UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL


# Assessing the impact of environmental changes in endemic freshwater fish populations through the temporal analysis of genetic diversity patterns

Joana Figueira Alves dos Anjos

Dissertação
Mestrado em Biologia Evolutiva e do Desenvolvimento
2013

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL


# Assessing the impact of environmental changes in endemic freshwater fish populations through the temporal analysis of genetic diversity patterns

Joana Figueira Alves dos Anjos
Dissertação de Mestrado coorientada por:
Professora Doutora Maria Manuela Coelho
Doutora Natacha Mesquita

Dissertação
Mestrado em Biologia Evolutiva e do Desenvolvimento
2013

## Preliminary Note

This Dissertation is presented including a scientific manuscript (chapter 2), enclosed by a general introduction (chapter 1) and final remarks (chapter 3), that will be submitted for publication in a scientific journal of the specialty, and follows the redaction rules of a standard journal.

**Table of Contents**

## Agradecimentos/Acknowledgements

A todos os que, de uma forma ou de outra, me apoiaram e incentivaram no decorrer da concretização desta Dissertação de Mestrado.

Ao Museu Nacional de História Natural e da Ciência (MNHNC; Lisboa, Portugal) pelo acesso às coleções zoológicas "Museu Bocage".

À Fundação para a Ciência e a Tecnologia (FCT) coparticipada pelo FEDER – Fundo Europeu de Desenvolvimento Regional, pelos projetos PTDC/BIA-BDE/69769/2006, PTDC/BIA-BDE/66519/2006 e PEst-OE/BIA/UI0329/2011 e pela bolsa de pós-doutoramento SFRH/BPD/14557/2003 atribuída à Natacha Mesquita.

À Autoridade Florestal Nacional (AFN) e ao Instituto da Conservação da Natureza e da Biodiversidade (ICNB) pelas licenças para o trabalho de campo efetuado.

À professora Maria Manuela Coelho por ter aceite orientar-me.

À Natacha Mesquita pela sua coorientação, pelo constante apoio, simpatia e disponibilidade para me acompanhar desde o início da parte prática ao final da parte teórica.

À Maria Drago e Catarina Mateus pelo apoio no laboratório, por todas as dúvidas que me tiraram e simpatia com que me acolheram a mim e à minha inexperiência.

Ao Filipe Ribeiro pela assistência e boa disposição no trabalho de campo.

À Tereza e à Tina, colegas da botânica, que todos os dias me fizeram companhia no laboratório sempre com bom humor e simpatia.

A todos os meus amigos, especialmente à Rita, à Joana e à Nídia, por me aturarem, pelos incansáveis incentivos e por terem estado sempre lá para mim.

Aos meus amigos de Évora, que, mesmo longe sempre se interessaram e me acompanharam neste percurso, em especial à Juliana, à Ana Cláudia, à Filipa e à Sofia.

À Raquel, à Sofia e à Joana por todos os desabafos e por tornarem o ano curricular bem mais animado!

A toda a minha família, pelo constante acompanhamento e interesse nesta Dissertação.

À minha irmã, pela amizade, apoio e constante companhia na secretária do lado, a animar os meus dias com as melhores bandas sonoras!

Finalmente, um agradecimento muito especial aos meus pais, pelo constante apoio em todas as decisões tomadas, pelo incentivo e encorajamento para continuar em frente.

## Resumo

Os impactos ambientais que afetam atualmente os ecossistemas resultam não só de alterações climáticas e desastres naturais, mas também de pressões de origem antropogénica. Presentemente, alguns dos ecossistemas mais afetados por estes impactos são os sistemas de água doce que, por sua vez, possuem uma biodiversidade bastante rica e característica que deverá ser conservada. Algumas das pressões de origem humana a que estes sistemas estão sujeitos residem na construção de barragens, extração de inertes, captação de água, poluição e introdução de espécies exóticas. Os impactos resultantes de alterações climatéricas e desastres naturais podem afetar estes ecossistemas através de alterações nos seus padrões de sazonalidade e temperatura, influenciando as relações interespecíficas.

Considerando os rios mediterrâneos, estes caracterizam-se por um ciclo anual previsível de alternância entre época de cheias (Outono-Inverno) e época seca (Verão). Dependendo da duração e intensidade dos períodos de chuva ao longo do ano, a intensidade destas perturbações hidrológicas pode variar consideravelmente entre anos, registando-se por vezes eventos de secas e cheias extremas. É também importante mencionar que a região mediterrânea pertence a um dos 25 *hotspots* mundiais, devendo ser considerada como prioritária para a conservação.

Mais especificamente, a bacia do Guadiana, habitat de diversas espécies endémicas de peixes de água doce, tem sofrido uma intensificação das pressões antropogénicas e climatéricas ao longo das últimas décadas. São de salientar, não só a construção da grande barragem do Alqueva, que iniciou o seu funcionamento em 2002/2004, mas também o período de seca severa, que ocorreu em 2004/2005, em que longas extensões dos cursos de água ficaram limitadas a alguns pegos isolados. Estas pressões terão tido um impacto negativo sobre as populações de espécies endémicas de peixes de água doce.

Com o objetivo de analisar os padrões genéticos de populações naturais, o uso de marcadores moleculares como os microssatélites é bastante vantajoso, visto que estes *loci* não são em geral afetados pela seleção, ou seja, a sua diversidade genética resulta da interação entre a deriva genética e mutações. Atualmente existem milhares de *primers* desenvolvidos para centenas de espécies de peixes.

As populações da espécie *Iberochondrostoma lemmingii* (Steindachner, 1866), da família Cyprinidae encontra-se distribuída pelas bacias do Guadiana, Tejo e Guadalquivir e pelas pequenas bacias hidrográficas de Odiel, Quarteira, Gilão e Almargem e, como consequência dos impactos supracitados sofreram importantes reduções de dimensão e ocupação ao longo de toda a sua área de distribuição, estando atualmente listadas como "Em Perigo" (EN) no Livro Vermelho dos Vertebrados de Portugal. Estas pressões deverão ter causado níveis elevados de mortalidade nas populações de *I. lemmingii*, níveis estes que deverão ter tido um impacto importante sobre a variabilidade e estrutura genética destas populações.

Com recurso a dois conjuntos de amostras diferenciados no tempo – amostras históricas depositadas nas coleções zoológicas "Museu Bocage" do Museu Nacional de História Natural e da Ciência (MNHNC; Lisboa, Portugal), constituídas por escamas secas recolhidas no fim dos anos noventa, e amostras contemporâneas capturadas em 2011-2012 – analisou-se a variação temporal da diversidade genética das populações de *I. lemmingii* para seis diferentes subsistemas da bacia do Guadiana (Caia, Degebe, Ardila, Chança, Vascão e Odeleite), utilizando-se nove microssatélites como marcadores moleculares.

O presente trabalho teve como objetivos: i) caracterizar os padrões de variação genética (diversidade e diferenciação) das populações de *I. lemmingii* para dois momentos no tempo; ii) através da comparação dos dois momentos no tempo, determinar a existência de alterações dos padrões de variação genética através do tempo; iii) relacionar as alterações observadas com a ação de possíveis fatores ambientais, e finalmente, iv) analisar a implicação dos resultados para a conservação da espécie *I. lemmingii*.

Os resultados alcançados apontam para uma diversidade genética da espécie relativamente elevada e inalterada ao longo do tempo e para um aumento geral da estrutura genética entre populações contemporâneas, tendo os maiores aumentos se verificado para as populações do Vascão e do Chança. Este aumento está provavelmente relacionado com impactos antropogénicos, tal como a construção da barragem do Alqueva, que poderá ter tido um efeito de barreira, alterando a distribuição, disponibilidade e conectividade de refúgios e reduzindo os níveis de migração destes indivíduos, aumentando a deriva genética entre populações, com consequente aumento da diferenciação entre as mesmas.

**Palavras-chave:** ADN histórico, genética populacional, *Iberochondrostoma lemmingii*, micossatélites, Península Ibérica

## Abstract

The freshwater systems contain an extremely rich fauna and, currently, are one of the most threatened ecosystems, suffering environmental impacts not only related with climacteric changes and natural hazards but also with anthropogenic pressures. One of the most stringent human pressure is the construction of dams, that act as a barrier, preventing gene flow among populations and also affecting natural processes, like the triggering of fish migration and spawning. As other Mediterranean systems, the Guadiana basin has been suffering increasing anthropogenic and climacteric pressures during the last decades, such as the construction of one of the largest reservoir in Western Europe, the Alqueva dam, and a period of severe draught in 2004/05. These two factors may have led to higher levels of mortality, as well as changes in the availability, connectivity and distribution of *refugia* and in the individual mobility, probably causing a reduction in migratory potential and consequent gene flow. These changes may have resulted in important alterations in the genetic patterns of variability and differentiation of the species that inhabit this basin. *Iberochondrostoma lemmingi* (Steindachner, 1866), the Iberian arched-mouth nase, is an endemic cyprinid species form the south of the Iberian Peninsula. This species is classified as threatened and constitutes a suitable natural model for studies addressing population genetics and conservation issues. The changes in patterns of genetic variability, population structure and demographic history were investigated comparing historical DNA samples dating back to the late 90's (before the severe draught event and the construction of the dam) with contemporary ones (2011-2012), using nine microsatellite *loci* as molecular markers. Data obtained suggested that, the genetic diversity of this population remained high and unchanged trough time, showing this species resilience to the extreme droughts. Moreover, there was a general increase in population structuring among populations, probably related with the construction of the dam.

**Keywords**: Historical DNA, Iberian Peninsula, *Iberochondrostoma lemmingii*, microsatellite *loci*, population genetics

# Chapter 1

## General Introduction

## 1.1. Environmental Impacts and Freshwater Systems

The environmental risks that the Earth faces on the present time are an outcome, not only from climacteric changes and natural hazards, but also from local human pressures. Governments in general have already established that biological diversity is crucial to make sure that ecosystems are resilient and continue to offer essential services, critical to the human well-being. These make the implementation of a global and determined response urgent, in order to stop the loss of biodiversity and the alterations in the ecosystems. (Settele *et al*., 2005)

It is known that, currently, the freshwater systems are one of the most threatened ecosystems, mainly because of their distinctive biodiversity and the essential ecosystem services they provide (Abell, 2002). One of the main human pressures that freshwater systems are subjected is the construction of dams, which are well recognized as one of the most striking and widespread intentional impacts caused by humans on natural environments (Petts, 1984). Dams influence the natural flow regime, which has had a growing recognition of influencing certain natural processes like the triggering of fish migration and spawning and the rejuvenation of floodplains (Bunn & Arthington, 2002). Other human impacts are related with water and sand abstraction, pollution and the introduction of alien species (Boix *et al*., 2010),

Concerning climacteric changes and natural hazards, on a local scale the alterations in seasonality and mean temperatures may modify local interactions of species (Burgmer *et al*., 2006). Whereas, on a larger scale, climate variables such as temperature or evapotranspiration are recognized as strong predictors of terrestrial and aquatic diversity (Hawkins *et al.,* 2003), causing changes on the levels of diversity of the communities.

The Iberian Peninsula is considered a conservation hotspot - places that hold especially high numbers of endemic species - included in the Mediterranean Basin Biodiversity Hotspot. This particular hotspot presents conservation priorities, since it has an exceptional concentration of endemic species undergoing great habitat losses. Moreover, the Iberian river fish fauna constitutes a natural asset largely dominated by native endemic species, with a considerable number of species presenting highly restricted distributions and conservation concerns. During the last decades freshwater systems in general, and in the Iberian Peninsula in particular, have been subject to increased human induced changes (e.g., pollution, dam construction and introduction of alien species) that represent important threats for the endemic species survival (Myers *et al*., 2000).

Considering the Mediterranean rivers, these are characterized by an annual cycle of seasonally predictable flooding (late autumn-winter) and drought (late summer-early autumn) events, which vary considerably in intensity among years in accordance with the rainfall levels. Depending on the duration and intensity of the warm and dry summers and of the wet winters, hydrologic disturbances, such as extreme droughts and floods, can occur (Coelho *et al*., 1995; Magalhães, 2002).

Especially when considering dry years, the drought period is considerably aggravated in length and intensity, with the dry up of large river sections and water restricted to the deepest pools during summer, exerting strong pressure on natural populations and community structure regulation. In accordance with the general trend observed in southern European basins, recent studies have shown a decrease in annual streamflow in most of the Iberian basins (especially those in the southern regions) steamed by climate variability, in particular decreasing precipitation. In addition, changes in the seasonality of the characteristic annual river regime have been observed, resulting from dam regulation and water management strategies. The decrease of streamflows during the second half of the twentieth century in the Iberian Peninsula may be accelerated in coming decades with the predicted climatic changes, including a generalized decrease in precipitation and increased evapotranspiration induced by higher temperatures (Lorenzo-Lacruz *et al*., 2012)

These disturbances on habitats cause great impacts on the local natural populations, namely on the populations of fishes. Droughts, in view of its effect on the hydrologic continuity/discontinuity (Lake, 2003), act like selection forces over fish populations, increasing the local density of individuals, changing the food resources available and affecting the interspecific interactions of the community of fishes (Magoulick & Kobza, 2003; Davey & Kelly, 2007). Further, droughts will also have an impact on the reproductive success and on the mortality levels of the individuals, which will cause changes in the size of the communities and in their specific richness. These alterations will affect mainly the native species, seen that they are less tolerant to disturbances on their habitat (Lake, 2003; Matthews & Marsh-Matthews, 2004; Magalhães *et al*., 2007).

All these factors combined will cause a reduction on population size, which will affect the temporal stability of allele frequencies in natural populations (Yongfeng & Jianwey, 2010). The loss of genetic variability can lead to an inbreeding depression and, consequently, debilitate the capacity of adaptation of the populations (Nielsen *et al*. 1999).

## 1.2. Natural History Collections

Natural history collections (NHC) are an extremely important resource for diverse areas of knowledge, as they constitute the most thorough and organized source of information about past and present Earth's diversity, providing irreplaceable evidence of long-term historical trends and providing solid grounds for researchers to better understand the environmental risks and to make scenarios into the future (Winker, 2004; Wissenschaftsrat, 2011). NHC are now widely recognized as playing a crucial role in biological and earth sciences and making innumerable contributions to society in areas as divergent as nature conservation, global climate change, public health and safety, bioprospecting of pharmaceuticals, forensic science and national security (Lane, 1996; Suarez *et al*., 2004; Lister *et al*., 2011)

Decades ago they were frequently used in morphology studies, and more recently they have been increasingly used in molecular biology studies (Chakraborty *et al*. 2006), seen that, with the continuous development of new techniques, new collections have been created alongside the traditional NHC, such as collections of tissue and DNA (Payne *et al*., 2002; Sheldon, 2006). Also, over the years, with the increase of molecular studies, natural history and other research institutions worldwide have registered an exponential growth of entries in their genetic resources' collections (Payne *et al*., 2002; Sheldon, 2006; Wandeler *et al*., 2007; Wolinski, 2010). One of the biggest advantages on the use of these collections in molecular biology is the possibility to build a genetic data base of great use to scientific research (Payne *et al*. 2002), furthermore, it is also important to note that, a systematic sampling enables the collection of data of the genetic variability of the populations through time, and consequently, obtaining a measure of the genetic "health" of those populations (Sheldon, 2006). Finally, and in view of the increasing difficulties for sampling natural populations, NHC are also very useful in studies regarding rare or extinct specimens, as well as in investigations addressing recent evolutionary changes, in view of the temporal samples available.

The major issue concerning the use of museum collections in molecular biology studies resides in the methods of fixation and preservation of the samples, seeing that these were not initially collected with the purpose of extracting DNA, making it more difficult to isolate it, to control its quantity and quality and, consequently, to determine the amount of DNA template to use in each PCR reaction (Nielsen *et al*., 1999).

