The University of Hull

Morphological and ecological divergence in the hybridogenic fish complex *Squalius alburnoides*

being a Thesis submitted for the Degree of

Doctor of Philosophy, Biological Sciences

in the University of Hull

by

Victoria Alexandra Joana Smith

B.Sc. (Hons)

April 2012

Abstract

Polyploidy in animals is rare and often associated with asexual reproduction in all-female lineages. Although some authors believe these lineages to be evolutionary dead ends there is increasing evidence that some species can adapt efficiently and ultimately evolve into a new sexually reproducing species. The hybrid fish complex *Squalius alburnoides*, endemic to the Iberian Peninsula, is one such complex, that is composed of several groups with varying genomic compositions of the maternal and paternal ancestral genome. This, coupled with novel reproductive tactics has lead to the presence of diploid, triploid and tetraploid individuals in a number of populations around the peninsula.

One of the objectives aimed to identify if morphological differences could be found with the species and whether it would be possible to identify the ploidy of an individual using morphological features alone. Differences between two ploidy groups, diploid males and triploids, were demonstrated using morphology of scales, head and body from fish collected from the Guadiana river basin in Portugal, using geometric morphometric methods. The morphological differences found using scale shape allowed the creation of a prediction function, capable of identifying the ploidy group of individuals of unknown ploidy using scale morphology alone, with a confidence rate of 75% and 92% respectively for diploid males and triploids. Diploid females were accurately classified 6.3% of the time, due to morphological similarities with triploids, with which they were often mistaken for. Ecological analyses on dietary and habitat selection were coupled with age and growth data to further discover if differences between ploidy levels existed. Results showed that there were differences found, and

that dietary composition could be linked to habitat selection and morphological adaptation.

This thesis has expanded the available knowledge of the species, specifically with relation to morphological differences. It has also provided a framework for future identification of *S. alburnoides* individuals without the need for invasive and expensive methodologies.

Acknowledgements

I would like to start by thanking my supervisors for their help and support. Prof. Ian Cowx for showing me how to electrofish, how to collect my data and also forgiving me for the small electric shock I gave him whilst teaching me to fish. Dr Stuart Humphries for taking me on, believing me when I said I probably would bother him too much. He has proved invaluable during my last year. Dr Leandro Monteiro for initially giving me the opportunity to fulfil this research and being forever patient when I knocked on the door, complaining of some computer malfunction or another. I must thank Dr Andy Nunn for his help when it came to dissecting fish and identifying their stomach contents. Thank you to NERC for providing the funding for all the research.

Prof. Maria João Collares-Pereira, “obrigada” for providing me with an invaluable collection of samples, without which this research would not have been possible, and helping to organise my sampling trips in Portugal. Thank you to Miguel Santos for sending me scales from his fish, and Luis da Costa for all the help during sampling. My thanks go to everyone at the Faculdade de Ciências at the Universidade de Lisboa for making me feel welcome during my stay there, and to Tomás Pais de Azevedo (priminho!) for introducing me to everyone before I even arrived!

In Hull, thank you to everyone in the lab that has provided help, encouragement and laughter over the past few years. Particular thanks go to Ryan, who helped me a lot when I had no clue on how to age my fish, always willing to come and help. To Alan, who will have to find someone else to make his coffee now, and to Julia for always putting a smile on my face.

To my friends and family that have always been there for me, no matter what is happening. To my parents, for encouraging and providing me with the opportunity to do all I have done. Thank you also for bringing over a suitcase full of fish scales, amongst your clothes, so that I could have even more data to analyse! Lastly, but by no means least, to Jonathan, for all your encouragement, support, love and even sometimes for being my P.A.!

Table of Contents

List of figures

Chapter 3: Scale differentiation using geometric morphometrics

List of tables

Chapter 1: Introduction

1.1 Hybridisation and polyploidy

Speciation, the evolution of a new species from an existing one, is a very complex process. Understanding speciation and the processes involved can allow us to infer how species in existence today developed. Conversely, by observing species present in a system, and monitoring and predicting how they may evolve, we can enhance the existing knowledge of speciation mechanisms. There are many ways in which speciation can arise such as parapatric speciation, where a small subsection of a population becomes isolated from the main population and evolves in its new environment, ultimately forming a new species that is reproductively isolated from the ancestral species (Thanukos, 2008). Sympatric speciation is also a common mode of speciation (Gavrilets, 2003), whereby reproduction isolation occurs in a species without the influence of geographical barrier, and it is dependent on factors including habitat differentiation and sexual selection (Coyne and Orr, 2004). A wide range of literature exists on speciation (Sousa-Santos, Collares-Pereira, and Almada, 2006b; Sobel et al., 2010; Mallet, 2007; Ungerer et al., 1998; Carroll et al., 2007), often regarding species or family specific speciation mechanisms (Zardoya and Doadrio, 1999). Hybridisation is one mechanism driving sympatric speciation and arises from interbreeding of two existing species (Schaefer et al., 2009). The occurrence of hybridisation is less than many other speciation processes due to a reduced sterility of the hybrids resulting (Schlupp, 2005) from disruptions to the normal meiotic process (Sobel et al., 2010). However, mechanisms have evolved in a number of species to overcome this difficulty. In fact, certain species may even benefit from the disruption, leading to the creation of new species that can be well adapted to their habitat (Joly

and Bruneau, 2004). Speciation by hybridisation can occur in two forms: polyploid speciation and recombinational speciation (Coyne and Orr, 2004). Recombinational speciation is generally perceived to be due to the hybridisation of two species, where unequal amounts of chromosomal material are provided by each ancestor (Mallet, 2007; Ungerer et al., 1998). The focus of this thesis will be on polyploidy, in which species have more than two copies of the parental genome. Not all polyploid species have arisen through hybridisation but knowledge of their biology and adaptations for survival can provide an insight into how polyploidy may lead to advantages for hybrid individuals. An example of a polyploid species is the treefrog *Hyla versicolor*, that shows individuals with two, three and four copies of the genome within the species (Ptacek et al., 1994). The loach *Misgurnus anguillicaudatus* has shown evidence of the existence of hexaploid individuals, with six copies of the genome present within cells of individuals samples (Abbas et al., 2009). The majority of evidence for polyploid species is found in examples from the plant world, where polyploidy is extensive (Joly and Bruneau, 2004; Sobel et al., 2010), however examples such as those mentioned above display that the process is not exclusive to flora.

Due to disruption in the normal meiotic process, where chromosomes are unable to pair up, many hybrids are sterile (Vrijenhoek, 1989). On the other hand, this disruption is associated with the origin of asexually reproducing all female hybrid lineages of which around 90 are known in vertebrates (Lampert, 2009; Vrijenhoek and Dawley, 1989; Cunha et al., 2011; Janko et al., 2007). In these groups, reproduction occurs without genetic recombination using one of three main reproductive pathways; parthenogenesis, gynogenesis and hybridogenesis (Dawley, 1989). In parthenogenesis, no sperm is required to stimulate embryogenesis of a clonal egg. Gynogenesis is a

similar process in which only the female genome is contributed to the offspring but sperm from a related species is required to initiate embryogenesis. Hybridogenesis is the mechanism by which the genome derived from one parental species is transmitted to the egg without recombination, while the genome from the other is discarded and replaced by fertilisation. As such, all offspring are produced clonally and the lineages are unisexual. From an ecological point of view, hybrid unisexual lineages are often very successful, outnumbering the parental species in their native habitat (Doeringsfeld et al., 2004). They may also be able to occupy intermediate niches as they can show characteristics from both parental species (Schultz, 1971; 1977; Vrijenhoek, 1989). The evolutionary impact of these species is however controversial. One view is that they face many constraints due to the lack of genetic recombination and a dependency on a bisexual species to continue clonal reproduction (Christiansen and Reyer, 2009). It has also been suggested that the unisexual lineage is a temporary vehicle in a process where normal meiosis and sexual reproduction can be restored by tetraploidization (Schultz, 1977). Once this is achieved, normal reproductive processes will be resumed and the resulting polyploids may rapidly outcompete the unisexual and parental lineages. A small number of species of cyprinid and salmonid fish as well as some frogs are all successful sexually reproducing polyploids although their evolutionary history is too ancient to determine whether they arose from hybrid crosses (Schultz, 1977; Lampert, 2009; Morozov-Leonov et al., 2009).

Although a number of polyploid hybrid species have been identified, especially in fish, and it is theorised that many of these species can provide explanations as to how speciation mechanisms occur, very few of these species are the target of research. This is often due to the difficulty of identifying the species in the field, where all ploidy

levels may look outwardly similar (e.g. Alves et al., 2002), or where the polyploid/ hybrid individuals may share many morphological and ecological characteristics with their ancestors (Schlosser et al., 1998). In many cases, the lack of basic ecological information about a species, and how individuals within the species occupy potentially different ecological niches, hinders research regarding speciation mechanisms and how polyploid asexual lineages eventually establish a successful sexually reproducing populations. In total, there are over 160 species of polyploid fish. Many of these species are well established polyploids, such as sturgeon, all members of the Acipenseridae family, which have normal reproduction processes despite the tetraploid genotype. *Poecilia formosa* and *Phoxinus eos-neogaeus* are two polyploid species, both displaying diploid and triploid genotypes that are the focus of research for some aspects of evolutionary mechanisms for polyploids. *P. formosa*, also known as the Amazon molly, was the first vertebrate discovered to reproduce clonally. Like some ploidy levels of *S. alburnoides*, they reproduce gynogenetically. *P. eos-neogaeus* is a polyploid hybrid species that regularly reproduces with its parental ancestors, *P. eos* and *P. neogaeus*. In polyploid fish species with several ploidy levels present within the species, it has been established that the hybrid lineages often have higher levels of genetic diversity than the parental species (Doeringsfeld, 2004). It is not clear however whether different ploidy levels within a species show high levels of genetic diversity. In *P. eos-neogaeus*, genetic diversity in diploids is low, contrasting with high genetic diversity within triploids (Lampert, 2009). In *S. alburnoides*, the unique reproductive modes allow for high genetic diversity within the species. Research has shown that the non-hybrid diploid males exhibit a lower level of genetic diversity than all other ploidy levels and that all *S. alburnoides* ploidy levels have a higher genetic diversity than the parental species *S. pyrenaicus*.

What research has been carried out on these and other polyploid species involves expensive genetic testing which can also be time consuming. There is a need for a cheaper and quicker alternative methodology to be developed, that can be used to better understand the ecological choices and preferences made by polyploid species.

1.2 Aims and objectives

Using the cyprinid *Squalius alburnoides* complex from the Iberian peninsula as an example of a polyploid species, this thesis will focus on how individuals in a polyploid complex can be discriminated, on a variety of levels, including detailed morphological and ecological segregation. The aim of the study is to better understand the biology of a polyploid species, so that it can be used in future to further knowledge on polyploid speciation. The study of the ecology of *S. alburnoides* may also provide valuable information should extensive conservation action become necessary for the species. *S. alburnoides* is currently classed as Vulnerable on both the Portuguese and Spanish Red Lists (Cabral et al., 2005; Doadrio, 2002) and the IUCN Red List (Crivelli, 2011). It is predicted that environments that the species currently inhabit will be subjected to a great deal of change over the coming years, due to the introduction of exotic species and the increase of water abstraction facilities and dams for human consumption and the exploitation of water resources for recreational and tourism purposes (Collares-Pereira et al., 2000; Morais et al., 2009; Chícharo et al., 2006). Management measures on how to cope with such changes whilst helping to further conservation of endangered species have been laid out in Collares-Pereira et al. (2000). It has been recognised that to achieve these aims, it is necessary to understand the ecological requirements of species in the system under threat.

The objectives of this thesis, on how further knowledge on *S. alburnoides* will contribute to the field of polyploid identification are:

- To develop a method of ploidy identification that does not involve traditional analysis of genomic composition and that can be used for a number of species of polyploidy fish in multiple river basins.

- To confirm that the method of identification developed is valid by using it to look at differences in ecological preferences within a species, specifically the *S. alburnoides* complex.

- To investigate whether any morphological features related to ploidy correlate with dietary composition and/or habitat use.

These will be achieved by fulfilling the following aims:

- To establish a detailed framework for the method of ploidy identification, using morphological identification of scale, head and body morphology, which are non-invasive methodologies.

- To determine if different ploidy groups of *S. alburnoides* exhibit differing age and growth characteristics.

- To determine if dietary composition differs between ploidy levels.

- To determine if microhabitat selection within a river is ploidy dependant and what ecological factors influence this selection.

1.3 Thesis structure

This thesis presents a variety of data collected from the same *S. alburnoides* individuals, starting off with the development of a new methodology for the identification of groups within a polyploid species. Further chapters attempt to confirm if the method established is valid, by examining aspects of the ecology of *S. alburnoides*, and whether ecological differences within the species can be related to ploidy level.

The thesis structures individual data collections in a format that is presented in a way that each chapter can be read as a stand-alone piece of work, with individual introduction, method, results and discussion sections. However, it has been taken into consideration that the individual chapters form part of a complete thesis. As such, general aspects of *S. alburnoides* biology are not mentioned in the introduction of each chapter, and is reliant upon the fact that the reader will have understood and remembered the general biology as outlined in Chapter 2.

The chapter structure of the rest of this thesis is detailed below:

- Chapter 2 - The *S. alburnoides* complex: This chapter details the general biology of *S. alburnoides* and reviews the work done so far on the species. It also introduces the methodology that is common to subsequent data chapters, detailing the datasets used in this thesis.

- Chapter 3 - Scale differentiation using geometric morphometrics: Using scales of *S. alburnoides* individuals of known ploidy, this chapter develops a method that can be used to predict the ploidy level of *S. alburnoides* individuals of unknown ploidy, dependant on river basin of origin.

- Chapter 4 - Body shape discrimination using geometric morphometrics: This chapter establishes whether body and head shape can also be used as tools to determine the ploidy of *S. alburnoides*, examining where differences can be found between ploidy levels.

- Chapter 5 - Growth and ecology of *S. alburnoides*: This chapter uses the information regarding ploidy from previous chapters and determines if differences in growth, diet and habitat exist between individuals of *S. alburnoides* with different ploidy levels.

- Chapter 6 - General discussion: The discussion will bring together the key findings from each data chapter and will attempt to find links between each chapter and how they may influence each other. It will also present the data found in a wider context, considering implications for conservation and furthering the knowledge for not only *S. alburnoides* but other polyploid species as well, making recommendations for future work.

Chapter 2: The *Squalius alburnoides* complex

2.1 Introduction

The family Cyprinidae has some of the highest species diversity levels of any freshwater fish and species are widespread throughout the world (Mesquita et al., 2007). Of the 15 representatives of the Squalius genus, 13 are endemic to the Mediterranean basin with the other two found in northern and central Europe (Sanjur et al., 2003; Mesquita et al., 2005). The Squalius alburnoides species complex (subsequently referred to as "species"), known as "Bordalo" in Portugal (Figure 2.1), is limited to the Iberian peninsula where it has been described as abundant and has been extensively studied in Portuguese rivers and their tributaries (Collares-Pereira and Moreira da Costa, 1999; Alves and Coelho, 2001; Alves et al., 2004; Morgado-Santos et al., 2010; Cunha et al., 2011; Figure 2.2). S. alburnoides is a small species, reaching a maximum of 13 cm, although the majority do not grow longer than 9 cm. The species tends to be found in water courses where emergent vegetation is abundance and where the river is narrow and not too deep. Within a stretch of river there are preferences found within the species for certain habitat characteristics. These will be discussed in detail further on. S. alburnoides are omnivores, although dipteran larvae make up a large part of their diet and once again, dietary differences exist within the species, not only within the same river basin, but also between river basins (Collares-Pereira et al., 2007).

Figure 2.1: Picture of a *S. alburnoides* individual in the laboratory and picture of a preserved fish, removed from the preservative and placed on graduated paper.

The Iberian peninsula is considered to be a biogeographical area distinct from the larger European and African regions due to the natural barriers of the Gibraltar Strait and the Pyrenees (Vargas et al., 1998). There are five major river systems in which *S. alburnoides* is present in the Iberian Peninsula; the Guadiana and Guadalquivir in the south, the Tejo in the centre and the Douro and Mondego rivers in the north. This complex has been show to be a fertile hybrid originating from two other endemic Iberian species (Alves et al., 2002). Mitochondrial DNA (mtDNA) analyses have shown the maternal ancestor to be *S. pyrenaicus*, a sympatric species in the Tejo and Guadiana basins in the centre and at the south of the peninsula (Alves et al., 1997; Alves and Coelho, 2001). Crossing experiments and studies of β-actin genes have revealed that the paternal ancestor is closely related to *Anaecypris hispanica*, a species with a distribution that partially overlaps that of *S. alburnoides* (Crespo-Lopez et al., 2006; Robalo et al., 2006). It has been hypothesised that the species occurred through multiple hybridisation events in different basins involving *S. pyrenaicus* females. A minimum of two hybridisation events occurred, one in the Sado basin and another in the Tejo and Guadiana basins (Alves et al., 1997). More recently studies have shown the likelihood of five independent hybridisation events. This is based on the fact that

mtDNA of *S. alburnoides* is more similar to that of *S. pyrenaicus* from the same river than the mtDNA of conspecifics from other basins (Alves and Coelho, 2001; Cunha et al., 2004). The first individuals to arise were most likely diploid hybrids and originated relatively recently, although more specific time frames cannot be determined due to the greater sequence divergence of hybrid lineages when compared to ancestral species (Alves et al., 1997; Sousa-Santos, Collares-Pereira, and Almada, 2006c). Due to the absence of *S. pyrenaicus* in the northern basins, it has been suggested that *S. alburnoides* dispersed to the Mondego and Douro from the Tejo river system (Alves and Coelho, 2001).

Figure 2.2: Distribution range of *S. alburnoides* in the Iberian Peninsula, shaded in grey, and location of the Iberian Peninsula in Europe. Distribution is based on potential presence of *S. alburnoides*, based on the drainage area of the river.

The hybrid complex is comprised of several ploidy levels and genomic compositions, where ploidy levels refer to the number of copies of the genome found in each cell of individual *S. alburnoides* and where genomic composition indicates the composition of the genome from maternal and paternal ancestry. Individuals may be diploid, triploid or tetraploid, with individuals of both sexes found in each level (Crespo-Lopez et al.,

2006). Reproduction occurs between all genomic possibilities, ensuring the continuing diversity of this lineage. The genomic composition of individuals is designated by varying combinations of the letters A and P, where A represents the genome of the paternal ancestor and P the genome of the maternal ancestor *S. pyrenaicus.* So far five forms have been found in nature (Figure 2.3), with another form (AAA) only found in controlled laboratory crosses (Alves et al., 1999). Reconstituted diploid non-hybrid individuals (AA) represent an all male lineage with only the genome of the paternal ancestor expressed. These individuals are reconstituted from the other hybrid forms and present *S. pyrenaicus* mitochondrial DNA (Alves et al., 2002). The variety of genomic compositions is ensured by the large number of different reproductive methods used by each ploidy level, some of which are ploidy-specific, explained in detail below.

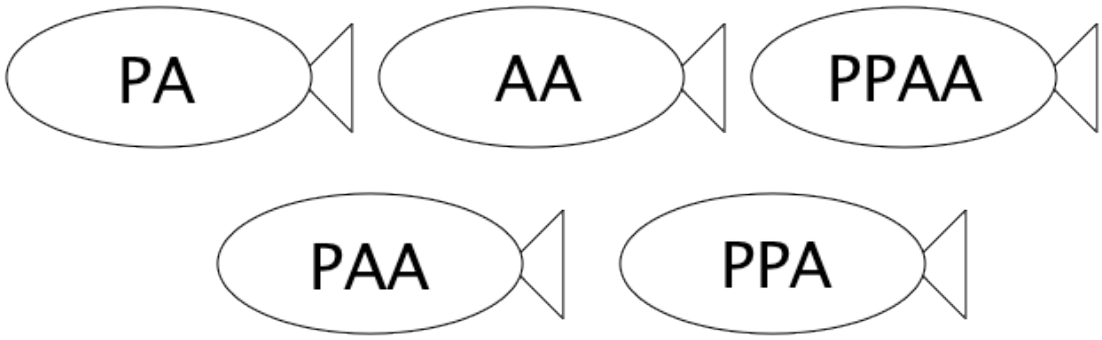


Figure 2.3: Diagram indicating the five *S. alburnoides* compositions found in nature, not including variations from the Northern basins of the Douro and Mondego, as in Alves et al. (2001). A represents the genome of the paternal ancestor and P the genome of the maternal ancestor *S.pyrenaicus*.

2.2 Reproduction

Diploid hybrid males (PA) do not undergo meiosis, thus producing clonal diploid gametes (Sousa-Santos, Collares-Pereira, and Almada, 2006c). However, non-hybrid diploid males (AA) undergo normal meiosis producing haploid sperm. Triploid males

(PAA or PPA) appear to produce unreduced sperm. Due to their rare occurrence in nature, it is still unknown whether they are fertile and contribute to reproduction within the species complex (Alves et al., 1999). As tetraploid males (PPAA) have two haploid genomes of the maternal and paternal ancestors, they undergo normal meiosis, producing PA gametes (Alves et al., 1999; Crespo-Lopez et al., 2006). It is however not expected for PAAA and PPPA tetraploids to produce diploid gametes. No experimental crosses have been carried out with individuals of this genomic composition and as such the outcome of the gametes remains unknown (Alves et al., 1999). The production of gametes in females is a more complicated process where female individuals can produce different types of gametes, including two at the same time (Alves et al., 2004). Diploid females always produce clonal gametes, which can be fertilised by a variety of mechanisms, including sperm from the sympatric species *S. pyrenaicus* (Alves and Coelho, 2001; Figure 2.4). The haploid and diploid sperm will respectively produce triploid and tetraploid individuals. A very small percentage (<3%) of females will reproduce by gynogenesis where the sperm will stimulate embryogenesis without contributing any genetic material (Alves and Coelho, 2001). Evidence from the Tejo and Guadiana river systems indicates that triploid females (PAA) produce haploid or diploid gametes by a modified form of hybridogenesis. The genome of the maternal species is discarded, sometimes followed by normal meiosis, producing haploid gametes ("meiotic hybridogenesis"). Diploid gametes are produced following the usual mechanism of hybridogenesis (Alves et al., 1998). Triploids may also produce gametes clonally which upon fertilisation will yield tetraploid offspring (Alves et al., 2004). Crespo-Lopez et al. (2006) also performed experimental crosses with a PPA triploid female and *A. hispanica* male. In this case, the paternal genome (A) was discarded and the normal meiotic process occurred to produce haploid gametes. As

with tetraploid males, tetraploid females undergo normal meiosis. It is postulated that these gametes may experience gynogenesis, resulting in diploid progeny (Crespo-Lopez et al., 2006).

In the northern basins of Mondego and Douro, where *S. pyrenaicus* is not present, the genome of *S. alburnoides* individuals incorporates the genome of *Squalius carolitertii*, a closely related sympatric species. Individuals undergo similar reproductive processes, however the absence of non-hybrid males (AA) and the rare occurrence of tetraploids reduces population genetic diversity (Pala and Coelho, 2005).

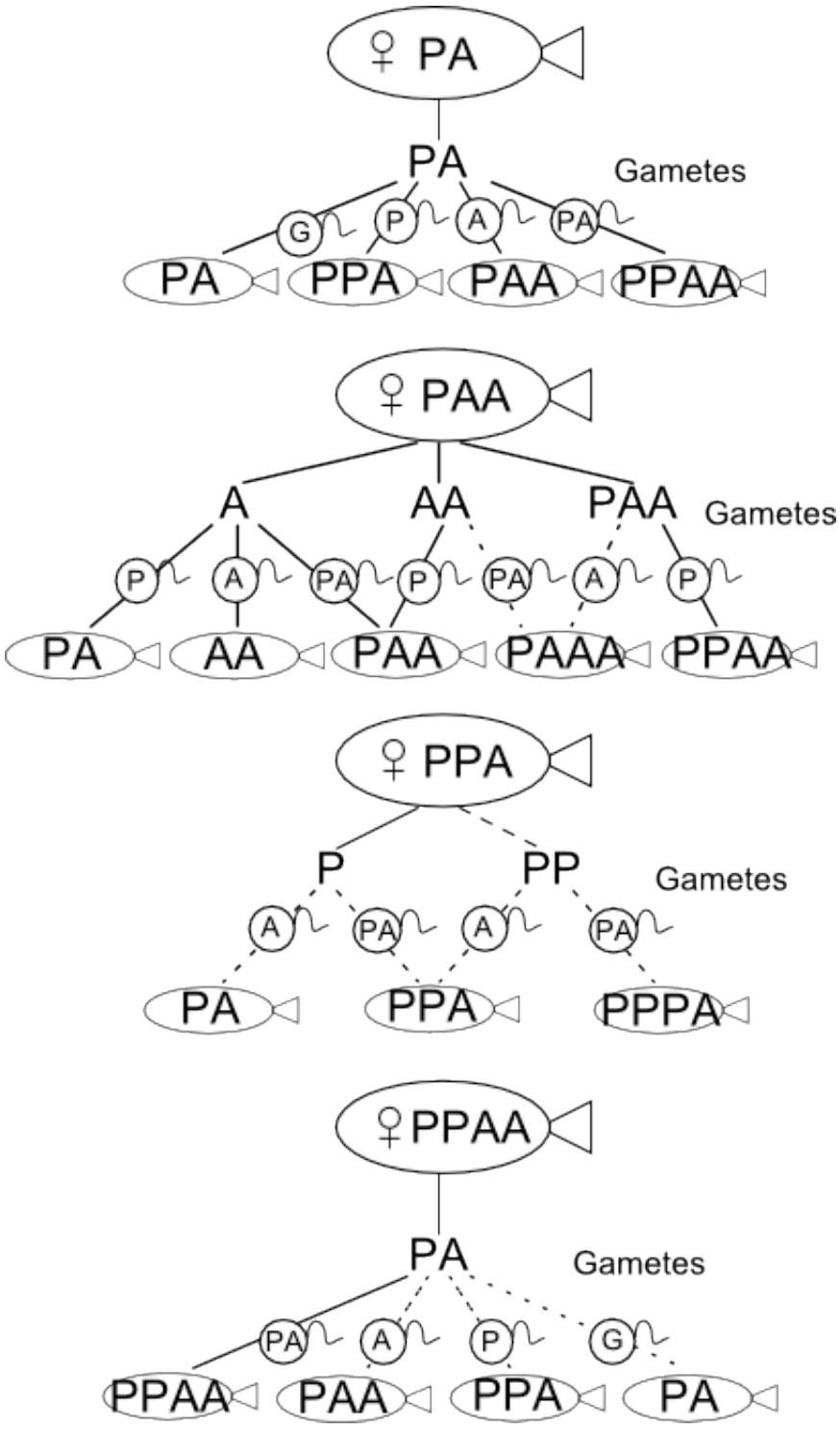


Figure 2.4: Reproductive methods of different ploidy levels and genomic compositions of *S. alburnoides*. A represents the genome of the paternal ancestor and P the genome of the maternal ancestor *S. alburnoides.* G represents fertilisation by gynogenesis, where no genetic information from the male is transferred to the offspring. Pathways in full lines represent reproductive methods shown in laboratory crosses. Dotted lines represent hypothetical pathways.

2.3 Genetic composition of populations

The genetic composition of *S. alburnoides* populations varies greatly between river basins. Female triploids appear the most abundant within all river systems studied in Portugal, ranging in frequency between 50 and 100% (Alves and Coelho, 2001); male triploid counterparts are however very rare. This is true for the majority of genomic compositions, where females are more common than males (Sousa-Santos, Collares-Pereira, and Almada, 2006a). Diploid individuals are an exception, showing a near equal sex distribution (Martins et al., 1998). Tetraploids are absent in the Guadiana basin and show low frequencies in all other basins except for the Tejo river system (Alves and Coelho, 2001; Crespo-Lopez et al., 2007). As previously mentioned, *S. alburnoides* is sympatric with *S. pyrenaicus* in the southern basins and with *S. carolitertii* in the north. In the independent Quarteira drainage in south-west Portugal, *S. alburnoides* shares sperm dependency with *S. aradensis* (Sousa-Santos, Collares-Pereira, and Almada, 2006a).

The frequency distribution of genomic compositions that occur may be linked to habitat preferences. It is not clear whether genomic frequencies are influenced by abundance of *S. alburnoides* in the river system. Even within the same river basin and at different periods of the year, genomic frequencies vary greatly (Martins et al., 1998; Alves and Coelho, 2001), which may be explained by the availability of certain habitats within the system.

The distributions of each ploidy level vary according to river systems. All ploidy levels have been reported for the Tejo in similar distributions. In the Guadiana river system, diploid and triploid females are the most abundant with tetraploid individuals rarely found (Gromicho and Collares-Pereira, 2004). In both the northern basins of Mondego

and Douro, triploid females are the most abundant and non-hybrid AA males are absent (Alves and Coelho, 2001). Due to the absence of *S. pyrenaicus* in the northern basins, sperm dependency has shifted to *S. carolitertii*, a sympatric species only found in the northern basins (Alves et al., 2004). This has led to suggestions that different ploidy levels occupy a preferred habitat, which may eventually lead to the evolution of a new species (Gomes-Ferreira et al., 2005). Despite the recognised need for the analysis of spatial variation (Gomes-Ferreira et al., 2005), few studies have been carried out, mostly due to the fact that different ploidy levels are not recognisable from visual inspection in the field. Diploid "non-hybrid" males are an exception to this, being smaller than hybrid individuals as well as resembling *A. hispanica* (Robalo et al., 2006). Sex recognition in the field is also difficult, unless caught during the reproductive period, which is generally around April for this species (Martins et al., 1998; Ribeiro et al., 2003).

2.4 Habitat and dietary preference

Martins et al. (1998) showed that non-hybrid diploid males, diploid and triploid females were spatially segregated. Diploid males were preferentially found in shallow waters with higher temperatures and organic material. The substratum of the habitat where diploid males were found was often silt and sand, which has also been confirmed by Sousa-Santos et al. (2006a). The few triploid males found also seemed to occupy a similar habitat (Martins et al., 1998). Most diploid females were found in deep waters, where the substrate was mostly composed of rocks and pebbles. Triploid females were found in fast flowing water with high pH and extensive cover. This spatial segregation most likely helps to reduce competitive interactions for the resources available. However, during the summer droughts, fish from different habitats may be

forced together in pools. Triploid females are however rarely found in these pools, reinforcing the notion of flowing water preference. The reproductive period has also been shown to minimise these spatial segregations as both diploid and triploid females prefer shallow waters during this period (Martins et al., 1998).

The preferences for habitat type are thought to be closely related to the feeding preferences for each ploidy level (Gomes-Ferreira et al., 2005). *S. alburnoides* is a generalist feeder, feeding mostly on macroinvertebrates such as Diptera and Ephemeroptera. However, differences between ploidy groups were shown for Diptera adults, Coleoptera adults and plant material. Diploid and triploid females eat more plant matter than diploid males. Diptera adults were preferred by triploid females and diploid males. Evidence has also been found that feeding regimes change throughout the life cycle, with smaller fish inhabiting the river bed and adults feeding from the surface and water column (Gomes-Ferreira et al., 2005).

