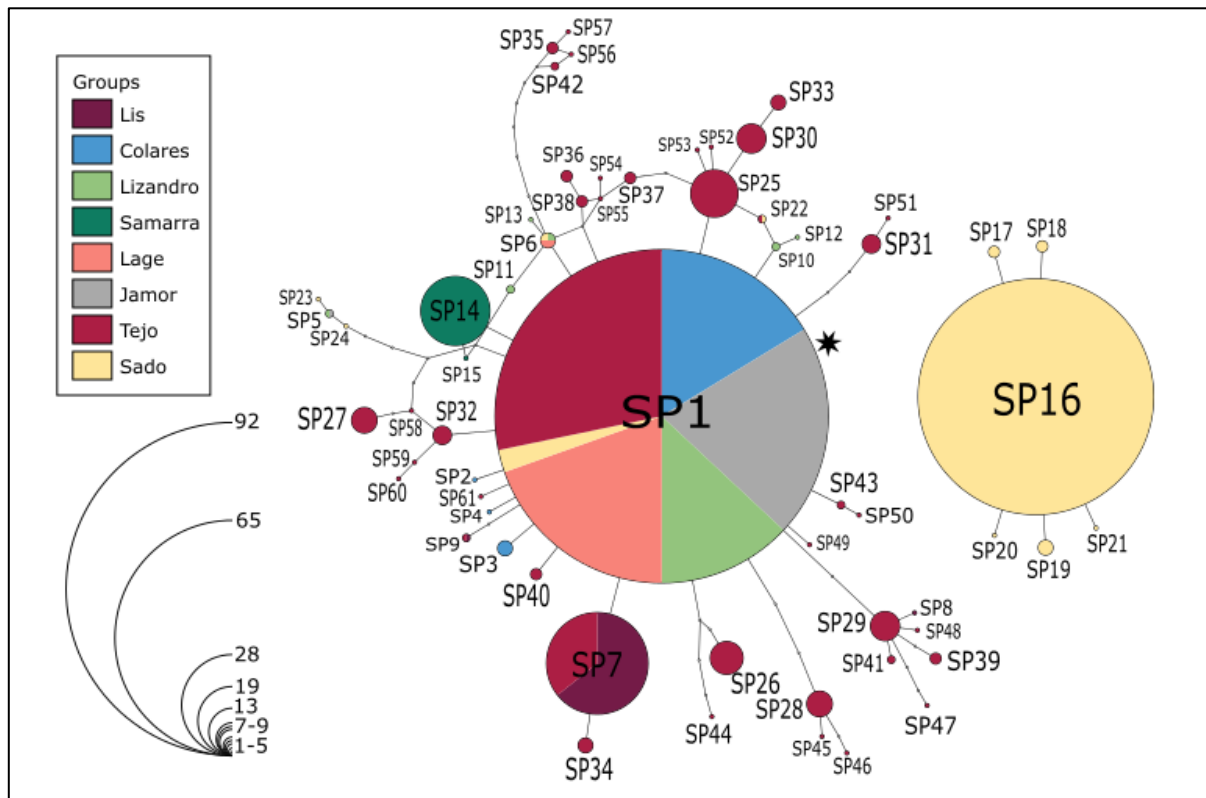
AMOVA results showed that the variation among and within *S. pyrenaicus* populations from the West region was similar (50.55% and 49.45%, respectively; Va=0.312 and Vb=0.305; p<0.001). The significance of the overall fixation index (FST=0.505, p-value<0.001) reflected the genetic structuring of the Western populations.

When assuming the West, Tejo and Sado as distinct populations, AMOVA results showed that 71.00% of the variation was explained among groups (FCT=0.710; *p*=0.248±0.004), 4.97% among populations within groups (FSC=0.171; *p*<0.001±0.000) and 24.03 within groups (FST=0.760 *p*<0.001±0.000), indicating that the populations were highly structured.

Average corrected pairwise differences were not significant between Lage, Jamor and Lizandro, and between Colares, Lage and Jamor populations (p>0.050), and the very low percentages of divergence indicate recent divergence times between these populations (*Table 5-9*). For the remaining population pairs, the differences were significant (p<0.050). Among pairs of Western populations, the percentage of divergence varied between 0.01% and 0.25%, corresponding to estimate times of divergence of 121 to 2388 years, respectively. The Western populations diverged more recently from the Tejo population, as expected due to its geographical proximity. The Sado population displayed a percentage of 1.90-2.20% of divergence from the Western populations, while the Tejo population only diverged 0.02% from Lizandro, Colares, Lage and Jamor populations and 0.10% from Lis and Samarra populations.

Table 5-9: Relationships between S. pyrenaicus populations from the West region and the Neighbour populations (Tejo and Sado) based on the average corrected pairwise distances between haplotypes. Above diagonal: corrected average pairwise differences; diagonal: average number of pairwise differences within populations; below diagonal: inter-population percentage of divergence between populations and respective estimated time of divergence based on the divergence rates calculated by Dowling et al. (2002) for the cytb gene.

	Lis	Lizandro	Samarra	Colares	Lage	Jamor	Tejo	Sado
Lis	0.689	0.868 (p<0.001)	1.805 (p<0.001)	0.834 (p<0.001)	0.810 (p<0.001)	0.805 (p<0.001)	0.768 (p<0.001)	15.516 (p<0.001)
Lizandro	0.12% (1149y)	1.574	1.048 (p<0.001)	0.092 (p<0.001)	0.028 (p=0.109)	0.023 (p=0.245)	0.156 (*p*=0.004)	13.995 (p<0.001)
Samarra	0.25% (2388y)	0.15% (1386y)	0.100	1.029 (p<0.001)	1.005 (p<0.001)	0.995 (p<0.001)	1.152 (p<0.001)	15.709 (p<0.001)
Colares	0.12% (1103y)	0.01% (121y)	0.14% (1361y)	0.514	0.034 (p=0.050)	0.029 (p=0.073)	0.181 (*p*=0.002)	14.833 (p<0.001)
Lage	0.11% (1072y)	0.00%	0.14% (1330y)	0.00%	0.189	0.005 (p=0.245)	0.140 (*p*=0.008)	14.621 (p<0.001)
Jamor	0.11% (1065y)	0.00%	0.14% (1316y)	0.00%	0.00%	0.600	0.138 (*p*=0.007)	14.512 (p<0.001)
Tejo	0.10% (1016y)	0.02% (207y)	0.10% (1524y)	0.02% (239y)	0.02% (185y)	0.02% (182y)	4.028	13.520 (p<0.001)
Sado	2.10% (20524y)	1.90% (18512y)	2.20% (20779y)	2.10% (19620y)	2.00% (19340y)	2.00% (19196)	1.90% (18281y)	2.585

No signature of expansion was detected for the Western populations (Tajima's D *p*>0.050, Fu *Fs* *p*>0.020 and SSD *p*<0.050; *Table 9-12 section 9*) and mismatch distributions were narrow and unimodal for Samarra, Lage and Colares populations and bimodal and wider for Lis, Lizandro and Jamor populations (*Figure 9-3 section 9*), reflecting the higher degree of differentiation of the latter.

High gene flow was detected between the Western populations and the Tejo population (M values varied between 2.624 and 32.103; *Table 9-18 section 9*).

5.5 Squalius carolitertii

The molecular diversity indices obtained for the only *Squalius carolitertii* population from the West region were much lower than those obtained for the Neighbour population from the Mondego basin (*Table 5-10*).

Table 5-10: Molecular diversity indices for S. carolitertii population from the West region and the Neighbour river basin of Mondego. Sample size (N); Number of haplotypes (NH); Number (and percentage) of haplotypes shared with the populations from the West region and with the remaining populations from the Neighbour rivers basins (SH); Gene diversity (h); Nucleotide diversity (π) and Mean number of pairwise differences (k) ± standard deviation (sd).

Population	N	NH	SH (%)	h±sd	π±sd	k±sd
Alcoa	20	2	1 (50%)	0.100±0.088	0.0001±0.0003	0.100±0.177
Mondego	93	14	1 (7%)	0.718±0.047	0.0025±0.0016	1.772±1.036

The haplotype network showed a star-like pattern centred on the ancestral and most frequent haplotype SC1 (shared by individuals from Alcoa and Mondego basins), from which the remaining 13 haplotypes were derived (*Figure 5-6*). Among these, 12 were private to the Mondego and one (SC2) was private to the Alcoa, showing that a recent diversification process must have occurred after the colonization of this river basin. Despite most of the haplotypes differed from the ancestor (SC1) by one mutational step, the presence of a branch of seven haplotypes private to the Mondego population which diverged from the ancestor by one (SC5) to nine mutations (SC4 and SC11) reflected a clear diversification process occurring within the Mondego basin.
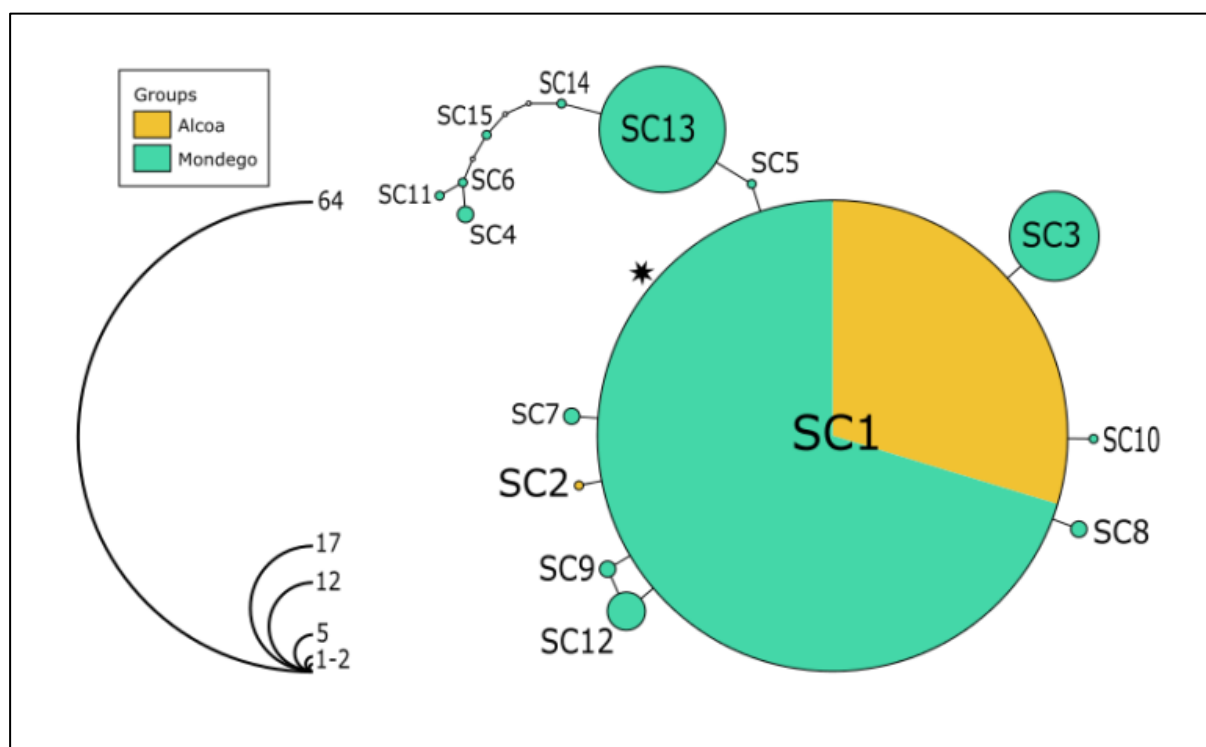
*Figure 5-6: Haplotype network of S. carolitertii populations from the West region and from the Neighbour Mondego river basin. *ancestral haplotype (outgroup weight=0.246).*

AMOVA results showed that most of the variation (93.56%) was explained by variation within populations (Vb=0.743; p=0.016±0.001) and only 6.44% was explained by variation among populations (Va=0.051; p=0.016±0.001). The significant but low overall fixation index (FST=0.064, p=0.016±0.001) indicated, as expected, a low degree of genetic structuring between both populations. However, due to the high intra variability of the Mondego population, a significant corrected pairwise difference was detected between the two populations (*Table 5-11*).

The estimated low percentage of divergence between populations suggested an extremely recent separation event (*Table 5-11*), which was corroborated by a high gene flow (M=7.268).

Table 5-11: Relationships between S. carolitertiii populations from the West region and from the Neighbour populations of Mondego based on the average corrected pairwise distances between haplotypes. Above diagonal: corrected average pairwise differences; diagonal: average number of pairwise differences within populations; below diagonal: inter-population percentage of divergence between populations and respective estimated time of divergence based on the divergence rates calculated by Dowling et al. (2002) for the cytb gene.

	Alcoa	Mondego
Alcoa	0.100	**0.135 (p=0.018)**
Mondego	0.02% (179y)	1.772

No signature of expansion was detected for the Alcoa population (Tajima's D *p*>0.050, Fu *Fs p*>0.020 and SSD *p*<0.050; *Table 9-13 section 9)* and the width of the mismatch distribution was very narrow, with a single pairwise difference (*Figure 9-4 section 9*).