However, the use of simply dried samples (e.g. dried fish scales), where no chemicals were used for fixation and preservation, makes these issues manageable through the use of optimized extraction and amplification protocols and also, through the implementation of preventive measures to avoid contamination: a bleach treatment on the extraction area between every set of extractions; the separation in time of DNA extraction and PCR amplifications of historic and contemporary samples; the use of negative controls (where no template has been added) to check for cross-and/or aerosol contaminations; and the use of a UV hood to prepare the PCR reactions. Further, a spectrophotometer is used to access the quantity and quality of the DNA extracted and to determine the amount of template DNA to use in each PCR reaction.

More specifically, regarding the use of dried fish scales, Nielsen *et al*. (1997) has showed that DNA extracted from scale samples is a useful source of information for tracking genetic changes in fish populations, seen that it can be used to analyze the impact of environmental changes such as epidemics, the building of dams and pollution, on the genetic structure of fish populations. As an example of this, there are other studies, such as Iwamoto *et al*. (2012), where salmon scale samples were used successfully to perform a temporal study where two moments in time are compared and their genetic variability is analyzed. In

conclusion, the study of genetic patterns using DNA from old scale samples adds a historical dimension to population genetic studies of fishes (Nielsen *et al*., 1997)

### 1.3. Microsatellites in Conservation Genetics

Microsatellites are repeating sequences of two to six base pairs of non-coding DNA. They present high rates of allelic polymorphism and it is considered that the majority of these *loci* suffer neutral evolution. In other words, they are not affected by selection, which means that their genetic diversity results from the interaction between genetic drift and mutations (Moran, 2002). However, there is the possibility that a microsatellite *locus* is in linkage with a different DNA region under selection, which causes the action of selection over the neighboring microsatellite *locus* as well (Chistiakov *et al*., 2006). The microsatellite *loci* can be found in coding and non-coding areas of the DNA (Toth *et al*., 2000), although, they are relatively rare in coding areas, and exist in only 9-15% of the vertebrate genome (e.g., Moran, 1993; Jurka & Pethiygoda, 1995).

All these characteristics make microsatellites one the most common markers used in conservation genetics studies, namely, in cases of temporal analysis of the genetic variability (e.g. Hutchinson *et al*., 2003, Holbrook *et al*., 2012, Iwamoto *et al*., 2012). More specifically, they constitute a tool of major importance in fishes' genetic studies (Wright & Bentzen, 1994), and presently there are thousands of primers pairs available for hundreds of fish species (Moran, 2002). However, one of the difficulties associated with the use of microsatellites is related with the development of species-specific primers, seen that, primers developed for a certain taxon may not work properly in a different/related *taxa*. However, the PCR (Polymerase Chain Reaction) amplification among related species is common (Chistiakov *et al*., 2006). For the Cyprinidae family and particularly for the Leuciscinae subfamily, there are numerous studies for different species in which primers were developed (e.g. Mesquita *et al*., 2003; Larno *et al*., 2005; Pala & Coelho, 2005; Vyskocilová *et al*., 2007; Muenzel *et al*., 2007).

Further, microsatellites also constitute a suitable marker for conservation genetic studies, seen that, it is only required a small amount of tissue/DNA, allowing the use of non-invasive sampling techniques and the restitution of the individual to its habitat, minimizing the disturbance of the natural populations.

The most common measures to analyze the levels of genetic diversity in natural populations are the heterozygosity (proportion of heterozygous individuals in a population), allelic diversity (number of alleles per *locus* in a population) and the proportion of polymorphic *loci* (Pujolar *et al*. 2005). Alteration in the expected values of heterozygosity are also used to test if the population was exposed to a severe bottleneck (Leberg, 1992), which consists in a severe decrease on population size and genetic variability (Kramer & Samelle, 2008), leading to a reduction on the occurrence of rare alleles on the population (Denniston, 1978).

Other important factor to understand and characterize populations is the analysis of their population structure, which is usually measured trough F-statistics (Wright, 1951). F-statistics allows to quantify the genetic differentiation among subpopulations and consequently to classify individuals in a sample to populations. By analyzing the population structure among different locations of the distribution area of a species, we can better understand their genetic differences and the levels of gene flow among them.

### 1.4. Characterization of the Species - *Iberochondrostoma lemmingii*

The species *Iberochondrostoma lemmingii* (Steindachner, 1866) belongs to the family Cyprinidae, subfamily Leuciscinae, it is endemic of the Iberian Peninsula (Almaça, 1995) and can be found, more specifically, at the large Guadiana, Tejo and Guadalquivir drainages and at the small basins of Odiel, Quarteira, Gilão and Almargem (Doadrio *et al*., 2001; Cabral *et al*., 2005).

Currently, due to the major reductions in area, quality and occupation of the available habitats, it only comprises an area of 150 $Km^2$ and the conservation status of this species is considered "endangered" in Portugal (Cabral *et al*. 2005). In Spain, the remaining populations are currently classified as "vulnerable" (Doadrio *et al*., 2001).

Morphologically, this species is characterized by a long body moderately flat, normally not surpassing the 15 cm of length. They have an arched mouth without horny layer on the lower lip, well developed and upward oriented process from premaxilla, dentary with a coronoid process posteriorly orientated and with a log anterior process, the length of the ethmoides is greater than the width. They have 46-60 canaliculated scales on the lateral line, 11-12 scales above the lateral line, 5-6 scales below the lateral line, 7-8 branched rays in the ventral 6–7 branched rays in the dorsal, 6–7 branched rays in the anal, 6–5/5 pharyngeal teeth, 24–27 gill rakers on the first branchial arch and the upper branch of ceratobranchial enlarged (Robalo *et al.* 2007). The coloration of the body is relatively dark, with small dots spread across the flanks (Collares-Pereira *et al*., 2007).

This species can be found in permanent or intermittent rivers and streams with abundant aquatic vegetation (Cabral *et al*., 2005). Its diet is mainly composed by algae and zooplankton and the breeding season occurs between April and May (Kottletat & Freyhof, 2007).

**Figure 1** ***Iberchondrosoma lemmingii*****, photo by FRibeiro**

The major threats to *I. lemmingii* populations are habitat degradation, due to dam construction, alterations in the natural flow regime, water abstraction, sand extraction, water pollution and the introduction of exotic species (Collares-Pereira *et al*., 2000). All this environmental concerns are particularly evident in the Guadiana basin, one of the largest Iberian river basins, with high diversity levels of endemic freshwater fish species and various listed as threatened. During the last decades the Guadiana basin has suffered increasing anthropogenic pressures, including the construction in the 1990s of the largest reservoir in Western Europe, the Alqueva dam, which started functioning in 2002-2004. Furthermore, a period of severe drought occurred in 2004/05, with extensive drought of the river courses during the dry season, probably leading to high mortality levels among the native freshwater fish populations and consequent changes in genetic variation patterns

On the Guadiana hydrographic basin, this species is sympatric with other Leuciscinae cyprinids: *Pseudochondrostoma willkommii* (Steindachner, 1866), *Squalius alburnoides* (Steindachner, 1866), *Squalius pyrenaicus* (Günter, 1868) and *Anaecypris hispanica* (Steindachner, 1866).

## 1.5. Thesis Aim

We aim to further understand the environmental risks that the Iberian endemic freshwater fish populations are, and will be facing in view of the observed climacteric changes, intensification of the natural hazards related with the characteristic annual river regime and human-driven environmental changes. To accomplish it we plan to study the genetic variability

patterns of the species *I. lemmingii,* which has been subjected to several environmental changes through time, some of human origin and some of natural origin.

For this purpose we aim to:

(1) Characterize the patterns of genetic variation (diversity and differentiation) of the *I. lemmingii* populations for two different moments in time. One of the data sets was obtained from the zoological collections "Museu Bocage" of the *Museu Nacional de História Natural e da Ciência* (MNHNC; Lisbon, Portugal), consisting in dried scales sampled in different areas of the Guadiana Basin between 1997 and 1999, the other data set consisted in fresh fin tips, collected between 2011 and 2012 through the use of electrofishing and stored in alcohol. For both moments in time, the subsystems sampled throughout the Guadiana basin were: Caia, Degebe, Ardila, Chança, Vascão and Odeleite;

(2) Determine the existence of alterations trough time in the patterns of population genetic variation by comparing the results obtained for the two different moments in time;

(3) Relate the observed changes with the action of possible environmental factors; and finally;

(4) Analyze the implication of the results for the conservation of the species *I. lemmingii*.

### 1.6. References


1. Abell R (2005). Conservation Biology for the Biodiversity Crisis: a Freshwater Follow-up. *Conservation Biology*, **16**:1435-1437.
2. Almaça C (1995). Freshwater fish and their conservation in Portugal. *Biological Conservation*, **72**:125-127.
3. Boix D, Garcia-Berthou E, Gascon S, Benejam L, Tornes E, Sala J, Benito J, Munne A, Sola C, Sabater S (2010). Response of community structure to sustained drought in Mediterranean rivers. *Journal of Hydrology,* **383**:135–146.
4. Bunn SE, Arthington AH (2002). Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management*, **30**(4):497-507.
5. Burgmer T, Hillebrand H, Pfenninger M (2006). Effects of climate-driven temperature changes on the diversity of freshwater macroinvertebrates. *Oecologia*, **151**:93-103.
6. Cabral MJ (Coord.), Almeida J, Almeida PR, Dellinger T, Ferrand de Ameida N, Oliveira ME, Palmeirim, JM, Queiroz AI, Rogado L, Santos-Reis M (eds) (2005). *Livro Vermelho dos Vertebrados de Portugal*. Instituto da Conservação da Natureza. Lisboa. 660pp.

7. Chakraborty A, Sakai M, Iwatsuki Y (2006). Museum fish specimens and molecular taxonomy: A comparative study on DNA extraction protocols and preservation techniques. *Journal of Applied Ichthyology*, **22**:160-166.
8. Chistiakov DA, Helleman B, Volckaert FAM (2006). Microsatellites and their genomic distribution, evolution, function and applications: a review with special reference to fish genetics. *Aquaculture,* **255**: 1–29.
9. Coelho MM, Brito RM, Pacheco TR, Figueiredo D, Pires AM (1995). Genetic variation and divergence of *Leuciscus pyrenaicus* and *L. carolitertii* (pysces, Cyprinidae). *Journal of Fish Biology*, **47**(Suppl. A):243-258.
10. Collares-Pereira MJ, Cowx IG, Ribeiro F, Rodrigues JA, 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.
11. Collares-Pereira MJ, Filipe AF, Moreira da Costa L (2007). *Os Peixes do Rio Guadiana: Que Futuro? Guia de Peixes do Guadiana Português*. Edições Cosmos, Alpiarça, 294 pp.
12. Davey AJH, Kelly DJ (2007). Fish community responses to drying disturbances in an intermittent stream: a landscape perspective. *Freshwater Biology*, **52**:1719–1733.
13. Denniston CD (1978). Small population size and genetic diversity: implications for endangered species. In: *Endangered Birds: Management techniques for Preserving Threatened Species* (eds. Temple SA), pp. 281–289. University of Wisconsin Press, Madison.
14. Doadrio I (2001). Atlas y libro rojo de los peces continentales de España. Dirección General de Conservación de la Naturaleza. Museo Nacional de Ciencias Naturales. Madrid. 364pp.
15. Hawkins BA, Field R, Cornell HV, Currie DJ, Guégan JF, Kaufman DM, Kerr JT, Mitellbach GG, Oberdorff T, O'Brien EM, Porter EE, Turner JRG (2003). Energy, Water, and broad-scale geographic patterns of species richness. *Ecology*, **84**(12):3105-3117.
16. Holbrook JD, DeYoung RW, Tewes ME, Young JH (2012). Demographic history of an elusive carnivore: using museums to inform management. *Evolutionary Applications*, **5**:619-628.
17. Hutchinson WF, Van Oosterhout C, Rogers SI, Carvalho GR (2003). Temporal analysis of archived samples indicates marked genetic changes in declining North Sea cod (*Gadus morhua*). *The Royal Society*, **270**:2125-2132.

18. Iwamoto EM, Myers JM, Gustafson RG (2012). Resurrecting an extinct salmon evolutionarily significant unit: archived scales, historical DNA and implications for restoration. *Molecular Ecology*, **21:**1567–1582.
19. Jurka J, Pethiyagoda C (1995). Simple repetitive DNA sequences from primates: compilation and analysis. *Journal of Molecular Evolution*, **40**:120–126.
20. Kottelat M, Freyhof J (2007). *Handbook of European freshwater fishes*. Kottelat, Cornol, Swirzterland and Freyhof, Berlin, Germany. 646pp.
21. Kramer A, Sarnelle O (2008). Limits to genetic bottlenecks and founder effects imposed by the Allee effect. *Oecologia*, **157**:561-569.
22. Lake PS (2003). Ecological effects of perturbation by drought in flowing waters. *Freshwater Biology*, **48**:1161–1172.
23. Lane M (1996). Roles of Natural History Collections. *Annals of the Missouri Botanical Garden*, **83**:536-545.
24. Larno V, Launey S, Devaux Alain, Laroche J (2005). Isolation and characterization of microsatellite *loci* from chub Leuciscus cephalus (Pisces: Cyprinidae). *Molecular Ecology Notes*, **5**:752-754.
25. Leberg PL (1992). Effects of population bottlenecks on genetic diversity as measured by allozyme electrophoresis. *Evolution* **46**:477-494.
26. Lister AM, Climate Change Research Group (2011). Natural history collections as sources of long-term data sets. *Trends in Ecology and Evolution*, **26**:153-154.
27. Lorenzo-Lacruz J, Vicente-Serrano SM, López-Moreno JI, Morán-Tejeda E, Zabalza J (2012). Recent trends in Iberian streamflow (1945-2005). *Journal of Hydrology*, **414-415**:463-475.
28. Magalhães MF (2002) Gradientes espaciais, variabilidade ambiental e estrutura de comunidades piscicolas em rios de tipo Mediterrânico. PhD Thesis, university of Lisbon, Portugal.
29. Magalhães MF, Beja P, Schlosser IJ, Collares-Pereira MJ (2007). Effects of multi-year droughts on fish assemblages of seasonally drying Mediterranean streams. *Freshwater Biology*, **52**:1494–1510.
30. Magoulick DD, Kobza RM (2003). The role of *refugia* for fishes during drought: a review and synthesis. *Freshwater Biology*, **48**:1186–1198.
31. Matthews WJ, Marsh-Matthews E (2003). Effects of drought on fish across axes of space, time and ecological complexity. *Freshwater Biology*, **48**:1232–1253.
32. Mesquita N, Cunha C, Hänfling B, Carvalho GR, Zé-Zé L, Tenreiro R, Coelho MM (2003). Isolation and characterization of polymorphic microsatellite *loci* in the endangered freshwater fish *Squalius aradensis* (Cyprinidae). *Molecular Ecology Notes*, **3**:572-574.