2.5 General methods

A number of datasets were available for analysis in this thesis, and were used in different sections of the thesis. All data chapters used data obtained from a set of fish sampled during the course of the research for this thesis. This sampling was carried out at one site on two separate occasions - 18th April 2009 and on the 9th October 2009. These sampling trips represented different time periods in the life cycle: April being at the start of the reproductive period for *S. alburnoides* and October the end of the first year growing season. The site sampled was Ribeira da Murtega in the Guadiana river basin in Portugal, a stream with a variety of different microhabitats, based on flow, depth, substrate composition, cover and vegetation (Figure 2.5). This site has been

sampled on numerous occasions for other research related to *S. alburnoides*, including research that has been directly compared to within this thesis (Figure 2.6).

Fish were caught by electrofishing, run off a portable generator (300-500V, 2-3A, DC). The sampling method used was similar to Point Abundance Sampling (PAS) which traditionally involves a discrete approach to a random chosen sampling site followed by rapid immersion of an anode, which stuns the fish in the chosen area so that they can be collected using dipnets (Brosse et al., 1999; Copp and Penaz, 1988; Copp and Garner, 1995). The method was modified slightly so that microhabitat units were sampled and the anode was left in for longer. The anode was submerged and swept around the sample area. A microhabitat unit consists of a smaller area within the stream, where characteristics such as substrate composition and vegetation presence can be quantified, identifying which areas of a river particular species inhabit. Microhabitats chosen for sampling measured about $1m^2$, based on substrate composition and vegetation presence, with approximately homogeneous flow rates across the selected area. The area of $1m^2$ sampled is an approximation based on the effective electric field created by the anode. At each location, an anode was briefly submerged and energised, so as to cover all areas of the selected sampling point, and all fish in the area were caught with a hand net. This process required two people, one to manipulate the electrode, the other to use the net. All sampling areas were approached with the minimum of disturbance to avoid scattering of fish and provide a reliable picture of the fish naturally present.

Figure 2.5: Map of the River systems in Portugal, with the Guadiana river and its tributaries highlighted and showing location of sampling at Ribeira da Murtega.

A total of 24 and 26 microhabitats were sampled in April and October respectively. The sampling points were selected according to characteristics such as flow, substrate and vegetation presence, with the total number of microhabitats selected covering the stream sampling area. For each habitat area depth (m) was measured at five random points within the $1m^2$ area. Substrate type was defined as percentage contribution of mud, sand, gravel, stones, boulder and bedrock (Mullen et al., 2011). The percentage of either emergent, floating or submerged vegetation was also noted as well as the presence or absence of overhanging branches. The current velocity was measured in cm/s at the centre of the sample area. In April, water temperature (˚C) and oxygen content (ppm) were measured at all sites and water conductivity (μScm^{-1}) and pH were

measured at one location in the stream and applied to all sites. These measurements were not available during sampling in October because of equipment malfunction.

Figure 2.6: Pictures of Ribeira da Murtega in April 2009, showing a number of the sampled microhabitats.

All fish from each microhabitat were identified to species level to determine community assemblage and measured to standard length (nearest mm) using a millimetered measuring board to avoid inaccurate measurements due to shrinking caused by preservation techniques. All *S. alburnoides* individuals were placed in individually labelled bags and then placed in 3% formalin solution to preserve shape and halt digestion of any dietary items. Once back at the laboratory, after photographs were taken, fish were transferred to ethanol for ease of use at a later date when imaging and dissection of the fish occurred. Dissection occurred within 3 months of being transferred to ethanol, to ensure preservation of stomach contents.

Information from these fish was used in each data chapter of this thesis. Chapter 3 used the data available in scale morphology, whilst Chapter 4 used body and head morphology of these fish to conduct analyses. Scales were also used in Chapter 5, where they were aged and measured for growth analysis. Information regarding dietary composition, resulting from the dissection of these fish was used for dietary

analyses in Chapter 5, and the microhabitat data collected during sampling was also used in Chapter 5 to investigate differences in habitat within the species. Detailed methodologies on how each dataset was used can be found in the relevant chapters. In addition to the data available from the fish sampled in 2009, Chapter 3 also used information regarding scale morphology from a number of fish collected for other research projects by the Universidade de Lisboa between 1994 and 2009.

Chapter 3: Scale differentiation using geometric morphometrics

3.1 Introduction

There are few visible differences between *S. alburnoides* of differing ploidy levels, and most differences have only been revealed previously with the use of genetic identification. This chapter aims to distinguish between ploidy groups by looking at scale shape using geometric morphometrics. If distinctions can be made, a prediction model will be created to enable identification of the ploidy level of an individual fish without the need for costly genetic analyses. This chapter will also aim to identify whether differences in scale shape exist between populations with variable compositions and evolutionary histories from different river basins in Portugal.

Few studies have attempted to discriminate between ploidy levels of *S. alburnoides* without using techniques such as flow cytometry which is costly and often requires for the fish to be killed. Although fish in this study were dissected for dietary analysis it is the aim of the author to establish techniques that can be used in future studies without the fish being killed. The handful of studies attempting to discriminate ploidy levels without genetic techniques have used direct measurements such as fork length, head length and depth (Martins et al., 1998) or the number of gillrakers (Collares-Pereira, 1984). More recently, Cunha et al. (2010) attempted to discriminate ploidy by using geometric morphometric techniques to describe the body shape of *S. alburnoides* individuals. However, there has been only mixed success in accurately showing differences between groups for the non-genetic based studies.

Over time, a large number of techniques have used aspects of morphology to identify and discriminate between species, populations or even genetically identical stocks of fish (Salini et al., 2004). Most of the work has been carried out on large-scale marine fisheries and uses traditional morphometric methods that focus on distances between fixed locations on the fish (such as fin tip to snout tip) or elliptical Fourier analysis of otolith shape (examples in Murta, 2000; Cardinale et al., 2004; Hermida et al., 2005; Stransky et al., 2008). However, one of the problems associated with these methods is that size differences between individuals are incorporated in the measurements, which may introduce differences in shape that are not present or meaningless. This problem can be overcome by standardising measured distances to a mean length (Hermida et al., 2005). Despite this solution, numerous other problems still persist, such as distances not being homologous among a species of fish, or difficulty in accurately reproducing the methods established by another team of researchers (Cadrin, 2000).

Many of the problems with distance-based techniques were overcome by the development of landmarked-based morphometrics, a technique that selects landmarks at points homologous to other individuals of the same or similar species. A landmark is "a point of correspondence on each object that matches between and within populations" (Zelditch, 2004). Accurate and rapid digitisation methods have also helped the development of this method as computer-based algorithms have allowed for the analysis of digitally processed landmarks. The box-truss method is one method that uses landmarks to infer linear distances (Cadrin, 2000). It has resulted in more accurate discrimination of groups when compared to traditional distance sampling. However, both the box-truss method and traditional distance measurements are still influenced by size, which may potentially discriminate between adults and juveniles

irrespective of any intrinsic shape differences. In addition, when using these methods it is not possible to establish whether shape changes are linked to growth or to an unknown environmental factor. As such, it is important to develop a method in which only shape is used for discrimination, and where size does not play a role.

Geometric morphometrics uses landmark coordinates, retaining the original geometry of the landmark configuration. The landmark coordinates are scaled to remove size differences and rotational and positional effects are also removed before statistical analysis (Zelditch, 2004). Geometric morphometrics can use a number of different criteria to classify the species under study, with those most commonly used for fish being the morphology of the whole fish or head, and otolith and scale shapes. Recently, studies have attempted to discriminate between populations, species or genera using one or more of these criteria (Ibañez et al., 2007; Silva, 2003). Geometric morphometric techniques have been successfully used to discriminate between two species of sturgeon and their hybrid (Costa et al., 2006) as well as between *S. alburnoides* and its sperm donor *S. carolitertii* in the Northern river basins of Portugal (Cunha et al., 2010).

All the landmark studies previously mentioned reinforce the idea that geometric morphometrics is a valid technique to discriminate both potential groups of *S. alburnoides* according to their ploidy level, and to identify differences in shape between individuals from different river basins sampled in the Iberian Peninsula. This chapter describes the methods and results of such a study.

3.2 Methods

To be able to identify the ploidy of fish sampled during this project a database of scales of known ploidy was needed. The database this was compiled from used fish collected

from the Tejo, Sado and Guadiana river basins in Portugal between 1994 and 2009 (details in Chapter 2: General Methods section). 362 fish were available from the Guadiana river basin and 23 and 21 respectively from the Tejo and Sado river basins. These fish were collected by the team at the Universidade de Lisboa for laboratory crosses and research concerning the genetic make-up of *S. alburnoides* (Alves et al., 1997; 1998; 1999; 2002; 2004; Ribeiro et al., 2003). The ploidy level of all fish was determined by flow cytometry (Ribeiro et al., 2003; Collares-Pereira and Moreira da Costa, 1999) and for certain individuals the exact genetic composition was also known.

Scales from each individual fish collected by the Universidade de Lisboa were examined to determine whether ploidy level influenced the phenotype of the fish. To achieve this, the shape of individual scales were determined using geometric morphometric methods. All scales were removed from the shoulder region of the fish, above the lateral line (Figure 3.1). This area was chosen as it is the area in which scales are first deposited in a number of cyprinid species, thus displaying the most growth history (Sire and Arnulf, 1990).

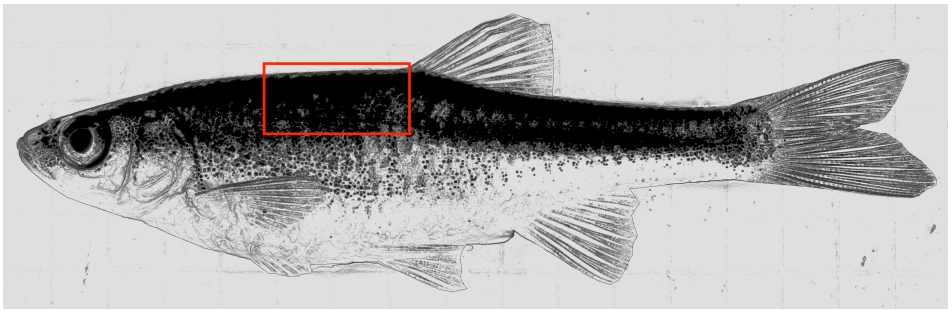


Figure 3.1: Photo of *S. alburnoides* with shoulder area outlined. All scale samples were removed from this area.

Where possible, five scales from each fish were photographed using a Leica microscope S8APO fitted with a Leica camera DFC290. All scales photographed were intact, any broken scales were rejected. The central point of the scale, called the focus, was plainly

visible on each photograph taken. These selection criteria meant that any replacement scales, that grow when the original scale is lost, were rejected (Figure 3.2). This was also considered necessary because replacement scales grow at a faster rate than other scales in the surrounding body areas (Ohira et al., 2007), and it was assumed that this rapid growth could alter the shape of the scale, thus modifying the average shape related to specific ploidy levels. Five scales per individual were chosen to provide an average scale shape in case of deformations that occurred in the normal development of the scale, which could also cause discrepancies in the final scale shape. As all replacement scales were discarded, the number of usable scales from some individuals was reduced, however the majority of fish sampled contained at least five good scales for analysis. The magnification at which all images were taken was noted and a scale bar of 0.5 mm was added to all images. This was later used to calibrate the images and compare shape whilst controlling for size.

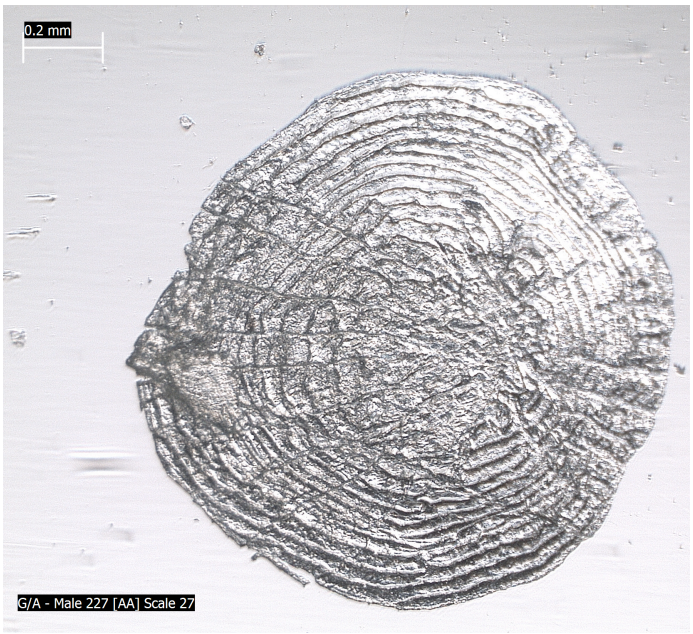


Figure 3.2: Replacement scale from a diploid male *S. alburnoides*. The central portion of the scale shows rapid growth, followed by normal growth thereafter.

Images were then exported to TPSDig2 (Rohlf, 2010) for morphometric analysis of scale shape. Three landmarks and sixteen semilandmarks were placed on each image to determine the shape of each scale. The number of semilandmarks was determined as the minimum number of semilandmarks that could accurately describe the shape of the scales of *S. alburnoides*. The three landmarks are repeatable points consisting of the centre of the posterior edge of the scale (1), the focus (2) and the centre of the anterior edge (3) (Figure 3.3 - A and B). In previous studies where scale shape was used in conjunction with geometric morphometrics, only landmarks were used, placed on clearly defined features along the scale, such as posterior and anterior edges (Ibañez et al., 2007; 2009). Due to the amount of variation present in the scale shape of *S. alburnoides* insufficient repeatable features could be reliably identified to fully describe the overall shape. As such, semilandmarks were used to define the contour of the scale. Eight semilandmarks were placed between landmarks 3 and 1 in a counter-clockwise fashion, equidistant from each other. The same process was repeated from landmarks 1 to 3, thus completing the outline of the entire shape of the scale.

Once landmarking was complete, with each scale being defined by 19 coordinates, these were superimposed using Generalised Least Squares Procrustes superimposition (GLS) in TPSRelw (Rohlf, 2010b), where all size, position and rotation information is removed, leaving only shape information (Zelditch, 2004). Semilandmarks were treated according to the sliding method developed by Bookstein (Bookstein, 1997) before GLS to take into account the reduced accuracy of semilandmarks compared with landmarks.

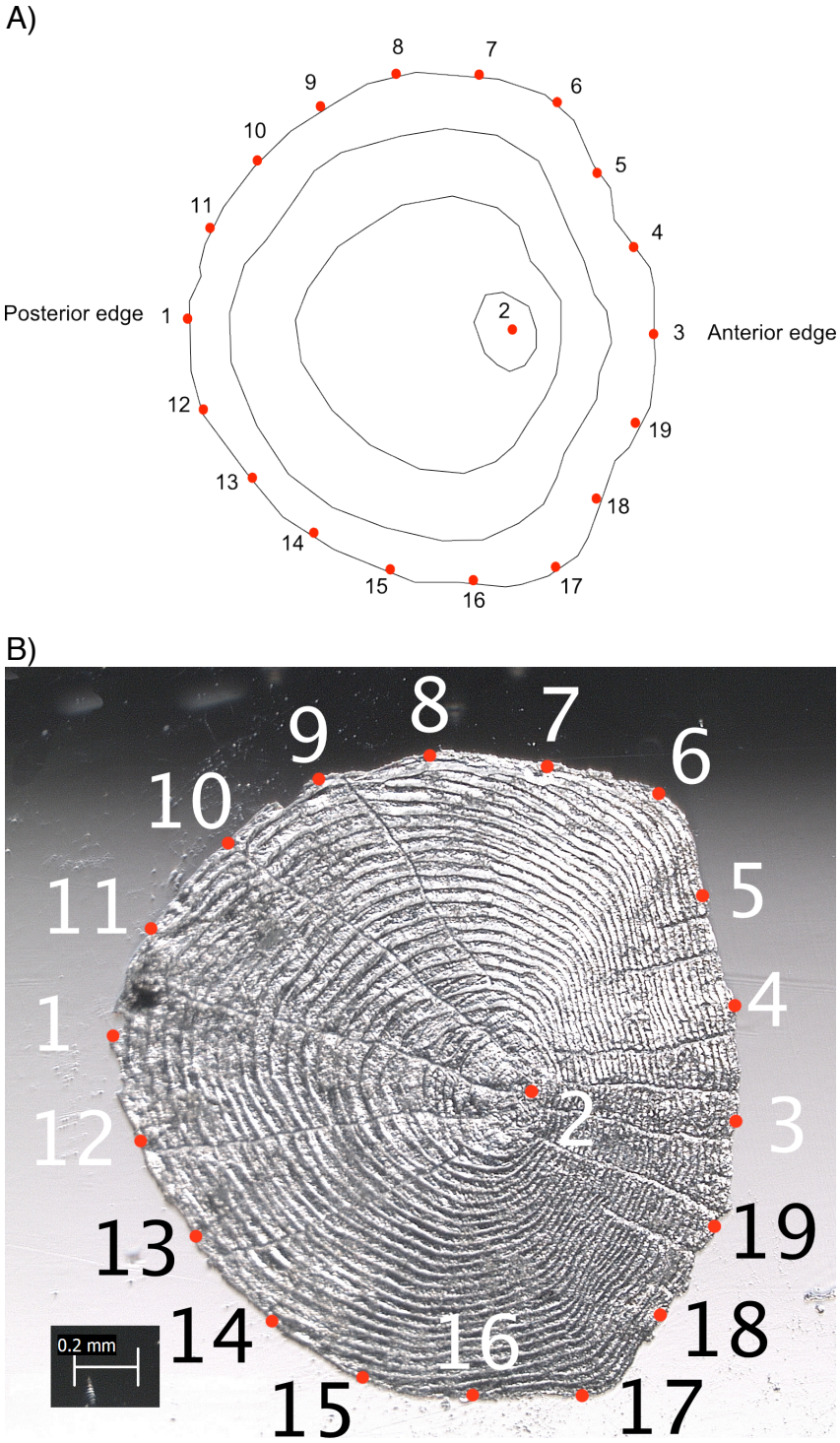


Figure 3.3: A) and B) Landmark and semilandmark positions on a *S. alburnoides* scale. Landmark size on image B) increased for clarity.

Coordinates extracted from GLS were imported to MorphoJ (Klingenberg, 2011) and a Procrustes fit performed to find the mean scale shape and how deviations from this shape occur (Klingenberg et al., 2002). The average scale shape for each individual fish

was determined, using the five sets of coordinates available. A pooled within-groups Principal Components Analysis (PCA) was performed. The pooled PCA analysis, where ploidy is the pooled variable, will allow any differences in shape within a group to be seen. It will thus show the amount of shape variation present within a group, without comparing shape differences between groups (Mitteroecker and Bookstein, 2011; Zelditch, 2004). Pooled results will then take into account any within group differences when between group discrimination is performed. All river basins were analysed separately to minimise within-group variation due to geographic differences within individual groups.

The R statistical software package (R Development Core Team, 2011) was used to perform a variety of analyses that were used to discriminate between ploidy levels as well as to predict the genomic composition of fish of unknown ploidy. The principal component scores were exported to R (R Development Core Team, 2011) for further analysis. Canonical Variate Analyses (CVA) were used to determine between-group variation, highlighting differences in the average scale shape of individuals with different ploidy levels. As a way of quantifying the differences, a Linear Discriminant Analysis (LDA) was performed, which created a function by which individuals could be classified according to scale shape. A leave-one-out cross validation was then performed which attempted to reclassify the ploidy of individuals using the linear discriminant function created (Cordeiro-Estrela et al., 2008; Mitteroecker and Bookstein, 2011).

3.3 Results

3.3.1 Guadiana

A number of discriminant analyses were performed on fish collected from the Guadiana river basin, using different criteria for grouping, with the aim of selecting the most effective for further studies and the prediction of fish of unknown ploidy.

3.3.1.1 Discriminant analysis 1

Fish of known ploidy were collected from a number of locations in the Guadiana river basin. In total 362 fish of known ploidy were collected, of which 163 were diploid and 199 were triploid. The first two principal component scores were plotted as a way of showing within-group variation (Figure 3.4). Both diploid and triploid groups showed a large amount of shape variation along both axes, although the diploid group appeared to demonstrate a wider variation along the first axis, indicating that scales within the group are highly variable, therefore the average shape for the group may not be representative of some individuals. Using the average scale shape for each fish, a Canonical Variates Analysis (CVA) was performed, with fish grouped by ploidy in an attempt to show the difference in shape between ploidy groups. For fish from this river basin only two ploidy levels were available and the CVA plot revealed some overlap between both groups (Figure 3.5). Despite the overlap, there remains still some distinction between the two groups, indicating that morphological differences exist.

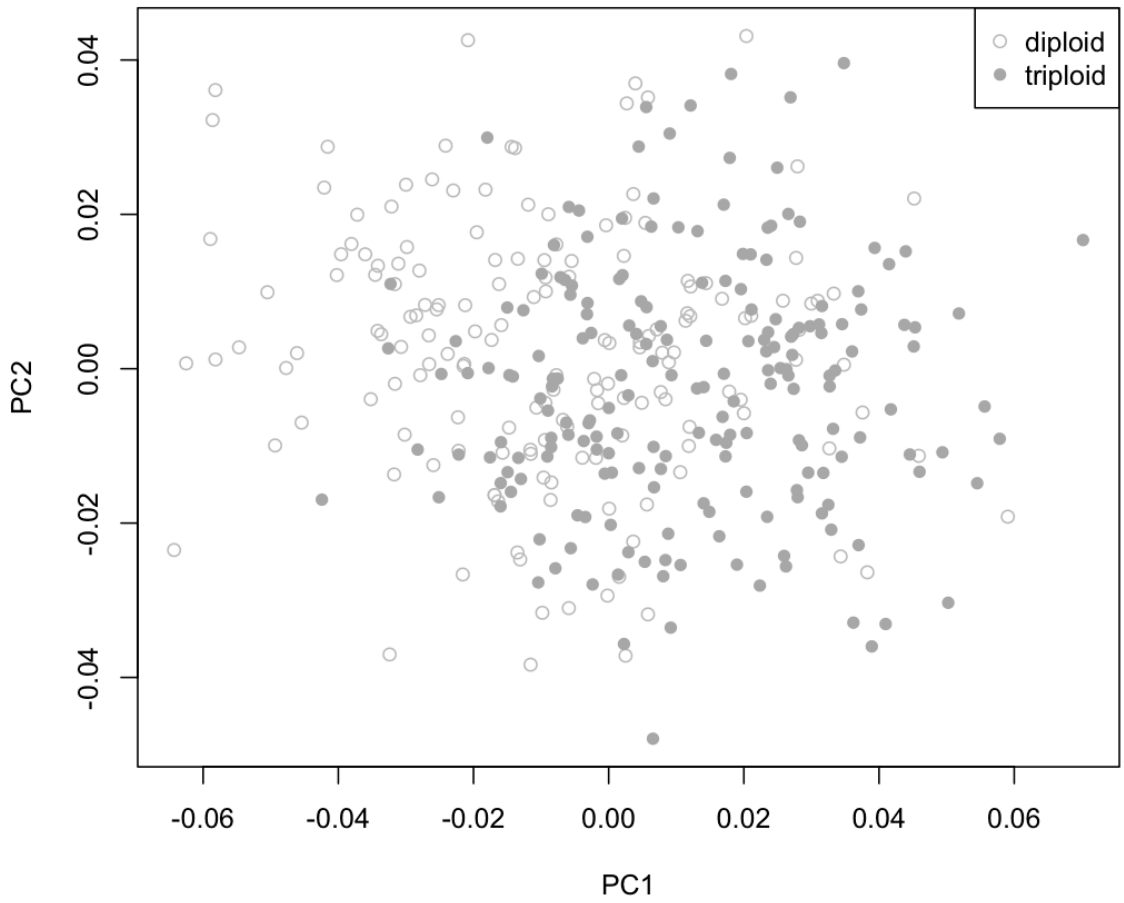
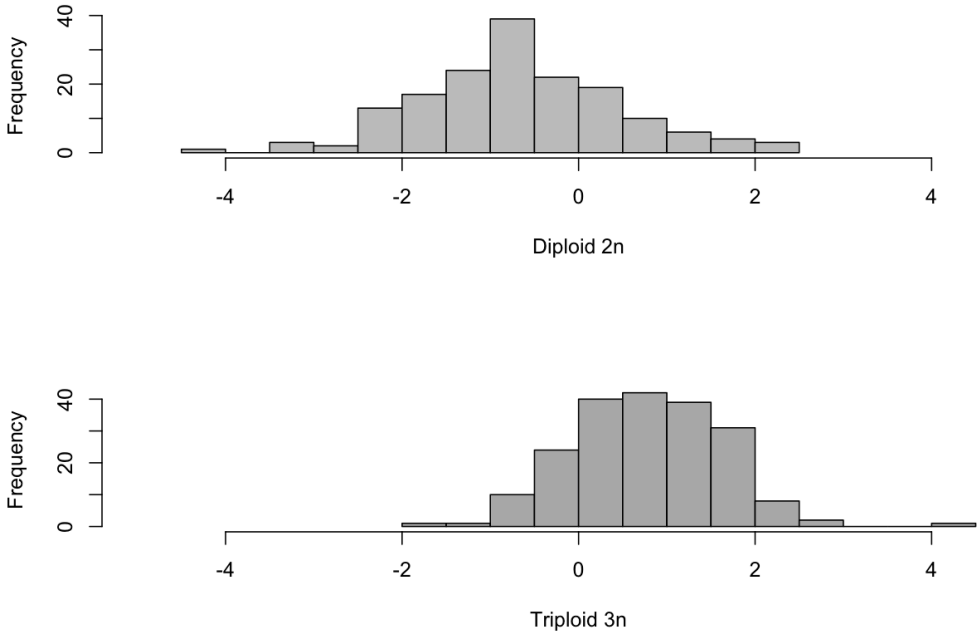


Figure 3.4: Principal components plot of the first two PC scores for fish of known ploidy from the Guadiana river basin, showing within-group shape variation.

A cross-validation table was created, using the LDA function to reclassify individuals to a specific group based on scale shape. The accuracy of the function could be determined by assessing the percentage of correct reclassifications (Table 3.1). The percentage of diploid and triploid individuals correctly classified using this method were 72% and 84% respectively.

A)



B)

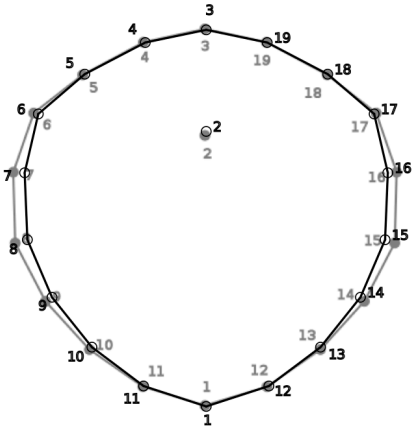


Figure 3.5: A) CVA plot of scales of *S. alburnoides* from the River Guadiana, showing shape differences between diploid (2n) and triploid (3n) individuals. B) Superimposed average shape of triploids and diploids. Black lines correspond to triploids, grey lines to diploids.

Table 3.1: Cross-validation of diploid (2n) and triploid (3n) individuals from the Guadiana river basin, where for a given row, the columns represent the number of individuals assigned to each group.

	2n	3n
2n (N = 163)	118	45
3n (N = 199)	30	169

3.3.1.2 Discriminant analysis 2

The diploid classification makes no distinction between diploid AA males and other diploid individuals. In an attempt to compensate for this, as diploid males are visibly different from other diploids, all diploid individuals of known male sex were classified as 2AA in preparation for a new analysis. This procedure produced three groups were created, with 79 diploids (2n), 84 diploid males (2AA) and 199 triploids (3n). The diploid group could potentially still contain diploid males as diploids of unknown sex were left in this group. The PCA plot shows that this is likely as variation is reduced on the X axis for diploid males and triploids, however diploids still show a large amount of variation, spread along the first principal component (Figure 3.6).

The CVA plot showed overlap between the three groups, with diploid males being slightly more distinct than the other two. The slight overlap of diploids with diploid males is probably caused by diploid males contributing towards other diploids as their sex was unknown. The diploid and triploid groups overlapped but the differing groups could still be partially determined (Figure 3.7). The cross-validation mirrored these findings (Table 3.2).

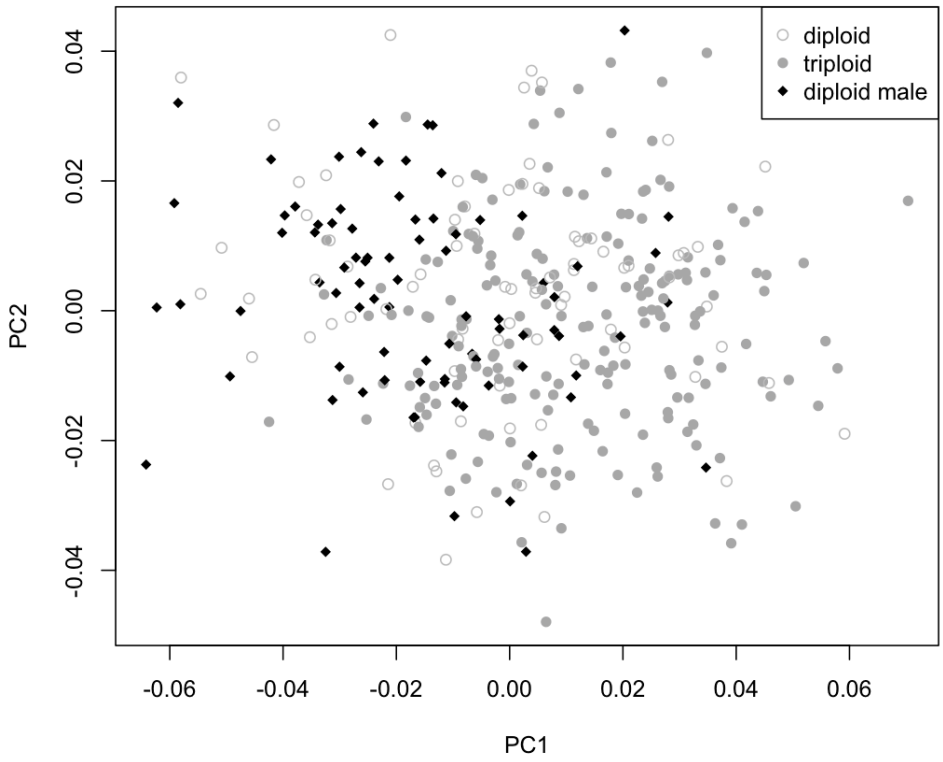


Figure 3.6: Principal components plot of diploid male, triploid and diploid individuals from the Guadiana river basin, showing within-group shape variation.