5.6 Squalius carolitertii and Squalius pyrenaicus nuclear DNA data

Although it is not possible to clearly distinguish *S. carolitertii* and *S. pyrenaicus* using the slow paced beta-actin gene (Almada and Sousa-Santos, 2010), the network of 21 haplotypes obtained showed some differentiation within this genus at the nuclear level. Indeed, most of the sampled haplotypes (N=13; 62%) were private to a single population, either from the West region (S3 in Jamor) or from the Neighbour river basins (S8, S10, S11 and S12 in Sado, S14, S15, S17 and S18 in Tejo and S19, S20 and S21 in Mondego) (*Figure 5-7*). The remaining seven haplotypes were shared by two to nine populations without an evident geographically-related pattern (*Figure 5-7*).

A single haplotype (S5) was found in Alcoa population, corresponding to the southern distribution limit of *S. carolitertii,* and was shared by individuals from Mondego, Lis and Tejo populations. In the Lis population, corresponding to the northern distribution limit of *S. pyrenaicus,* three haplotypes were found (S2, S4 and S5), all of which were shared by individuals from others populations (*Figure 5-7).*

*Figure 5-7: Haplotype network of Squalius populations from the West region and Neighbour rivers basins of Mondego, Tejo and Sado. *ancestral haplotype (outgroup weight=0.178)*

5.7 Pseudochondrostoma polylepis

The mitochondrial dataset for *Pseudochondrostoma polylepis* from the West region comprised five distinct haplotypes and moderate overall levels of genetic diversity: h=0.516±0.070, π=0.0010±0.0008 and k=0.710±0.542. The molecular diversity indices obtained for Western and Neighbour populations are presented in *Table 5-12*. Concerning the Western populations, Lis and Colares showed a higher number of haplotypes and higher values for all the molecular indices than those of Alcoa, (in which a single haplotype was found,) and of the Neighbour populations of Mondego and Sado (*Table 5-12*). The Lis population was the only Western population showing private haplotypes, while in the Alcoa and Colares 100% of the haplotypes were shared (*Table 5-12*). Overall, the highest values of genetic diversity and the highest number of total and private haplotypes were found in the Neighbour population of Tejo (*Table 5-12*).

Table 5-12: Molecular diversity indices obtained for P. polylepis populations from the West region and the Neighbour rivers basins of Mondego, Tejo and Sado. Sample size (N); Number of haplotypes (NH); Number (and percentage) of haplotypes shared with other populations from the West region (SH); Gene diversity (h); Nucleotide diversity (π) and Mean number of pairwise differences (k) ± standard deviation (sd). The blank cells correspond to null values.

Population	N	NH	SH (%)	h±sd	π±sd	k±sd
Lis	19	4	2 (50%)	0.380±0.134	0.0006±0.0006	0.409±0.393
Alcoa	20	1	1 (100%)			
Colares	16	3	3 (100%)	0.633±0.074	0.0010±0.0009	0.733±0.574
Mondego	108	3	1 (33%)	0.140±0.044	0.0002±0.0003	0.141±0.207
Tejo	192	19	4 (21%)	0.737±0.030	0.0019±0.0013	1.378±0.853
Sado	20	3	1 (33%)	0.279±0.123	0.0004±0.00050	0.289±0.320

The network for all haplotypes found in the Western and Neighbour populations showed two common haplotypes: PP2, the ancestral haplotype, shared by 167 (45%) individuals from Lis, Colares and Tejo populations; and PP1, distinct from the PP2 by a single mutation, which was shared by 101 (27%) individuals from Lis, Alcoa, Colares, Mondego and Tejo populations. Besides these two frequent haplotypes, only haplotype PP3 was shared by individuals from Western and Neighbour populations (Colares and Tejo) and haplotype PP8 was shared between the two Neighbour populations (Tejo and Sado; *Figure 5-8*). The remaining 25 haplotypes were private of a specific population: 20 haplotypes were private to Tejo, 1 to Lis, 2 to Mondego and 2 to Sado populations.

*Figure 5-8: Haplotype network of P. polylepis populations from the West region and from the Neighbour rivers basins of Mondego, Tejo and Sado. *ancestral haplotype (outgroup weight=0.274)*

AMOVA results concerning the Western populations showed that 59.89% of the variation can be explained among populations (Va=0.684; p<0.001) and 40.11% within populations (Vb=0.355; p<0.001). The significance of the overall fixation index (FST=0.599, p-value<0.001) reflected the genetic structuring of the Western populations.

When considering the Western, Mondego, Tejo and Sado as distinct populations, AMOVA results show that 21.56% of the variation is explained among groups (FCT=0.216; p=0.452±0.004), 29.81% among populations within groups (FSC=0.380; p<0.001±0.000) and 48.63% within groups (FST=0.514; p<0.001±0.000), indicating that the populations are not highly structured.

Corrected pairwise differences were significant (p<0.05) between all pair of populations except between Lis and Alcoa, the most geographically close populations, and between Alcoa and Mondego (*Table 5-13*). The divergence estimated for the significantly different populations was very low (0-0.10%) between Western populations, corresponding to recent divergence times (under 1200 years ago; *Table 5-13*). Regarding the divergence between Western and Neighbour populations, the calculated

percentages of divergence showed that separation events were extremely recent and that Lis and Alcoa were more closely related to Mondego (0.00% of divergence), while Colares was closer to Tejo (0.01%). As expected, higher divergence times were estimated between the Western and Sado population, the most geographically distant Neighbour population (0.30% to 0.40%, corresponding to estimated times of divergence between 2416 and 3975 years).

Table 5-13: Relationships between P. polylepis populations from the West region and the Neighbour populations of Mondego, Tejo and Sado based on the average corrected pairwise distances between haplotypes. Above diagonal: corrected average pairwise differences; diagonal: average number of pairwise differences within populations; below diagonal: inter-population percentage of divergence between populations and respective estimated time of divergence based on the divergence rates calculated by Dowling et al. (2002) for the cytb gene.

	Lis	Alcoa	Colares	Mondego	Tejo	Sado
Lis	0.409	0.006 (*p*=0.236)	**0.797** (***p*<0.001**)	**0.009** (***p*=0.027**)	**0.645** (***p*<0.001**)	**2.906** (***p*<0.001**)
Alcoa	0.00%	0.000	**0.883** (***p*<0.001**)	0.004 (*p*=0.364)	**0.728** (***p*<0.001**)	**3.005** (***p*<0.001**)
Colares	0.10% (1107y)	0.10% (1168y)	0.733	**0.887** (***p*<0.001**)	**0.101** (***p*<0.001**)	**2.199** (***p*<0.001**)
Mondego	0.00% (12y)	0.00%	0.10% (1173y)	0.141	**0.731** (***p*<0.001**)	**3.009** (***p*<0.001**)
Tejo	0.09% (854y)	0.10% (963y)	0.01% (133y)	0.10% (967y)	1.378	**1.827** (***p*<0.001**)
Sado	0.40% (3844y)	0.40% (3975y)	0.30% (2829y)	0.40% (3980y)	0.20% (2416y)	0.289

No signature of expansion was detected for Western populations (Tajima's D p>0.050, Fu Fs p>0.020 and SSD p<0.050; *Table 9-14 section 9*) and mismatch distribution were narrow and unimodal for Lis and Colares populations (*Figure 9-5 section 9*).

High gene flow (M>1) was detected for Lis-Mondego, Lis-Tejo, Alcoa-Mondego and Colares-Tejo pairs (*Table 9-19 section 9*), corroborating the above mentioned extremely low divergences between these populations. The highest gene flow was estimated to have occurred between Alcoa and Mondego populations (*Table 9-19 section 9*).

5.8 *Luciobarbus bocagei*

The mitochondrial cyt*b* dataset for *Luciobarbus bocagei* from the West region comprised only two haplotypes and moderate overall levels of genetic diversity: h=0.468±0.031, π=0.0007±0.000636 and k=0.468±0.413. The molecular diversity indices obtained for each population are presented in *Table 5-14*. Due to the low number of haplotypes, the values of genetic diversity are lower for Western than the Neighbour populations (except for the h index, which is higher in Lis than in the Mondego) (*Table 5-14*). No private haplotypes were detected in the Western populations (100% of shared haplotypes), contrasting with the 22% to 67% of shared haplotypes obtained for the Neighbour populations (*Table 5-14*).

Population	N	NH	SH (%)	*h±sd*	*π±sd*	*k±sd*
Lis	20	2	2 (100%)	0.521±0.042	0.0007±0.0007	0.521±0.456
Alcoa	20	1	1 (100%)			
Colares	19	1	1 (100%)			
Lizandro	20	1	1 (100%)			
Mondego	110	3	2 (67%)	0.513±0.012	0.0007±0.0007	0.522±0.441
Tejo	200	9	2 (22%)	0.598±0.038	0.0011±0.0009	0.782±0.572
Sado	158	2	1 (50%)	0.222±0.040	0.0009±0.0008	0.668±0.516

The global network of haplotypes showed a star-like pattern, centred on the ancestral and most frequent haplotype LB2, from which the remaining 10 haplotypes were derived, most of them by a single mutation (*Figure 5-9*). Both haplotypes found in the populations of the West region were shared with the Neighbour populations: LB1was shared with individuals from Tejo and Mondego populations and LB2 with individuals from Tejo, Mondego and Sado populations. These haplotypes were also the most common, representing 67% (LB2; N=369) and 17% (LB1; N=94) of the individuals sampled, (*Figure 5-9*). The remaining haplotypes were private to Tejo (LB5 and LB11) and to Sado (LB3 and LB4), pointing to diversification processes occurring within these river basins (*Figure 5-9*).
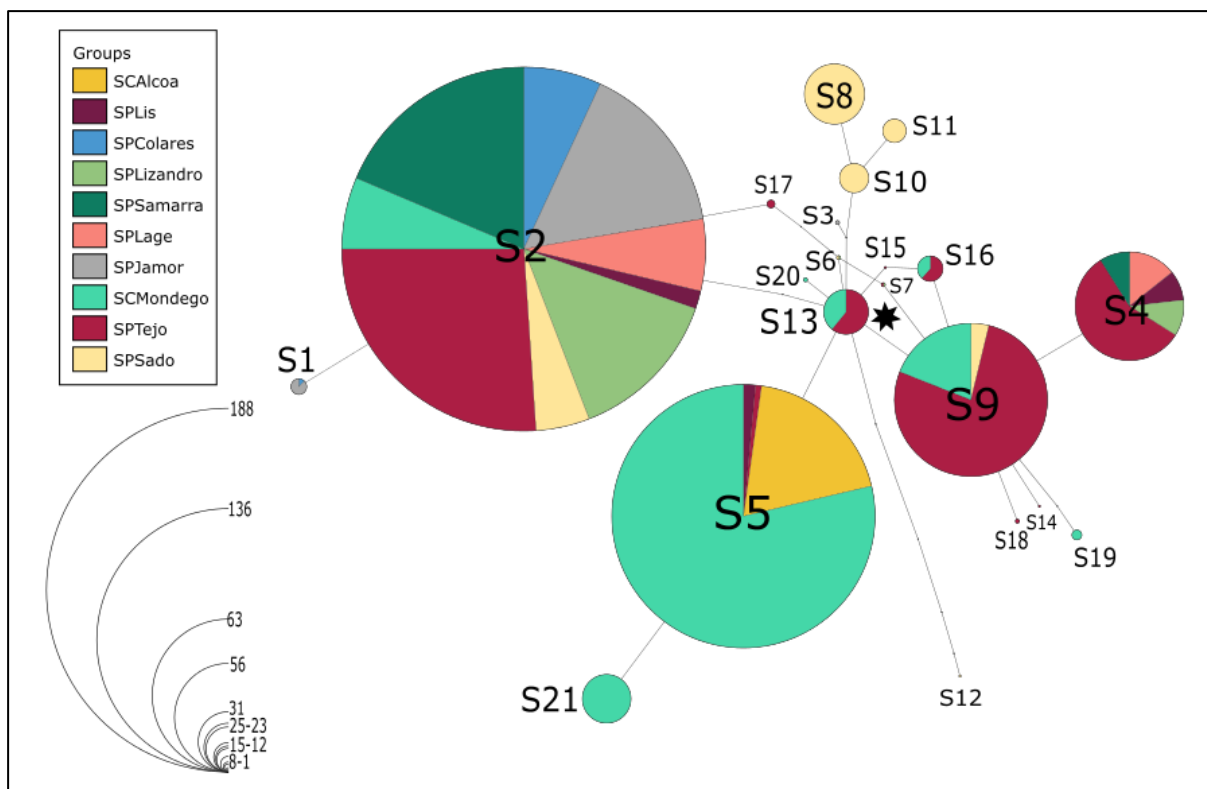