33. Moran C (1993). Microsatellite repeats in pig (*Sus domestica*) and chicken (*Gallus domesticus*) genomes. *Journal of Heredity*, **84**:274–280.
34. Moran P (2002). Current conservation genetics: building an ecological approach to the synthesis of molecular and quantitative genetic methods. *Ecology of Freshwater Fish*, **11**:30-55.
35. Muenzel FM, Sanetra M, Sakburger W, Meyer A (2007). Microsatellites from the vairone *Leuciscus souffia* (Pisces: Cyprinidae) and their application to closely related species. *Molecular Ecology Notes*, **7**:1048-1050.
36. Myers N, Mittermeier RA, Mittermeier CG, Fonseca GAB, Kent J (2000). Biodiversity hotspots for conservation priorities. *Nature*, **403**:853-858.
37. Nielsen EE, Hansen MM, Loeschcke V (1997). Analysis of microsatellite DNA from old scale samples of Atlantic salmon S*almo salar*: a comparison of genetic composition over 60 years. *Molecular Ecology*, **6**:487-492.
38. Nielsen EE, Hansen MM, Loeschcke V (1999). Analysis of DNA from old scale samples: technical aspects, applications and perspectives for conservation. *Hereditas,* **130**:264-276.
39. Pala I, Coelho MM (2005). Contrasting views over a hybrid complex: Between speciation and evolutionary dead-end. *Gene*, **347**:283-294.
40. Payne RB, Sorenson MD (2003). Museum Collections as Sources of Genetic Data. *Bonner zoologische Beiträge*. **51**:97-104.
41. Petts GE (1984). Impounded Rivers: Perspectives for ecological management. Wiley, Chichester, UK.
42. Pujolar JM, Maes GE, Vancoillie C, Volckaert FAM, (2005). Growth rate correlates to individual heterozygosity in European eel, *Anguilla anguilla* L. *Evolution* **59**:189–199.
43. Robalo JI, Almada VC, 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.
44. Settele J, Hammen V, Hulme P, Karlson U, Klotz S, Kotarac M, Kunin W, Marion G, O'Connor M, Petanidou T, Peterson K, Potts S, Pritchard H, Pysek P, Rounsevell M, Spangenberg J, Steffan-Dewenter I, Skyes M, Vighi M, Zobel M, Kühn I (2005). Alarm: Assessing Large-scale environmental Risks for biodiversity with tested Methods. *GAIA – Ecological Perspectives for Science and Society*, **14**:60-72.
45. Sheldon FH (2006) Molecular Collections for basic research: Museums, Methods, and Mortality. In: Proceedings of the International Conference on In-situ and Ex-situ Biodiversity Conservation in the New Millennium (eds. Jaacob Z, Moo-Tan S, Jorath S), pp. 331-346. Sabah, Malásia.

46. Suarez AV, Tsutsui ND (2004). The value of museum collections for research and society. *Bioscience*, **54**:66–74.
47. Toth G, Gaspari Z, Jurka J (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research*, **10**:967–981.
48. Vyskocilová M, Simková A, Martin JF (2007). Isolation and characterization of microsatellites in *Leuciscus cephalus* (Cypriniformes, Cyprinidae) and cross-species amplification within the family Cyprinidae. *Molecular Ecology Notes*, **7**:1150-1154.
49. Wandeler P, Hoeck PEA, Keller LF (2007). Back to the future: museum specimens in population genetics. *Trends in Ecology and Evolution*, **22:**634-642.
50. Winker K (2004). Natural History Museums in a Postbiodiversity Era. *BioScience*, **54**:455-459.
51. Wissenschaftsrat (2011). Recommendations on Scientific Collections as Research Infrastructures. Berlin.
52. Wolinsky H (2010) History in a single hair. EMBO reports, **11**:427-430.
53. Wright JM, P Bentzen, (1994). Microsatellites: genetic markers of the future. *Reviews in Fish Biology and Fisheries*, **4**:384–388.
54. Wright S (1951). The genetical structure of populations. *Annals of Eugenetics*, **15**:323-354.
55. Yongfeng H, Jianwei W (2010). Temporal variation in genetic structure of the Chinese Rare Minnow (Gobiocypris rarus) in its type locality revealed by microsatellite markers. *Biochemical Genetics,* **48**:312-325.
56. Ywamoto EM, Myers JM, Gustafson RG (2012). Resurrecting an extinct salmon evolutionary significant unit: archived scales, historical DNA and implications for restoration. *Molecular Ecology*, **21**:1567-1582.

# Chapter 2

## Assessing the Impact of Environmental Changes in Endemic Freshwater Fish Populations Through the Temporal Analysis of Genetic Diversity Patterns

# Assessing the impact of environmental changes in endemic freshwater fish populations through the temporal analysis of genetic diversity patterns

J. Anjos*[1,2], M. J. Alves[1,2], M. M. Coelho[2], N. Mesquita[1,2]

[1]Museu Nacional de História Natural e da Ciência, Universidade de Lisboa, Rua da Escola Politécnica 56/58, 1250-102 Lisboa, Portugal

[2]Centro de Biologia Ambiental, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

*Corresponding author. E-mail: joananjos@hotmail.com

## 2.1. Abstract

Over the last decades, Iberian river fish fauna included in the Mediterranean Basin Biodiversity Hotspot, have suffered increasing pressures, due to anthropogenic causes, like the construction of dams, water and sand abstraction, pollution and introduction of alien species, and also due to natural causes like the alterations in seasonality and mean temperatures. These factors are probably negatively influencing the levels of gene flow and mortality of these species. In particular, the Guadiana Basin has suffered increasing pressures during the last decades, including the construction of the Alqueva Dam in the 1990s, one of the largest reservoirs in Western Europe, which started functioning in 2002-2004. Further, the populations in Guadiana have suffered a period of severe draught in 2004/05. These two factors may have led to high levels of mortality for *Iberochondrostoma lemmingi* (Steindachner, 1866), and consequently to important changes in the genetic patterns of variability and differentiation of this species. To access the temporal genetic variability patterns of this species, microsatellite analysis was applied to scale samples collected around 15 years ago (late 90's) as well as to fresh fin samples collected in 2011-2012. These two data sets over time were compared using different measures of genetic diversity, population structure and demographic analysis. Our finding suggest that, although the genetic diversity of the species has been kept high and unchanged over time, their population structure has been intensified through time, mainly for the populations of Vascão and Chança. This increased differentiation is probably related with

anthropogenic impacts, mainly the construction of the Alqueva dam, which may have caused the reduction of the gene flow among populations, acting as a barrier.

**Keywords**: conservation genetics, historical DNA, Iberian Peninsula, *Iberochondrostoma lemmingii*, microsatellite *loci*, population genetics

## 2.2. Introduction

In the present days, it is well known by Governments that biological diversity is essential to maintain the ecosystems, which are critical to the human well-being. Also, these ecosystems are currently facing environmental risks, not only from climacteric changes and natural hazards, but also from local human pressures (Settele *et al*., 2005).

More specifically, concerning the freshwater systems, they possess a distinctive biodiversity and provide essential ecosystem services, making them one of the most threatened ecosystems (Abell, 2002). Moreover, freshwater systems have been suffering increased human pressures, representing important threats to the endemic species survival (Myers *et al*., 2000) and also climacteric changes and natural hazards. One of the main human pressures is the construction of dams, which is well recognized as one of the most crucial and widespread intentional impacts caused by humans on natural environments (Petts, 1984). Dams influence the natural flow regime, which has had a growing recognition of influencing various natural processes like the triggering of fish migration and spawning and the rejuvenation of floodplains (Bunn & Arthington, 2002). Other human impacts are related with the water and sand abstraction, pollution and the introduction of alien species (Boix *et al*., 2010). Some of the main climacteric changes and natural hazards are related with the alterations in seasonality and mean temperatures and its' effects on the interactions of species (Burgmer *et al*., 2006) also, on a larger scale, climate variables such as temperature or evapotranspiration cause changes on the levels of diversity and are very strong predictors of terrestrial and aquatic diversity (Hawkins *et al.,* 2003).

Furthermore, some of the richest freshwater systems are present in the Iberian Peninsula which belongs to one of the 25 biodiversity hotspots for conservation priorities (Myers *et al*., 2000). The Iberian fish fauna is largely dominated by native endemic species, with various presenting limited distributions and conservation concerns.

Concerning the Mediterranean rivers, these are characterized by an annual cycle of seasonally predictable flooding (late autumn-winter) and drought (late summer-early autumn) events, which vary considerably in intensity among years according to the rainfall levels. Depending on the duration and intensity of the warm and dry summers and of the wet winters, extreme hydrologic disturbances, such as severe droughts, can occur.

These disturbances on habitats cause great impacts on the local populations, namely on the populations of native fish species. For example, droughts act like selection forces over these populations, influencing population density, the availability of food resources, interspecific interactions, reproductive success and mortality levels, ultimately causing changes in the size and specific richness of the communities (Magoulick & Kobza, 2003; Davey & Kelly, 2007). Also, the reduction on population size will affect the temporal stability of allele frequencies in natural populations (Yongfeng & Jianwey, 2010), meaning that the study of genetic characteristics using highly polymorphic molecular markers will access these changes in allelic frequencies.

An important repository to study these impacts resides in Natural History collections which are currently being rediscovered as critical research infrastructures (Wissenschaftsrat, 2011; Lane, 1996; Suarez *et al*., 2004; Winker, 2004; Lister, 2011). They constitute a permanent record, tangible and timeless, of the biodiversity of different biogeographic areas, and are being used as fundamental resources in different research areas, including evolutionary biology, conservation of biodiversity and management of wild life, toxicology, epidemiology and forensic sciences (reviewed by Stuebing & Wong, 2000). They represent a crucial source of data in studies of rare or even extinct species, and to access recent evolutionary changes. In the last few years they have been increasingly used in molecular biology studies (Chakraborty *et al*. 2006). However, the use of museum specimens to recover historical DNA faces methodological difficulties, brought by the methods of fixation and preservation of samples, as these were not initially collected and processed with the purpose of DNA extraction. The use of dried scale samples, which were not fixated or preserved, makes it possible to extract and amplify it with a relative success, through the use of optimized protocols and preventive measures to avoid contamination. Still, DNA in dried historical scales is in general degraded, with most fragments smaller than 400 base pairs, which means that only microsatellite *loci* with alleles smaller than 200-250 base pairs should be chosen for amplification, since longer fragments are more difficult to amplify, especially in scale samples with more than 10 years old (Nielsen *et al*., 1999).

In addition, Nielsen *et al*. (1997) has showed that DNA extracted from scale samples is a useful source of information for tracking genetic changes in fish populations, seen that it can be used to analyze the impact of environmental changes, such as epidemics, the building of dams and pollution, on the genetic structure of fish populations As an example of this, where scales were used successfully, is the Iwamoto *et al*. (2012) study, where dried salmon scales were analyzed and compared with recent ones, in order to test the genetic variability of the species. In conclusion, the study of microsatellite variation in DNA from old scale samples adds a historical dimension to population genetic studies of fishes (Nielsen *et al*., 1997).

The *Iberchondrostoma lemmingii* is a iberian endemic freshwater fish species (Almaça, 1995) present in the large Guadiana, Tejo and Guadalquivir drainages and in the small basins of

Odiel, Quarteira, Gilão and Almargem (Doadrio *et al*., 2001; Cabral *et al*., 2005). In Portugal the species is currently classified as endangered by the Red Book of Vertebrates of Portugal (Cabral *et al*., 2005) due to major occupation and a reduced quality of its habitat. The main factors threatening this species' populations and promoting their habitat degradation are the construction of dams, alteration of the natural flow regime, water and sand abstraction, poor water quality and the introduction of exotic species (Collares-Pereira *et al*., 2000).

In particular in the Guadiana basin, during the last decades, *I. lemmingii* populations have been suffering increasing anthropogenic pressures, including the construction in the 1990s of the largest reservoir in Western Europe, the Alqueva dam which started functioning in 2002-2004. In addition, a period of severe draught occurred in 2004/05, with extensive drought of the river courses during the dry season, which may have led to an increase in mortality levels and consequent changes in genetic variation patterns. Changes in population size and allelic richness are expected, as well as an increase in population genetic differentiation, resulting from increased isolation among populations. In addition, seen that native species, such as *I. lemmingii,* are less tolerant to disturbances on their habitat (Lake, 2003; Matthews & Marsh-Matthews, 2004; Magalhães *et al*., 2007), the effect of environmental changes could be further increased.

All these make *I. lemmingii* an ideal natural model to further understand the impact of environmental changes in shaping population genetic patterns through time.

Various authors have previously investigated population genetics in different endangered Iberian endemic cyprinids (e.g.: Salgueiro *et al*., 2003; Mesquita *et al*., 2005; Sousa *et al*., 2008; Sousa *et al*., 2010). However, this is the first time that two different moments in time are analyzed and compared for an endemic cyprinid species, with the purpose of understanding genetic variation through time. These kind of temporal studies are relatively rare, seeing that they rely on the availability of genetic material pre-dating the effects of the environmental factors (e.g. museum specimens, subfossil and permafrost samples) (Martínez-Cruz, *et al*. 2007).

In the current study, the temporal changes in genetic variation for the species *I. lemmingi* was evaluated, accessing genetic diversity measures, population structure and demographic analysis. To accomplish this, nine microsatellite *loci* were used as molecular markers and two comparable data sets from six different locations along the Guadiana basin, subsystems Caia, Degebe, Ardila, Chança, Vascão and Odeleite, were analyzed. The historical data set included dried scale samples from museum collections collected by the end of the 90's; and the contemporary data set included fresh fin samples caught in 2011-2012. We aim to i) characterize the patterns of genetic variation (diversity and differentiation) of the *I. lemmingii* populations for two different moments in time; ii) determine the existence of alterations trough time in the patterns of population genetic variation by comparing the results obtained for the

two different moments in time; iii) relate the observed changes with the action of possible environmental factors; and finally, iv) analyze the implication of the results for the conservation of the species *I. lemmingii*.

## 2.3. Material and Methods

### 2.3.1 Sampling and DNA Extraction

Historical samples consisted of dried scales on notebooks, collected in the Guadiana basin between 1997 and 1999 within ecology studies, with the purpose of individual age determination. The samples not used for such were deposited afterwards in the zoological collections "Museu Bocage" of the Museu Nacional de História Natural e da Ciência (MNHNC; Lisbon, Portugal). A total of 242 scales' samples were used, covering six subsystems throughout the Guadiana basin (Figure 2): 45 from Caia (MB85-010246-MB85-010291); 46 from Degebe (MB85-010339-MB85-010345, MB85-010383-MB85-010421); 30 from Ardila (MB85-011512-MB85-011517, MB85-011518-MB85-011531, MB85-011676-MB85-011687); 30 from Chança (MB85-010466-MB85-010482, MB85-011501-MB85-011513); 44 from Vascão (MB85-010422-MB85-010465) and 46 from Odeleite (MB85-010292-MB85-010338).

In accordance with the location of the historic samples, the collection of the contemporary samples was conducted between 2011 and 2012, throughout the Guadiana basin and using electro fishing as sampling method. The individuals were fin clipped and returned to the river. The tissues collected were preserved in 100% ethanol at 4ºC and deposited in the zoological collections "Museu Bocage" of the Museu Nacional de História Natural e da Ciência (MNHNC; Lisbon, Portugal). A total of 272 individuals were sampled for the six subsystems under study (Figure 2): 39 from Caia (MB85-010706-MB85-010743); 32 from Degebe (MB85-010650-MB85-010681); 55 from Ardila (MB85-010766-MB85-010800, MB85-010802-MB85-010821); 50 from Chança (MB85-010828-MB85-010877); 54 from Vascão (MB85-011587-MB85-011602, MB85-011636-MB85-011673) and 42 from Odeleite (MB85-011531-MB85-011573).

Genomic DNA extraction from fresh samples and dried historical samples was made following a standard phenol-chloroform protocol adapted from Sambrook *et al*. (1989) and was stored afterwards at -20ºC. DNA concentration and quality was measured using NanoDrop Spectophotometer ND-1000 v3.2.1. and standardized to 50 ng/µl per sample.

Due to the fact that contamination is a common risk when working with older/degraded tissues (Paabo *et al*. 2004, Wandeler *et al*. 2007), several preventive measures were taken: a bleach treatment was made on the extraction area between every set of extractions; DNA extraction and PCR amplification of the historic samples was preformed strictly separated in time from the fresh ones; and the preparation of the PCR reactions took place in a UV hood.

Besides the risk of contamination when working with historical scales, there are other difficulties associated with this, such as the levels of DNA quality and concentration, which can vary considerably among individuals samples, making it hard to determine the amount of DNA template to use in each PCR reaction (Nielsen *et al*., 1999). However, the use of the Nanodrop Spectophotometer ND-1000 v3.2.1. allowed the circumvention of these issues.