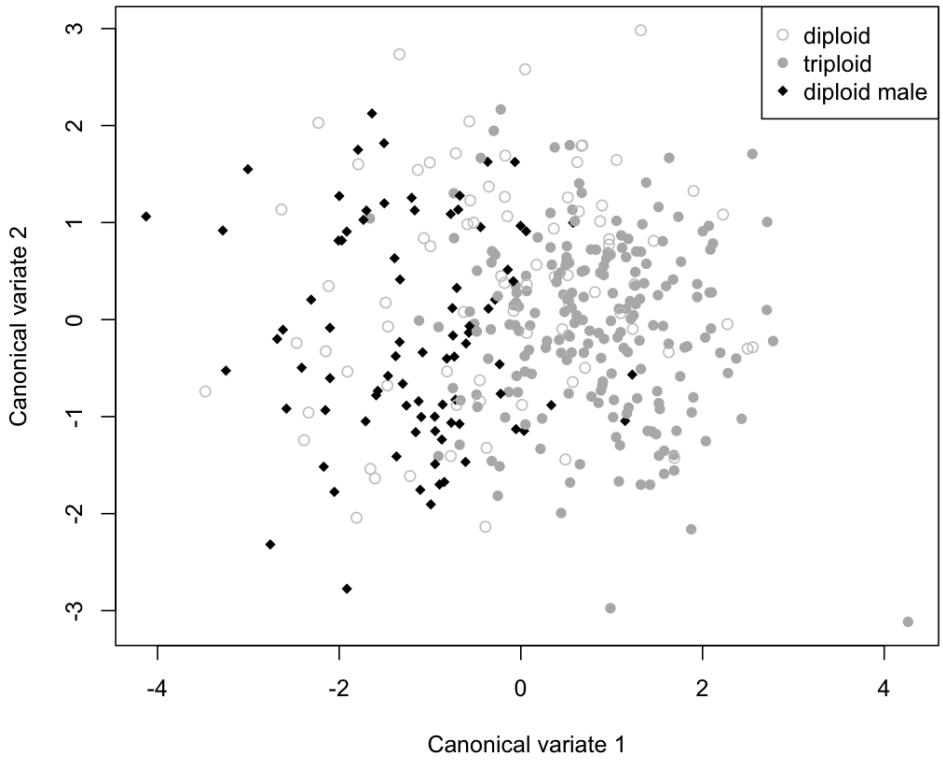


Figure 3.7: CVA plot of Guadiana river, showing shape differences between diploid (2), diploid males (2AA) and triploid (3) individuals.

Table 3.2: Cross-validation of diploid (2n), diploid males (2AA) and triploid (3n) individuals from the Guadiana river basin.

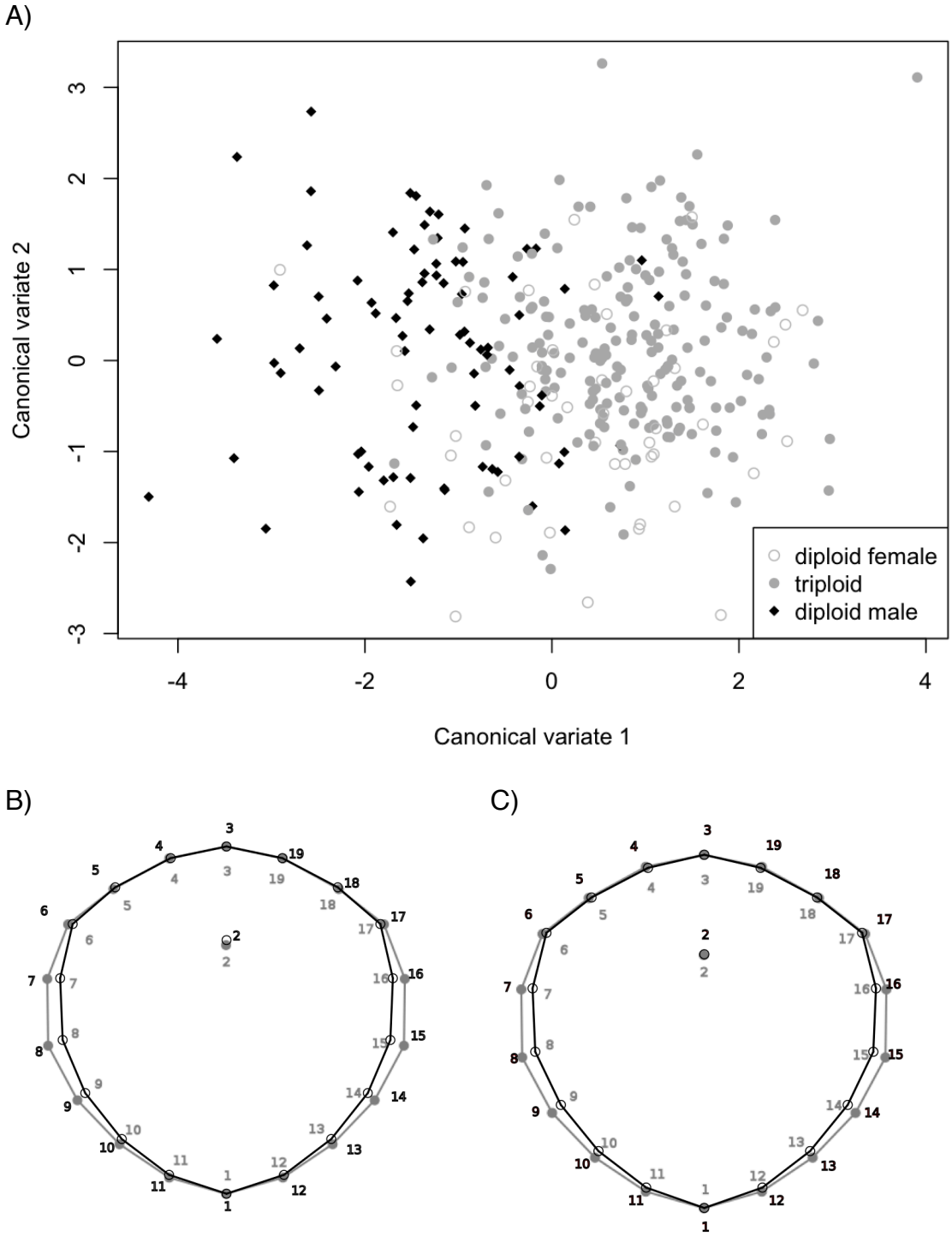
	2n	2AA	3n
2n (N = 79)	11	23	45
2AA (N = 84)	11	56	17
3n (N = 199)	7	7	185

The resulting function was very poor at correctly classifying diploid individuals (13.9% correct), however performed much better for both male diploid (66.7% correct) and triploid individuals (93% correct). The results show that the variation within diploid individuals is large, which may be caused by an initial misclassification of diploid male individuals (genomic composition AA) as diploids (genomic composition PA), explaining the large number of diploid individuals that both the plot and cross-validation appear to class as diploid males.

3.3.1.3 Discriminant analysis 3

The problem of misclassified diploids was avoided by reducing the sample size so that only diploid individuals of known sex were included. Female diploids were assumed to have a PA genomic composition and all diploid males were assumed to have an AA genomic composition. This modified the dataset so that it contained 84 diploid male, 48 diploid female and 199 triploid individuals. A PC analysis shows smaller within-group variation for diploid females when compared to discriminant analysis 2, indicating that the assumption of misclassified diploid males is likely to be accurate (PCA plot not shown). Differences can also be seen using the CVA plot (Figure 3.8), where the overlap between diploid males and diploid females is substantially reduced. A large degree of overlap can still be found between triploid and diploid female individuals, that the cross-validation would be expected to also show. This is the case

(Table 3.3), with high accuracy rates for both diploid males and triploids (75% and 92% respectively). However, once again the accuracy for the remaining diploid individuals remains low, achieving only 6.3% in this analysis.



D)

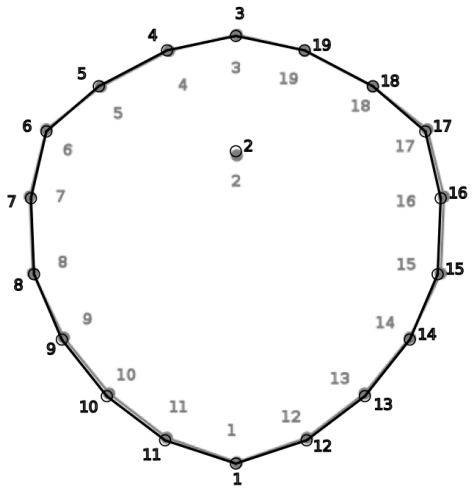


Figure 3.8: A) CVA plot of Guadiana river, showing shape differences between diploid female, diploid male and triploid individuals B) Superimposed average shape of triploids (black lines) and diploid males (grey lines) C) Superimposed average shape of diploid females (black lines) and diploid males (grey lines) D) Superimposed average shape of triploids (black lines) and diploid females (grey lines).

Table 3.3: Cross-validation of diploid females (2n), diploid males (2AA) and triploid (3n) individuals from the Guadiana river basin.

	2n	2AA	3n
2n (N = 48)	3	8	37
2AA (N = 84)	1	63	20
3n (N = 199)	3	13	183

3.3.2 Prediction of Guadiana fish

The last LDA (Discriminant function 3) is the most accurate both in biological and statistical terms and as such is used to predict the ploidy of fish collected in the Ribeira da Murtega in 2009, as well as fish collected between 1999 and 2001 where ploidy was not determined during the original genetic analyses, either due to uncertainty or because a subset of fish were not sampled for genetic purposes. A prediction function was created using the LDA, which assigned a probability for each individual of

belonging to each group. A total of 178 fish were classified in one of three groups (2AA, 2n, 3n), where 2AA corresponds to diploid males and 2n are assumed to be diploid females. The initial classification classed only four individuals as diploid female, 63 were classed as diploid male and the remaining 111 were triploids. However, the probability of belonging to a specific group for some individuals was not much higher than the expected probability due to chance (33.3%). Thus, an assignment cutoff of 60% was used such that individuals were only assigned to a group if the prediction probability was 60% or higher, almost twice that expected by chance. Individuals excluded by the cutoff were classed as inconclusive genomic composition and were discarded from any subsequent analyses where the ploidy could be a potential grouping factor.

The subset selected contained 87 triploid individuals and 50 diploid male individuals. None of the four diploid female individuals achieved probabilities of over 60%. The probabilities for these ranged between 45% and 59%. The full list of individuals where ploidy was predicted, as well as the probabilities, can be found in Appendix 1.

Figure 3.9 provides a graphical representation of the shape variation present for all ploidy levels known. A subset of individuals whose ploidy was predicted is also plotted, with each individual represented by a pie chart of the probabilities of belonging to a particular group. The plot shows that the probability of belonging to a group is directly related to the shape of the scale analysed for prediction, where both male diploids and triploids and much easier to identify than female diploids. The fish whose scale shape are in the overlapping region of the three ploidy levels have probabilities that represent this, making some difficult to confidently identify.

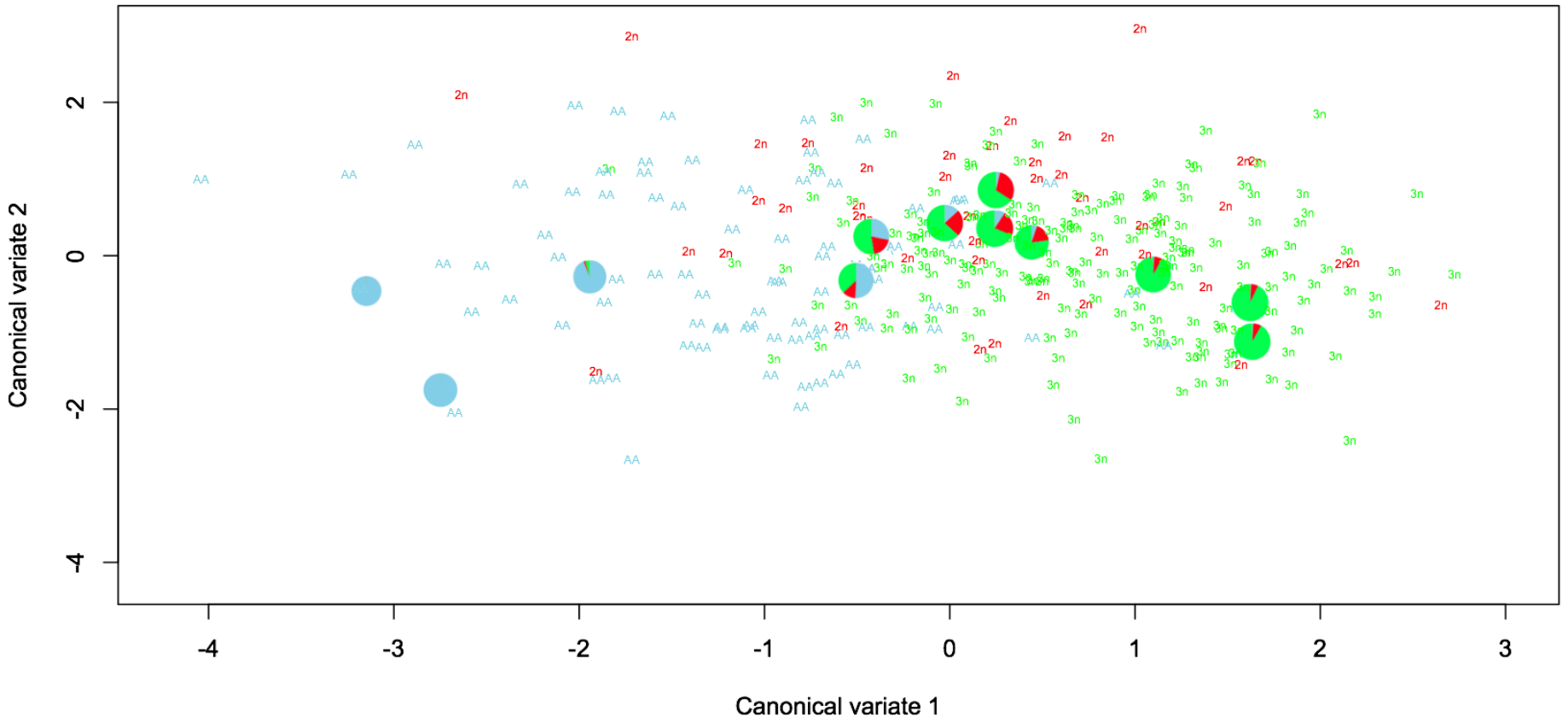


Figure 3.9: Canonical variate plot of fish of known ploidy as well as a random subsample of 12 fish of unknown ploidy. Blue represents diploid males, red diploid females and triploid males are represented by green, both of the CVA plot as well as the overlaying piecharts. So as to better visualise the predictions, the plot was cropped so as to remove a single triploid outlier, of coordinates (5.352, -4.259).

3.3.3 Tejo

No fish from the Tejo river basin were collected during this study, and as such scales available from the Tejo are used only to visualise differences between ploidy groups and not to predict the ploidy of unknown individuals. As with the fish from the Guadiana basin, fish from the Tejo river basin were classified to ploidy level and where possible, genomic composition and sex were identified. The average scale shape from a maximum of five scales was used for all subsequent analyses. A total of 23 fish were analysed, from three distinct ploidy groups (six diploid, 11 triploid, six tetraploid). The exact genotypic composition of all diploid individuals was known and all were found to be of PA genomic composition. There were no diploid males present in the samples collected from the Tejo. As such, it was not necessary to discriminate between both types of diploid individuals in all analyses.

The first four PC scores, representing 95.7% of the total variance, were used for the linear discriminant analysis and CVA plot (Figures 3.10 and 3.11). Analysis was performed on all groups together and the CVA plot showed some discrimination between groups, especially for triploids. However, the scatter of individuals for any one group was wide, therefore any clear and accurate discrimination was not possible.

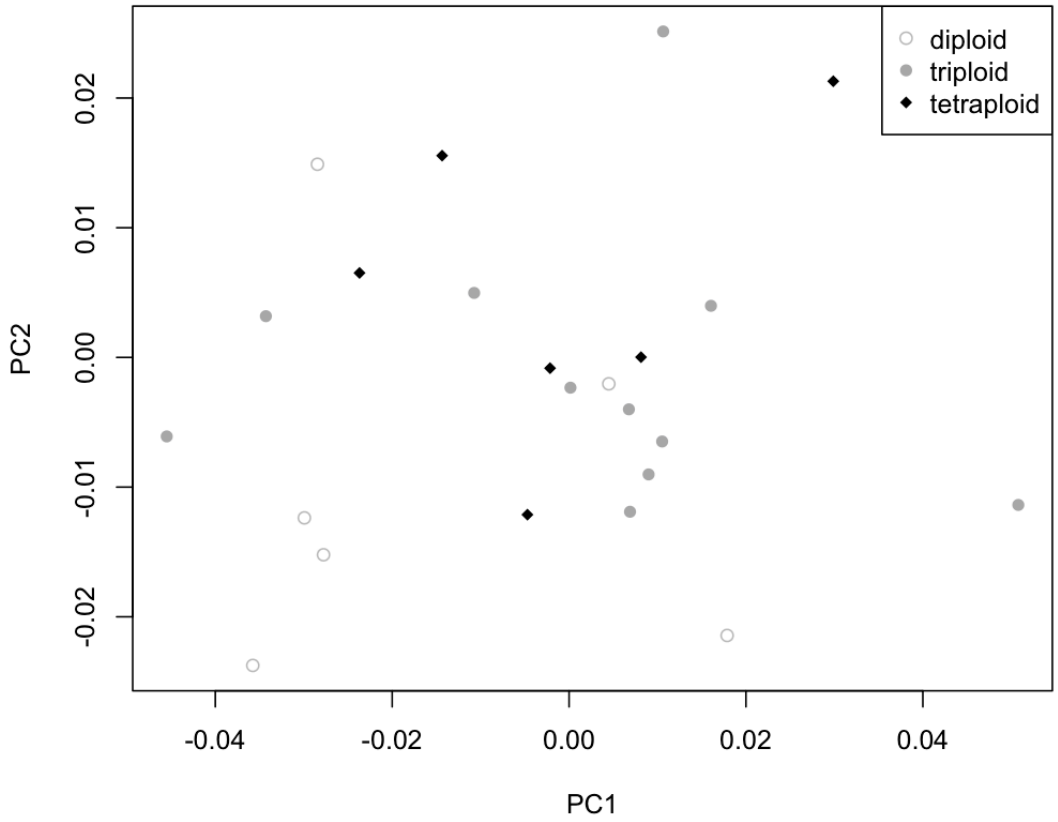


Figure 3.10: PCA plot of Tejo river, showing variation in shape within ploidy levels.

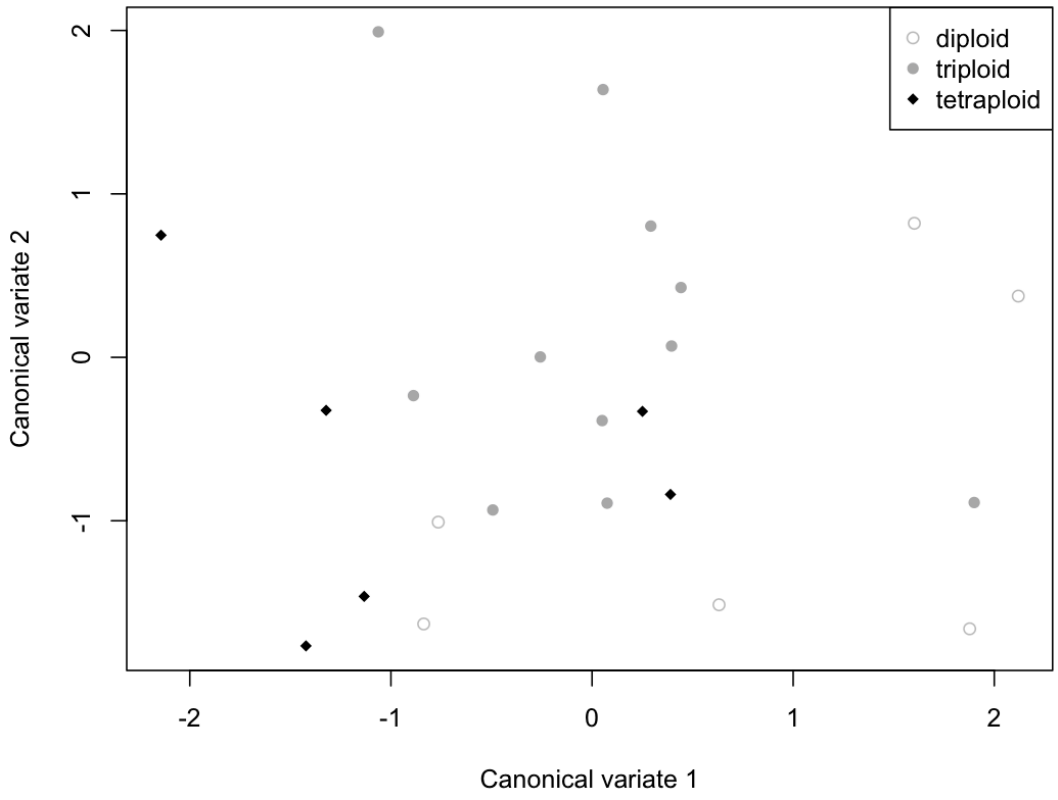


Figure 3.11: CVA plot of Tejo river, showing shape differences between diploid, triploid and tetraploid individuals.

The discriminant analysis produced a similar pattern, with triploids accurately classified 63.6% of the time, a much higher percentage than both diploid and tetraploid (33.3% each of correct classification). The classification was not accurate for all groups, and the percentage of accuracy varied greatly between groups (Table 3.4).

Table 3.4: Cross-validation of diploid (2n), triploid (3n) and tetraploid (4n) individuals from the Tejo river basin.

	2n	3n	4n
2n (N = 6)	2	2	2
3n (N = 11)	2	7	2
4n (N = 6)	0	4	2

Neither the graph nor the cross-validation showed a higher overlap in shape between any two groups when compared to the other. As the percentages of correct reclassification for certain groups was low, it was deemed that the data were insufficient to be able to create a prediction function that would be able to classify unknown individuals in future. It was however possible to view the average shape for all groups, and identify where differences between groups were found on the scale (Figure 3.12). Triploid and tetraploid individuals have the most similar average shapes, whilst diploid individuals display broader average shapes than both other ploidy levels, evident by the position of landmarks 7 to 8 and 15 to 16.

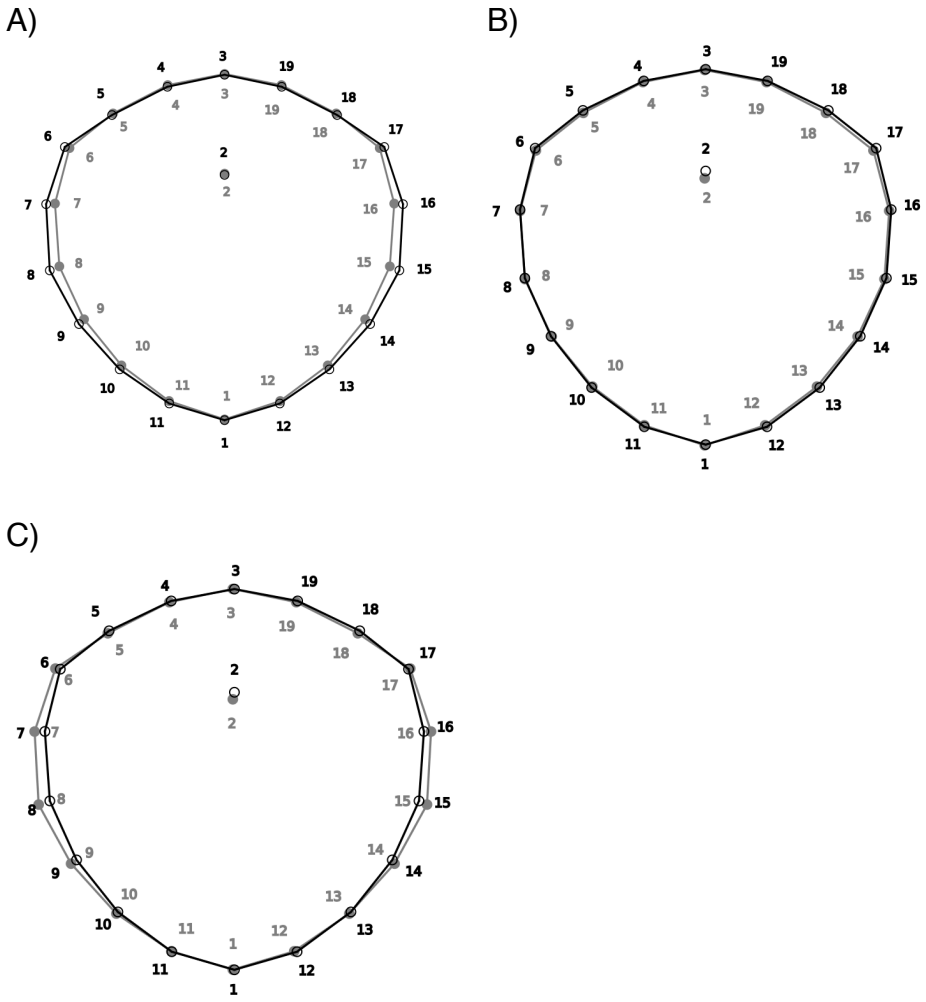


Figure 3.12: Average shape of all ploidy levels from the Tejo river basin, superimposed. A) Superimposed average shape of triploids (grey lines) and diploids (black lines) B) Superimposed average shape of triploids (grey lines) and tetraploids (black lines) C) Superimposed average shape of tetraploids (black lines) and diploids (grey lines).

3.3.4 Sado

As with the Tejo dataset, individuals from the Sado were only used to visualise potential shape differences between ploidy groups in this river basin. No new fish were sampled for this study, therefore a prediction function was not created. The sample size of fish originating from the Sado river basin was very small, being made up of two diploid individuals, 18 triploids and one tetraploid individual. Due to the discrepancy in sample sizes between the three groups it was expected that any analysis would be skewed in favour of triploid individuals, where more information was available. The

CVA plot showed a good separation of the three groups (Figure 3.13). However, the variation in shape present within both diploid and triploid groups was large (Figure 3.14), which may reduce the accuracy of any subsequent cross-validation or predictions. As only one individual represented the tetraploid group it was not possible to determine whether this was a representative shape of this group.

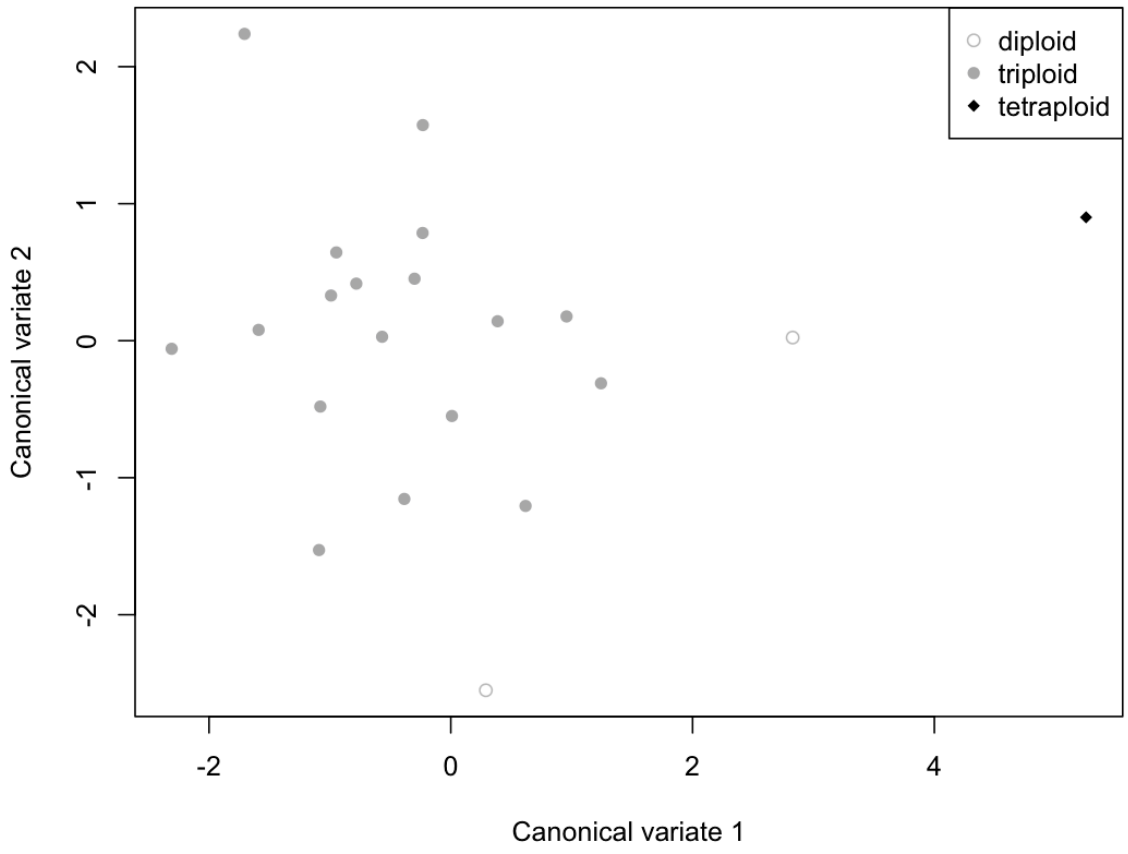


Figure 3.13: Canonical variate analysis plot of diploid, triploid and tetraploid individuals from the Sado river basin.

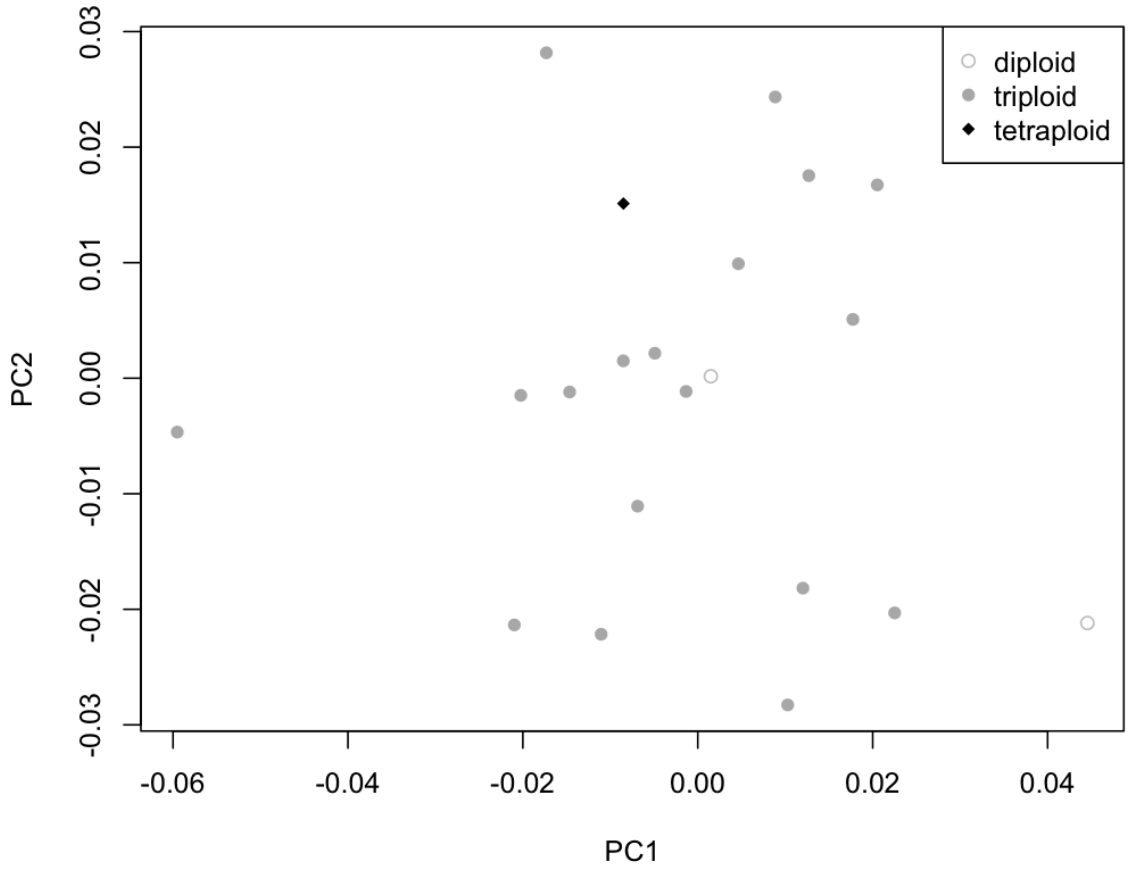


Figure 3.14: PCA plot of Sado river, showing variation in shape within ploidy levels.