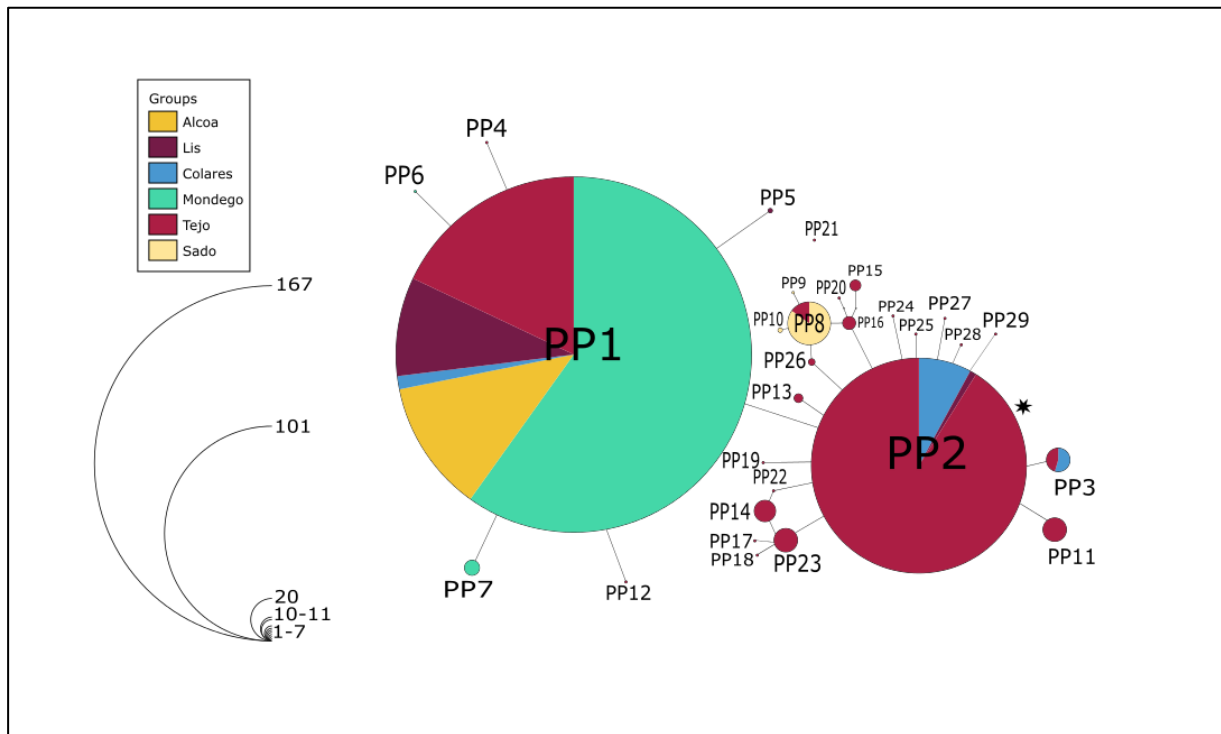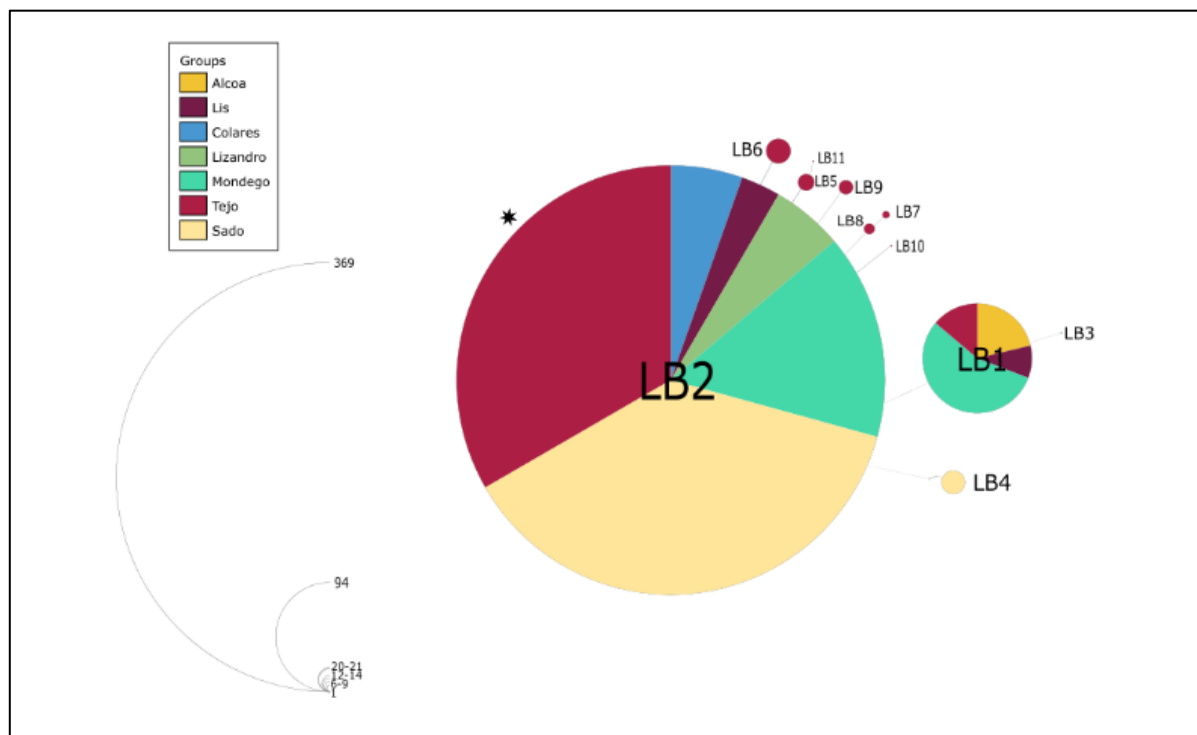*Figure 5-9: Haplotype network of L. bocagei populations from the West region and the Neighbour river basins of Mondego, Tejo and Sado. *ancestral haplotype (outgroup weight=0.292)*

AMOVA results concerning the Western populations showed that variation was higher among (77.35%; Va=0.222, p<0.001) than within populations (22.65%; Vb=0.065, p<0.001). The overall fixation index

was high and significant (F_{ST}=0.773, *p*<0.001), indicating that although they share the same two haplotypes, Western populations are structured.

When considering the West, Mondego, Tejo and Sado populations as disctint, AMOVA results showed that only 17.99% of the variation is explained among populations (Va=0.071) and most of the variation (82.01%) was explained within populations (Vb=0.326). The overall fixation index was low but significant (F_{ST}=0.180, *p*<0.001), pointing to the existence of genetic structuring.

Corrected pairwise differences were significant (p<0.050) between all pairs of populations except Lizandro-Colares, Mondego-Lis, Tejo-Lizandro, Sado-Colares and Sado-Lizandro (*Table 5-15*). The divergence values between Western populations were very low (between 0.00% and 0.14%), corresponding to extremely recent divergence times between populations (under 1500 years; *Table 5-15*). As expected, geographically close populations were the ones that diverged more recently: Alcoa and Lis in the northernmost part of the West region, and Lizandro and Colares in the southernmost part. The divergence between the Western populations and those of the Neighbour river basins was also very recent: Lis and Alcoa diverged from the Mondego less than 352 years ago, and Colares and Lizandro populations were virtually identical to those from the Tejo and Sado basins.

Table 5-15: Relationships between L. bocagei populations from the West region and from Neighbour river basins of Mondego, Tejo and Sado based on the average corrected pairwise distances between haplotypes. Above diagonal: corrected average pairwise differences; diagonal: average number of pairwise differences within populations; below diagonal: inter-population percentage of divergence between populations and respective estimated time of divergence based on the divergence rates calculated by Dowling et al. (2002) for the cytb gene.

	Lis	Alcoa	Colares	Lizandro	Mondego	Tejo	Sado
Lis	0.521	**0.289** **(p<0.001)**	**0.189** **(p<0.001)**	**0.189** **(p<0.001)**	-0.014 (p=0.659)	**0.160** **(p<0.001)**	**0.235** **(p<0.001)**
Alcoa	0.04% (383y)	0.000	**1.000** **(p<0.001)**	**1.000** **(p<0.001)**	**0.266** **(p<0.001)**	**0.899** **(p<0.001)**	**1.046** **(p<0.001)**
Colares	0.14% (1323y)	0.03% (251y)	0.000	0 (p=1.000)	**0.230** **(p<0.001)**	**0.029** **(p=0.036)**	0.046 (p=0.082)
Lizandro	0.14% (1323y)	0.03% (251y)	0.00%	0.000	**0.230** **(p<0.001)**	0.029 (p=0.055)	0.046 (p=0.100)
Mondego	0.04% (352y)	0.00%	0.03% (304y)	0.03% (319y)	0.522	**0.1963** **(p<0.001)**	**0.276** **(p<0.001)**
Tejo	0.12% (1189y)	0.02% (212y)	0.00% (38y)	0.00%	0.03% (260y)	0.782	**0.075** **(p<0.001)**
Sado	0.14% (1384y)	0.03% (311y)	0.00%	0.00%	0.04% (365y)	0.01% (99y)	0.668

Regarding the Lis population, the mismatch distribution was unimodal (*Figure 9-6 section 9*) and no signal of expansion was detected (Tajima's D=1.531 *p*>0.050, Fu *Fs*=1.467 *p*>0.020 and SSD<0.050; *Table 9-15 section 9*). For the remaining Western populations, mismatch analyses were not conducted due to the presence of a single haplotype.

High gene flow (M>1) was detected for all Western and Neighbour population pairs, except between Alcoa and the three Neighbour Mondego, Tejo and Sado basins and between Lis and Mondego basins (*Table 9-20 section 9*). The highest gene flow was estimated to have occurred for Tejo-Colares and Tejo-Lizandro pairs (*Table 9-20 section 9*).

6. <u>**DISCUSSION**</u>

Given the high anthropogenic pressure which affects river basins (e.g. Allan and Flecker, 1993; Malmqvist and Rundle, 2002), the minimization of the extinction risk for freshwater fishes (one of the most threatened taxonomic groups at a global scale; Nelson *et al.*, 2016) requires a set of multidisciplinary data regarding the threatened populations. Thus, genetic and biogeographic data are important, as well as ecological data, to define evolutionary significant units, establish conservation priorities and propose well-founded conservation management practices.

In this context, it was known that the small river basins of the West region of Portugal harbour seven Iberian and Portuguese endemic cyprinids (*Iberochondrostoma lusitanicum, Achondrostoma occidentale, Achondrostoma oligolepis, Squalius pyrenaicus, Squalius carolitertii, Pseudochondrostoma polylepis* and *Luciobarbus bocagei*), most of which with a high conservation status and medium to low levels of genetic diversity (Cabral *et al.* 2005, Sousa-Santos *et al.* 2016). However, a complete analysis of all the populations from the West region regarding genetic diversity was still lacking, and there was no evidence allowing the reconstitution of the colonization routes used by the species and the assessment of the eventual barrier effect imposed by the mountain ranges isolating the Western river basins from the larger Neighbour river basins.

Thus, in order to provide genetic and biogeographic data for a more efficient conservation and management of the native cyprinid fish species from the West region of Portugal, their evolutionary history was drawn with the determination of the colonization source of each population, the divergence times between populations and their current levels of genetic diversity.

Finally, the importance of genetic and biogeographic analyses for conservation management plans is discussed according to the observed patterns, units of conservation and evolutionary significant units (ESUs) are suggested and future concerns for freshwater fish conservation are highlighted.

6.1 Evolutionary history of cyprinid populations from the West region of Portugal

As mentioned in the first section of this thesis, Cyprinid fish colonized the Iberian Peninsula during the Oligocene (Doadrio *et al.*, 2011) and dispersed through connections between endorheic Miocene lakes that only began to be exorheic in the Neogene (Andeweg, 2002). These endorheic basins were the precursors of current large river basins, such as those of the Douro, Ebro, Guadiana and Tejo, and acted as sources of colonizers when the Iberian river network started its establishment (Almada and Sousa-Santos, 2010). Then, as the hydrographical network developed, ancient fish were allowed to disperse throughout the Iberian territory and eventually reached the Western border of the Portuguese territory. According to the findings presented in this work, the colonization of the West region by cyprinids was posterior to the formation of the surrounding Montejunto, Sicó-Aire-Candeeiros and Sintra mountain systems of Jurassic age (Kullberg *et al.*, 2013).

Globally, the obtained results suggest that cyprinids colonized the West region during the Holocene, more recently than the Pleistocenic age previously proposed by Sousa-Santos *et al.* (2007) and Waap *et al.* (2011) for the Western populations of *S. pyrenaicus*. The estimated divergence times between the Western and Neighbour populations of *A. oligolepis* and *I. lusitanicum* indicate that the colonization of the West region ceased 7000 to 2500 years ago, coinciding with a marine regression and the reaching of the current sea level (Cunha and Gouveia, 2015). The estimated divergence times between the Western and Neighbour populations obtained for *S. pyrenaicus*, *S. carolitertii*, *L. bocagei* and *P. polylepis* are

lesser than 1500 years, reinforcing the view that the Western basins were colonized very recently, posteriorly to all the major geomorphological rearrangements that occurred in the region. Moreover, within the West region, the estimated divergence times between populations was null to low (less than 2500 years ago) for all the species.

Due to their intolerance to marine salinity, which prevented the migration between river mouths by the sea, the colonization and dispersion of cyprinids in the West region could only have occurred through connections between coastal paleobasins, fluvial headwater captures, confluence of river mouths during marine regressions and/or subterranean streams linking river basins. Dendritic subterranean drainages are common in karstic mountains such as Sicó-Aire-Candeeiros (Cunha, 1990) and underground dispersal of fish between river basins of both sides of a mountain range has been suggested in other regions of the globe (Kim *et al.*, 2016). Although cavernous cyprinids have only been found in South Spain (Doadrio *et al.,* 2016), the colonization of Western rivers by fish from the river basins located on the opposite slope of karstic mountains but linked though subterranean streams is plausible, especially during transgressions and intense pluvial periods. This hypothesis might explain the colonization of the rivers from the upper West region by fish from Mondego and Tejo basins, through subterranean streams present in the karstic Sicó-Aire-Candeeiros mountains, as will be detailed below.