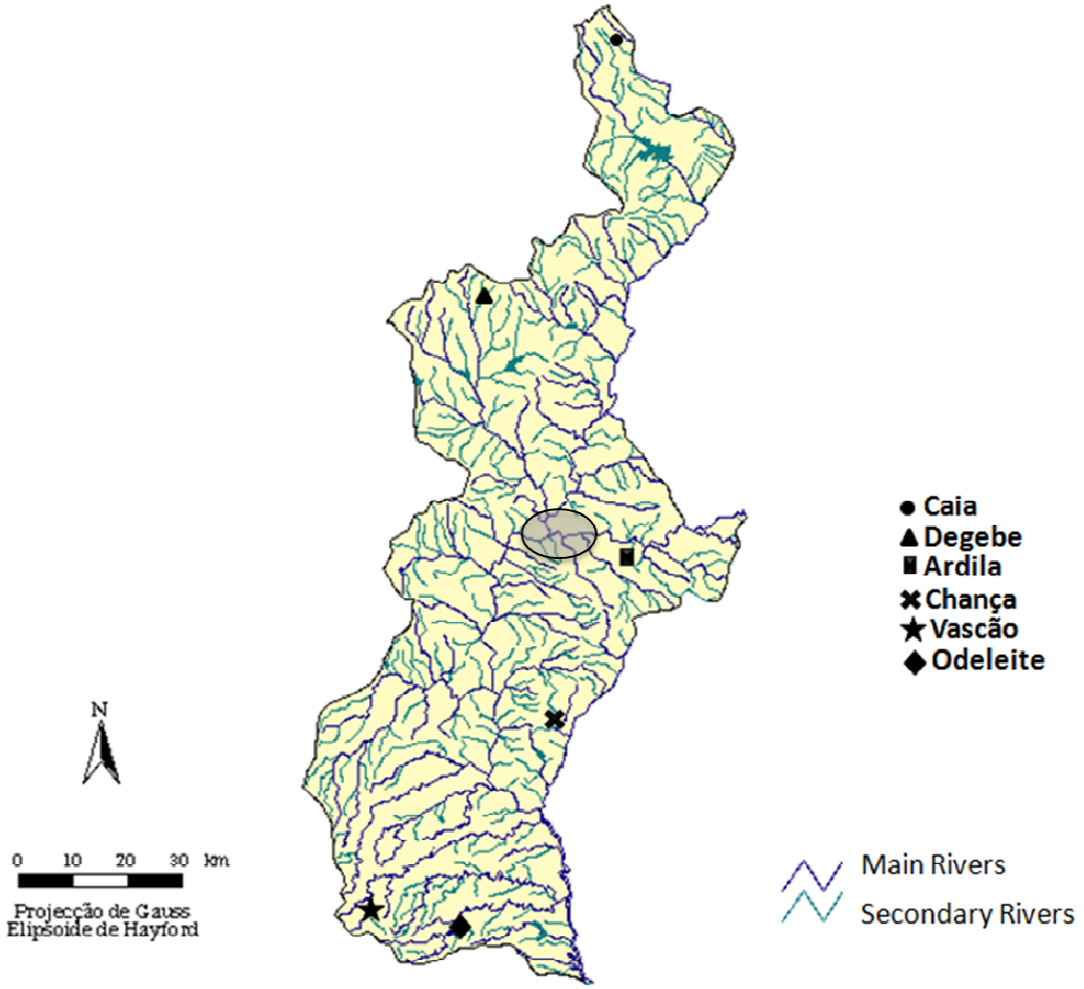


**Figure 2 Guadiana basin in Portugal, the Alqueva dam ( ) and sampling locations: Caia (●); Degebe (▲); Ardila (■); Chança ( ✖); Vascão ( ★); and Odeleite (♦). Adapted from: Ministério do Ambiente, INAG-DSRH, SNIRH in Altlas do Ambiente.**

### 2.3.2. Microsatellites Amplification and Allele Scoring

There were no species-specific microsatellite primers available in literature for *I. lemmingii*. However, various microsatellite primers developed for other close related species (Leuciscinae subfamily) were available ( e.g. Mesquita *et al*., 2003; Larno *et al*., 2005; Pala & Coelho, 2005; Vyskocilová *et al*., 2007; Muenzel *et al*., 2007). Some of these have been successfully used for the target species in a previous study (Lopes-Cunha *et al*., 2012) and for

other close species (*Pseudochondrostoma polylepis*; Mesquita *et al.* personal communication), allowing the development of a cross-species amplification strategy.

Microsatellite *loci* were chosen according to previous information available regarding their PCR (Polymerase Chain Reaction) amplification conditions and allele size (Table 1). Optimization of PCR reactions for the 15 microsatellites *loci* initially tested was conducted in 12µl reactions, containing 50 ng of template DNA, 0.5 µM of forward and reverse primers, 0.2 µM dNTP's, 1.50 mM $MgCl^2$ and 1 u/µL of Taq Polymerase (Dream Taq®) in 1xMH4 reaction buffer. Reactions were conducted on a gradient Bio-Rad® thermal cycler under the following amplification conditions: 95ºC for 15 min.; 30 cycles of 94ºC for 30 s, annealing temperatures (Ta) tested from 52ºC to 62ºC in 2ºC intervals for 1 min 30 s and 72ºC for 1 min; and 72ºC for 10 min (adapted from Mesquita *el al*., 2003).

**Table 1 Microsatellite *loci* that amplified for *I. lemmingii*, and successful annealing temperature.**

| Microsatellite | Annealing Temperature (ºC) | Author |
|---|---|---|
| CypG24 | 60 | Baerwald & May (2004) |
| CypG3 | 60 | Baerwald & May (2004) |
| CypG48 | 60 | Baerwald & May (2004) |
| LceA 149 | 60 | Larno *et al*. (2005) |
| Lcel 100 | 54 | Larno *et al*. (2005) |
| LCO 1 | 60 | Turner *et al*. (2004) |
| LCO3 | 60 | Turner *et al*. (2004) |
| LCO5 | 58 | Turner *et al*. (2004) |
| Lsou 05 | 60 | Muenzel *et al.* (2007) |
| Lsou 08 | 60 | Muenzel *et al.* (2007) |
| Lsou 21 | 60 | Muenzel *et al.* (2007) |
| SarN2F11b | 58 | Mesquita *et al*. (2003 |
| SarN7F8 | 58 | Mesquita *et al*. (2003) |
| SarN7G5 | 60 | Mesquita *et al*. (2003) |
| SarN7K4 | 58 | Mesquita *et al*. (2003) |

After optimization of the isolated amplification of each microsatellite *locus*, two multiplex PCR sets were established, combining identical PCR programs and non-overlapping

allele sizes. However for some of the *loci*, due to different annealing temperatures or different number of PCR cycles, the PCR products were obtained by isolated amplification and afterwards resolved, in combination with the two multiplex sets, by multiloading. The first multiplex/multiloading set included five microsatellite *loci*: CypG3, Lsou 08, Lsou 21, sarN7G5 and Lcel100; and the second included six microsatellite *loci*: LceA149, Lsou 05, CypG48, CypG24, LCO1 and LCO3.

Successful PCR amplification reactions were confirmed through electrophoresis in a 2% agarose gel stained with Sybr Green and posterior visualization under ultra-violet light with a DC290 Kodak digital camera. All tested primers that resulted in a successful amplification for *I. lemmingii* historical and fresh samples, and the respective annealing temperatures, are presented in Table 2.

PCR products were resolved on an ABI Prism – 310 Genetic Analyzer® (Applied Biosystems) in the preliminary study for the optimization of the microsatellite *loci*. For a better resolution and visualization of the allele peaks, the number of PCR cycles was reduced to 21 or 23 for the fresh samples, and 25 or 27 cycles for the historical samples. The PCR products were diluted 1:2 in ultra-pure water, 1µl of the dilution was suspended in 15µl of Formamide Hi-Di and in 0,15µl of LIZ 500 (-250) size standard, and afterwards the samples were denatured for 3 minutes at 95ºC.

After the preliminary study, all the samples were sent to the company Secugen for genotyping under the optimized conditions. To determine the allele sizes the software GeneMapper 3.1 (Applied Biosystems®) was used.

**Table 2 Microsatellite *loci* characterization: repeat motif, primer sequences and fluorochrome labeling, number of observed alleles and respective allelic range.**

| Microsatellite *Loci* | Repeat Motif | Primer Sequence | Nº Alleles | Allelic Size Range (bp) |
|---|---|---|---|---|
| CypG3 (M1) | $(CAGA)_n(TAGA)_n$ | R: GACTGGACGCCTCTACTTTCATA<br>F: AGTAGGTTTCCCAGCATCATTGT NED | 18 | 195-263 |
| CypG24 (M2) | $(CAGA)_n$ | R: TGGCGGTAAGGGTAGACCAC<br>F: CTGCCGCATCAGAGATAAACACTT 6-FAM | 8 | 138-170 |
| CypG48 (M2) | $(TAGA)_nTACGG(TAGA)_n$ | R: CGCACGAGGCACCCACTA<br>F: GTGCTCATGGGACCAAACTGTA PET | 30 | 157-277 |
| Lcel 100 (M1) | $(GATA)_n$ | R: CCAAAAGCACAAGAATAAAAC<br>F: CTTCTTTCAAATCCATCAAAG VIC | 8 | 176-212 |
| Lsou 05 (M2) | $(CA)_n$ | R: CCCACATCTGCTGACTCTGAC<br>F: CTGAAGAAGACCCTGGTTCG VIC | 20 | 174-292 |
| Lsou 08 (M1) | $(GT)_n$ | R: TAGGAACGAAGAGCCTGTGG<br>F: GCGGTGAACAGGCTTAACTC PET | 10 | 181-199 |
| Lsou 21 (M1) | $(GT)_n\ N_{34}(GT)_n\ N_{28}(GT)_n$ | R: TCATGAAGTCGCTGTGGTTC<br>F: GGCAGGAGGACGTCTATGAG 6-FAM | 6 | 269-283 |
| LCO3 (M2) | $(TG)_n$ | R: AAA CAG GCA GGA CAC AAA GG<br>F: GCA GGA GCG AAA CCA TAA AT 6-FAM | 13 | 222-276 |
| SarN7G5 (M1) | $(TG)_nCG(TG)_nCGTGCG(TG)_n(TG)_n$ | R:CTGGAGACCACAGTCAGGTAAATC<br>F: GAGCTTCAGCACCGAGGAC VIC | 12 | 75-109 |

M1, multiplex/multiloading set 1; M2, multiplex/multiloading set 2.

### 2.3.3. Genetic Diversity Analysis

The program MICRO-CHECKER VERSION 2.2.03 (Van Oosterhout *et al*., 2004) was used to check for scoring errors due to null alleles, stuttering and large allele drop out. This application uses a Monte Carlo stimulation (bootstrap) method to generate expected homozygote and heterozygote allele size difference frequencies.

GENEPOP VERSION 4.1.4 (Raymond & Rousset, 1995) was used to test deviations from Hardy-Weinberg equilibrium, which states that allelic frequencies will remain constant from generation to generation in a randomly mating large population with no immigration, mutation or natural selection. GENEPOP VERSION 4.1.4 (Raymond & Rousset, 1995) was also used to test the existence of linkage disequilibrium between all pairs of *loci*. Since the gametic phase is unknown this test is based on a likelihood ratio test, where the likelihood of the sample evaluated under the hypothesis of no association among *loci* (linkage disequilibrium) is compared to the likelihood of the sample when association is allowed (Slatkin & Excoffier, 1996). The significance of the observed likelihood ratio is found by computing the null distribution of this ratio under the hypothesis of linkage disequilibrium, using a permutation procedure.

To access the levels of genetic diversity of both historical and contemporary samples several measures were tested. The number of alleles per *locus* (A), the observed heterozygosity which is the proportion of heterozygous individuals that can be found in a single population, and the expected heterozygosity, which is the prediction of the proportion of heterozygous individuals in a single population based on the known allele frequencies of a sample, were all tested using the program ARLEQUIN 3.11 (Excoffier *et al*., 2005). The allelic richness, average number of alleles by *locus* (Ar) corrected to take into account population size, was tested through a rarefaction procedure as implemented in the software HP-RARE 1.0 (Kalinoswki, 2005; the suggested default of 100 genes in each sample was used to account for rarefaction). Also implemented using this program was the test for the number of private alleles (PA): the number of alleles that can only be found in a single population. Finally, Wright's fixation index (Wright, 1951), or inbreeding coefficient ($F_{IS}$), that measures the decrease in heterozygosity based on the relation between the observed heterozygous with the proportion of heterozygous that would occur in Hardy-Weinberg equilibrium, was also determined using FSTAT 2.9.3.2 (Goudet, 2001). Inbreeding has the direct consequence of homozygosity increase and consequently positive values of $F_{IS}$, which varies between -1 (heterozygous excess) and 1 (homozygous excess).

In order to compare the two data sets with the purpose of testing possible genetic diversity reductions through time, considering allelic richness, heterozygosity and Wright's fixation index, the IBM® SPSS Statistics for Windows version 20.0 (released 2011) software

was used to perform a pairwise Wilcoxon signed-rank test. Sequential Bonferroni corrections were applied when appropriate. The Bonferroni correction is applied when multiple tests are performed in order to contradict the problem of multiple comparisons, avoiding false positives.

### 2.3.4. Population Structure Analysis

Different approaches were used to determine population genetic structure for each moment in time in order to compare them. Firstly, for each set of samples, the genetic differentiation among samples was studied through Wright's F-statistics (1951). Pairwise $F_{ST}$ (Weir & Cockeram, 1984) and $R_{ST}$ (Rousset, 1996) were calculated with the use of the software ARLEQUIN 3.11 (Excoffier *et al*., 2005) with a significance level of 0.05 after 10 000 permutations. $R_{ST}$ is a measure analogous to Wright's $F_{ST,}$ specific for microsatellite *loci* data that differs in taking explicit account of the mutation process in these markers. This measure will generally provide less biased estimates of demographic parameters than will $F_{ST}$, although, comparing both estimates might provide further information (Stalkin, 1995). Significant levels were then adjusted with sequential Bonferroni correction.

The second approach was the factorial correspondent analysis (FCA) was applied using the software GENETIX v. 4.05 (Belkhir *et al*., 2004), which is used to investigate the genetic similarity of individuals based on allele frequencies and genotypes. FCA by population compares genotypes of individuals within the dataset and produces a 2D/3D scatter plot that allows the identification of possibly related groups and the presence of outliers.

Next, the isolation by distance analysis, which was tested on both data sets using Mantel tests on the software GenAlEx 6.5 (Peakall and Smouse, 2012), with the purpose of testing whether genetic differentiation [$F_{ST}$ (1- $F_{ST}$)] among subsystems within Guadiana basin can be simply explained by geographical distance [Ln(Km)] – riverine distance among sampling sites.

In order to assess the population structure, without any *a priori* information about the origin of the individuals, the software STRUCTURE 2.2 (Pritchard *et al*., 2000) was used. This program uses a model-based method developed by Pritchard *et al*. (2000) that has the purpose to determine clusters of individuals on the basis of their genotypes at multiple *loci*, using a Bayesian approach. The model assumes that the populations are in Hardy-Weinberg equilibrium and in linkage equilibrium, attempting to find population groupings that (as far as possible) are not in disequilibrium (Pritchard *et al*., 2000). The software also gives us the estimated log probability data Pr (X|K) for each value of K (number of population clusters created by the software), which allows the estimation of the more likely number of clusters. Information about the quantification of how likely each individual is to belong to each group is also given, in order to assign the individuals to their respective population. The

software was used under the "admixture" model with allelic frequencies independent and values of K from 1 to 7. For each K, 20 simulations were performed with a burning period of 50 000, followed by 100 000 Markov steps (MCMC). To choose the value of K that best suited the data, the protocol by Evanno *et al*. (2005) was followed. This protocol states that the log probability of data Pr (X|K) does not provide a correct estimation of the number of clusters (K), which makes it necessary to perform an *ad hoc* statistics based on the rate of change in the log probability of data between successive K values (ΔK; Evanno *et al*., 2005).