The linear discriminant analysis and subsequent cross-validation were unable to reclassify the tetraploid individual, as more than one individual per group is needed for the analysis. As such, the CVA and linear discriminant analyses were repeated, removing the tetraploid individual (Figure 3.15).

The first three PC scores were taken into account, incorporating 95.1% of the variance. As expected, the function was not accurate, classifying all 20 individuals as triploids (Table 3.5).

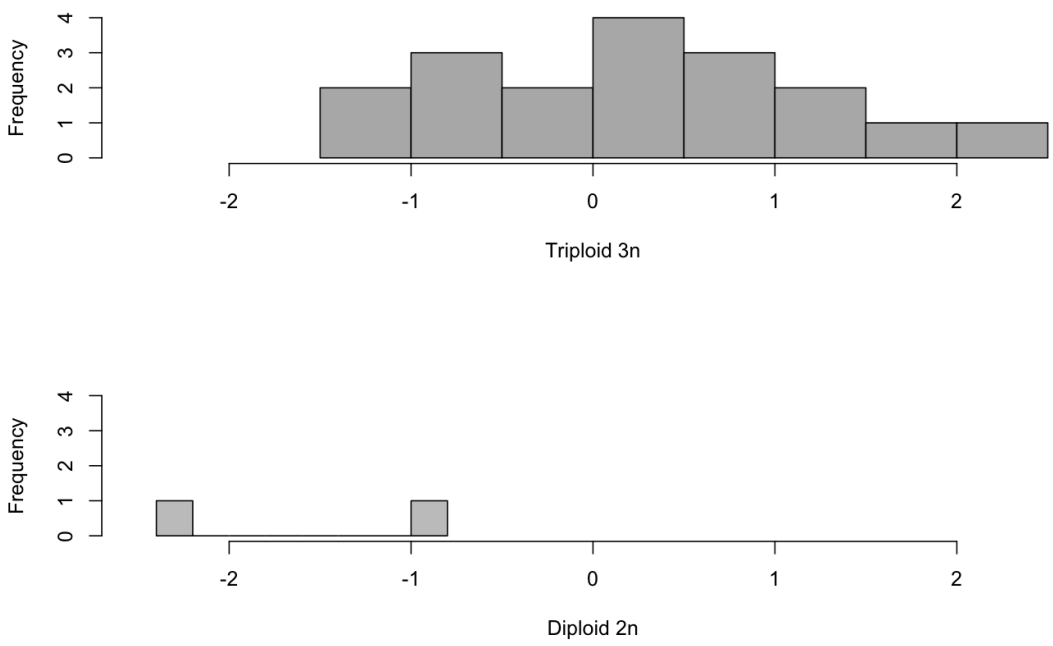


Figure 3.15: CVA plot of diploid and triploid individuals from the Sado river basin.

Table 3.5: Cross-validation of diploid (2n) and triploid (3n) individuals from the Sado river basin.

	2n	3n
2n (N = 2)	0	2
3n (N = 18)	0	18

The results were not sufficiently accurate to potentially use in the prediction of fish of unknown ploidy. In addition, the Sado basin has tetraploid individuals present (Sousa-Santos, Collares-Pereira, and Almada, 2007b) and any prediction function not incorporating tetraploid information would undoubtedly misclassify a large proportion of individuals. Differences between the average morphology of tetraploid scales and those of other ploidy levels are evident in Figure 3.16. Tetraploids display a broader posterior edge when compared to diploid individuals and a broader anterior edge when compared to triploids. Diploid scales are narrower along the length of the scale when compared to both other ploidy levels.

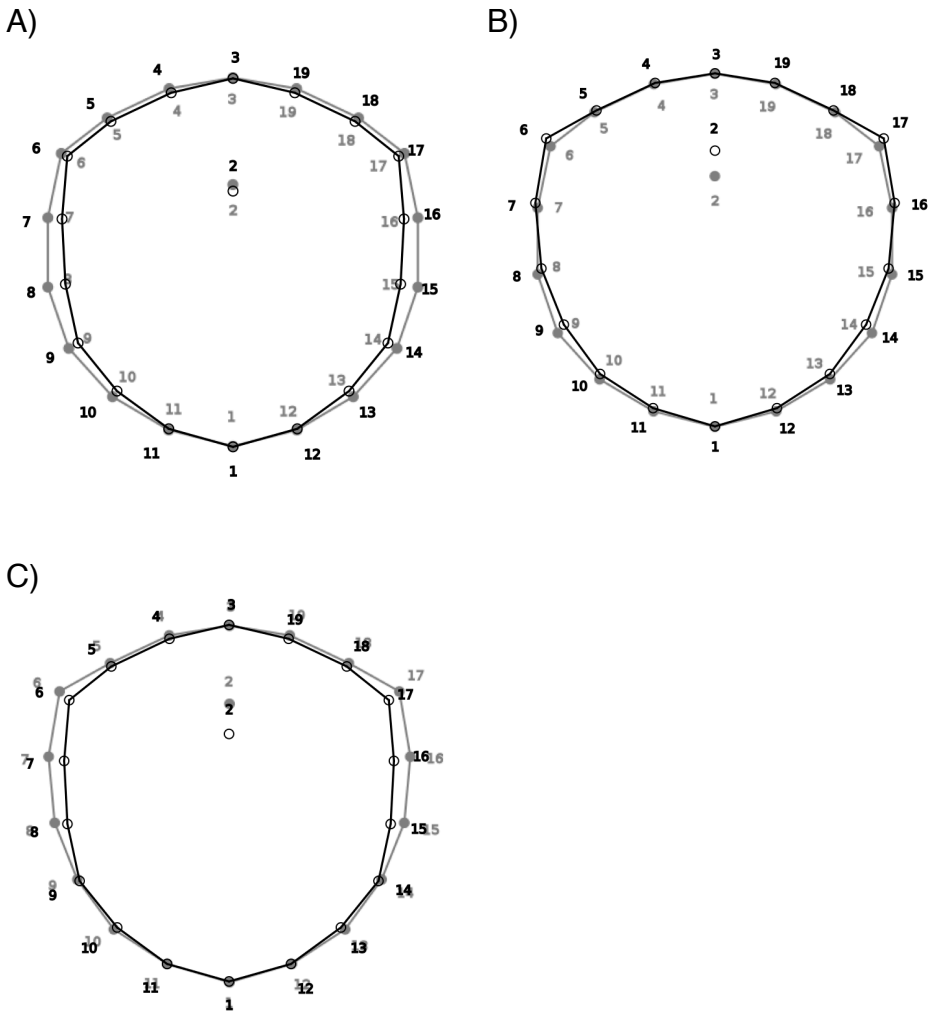


Figure 3.16: Average shape of all ploidy levels from the Sado river basin, superimposed. A) Superimposed average shape of triploids (grey lines) and diploids (black lines) B) Superimposed average shape of triploids (grey lines) and tetraploids (black lines) C) Superimposed average shape of tetraploids (grey lines) and diploids (black lines)

3.3.5 All river basins - Guadiana, Tejo and Sado

In an attempt to compensate for the lack of information from the Tejo and Sado datasets, shape information from all three river basins was pooled together and the same analyses as previously were performed. Diploid fish of unknown sex were not included in the analysis as the Guadiana dataset showed that this could introduce errors in the classification of diploid individuals. A total of 86 diploid males (2AA), 50

diploid females, 228 triploids and seven tetraploids were included. It was expected that accurate results would be achieved, as the majority of fish were from the Guadiana basin. However, it is important to remember that this may not be biologically significant or accurate due to the difference in evolutionary histories between fish of different basins (Cunha et al., 2004; see Chapter 2.1). This was demonstrated by the principal components plot (Figure 3.17), where the variation within-groups appeared to be larger when compared to the PCA plot for the Guadiana river basin, especially for triploid individuals which are present in all three rivers sampled. The other groups present had a similar spread of shapes within-groups as found when the basins were analysed separately.

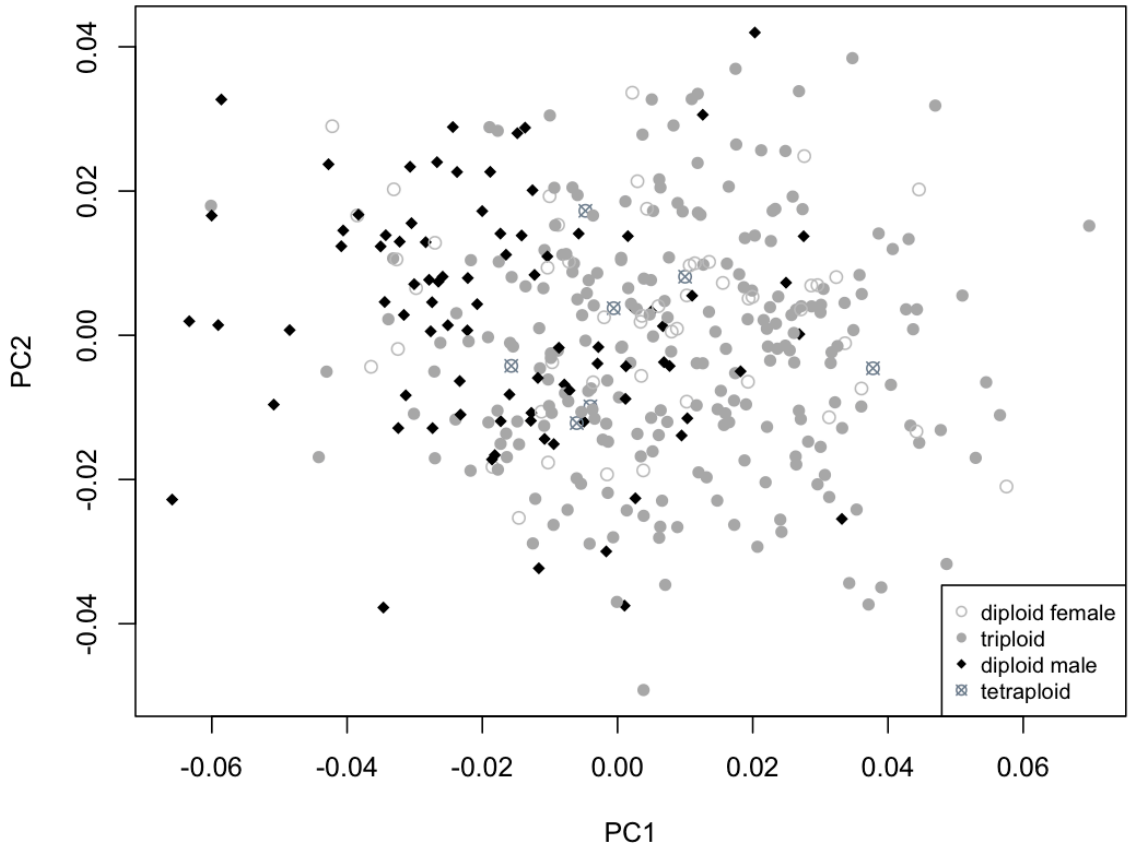


Figure 3.17: PCA plot of individuals from the Guadiana, Tejo and Sado river basins, grouped by ploidy, showing the within-group shape variation.

The differences in average shape of each ploidy between river basins may help explain classification errors but may also provide answers as to whether differences in shape and evolutionary histories exist between river basins and if it would be feasible to discriminate and classify fish irrespective of the river basin they originated from. For diploid female individuals the most visible difference in average shape occurred in the Tejo river basin, which was wider relative to the focus of the scale (landmark 2) than both the shapes from the Guadiana and the Sado (Figure 3.18).

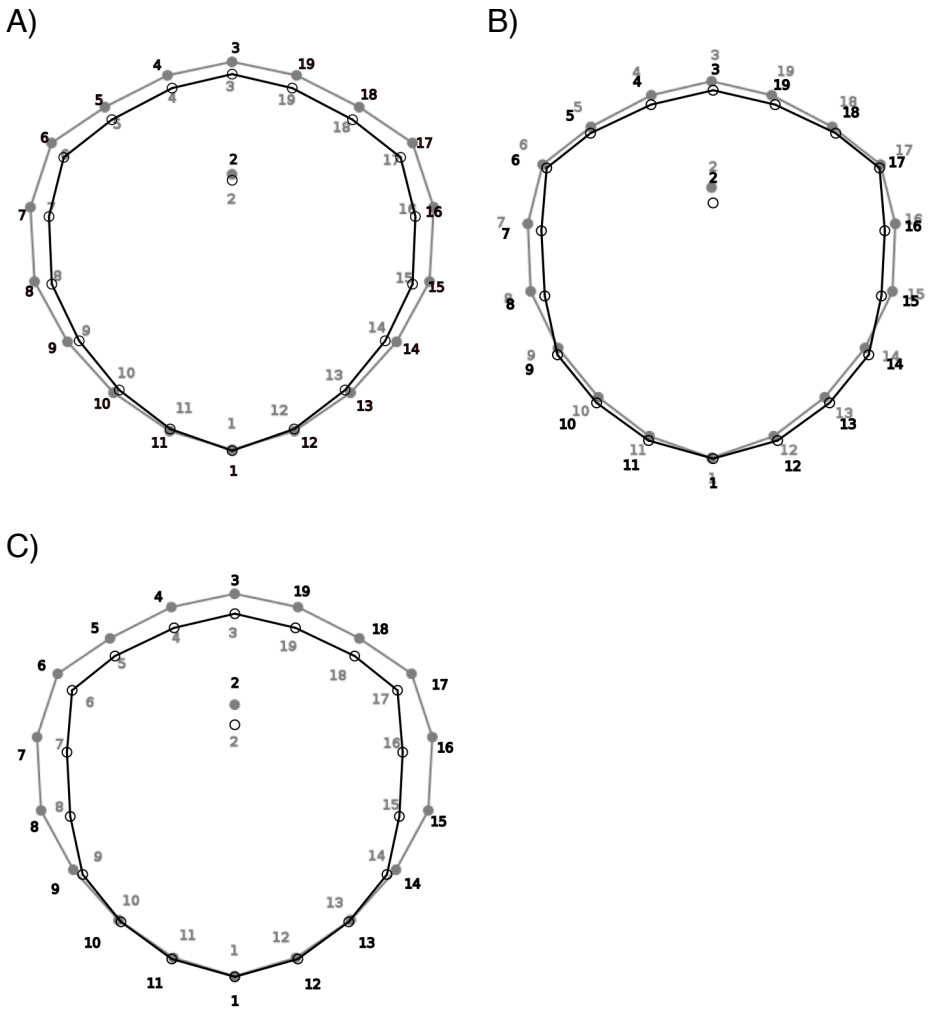


Figure 3.18: Differences in average scale shape of diploid female individuals from the Guadiana, Tejo and Sado river basins. A) Guadiana is in black, Tejo in grey B) Guadiana is in grey, Sado in black C) Tejo is in grey, Sado in black.

In triploid individuals, differences between basins were less visible, however variation exist. Tejo triploids were longer at the posterior end of the scale than Guadiana and Sado triploids. Sado triploids were wider at the anterior edges than both individuals from the Guadiana and Tejo rivers (Figure 3.19).

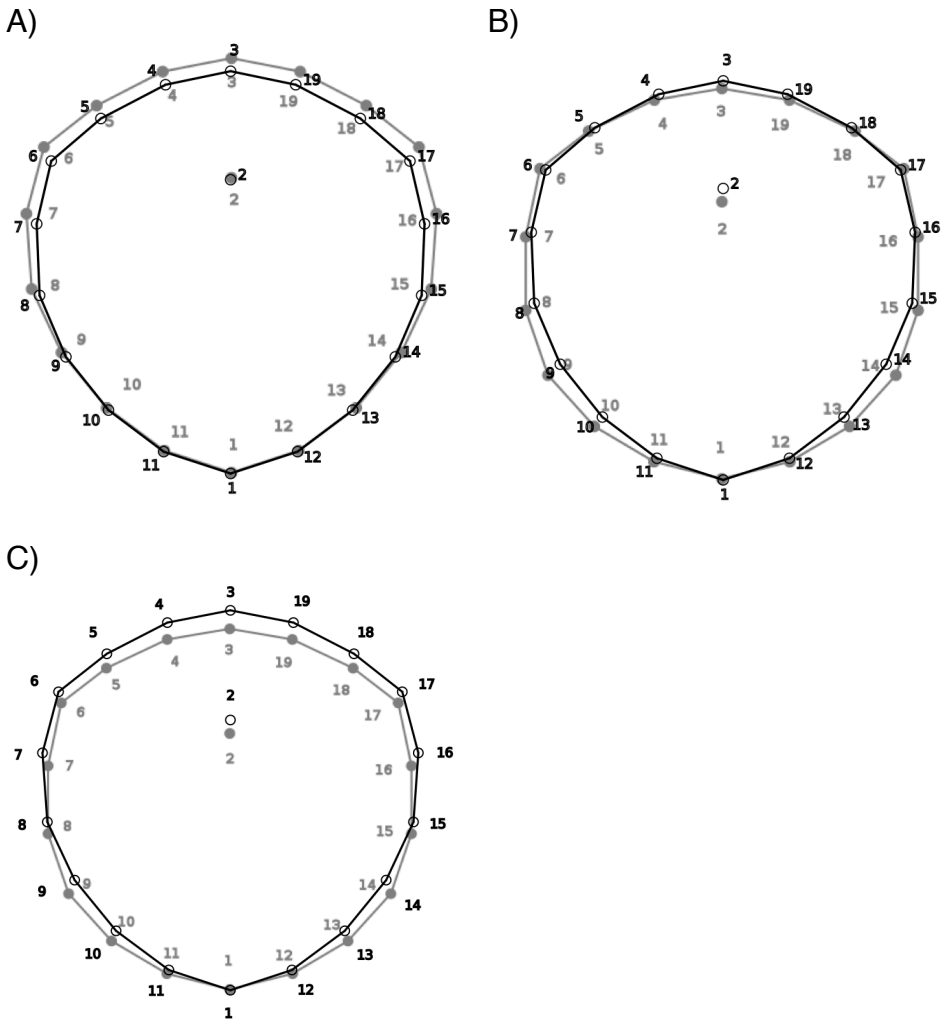


Figure 3.19: Differences in average scale shape of triploid individuals from the Guadiana, Tejo and Sado river basins. A) Guadiana is in black, Tejo in grey B) Guadiana is in black, Sado in grey C) Tejo is in black, Sado in grey.

For tetraploid individuals, the only possible comparison was between the average shape of individuals from the Tejo and Sado basins, as tetraploids were not present in the Guadiana river. Scales from the Tejo are longer than those originating from the Sado basin, visible around the posterior edge of the scale (Figure 3.20).

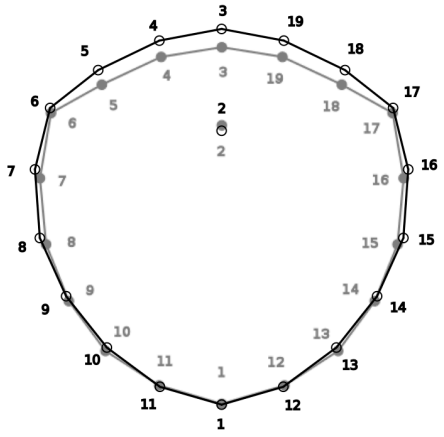


Figure 3.20: Difference in average scale shape of tetraploid individuals from the Tejo and Sado river basins. Tejo is represented by black lines, Sado by grey lines.

The cross-validation was extremely successful at classifying triploid individuals (93.9%) and moderately successful at classifying male diploids (69.8%). However, for both female diploids and tetraploids, the function was poor at discrimination (4% and 0% respectively). This may be due in part to the small number of tetraploids present in the sample (Table 3.6).

Table 3.6: Cross-validation of diploid male (2AA), diploid female (2n), triploid (3n) and tetraploid (4n) individuals from the Guadiana, Tejo and Sado river basins combined.

	2AA	2n	3n	4n
2AA (N = 86)	60	0	26	0
2n (N = 50)	11	2	37	0
3n (N = 228)	10	3	214	1
4n (N = 7)	0	0	7	0

As with the Sado dataset, it was deemed necessary to remove the small tetraploid sample in an attempt to improve results. The analysis comprised a total of 364 fish of known ploidy. Discrimination between groups was effective, with the clearest group being the triploid individuals. There was some overlap of female diploids both with male diploids and triploids, however it appeared to overlap most with male diploids

(Figure 3.21). The cross-validation agreed in part with this, accurately classifying both male diploids and triploids (69.8% and 94.3% respectively). Once again, only 4% of female diploids were correctly classified (Table 3.7).

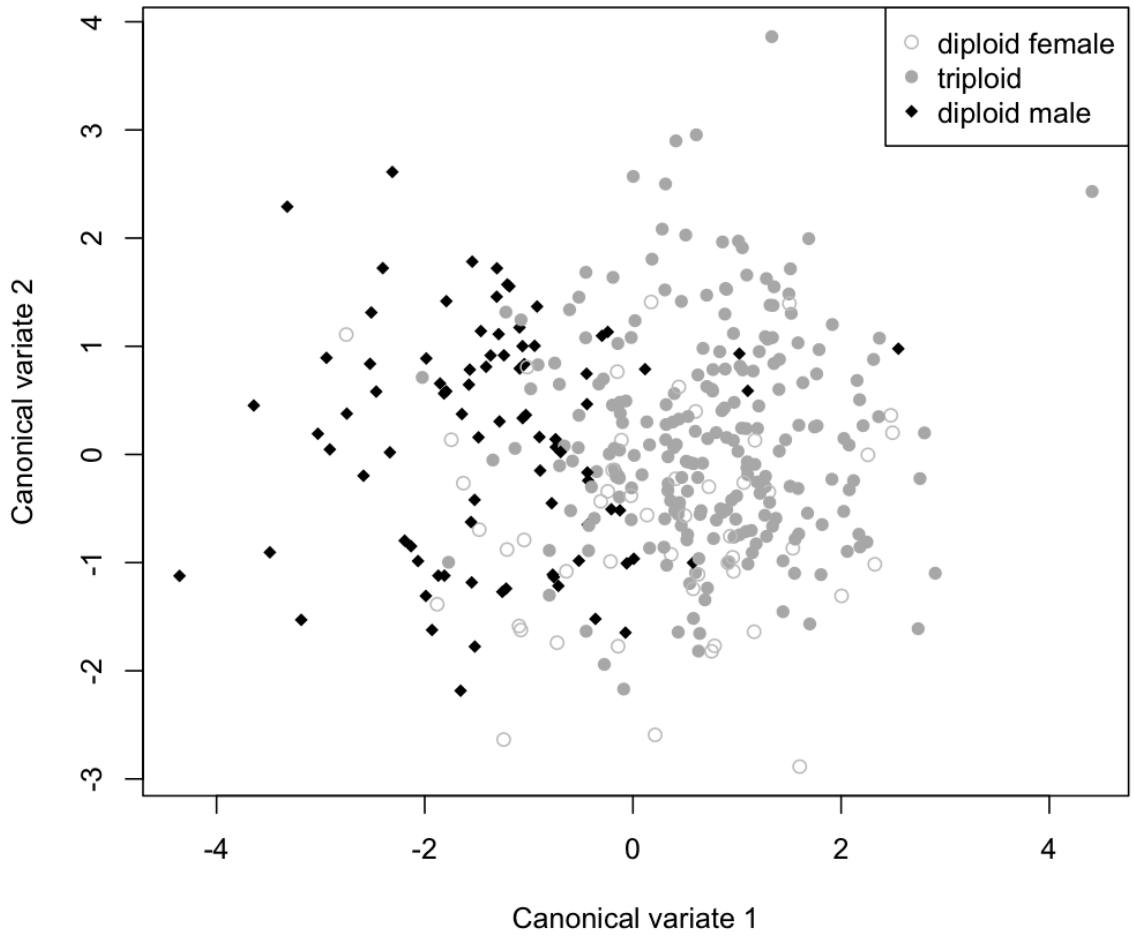


Figure 3.21: CVA plot of male diploid, female diploid and triploid individuals from the Guadiana, Tejo and Sado river basins combined, showing between-group shape variation.

Table 3.7: Cross-validation of diploid male (2AA), diploid female (2n) and triploid (3n) individuals from the Guadiana, Tejo and Sado river basins combined.

	2AA	2n	3n
2AA (N = 86)	60	0	26
2n (N = 50)	11	2	37
3n (N = 228)	10	3	215

3.4 Discussion

3.4.1 Guadiana

The results obtained from the Guadiana river basin showed that differences in scale shape between ploidy groups exist and are identifiable using a range of geometric morphometric techniques. It was previously expected that differences could be identified because of the variation known to be present within the genome of *S. alburnoides* individuals, as well as the visible differences in body shape between diploid male and individuals belonging to other ploidy groups. The analyses performed showed varying success in differentiating between groups, dependent on the choice of grouping and method used. The plot showing shape differences between diploid males, diploids of unknown sex and triploids (Figure 3.7) highlighted the differences that can be found within the diploids. The overlap of diploids of unknown sex with AA males indicated that some male diploids must have been misclassified as PA diploids when the initial grouping was performed. As such, for diploid individuals, sex plays a crucial role in correctly identifying them. Although successful, the first discriminant analysis has little biological meaning as all diploid fish are grouped together and discriminated as such. It is for this reason that the most successful function contained only diploids of known sex (discriminant analysis 3), thus removing any of the bias previously encountered. Despite the reclassification errors in this analysis, it was deemed sufficiently accurate for the prediction of fish of unknown ploidy. Overall, 249 individuals from 331 were correctly classified using the discriminant function, which equals to 75.2% of individuals reassigned to their respective groups. This figure is within similar range of other studies that have used discriminant function analyses to reclassify individuals of known grouping as well as to predict the grouping of unknown

individuals (Helland et al., 2009; Ibañez et al., 2007). These studies used fish from the Mugilidae and Salmonidae families, distinguishing between two species from each family. The majority of diploid females that were misclassified were classified as triploid individuals. This supports the plots that showed marginal difference in scale shape between diploid females (genomic composition PA) and triploids. Indeed, this was be further supported because it has not been possible to tell these two groups apart beforehand without detailed molecular genotyping (Gomes-Ferreira et al., 2005). The genome of diploid females (PA) and triploids (PAA) is similar, as both the genome of maternal and paternal ancestors are present, contrary to diploid male (AA) individuals. This similarity is reflected in scale shape, which means that accurate discriminations may never be possible. Discriminations between these two groups may be possible if further information is taken into account, such as size or additional morphological information relating to body shape. Age may also be a factor affecting scale shape as some studies have show that discrimination between different populations becomes more difficult with individuals of older ages (Doering-Arjes et al., 2008). This aspect was not taken into account in the study performed here as the age differences between individuals of *S. alburnoides* varied only by one or two years, therefore it was not considered a greatly influencing factor. A number of studies with low discrimination accuracy between two groups have nevertheless confidently shown distinctions by using more than one factor when discriminating between groups, such as further shape information or habitat preferences (Aguirre, 2009; Friess and Baylac, 2003).

As the methodologies used do not accurately separate all individuals, ecological analyses in later chapters combine the information for both groups into one, which is

represented by triploids. If ecological differences do exist between PA and PAA individuals, a distinction within the combined group will be visible and this may prove a useful tool in further correctly identifying diploid females.

3.4.2 Guadiana prediction

The prediction function was unable to provide support for the correct classification of the four individuals assumed to be diploid females. In hindsight, this was to be expected as the canonical variates plot showed a large amount of overlap between the shape of diploid females and triploids, which was also confirmed by the cross-validation of the LDA function. Several explanations can be put forward as to the lack of diploid females present in the samples collected. As mentioned, discrimination between diploid and triploid individuals is not accurate, therefore any female diploid individuals present may have been incorrectly classified as triploids. If this is the case, it will be important to bear this in mind when performing further analyses that attempt to show ecological differences between groups. Another explanation may be simply that in the samples collected few diploid females were actually present, making it impossible to identify them.

Ribeiro et al. (2003) showed that diploid females are less abundant throughout most of the year, except during the spawning season around April and May, depending on the year sampled. This implies that some migration by diploid females occurs, with evidence to suggest that they prefer deep water and coarse substratum material, except during the reproductive period when shallow waters are preferred (Martins et al., 1998). This could not be confirmed in the present study as it was not possible to distinguish female diploids from triploid individuals accurately enough. However, the results presented here do indicate that there was a larger proportion of triploids, which

may include some diploid females, during the April sampling, when compared to the October sampling. The study site sampled contained a number of different microhabitats, some of which may not have been suitable for diploid females. The microhabitat sites is discussed in detail in Chapter 5, which may further explain the lack of diploid female samples that were predicted.

It is feasible that prediction could also be skewed in favour of a particular group due to the percentages of each ploidy present in the initial database used to create the prediction function. However, this was not the case, as the initial dataset had similar proportions of abundance as samples collected in previous years from the same river basin (Figure 3.22), as mentioned in Martins et al. (1998) and Ribeiro et al. (2003).

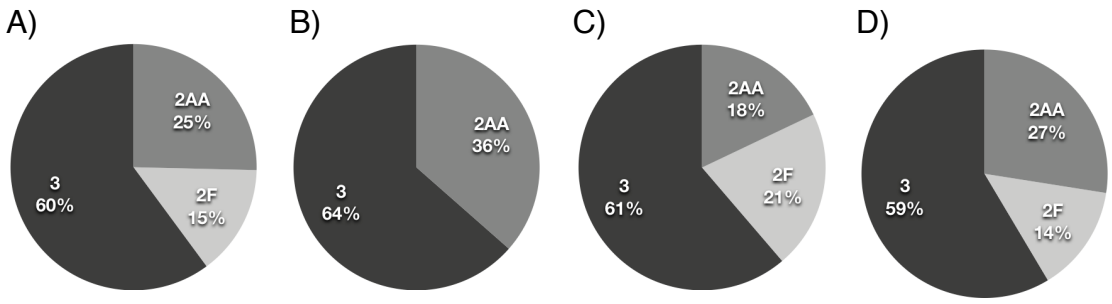


Figure 3.22: A) Proportions of ploidy levels used for creation of prediction LDA. B) Proportions of ploidy levels predicted by LDA (prediction probabilities over 60% only). C) Proportions of ploidy levels as in Ribeiro et al. (2003). D) Proportions of ploidy levels as in Martins et al. (1998).

Despite the lack of diploid females, total proportions remained consistent, with triploids always more abundant than diploid males, as has been previously recorded in all studies in the Guadiana basin (Alves and Coelho, 2001). The literature suggested great variability in the proportions present at specific months between rivers and years sampled. It is thus difficult to compare predictions for fish sampled in April and October. In the samples collected, triploids appeared to be more abundant in April than October, which once again could be an indication of misclassification of diploid

females, of which more would be expected during the reproductive season in April. In the year 2009, the reproductive season was starting when collecting occurred in April, as evident from the presence of eggs in all female fish collected.