The genetic validation of the taxonomy of the cyprinids present in the West region retrieved three distinct association of species: 1) *I. lusitanicum*, *S. pyrenaicus*, *P. polylepis* and *L. bocagei* in the southern part of the West region, including in Jamor, Ossos, Lage, Colares, Samarra and Lizandro river basins, 2) *A. occidentale,* in the intermediate part, including Safarujo, Sizandro and Alcabrichel river basins, and 3) *A. oligolepis*, *S. carolitertii*, *S. pyrenaicus*, *P. polylepis* and *L. bocagei* in the northern part of the West region, including Real, Tornada, Alcoa, São Pedro and Lis river basins.
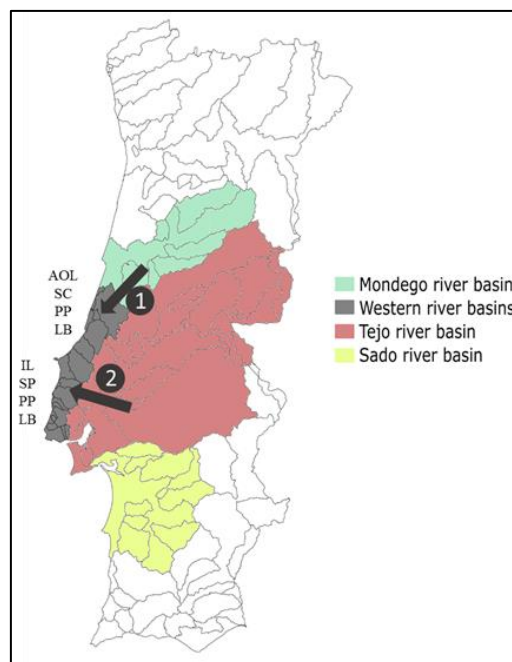
Figure 6-1: Diagram illustrating the West region colonization routes: 1) from the Mondego southwards by A. oligolepis (AOL), S. carolitertii (SC), P. polylepis (PP) and L. bocagei (LB); and 2) from the Tejo westwards by I. lusitanicum (IL), S. pyrenaicus (SP), P. polylepis (PP) and L. bocagei (LB).´

Due to their geographical positioning, the river basins of Mondego, Tejo and Sado were the three possible sources of colonizers for the Western basins. According to the higher divergence times found between Sado and the Western populations, Sado was not a direct source of colonizers. Contrastingly,

low divergence times were found between the populations from the northern part of the West region and Mondego, and between the populations from the southern part of the West region and Tejo. Thus, two independent routes are proposed to explain the colonization of the West region: 1) *S. carolitertii*, *A. oligolepis*, *L. bocagei* and *P. polylepis* followed a southward route from Mondego to the West region; and 2) *S. pyrenaicus*, *I. lusitanicum*, *L. bocagei* and *P. polylepis* followed a West-Northwestward route from Tejo to the West region (*Figure 6-1*). According to this scenario, *L. bocagei* and *P. polylepis* colonized both ends of the West region by distinct routes, one from the Mondego southwards and the other from the Tejo Westwards.

However, three exceptional occurrences are apparently inconsistent with the proposed colonization pattern: a) *A. oligolepis* population from the Nabão, the only Tejo sub-basin where this species occurs; b) *S. pyrenaicus* enclave in River Lis, flanked by *S. carolitertii* populations from the Mondego and Alcoa; and c) the three populations of *A. occidentale*, isolated in the intermediate part of the West region, where no other cyprinid species occur.

Regarding the *A. oligolepis* population from Nabão, a tributary of the right bank of the Tejo which is geographically close to a left bank sub-basin of the Mondego (River Arunca), its presence might be explained by one of the following hypothesis: 1) fish from the Mondego colonized the Nabão through an headwater capture in the Sicó mountain (currently Arunca and Nabão headwaters are only approximately 25 km apart); 2) flooded karstic subterranean drainages of the Sicó mountain allowed the passage of fish from the Mondego to the Nabão, and 3) anthropogenic introduction. Among these hypotheses, the first two appear more plausible than the last one, since the Tejo population shows private haplotypes derived from haplotypes present in the Mondego population (*Figure 6-3*). Although the third hypothesis may not be completely discarded, if the *A. oligolepis* population from Nabão was introduced its haplotypes would be shared with the population from where the introduced stock came from and, most likely, no private haplotypes of the cyt*b* gene would be found. Future studies conducted with fast paced markers, such as microsatellites, may help to disentangle this issue.

Secondly, considering the distribution ranges of *S. carolitertii* and *S. pyrenaicus* (*Figure 3-2*), it was not expected that *S. carolitertii* occurred in the Alcoa river basin and not in the Lis river basin, which is located immediately next to the Mondego (the previously assumed southernmost distribution limit of the species). The exceptional occurrence of a *S. pyrenaicus* population in the River Lis, flanked by two *S. carolitertii* populations (Mondego and Alcoa) might involve human-mediated introductions. Three alternative scenarios may be drawn to explain the *S. pyrenaicus* enclave in the River Lis: 1) *S. carolitertii* naturally dispersed from the Mondego southwards, reaching Lis and Alcoa rivers, but was latter displaced by *S. pyrenaicus* introduced in the Lis; 2) *S. pyrenaicus* from the Tejo colonized the River Lis through river capture (the headwaters of Lis are geographically close to those of the Alviela and Nabão tributaries) and *S. carolitertii* was introduced in the River Alcoa; and 3) *S. carolitertii* dispersed naturally from Mondego to Lis and to Alcoa when these two latter rivers were still connected and, more recently, a headwater capture between Lis and Tejo allowed the colonization by *S. pyrenaicus*, displacing (or largely supplanting in density) the former *S. carolitertii* population from the Lis. As the results obtained for the mitochondrial and nuclear genes were inconclusive, these hypotheses should be addressed in future studies to be conducted in the Alcoa and Lis populations, using larger samples, collected in multiple sites across the river basins, and genotyping polymorphic microsatellites.

And, finally, the exceptional occurrence of *A. occidentale*, phylogenetically closer to Spanish populations of *A. arcasii* than to the geographically closer *A. oligolepis* (*Figures 3-2 and 3-3*) was already explained by Robalo *et al.* (2006a). According to these authors, the *A. occidentale* populations are considered relic pockets, vicariantly separated from *A. arcasii*, with which the species shared a common ancestor, likely with a wide distribution range (Robalo *et al.*, 2006a). The radiation of *A.*

oligolepis along the Portuguese Atlantic border might have been related to the vicariant event which isolated *A. occidentale* in the intermediate part of the West region more than 7 Mya (Robalo *et al.*, 2006a), long before the estimated colonization of the region by the remaining cyprinid species. Thus, it seems that *A. occidentale* was the first species colonizing the West region or at least the most ancient species persisting in the region until the present. Also, since no other cyprinid occurs in sympatry with *A. occidentale*, it seems plausible that the above mentioned routes advanced to explain the colonization of the West region were not extendable to the Alcabrichel, Sizandro and Safarujo river basins. The route from Mondego southwards likely ended in the Real river basin (the southernmost distribution limit of *A. oligolepis*) and the route from the Tejo West-Northwestwards reached Lizandro river basin (the northernmost distribution limit of *I. lusitanicum*). Considering the geographical location of the river basins where it occurs, the isolation of *A. occidentale* in the intermediate part of the West region may be related to a barrier effect imposed by the Montejunto mountain range, preventing contacts with Tejo tributaries, and by the absence of recent connections between Neighbour river basins. Furthermore, the submarine canyons such as the Nazaré canyon could have acted as isolation barriers during regressions, preventing colonization by confluence of river mouths, as mention in section 1.2.

Assuming that the colonization routes were the same and were available at the same time for all the species, it was expected that the divergence times were similar between the Western populations and the Mondego and Tejo populations which acted as colonization sources for all the species. However, different patterns of differentiation between Western and Neighbour populations were found among species (*Table 6-1*). There was a moderated to low differentiation for *A. oligolepis* and *I. lusitanicum* and an extremely low differentiation for *S. carolitertii, S. pyrenaicus, P. polylepis* and *L. bocagei*. Thus, higher differentiation from the colonization source was detected for the species with smaller size, lower generation time and little or no migratory ability (*Table 6-1;* Ribeiro *et al.*, 2007). Although the difference is not substantial, these results point to the eventual existence of intrinsic drivers of genetic variation related to the life-history traits of the species, as already suggested by Sousa-Santos *et al.* (2016). Further studies are needed to clarify this issue.

Table 6-1: Summary results for each species (e.g. Doadrio et al., 2011; Ribeiro et al., 2007; Rodrigues, 1999)

Species	Max. size (cm)	Generation time (years)	Colonization source	Differentiation between Western and Neighbour populations	Intrapopulation diversity Average ± sd (Nb. populations)	% Variation among populations (AMOVA)
I. lusitanicum	15	2	Tejo	0.30%	0 (N=6)	-
A. oligolepis	15	2	Mondego	0.80%±0.07	0.69±0.35 (N=5)	74.96%
A. occidentale	15	2	-	-	0.57±0.75 (N=3)	65.85%
S. pyrenaicus	26	3.5	Tejo	0.05%±0.04	0.61±0.53 (N=6)	50.55%
S. carolitertii	26	3.5	Mondego	0.02%	0.10 (N=1)	-
P. polylepis – northern part	40	3.5	Mondego	0%	0.20±0.29 (N=2)	59.89%
P. polylepis – southern part	40	3.5	Tejo	0.01%	0.73 (N=1)	59.89%
L. bocagei – northern part	100	6	Mondego	0.02%±0.03	0.26±0.37 (N=2)	77.35%
L. bocagei – southern part	100	6	Tejo	0%	0 (N=2)	77.35%

6.2 Genetic Diversity of cyprinid populations from the West region of Portugal

Freshwater fish populations are facing a decrease in genetic diversity at the global scale, due to multiple, often cumulative, threats, such as habitat degradation and fragmentation, pollution, biological invasions, damming, water abstraction and flow regulation (e.g. Allan and Flecker 1993; Malmqvist and Rundle,

2002; Macedo-Veiga *et al*., 2013). The populations of cyprinids native to Portugal are not an exception (e.g. Collares-Pereira *et al.*, 2000; Ribeiro *et al.,* 2007; Sousa-Santos *et al.*, 2016).

Cyprinids from the West region of Portugal display low intra-population variability and moderate to low levels of molecular diversity, which are slightly lower than those found in the contiguous larger river basins of Mondego, Tejo and Sado. Moreover, the overall genetic diversity values obtained for the species occurring in the West region are slightly lower than those obtained for the same species at all national range scale (Sousa-Santos *et al.,* 2016).

Although this low diversity may be associated with the recent colonization of the West region, it can also result, at least in part, from the occurrence of successive bottlenecks caused by marine transgressions after the last glaciation and, more recently, by events imposing severe reductions of the population effective size (e.g. discharges of pollutants and cyclical summer droughts; Macedo-Veiga *et al*., 2013). Distinct populations from different river basins may had been differently affected by extrinsic drivers due to the different geomorphological characteristics of each river basins. This view is corroborated by the fact that, although the colonization of distinct Western river basins occurred within the same time frame, distinct levels of intra-population variability were detected for each species (*Figures 5-2, 5-4, 5-6, 5-8, 5-10, 5-12* and *5-14*).

Thus, the results indicate that the genetic variation may be constrained by environmental conditions, which may explain the distinct levels of genetic variation for each population of a given species, and also by species-specific traits, as mention in the previous section, which may explain differences between species occurring in the samebasins. An example of these differences is the fact that while populations of *I. lusitanicum* and *L. bocagei* from the southern part of the West region are identical (null intra-population variability), the sympatric *P. polylepis* and *S. pyrenaicus* showed moderate levels of intra-population variability (Table 6-1).

Globally, it thus appears that genetic diversity may be conditioned by multiple intrinsic and extrinsic drivers such as tolerance to environmental variations, habitat preferences, reproductive strategies, and the environmental variations sourced in geological, climatic and anthropological events that affect habitat availability, quality, and connectivity (Osborne *et al.,* 2014; Sousa-Santos *et al.,* 2016). Moreover, it can also be highlighted that, besides being recently separated, most of the Western populations are clearly differentiated from each other and from the Neighbour populations from which they were founded.

6.3 Management and conservation of cyprinid populations from the West region of Portugal

Taken together, the results obtained in this thesis have important implications for the conservation management of native cyprinids from the West region of Portugal. Although resulting from recent colonization events from the contiguous Tejo and Mondego river basins, the Western populations of each species are clearly differentiated. The observed patterns suggest the presence of accumulated differences within each population, reflecting their unique evolutionary histories after the cessation of past connections between distinct rivers and the consequent isolation of populations, but also the adaptation to different environmental conditions, as suggested by Frankham *et al.* (2010). Thus, for conservation and management purposes, the Western populations should be considered distinct conservation units and, thus, managed separately, with specific concerns and conservation measures. Indeed, the results obtained in this thesis reinforce the view that the conservation of primary freshwater fish should be considered an exceptional case within the conservation of fish *sensu lato*, similarly to what happens with the conservation of island species, due to their isolation and consequent particular

evolution. As such, all populations must also be viewed as ESUs since their evolutionary history is unique. However, if it is not possible to establish conservation plans encompassing all populations and some prioritization is necessary, the first targets for conservation should be the populations of *A. occidentale,* due to their high differentiation from the most common and geographically close *Achondrostoma* species (*A. oligolepis*), limited range and ancient vicariant origin. The populations of *I. lusitanicum* and *S. pyrenaicus* should be placed in a second line of priority due to their high conservation status and for being the only populations of these species occurring outside large sized river basins, thus, representing the edges of their respective radiations. Thirdly, and although the conservation status of the species is not high (Vulnerable), Western populations of *A. oligolepis* should also be prioritized due to their high divergence from the colonization source.