Finally, to test the significance of the different structures suggested by STRUCTURE, an Analysis of Molecular Variance (AMOVA; Excoffier *et al*., 1992) was implemented in ARLEQUIN 3.11 (Excoffier *et al*., 2005). This analysis estimates the genetic structure indices using information on the allelic content of genotypes, as well as their frequencies (Excoffier *et al*., 1992). This test was performed with a significance level of 0.05 and a total of 10 000 permutations.

### 2.3.5. Demographic Analysis

The demographic history of these populations was accessed by investigating if recent reductions in effective population size have occurred. BOTTLENECK (Piry *et al*. 1999) software was used to test each set of samples for demographic signatures of bottleneck effect, which were identified through the deviation in the number of heterozygous individuals expected, based on the allelic frequencies of the used *loci* (Piry *et al*.,1999). Bottleneck events that result in a decrease in the population effective size usually cause a considerable reduction of heterozygous individuals and also in the number of rare alleles, since these are easily lost. Seeing that the rare alleles have little effect in the heterozygosity of the population, their decrease will happen faster than the decrease in the amount of heterozygosity of the population (Piry *et al*. 1999). To determine the existence of a bottleneck effect, the observed and expected heterozygous individuals are compared, taking in consideration the sample size, the number of alleles per *locus* and the assumption of mutation-drift equilibrium (Piry *et al*., 1999). The three evolutionary/mutational models available in the software BOTTLENECK (Piry *et al*. 1999) were used to test possible signs of bottleneck: infinite allele model (IAM), stepwise mutation model (SMM) and two-phase model (TPM). The models were tested with 1000 replications and the TPM model was tested with 70% SMM and 30% IAM.

## 2.4. Results

### 2.4.1 *Loci* Amplification, Allele Scoring, Hardy-Weinberg and Linkage Equilibrium

Two of the 11 microsatellite *loci* optimized were excluded from the analysis, making nine the final number of microsatellites analyzed. The alleles excluded were Lce 149, because it was monomorphic, and LCO1, due to difficulties in allele scoring and genotyping. All the 272 contemporary DNA samples were successfully extracted (100% extraction success) and presented a 95.9% amplification success. Of the 240 historic samples, 99.6% were successfully extracted and 85.1% successfully amplified.

Regarding both sample sets, MICROCHECKER v.2.2.03 (Van Oosterhout *et al.*, 2004) analysis revealed in general no signs of scoring errors due to null alleles, stuttering or large allele drop-out for the nine microsatellite *loci*. Only in four punctual situations, historical samples exclusively (LCO3 for Degebe, Ardila and Vascão; and CypG48 for Vascão) evidences of null alleles where shown, which can be explained by the increased difficulties when extracting and amplifying the historical DNA.

Concerning the two sets of samples, historical and contemporary, and after Bonferroni correction, no evidences of linkage disequilibrium were found. Further, all samples were in Hardy-Weinberg equilibrium, with the exception of two punctual cases for the historical sample set: LCO3 *locus* in Ardila and CypG48 *locus* in Odeleite.

All *loci* were polymorphic and a total of 125 alleles were detected across all *loci* and samples. The less polymorphic *locus* was Lsou 21 with six alleles ranging from 269bp to 283bp, while the most polymorphic *locus* was CypG48 with a total of 30 alleles ranging from 157bp to 277bp (Table 2).

### 2.4.2 Genetic Diversity

In respect to genetic diversity measures, the mean values of allelic diversity obtained through a rarefaction procedure ($A_r$) per *locus* per population for the historical samples (Table 3) were low and varied between 3.22 in Chança and 3.87 in Degebe. Concerning the contemporary samples (Table 4) $A_r$ values ranged from 3.22 in Vascão to 4.00 in Odeleite. Wilcoxon signed-rank test performed for this diversity measure indicated there were no significant differences between the two data sets ($P>0.05$).

Estimates of $H_O$ and $H_E$ for each temporal period indicated only minor changes over time. Regarding the historical populations $H_O$ values varied between 0.482 in Ardila and 0.616 in Caia, and $H_E$ values varied from 0.504 in Ardila till 0.622 in Degebe. Regarding the contemporary samples, $H_O$ values ranged from 0.529 in Ardila till 0.678 in Degebe, and $H_E$ values ranged from 0.528 in Ardila till 0.681 in Odeleite. Wilcoxon signed-rank tests for the $H_O$ and $H_E$ showed that there were no significant differences through time ($P>0.05$).

The fixation index values ($F_{IS}$) on the historical samples ranged from -0.003 in Vascão till 0.099 in Odeleite, and on the contemporary samples range from -0.003 in Ardila till 0.097 in Vascão. When analyzing the variation of $F_{IS}$ values through time using the Wilcoxon signed-rank test, we concluded that the differences between the two data sets were not significant (P>0.05).

**Table 3 Microsatellite diversity measures for the historical samples, per *locus* and overall.**

| | | Caia | Degebe | Ardila | Chança | Vascão | Odeleite |
|---|---|---|---|---|---|---|---|
| CypG3 | N | 45 | 46 | 29 | 30 | 44 | 46 |
| | A | 7 | 8 | 7 | 4 | 7 | 10 |
| | $A_r$ | 4.845 | 5.476 | 5.074 | 2.098 | 4.176 | 5.160 |
| | $P_A$ | 0.065 | 0.644 | 0.066 | 0.080 | 1.544 | 1.174 |
| | $H_O$ | 0.844 | 0.810 | 0.929 | 0.276 | 0.690* | 0.800 |
| | $H_E$ | 0.826 | 0.867 | 0.844 | 0.279 | 0.726* | 0.840 |
| | $F_{IS}$ | -0.022 | 0.068 | -0.102 | 0.011 | 0.049 | 0.048 |
| CypG24 | N | 45 | 46 | 29 | 30 | 44 | 46 |
| | A | 5 | 5 | 5 | 5 | 5 | 5 |
| | $A_r$ | 3.022 | 3.213 | 3.521 | 2.568 | 2.210 | 3.068 |
| | $P_A$ | 0.260 | 0.263 | 0.038 | 0.287 | 0.208 | 0.599 |
| | $H_O$ | 0.622 | 0.605 | 0.448 | 0.600 | 0.364 | 0.595 |
| | $H_E$ | 0.567 | 0.605 | 0.445 | 0.552 | 0.317 | 0.618 |
| | $F_{IS}$ | -0.098 | 0.000 | -0.008 | -0.089 | -0.148 | 0.038 |
| CypG48 | N | 45 | 46 | 29 | 30 | 44 | 46 |
| | A | 20 | 20 | 15 | 14 | 15 | 12 |
| | $A_r$ | 7.183 | 7.241 | 7.005 | 6.839 | 6.659 | 6.736 |
| | $P_A$ | 1.919 | 1.469 | 2.330 | 0.750 | 1.071 | 0.927 |
| | $H_O$ | 0.902 | 0.900 | 1.000 | 0.875 | 0.955 | 0.539 |
| | $H_E$ | 0.942 | 0.944 | 0.937 | 0.927 | 0.921 | 0.926 |
| | $F_{IS}$ | 0.042 | 0.047 | -0.069 | 0.058 | -0.036 | 0.429 |
| Lcel 100 | N | 45 | 46 | 29 | 30 | 44 | 46 |
| | A | 4 | 4 | 2 | 2 | 3 | 2 |
| | $A_r$ | 2.043 | 2.028 | 1.544 | 1.985 | 2.139 | 1.840 |
| | $P_A$ | 0.226 | 0.771 | 0.000 | 0.000 | 0.312 | 0.000 |
| | $H_O$ | 0.242 | 0.212 | 0.077 | 0.111 | 0.357 | 0.308 |
| | $H_E$ | 0.250 | 0.250 | 0.145 | 0.425 | 0.326 | 0.271 |
| | $F_{IS}$ | 0.032 | 0.153 | 0.474 | 0.750 | -0.097 | -0.143 |
| Lsou 05 | N | 45 | 46 | 29 | 30 | 44 | 46 |
| | A | 9 | 11 | 5 | 7 | 10 | 10 |
| | $A_r$ | 4.433 | 5.353 | 3.986 | 4.253 | 4.074 | 4.852 |
| | $P_A$ | 0.304 | 0.763 | 0.351 | 0.735 | 0.876 | 1.499 |
| | $H_O$ | 0.778 | 0.738 | 0.750 | 0.690 | 0.727* | 0.718 |
| | $H_E$ | 0.788 | 0.853 | 0.731 | 0.750 | 0.735* | 0.807 |
| | $F_{IS}$ | 0.013 | 0.136 | -0.026 | 0.082 | 0.010 | 0.111 |
| Lsou 08 | N | 45 | 46 | 29 | 30 | 44 | 46 |
| | A | 7 | 7 | 5 | 5 | 4 | 5 |
| | $A_r$ | 3.822 | 3.625 | 1.997 | 3.463 | 2.521 | 3.110 |
| | $P_A$ | 1.021 | 0.2999 | 0.392 | 0.256 | 0.138 | 0.752 |
| | $H_O$ | 0.733* | 0.738 | 0.250 | 0.750 | 0.605 | 0.568 |
| | $H_E$ | 0.715* | 0.719 | 0.233 | 0.699 | 0.543 | 0.556 |
| | $F_{IS}$ | -0.026 | -0.027 | -0.077 | -0.075 | -0.115 | -0.021 |
| Lsou 21 | N | 45 | 46 | 29 | 30 | 44 | 46 |
| | A | 4 | 4 | 2 | 3 | 3 | 2 |
| | $A_r$ | 2.629 | 2.973 | 1.415 | 2.269 | 2.188 | 1.991 |
| | $P_A$ | 0.412 | 0.919 | 0.000 | 0.634 | 0.060 | 0.000 |
| | $H_O$ | 0.568 | 0.513* | 0.107 | 0.191* | 0.535 | 0.345 |
| | $H_E$ | 0.577 | 0.522* | 0.103 | 0.335* | 0.501 | 0.479 |
| | $F_{IS}$ | 0.015 | 0.018 | -0.038 | 0.437 | -0.068 | 0.284 |
| LCO3 | N | 45 | 46 | 29 | 30 | 44 | 46 |
| | A | 3 | 3 | 3 | 3 | 4 | 4 |
| | $A_r$ | 2.101 | 1.790 | 2.557 | 2.709 | 2.789 | 1.450 |
| | $P_A$ | 0.125 | 0.150 | 0.590 | 0.711 | 1.682 | 0.300 |
| | $H_O$ | 0.389 | 0.100* | 0.222 | 0.500 | 0.405* | 0.100 |
| | $H_E$ | 0.460 | 0.213* | 0.522 | 0.611 | 0.543* | 0.098 |
| | $F_{IS}$ | 0.156 | 0.535 | 0.582 | 0.189 | 0.257 | -0.018 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| SarN7G5 | N | 45 | 46 | 29 | 30 | 44 | 46 |
| | A | 5 | 5 | 4 | 6 | 3 | 4 |
| | $A_r$ | 2.602 | 3.147 | 3.003 | 2.802 | 2.820 | 2.759 |
| | $P_A$ | 0.107 | 0.107 | 0.015 | 0.339 | 0.059 | 0.982 |
| | $H_O$ | 0.467 | 0.733 | 0.552 | 0.778* | 0.605 | 0.711 |
| | $H_E$ | 0.456 | 0.625 | 0.579 | 0.592* | 0.614 | 0.583 |
| | $F_{IS}$ | -0.023 | -0.176 | 0.048 | -0.320 | 0.016 | -0.223 |
| ALL | N | 45 | 46 | 29 | 30 | 44 | 46 |
| | A | 7.111 | 7.444 | 5.333 | 5.444 | 6.000 | 6.000 |
| | $A_r$ | 3.63 | 3.87 | 3.23 | 3.22 | 3.29 | 3.44 |
| | $P_A$ | 0.49 | 0.60 | 0.42 | 0.42 | 0.66 | 0.69 |
| | $H_O$ | 0.616 | 0.594 | 0.482 | 0.530 | 0.583 | 0.520 |
| | $H_E$ | 0.620 | 0.622 | 0.504 | 0.574 | 0.581 | 0.575 |
| | $F_{IS}$ | 0.006 | 0.045 | 0.046 | 0.082 | -0.003 | 0.099 |

N: Number of individuals; A: Number of alleles; $A_r$: Allelic diversity obtained through a rarefaction procedure; $P_A$: Private alleles; $H_O$: Observed heterozygosity; $H_E$: Expected heterozygosity; $F_{IS}$: Fixation Index. Hardy-Weinberg desiquilibrium after Bonferroni correction marked with an * ($p$-value<0.005)