In summary, the prediction of fish sampled from the River Murtega was sufficiently accurate for further studies assuming some conditions, such as prediction probability cut-offs, have been met. Any misclassifications will have grouped diploid females as triploids, which is consistent with documented evidence that there is little morphological and ecological difference between the two groups. Increasing the number of diploid females in the original dataset may improve the function, however it is sufficiently robust to be used for any further studies involving ploidy determination of *S. alburnoides*. The advantage of this methodology is principally that fish do not need to be killed to be identified. Furthermore, the financial cost of such discrimination is greatly reduced when compared with genetic methodologies, and once the user is familiar with the process, landmarking and analysing scales is relatively fast, especially with small sample sizes. If genetic analyses are performed on fish collected, it is recommended the information is added to the existing database, further improving reliability.

3.4.3 Tejo

Few conclusions can be reliably drawn from the Tejo dataset due to its small size. The results showed it may be possible to discriminate tetraploids from both diploids and triploids, if the sample size was larger. The current sample is potentially statistically inaccurate as it has been shown that CVA and LDA are unreliable in certain cases. This generally happens where there are fewer cases (specimens) than there are variables (landmarks) (Mitteroecker and Bookstein, 2011).

As in the Guadiana river basin, there appears to be a large amount of overlap between triploids and PA diploids, which is most likely caused by the similarities in their genetic structure, having at least one copy of both the maternal and paternal ancestor's genome.

The triploid individuals collected in the Tejo river were all PAA individuals, thus it was expected that there would be little morphological variation within this group. It would have been of interest to obtain triploid individuals that had another genome combination, so as to determine whether variation is also present within a ploidy level.

One of the other problems faced when attempting to draw conclusions from this dataset was the lack of diploid males present in the sample. Diploid males have repeatedly been found in the tributaries of the Tejo, and account for up to 15% of the total number of fish present in a typical *S. alburnoides* population (Sousa-Santos, Collares-Pereira, and Almada, 2006a; Crespo-Lopez et al., 2007; Alves and Coelho, 2001). Even if the sample size was larger and was able to discriminate between the three forms present, it would not be able to be possible to use for any predictions as it would potentially misclassify 15% of the sample.

3.4.4 Sado

As with the Tejo dataset it was not possible to draw many conclusions from the Sado drainage, due to the sample being almost solely composed of triploid individuals. This dataset was however included as a way of demonstrating the differences in population structure and potential variation in shape between all river basins. In the event of this technique being used in the Sado basin, it would be first necessary to gather a much larger sample, which also includes diploid males that are normally present in the Guadiana basin (Sousa-Santos, Collares-Pereira, and Almada, 2006a).

3.4.5 All datasets

The results found that when the Tejo, Guadiana and Sado river basins were combined, results were similar to those from the Guadiana river, where diploid males are easily identifiable by both the plot and the LDA function as well as significant overlap between the remaining three groups. By examining the average scale shapes of ploidy levels for each basin, it was possible to determine whether it would be accurate to group the Southern rivers of Portugal when performing morphological identification. The superimposition plots of ploidy levels from the Guadiana and Tejo basins showed that for comparable levels, the average shapes were similar but notable differences did exist. For all ploidy levels, scales from the Tejo river basin were the most different to Guadiana and Sado scales, generally showing more elongation of the scale towards the posterior edge. Differences between Guadiana and Sado scales were not as pronounced but were still clearly visible. The findings here support the idea of previous work that hypothesised independent hybridisation events for different geographical areas of the Iberian Peninsula. Cunha et al. (Cunha et al., 2004) suggested five independent origins, one being present in each of the basins studied here. Contrary to this, Alves et al. (Alves et al., 1997) had previously suggested that at least two separate events had occurred, one in the Sado and the other in the Tejo drainage, which then dispersed into other basins, including those in the north of the Peninsula. The data collected here did not support this hypothesis, due to the great differences found between fish from the Tejo and the other drainages. Differing evolutionary origins help to confirm the hypothesis that fish of different basins should be analysed separately with regards to scale shape.

Another factor that must be considered, is the diversity present within the drainage as well as the inclusion of the genome of potential sperm donors that to not always belong to the same species (Pala and Coelho, 2005). A number of studies have shown that genotypic diversity varies greatly from basin to basin (Pala and Coelho, 2005; Sousa-Santos, Collares-Pereira, and Almada, 2006a; Crespo-Lopez et al., 2007), increasing the chance that phenotypic diversity will also be present. A study of rostrum dace, *Leuciscus leuciscus burdigalensis*, on an isolated section of river in France, found that scale shape could differ within a river basin, based on genetic factors (Poulet et al., 2005). To add to this, it has been shown that environmental conditions also play a role in scale characteristics (Ihssen et al., 1981), further enhancing differences present between river basins.

As such, the identification of *S. alburnoides* ploidy levels must be done among individuals originating from the same river basin, due to genetic and environmental influences. Not only has this chapter shown that it is feasible to differentiate ploidy levels by looking at scale shape, it also provides the possibility of identifying regional differences that are present within the *S. alburnoides* species complex.

Chapter 4: Body shape discrimination using geometric morphometrics

4.1 Introduction

It has been previously shown, both in this thesis and in the literature, that the identification of *S. alburnoides* ploidy levels can be achieved using genetic methodologies and using scale morphology. This chapter examines whether discrimination between groups can be achieved using other aspects of morphology, specifically by observing the shape of the head and the whole body of the fish. If successful, this will provide another method that can be used to identify individuals, without looking at the genome, which can be a costly process. This method will also be a cheap and non-destructive method, by taking photos of individuals in the field and returning them to their original habitat. The photos taken can then be used to identify ploidy level in the field or in the laboratory. An attempt will also be made to identify seasonal differences in morphology, more specifically relating to the reproductive season, where variations within ploidy groups may occur.

Only one study has previously described the morphology of *S. alburnoides* using geometric morphometric techniques, where an attempt was made to differentiate between ploidy groups and the sperm donor, *S. carolitertii*, using the overall body shape (Cunha et al., 2010). Despite successful discrimination, the results are not applicable to the southern river basins of Portugal due to the varying population composition as well as the presence of the maternal ancestor *S. pyrenaicus* as the principal sperm donor.

The study on *S. alburnoides* joins a number of studies that have successfully used geometric morphometrics to identify differences in body shape within species. Several have been carried out on *Poecilia vivipara*, a live-bearing guppy, investigating whether shape differences are influenced by environmental factors (Neves and Monteiro, 2003; Gomes and Monteiro, 2008) or using the differences found to elucidate possible causes for evolution among the species (Monteiro and Gomes, 2005). It is also possible to identify subtle differences between species occupying the same habitat using body morphology (Helland et al., 2009; Cavalcanti et al., 1999). Head shape alone has been less used in geometric morphometrics on fish, however it has been used to identify differences within the species *Coregonus lavaretus*, commonly known as powan, in response to ecological shifts and feeding regimes (Etheridge et al., 2010). In fact, one study has managed to accurately discriminate between two closely related species of Acipenseridae sturgeon and their hybrid using head shape alone (Costa et al., 2006).

The literature suggests that geometric morphometric methods are effective at picking up subtle differences in shape that are often difficult to distinguish in the field. This chapter relies extensively on geometric techniques laid out in Chapter 3 to highlight differences in *S. alburnoides* populations, with the aim of facilitating the identification of ploidy groups of this species for future studies relating to conservation efforts.

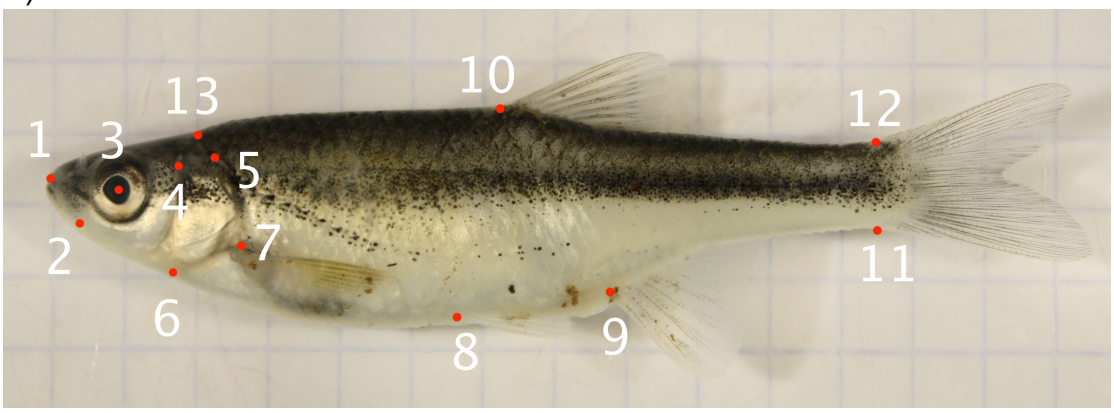
4.2 Methods

Fish used for analyses of shape for this chapter were collected from Ribeira da Murtega in 2009 of which 74 were collected in April and a further 89 were collected in October. Details regarding sampling can be found in the General Methods section of this thesis. Where the ploidy of the fish was known, it was predicted from scale shape analysis as described in Chapter 3. Any fish for which ploidy was not predicted with over 60%

accuracy previously was categorised as being of unknown ploidy. Images used to collect morphometric data were taken by placing all fish collected on graduated paper with their left side facing upwards and photographed from a fixed distance using a Canon EOS 1000D digital SLR camera. The distance used was different for the April and October samples due to equipment availability, however the recommended distance should allow the photographs to be taken of the entire fish, with minimal white space around the fish. Photographs were then imported into TPSDig2 (Rohlf, 2010a) and calibrated for size using the standard length for each fish, measured in the field, to avoid errors due to shrinking during preservation.

Two geometric morphometric analyses were performed in this chapter, using body shape and head shape to discriminate between ploidy groups. For analyses involving body shape, a total of thirteen landmarks were placed in repeatable points along the body of the fish (Figure 4.1 - A). The analyses involving head shape used the same images but focus only on landmarks around the head. In total, eight landmarks were used to define the shape of the head (Figure 4.2 - B). For both analyses the landmark coordinates were exported to MorphoJ (Klingenberg, 2011) and a Procrustes fit performed to find the mean shape of fish sampled and how deviations from this shape occurred (Klingenberg et al., 2002). A pooled PCA analysis, with ploidy level as the pooled variable, was performed and the resulting principal components exported to R (R Development Core Team, 2011) for further statistical analyses. Analyses performed were CVA, to determine differences in shape between the groups analysed, and LDA which created a classification function. The classification function was then used to predict the ploidy of individuals whose ploidy could not be determined using scale shape.

A)



B)

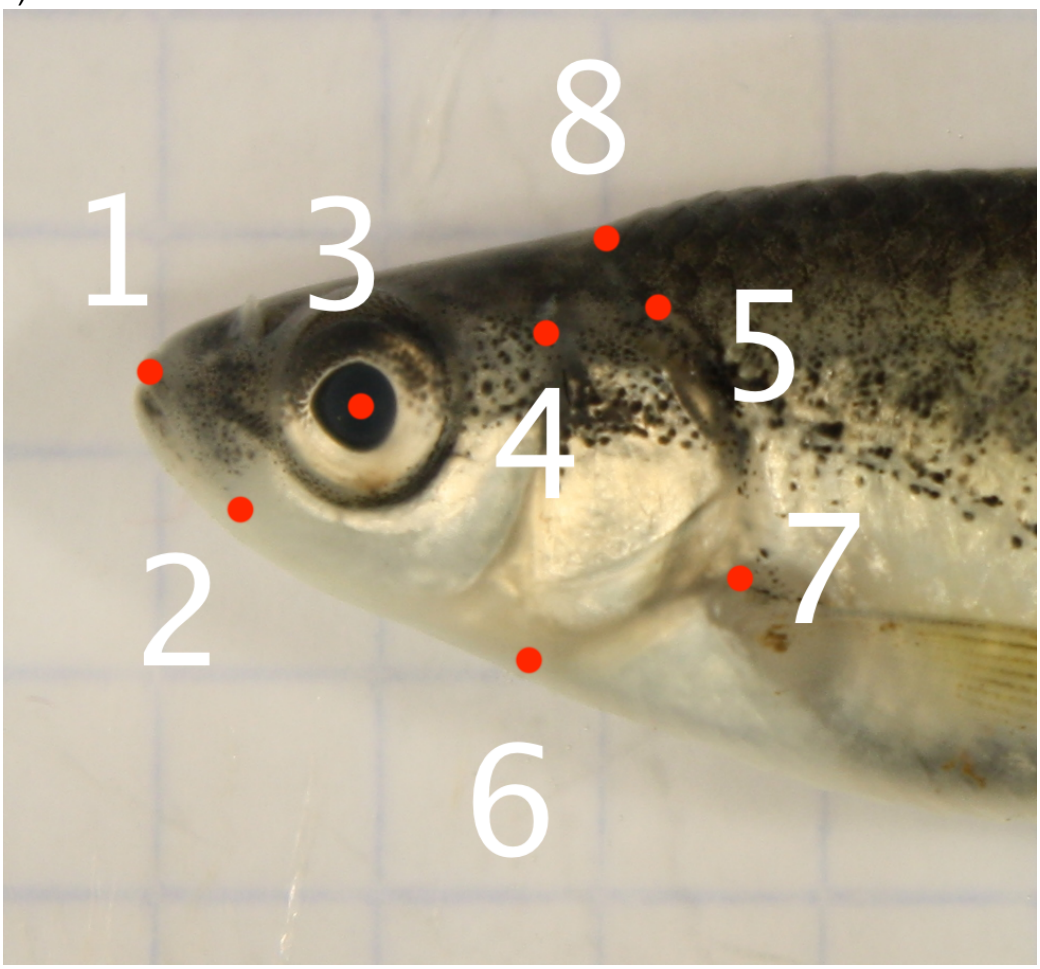


Figure 4.1: A) Position of landmarks on individuals of *S. alburnoides* to describe body morphology. B) Position of landmarks to describe head morphology of *S. alburnoides* individuals. Size of landmarks have been increased on both images for clarity.

4.3 Results

A total of 163 fish were collected for shape analysis, during two different time periods, April and October. The intention of this chapter was to analyse both seasons together, and include distinctions within ploidy groups at different time periods, as well as distinctions between ploidy levels. Unfortunately, due to equipment availability, photographs of fish from different seasons were not taken under the same conditions. This discrepancy in sampling introduced supplementary variation in shape when the two seasons were analysed together. It was therefore decided that the two seasons would be analysed separately, so as to not introduce unnecessary bias in the discrimination between ploidy groups and seasons. As a result, the dataset for April consisted of 74 individuals, of which 19 were classed as diploid male, 37 as triploid and the remaining 18 as being of unknown ploidy. The October dataset was made up of 30 diploid males, 37 triploid individuals and 22 of unknown ploidy.

4.3.1 Body shape

4.3.1.1 April

Once landmarked and rotated for GLS Procrustes superimposition, a pooled within-group PCA was performed on fish collected in April, to determine variation present within both ploidy groups present (Figure 4.2). Both groups show similar amounts of variation within each group, with triploid individuals generally having higher scores along the second principal component.

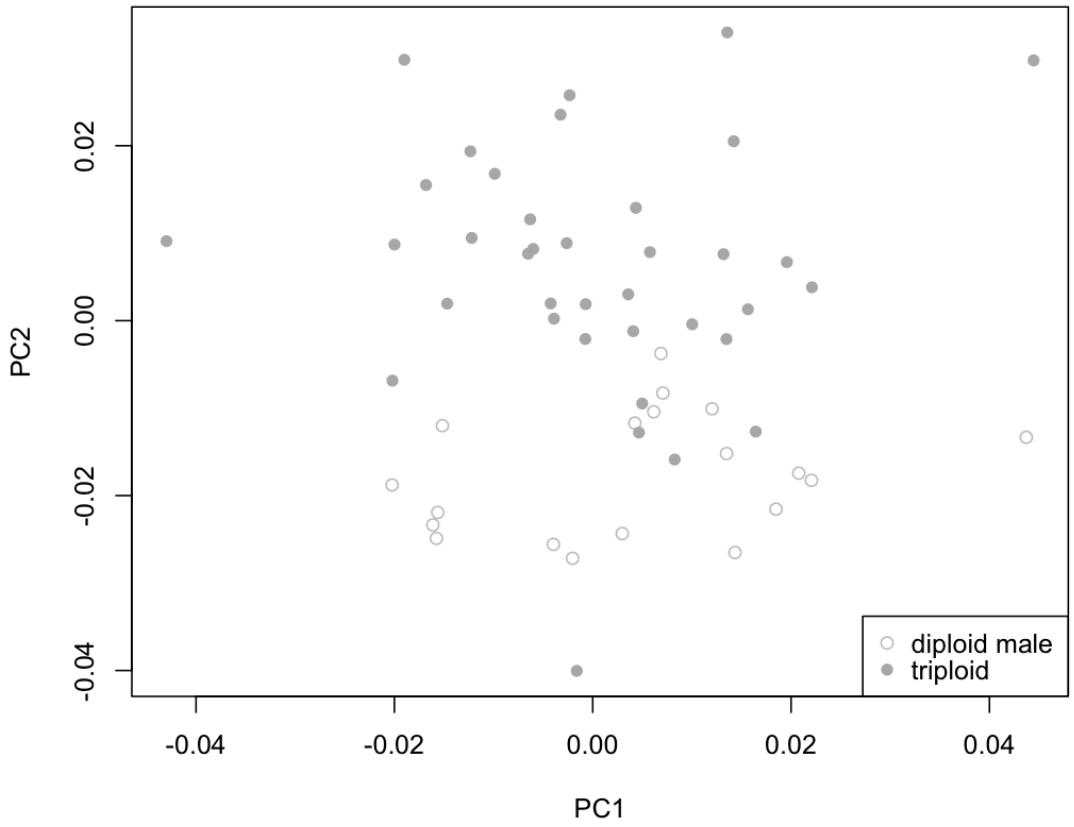


Figure 4.2: PCA plot of individuals of known ploidy from Ribeira da Murtega collected in April, grouped by ploidy, showing within-group shape variation. Triploids appear to have higher PC2 scores in general.

Using the first 13 principal components, which account for 95.8% of the total variance, a linear discriminant function and leave-one-out cross-validation were performed, to determine whether diploid male and triploid individuals could be separated by shape alone. Both the CVA plot (Figure 4.3) and cross-validation (Table 4.1) show excellent distinction between the two groups, with 100% and 97.3% of diploid male and triploids respectively being correctly identified.

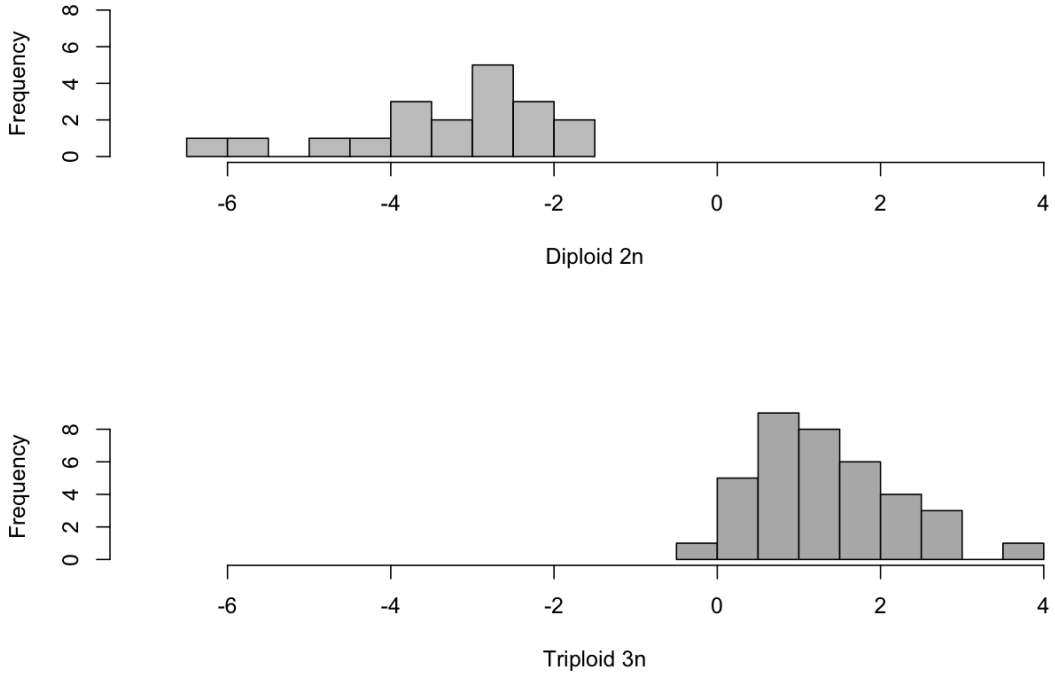


Figure 4.3: CVA plot of fish of known ploidy from Ribeira da Murtega collected in April, showing variation in shape present between diploid male and triploid individuals.

Table 4.1: Cross-validation of diploid male (2AA) and triploid (3n) individuals collected in April from Ribeira da Murtega.

	2AA	3n
2AA (N = 19)	19	0
3n (N = 37)	1	36

A transformation grid (Figure 4.4), displaying shape changes between the two groups, show that triploids had less slender bodies than their diploid male counterparts, who despite having smaller average lengths had a longer tail relative to the rest of the body.

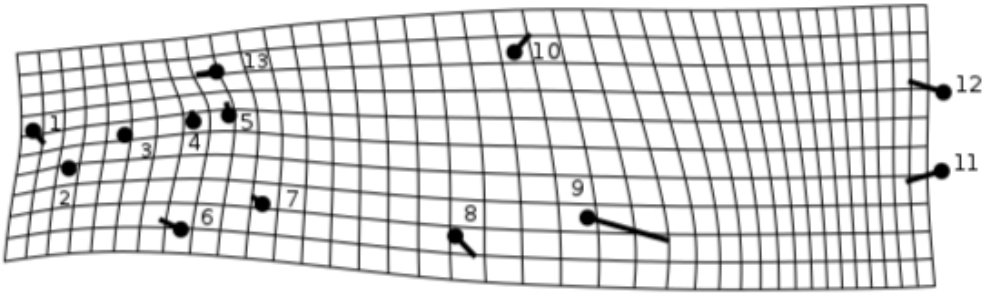


Figure 4.4: Comparison of average shape of diploid males and triploids collected in April, represented by a transformation grid, displaying the shape deformation of triploids, when applied to diploid males. Numbers correspond to landmarks, the lines show the direction of the shape deformation.

Using the linear discriminant function, the remaining 18 individuals of unknown ploidy were assigned a ploidy group based on their body shape. All achieved over 60% confidence of accuracy, with most achieving over 95% (see Appendix 2 for prediction values).

4.3.1.2 October

The same analyses were performed on the 67 fish of known ploidy from October as those collected in April. The pooled-within groups PCA showed slightly more variation present within diploid males than seen in the April sample (Figure 4.5). As a result, both the CVA (Figure 4.6) and the cross-validation (Table 4.2) did not discriminate as clearly between groups as for April. However the results were still encouraging, with 83.3% of diploid males correctly reclassified and 94.6% of triploid individuals assigned to the correct group.

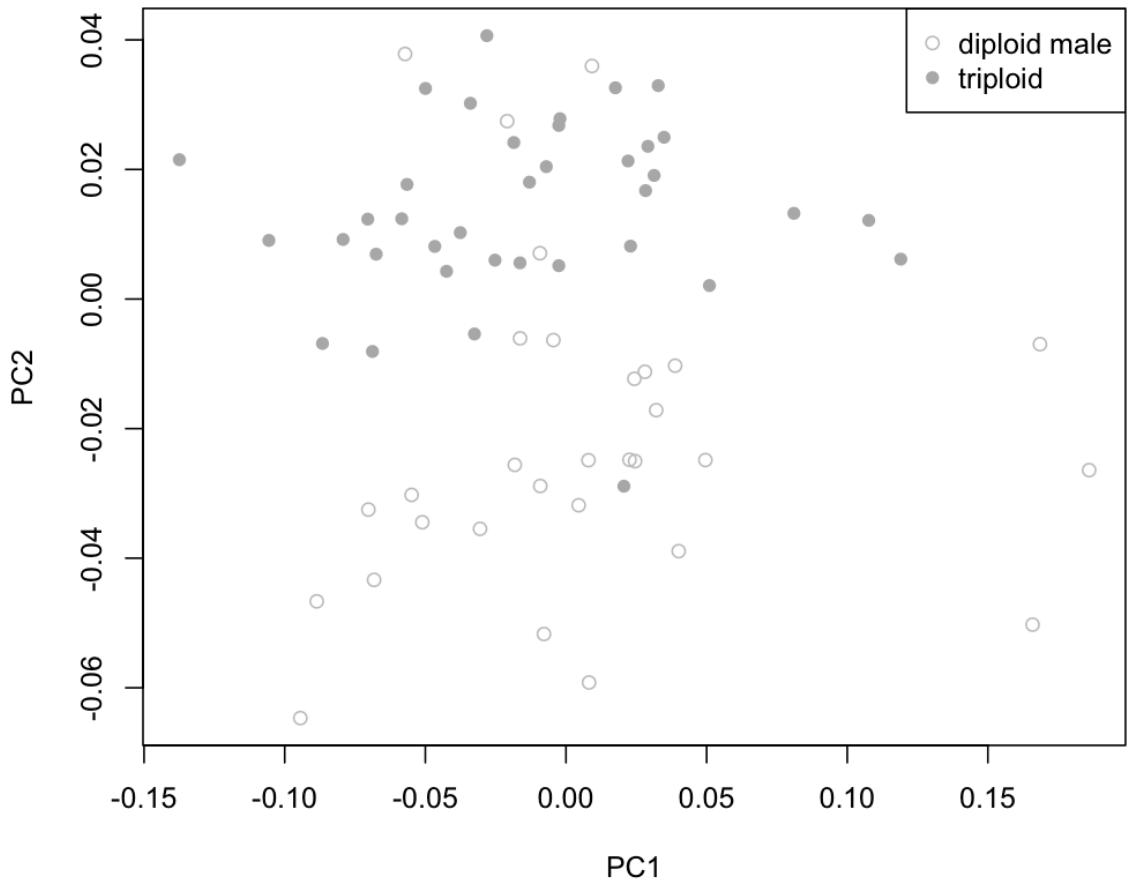


Figure 4.5: PCA plot of individuals of known ploidy from Ribeira da Murtega collected in October, grouped by ploidy, showing the within-group shape variation. As with the April sample, triploids appear to show higher values on the second PC axis.

Table 4.2: Cross-validation of diploid male (2AA) and triploid (3n) individuals collected in October from Ribeira da Murtega.

	2AA	3n
2AA (N = 30)	25	5
3n (N = 37)	2	35

The transformation grid displayed similar differences in shape as for fish collected in April, however these differences were less pronounced (Figure 4.7). The ploidy level of 22 individuals was predicted using the function created, with only one individual not achieving the 60% confidence cut-off.

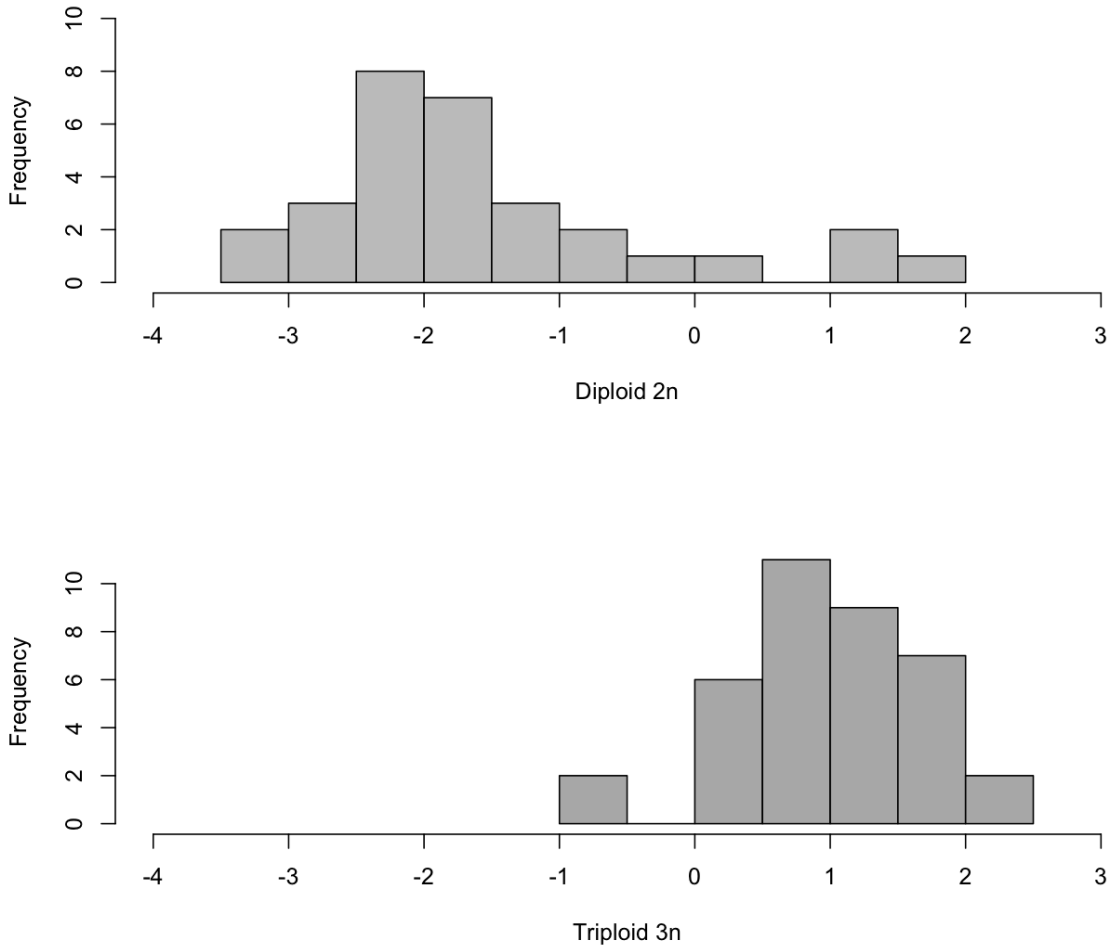


Figure 4.6: CVA plot of fish of known ploidy from Ribeira da Murtega collected in October, showing variation in shape present between diploid male and triploid individuals.

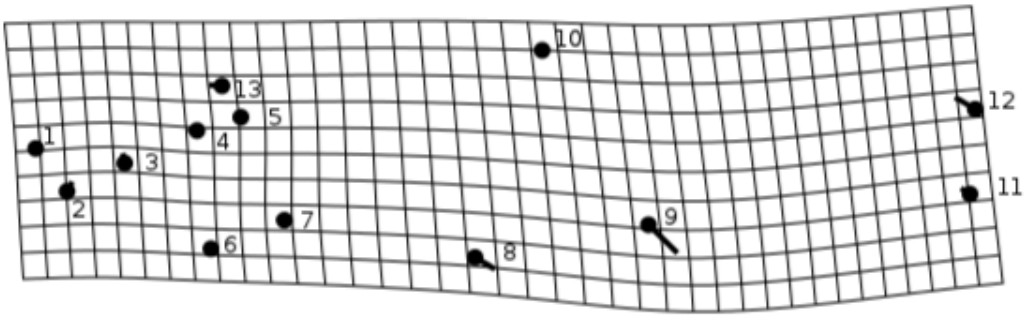


Figure 4.7: Comparison of average shape of diploid males and triploids collected in October, represented by a transformation grid, displaying the shape deformation of triploids, when applied to diploid males.