Contrasting the results obtained for the different species, intrinsic ecological and biological traits seem to be influencing their genetic variation. More specifically, it seems plausible to assume that the species which attain its first maturation at older ages (such as *L. bocagei* and *P. polylepis*) have less chances to accumulate mutations than more precocious species (such as *I. lusitanicum*, *A. oligolepis* and *A. occidentale*) which will have more generations within the same time frame and, thus, more chances to accumulate novel mutations. Similarly, populations of the larger species with higher migratory ability (such as the potamodromous *L. bocagei* and *P. polylepis;* Ribeiro *et al*., 2007) are expected to show more genetic homogenization due to higher gene flow than smaller sedentary species (such as *I. lusitanicum*, *A. oligolepis* and *A. occidentale*) which, due to their different ecology, are more prone to restricted gene flow between demes. This hypothesis was already suggested to explain the high genetic diversity of the small, sedentary, and fragmented populations of *Anaecypris hispanica* (Sousa-Santos *et al.*, 2014b). Thus, conservation management plans should take into account this variability and the measures to implement should be species-specific. For instance, for species with higher migratory ability, it is crucial to consider management measures that prevent habitat fragmentation and that restore fluvial connectivity, such as eliminating transversal barriers and limiting water abstraction during the summer. On the other hand, for smaller species, such as *I. lusitanicum*, *A. oligolepis* and *A. occidentale*, the management measures should be focused on habitat rehabilitation, namely concerning the preservation of aquatic vegetation and the restoration of the riparian gallery, and on increasing the amount of refuges available during the summer (Mameri, 2015).

Finally, it was clear that native cyprinids from the West region show low levels of genetic diversity, which theoretically will decrease the species ability to overcome environmental changes and, consequently, increase their risk of extinction (Frankham *et al.*, 2010). These low levels of genetic diversity may be related to the loss of adequate spawning grounds, habitat fragmentation and poor conditions for the persistence of populations. These environmental negative impacts would lead to the reduction of the effective population size and, ultimately, to the loss of genetic diversity by inbreeding and lineage sorting (e.g. DeSalle and Amato, 2004; Allendorf *et al.,* 2010; Frankham *et al.*, 2010). Overall solutions include conservation efforts such as habitat restoration, creation of refuges, population reinforcements by ex-situ conservation programs, and re-introductions and translocations (Ginson, 2012). Despite the existence of international guidelines (McGowan *et al.*, 2016), the re-introduction and translocation of species are often controversial (Maceda-Veiga, 2013), mainly because of the outbreeding depression risk (Week *et al.,* 2011; Frankham *et al.*, 2010) and the undesirable mixing of different conservation units (Moritz, 1994), as mention in the section 1.1. As it was evident from this study, in the particular case of native cyprinids, distinct populations have unique and unrepeatable evolutionary histories due to the permanent confinement to their river basins, thus, the translocation of individuals between populations to increase genetic diversity should be avoided at all cost. Alternatively, population reinforcements with captive bred fish (ongoing since 2011 in five Western river basins) are

a valuable conservation tool to increase the effective population size and minimize the risk of extinction, and should be conducted in parallel with in-situ conservation measures (Sousa-Santos *et al.* 2014b).

Another common conservation tool is the delineation of protected areas. Although small specific protection areas can act as sanctuaries for the endangered fish fauna and be a stepping stone for future recolonization of the remainder of the river basin, in order to take full effect of this measure, global protection of the river basin would be necessary since this ecosystem is a continuum (Dudgeon *et al*., 2006). In fact, if the environmental conditions in the remainder of the river basin are not improved, the populations might become reduced to the sanctuaries. Moreover, the pressures that threaten the remainder of the river basin might even lead to the degradation of the sanctuaries' conditions. Thus, global measures that aim to protect the entire river basin should be prioritized and can be considered as a better cost-efficiency option (Griffiths and Pavajeau 2008), increasing reproductive success and enabling the natural regeneration of the populations (Frankham *et al.*, 2010). Furthermore, these measures could also prevent the aggravation of vulnerable species status, avoiding the loss of genetic diversity and the consequent loss of adaptation ability to environmental variations (Griffiths and Pavajeau, 2008).

Likewise, introduction of global conservation measures in the river basins of the West region, would likely result in substantial improvements of the local fish communities. As such, future management projects targeting the native fish fauna in the West region should also focus on preventing the discharge of pollutants, limiting water abstraction, eradicating invasive fauna and flora such as *Arundo donax* and *Procambarus clarkii* and restoring the native riparian gallery.

7. <u>FINAL REMARKS</u>

In conclusion, this study underlines the importance of conservation genetics to the preservation of endemic cyprinid species native to Portugal. Furthermore, it highlights the urgent need for multidisciplinary studies, encompassing among others genetic and phylogenetic data, and resulting in more well-reasoned decisions. Fine scale studies, such as this one, are essential to provide detailed information about freshwater fish populations and complement broad scale genetic studies (Wu *et al.*, 2016).

In summary, Western populations of native cyprinids display low differentiation from the Neighbour river basins of Mondego and Tejo, reflecting recent colonization events (dated from the Holocene), most likely through headwater fluvial captures, connection between coastal paleobasins and/or confluence of river mouths during marine regressions. Despite the recent colonization and the low diversity found, populations of most of the native cyprinid species inhabiting the West region have already accumulated differences and show clear differentiation between them. The Western populations of native cyprinids should thus be considered as independent conservation units, due to their island-like idiosyncrasies, and, moreover, effective management and conservation measures should be species-specific.

Future studies should specifically address ecological, biological, and environmental drivers of genetic diversity for a better comprehension of the patterns highlighted in this study. Also, further knowledge on the availability and connectivity of suitable habitats for fish and the way this may vary under future climate scenarios would be critical for the development of effective conservation management plans for river basins in the Western region.

8. REFERENCES

Aboim, M.A., Mesquita, N., Drago, M., Coelho, M.M., Alves, M.J., 2013. Assessing inter-drainage connections: patterns of genetic diversity in an Iberian cyprinid fish. *Biological Journal of the Linnean Society*, 109, 656–669;

Agência Portuguesa do Ambiente, I.P., 2016 Sistema Nacional de Informação de Ambiente. Electronic publishing available in: <http://sniamb.apambiente.pt/pgrh/> accessed in 20 September 2016;

Allan, J.D., Flecker, A.S., 1993. Biodiversity conservation in running waters. *BioScience* 43, 32–43;

Allendorf, F. W., Hohenlohe, P. A., Luikart, G., 2010. Genomics and the future of conservation genetics. *Nature reviews genetics*, 11, 697-709.

Almada, V., Sousa-Santos, C., 2010. Comparisons of the genetic structure of Squalius populations (Teleostei, Cyprinidae) from rivers with contrasting histories, drainage areas and climatic conditions based on two molecular markers. *Molecular Phylogenetics and Evolution*, 57, 924-931;

Alves, T.M., Gawthorpe, R.L., Hunt, D.W., Monteiro, J.H., 2003. Cenozoic tectono-sedimentary evolution of the Western Iberian margin. *Marine Geology*, 195, 75-108;

Andeweg, B., 2002. Cenozoic tectonic evolution of the Iberian Peninsula. PhD Thesis. Vrije Universiteit;

Banarescu, P., 1992. Zoogeography of fresh waters Distribution and dispersal of freshwater animals in North America and Eurasia (Vol. II). Aula Verlag, Wiesbaden;

Bianco, P.G., 1990. Potential role of the palaeohistory of the Mediterranean and Paratethis basins on the early dispersal of Euro-Mediterranean freshwater fishes. *Ichthyological Exploration Freshwaters,* 1, 167–184;

Briggs, J.C., 1995. Global biogeography. In Developments in paleontology and stratigraphy, V. 14, Oxford, UK, Elsevier;

Cabral, M.J., Almeida, J., Almeida, P.R., Dellinger, T.R., Ferrand de Almeida, N., Oliveira, M.E., Palmeirim, J.M., Queiroz, A.I., Rogado, L., Santos-Reis, M., 2005. Livro Vermelho dos Vertebrados de Portugal. Instituto da Conservação da Natureza, Lisboa;

Carmona, J.A., Doadrio, I., 2000. Threatened fishes of the world: Leuciscus carolitertii Doadrio, 1988 (Cyprinidae). *Environmental Biology of Fishes*, 57, 96-96;

Casas-Sainz, A.M., De Vicente, G., 2009. On the tectonic origin of Iberian topography. *Tectonophysics*, 474, 214-235;

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, 9, 1657-1659;

Collares-Pereira, M.J., Cowx, I.G., Ribeiro, F., Rodrigues, J.A., Rogado, L., 2000. Threats imposed by water resource development schemes on the conservation of endangered fish species in the Guadiana River Basin in Portugal. *Fisheries Management and Ecology*, 7, 167-178;

Cunha, L., 1990. Serras Calcárias de Condeixa-Sicó-Alvaiázare. Estudo de geomorfologia. Col. Geografia Física, 1 Instituto Nacional de Investigação Cientifica, Coimbra;

Cunha, P.P., Gouveia, M.P., 2015. The Nazaré coast, the submarine canyon and the giant waves: a synthesis. Marine and Environmental Sciences Centre. University of Coimbra, Coimbra;

Cunha, P.P., Martins, A.A., 2004. Principais aspectos geomorfológicos de Portugal central, sua relação com o registo sedimentar e a relevante importância do controlo tectónico in Araújo, M.A., Gomes A. (eds.) Geomorfologia do NW da Península Ibérica. Faculdade de Letras da Universidade do Porto, 155-182;

Darlington, PJ., 1957. Fresh-water fishes zoogeograhy: the geographical distribution of animals. John Wiley & Sons, Inc., New York;

DeSalle, R., Amato, G., 2004. The expansion of conservation genetics. *Nature Reviews Genetics*, 5, 702-712.

Dekar, M.P., Magoulick, D.D., 2007. Factors affecting fish assemblage structure during seasonal stream drying. *Ecology of Freshwater Fish,* 16, 335-342;

Dias, J.M.A., Boski, T., Rodrigues, A., Magalhães, F., 2000. Coast line evolution in Portugal since the Last Glacial Maximum until present: a synthesis. *Marine Geology*, 170, 177-186;

Doadrio, I., Carmona, J.A., 2004. Phylogenetic relationships and biogeography of the genus Chondrostoma inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, 33, 802-815;

Doadrio, I., Perea, S., Garzón-Heydt, P., González, J.L., 2011. Ictiofauna Continental Española. Bases para su seguimiento. Dirección General Medio Natural y Política Forestal, Ministerio de Medio Ambiente y Medio Rural y Marino, Madrid;

Doadrio, I., Valladolid, M., Carmona, J.A., Corona-Santiago, D.K., Perea, S., Cunha, C., Boto, L., 2016. Sobre la presencia del complejo híbrido *Squalius alburnoides* (Steindachner, 1866) (Cyprinidae, Actinopterygii) en un sistema de cuevas situado en el sur de España. *Actas EspeleoMeting Ciudad de Villacarrillo*, 9-15;

Dowling, T., Tibbets, C.A., Minckley, W.L., Smith, G.R., 2002. Evolutionary relationships of the Plagopterins (Teleostei: Cyprinidae) from cytochrome b sequences. *Copeia,* 3, 665-678;

Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.I., Knowler, D.J., Lévêque, C., Naiman, R.J., Prieur-Richard, A., Soto, D., Stiassny, M.L.J., Sullivan, C.A., 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews*, 81, 163-182;

Durand, J.D., Tsigenopoulos, C.S., Ünlü, E., Berrebi, P., 2002. Phylogeny and biogeography of the family Cyprinidae in the Middle East inferred from cytochrome b DNA—evolutionary significance of this region. *Molecular Phylogenetics and Evolution*, 22, 91-100;

Excoffier, L., Lischer, H., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources,* 10, 564-567;

Feio, M. J., Norris, R. H., Graça, M. A. S., Nichols, S., 2009. Water quality assessment of Portuguese streams: Regional or national predictive models?. *Ecological Indicators*, 9, 791-806;

Fenolio, D.B., Zhao, Y., Niemiller, M.L., Stout, J.F. (2013). In-situ observations of seven enigmatic cave loaches and one cave barbel from Guangxi, China, with notes on conservation status. *Speleobiology Notes*, 5, 19-33;

Frankham, R., 2003. Genetics and conservation biology. *Comptes Rendus Biologies*, 326, 22-29;

Frankham, R., 2005. Genetics and extinction. *Biological Conservation*, *126*, 131-140;

Frankham, R., 2010. Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation*, 143, 1919-1927;

Frankham, R., Briscoe, D.A., Ballou, J.D., 2010. Introduction to conservation genetics. (Second Edition) Cambridge University Press, New York;

Freyhof, J. & Kottelat, M. 2008. *Achondrostoma occidentale*. The IUCN Red List of Threatened Species 2008: e.T135683A4180158. available in: <http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T135683A4180158.en.> accessed in 20 September 2016;

Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147, 915-925;

Gante, H.F., Micael, J., Oliva-Paterna, F.J., Doadrio, I., Dowling, T.E., Alves, M.J., 2009. Diversification within glacial refugia: tempo and mode of evolution of the polytypic fish *Barbus sclateri*. *Molecular Ecology*, 18, 3240-3255;

Gante, H.F., Santos, C.D., Alves, M.J., 2010. Phylogenetic relationships of the newly described species *Chondrostoma olisiponensis* (Teleostei: Cyprinidae). *Journal of Fish Biology*, 76, 965-974;