**Table 4 Microsatellite diversity measures for the contemporary samples, per *locus* and overall.**

| | | **Caia** | **Degebe** | **Ardila** | **Chança** | **Vascão** | **Odeleite** |
|---|---|---|---|---|---|---|---|
| CypG3 | N | 39 | 32 | 55 | 50 | 54 | 42 |
| | A | 7 | 7 | 7 | 7 | 9 | 11 |
| | $A_r$ | 4.549 | 5.220 | 4.726 | 2.364 | 3.894 | 5.449 |
| | $P_A$ | 0.051 | 0.837 | 0.083 | 0.140 | 1.435 | 1.584 |
| | $H_O$ | 0.895 | 0.936 | 0.836 | 0.333 | 0.611 | 0.872 |
| | $H_E$ | 0.804 | 0.854 | 0.817 | 0.331 | 0.641 | 0.858 |
| | $F_{IS}$ | -0.115 | -0.097 | -0.024 | -0.008 | 0.047 | -0.017 |
| CypG24 | N | 39 | 32 | 55 | 50 | 54 | 42 |
| | A | 4 | 4 | 4 | 6 | 3 | 5 |
| | $A_r$ | 2.613 | 3.248 | 2.550 | 2.709 | 1.739 | 3.623 |
| | $P_A$ | 0.184 | 0.671 | 0.201 | 0.471 | 0.053 | 1.070 |
| | $H_O$ | 0.410 | 0.700 | 0.472 | 0.440 | 0.173 | 0.684 |
| | $H_E$ | 0.439 | 0.619 | 0.459 | 0.499 | 0.194 | 0.716 |
| | $F_{IS}$ | 0.065 | -0.134 | 0.028 | 0.119 | 0.110 | 0.045 |
| CypG48 | N | 39 | 32 | 55 | 50 | 54 | 42 |
| | A | 18 | 13 | 18 | 15 | 18 | 18 |
| | $A_r$ | 6.981 | 6.779 | 6.214 | 5.711 | 6.764 | 6.663 |
| | $P_A$ | 2.049 | 0.889 | 0.991 | 0.884 | 2.359 | 1.042 |
| | $H_O$ | 0.947 | 1.000 | 0.942 | 0.833* | 0.870 | 0.925 |
| | $H_E$ | 0.933 | 0.928 | 0.897 | 0.852* | 0.925 | 0.918 |
| | $F_{IS}$ | -0.016 | -0.081 | -0.052 | 0.022 | 0.060 | -0.008 |
| Lcel 100 | N | 39 | 32 | 55 | 50 | 54 | 42 |
| | A | 3 | 3 | 2 | 7 | 4 | 3 |
| | $A_r$ | 2.122 | 1.652 | 1.158 | 2.898 | 2.011 | 2.239 |
| | $P_A$ | 0.120 | 0.044 | 0.000 | 0.937 | 0.162 | 0.031 |
| | $H_O$ | 0.306 | 0.161 | 0.036 | 0.460* | 0.259 | 0.439 |
| | $H_E$ | 0.295 | 0.154 | 0.036 | 0.510* | 0.252 | 0.438 |
| | $F_{IS}$ | -0.035 | -0.049 | -0.009 | 0.099 | -0.028 | -0.002 |
| Lsou 05 | N | 39 | 32 | 55 | 50 | 54 | 42 |
| | A | 6 | 8 | 6 | 8 | 6 | 8 |
| | $A_r$ | 4.168 | 4.692 | 4.322 | 4.258 | 3.802 | 4.847 |
| | $P_A$ | 0.146 | 0.237 | 0.573 | 0.779 | 0.830 | 1.087 |
| | $H_O$ | 0.790 | 0.833 | 0.712 | 0.816 | 0.615 | 0.711 |
| | $H_E$ | 0.766 | 0.810 | 0.766 | 0.758 | 0.731 | 0.811 |
| | $F_{IS}$ | -0.031 | -0.029 | 0.072 | -0.078 | 0.160 | 0.126 |
| Lsou 08 | N | 39 | 32 | 55 | 50 | 54 | 42 |
| | A | 5 | 4 | 5 | 5 | 4 | 4 |
| | $A_r$ | 3.728 | 3.170 | 1.905 | 3.617 | 2.511 | 3.054 |
| | $P_A$ | 0.741 | 0.146 | 0.217 | 0.432 | 0.348 | 0.756 |
| | $H_O$ | 0.711 | 0.807 | 0.222 | 0.617 | 0.426* | 0.625 |
| | $H_E$ | 0.718 | 0.679 | 0.208 | 0.715 | 0.558* | 0.563 |
| | $F_{IS}$ | 0.010 | -0.192 | -0.069 | 0.138 | 0.239 | -0.112 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Lsou 21 | N | 39 | 32 | 55 | 50 | 54 | 42 |
| | A | 3 | 5 | 4 | 4 | 3 | 4 |
| | $A_r$ | 2.322 | 3.113 | 2.488 | 2.970 | 2.079 | 2.597 |
| | $P_A$ | 0.037 | 0.537 | 0.275 | 0.904 | 0.007 | 0.417 |
| | $H_O$ | 0.703 | 0.667 | 0.453 | 0.500* | 0.389 | 0.525 |
| | $H_E$ | 0.543 | 0.633 | 0.418 | 0.609* | 0.507 | 0.528 |
| | $F_{IS}$ | -0.299 | -0.054 | -0.085 | 0.181 | 0.234 | 0.005 |
| LCO3 | N | 39 | 32 | 55 | 50 | 54 | 42 |
| | A | 5 | 7 | 6 | 5 | 5 | 8 |
| | $A_r$ | 2.860 | 3.587 | 2.707 | 3.056 | 3.132 | 3.778 |
| | $P_A$ | 0.493 | 0.888 | 0.316 | 0.617 | 0.730 | 0.921 |
| | $H_O$ | 0.632 | 0.526 | 0.491 | 0.653 | 0.540 | 0.610 |
| | $H_E$ | 0.540 | 0.627 | 0.655 | 0.600 | 0.544 | 0.618 |
| | $F_{IS}$ | -0.172 | 0.165 | 0.004 | -0.089 | 0.007 | 0.014 |
| SarN7G5 | N | 39 | 32 | 55 | 50 | 54 | 42 |
| | A | 4 | 4 | 4 | 5 | 5 | 8 |
| | $A_r$ | 2.508 | 2.826 | 3.071 | 3.070 | 3.028 | 3.726 |
| | $P_A$ | 0.012 | 0.020 | 0.004 | 0.097 | 0.211 | 1.265 |
| | $H_O$ | 0.421 | 0.469 | 0.600 | 0.633 | 0.604 | 0.634 |
| | $H_E$ | 0.399 | 0.523 | 0.655 | 0.645 | 0.615 | 0.679 |
| | $F_{IS}$ | -0.057 | 0.106 | 0.084 | 0.019 | 0.018 | 0.066 |
| ALL | N | 39 | 32 | 55 | 50 | 54 | 42 |
| | A | 6.111 | 6.111 | 6.222 | 6.889 | 6.333 | 7.667 |
| | $A_r$ | 3.54 | 3.81 | 3.24 | 3.41 | 3.22 | 4.00 |
| | $P_A$ | 0.43 | 0.47 | 0.30 | 0.58 | 0.68 | 0.91 |
| | $H_O$ | 0.646 | 0.678 | 0.529 | 0.587 | 0.499 | 0.669 |
| | $H_E$ | 0.604 | 0.647 | 0.528 | 0.613 | 0.552 | 0.681 |
| | $F_{IS}$ | -0.070 | -0.047 | -0.003 | 0.043 | 0.097 | 0.017 |

N: Number of individuals; A: Number of alleles; $A_r$: Allelic diversity obtained through a rarefaction procedure; $P_A$: Private alleles; $H_O$: Observed heterozygosity; $H_E$: Expected heterozygosity; $F_{IS}$: Fixation Index. Hardy-Weinberg Desiquilibrium after Bonferroni correction marked with an * ($p$-value<0.005)

### 2.4.3 Population Structure

Genetic differentiation among samples from each data set, accessed through Wright's F statistic (Wright, 1951), showed significant values overall comparisons even after Bonferroni correction ($p$-value<0.05) (Table 5 and Table 6): average $F_{ST}$=0.077 for historical samples, $p$-value<0.001; average $F_{ST}$=0.156 for contemporary samples, $p$-value<0.001). Pairwise $F_{ST}$ values ranged from 0.013 between Caia and Degebe till 0.123 between Chança and Vascão for the historical samples, and from 0.025 between Caia and Degebe till 0.293 between Ardila and Vascão for the contemporary samples. Overall, these tests suggest that each of the six localities analyzed represent a genetically definable population. However, the pairwise $F_{ST}$ values among contemporary populations reached higher levels when compared with the $F_{ST}$ values observed among historical populations, suggesting an increase in population genetic differentiation/structure through time. Especially notable is the higher increase in $F_{ST}$ values registered for the pairwise comparisons involving Chança and Vascão and the remaining populations, where these values reached their higher numbers.

Concerning Roussest's R statistics (Rousset, 1996), $R_{ST}$ values presented a similar pattern to $F_{ST}$ comparisons, although the absolute values were in average lower (average $R_{ST}$=0.039 for historical samples, $p$-value<0.001; and average $R_{ST}$=0.128 for contemporary samples, $p$-value <0.001). For the historical data set the values range from -0.038 between

Odeleite and Vascão till 0.136 between Caia and Odeleite. Regarding the contemporary data set, the $R_{ST}$ values ranged from 0.027 between Chança and Vascão till 0.275 between Caia and Odeleite (Table 5 and Table 6). Similarly to $F_{ST}$ pairwise comparisons, an increase through time in the pairwise $R_{ST}$ values was also observed, except for punctual cases were a small decrease in pairwise $R_{ST}$ value is registered: Caia *vs.* Ardila, Caia *vs.* Degebe.

**Table 5 $F_{ST}$ (bellow diagonal) and $R_{ST}$ (above diagonal) microsatellite pairwise values among historical samples.**

| | **Caia** | **Degebe** | **Ardila** | **Chança** | **Vascão** | **Odeleite** |
|---|---|---|---|---|---|---|
| **Caia** | - | 0.036 | 0.101* | 0.027* | 0.060* | 0.136* |
| **Degebe** | 0.013* | - | 0.035* | 0.009* | -0.019* | 0.043 |
| **Ardila** | 0.053* | 0.040* | - | 0.049 | -0.025* | 0.110* |
| **Chança** | 0.088* | 0.053* | 0.085* | - | 0.0005* | 0.067* |
| **Vascão** | 0.070* | 0.068* | 0.028* | 0.123* | - | -0.038* |
| **Odeleite** | 0.097* | 0.089* | 0.117* | 0.086* | 0.150* | - |

* Significant values after sequential Bonferroni correction (*p*-value<0.05)

**Table 6 $F_{ST}$ (bellow diagonal) and $R_{ST}$ (above diagonal) microsatellite pairwise values among contemporary samples.**

| | **Caia** | **Degebe** | **Ardila** | **Chança** | **Vascão** | **Odeleite** |
|---|---|---|---|---|---|---|
| **Caia** | - | 0.034* | 0.095* | 0.127* | 0.144* | 0.275* |
| **Degebe** | 0.025* | - | 0.043* | 0.115* | 0.123* | 0.225* |
| **Ardila** | 0.116* | 0.093* | - | 0.119* | 0.147* | 0.272* |
| **Chança** | 0.150* | 0.025* | 0.190* | - | 0.027* | 0.097* |
| **Vascão** | 0.188* | 0.211* | 0.293* | 0.235* | - | 0.086* |
| **Odeleite** | 0.098* | 0.076* | 0.128* | 0.168* | 0.203* | - |

* Significant values after sequential Bonferroni correction (*p*-value<0.05)

To further access genetic differentiation among populations, for each data set and through time, the factorial correspondence analysis (FCA) was plotted for each time period (Figure 3 and Figure 4). Historically, population genetic structure is observed including a greater differentiation among the populations of Vascão and Odeleite and the remaining four populations, Ardila, Caia, Chança and Degebe, which, as seen in the scatter plot, stay close together in a blur. Contemporarily (Figure 4), there is an increase in the population structure observed and also, the Chança population shows higher differentiation from the others, while the individuals from the Odeleite population get closer to the blur mentioned previously, that is now formed by Ardila, Caia and Degebe. Finally, Vascão population remains the most

differentiated population. This pattern confirms the trends observed for the Wright's F statistic (Wright, 1951) analysis.

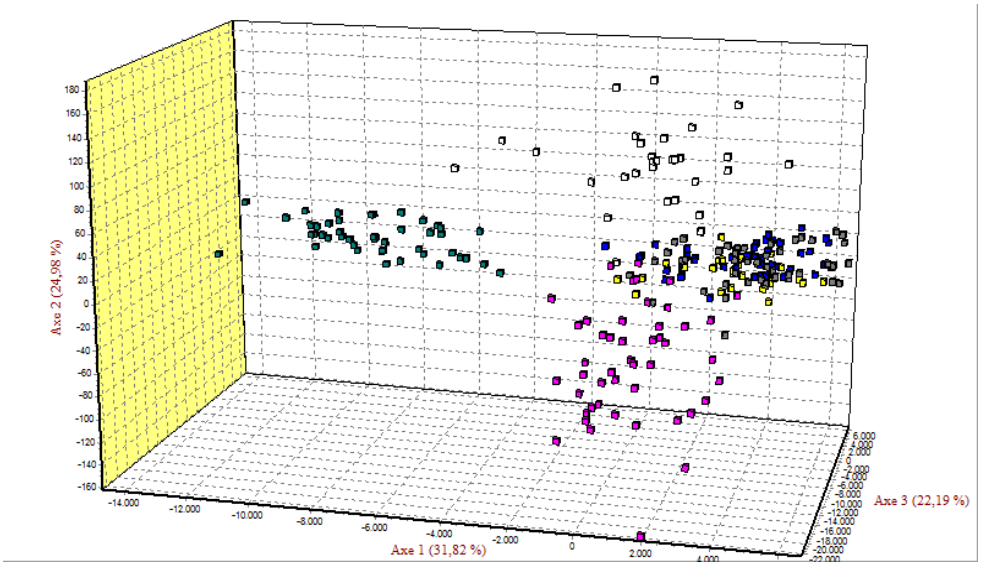


**Figure 3 FCA for the historical data set. Sample localities are: Caia, blue; Degebe, grey; Ardila, yellow; Chança, white; Vascão, dark green and Odeleite, pink.**

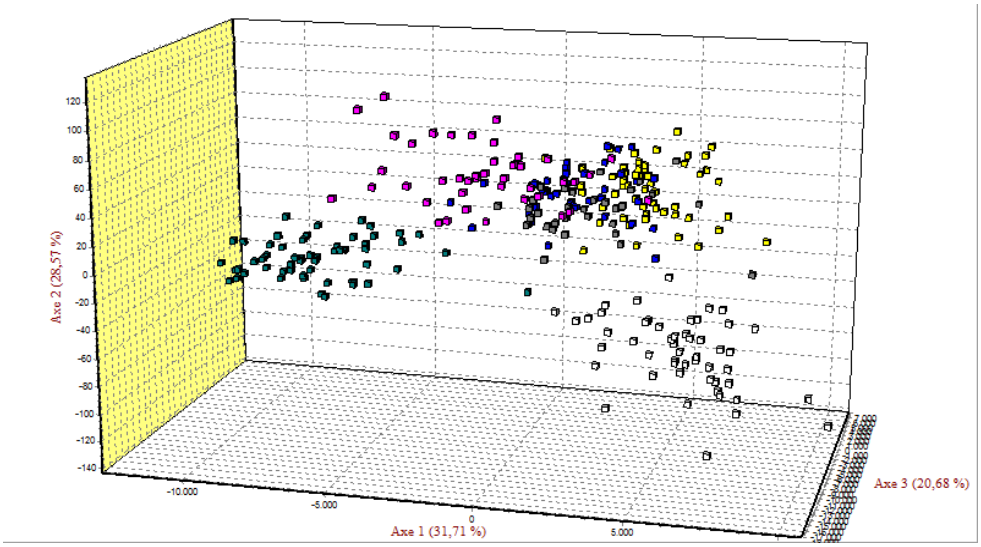


**Figure 4 FCA for the contemporary data set. Sample localities are: Caia, blue; Degebe, grey; Ardila, yellow; Chança, white; Vascão, dark green and Odeleite, pink.**

The analysis of isolation by distance (Figure 5) performed through Mantel test on GenAlEx 6.5 (Peakall and Smouse, 2012), showed that the genetic differentiation registered cannot be explained simply by riverine distances among localities. The genetic differentiation observed among contemporary samples is even less explained by river course distance, seeing that the slope is slightly more negative, than for the historical ones.

Also, this analysis confirmed the pairwise $F_{ST}$ patterns obtained on previous analysis, where the contemporary populations reached higher values than the historical ones.

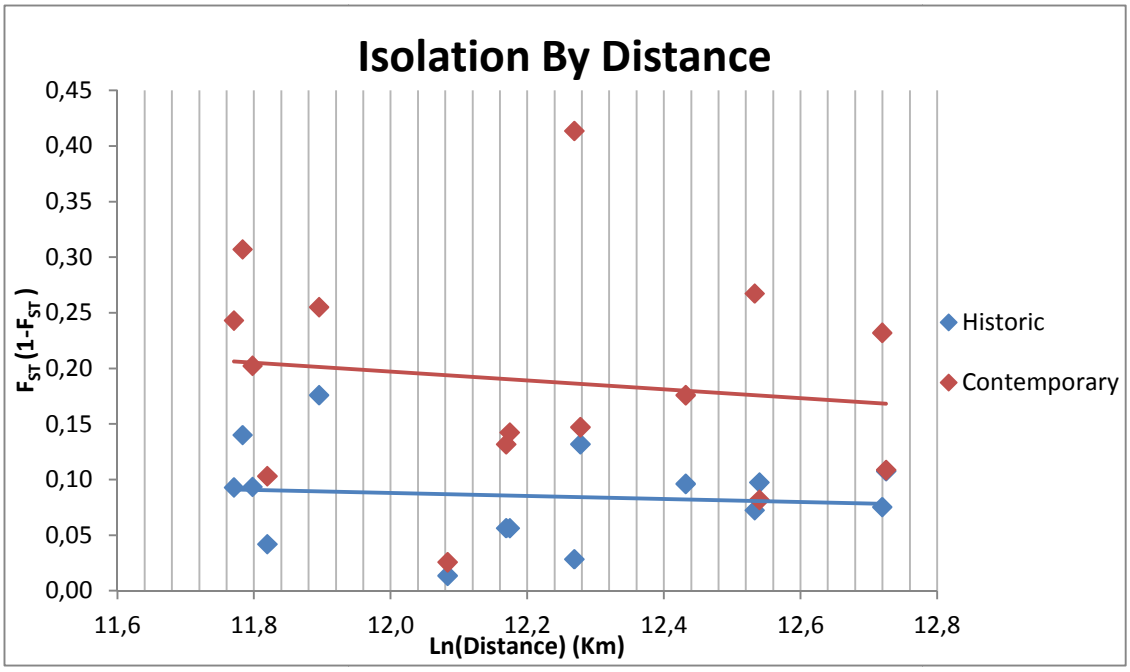


**Figure 5 Isolation by distance patterns for the contemporary and historical samples. Contemporary: $R^2$= 0.0184, P=0.348. Historical: $R^2$= 0.0112, P=0.380.**

The population structure analysis with STRUCTURE 2.2 software (Pritchard *et al.*, 2000), without prior knowledge of the origin of the individuals, and a posterior analysis following Evanno *et al.* (2005) protocol, showed that for the historical samples the maximum value for the estimated likelihood was at K=2 (Figure 6), where the two groups defined are formed by Ardila+Caia+Chança+Degebe and by Odeleite+Vascão populations.