4.3.2 Head shape

4.3.2.1 April

The 56 fish collected in April displayed a large amount of within-group variation, especially on the first principal component (Figure 4.8), and the first 8 principal components accounted for 95.7% of the variance present within the samples collected.

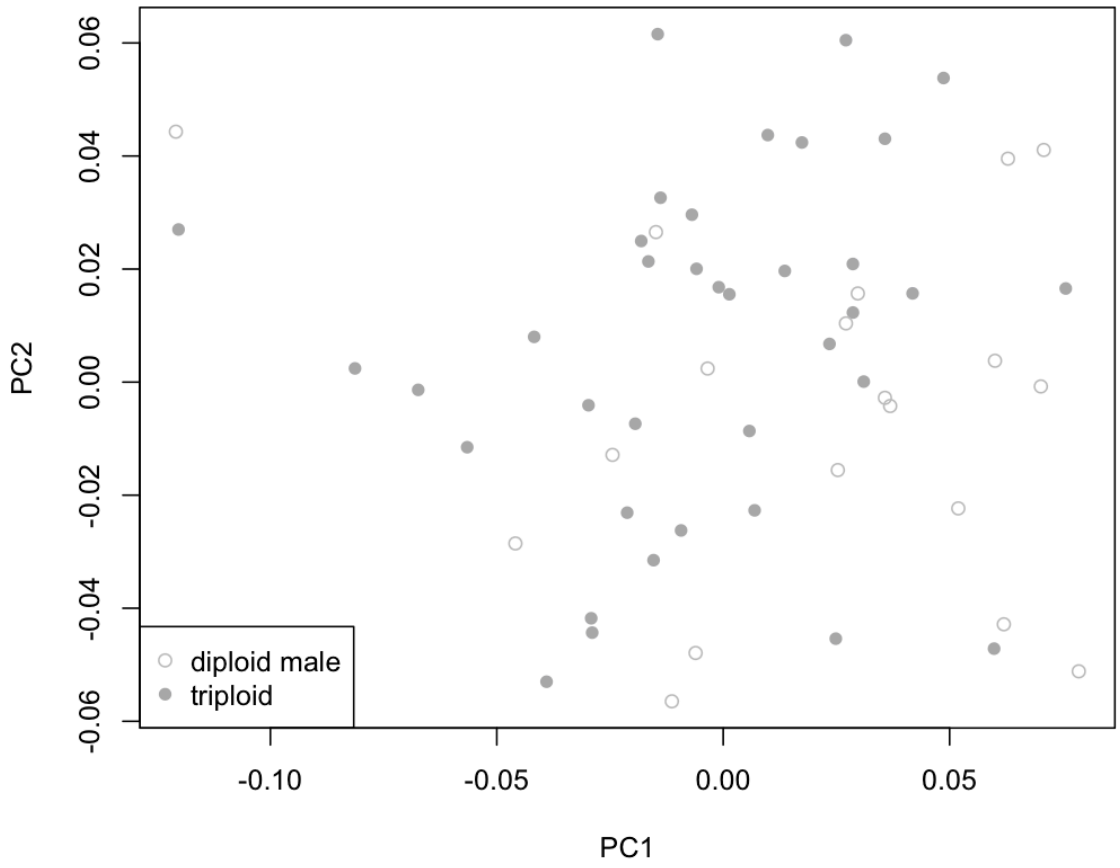


Figure 4.8: PCA plot of individuals of known ploidy from Ribeira da Murtega collected in April, grouped by ploidy, showing the within-group shape variation.

When a CVA was performed, the plot showed that some discrimination was possible, however a substantial amount of overlap was present between the two groups (Figure 4.9). The cross-validation (Table 4.3) mirrored these findings, by effectively identifying triploid individuals (81.1%) but did not perform as well in identifying diploid males (57.7%).

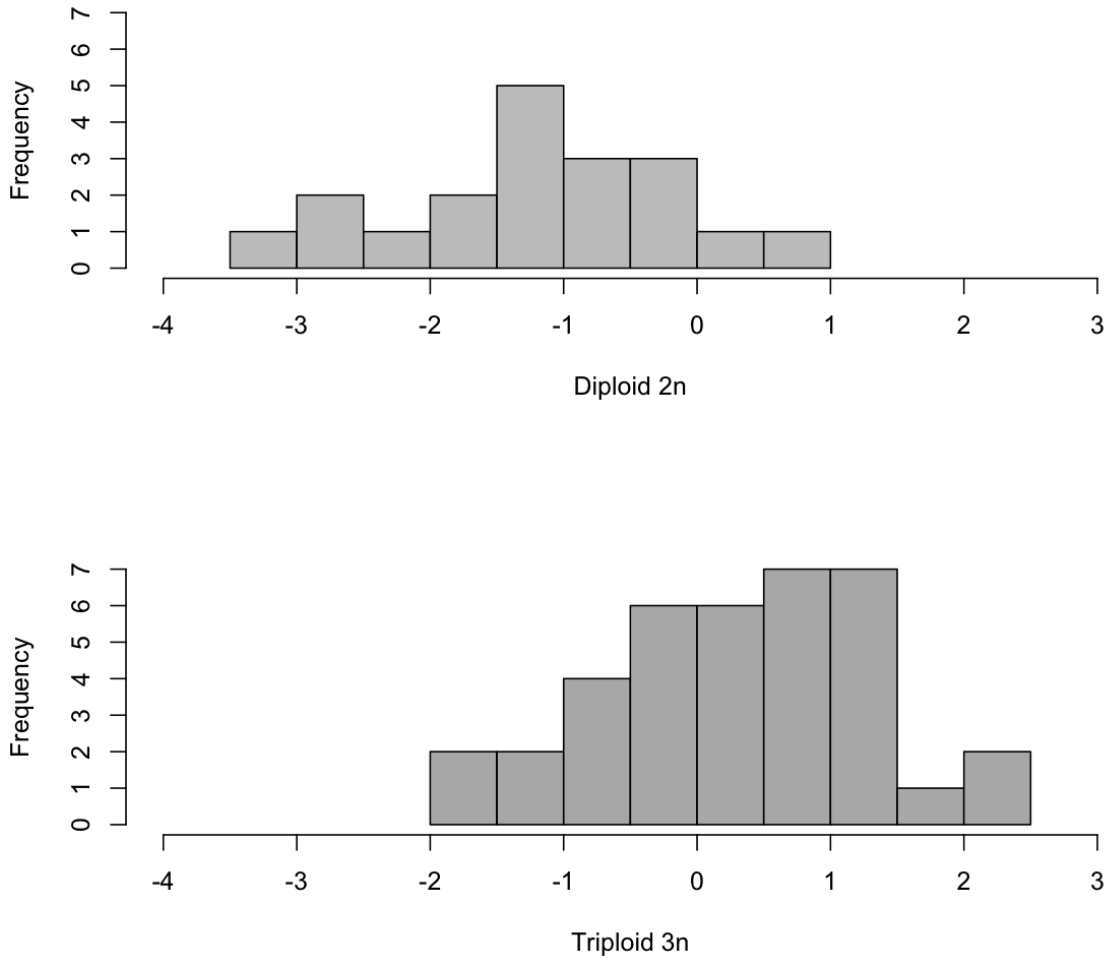


Figure 4.9: CVA plot of fish of known ploidy from Ribeira da Murtega collected in April, showing variation in shape present between diploid male and triploid individuals.

Table 4.3: Cross-validation of diploid male (2AA) and triploid (3n) individuals collected in April from Ribeira da Murtega.

	2AA	3n
2AA (N = 19)	11	8
3n (N = 37)	7	30

The transformation grid, displaying the shape differences between the two groups, showed that triploids had a slightly more elongated head than diploid males, and that the top of the mouth was situated further down relative to the rest of the head on triploids than on diploid males (Figure 4.10).

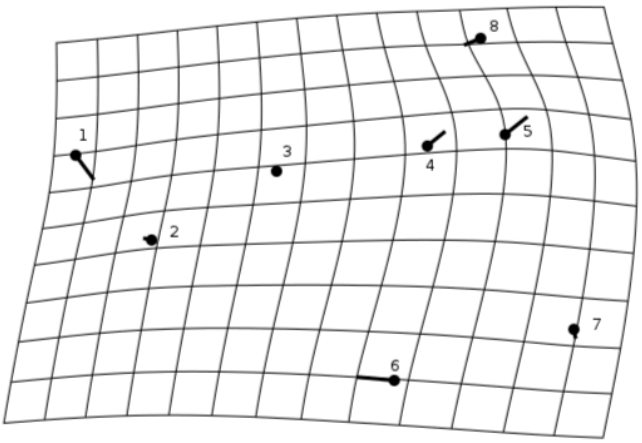


Figure 4.10: Comparison of average shape of diploid males and triploids collected in April, represented by a transformation grid, displaying the shape deformation of triploids, when applied to diploid males.

Despite confidence of discrimination being lower than had previously been achieved using other methods, the percentage of discrimination achieved was deemed sufficient to create a prediction function and predict the ploidy of fish of unknown ploidy levels. This was confirmed as 17 of the 18 individuals analysed achieved over 60% confidence in the probability of prediction.

4.3.2.2 October

Fish sampled in October showed similar within-group variation, irrespective of their ploidy level, and the first 9 principal components accounted for 95.3% of the total variance (Figure 4.11).

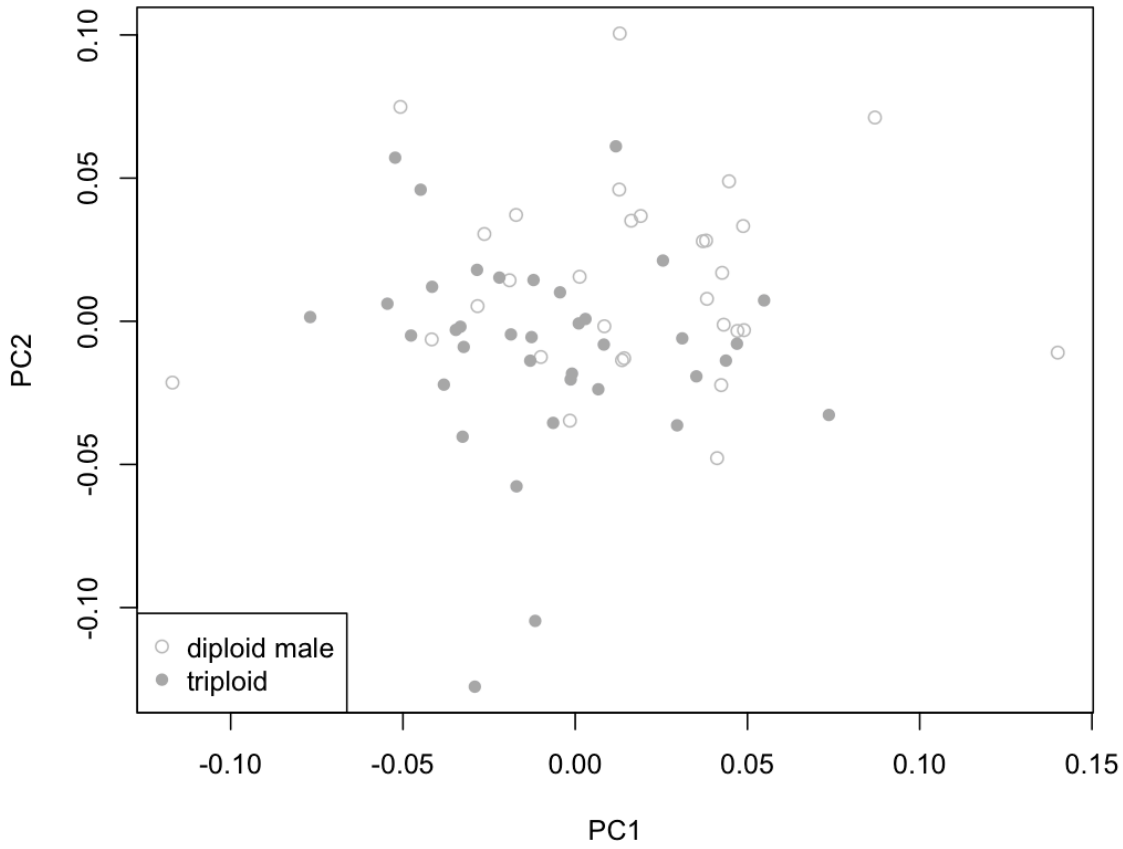


Figure 4.11: PCA plot of individuals of known ploidy from Ribeira da Murtega collected in October, grouped by ploidy, showing the within-group shape variation.

The CVA plot showed that a significant overlap in shape between diploid males and triploids did occur (Figure 4.12 - A), which was also displayed in the transformation grid showing differences in shape (Figure 4.12 - B). This grid showed that triploid individuals had longer heads and less upturned mouths than diploid male individuals, however differences were not as large as those present in fish collected in April.

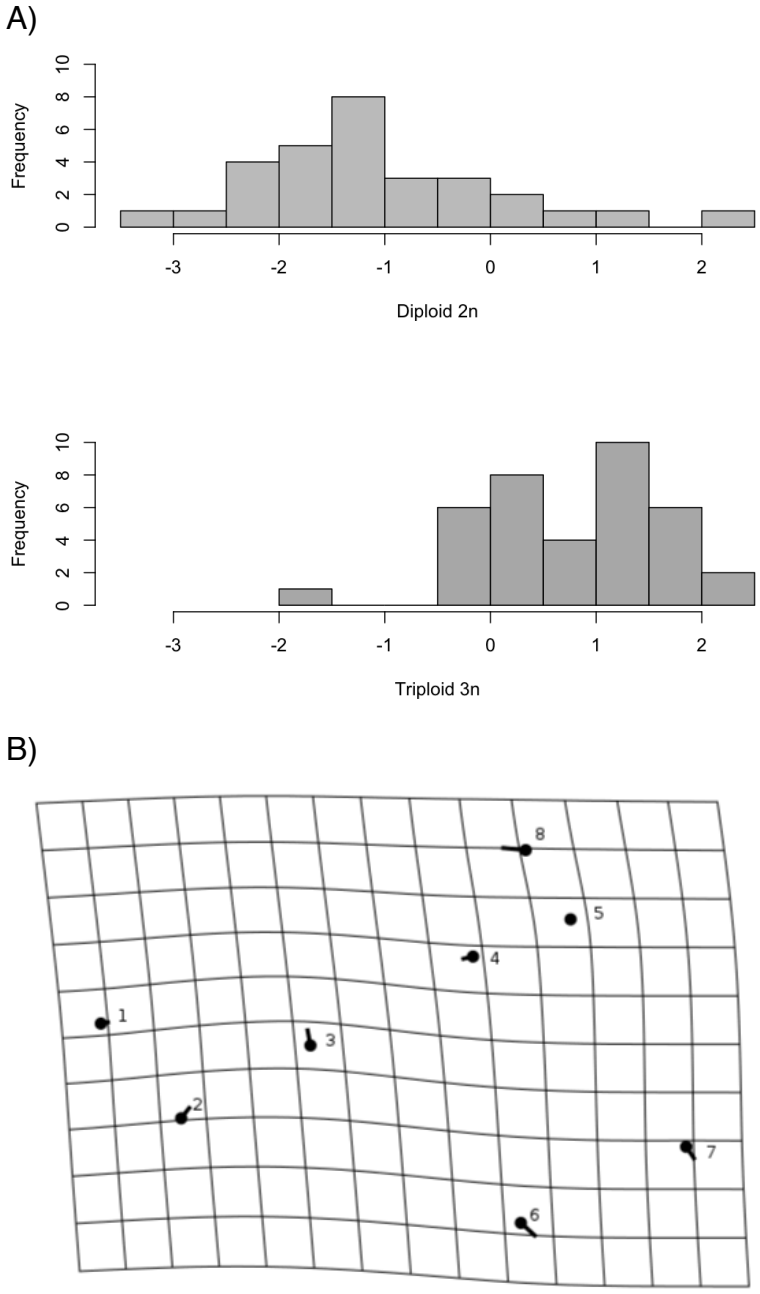


Figure 4.12: A) CVA plot of fish of known ploidy from Ribeira da Murtega collected in October, showing variation in shape present between diploid male and triploid individuals. B) Comparison of average shape of diploid males and triploids collected in October, represented by a transformation grid, displaying the shape deformation of triploids, when applied to diploid males.

Despite the overlap present, the linear discriminant analysis did manage to effectively discriminate between the two groups, accurately classifying 73.3% and 86.5% of diploid males and triploids respectively (Table 4.4). The LDA function was then used to

predict the ploidy of 22 individuals of unknown ploidy, of which 21 achieved over 60% confidence of the prediction being accurate.

Table 4.4: Cross-validation of diploid male (2AA) and triploid (3n) individuals collected in October from Ribeira da Murtega.

	2AA	3n
2AA (N = 30)	22	8
3n (N = 37)	5	32

4.4 Discussion

Discrimination between ploidy levels using either body shape or head shape appears to accurate show where differences can be found on the body, contrary to using scale shape, and confirms previous findings that illustrate that measurable differences in body shape of *S. alburnoides* between ploidy levels exist, notably in position of the mouth and body depth (Cunha et al., 2010). However, the study by Cunha et al. (2010) did not include non-hybrid diploid male individuals, as these are not found in the Northern basins of Portugal where the study was carried out. The findings presented here show that there is a very clear distinction between the body shape of diploid male when compared to other available ploidy levels. This distinction can sometimes be made in the field, however is reliant on examining the pattern of coloration on the fish rather than the body shape of the fish itself (pers. observation). It is unfortunate that no individuals were present where it was certain that their genetic makeup corresponded to diploid females. The discriminate function could have potentially also been able to discriminate between diploid females and triploids. The lack of diploid females is most likely due to uncertainties in the prediction of fish ploidy using scale

shape. It was surmised in Chapter 3 that some diploid female fish may have been classified as triploids, thus obscuring differences in this chapter. It is unknown whether a distinction could be made, but the presence of diploid females may account for some of the variation in shape that is found amongst the triploid individuals analysed.

Overall, body and head shape show similar distinctions and prediction of ploidy (Appendix 2), despite body shape showing more effective discrimination. This is due to some of the biggest differences between the two groups being present in the tail area of the fish and diploid males having more slender bodies. Differences between the two groups were also visible by only examining the head of *S. alburnoides,* where differences in shape were concentrated around the position of the mouth. Diploid males displayed a slightly more superior mouth, which could be related to feeding preferences (Karpouzi and Stergiou, 2003). The slender form of diploid males might be due to differences in habitat preferences and current velocity in *S.* alburnoides (Martins et al., 1998). However, this tends to contradict the majority of the literature that suggests that streamlined fish are more likely to be found in areas of higher water velocity (Morinville and Rasmussen, 2008). Despite this statement being largely true, other reports have found that fish present in higher currents have shorter caudal regions (Dieterman and Galat, 2005), as is the case with triploid individuals of *S. alburnoides*. Due to reports on optimum body morphology largely contradicting the known habitat preferences of *S. alburnoides*, it is suggested that the variation in body shape may also be due to more specific factors. Non-hybrid diploid males do not possess the genome of their maternal ancestor *S. pyrenaicus*, a much larger and generally broader fish, but only the genome of their paternal ancestor. Despite this paternal ancestor not being known, as it is most likely now extinct, it has been shown

that *Anaecypris hispanica* is a close relative to this ancestor (Robalo et al., 2006). In fact, the body shape of diploid male *S. alburnoides* present many similarities with that of *A. hispanica*, including the small size and slender body. It is to be concluded that although environmental factors may play a part in determining shape of the fish, the majority of differentiation must be explained by genetic factors.

Having established the differences in morphology between ploidy groups it is necessary to examine potential seasonal differences. It has not been possible to directly compare the morphology of individual ploidy levels across season due to sampling difficulties with photographs of the fish as described above. However, the data does demonstrate that distinctions between the two ploidy levels were much more pronounced in April, during the reproductive season. It is believed that this distinction is largely due to the presence of eggs in females, which have been previously linked to a greater body depth in other species of fish (Wesner et al., 2011). By definition, diploid males do not have eggs present during the sampling period, thus potentially greatly increasing existing differences in slenderness between the two groups. As the majority of triploid individuals collected were more than likely to be female (Crespo-Lopez et al., 2007), any differences found between female and male diploids have more than likely been incorporated into the total shape variation present within the sample. The hypothesis that the more accurate distinction is due to the presence of eggs is further supported by previous reports whereby non-invasive sexing of *S. alburnoides* is not possible outside the reproductive period (Collares-Pereira et al., 2007).

In conclusion, body and head shape analyses are effective in discriminating between ploidy levels of *S. alburnoides*. It should be combined with scale shape analyses, which

would provide a large degree of accuracy in the prediction of unknown ploidy. Future work should focus on improving the functions associated with body and head shape by including individuals where ploidy has been determined by genetic methods and not inferred solely from scale shape morphology. As with scale shape, it is likely that separate functions will have to be created to distinguish between river basins and taking into account the varied evolutionary history and reproductive preferences of *S. alburnoides*.

Chapter 5: Growth and ecology of *S. alburnoides*

5.1 Introduction

It has been established in Chapters 3 and 4 that the different ploidy levels of *S. alburnoides* have different morphological features by which they can be discriminated. Little is known as to how the differences found in *S. alburnoides* may be related to ecological conditions such as dietary and habitat preference as morphological features are known to be associated with feeding adaptations (Karpouzi and Stergiou, 2003) and habitat selection (Winemiller, 1991; Carlson and Wainwright, 2010), in many animals, and fish in particular. Size differences have been shown to influence predator avoidance tactics, with regards to microhabitat selection (Burns et al., 2009; Coyne and Orr, 2004). It therefore seems plausible that the morphological differences identified in *S. alburnoides* may be associated with a separation of ecological niches between ploidy levels, specifically with respect to dietary composition and habitat preference. If differences in resource use are to be found, they will help contribute towards the understanding of a species with a complex life and reproductive cycle. Individuals of different ploidy levels contribute varying amounts of genetic material to *S. alburnoides* offspring, influencing the evolutionary pattern of the species complex and the continuation of the ploidy levels found therein. If differing ploidy levels demonstrate varying levels of adaptability to changing environmental demands, specifically those related to habitat and dietary requirements, then these may influence reproductive choices and outcomes within the species. Ecological choices, such as habitat or dietary selection, may also be dependent on what stage of life an individual is at. Therefore, knowing the age and growth patterns of a species can help to better identify ecological preferences and where potential intra-specific competition can occur.

Analyses of growth rates and length at age values are a useful tool to distinguish between populations within species of fish, both for species that have a large range (Restrepo et al., 2010; Kendall and Quinn, 2009) and species confined to small riverine and lake systems (Robinson et al., 2010). Many methods have been used to determine growth rates of fish. Direct observation of fish age is the simplest method of data collection, often used in aquaculture, where fish of known age are measured in both length and weight (Hurvitz et al., 2008). However, this method cannot be used for the majority of ecological studies. Instead, length-frequency methods are more commonly used in fisheries biology. By looking at length-distribution patterns and their distribution peaks, different cohorts (years in which fish were born) can be potentially identified (e.g. Alarcón et al., 2011; Nunn et al., 2002). This method can also have its drawbacks as it may not be able to distinguish between different populations or may miss out cohorts from within the population distribution (Heery and Berkson, 2009). As such, the most widely used method is the estimation of age by observing the hard parts of fish. The hard parts of fish show patterns that can be used to age individual fish and determine their growth rate. A number of different parts have been used for many species, ranging from scales (Britton, 2007) and otoliths (Mercier et al., 2011), to vertebrae (Hale and Lowe, 2008) and fin rays (Weber and M. L. Brown, 2011). The structure used may often depend on the species sampled and ease of collecting the necessary sample. Comparisons have been made to determine which part is the most accurate for age estimation. Most methodologies show comparable results within a species, but there is no hard part that is more informative across the majority of species, as different species have been reliably aged using different parts (Spiegel et al., 2010; Zymonas and McMahon, 2009). The analysis of scales has advantages in that it does not require that the fish be killed, and that any scales removed will regenerate

(Sire and Akimenko, 2004). As such, this method is likely to cause minimal harm and distress to the fish. Circular patterns can be seen on scales that correspond to the deposition of collagenous material at different stages of an individual's life (Sire and Akimenko, 2004). To a trained observer, these patterns can be identified as year marks (annuli), providing a representation of the age of the fish. Patterns of growth and food availability can also be seen on scales, which in some cases may lead to inaccurate aging (Zymonas and McMahon, 2009). However, these patterns also have their use when identifying food availability for a population (Ibáñez et al., 2008). The combination of the above factors mean that the hard parts of fish, whether from scales as described or from other available structures, are a useful tool to piece together the life history and specific growth patterns of individual fish and the populations they come from.

Ribeiro et al. (2003) examined the growth rate of *S. alburnoides* from the Guadiana river basin in Portugal using scales in an attempt to identify whether different ploidy groups had different growth rates. They found that diploid males have slower growth rates than both diploid females and triploid individuals. However, little work has been done since to obtain more data on the growth of these fish. Cohorts may vary in their growth patterns significantly and environmental changes, on habitat and diet, may also modify the growth rate, when compared to previous research. A section of this chapter will attempt to identify whether diploid males have a slower growth rate than other ploidy groups, and to corroborate the findings of Ribeiro et al. (2003). If possible, distinctions in growth rates between diploid females and triploids will also be made, with the aim that this may help reliably identify genetic groups without the need for genetic analyses. The chapter will also attempt to identify whether ploidy groups have

different life expectancies as it is anticipated (Collares-Pereira et al., 2007) that diploid males will not live as long as other ploidy levels, in part due to their small size which may increase the risk of predation, inclusive by conspecifics, as well as inherent life expectancies related to ploidy.

Niche partitioning in *S. alburnoides* is also likely to exist, except during the reproductive period when ploidy levels share habitat resources when breeding (Martins et al., 1998; Gomes-Ferreira et al., 2005). Conditions in the rivers for the rest of the year are highly variable, with different ploidy levels of *S. alburnoides* reacting differently to the changes. The Guadiana river basin has Mediterranean climatic conditions, characterised by seasonal fluctuations in river flow, which causes the rivers in summer to almost dry out. Remaining habitats are reduced to pools where high densities of fish can be found (Godinho et al., 2000). In winter, the rivers are subjected to large winter floods (Magalhaes, 1993b). The identification of niche requirements in a relatively untouched system, such as that of the Ribeira da Murtega, could be then applied to affected systems, ensuring that populations from other river systems have better chances of continued survival. The increasing number of dams and water abstraction facilities, specifically the creation of the Alqueva dam, has modified the natural flow pattern of the river (Martínez-Capel et al., 2009; Chícharo et al., 2006) and may have modified the number of pools available during the summer season, reducing the number and type of available habitats for all species present in affected parts of the river system. Increasing the knowledge on the biology of the species is thus imperative, as many of the habitats occupied by *S. alburnoides* are likely to be affected by the creation of new structures on and along the river. Such conservation measures can only be implemented when the biology of the species is fully understood, relating to

growth, reproductive and feeding characteristics. This chapter will attempt to enhance the knowledge relating to growth and feeding characteristics.

Clear separations have been shown between ploidy levels in *S. alburnoides* with regards to feeding and habitat, but such separation also exists between male and female individuals (Gomes-Ferreira et al., 2005). This distinction between the sexes is in part determined by the grouping of AA diploid males with PA diploid individuals in the study by Gomes-Ferreira et al. (2005), potentially attributing differences to sex that are caused by ploidy variation, despite the rarity of male PA diploids (Alves et al., 2002). Differences in habitat and feeding among species of fish does not have to occur over a large scale, and examples have been found of differences in both habitat and diet within a small stretch of river system (Morán-López et al., 2006). Such small scale differences can be determined by sampling microhabitat units, which represent distinct habitats within a stream or river system (Copp and Penaz, 1988). The differences between these habitats have been related to differences in fish and prey community assemblages (Pires, Cowx, and Coelho, 2000a; Pires et al., 1999).

This chapter also aims to identify further differences in diet between the ploidy levels by determining whether certain prey items play a key role in the food consumed by different *S. alburnoides* groups and whether prey selection is modified by seasonal variations. It will also aim to establish where, if they exist, differences in habitat preferences lie, and which factors, such as current velocity and substrate composition, influence *S. alburnoides* the most. In addition, associations will be made between prey items and habitats.

5.2 Methods

5.2.1 Age and growth

As a large number of scales were available and have been reliably used in previous aging studies, these were the parts chosen to be used for aging. Scales were always taken from the shoulder area of the fish (See Chapter 3 for details). Only fish collected from the Murtega river in April and October 2009 were used for scale aging, as length data was not available for all other fish involved in the previous analyses. The standard length of all fish was measured and the ploidy was determined by using the prediction function created in the scale shape chapter (Chapter 3). Only fish whose ploidy had been reliably predicted (using the 60% cutoff point established previously) were included in scale aging analyses. In total 121 fish were analysed, including 75 triploids and 46 diploid males. Age was determined by observing the scale directly under a Leica S8APO microscope or at a later date from the pictures taken. Defined lengths (total radius or radius to annuli) on the scale were measured either in Leica Application Suite (Leica Microsystems, 2008) or in TPSDig2 (Rohlf, 2010a), both of which have the ability to calibrate distance measured if the magnification at which the picture is taken is known. The total radius was measured using landmarks used in Chapter 3 as a reference guide, measuring between landmarks 2 and 1 (Figure 5.1). The age of each fish was determined by counting the number of annuli on each scale. Age and radial distances were determined by measuring one scale per fish, except in the case of uncertain aging, where further scales were used to confirm age. This was done by identifying patterns within the scale (checks) that represent periods of fast and slow growth (Bagenal, 1978). For each scale, the radial distance from the focus to each annulus was also measured, along the line defined for total radius.