Gante, H.F., Doadrio, I., Alves, M.J., Dowling, T.E., 2015. Semi-permeable species boundaries in Iberian barbels (Barbus and Luciobarbus, Cyprinidae). *BMC Evolutionary Biology*, 15, 1;

Ginson, R., 2012. Population and conservation genetics of a habitat-specific riverine fish species, the eastern sand darter (Ammocrypta pellucida). M. Sc. Dissertation, University of Windsor;

Griffiths, R.A., Pavajeau, L., 2008. Captive breeding, reintroduction, and the conservation of amphibians. *Conservation Biology*, 22, 852-861;

Hewitt, G., 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society,* 68, 87-112;

Houston, D.D., Shiozawa, D.K., Riddle, B.R., 2010. Phylogenetic relationships of the Western North American cyprinid genus Richardsonius, with an overview of phylogeographic structure. *Molecular Phylogenetics and Evolution*, 55, 259-273;

Jones, P. D., Mann, M. E., 2004. Climate over past millennia. *Reviews of Geophysics*, 42;

Kim, D., Hirt, M. V., Simons, A. M., Won, Y. J., 2016. Small fishes crossed a large mountain range: Quaternary stream capture events and freshwater fishes on both sides of the Taebaek Mountains. *Integrative Zoology*; Accepted Author Manuscript;

Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J., Wilson, D.S., 1999. Chronology, causes and progression of the Messinian salinity crisis. *Nature*, 400, 652-655;

Kotlik, P., Markova, S., Choleva, L., Bogutskaya, N.G., Ekmekci, F.G., Ivanova, P.P., 2008. Divergence with gene flow between Ponto-Caspian refugia in an anadromous cyprinid Rutilus frisii revealed by multiple gene phylogeography. *Molecular Ecology*, 17, 1076-1088;

Kottelat, M., Freyhof, J., 2007. Handbook of European freshwater fishes. Publications Kottelat, Cornol and Freyhof, Berlin;

Kullberg, J.C., Rocha, R.B., Soares, A.F., Rey, J., Terrinha, P., Azerêdo, A.C., Callapez, P., Duarte, L.V., Kullberg, M.C., Martins, L., Miranda, R., Alves, C., Mata, J., Madeira, J., Mateus, O., Moreira, M., Nogueira, C.R., 2013. A Bacia Lusitaniana: Estratigrafia, Paleogeografia e Tectónica, in Dias, R., Araújo, A., Terrinha, P., Kullberg, J.C. (eds), Geologia de Portugal Volume II Geologia Meso-cenozóica de Portugal, Escolar Editora, Lisboa;

Levy, A., Doadrio, I., Almada, V.C., 2009. Historical biogeography of European leuciscins (Cyprinidae): evaluating the Lago Mare dispersal hypothesis. *Journal of biogeography*, 36, 55-65;

Maceda-Veiga, A., 2013. Towards the conservation of freshwater fish: Iberian Rivers as an example of threats and management practices. *Reviews in Fish Biology and Fisheries*, 23, 1-22;

Machordom, A., Doadrio, I., 2001. Evidence of a Cenozoic Betic–Kabilian connection based on freshwater fish phylogeography (Luciobarbus, Cyprinidae).*Molecular Phylogenetics and Evolution*, 18, 252-263;

Magalhães, M.F., Beja, P., Canas, C., Collares-Pereira, M., 2002. Functional heterogeneity of dry-season fish refugia across a Mediterranean catchment: the role of habitat and predation. *Freshwater Biology,* 47, 1919-1934;

Magoulick, D.D., Kobza, R.M., 2003. The role of refugia for fishes during drought: a review and synthesis. *Freshwater Biology*, 48, 1186-1198;

Malmqvist, B., Rundle, S., 2002. Threats to the running water ecosystems of the world. *Environmental Conservation,* 29, 134-153;

Mameri, D. 2015. Habitat use, growth and reproductive behavior of the western ruivaco (Cyprinidae): contributes to the restocking of wild populations. M. Sc. Dissertation, Universidade de Lisboa;

Mateus, C.S., Quintella, B.R., Almeida, P.R., 2008. The critical swimming speed of Iberian barbel Barbus bocagei in relation to size and sex. *Journal of Fish Biology*, 73, 1783-1789;

Mateus, C., Alves, J., Quintella, B., Almeida, P.R., 2013. Three new cryptic species of the lamprey genus Lampetra Bonnaterre, 1788 (Petromyzontiformes: Petromyzontidae) from the Iberian Peninsula. *Contribution to Zoology*, 82, 37-53;

McGowan, P.J., Traylor-Holzer, K., Leus, K., 2016. IUCN Guidelines for Determining When and How Ex Situ Management Should Be Used in Species Conservation. *Conservation Letters*, Accepted Author Manuscript;

Moritz, C., 1994. Defining'evolutionarily significant units' for conservation. *Trends in Ecology and Evolution*, 9, 373-374;

Moritz, C., 1999. Conservation units and translocations: strategies for conserving evolutionary processes. *Hereditas*, 130, 217-228;

Nesbø, C.L., Fossheim, T., Vøllestad, L.A., Jakobsen, K.S., 1999. Genetic divergence and phylogeographic relationships among European perch (*Perca fluviatilis*) populations reflect glacial refugia and postglacial colonization. *Molecular Ecology*, 8, 1387-1404;

Nelson, J. S., Grande, T.C., Wilson, M.V., 2016. Fishes of the World. John Wiley & Sons, Hoboken, New Jersey;

Osborne, M.J., Perkin, J.S., Gido, K.B., Turner, T.F., 2014. Comparative riverscape genetics reveals reservoirs of genetic diversity for conservation and restoration of Great Plains fishes. *Molecular Ecology,* 23, 5663-5679;

Pais, J., Cunha, P.P., Pereira, D., Legoinha, P., Dias, R., Moura, D., da Silveira, A.B., Kullberg, J.C., González-Delgado, J.A., 2012. The Paleogene and Neogene of Western Iberia (Portugal). Springer Briefs in Earth Sciences. Springer, New York;

Perea, S., Böhme, M., Zupančič, P., Freyhof, J., Šanda, R., Özuluğ, M., Abdoli, A., Doadrio, I., 2010. Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean subfamily Leuciscinae (Teleostei, Cyprinidae) inferred from both mitochondrial and nuclear data. *BMC Evolutionary Biology*, 10, 1;

Perea, S., Doadrio, I., 2015. Phylogeography, historical demography and habitat suitability modelling of freshwater fishes inhabiting seasonally fluctuating Mediterranean river systems: a case study using the Iberian cyprinid Squalius valentinus. *Molecular Ecology*, 24, 3706-3722;

Perea, S., Cobo-Simon, M., Doadrio, I., 2016. Cenozoic tectonic and climatic events in southern Iberian Peninsula: Implications for the evolutionary history of freshwater fish of the genus Squalius (Actinopterygii, Cyprinidae). *Molecular Phylogenetics and Evolution*, 97, 155-169;

Ribeiro, F., Beldade, R., Dix, M., Bochechas, J. 2007. Carta Piscícola Nacional. Direção-Geral dos Recursos Florestais - Fluviatilis, Lda. Electronic publishing (versão 01/2007) available in: < http://www.cartapiscicola.org/ > accessed in 11 October 2016;

Robalo, J.I., Sousa-Santos, C., Almada, V.C., Doadrio, I., 2006a. Paleobiogeography of Two Iberian Endemic Cyprinid Fishes (*Chondrostoma arcasii – Chondrostoma macrolepidotus*) Inferred from Mitochondrial DNA Sequence Data. *Journal of Heredity*, 97, 143-149;

Robalo, J.I., Sousa-Santos, C., Levy, A., Almada, V.C., 2006b. Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybrigogene complex. *Molecular Phylogenetics and Evolution,* 39, 276-281;

Robalo, J., 2007. Filogenia, filogeografia e comportamento dos pequenos ciprinídeos do género *Chondrostoma Agassiz, 1832 (Actinopterygii: Cyprinidae)*. PhD Thesis, Universidade do Porto;

Robalo, J.I., Almada, V.C., Levy, A., Doadrio, I., 2007. Re-examination and phylogeny of the genus Chondrostoma based on mitochondrial and nuclear data and the definition of 5 new genera. *Molecular Phylogenetics and Evolution,* 42, 362-372;

Robalo, J.I., Sousa-Santos, C., Doadrio, I., Almada, V.C., 2008. Threatened fishes of the world: *Achondrostoma occidentale* Robalo, Almada, Sousa-Santos, Moreira & Doadrio 2005 (Cyprinidae). *Environmental Biology of Fishes,* 83, 347-347;

Rodrigues, A., Dias, J.A., 1989. Evolução pós-glaciária da plataforma continental Portuguesa a norte do cabo Mondego. *Anais do Instituto Hidrográfico*, 10, 39-50;

Rodrigues, J. A., 1999. Aspectos da bio-ecologia das populações de Leuciscus pyrenaicus Günther, 1868 (Pisces, Cyprinidae) na Bacia Hidrográfica do rio Tagus. PhD Thesis. Universidade de Lisboa, Lisboa;

Rogers, A.R., Harpending, H., 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution,* 9, 552-569;

Santos, A.M., Cabezas, M.P., Tavares, A.I., Xavier, R., Branco, M., 2015. TcsBU: a tool to extend TCS network layout and visualization. *Bioinformatics*;

Schmidt, T.R., Gold, J.R., 1993. Complete sequences of the mitochondrial cytochrome *b* gene in the Cherryfin Shinner, *Liturus roseipinnis* (Teleostei: Cyprinidae). *Copeia,* 3, 880-883;

Sousa-Santos, C., Robalo, J., Collares-Pereira, M.J., Almada V., 2005. Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genomic composition of hybrid organisms. *DNA Sequence,* 16, 462–467;

Sousa-Santos, C., 2007. Reproductive behaviour and the evolutionary history of the hybridogenetic complex *Squalius alburnoides* (Pisces, Cyprinidae). PhD thesis. Universidade de Lisboa;

Sousa-Santos, C., Collares-Pereira, M. J., Almada, V., 2007. Reading the history of a hybrid fish complex from its molecular record. *Molecular Phylogenetics and Evolution*, 45, 981-996;

Sousa-Santos, C., Robalo, J., Santos, J.M., Branco, P., Ferreira, T., Sousa, M., Ramos, A., Castilho, R., Doadrio, I., Almada, V., 2013. Atlas Genético Nacional dos peixes ciprinídeos nativos. Electronic publishing available in: <http://www.fishatlas.net> accessed in 06 Junho 2016;

Sousa-Santos, C., Gante, H.F., Robalo, J., Cunha, P.P., Martins, A., Arruda, M., Alves, M.J., Almada, V., 2014a Evolutionary history and population genetics of a cyprinid fish (*Iberochondrostoma olisiponensis*) endangered by introgression from a more abundant relative. *Conservation Genetics,* 15, 665–677;

Sousa-Santos, C., Robalo, J.I., Francisco, S.M., Carrapato, C., Cardoso, A.C., Doadrio, I., 2014b Metapopulations in temporary streams – The role of drought–flood cycles in promoting high genetic diversity in a critically endangered freshwater fish and its consequences for the future. *Molecular Phylogenetics Evololution*, 80, 281-296;

Sousa-Santos, C., Robalo, J.I., Pereira, A.M., Branco, P., Santos, J.M., Ferreira, M.T., Sousa, M., Doadrio, I., 2016. Broad-scale sampling of primary freshwater fish populations reveals the role of intrinsic traits, inter-basin connectivity, drainage area and latitude on shaping contemporary patterns of genetic diversity. *Peer Journal*, 4, 1694;

Taberlet, P., Fumagalli, L., Wust-saucy, A.G., Cosson, J.F., 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, 7, 453-464;

Tajima, F., 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105: 437-460;

Teixeira, C., Gonçalves, F., 1980. Introdução à geologia de Portugal. Instituto Nacional de Investigação científica, Universidade de Lisboa, Lisboa;

Tsigenopoulos, C.S., Durand, J.D., Unlu, E., Berrebi, P., 2003. Rapid radiation of the Mediterranean Luciobarbus species (Cyprinidae) after the Messinian salinity crisis of the Mediterranean Sea, inferred from mitochondrial phylogenetic analysis. *Biological Journal of the Linnean Society*, 80, 207–222;

Waap, S., Amaral, A.R., Gomes, B., Coelho, M.M., 2011. Multi-locus species tree of the chub genus Squalius (Leuciscinae: Cyprinidae) from Western Iberia: new insights into its evolutionary history. *Genetica*, *139*, 1009-1018;

Weeks, A.R., Sgro, C.M., Young, A.G., Frankham, R., Mitchell, N.J., Miller, K.A., Byrne, M., Coates, D.J., Eldridge, M.D.B., Sunnucks, P., Breed, M.F., James, E.A., Hoffmann, A.A., 2011.

Assessing the benefits and risks of translocations in changing environments: a genetic perspective. *Evolutionary Applications*, 4, 709–725.

Wu, T.H., Tsang, L.M., Chen, I-S., Chu, K.H., 2016. Multilocus approach reveals cryptic lineages in the goby Rhinogobius duospilus in Hong Kong streams: Role of paleodrainage systems in shaping marked population differentiation in a city. *Molecular Phylogenetics and Evolution*, 104, 112-122.

9. ANEXOS

I: Haplotype network frequencies

Table 9-1: Frequency of the haplotypes found in I. lusitanicum cyt b haplotype network.