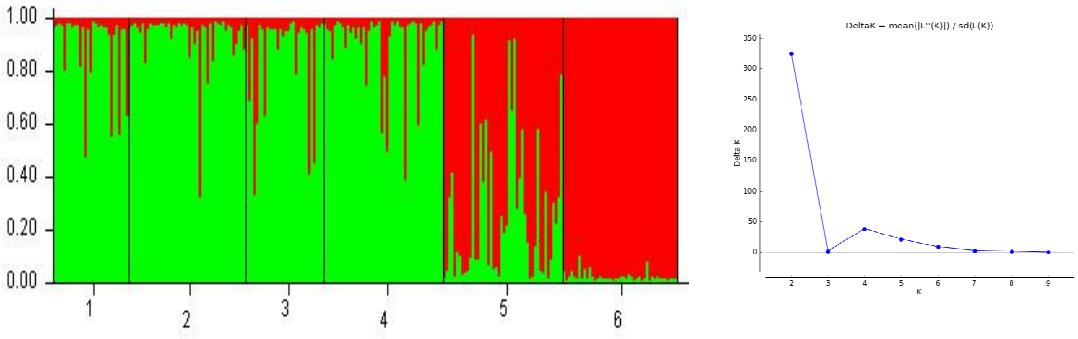


**Figure 6 Left: Most likely population structure for the historical samples obtained under the admixture model. Populations are separated by black lines and labeled as follows: 1 – Ardila; 2 – Caia; 3 – Chança; 4 – Degebe; 5 – Odeleite; 6 - Vascão. Each vertical line corresponds to one individual. Each color represents one structure group. Right: Distribution of the rate of change in the log probability of data between successive values of K accordingly to Evanno et al. (2005). The modal value of this distribution represents the true value of K.**

For the contemporary samples the most likely number of inferred populations' groups was K=3 (Figure 7). The three groups identified are constituted by Ardila+Caia+Degebe+Odeleite, by Chança and by Vascão populations. Here, it is important to

point out the high differentiation through time observed for Chança. Contrarily, the Odeleite population has become more similar to Ardila+Caia+Degebe.

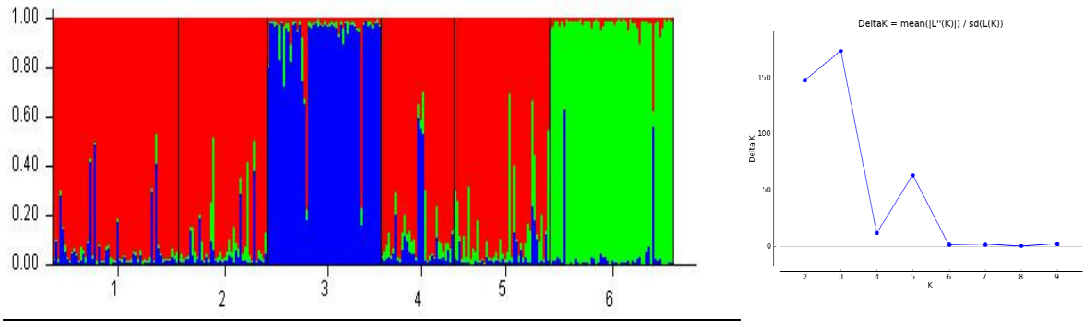


**Figure 7 Left: Most likely population structure for the contemporary samples obtained under the admixture model. Populations are separated by black lines and labeled as follows: 1 – Ardila; 2 – Caia; 3 – Chança; 4 – Degebe; 5 – Odeleite; 6 - Vascão. Each vertical line corresponds to one individual. Each color represents one structure group. Right: Distribution of the rate of change in the log probability of data between successive values of K accordingly to Evanno et al. (2005). The modal value of this distribution represents the true value of K.**

For each data set, the AMOVA analysis was performed for both group structures defined by STRUCTURE 2.2 (Pritchard *et al*., 2000) (Table 7 and Table 8): two groups structure, Ardila+Caia+Chança+Degebe and Vascão+Odeleite; and three groups structure, Ardila+Caia+Degebe+Odeleite, Chança and Vascão. The AMOVA analysis did not support the STRUCTURE 2.2 (Pritchard *et al*., 2000) results for the historical populations, seen that the variation among groups is higher on the alternative structure (three groups structure; $F_{CT}$=0.084) than on the one provided by the software (two groups structure; $F_{CT}$= -0.004). In addition, the differentiation among populations/within groups is lower on the three group structure ($F_{SC}$= 0.0702) than on the two group structure ($F_{SC}$=0.081). Contrarily, for the contemporary populations, the three group clustering defined by STRUCTURE, Ardila+Caia+Degebe+Odeleite, Chança and Vascão, is supported by the AMOVA results.

**Table 7 AMOVA results for the historical dataset.**

| | **Historical dataset (k=2)** | | | | |
|---|---|---|---|---|---|
| Source of Variation | d.f. | Sum of Squares | Variance Components | Percentage of Variation | Fixation Index |
| Among Groups | 1 | 6.623 | -0.004 | -0.37 | $F_{CT}$=-0.004 |
| Among Populations Within Groups | 4 | 27.353 | 0.077 | 8.08 | $F_{SC}$=0.081 |
| Within Populations | 470 | 413.743 | 0.880 | 92.29 | $F_{ST}$=0.077 |
| Total | 475 | 447.718 | 0.954 | 100 | - |
| | **Historical dataset (k=3)** | | | | |
| Source of Variation | d.f. | Sum of Squares | Variance Components | Percentage of Variation | Fixation Index |

| | | | | | |
|---|---|---|---|---|---|
| Among Groups | 2 | 15.153 | 0.1505 | 1.57 | $F_{CT}$=0.084 |
| Among Populations Within Groups | 3 | 18.823 | 0.06641 | 6.91 | $F_{SC}$=0.0702 |
| Within Populations | 470 | 413.743 | 0.88030 | 91.53 | $F_{ST}$=0.016 |
| Total | 475 | 447.718 | 0.96177 | 100 | - |

**Table 8 AMOVA results for the contemporary dataset.**

| | **Contemporary Dataset (k=2)** | | | | |
|---|---|---|---|---|---|
| Source of Variation | d.f. | Sum of Squares | Variance Components | Percentage of Variation | Fixation Index |
| Among Groups | 1 | 54.427 | 0.08666 | 3.87 | $F_{CT}$=0.039 |
| Among Populations Within Groups | 4 | 122.781 | 0.32548 | 14.54 | $F_{SC}$=0.151 |
| Within Populations | 538 | 982.344 | 1.82592 | 81.59 | $F_{ST}$=0.184 |
| Total | 543 | 1159.551 | 2.23805 | 100 | - |
| | **Contemporary Dataset (k=3)** | | | | |
| Source of Variation | d.f. | Sum of Squares | Variance Components | Percentage of Variation | Fixation Index |
| Among Groups | 2 | 123.350 | 0.271 | 11.84 | $F_{CT}$=0.118 |
| Among Populations Within Groups | 3 | 53.858 | 0.195 | 8.49 | $F_{SC}$=0.096 |
| Within Populations | 538 | 983.344 | 1.826 | 76.67 | $F_{ST}$=0.203 |
| Total | 543 | 1159.551 | 2.292 | 100 | - |

## 2.4.4. Demographic Analysis

Regardless of the mutation model taken into account, none of the historical and contemporary data sets revealed signs of bottleneck departures from the mutation drift equilibrium, given that all *p*-values were superior to 0.05 (Table 9 and Table 10).

**Table 9 Bottleneck analysis results for the historical samples.**

| | **IAM** | **TPM (70%)** | **SMM** |
|---|---|---|---|
| **Caia** | 0.06445 | 0.36719 | 0.97559 |
| **Degebe** | 0.12500 | 0.28516 | 0.91797 |
| **Ardila** | 0.17969 | 0.32617 | 0.67383 |
| **Chança** | 0.17969 | 0.67383 | 0.89844 |
| **Vascão** | 0.08203 | 0.41016 | 0.89844 |
| **Odeleite** | 0.06445 | 0.24805 | 0.91797 |

IAM, Infinite Allele Model; TPM 70%, Two-Phase mutation Model with 70% of SMM; SMM, Stepwise Mutation Model.

**Table 10 Bottleneck analysis results for the contemporary samples.**

| | IAM | TPM (70%) | SMM |
|---|---|---|---|
| **Caia** | 0.06445 | 0.32617 | 0.97559 |
| **Degebe** | 0.12500 | 0.28516 | 0.91797 |
| **Ardila** | 0.15039 | 0.32617 | 0.67383 |
| **Chança** | 0.17969 | 0.58984 | 0.89844 |
| **Vascão** | 0.08203 | 0.36719 | 0.87500 |
| **Odeleite** | 0.06445 | 0.24805 | 0.91797 |

IAM, Infinite Allele Model; TPM 70%, Two-Phase mutation Model with 70% of SMM; SMM, Stepwise Mutation Model.

## 2.5. Discussion

### 2.5.1. Continuous Genetic Diversity Through Time

Overall, we found no significant changes in the genetic diversity patterns of *I. lemmingii* between the two moments in time. Estimates of observed heterozygosity ($H_O$) and expected heterozygosity ($H_E$) for each data set were high and indicated no change over time, seeing that mean $H_E$ in historical populations ranged from 0.504 till 0.622 and in contemporary populations from 0.527 till 0.681. The overall values of genetic diversity were generally higher than those observed for other *Iberchondrostoma* species, like *I. lusitanicum* (Sousa *et al*. 2008) and *I. almacai* (Sousa *et al*. 2010). Also, these values were similar to the ones obtained by Lopes-Cunha *et al*. (2012) for *I. lemmingii*. However, these comparisons must be addressed with caution, since they might be biased due to differences among markers.

In view of the known environmental threats and predictable population impacts, a reduction in genetic diversity through time was expected, since microsatellites markers are highly polymorphic, usually including a high number of rare alleles that potentially makes them very sensitive to the detection of genetic diversity losses (Nielsen *et al*. 1999). The lack of observed changes in heterozygosity values may be due to the short period of time that passed between sampling sets, seen that about 15 years may not be sufficient to cause significant changes in genetic diversity values. Also, *I. lemming* has a wide distribution range and inhabits a diverse number of habitats (Lopes-Cunha *et al*., 2012), which may also explain the high and unchangeable values of genetic diversity through time. Moreover, Robalo *et al*. (2008), explained the observed pattern of higher genetic variability in *I. lemmingii* as an outcome of its' broad distribution: species with a wider distribution (like *I. lemmingii*) are more likely to maintain several ancestral polymorphisms, while species with more limited ranges (like *I. almacai* and *I. lusitanicum*) might incur in higher losses of genetic variability due to stochastic events.

In addition, the expected loss of diversity would be related with the increase in mortality, which would be caused by the extreme drought event. However, since this species has

high population sizes and high levels of fertility, and is adapted to the annual cycle of seasonal predictable floods and droughts that characterize the Mediterranean rivers, it was able to respond to the extreme drought event and maintain its high genetic diversity values through time. Nevertheless, the increase in anthropogenic pressures might compromise this resilience capacity of the species.

### 2.5.2. Increased Population Differentiation Through Time

The analyses performed in order to understand the population structure of the two moments in time: the FCA (Figure 3 and Figure 4), the $F_{ST}$ and $R_{ST}$ values, (Table 5 and Table 6) and STRUCTURE 2.2 (Pritchard *et al*., 2000; Figure 6 and Figure 7) suggest that overall, there was an increase among populations' differentiation through time. The highest increases in genetic differentiation were observed in the populations of Chança and Vascão with the other populations.

The STRUCTURE (Pritchard *et al*., 2000) analysis showed that, on the historical samples two clusters were formed, with Vascão and Odeleite forming one and Ardila, Caia, Chança and Degebe forming the other. On the contemporary samples the number of clusters formed was three, which shows the increase in population structure, with Chança and Vascão forming two different clusters, and the remaining populations, Ardila, Caia and Degebe, forming the other one.

These results regarding the historical populations suggest that the genetic structure was related with the geographical distribution of the sub-basins, since the ones further south, Vascão and Odeleite, clustered together. The results are concordant with other studies, given that in general, freshwater fish differentiation patterns are usually associated with river courses differentiation and isolation, like is seen in Sousa *et al.* (2008) for *I. lusitanicum* and *I. almacai*.

Concerning the isolation by distance test (Figure 5), we found that population differentiation patterns on both moments in time is not explained exclusively by riverine distance and that consequently there are other factors influencing population structuring.

Regarding the contemporary populations, it is important, for conservation purposes, to determine the origins of this higher differentiation pattern, whether it is a natural remnant of historic population structure or whether it is the result of more recent human induced changes (Tracy and Jamieson, 2011). Furthermore, as seen for the historic samples, the isolation by distance analysis for this data set also indicates that other factors than riverine distance should be influencing the population structuring, meaning that the alterations observed for the contemporary samples in relation to the historical ones, should have more recent mechanisms influencing it and shaping the species' population structure.

In the case of *I. lemmingii* these more recent mechanisms are most likely due to anthropogenic pressures, which can be explained by the construction of the Alqueva dam that

may have caused reduced availability, connectivity and distribution of *refugia* and reduced mobility. As a consequence of this action as barrier, the migratory potential was reduced and thus, in view of the reduction in gene flow among more isolated populations, the increased differentiation among populations was promoted. All the populations suffered an increase in genetic structuring, although the highest was observed for Chança, which was differentiated from all the others forming a cluster. This sub-basin is located downstream from the Alqueva dam, with a riverine distance of approximately 72 km from it, probably becoming isolated from all the other upstream sub-basins. Moreover, the Vascão, also located downstream from the Alqueva dam, has similarly experienced one of the greatest differentiations, also forming a cluster.

On the other hand, natural environmental changes, like severe droughts, are less likely to be the cause of these alterations in population structure, since usually their effects are not localized, affecting populations in all sub-basins equally. Further, due to presenting an action that is more limited in time, their primary effect would be a reduction in genetic diversity (due to increased mortality) and less probably an effect over population structure.

### 2.5.4 Implications for Conservation

Our results support that, although the values of *I. lemmingii* genetic diversity remain high and similar through time, a recent increase in population structure was observed, probably related to anthropogenic pressures. Also, due to the intermittent character of this species' habitat, populations are more vulnerable to stochastic events, and consequently to sudden population collapses and losses of genetic diversity, increasing their extinction risk (Lopes-Cunha *et al*., 2012). In addition, anthropogenic pressures might also cause these stochastic events, decreasing the species ability to adapt to the cyclic character of their habitat.

No evidences of bottleneck departures from mutation drift equilibrium under the different mutation models were found. This lack of evidences of bottleneck is contrary to what has been described for other endangered cyprinids, like *I. almacai* and *I. lusitanicum* (Sousa *et al*., 2008). Though, it might be possible that there were not enough statistical power to detect recent demographic events, since a higher statistical power is only achieved with 10 to 15 *loci* (Cornuet & Luikart, 1996).

As seen in the present study, it is important to keep in mind that some of the main threats to the conservation of freshwater fishes are related with human impacts on their habitats. Reduction of water quality, water and sand extractions, construction of dams and habitat degradation, which cause alterations on the natural flow regime and on the physical structure of the ecosystem, are some examples of the main threats affecting these habitats. Other important threats responsible for freshwater fish species decline and extinction are the introduction of

exotic species and translocations (Collares-Pereira *et al*., 2000). Therefore, the recovering of a species only makes sense if its habitat is also properly recovered to assure their survival. In accordance, an efficient management strategy should have one of its main focuses on the habitat. In conclusion, a management program with river vigilance regarding the previously mentioned impacts, would allow an earlier detection and reaction, benefiting not only *I. lemmingii*, but also all the other species that share the same habitat.