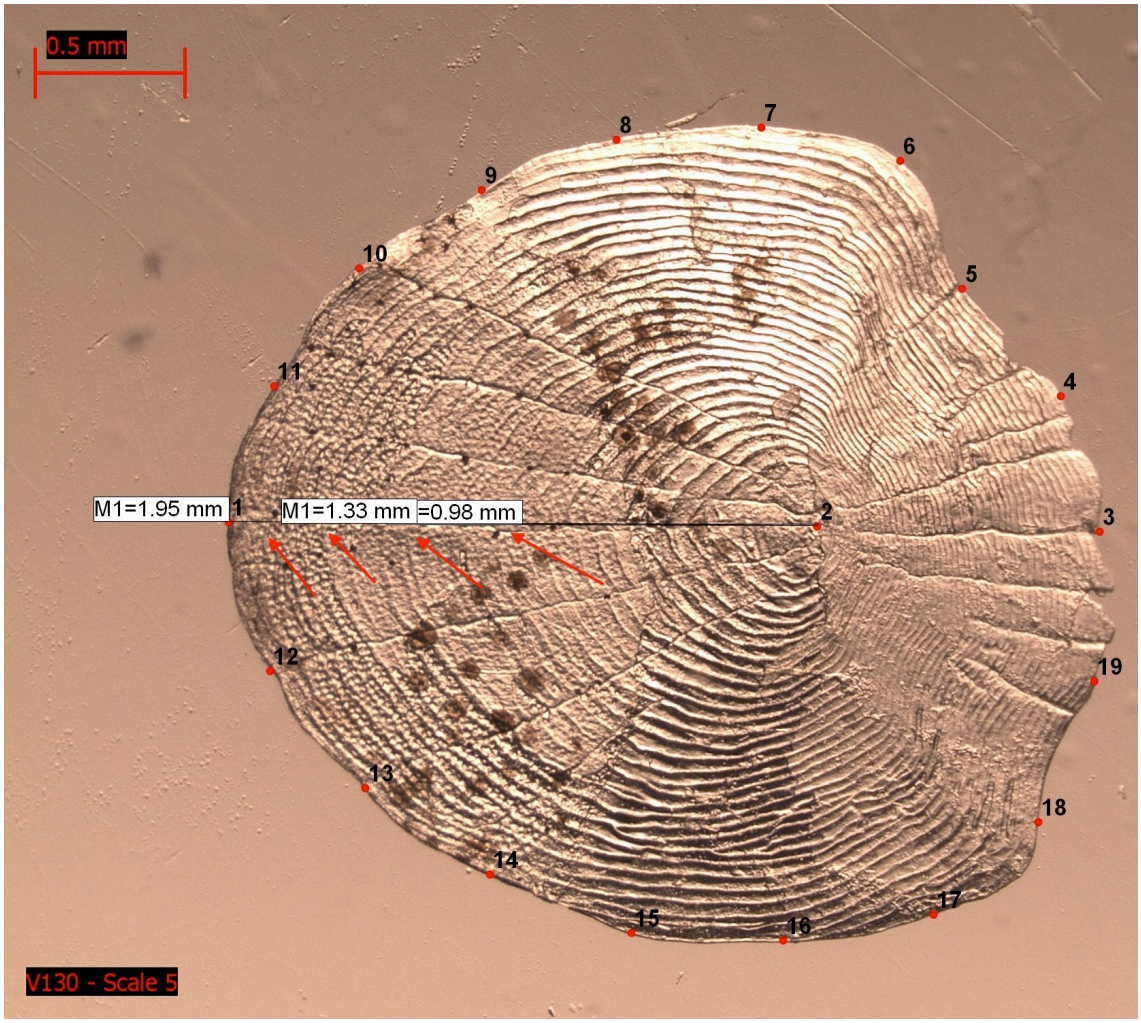


Figure 5.1: Example scale from a *S. alburnoides* individual aged 4+. All four checks are shown, however for clarity only the first two are measured, as well as the total scale radius, between landmarks 2 and 1.

Standard length of fish was plotted against total scale radius to determine whether all fish sampled originated from the same population and whether different ploidy groups should be analysed separately. This was determined by identifying if all fish sampled followed the same growth trend. Two methods were used to back-calculate length at age values for each individual fish: the Dahl-Lea equation (Lea, 1910; Francis, 1990) and ordinary least-squares linear regression. The Dahl-Lea equation (1) uses the fish

length at capture (L_c), the total scale radius at capture (S_c) and the scale length at age *i* (S_i) to determine lengths at age *i* (L_i).

(1) $$L_i = \left(\frac{S_i}{S_c}\right) L_c$$

The regression method uses the linear relationship between fish standard length and total scale radius (Equation 2) for each group to estimate the fish length at the point of laying down of each annulus along the scale radius:

(2) $$L_i = aS_i + b$$

where *a* represents the slope and *b* the intercept. In biological terms, the intercept is the length at which scales are first deposited on the fish, which is always greater than 0, as squamation of scales occurs sometime after hatching (Robinson et al., 2010).

Using the Dahl-Lea equation for growth and the regression equations, mean lengths for each age group were calculated. Growth curves were plotted for each ploidy type and an ANOVA, performed in R (R Development Core Team, 2011), was used to test for differences in length between groups of the same age and whether different ploidy levels had different growth rates. A Von Bertalanffy growth model was applied using back-calculated values from both methodologies, which created a growth curve (Britton, 2007; Bertalanffy, 1938), describing the growth of *S. alburnoides*, according to Equation 3. L_∞ represented the maximum theoretical length in mm, K represented the growth coefficient per year and t_0 was the time when the length theoretically was equal to 0mm. These values were estimated according to the methodology described in Hickley and Dexter (1979).

(3) $L_t = L_\infty\left(1 - e^{-K(t-t_0)}\right)$ as in Hale and Lowe (2008).

5.2.2 Dietary composition

All *S. alburnoides* collected in 2009 were dissected to determine sex and the intestines removed for dietary analysis. As cyprinids do not have a differentiated stomach, all gut contents were placed under a cold light source attached to a Leica S8APO microscope. The gut contents were not allowed to dry out whilst identification of different food items was taking place.

All food items were identified to the lowest taxonomic level possible, with most organisms being identified to the family level using keys found in Fitter (1986). Several methods to determine dietary composition were used, all of which, as well as other dietary counting options, are described in further detail in Hynes (1950). Initially stomach fullness was assessed to determine the intensity of feeding of each individual. The fullness method demonstrates how full an individual's stomach is at the time of capture, ranging from empty to distended, examined upon dissection of the stomach (Table 5.1).

Table 5.1: Description of values used in the fullness method.

Ranking	Description of stomach
0	Empty stomach
2	Quarter full stomach
4	Half full stomach
6	Three-quarter full stomach
8	Full stomach
10	Distended stomach

The volumetric method was used to determine the quantity of each food item present in the stomach. A percentage value, directly relating to volume, was estimated by eye,

ensuring the total volume added up to 100%. In addition to this, the numerical method was initially used, where the number of each food item eaten was determined. However, this method was abandoned as the presence of large quantities of filamentous algae in certain individuals did not lead itself to be counted.

The length of all fish dissected was measured to the nearest millimetre whilst still alive, during the sampling field trips. Ploidy level was determined by scale shape (See Chapter 3), therefore only fish whose ploidy was predicted to at least 60% accuracy were included in the analysis (121 individuals, 47 of which were diploid male and 74 that were triploid). Forty distinct diet items were identified, of which a large number were found in small amounts. As such, following the methodology described in Gomes-Ferreira et al. (2005), items that had an occurrence of less than 5% in both diploid male and triploid groups were pooled to form a group named "Other". This pooling resulted in 22 prey groups being used for dietary analyses, according to the following categories: Simuliidae larvae, Hydropsyche larvae, Chironomid larvae, Baetidae larvae, Corixidae, Simuliidae adult, Formicidae, benthic algae, Diptera adult, Culicidae adult, Coleoptera adult, Elmidae larvae, Chydoridae, Ostracoda, Ephemerellidae nymphs, Caenidae nymphs, Ephemeroptera nymphs, eggs, Harpacticoida, fish, Dixidae adult and Other.

Diversity of organisms consumed was estimated using the Shannon-Wiener index of diversity (*H'*, Equation 4), which allows an assessment of the range of organisms consumed, taking into account both the quantity of prey consumed and evenness of prey among the sample (Merkel et al., 2007; Zar, 1999). The higher the value of *H'*, the greater the diversity in prey eaten by *S. alburnoides* individuals. Values of *H'* were log-transformed for normality before t-tests were performed to assess whether significant

differences in dietary diversity existed between ploidy levels, as in Jiangfeng et al (2011). Overlap in dietary choice between the two ploidy levels, as well as between different time periods was estimated using Schoener's Index of overlap (Schoener, 1970), described by Equation 5. This index ranges from 0 to 1, where 0 represents no overlap and 1 represents identical dietary preferences. It is often expressed as a percentage of shared resources and considered to be biologically significant if the overlap is greater than 60% (Wallace, 1981).

(4) $$H' = -\sum_i p_i \ln p_i$$ as in Washington (1984).

(5) $$C_{xy} = 1 - 0.5\left(\sum_{i=1}^{n} \left| p_{xi} - p_{yi} \right|\right)$$ as in Madurell and Cartes (2005).

A number of ordination methods are used in studies of both diet and habitat choice (e.g. Beamish et al., 2006; Bolger et al., 1990; Vilizzi and Copp, 2008; Pusey et al., 2010) to highlight factors that may influence the studied species. Canonical correspondence analysis (CCA) is one such method that allows the extraction of environmental variables that maximise niche separation among species. Data used in a CCA is typically occurrence data of species (often counts of individuals) at a number of different "sites" and the environmental variables taken at each site to explain the variation present among species (Braak and Verdonschot, 1995). In the case of dietary analysis, site scores correspond to individual fish and species scores are representative of the prey items consumed. The magnitude of the influence of environmental variables on species and site variables is determined by the length of the arrows displayed in the CCA plot, indicating where differences in diet lie.

5.2.3 Habitat preference

A total of 50 distinct sampling points, of approximately 1m^2 each, were sampled in Ribeira da Murtega in April and October 2009. Detailed sampling methodology can be found in the General Methods section of this thesis. Environmental variables were measured and community composition was assessed by identifying all fish collected to species level. All fish were measured (standard length) in the field, and all fish other than *S. alburnoides* were subsequently returned to the stream as close as possible to the microhabitat where they were captured. All *S. alburnoides* were removed for subsequent examination and the majority were then identified to ploidy level using the techniques described in Chapter 3. Seventeen *S. alburnoides* were not identified to ploidy level as they were used in another study at the Universidade de Lisboa in Portugal (Morgado-Santos et al., 2010) and no scales were removed for morphological identification. CCA was performed to identify which environmental variables, if any, played a role in habitat preference of the *S. alburnoides* species complex. Distinctions within *S. alburnoides* were made, as previous work has shown habitat segregation within the species (Martins et al., 1998), therefore individuals were split into six distinct groups. Diploid males were separated from triploids and within these groups individuals were separated according to age classes. Individuals less than one year old were considered to be juveniles, as these were only found in fish sampled in October, which would have been born in April or May of the same year. A distinction was also made between individuals over one year old and those of three years and above. The majority of individuals reach sexual maturity in their second year of life, but many diploid males do not live longer than three years (Collares-Pereira et al., 2007). It was thus felt important that a distinction be made between individuals at the beginning and end of their reproductive life cycle. Individuals that could not be identified to

ploidy level (See Chapter 3) were grouped as "Unidentified *S. alburnoides*" and may have included fish of both ploidy levels and varying ages.

All analyses to determine ecological preferences and variations were performed using R (R Development Core Team, 2011).

5.3 Results

5.3.1 Age and growth

The standard length of each fish was plotted against scale radius to determine whether all fish sampled followed the same relationship between scale dimension and length (Figure 5.2). In general, diploid males were of shorter length than the majority of triploids, with a mean length of 45mm (± 1.3 mm SE), as opposed to 72mm (± 1.8 mm SE) for triploids. Length-frequency histograms for each sampling period showed a wider range of lengths in October, including juveniles smaller than 40 mm, which were not present in April (Figure 5.3). The sample collected in October also contained a greater number of larger fish, with a length greater than 80 mm. The length-frequency distributions indicated that the majority of the fish sampled in April were probably in their third year, with few individuals from that sampling trip reaching ages greater than four.

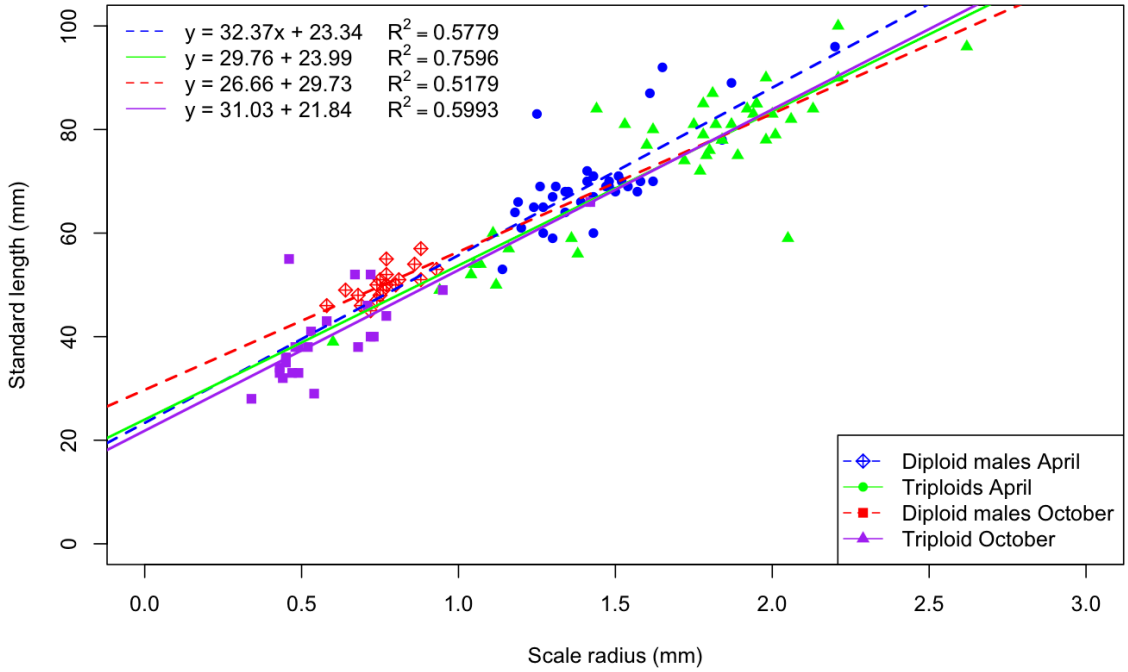


Figure 5.2: Standard length of individuals at capture against scale radius at capture. Regression lines are calculated according to ploidy and sampling period, and all R^2 values have $p < 0.05$.

For the Dahl-Lea method, the mean length at age was compared between triploids and male diploids using ANOVA. The average yearly growth for each cohort was calculated, showing that in the sample available there has been an increase in the size of fish at age one between cohorts from 2004 up to 2008, irrespective of ploidy (Table 5.2). No difference was found in the length of fish at age 1 ($F_{1,107} = 0.637$, $p = 0.427$). However, differences were found between the two groups at age 2 ($F_{1,84} = 9.067$, $p = 0.003$) and at age 3 ($F_{1,64} = 4.504$, $p = 0.038$). There was insufficient data for fish at ages 4 and 5 to accurately compare differences.

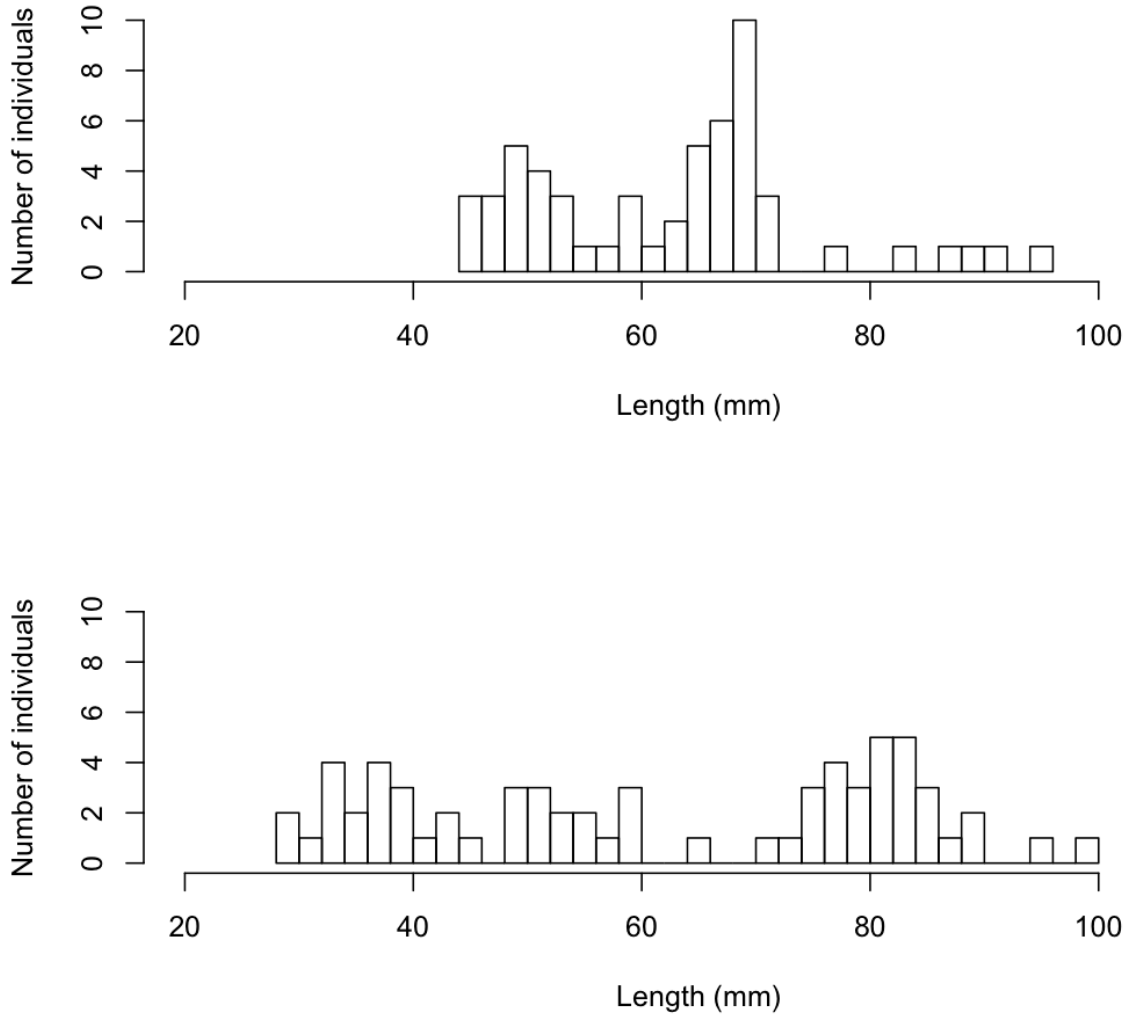


Figure 5.3: Length-frequency histogram of fish collected from Ribeira da Murtega in April (top) and October (bottom). Each break corresponds to a length class difference of 2 mm.

Figure 5.2 indicated that the Dahl-Lea equation for back-calculated growth was valid, as both ploidy levels displayed isometric growth. Regressions, calculating growth for each of the two ploidy groups were also calculated and the length for age using both methods of back-calculation were compared.

The back-calculated values were used to calculate Von Bertalanffy growth curves for diploid males and triploids (Figure 5.4). Estimated L_∞, the theoretical maximum length, was 55 mm for diploid males and 90 mm for triploid individuals. The growth coefficient

for diploid males was 0.95 years^{-1}, which was larger than that found for triploids, which was estimated at 0.45 years $^{-1}$.

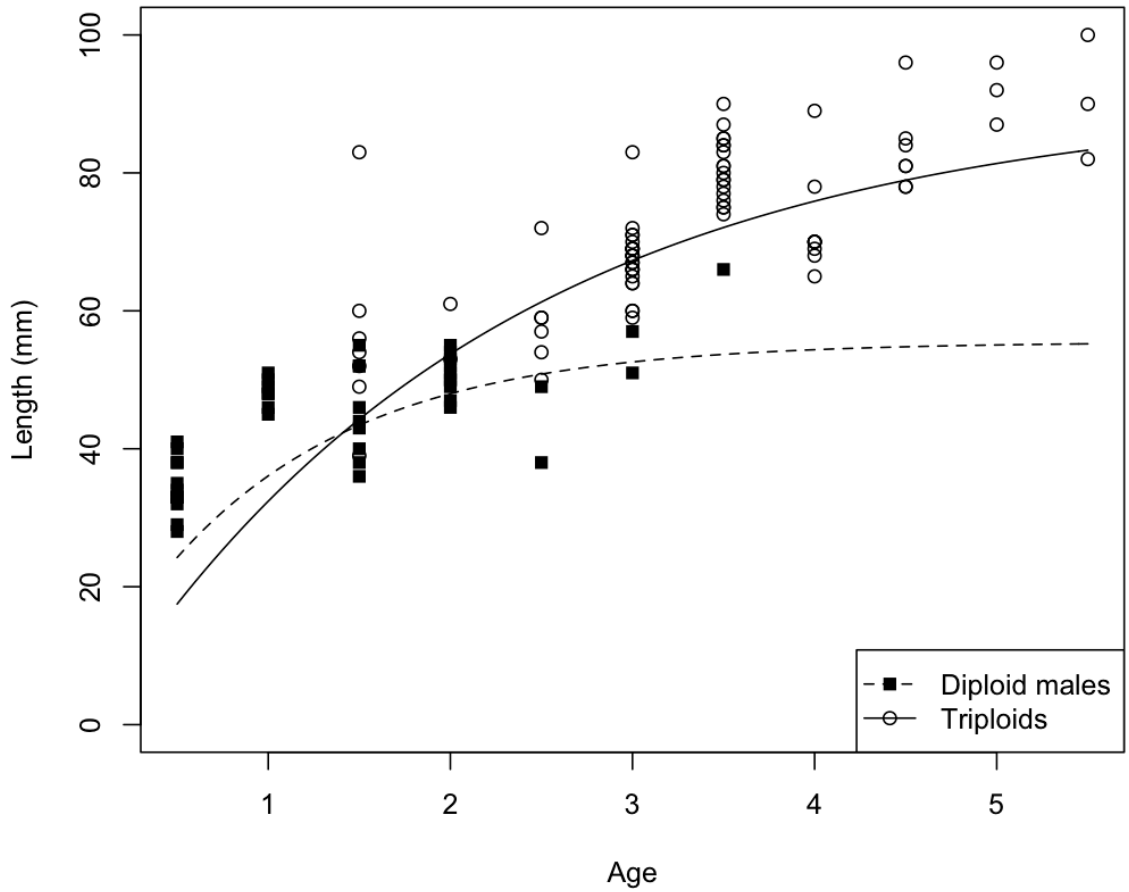


Figure 5.4: Von Bertalanffy growth curves for diploid males and triploids, calculated using Dahl-Lea back-calculated values, superimposed on measured length-at-age values from fish collected in April and October in Ribeira da Murtega. Fish collected in October were offset by 0.5 years to account for plus growth and are represented by the points between age values. This was done to account for fish born that same year that had not yet reached one year in age.

Table 5.2: Mean back-calculated standard lengths and standard errors (S.E.) at age, using the Dahl-Lea equation, and annual increment for each cohort from diploid male and triploid individuals, from the Ribeira da Murtega.

Cohorts	Age Groups					
	N	1	2	3	4	5
		SL ± SE (mm)	SL ± SE (mm)	SL ± SE (mm)	SL ± SE (mm)	SL ± SE (mm)
Diploid males	43					
2006	3	31.1 ± 4.2	41.83 ± 5.2	53.67 ± 4		
2007	12	33.24 ± 1.2	44.88 ± 1.1			
2008	16	36.56 ± 1				
Average		34.75 ± 0.9	44.27 ± 1.3	53.67 ± 4		
Annual increment			9.52	9.40		
Triploids	78					
2004	6	33.33 ± 3	47.85 ± 2.8	65.25 ± 2	74.57 ± 1.9	83.93 ± 2.2
2005	16	32.62 ± 2.1	47.06 ± 2.4	60.79 ± 2.4	71.29 ± 2.2	
2006	41	35.28 ± 1.1	52.83 ± 1.3	67.32 ± 1.4		
2007	8	39.40 ± 2.1	52.04 ± 1.1			
2008	7	46.59 ± 3.6				
Average		36.02 ± 0.9	51.02 ± 1	65.18 ± 1.2	72.18 ± 1.7	83.93 ± 2.2
Annual increment			15	14.06	7	11.75

The regression method (Equation 2) provided slightly different results, with ANOVA showing differences between ploidy levels at all ages (Table 5.3). Mean lengths-at-age for each cohort are displayed in Table 5.4, where the greatest variation in length between ages was found between ages 1 and 2, showing that the most rapid period of growth is in the early stages of life.

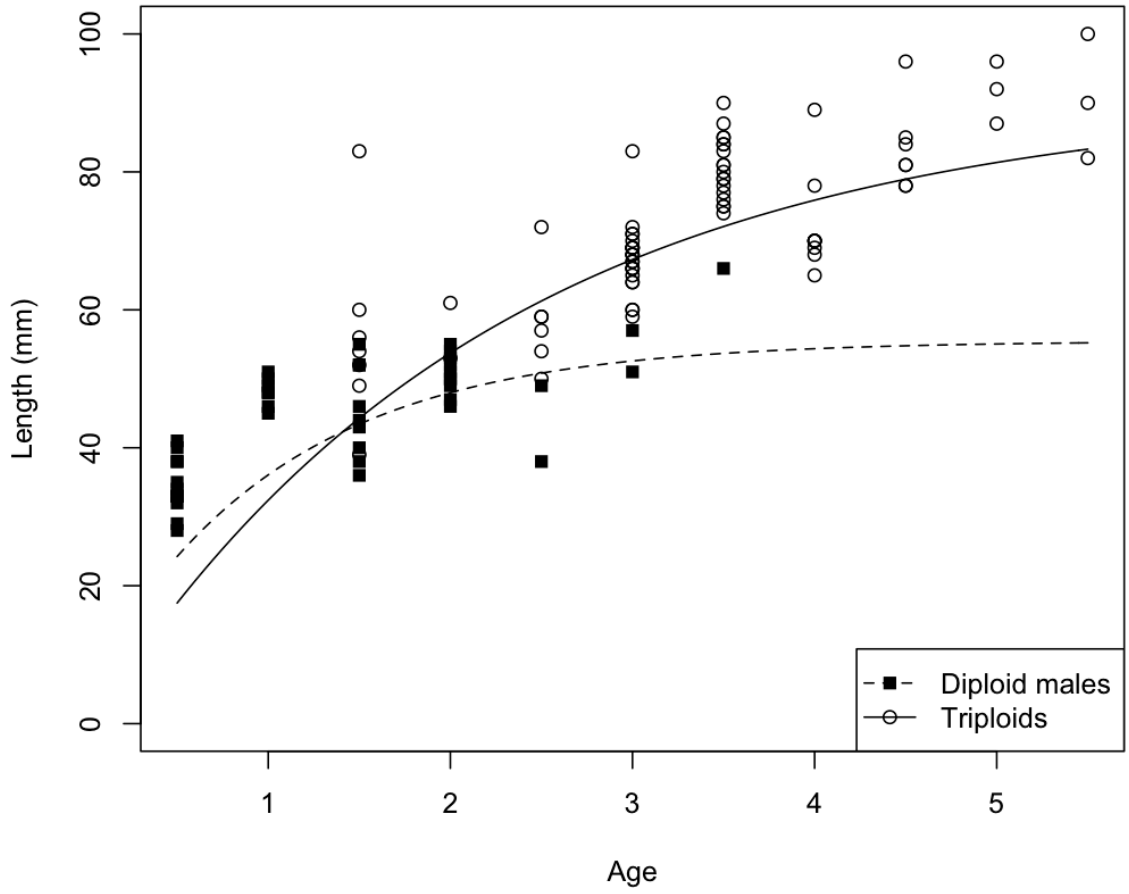


Figure 5.5: Von Bertalanffy growth curves for diploid males and triploids, calculated using linear regression back-calculated values, superimposed on measured length-at-age values from fish collected in April and October in Ribeira da Murtega. Fish collected in October were offset by 0.5 years to account for plus growth.

Table 5.3: Results from the ANOVA, comparing mean lengths calculated using the regression line between ploidy levels (diploid males and triploids).

	Age 1	Age 2	Age 3
ANOVA	F = 20.27	F = 24.58	F = 8.21
	d.f. = 1, 107	d.f. = 1, 84	d.f. = 1, 64
	$p < 0.001$	$p < 0.001$	$p < 0.05$

The Von Bertalanffy growth model was also performed on back-calculated values from the regression equation (Figure 5.5). Diploid males were calculated to have a L_∞ of 54 mm, a growth coefficient K of 1.46 years^{-1} and t_0 of 0.20 years. Triploids had a greater L_∞ at 77 mm, a lower growth coefficient (K) of 0.88 years^{-1} and a value of 0.29 for t_0.

Table 5.4: Mean back-calculated standard lengths and standard errors (S.E.) at age, using the regression method, and annual increment for each cohort from diploid male and triploid individuals, from the Ribeira da Murtega.

Cohorts	Age Groups					
	N	1	2	3	4	5
		SL ± SE (mm)	SL ± SE (mm)	SL ± SE (mm)	SL ± SE (mm)	SL ± SE (mm)
Diploid males	43					
2006	3	43.57 ± 2.5	49.12 ± 4.1	54.85 ± 4.4		
2007	12	43.08 ± 0.8	48.07 ± 1.1			
2008	16	40.26 ± 1.1				
Average		41.67 ± 0.7	48.28 ± 1.1	54.85 ± 4.4		
Annual increment			6.61	6.57		
Triploids	78					
2004	6	45.92 ± 2.9	55.75 ± 2.9	65.50 ± 2.6	73.87 ± 2.8	80.20 ± 2.8
2005	16	46.09 ± 1.6	56.48 ± 2	66.30 ± 2.1	73.86 ± 2.1	
2006	41	46.58 ± 0.9	58.43 ± 1.1	68.11 ± 1.1		
2007	8	51.56 ± 2.9	60.57 ± 3.2			
2008	7	52.04 ± 3.2				
Average		47.43 ± 0.8	58.00 ± 0.9	67.40 ± 0.9	73.87 ± 1.7	80.20 ± 2.8
Annual increment			10.57	9.40	6.47	6.33

When comparing the regression method to the Dahl-Lea method, length at age values were higher for all groups when calculated using the regression method. However, results from both methods indicated that triploids are bigger at all ages than diploid males (Figure 5.6). The Von Bertalanffy growth models displayed similar results using both methods, where the growth rate for diploid males was higher than that for triploids. However, neither model accurately represented the growth pattern of *S. alburnoides* in the first two years of life.

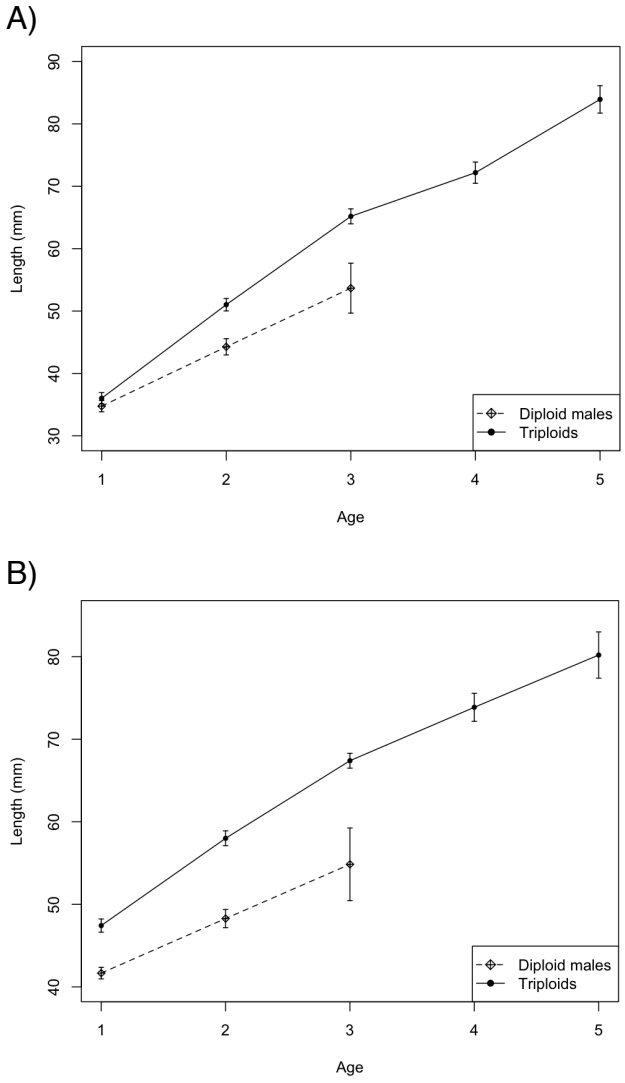


Figure 5.6: Length at age of Murtega populations calculated using Dahl-Lea (A) and regression methods (B).