Haplotype	Frequency	Haplotype	Frequency
IL1	143	IL14	2
IL2	2	IL15	35
IL3	41	IL16	23
IL4	16	IL17	2
IL5	2	IL18	1
IL6	1	IL19	37
IL7	1	IL20	11
IL8	2	IL21	7
IL9	1	IL22	2
IL10	1	IL23	13
IL11	3	IL24	8
IL12	1	IL25	1
IL13	1		

Table 9-2: Frequency of the haplotypes found in A. occidentale cyt b haplotype network.

Haplotype	Frequency
AOC1	1
AOC2	31
AOC3	4
AOC4	5

Table 9-3: Frequency of the haplotypes found in A. oligolepis cyt b haplotype network.

Haplotype	Frequency	Haplotype	Frequency	Haplotype	Frequency
AOL1	42	AOL12	1	AOL23	11
AOL2	1	AOL13	1	AOL24	6
AOL3	2	AOL14	49	AOL25	4
AOL4	15	AOL15	1	AOL26	2
AOL5	1	AOL16	1	AOL27	9
AOL6	23	AOL17	1	AOL28	3
AOL7	1	AOL18	4	AOL29	3
AOL8	2	AOL19	3	AOL30	15
AOL9	1	AOL20	6	AOL31	1
AOL10	1	AOL21	9	AOL32	1
AOL11	1	AOL22	3	AOL33	2

Table 9-4: Frequency of the haplotypes found in S. pyrenaicus cyt b haplotype network.

Haplotype	Frequency	Haplotype	Frequency	Haplotype	Frequency	Haplotype	Frequency
SP1	92	SP17	3	SP33	4	SP49	1
SP2	1	SP18	3	SP34	4	SP50	1
SP3	4	SP19	4	SP35	3	SP51	1
SP4	1	SP20	1	SP36	3	SP52	1

Haplotype	Frequency	Haplotype	Frequency	Haplotype	Frequency	Haplotype	Frequency
SP5	2	SP21	1	SP37	3	SP53	1
SP6	4	SP22	2	SP38	3	SP54	1
SP7	28	SP23	1	SP39	3	SP55	1
SP8	1	SP24	1	SP40	3	SP56	1
SP9	2	SP25	13	SP41	2	SP57	1
SP10	2	SP26	9	SP42	2	SP58	1
SP11	2	SP27	7	SP43	2	SP59	1
SP12	1	SP28	7	SP44	1	SP60	1
SP13	1	SP29	8	SP45	1	SP61	1
SP14	19	SP30	8	SP46	1		
SP15	1	SP31	5	SP47	1		
SP16	65	SP32	5	SP48	1		

Table 9-5: Frequency of the haplotypes found in S. carolitertii cyt b haplotype network.

Haplotype	Frequency	Haplotype	Frequency
SC1	64	SC9	2
SC2	1	SC10	1
SC3	12	SC11	1
SC4	2	SC12	5
SC5	1	SC13	17
SC6	1	SC14	1
SC7	2	SC15	1
SC8	2		

Table 9-6: Frequency of the haplotypes found in P. polylepis cyt b haplotype network.

Haplotype	Frequency	Haplotype	Frequency	Haplotype	Frequency
PP1	167	PP11	11	PP21	1
PP2	101	PP12	1	PP22	1
PP3	11	PP13	4	PP23	11
PP4	1	PP14	10	PP24	1
PP5	2	PP15	5	PP25	1
PP6	1	PP16	6	PP26	3
PP7	7	PP17	1	PP27	1
PP8	20	PP18	1	PP28	1
PP9	1	PP19	1	PP29	1
PP10	2	PP20	1		

Table 9-7: Frequency of the haplotypes found in L. bocagei cyt b haplotype network.

Haplotype	Frequency	Haplotype	Frequency
LB1	94	LB7	6
LB2	369	LB8	9
LB3	1	LB9	12
LB4	20	LB10	1
LB5	14	LB11	1
LB6	21		

Table 9-8: Frequency of the haplotypes found in A. oligolepis beta-actine haplotype network.

Haplotype	Frequency	Haplotype	Frequency
AOB1	41	AOB7	8
AOB2	32	AOB8	5
AOB3	15	AOB9	5
AOB4	15	AOB10	3
AOB5	10	AOB11	1
AOB6	8	AOB12	1

Table 9-9: Frequency of the haplotypes found in Squalius beta-actine haplotype network.

Haplotype	Frequency	Haplotype	Frequency	Haplotype	Frequency
S1	8	S8	31	S15	1
S2	188	S9	63	S16	12
S3	2	S10	15	S17	4
S4	56	S11	13	S18	2
S5	136	S12	1	S19	5
S6	2	S13	23	S20	2
S7	2	S14	1	S21	25

II: Tajima's D and Fu's Fs neutrality tests results

Table 9-10: Tajima's D and Fu's Fs neutrality tests results; observed mismatch distributions; and significant values of the goodness-of-fit sum of squares deviations (SSD) test for the sudden and spatial expansion models of A. occidentale populations from the West region.

Population	Neutrality tests				Mismatch		Sudden expansion		Spatial expansion	
	Tajima's D	Tajima's D p	Fu Fs	Fu Fs p	Observed mean	Observed variance	SSD	P	SSD	P
Alcabrichel	0.148	0.605	-2.003	0.091	1.423	0.769	0.0441	0.033	0.0441	0.000
Safarujo	-1.727	0.018	0.494	0.370	0.286	0.779	0.0133	0.065	0.0058	0.173
Sizandro	0.000	1.000	0.000	N.A.	0.000	0.000	0.0000	0.000	0.0000	0.000

Table 9-11: Tajima's D and Fu's Fs neutrality tests; observed mismatch distributions; and significant values of the goodness-of-fit sum of squares deviations (SSD) test for the sudden and spatial expansion models of A. oligolepis populations from the West region.

Population	Neutrality tests				Mismatch		Sudden expansion		Spatial expansion	
	Tajima's D	Tajima's D p	Fu Fs	Fu Fs p	Observed mean	Observed variance	SSD	P	SSD	P
Lis	-0.760	0.266	-1.592	0.081	0.832	0.448	**0.0277**	**0.064**	0.0277	0.028
São Pedro	**-1.630**	**0.037**	1.211	0.757	0.995	3.339	**0.0180**	**0.153**	0.0180	0.045
Alcoa	**-1.513**	**0.047**	**-1.863**	**0.011**	0.200	0.171	**0.0016**	**0.396**	0.0005	0.459
Tornada	**-1.652**	**0.032**	**-2.206**	**0.002**	0.462	0.330	**0.0068**	**0.502**	0.0068	0.366
Real	-1.923	0.106	**-3.669**	**0.002**	0.968	1.301	0.3624	0.000	**0.0007**	**0.900**

Table 9-12: Tajima's D and Fu's Fs neutrality tests; observed mismatch distributions; and significant values of the goodness-of-fit sum of squares deviations (SSD) test for the sudden and spatial expansion models of S. pyrenaicus populations from the West region.

Population	Neutrality tests				Mismatch		Sudden expansion		Spatial expansion	
	Tajima's D	Tajima's D p	Fu *Fs*	Fu *Fs p*	Observed mean	Observed variance	SSD	*P*	SSD	*P*
Lis	-1.888	0.012	0.413	0.546	0.689	2.035	0.0345	0.064	0.0078	0.375
Lizandro	**-1.545**	**0.047**	-1.957	0.083	1.574	3.537	0.0218	0.426	0.0179	0.491
Samarra	-1.164	0.144	-0.879	0.085	0.100	0.090	0.0001	0.306	0.0001	0.318
Colares	-1.007	0.188	-1.456	0.054	0.514	0.337	0.0125	0.261	0.0125	0.115
Lage	-0.592	0.245	-0.097	0.212	0.189	0.154	0.0152	0.272	0.0002	0.287
Jamor	**-2.056**	**0.005**	1.743	0.771	0.600	3.257	0.0147	0.056	0.0074	0.203

Table 9-13: Tajima's D and Fu's Fs neutrality tests results; observed mismatch distributions; and significant values of the goodness-of-fit sum of squares deviations (SSD) test for the sudden and spatial expansion models of S. carolitertii population from the West region.

Population	Neutrality tests				Mismatch		Sudden expansion		Spatial expansion	
	Tajima's D	Tajima's D p	Fu *Fs*	Fu *Fs p*	Observed mean	Observed variance	SSD	*P*	SSD	*P*
Alcoa	-1.164	0.145	-0.879	0.083	0.100	0.09	0.0001	0.295	0.0001	0.336

Table 9-14: Tajima's D and Fu's Fs neutrality tests; observed mismatch distributions; and significant values of the goodness-of-fit sum of squares deviations (SSD) test for the sudden and spatial expansion models of P. polylepis populations from the West region.

Population	Neutrality tests				Mismatch		Sudden expansion		Spatial expansion	
	Tajima's D	Tajima's D p	Fu *Fs*	Fu *Fs p*	Observed mean	Observed variance	SSD	*P*	SSD	*P*
Lis	-1.422	0.070	-2.070	0.010	0.409	0.302	0.0046	0.467	0.0046	0.259
Alcoa	0.000	1.000	0.000	N.A.	0.000	0.000	0.0000	0.000	0.0000	0.000
Colares	0.555	0.751	0.348	0.521	0.733	0.399	0.0334	0.074	0.0334	0.029

Table 9-15: Tajima's D and Fu's Fs neutrality tests; observed mismatch distributions; and significant values of the goodness-of-fit sum of squares deviations (SSD) test for the sudden and spatial expansion models of L. bocagei populations from the West region.

Population	Neutrality tests				Mismatch		Sudden expansion		Spatial expansion	
	Tajima's D	Tajima's D p	Fu *Fs*	Fu *Fs p*	Observed mean	Observed variance	SSD	*P*	SSD	*P*
Alcoa	0.000	1.000	0.000	N.A.	0.000	0.000	0.000	0.000	0.000	0.000
Lis	1.531	0.966	1.467	0.710	0.521	0.521	0.027	0.074	0.027	0.010
Colares	0.000	1.000	0.000	N.A	0.000	0.000	0.000	0.000	0.000	0.000
Lizandro	0.000	1.000	0.000	N.A	0.000	0.000	0.000	0.000	0.000	0.000

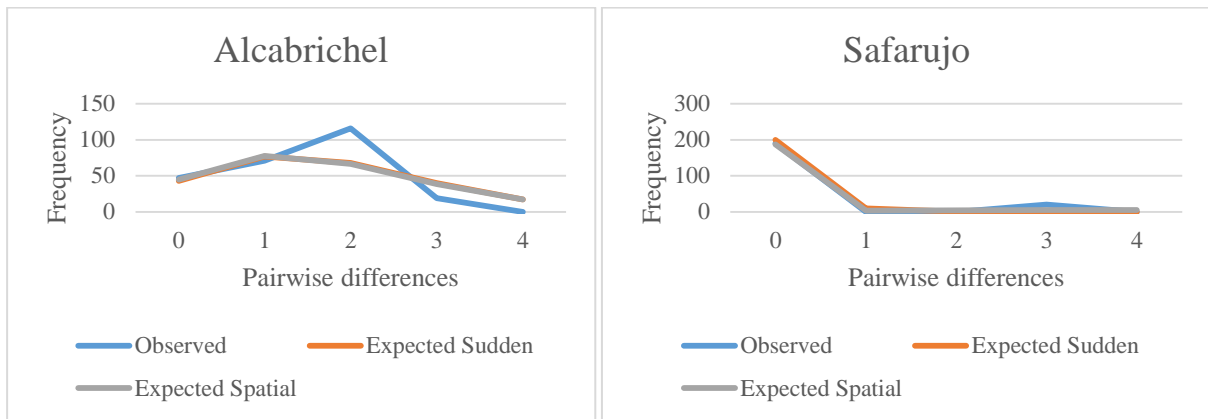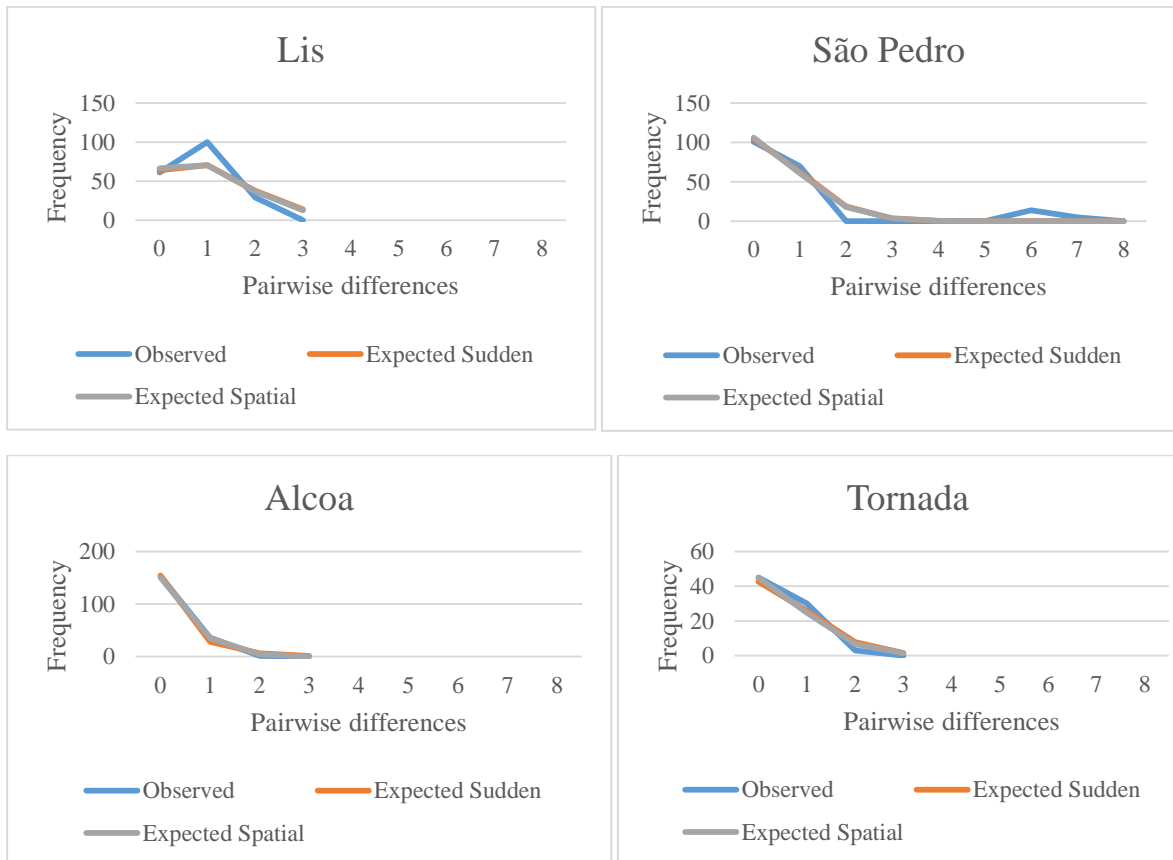# III: Observed and expected mismatch distributions under sudden and spatial expansion models