Finally, results have shown that *I. lemmingii* has a rather complex population structure that has been changing through time and that could only benefit from a deeper knowledge of its current structure. In order to create a good conservation program for this species, the inclusion of samples from other subsystems of the Guadiana basin, in order to cover a wider and detailed geographic range, and an increase in the number of microsatellite loci used would be of great benefit. These analyses would complement the results already obtained, increasing their statistical power and providing a better understanding of the species.

## Acknowledgments

We would like to thank Filipe Ribeiro for assistance in fieldwork, and Maria Drago and Catarina Mateus for assistance in lab work. We thank Museu Nacional de História Natural e da Ciência (MNHNC; Lisbon, Portugal) for the access to the zoological collections "Museu Bocage". This study was founded by Fundação para a Ciência e a Tecnologia (FCT) co-participated by FEDER – Fundo Europeu de Desenvolvimento Regional, under the framework of the PTDC/BIA-BDE/69769/2006, PTDC/BIA-BDE/66519/2006 and PEst-OE/BIA/UI0329/2011 projects and the postdoctoral grant SFRH/BPD/14557/2003 awarded to N. Mesquita The Portuguese agencies Autoridade Florestal Nacional (AFN) and Instituto da Conservação da Natureza e da Biodiversidade (ICNB) provided the necessary permits for the fieldwork.

### 2.6. References

1. Abell R (2005). Conservation Biology for the Biodiversity Crisis: a Freshwater Follow-up. *Conservation Biology*, **16**:1435-1437.
2. Almaça C (1995). Freshwater fish and their conservation in Portugal. *Bioogicall Conservation*, **72**:125-127.
3. Baerwald MR & May B (2004). Characterization of microstaellite *loci* for five members of the minnow family Ciprynidae found in the Sacramento-San Joaquim Delta and its tributaries. *Molecular Ecology Notes*, **4**:385-390.
4. Belkhir K, Borsa P, Chiki L, Refaust N, Bonhomme F (2000). GENETIX 4.01; logiciel sous Windows pour la génétique des populations. Laboratoire Génome, Populations,

Interactions, Université de Montpellier, Montpellier. Available in: http://www.univmontp2.fr/~GENETIX 4.2/GENETIX 4.2/GENETIX 4.2.html.

5. Boix D, Garcia-Berthou E, Gascon S, Benejam L, Tornes E, Sala J, Benito J, Munne A, Sola C, Sabater S (2010). Response of community structure to sustained drought in Mediterranean rivers. *Journal of Hydrology,* **383**:135–146.
6. Bunn SE & Arthington AH (2002). Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management*, **30**(4):497-507.
7. Burgmer T, Hillebrand H, Pfenninger M (2006). Effects of climate-driven temperature changes on the diversity of freshwater macroinvertebrates. *Oecologia*, **151**:93-103
8. Cabral MJ (Coord.), Almeida J, Almeida PR, Dellinger T, Ferrand de Ameida N, Oliveira ME, Palmeirim, JM, Queiroz AI, Rogado L, Santos-Reis M (eds) (2005). *Livro Vermelho dos Vertebrados de Portugal*. Instituto da Conservação da Natureza. Lisboa. 660pp.
9. Chakraborty A, Sakai M, Iwatsuki Y (2006). Museum fish specimens and molecular taxonomy: A comparative study on DNA extraction protocols and preservation techniques. *Journal of Applied Ichthyology*, **22**:160-166.
10. Collares-Pereira MJ, Cowx IG, Ribeiro F, Rodrigues JA, 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.
11. Cornuet JM & Luikart G (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency date. *Genetics*, **144**:2001-2014.
12. Davey AJH & Kelly DJ (2007). Fish community responses to drying disturbances in an intermittent stream: a landscape perspective. *Freshwater Biology*, **52**:1719–1733.
13. Doadrio I (2001). Atlas y libro rojo de los peces continentales de España. Dirección General de Conservación de la Naturaleza. Museu Nacional de Ciencias Naturales. Madrid. 364pp.
14. Evanno E, Regnaut S, Goudet E (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**:2611-2620.
15. Excoffier L, Laval G, Schneider S (2005). Arlequin 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**:47-50.
16. Excoffier L, Smouse P, Quattro J (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, **131**:479-491.

17. Goudet J (2001). FSTAT: a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available in: http://www.unil.ch/izea/softwares/fstat.html.
18. Hawkins BA, Field R, Cornell HV, Currie DJ, Guégan JF, Kaufman DM, Kerr JT, Mitellbach GG, Oberdorff T, O'Brien EM, Porter EE, Turner JRG (2003). Energy, Water, and broad-scale geographic patterns of species richness. *Ecology*, **84**(12):3105-3117.
19. Kalinowski ST (2005). HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**:187-189.
20. Lake PS (2003). Ecological effects of perturbation by drought in flowing waters. *Freshwater Biology*, **48**:1161–1172.
21. Lane M (1996). Roles of Natural History Collections. *Annals of the Missouri Botanical Garden*, **83**:536-545.
22. Larno V, Launey S, Devaux Alain, Laroche J (2005). Isolation and characterization of microsatellite *loci* from chub Leuciscus cephalus (Pisces: Cyprinidae). *Molecular Ecology Notes*, **5**:752-754.
23. Lister AM, Climate Change Research Group (2011). Natural history collections as sources of long-term data sets. *Trends in Ecology and Evolution*, **26**:153-154.
24. Lopes-Cunha M, Aboim MA, Mesquita N, Alves MJ, Doadrio I, Coelho MM (2012). Population genetic structure in the Iberian cyprinid fish Iberochondrostoma lemmingii (Steindachner, 1866): disentangling species fragmentation and colonization processes. *Biological Journal of the Linnean Society*, **105:**559–572.
25. Magalhães MF, Beja P, Schlosser IJ, Collares-Pereira MJ, (2007). Effects of multi-year droughts on fish assemblages of seasonally drying Mediterranean streams. *Freshwater Biology*, **52**:1494–1510.
26. Magoulick DD & Kobza RM (2003). The role of *refugia* for fishes during drought: a review and synthesis. *Freshwater Biology*, **48**:1186–1198.
27. Martínez-Cruz B, Godoy JA, Negro JJ (2007). Population fragmentation leads to spatial and temporal genetic structure in the endangered Spanish imperial eagle. *Molecular Ecology*, **16**:477-486.
28. Matthews WJ & Marsh-Matthews E, (2003). Effects of drought on fish across axes of space, time and ecological complexity. *Freshwater Biology*, **48**:1232–1253.
29. Mesquita N, Cunha C, Hänfling B, Carvalho GR, Zé-Zé L, Tenreiro R, Coelho MM (2003). Isolation and characterization of polymorphic microsatellite *loci* in the endangered freshwater fish *Squalius aradensis* (Cyprinidae). *Molecular Ecology Notes*, **3**:572-574.

30. Mesquita N, Hänfling, Carvalho GR, Coelho MM (2005). Phylogeography of the cyprinid Squalius aradensis and implications for conservation of the endemic freshwater fauna of southern Portugal. *Molecular Ecology,* **14**:1939-1954.
31. Muenzel FM, Sanetra M, Sakburger W, Meyer A (2007). Microsatellites from the vairone Leuciscus souffia (Pisces: Cyprinidae) and their application to closely related species. *Molecular Ecology Notes*, **7**:1048-1050.
32. Myers N, Mittermeier RA, Mittermeier CG, Fonseca GAB, Kent J (2000). Biodiversity hotspots for conservation priorities. *Nature*, **403**:853-858.
33. Nielsen EE, Hansen MM, Loeschcke V (1997). Analysis of microsatellite DNA from old scale samples of Atlantic salmon salmo salar: a comparison of genetic composition over 60 years. *Molecular Ecology*, **6**:487-492.
34. Nielsen EE, Hansen MM, Loeschcke V (1999). Anlysis of DNA from old scale samples: technical aspects, applications and perspectives for conservation. *Heredity,* **130**:264-276.
35. Paabo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreite M (2004) Genetic analyses from ancient DNA. *Annual Review of Genetics*, **38**:645-679.
36. Pala I & Coelho MM (2005). Contrasting views over a hybrid compex: Between speciation and evolutionary "dead-end". *Gene*, **347**:283-294.
37. Peakall R and Smouse PE (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**:2537-2539
38. Petts GE (1984). Impounded Rivers: Perspectives for ecological management. Wiley, Chichester, UK.
39. Piry S, Luickart G, Cornuet JM (1999). BOTTLENECK: a computer program for determining recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**:248-249.
40. Pritchard JK, Stephens M; Donnely P (2000). Inference of population structure using multilocus genotype data. *Genetics*, **155**:945-959.
41. Raymond M & Rousset F (1995). GENEPOP (version 4.1.4): a population genetics software for exact tests and acumenicism. *Journal of Heredity*, **86**:248-249.
42. Robalo JI, Doadrio I, Valente A, Almada VC (2008). Identification of ESUs in the critically endangered Portuguese minnow *Chondrostoma lusitanicum* Collares-Pereira 1980, based on a phylogeographical analysis. *Conservation Genetics*, **8**:1225-1229.
43. Rousset F (1996). Equilibrium values for measures of population subdivision for setpwise mutation processes. *Genetics*, **145**:1219-1228.

44. Salgueiro P, Carvalho G, Collares-Pereira MJ, Coelho MM (2003). Microsatellite analysis of genetic population structure of the endangered cyprinid Anaecypris hispanica in Portugal: implications for conservation. *Biological Conservation*, **109**:47-56.
45. Sambrook J, Fritsch EF, Maniatis T (1989). Molecular cloning: a laboratory manual. New York, NY: Cold Spring Harbor Laboratory Press.
46. Settele J, Hammen V, Hulme P, Karlson U, Klotz S, Kotarac M, Kunin W, Marion G, O'Connor M, Petanidou T, Peterson K, Potts S, Pritchard H, Pysek P, Rounsevell M, Spangenberg J, Steffan-Dewenter I, Skyes M, Vighi M, Zobel M, Kühn I (2005). Alarm: Assessing Large-scale environmental Risks for biodiversity with tested Methods. *GAIA – Ecological Perspectives for Science and Society*, 14:60-72.
47. Slatkin M & Excoffier L (1996) Testing for linkage disequilibrium in genotypic data using the EM algorithm. *Heredity*, **76**:377-383.
48. Slatkin M & Excoffier L (1996) Testing for linkage disequilibrium in genotypic data using the EM algorithm. *Heredity*, **76**:377-383.
49. Sousa V, Penha F, Collares-Pereira MJ, Chikhi L, Coelho MM (2008). Genetic structure and signature of population decrease in the critically endangered freshwater cyprinid *Chondrostoma lusitanicum*. *Conservation Genetics,* **9**:791-805.
50. Sousa V, Penha F, Pala I, Chikhi L, Coelho MM (2010). Conservation genetics of a critically endangered Iberian minnow: evidence of population decline and extirpations. *Animal Conservation*, **13**:162-171.
51. Stuebing RB, Wong A (2001). The role of Sabah's Zoological Collections in the Conservation of protected areas. In: Proceedings of the International Conference on In-situ and Ex-situ Biodiversity Conservation in the New Millennium (eds. Jaacob Z, Moo-Tan S, Jorath S) Sabah, Malásia.
52. Suarez AV, Tsutsui ND (2004). The value of museum collections for research and society. *Bioscience*, **54**:66–74.
53. Tracy LN & Jamieson IG (2011). Historic DNA reveals contemporary population structure results from anthropogenic effects, not pre-fragmentation patterns. *Conservation Genetics*, **12**:517-526.
54. Turner TF, Dowling TE, Broughton RE, Gold JR (2004). Variable microsatellite markers amplify across divergent lineages of cyprinid fishes (subfamily Leuscinae). *Conservation Genetics*, **5**:279-281.
55. Van Oosterhout, C, Hutchinson WF, Willis O M, Shipley PF (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**:535-538.

56. Vyskocilová M, Simková A, Martin JF (2007). Isolation and characterization of microsatellites in Leuciscus cephalus (Cypriniformes, Cyprinidae) and cross-species amplification within the family Cyprinidae. *Molecular Ecology Notes*, **7**:1150-1154.
57. Wandeler P, Hoeck PEA, Keller LF (2207). Back to the future: museum specimens in population genetics. *Trends in Ecology and Evolution*, **22**:634-642
58. Weir BS & Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, **38**:1358-1370.
59. Winker K (2004). Natural History Museums in a Postbiodiversity Era. *BioScience*, **54**:455-459.
60. Wissenschaftsrat (2011). Recommendations on Scientific Collections as Research Infrastructures. Berlin.
61. Wright S (1951). The genetical structure of populations. *Annals of Eugenetics*, **15**:323-354.
62. Yongfeng H, Jianwei W (2010). Temporal variation in genetic structure of the Chinese Rare Minnow (*Gobiocypris rarus*) in its type locality revealed by microsatellite markers. *Biochemical Genetics,* **48**:312-325.
63. Ywamoto EM, Myers JM, Gustafson RG (2012). Resurrecting an extinct salmon evolutionary significant unit: archived scales, historical DNA and implications for restoration. *Molecular Ecology*, **21**:1567-1582.

# Chapter 3

## Final Remarks

## 3.1. Final Remarks

This study allows a better understanding of the dynamics and evolution of the species *Iberochondrostoma lemmingii* through time in the Guadiana Basin and a better comprehension of the environmental impacts over natural populations. Also, it illustrates the value of using archived museum samples for genetic analysis through time, providing a valuable temporal dimension when examining evolutionary processes.

Firstly, data showed that the genetic diversity of *I. lemmingii* remained high and unchanged through time between the two data sets. It is important to point out that this species has a wide distribution range, occupying diverse habitats (Robalo *et al*. 2008, Lopes-Cunha *et al*., 2012). Also, its' high population sizes, high levels of fertility and adaptation to the annual cycle of seasonally predictable floods and droughts, may explain this species resilience to the extreme drought. Further, the 14/15 years that passed between sampling dates may not be enough to cause a significant change on these populations' genetic diversity.

Secondly, it is also important to stress out that an increase through time in population structure was observed for this species, which is likely to be related with anthropogenic pressures. The Alqueva dam was built and started working within the time period addressed by this investigation and dams are widely recognized as one of the most crucial and widespread intentional impacts caused by humans on natural environments (Petts, 1984), since they may affect natural processes of fish populations, like the triggering of fish migration and spawning and the rejuvenation of floodplains (Bunn & Arthington, 2002). Also, they act as a barrier, causing reduced availability, connectivity, and distribution of *refugia* and reduced mobility, which origins the reduction migratory potential and thus, the decrease gene flow among the most isolated populations.

Finally, it is essential to have in mind that although this species has shown to maintain its high genetic diversity values, it is still crucial to develop a conservation program making this species a target of concern. Mainly, because of the intermittent character of this species' habitat, their populations are more vulnerable to stochastic events, such as the environmental threats known, and consequently to sudden population collapses and losses of genetic diversity, increasing their extinction risk (Lopes-Cunha *et al*., 2012).

## 3.2. References


1. Bunn SE, Arthington AH (2002). Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management*, **30**(4):497-507.
2. Lopes-Cunha M, Aboim MA, Mesquita N, Alves MJ, Doadrio I, Coelho MM (2012). Population genetic structure in the Iberian cyprinid fish *Iberochondrostoma lemmingii*

(Steindachner, 1866): disentangling species fragmentation and colonization processes. *Biological Journal of the Linnean Society*, **105:**559–572.

3. Petts GE (1984). Impounded Rivers: Perspectives for ecological management. Wiley, Chichester, UK.
4. Robalo JI, Doadrio I, Valente A, Almada VC (2008). Identification of ESUs in the critically endangered Portuguese minnow *Chondrostoma lusitanicum* (Collares-Pereira 1980), based on a phylogeographical analysis. *Conservation Genetics*, **8**:1225-1229.