Figure 9-1: Observed and expected mismatch distributions under sudden and spatial expansion models for A. occidentale populations from Alcabrichel and Safarujo Rivers.
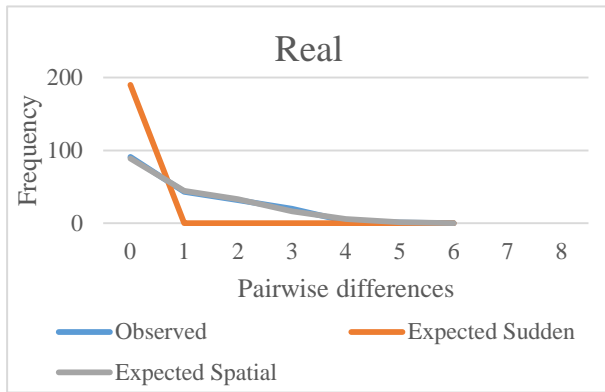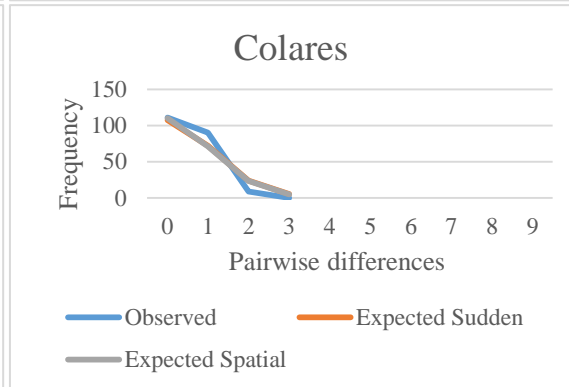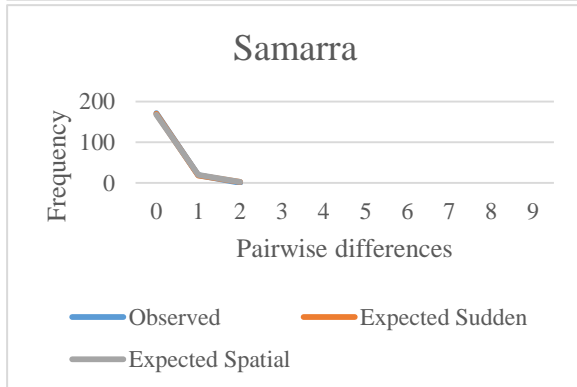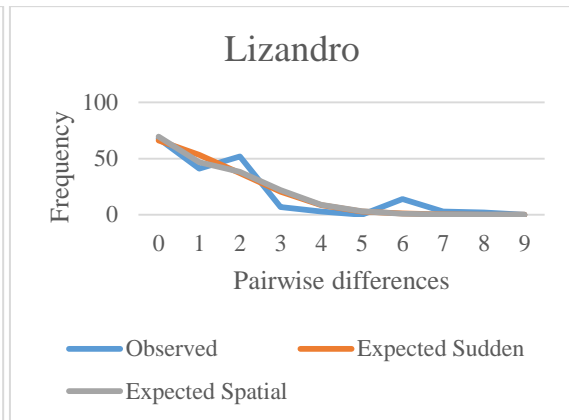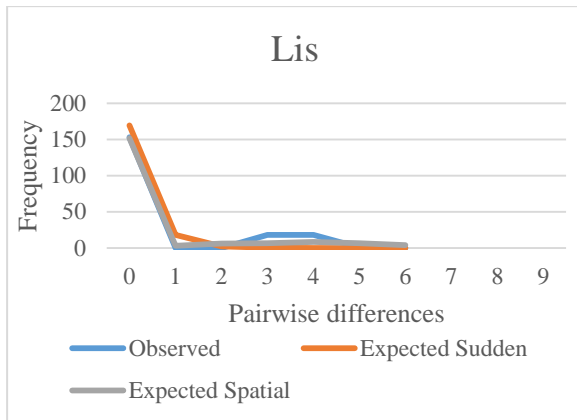
Figure 9-2: Observed and expected mismatch distributions under sudden and spatial expansion models for A. oligolepis from Lis, Lizandro, Samarra, Colares, Lage and Jamor Rivers.

Figure 9-3: Observed and expected mismatch distributions under sudden and spatial expansion models for S. pyrenaicus populations from Lis, Lizandro, Samarra, Colares, Lage and Jamor Rivers.

Figure 9-4: Observed and expected mismatch distributions under sudden and spatial expansion models for S. carolitertii population from Alcoa River.

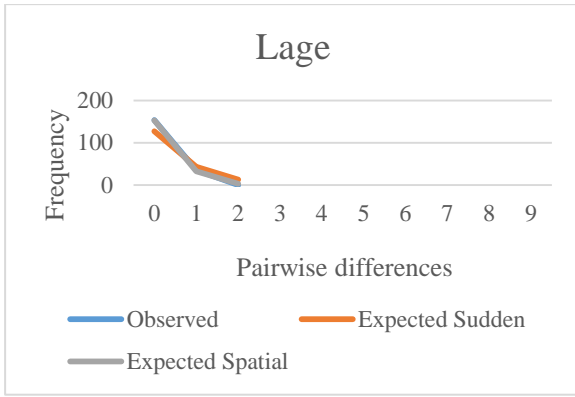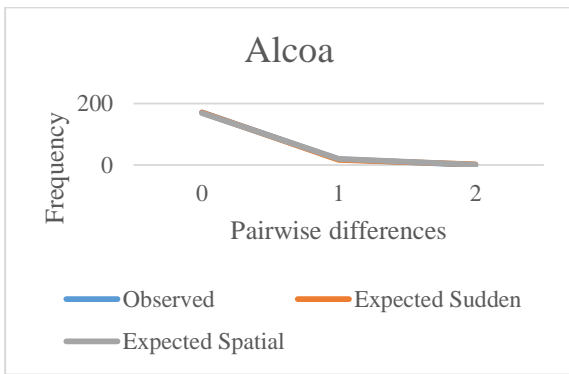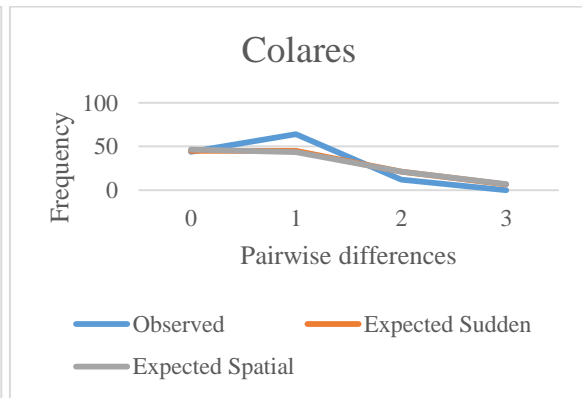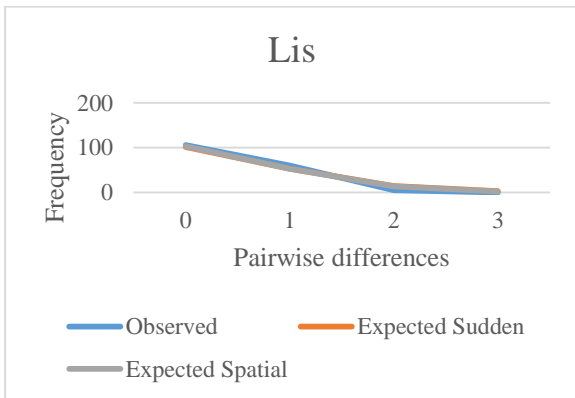*Figure 9-5: Observed and expected mismatch distributions under sudden and spatial expansion models for P. polylepis populations from Lis and Colares Rivers.*
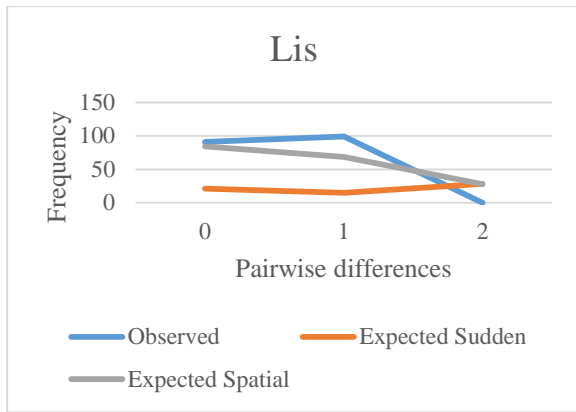
Figure 9-6: Observed and expected mismatch distributions under sudden and spatial expansion models for L. bocagei populations from Lis River.

IV: Number of migrants

Table 9-16: Absolute number of migrants (M=Nm) exchanged between I. lusitanicum populations from West region and the Neighbour river basins of Tejo and Sado.

	Lizandro	**Samarra**	**Colares**	**Lage**	**Ossos**	**Jamor**	**Tejo**	**Sado**
Lizandro								
Samarra								
Colares								
Lage								
Ossos								
Jamor								
Tejo	0,554	0,562	0,558	0,547	0,562	0,558		
Sado	0,054	0,055	0,055	0,053	0,055	0,055	0,092	

Table 9-17: Absolute number of migrants (M=Nm) exchanged between A. oligolepis populations from the West region and the Neighbour river basins of Mondego and Tejo.

	Lis	**São Pedro**	**Alcoa**	**Tornada**	**Real**	**Mondego**	**Tejo**
Lis							
São Pedro							
Alcoa							
Tornada							
Real							
Mondego	0,117	0,142	0,118	0,126	0,134		
Tejo	0,174	0,224	0,150	0,190	0,199	**2,506**	

Table 9-18: Absolute number of migrants (M=Nm) exchanged between S. pyrenaicus populations from West region and the Neighbour river basins of Tejo and Sado.

	Lis	Lizandro	Samarra	Colares	Lage	Jamor	Tejo	Sado
Lis								
Lizandro								
Samarra								
Colares								
Lage								
Jamor								
Tejo	**2,624**	**18,200**	**1,680**	**16,636**	**32,103**	**28,895**		
Sado	0,072	0,086	0,068	0,074	0,073	0,076	0,127	

Table 9-19: Absolute number of migrants (M=Nm) exchanged between P. polylepis populations from West region and the Neighbour river basins of Mondego, Tejo and Sado.

	Lis	Alcoa	Colares	Mondego	Tejo	Sado
Lis						
Alcoa						
Colares						
Mondego	**5,848**	**80,624**	0,118			
Tejo	**1,040**	0,899	**8,101**	0,640		
Sado	0,060	0,024	0,113	0,027	0,355	

Table 9-20: Absolute number of migrants (M=Nm) exchanged between L. bocagei populations from the West region and the Neighbour river basins of Mondego, Tejo and Sado.

	Alcoa	Lis	Colares	Lizandro	Mondego	Tejo	Sado
Alcoa							
Lis							
Colares							
Lizandro							
Mondego	0,870	-	**1,015**	**1,015**			
Tejo	0,405	**2,462**	**31,235**	**31,235**	**1,762**		
Sado	0,289	**1,403**	**9,514**	**9,514**	**1,103**	**4,878**	

V: Demographic and Spatial expansion

Table 9-21: Demographic and Spatial expansion of Alcoa and Tornada A.oligolepis populations: effective population sizes before (N₀) and after (N₁) a sudden expansion; effective population size prior to a spatial expansion event (N); time since the expansion (t; in years), and the migration rate (m).

Demographic expansion		
	Alcoa	Tornada
Tau	3.000	0.600
Theta 0	0.000	0.007
Theta 1	0.254	99999.000
Time (generations)	396825.400	79365.079
Time (years)	396825.400	79365.079
N_0	0.000	929.894
N_1	33689.153	13227380952.000
Spatial expansion		
	Alcoa	Tornada
Tau	0.148	0.558
Theta	0.081	0.004
N	5324.735	36916.005
T	9766.534	36916.005
M	9.390	1.354