5.3.2 Dietary composition

Dietary composition was determined for each ploidy irrespective of season but also by identifying seasonal differences in dietary intake. Simuliidae larvae predominated in the diets of both diploid males and triploids (22.31% and 19.17% of total stomach contents respectively). However, the most abundant item found in the gut of diploid males were Chydoridae (25.56%), which only accounted for 2.23% in triploids. Diploid males also consumed large quantities of Harpacticoida (22.13%) over both seasons, as did triploids (18.26%). Triploids had a high proportion of benthic algae (19.18%) in their

gut contents, which were rarely present in diploid males (0.68%). Triploid individuals showed preference for a wider range of diet items than diploid males, consuming Coleoptera adults (2.36%), Ephemeroptera nymphs (1.25%) and other fish (3.26%), all of which were not consumed by diploid males.

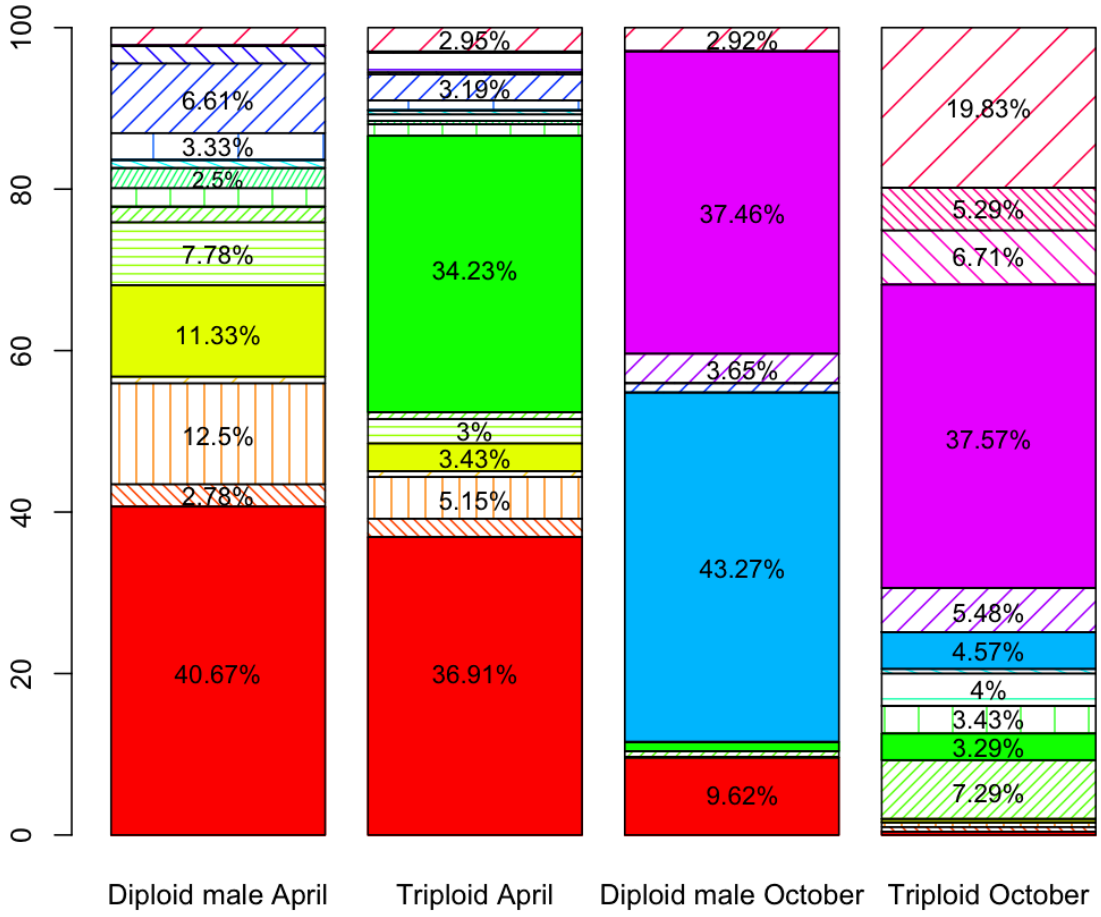


Figure 5.7: Dietary composition of *S. alburnoides* diploid males and triploids from two sampling seasons, April and October. Simuliidae larvae Hydropsyche larvae Chironomidae larvae Baetidae larvae Corixidae Simuliidae adult Formicidae benthic algae Diptera adult Culicidae adult Coleoptera adult Elmidae larvae Chydoridae Ostracoda Ephemerellidae nymphs Caenidae nymph Ephemeroptera nymph Eggs Harpacticoida Fish Dixidae adult Other. Percentage occurrence only indicated on items over 2.5%, for clarity.

When seasonal variation was taken into account (Figure 5.7), in April 40.60% of the items consumed by diploid males were Simuliidae larvae, followed by Chironomidae

larvae (12.50%). Triploids consumed mostly Simuliidae larvae (36.90%) and benthic algae (34.22%), followed by Chironomidae larvae (5.14%). In October, the gut contents of diploid male individuals contained mostly Chydoridae (43.26%) and Harpacticoida (37.46%). Triploid individuals also consumed a large quantity of Harpacticoida (37.57%) but did not show any preference for any other individual food item, eating small quantities of a wide variety of organisms, including other fish (6.71%). Not all fish found in the stomach contents could be reliably identified to species level, due to varying stages of decomposition, however those that could be identified were *Gambusia holbrookii* and *S. alburnoides*. One sample was in very good condition and was identified as a diploid male *S. alburnoides* due to size and markings present on the body.

Triploids consumed a wider range of prey items than diploid males, even in October, where prey diversity consumed was greatly reduced when compared to April. This was demonstrated by the Shannon-Wiener index of diversity, with higher values for triploids when both months sampled were considered together and for sampling from October only (Table 5.5).

Table 5.5: Shannon-Wiener dietary diversity index for prey items consumed by *S. alburnoides.*

	Diploid male	Triploid	Diploid male April	Triploid April	Diploid male October	Triploid October
Index	0.703	0.775	1.246	1.013	0.327	0.524

In April, diploid males showed a higher diversity than triploids, most likely due to a tendency for triploids to consume benthic algae, which generally accounted for at least 50 % of the total gut contents. Many individuals that ate algae also presented full or

distended stomachs (8 or 10 on the fullness method scale). The greatest difference in prey diversity was between seasons, with diversity of prey consumed in April much higher than in October. A t-test was performed to demonstrate if and where differences were found between ploidy groups. When both seasons were pooled together, there was no difference found in the range of items eaten by diploid male and triploid individuals (t = -0.871, *p* = 0.386). No differences were found between ploidy levels sampled during the same season (April: t = 1.381, *p* = 0.177; October: t= -1.616, *p* = 0.112). Differences were however found between different seasons, both within the same ploidy and between both ploidy levels (Table 5.6).

Table 5.6: Results of t-tests showing differences in diet diversity between and within ploidy groups sampled at different seasons.

	t - value	df	Significance (*p)*
Diploid male April vs Diploid male October	6.250	36.341	< 0.001
Diploid male April vs Triploid October	4.748	39.418	< 0.001
Triploid April vs Diploid male October	6.117	52.486	< 0.001
Triploid April vs Triploid October	4.265	65.46	< 0.001

Schoener's index of dietary overlap showed that there was little similarity (<60% overlap) in the species of prey eaten by both ploidy levels between April and October. When both sampling seasons were analysed together, the overlap in dietary composition represented just over half of the prey consumed. The same result was found when seasons were analysed separately and dietary overlap between ploidy levels examined (Table 5.7).

Table 5.7: Dietary overlap between varying ploidy levels and seasons of *S. alburnoides.*

	Schoener's index		Schoener's index
Triploid vs Diploid male	0.583	Diploid April vs Diploid October	0.138
Triploid April vs Diploid April	0.616	Triploid April vs Diploid October	0.158
Triploid October vs Diploid October	0.510	Diploid April vs Triploid October	0.090
Triploid April vs Triploid October	0.120		

CCA was performed to show where variability existed in the sampled data and demonstrate how prey items were correlated with individuals and their characteristics. The CCA model accounted for 17% of the total variability in the sample, using ploidy, length, age and season as explanatory variables. The plot showed differences in the diet of *S. alburnoides* between seasons, represented by variation along the first axis (Figure 5.8). There was also a separation along the second axis that correlated to ploidy, age and length, creating a distinction between the two ploidy levels. The plot also highlighted food items such as fish that were exclusive to a particular group or items like Simuliidae and Chironomidae larvae that were consumed by both ploidy levels in similar amounts.

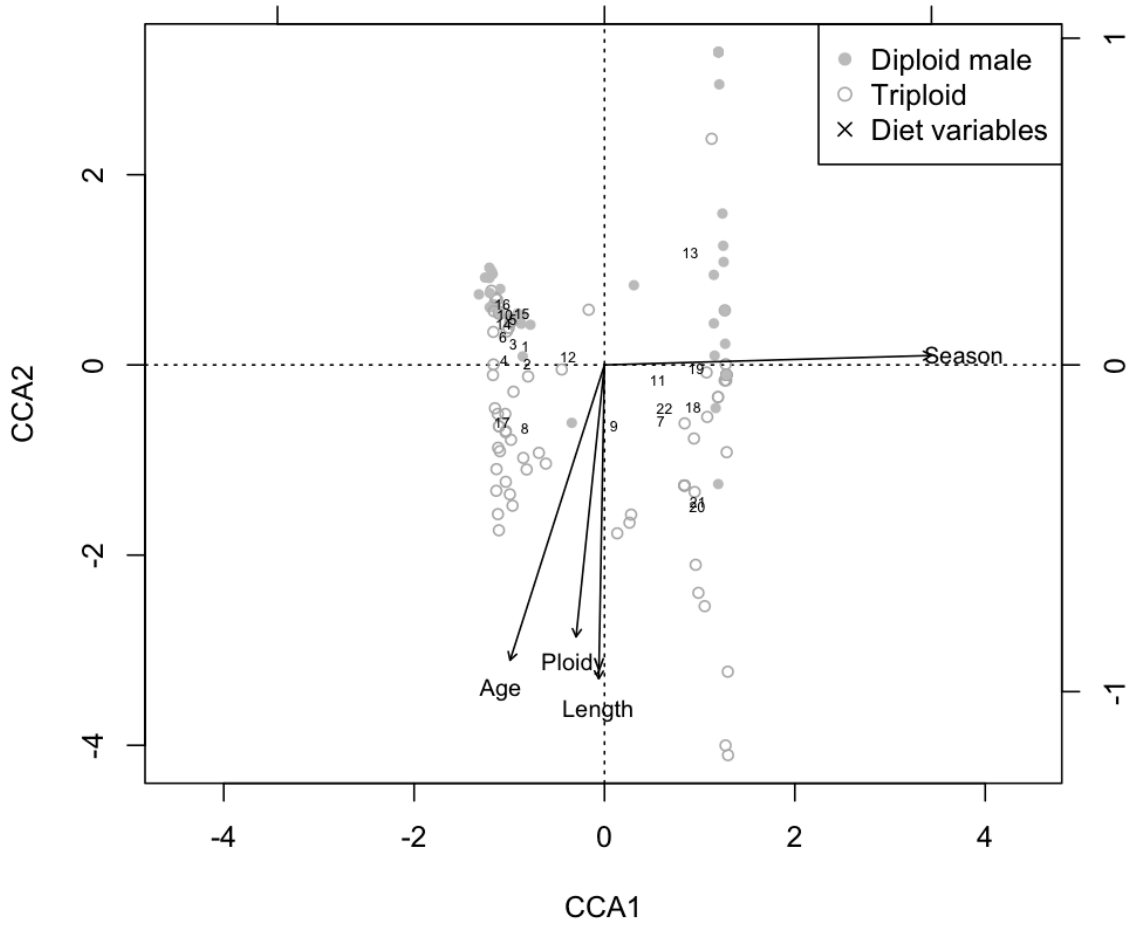


Figure 5.8: Representation of the CCA performed on the diet of *S. alburnoides*, using ploidy, length, age and season as factors to explain the variation in the sample. Diet items represented by numbers are: 1. Simuliidae larvae 2. Hydropsyche larvae 3. Chironomidae larvae 4. Baetidae larvae 5. Corixidae 6. Simuliidae adult 7. Formicidae 8. Benthic algae 9. Diptera adult 10. Culicidae adult 11. Coleoptera adult 12. Elmidae larvae 13. Chydoridae 14. Ostracoda 15. Ephemerellidae nymph 16. Caenidae nymph 17. Ephemeroptera nymph 18. Eggs 19. Harpacticoida 20. Fish 21. Dixidae adult 22. Other.

5.3.3 Habitat preferences

Of the 50 microhabitats sampled, 47 contained fish and were used to assess species preferences according to a number of environmental variables. In addition to *S. alburnoides*, a further six species were found in Ribeira da Murtega: *Squalius pyrenaicus, Lepomis gibbosus, Gambusia holbrookii, Barbus* sp.*, B. microcephalus* and *B. steindachneri*. *Barbus* sp. encompasses individuals that could not be identified to the species level within the genus *Barbus* due to the morphological similarity between

species. Table 5.8 provides the number of individuals found in all microhabitats for each species taxon.

Table 5.8: Number of each species found in Ribeira da Murtega, irrespective of microhabitat and season, and maximum and minimum lengths for each species.

Species grouping	N	L_{min} (mm)	L_{max} (mm)
Squalius alburnoides			
Diploid male < 1 year	11	28	41
Diploid male 1 -2 years	29	36	55
Diploid male ≥ 3 years	6	51	78
Triploid 1 -2 years	13	49	83
Triploid ≥ 3 years	60	59	100
Unidentified	66	27	99
Squalius pyrenaicus	1	108	108
Lepomis gibbosus	5	42	60
Gambusia holbrookii	57	19	42
Barbus sp.	117	55	248
Barbus microcephalus	40	101	182
Barbus steindachneri	40	94	182

The ordination of species and environmental variables according to the CCA showed that the variables explained 26.31% of the total variation within the sample, with the first two axes representing 67.6% of that variation. Younger diploid male *S. alburnoides* appeared to occupy different habitats from older diploid males, triploid *S. alburnoides* and all other species present in the Ribeira da Murtega (Figure 5.9). The percentage of gravel present in the substrate showed a very strong influence on the microhabitat choice of species sampled. High current velocities and depths also influenced the type of fish present, with particular attention to *S. alburnoides* individuals older than three

years, irrespective of ploidy. Therefore, the majority of diploid males preferred areas of low depth and flow, as well as the presence of cover and coarse substratum such as gravel and stones. Triploid individuals had a marked preference for deeper areas and greater current velocities than their diploid male counterparts.

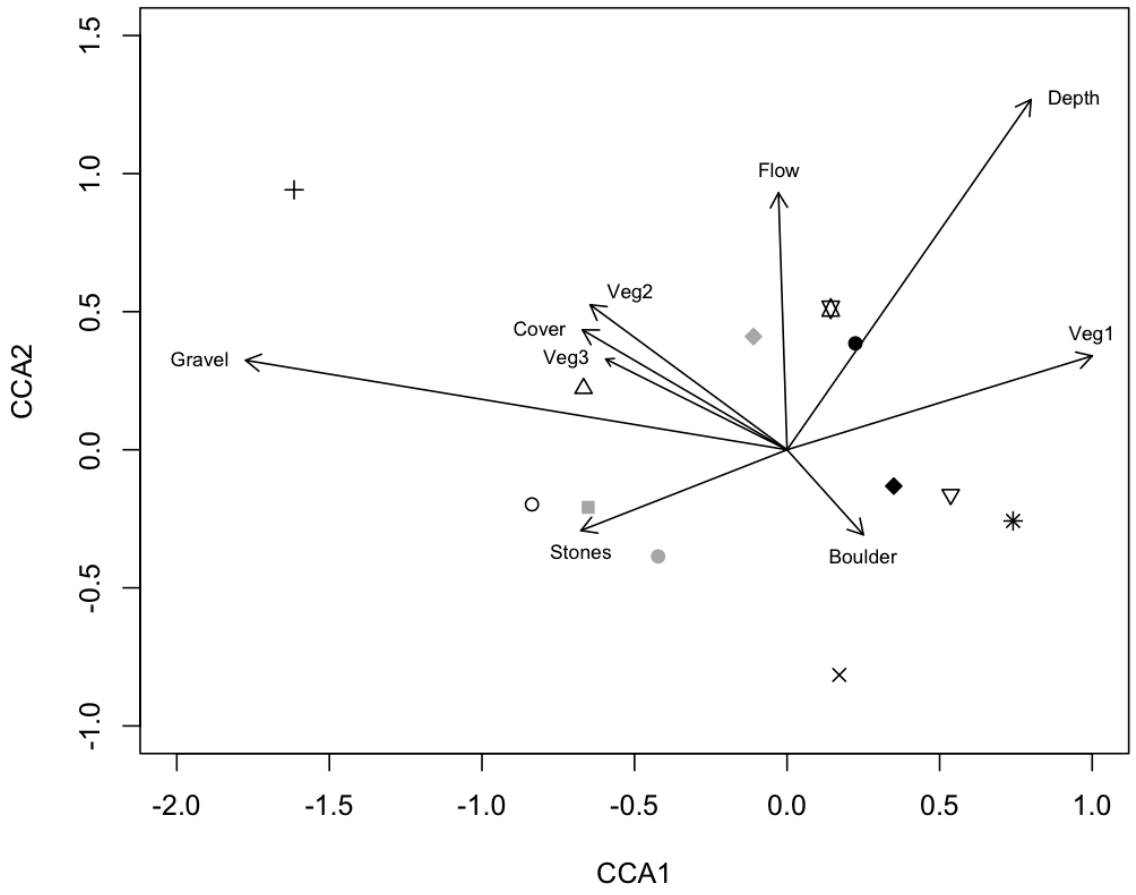


Figure 5.9: Representation of the CCA showing influence of habitat variables on the presence of species. ● Diploid male <1 year ○ Diploid male 1-2 years ● Diploid male ≥ 3 years ◆ Triploid 1-2 years ◆ Triploid ≥ 3 ■ Unidentified *S. alburnoides* ┼ *S. pyrenaicus* △ *L. gibbosus* ✡ *Barbus* sp. × *G. holbrookii* ▽ *B. microcephalus* ✳ *B. steindachneri.* Veg1 is emergent vegetation, Veg2 represents floating vegetation and Veg3 represents submerged vegetation.

5.3.4 Diet and habitat correlations

CCA was performed to determine associations between food items consumed and the characteristics of the microhabitat that the fish were sampled from. The first axis was most highly correlated with depth, showing that surface dwelling and non-aquatic

organisms were consumed by fish found in shallow waters and a high degree of tree cover (Figure 5.10). The second axis was mostly influenced by the proportion of stones found in the substrate. The presence and quantity of vegetation, either emergent, floating or submerged, played a role on both axes, influencing prey preference on a large number of organisms, including prey such as Simuliidae and Chironomid larvae that were found in large quantities in all ploidy groups. Terrestrial insects were also consumed in microhabitat areas with little or no vegetation.

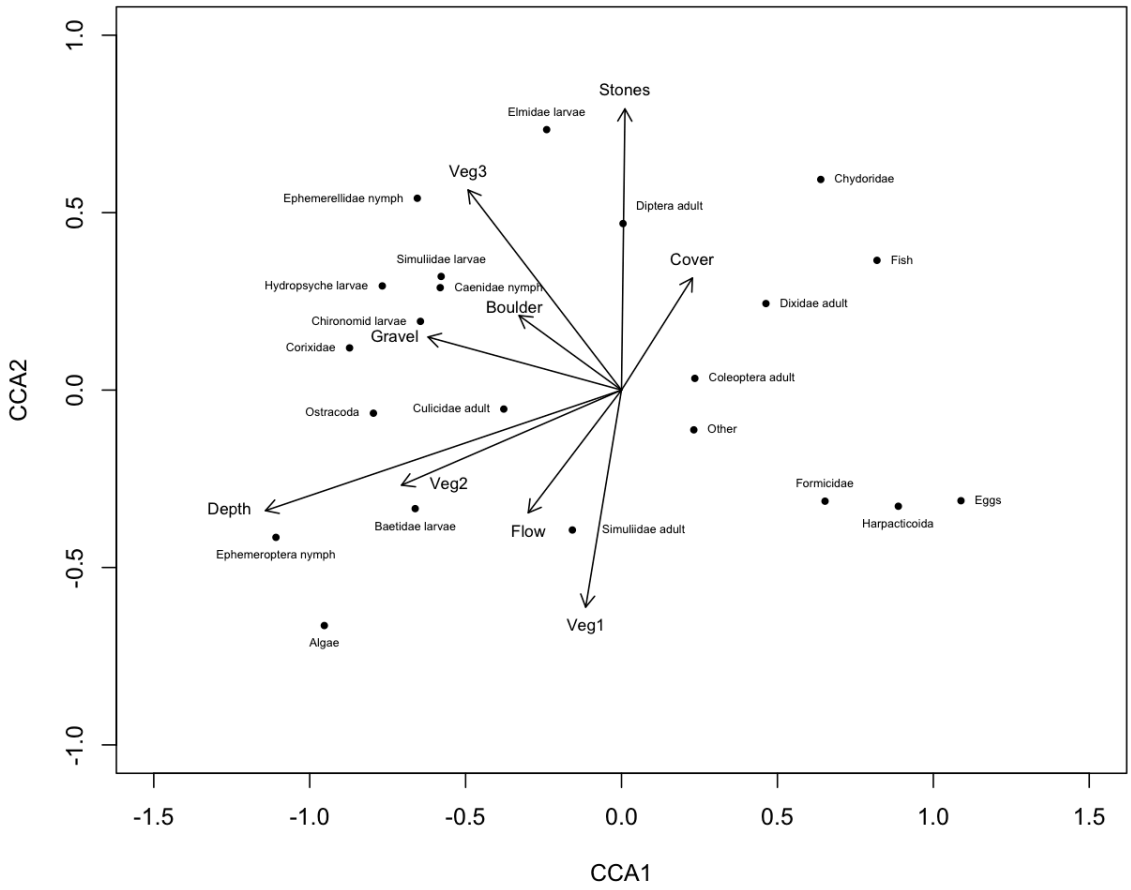


Figure 5.10: Representation of the CCA performed on the diet of *S. alburnoides*, using habitat factors to explain the variation in the sample. Veg1 is emergent vegetation, Veg2 represents floating vegetation and Veg3 represents submerged vegetation.

5.4 Discussion

Ecological and growth data of *S. alburnoides* showed that differences between the two ploidy levels studied, diploid males and triploids, exist beyond those identified in their morphologies (Chapters 3 and 4).

Previous studies concluded that growth rates between ploidy levels differed significantly (Collares-Pereira et al., 2007; Ribeiro et al., 2003), therefore it was considered necessary to continue making the distinction between ploidy levels when analysing *S. alburnoides*, as inherent size differences existed between diploid males and other ploidy levels, with diploid males being of much smaller size, and only growing to a maximum of 90 mm compared to 130 mm for triploids (Collares-Pereira et al., 2007). The results from aging the fish also provide another reason to distinguish between ploidy levels, as diploid males do not live as long as other ploidy levels and no individuals of more than four years were found. This finding was also confirmed by previous studies, such as by Ribeiro et al. (2003), where the oldest diploid males found were three years old. The reasons why diploid males might tend to have a shorter lifespan are numerous. Genetic differences are implicit, the genome of diploid males being different to its hybrid diploid and triploid counterparts. In *S. alburnoides,* diploid individuals have smaller genome sizes than triploids (Prospero and Collares-Pereira, 2000) and the quantity of DNA in a cell can significantly influence the longevity of a number of animal orders (Griffith et al., 2003). Another reason for differences in lifespan may be related to feeding behaviour, especially during dry summers, where there is increased competition for food and alterations in the community structure of fish present (Collares-Pereira et al., 2000; Pires et al., 1999). Triploid individuals have a much broader diet, which may increase chances of survival in conditions of drought

(Gomes-Ferreira et al., 2005). The small size of diploid males could be disadvantageous in intraspecific competition for food, but it may also increase the risk of predation, notably from introduced species such as the largemouth bass *Micropterus salmoides* (Godinho et al., 1997). Risk of interspecific predation may also occur in periods when food is scarce, as large *S. alburnoides* consume fish, including diploid male *S. alburnoides* (details in this chapter, section 5.3.2).

Significant differences in size between ploidy groups become apparent beyond the first year of growth, irrespective of the method with which average lengths-at-age were calculated. This reinforces the idea that growth for ploidy levels should be viewed separately, as they lead different life histories that have a significant effect on overall growth. These results are comparable to those found by Ribeiro et al. (2003), who found differences in growth between diploid males and other ploidy levels. They also demonstrated that growth slowed for triploids after the second year, hypothesised to be due to increased reproductive effort. A similar trend was found in the populations studied here, in which growth slowed later on in life.

Stomach contents analysis was used to examine the diversity of food intake in the species, which may represent diversity due to definable morphological and genetic features or may represent diversity due to factors that influence prey availability such as season or riverine habitat characteristics (Pires, Cowx, and Coelho, 2000a; Esteves et al., 2000). In the study performed here, it is possible that the diet diversity found in *S. alburnoides* is due to seasonal differences in prey availability. Although confirmation of this is not possible, due to a lack of macroinvertebrate sampling in this thesis, the items that are the most commonly consumed are congruent with a previous report of prey availability in the same study area (Pires, Cowx, and Coelho, 2000a).

In addition to variation between seasons, *S. alburnoides* dietary data displayed a measure of diversity between ploidy groups within a season. The overlap between diets of both ploidy levels showed that around half of the diet was constituted by the same items, the other half representing prey consumed only by one group. One of the most notable differences in dietary composition was the consumption of fish by triploid individuals in October when the overall diversity and abundance of food was no doubt lower (Pires, Cowx, and Coelho, 2000a). This suggests that when food is scarce, triploids, due to their larger size and more aggressive nature (Gomes-Ferreira et al., 2005), are able to consume a wider variety of prey.

Results of habitat preferences of *S. alburnoides* have not only shown segregation based on ploidy, but also one based on size and age, with older diploid male individuals sharing some habitat characteristics with triploid individuals. This mirrors findings from *S. pyrenaicus* that display varied dietary and habitat preferences relating to size (Magalhaes, 1993a). In many respects, such as morphology, dietary composition and habitat selection, triploid *S. alburnoides* could be associated with their maternal ancestor *S. pyrenaicus* and diploid males with *A. hispanica*, presumed to be similar to the paternal ancestor of the species complex (Collares-Pereira and Coelho, 2010). Niche segregation may have occurred between the ancestors of *S. alburnoides* and such segregations may have been transferred to the species complex, influencing niche selection. This could help to determine whether differences in ecological niches are due to behavioural preferences inherent to the ploidy level or determined by morphological features, including size.

Triploids and diploid males appeared to occupy different microhabitats, which were also occupied by other species. A review of the literature of the preferred habitats of

species found in the Murtega suggested that triploids used similar habitat characteristics to those of several species of the genus *Barbus*, especially fast flowing and deep water conditions, which was also confirmed by the CCA analysis (Table 5.9). Diploid males occupied similar habitats to those of introduced species *L. gibbosus* and *G. holbrookii,* which could be a major source of competition regarding dietary composition. *L. gibbosus* has a very wide dietary range, feeding extensively on any prey items available (Blanco-Garrido et al., 2009), which could lead to extensive competition with native species such as *S. alburnoides,* dependent on food availability (Garcia-Berthou and Moreno-Amich, 2000; Fox et al., 2007). The overlap in microhabitat means that diploid males are more likely to be influenced by *L. gibbosus*, however it does consume similar prey to that preferred by triploids as well (Table 5.10).

Table 5.9: Habitat preferences for fish species found in the community assemblage of the Murtega river.

Species	Depth	Flow	Substrate	Other characteristics	Species origin	Source
Squalius alburnoides diploid male	Shallow	Low	Gravel and stones	Presence of aquatic vegetation	Native/ Endemic to Iberian Peninsula	See Section 5.3.3; Collares-Pereira et al. (2007)
Squalius alburnoides triploids	Deep	High	Stones and boulder		Native/ Endemic to Iberian Peninsula	See Section 5.3.3; Collares-Pereira et al. (2007)
Squalius pyrenaicus	Rivers with shallow depths	Low - medium	Stones and boulder	Presence of aquatic vegetation	Native/ Endemic to Iberian Peninsula	Collares-Pereira et al. (2007), Magalhaes (1993)
Lepomis gibbosus	Shallow	Low	Sand and mud	Murky and brackish waters	Introduced exotic	Collares-Pereira et al. (2007), Maitland and Campbell (1992)
Barbus microcephalus	Deep	High	Gravel and stones		Native/ Endemic to Iberian Peninsula	Collares-Pereira et al. (2007), Cabral et al. (2005)
Barbus steindachneri	Deep	High	Boulder and bedrock	Smaller individuals prefer areas with high vegetation abundance	Native/ Endemic to Iberian Peninsula	Collares-Pereira et al. (2007), Cabral et al. (2005)
Gambusia holbrookii	Shallow	Low	All types of substrate	Lots of aquatic vegetation	Introduced exotic	Collares-Pereira et al. (2007)

Table 5.10: Dietary preferences for fish species found in the community assemblage of the Murtega river.

Species	Typical diet consumption	Source
Squalius alburnoides diploid male	Insects, including many dipteran adults	See Section 5.3.2; Collares-Pereira et al. (2007)
Squalius alburnoides triploids	Insect larvae, occasionally opportunist	See Section 5.3.2; Collares-Pereira et al. (2007)
Squalius pyrenaicus	Dipteran larvae, Ephemeroptera	Collares-Pereira et al. (2007)
Lepomis gibbosus	Dipteran larvae, occasionally opportunist	Collares-Pereira et al. (2007), Godinho et al. (1997)
Barbus microcephalus	Generalist feeder, algae and macroinvertebrates	Collares-Pereira et al. (2007), Pires et al. (2001)
Barbus steindachneri	Algae and crustaceans	Collares-Pereira et al. (2007), Pires et al. (2001)
Gambusia holbrookii	Crustaceans (October to May), Dipteran larvae (June to September)	Collares-Pereira et al. (2007)

As with triploid *S. alburnoides*, *S. pyrenaicus* can be found in many different microhabitats but prefers regions of moderate water flow and coarse substrate (Pires, Cowx, and Coelho, 2000b; Morán-López et al., 2006). Although microhabitat characteristics will affect prey species present, it is possible to compare dietary preferences between species. Little data has been achieved on the dietary patterns of *S. pyrenaicus*, but that available does bear some resemblance to dietary preferences found in *S. alburnoides* triploids, especially for the autumn and winter months, when prey diversity present in streams is reduced (Pires, Cowx, and Coelho, 2000a). Both *S. alburnoides* triploids collected in October and *S. pyrenaicus* sampled in the same

season showed that although overall dietary diversity decreased when compared to summer and spring, the consumption of species less abundant in the gut increased, indicating generalist and opportunistic feeding behaviour (Magalhaes, 1993a). This opportunistic nature is most readily seen in *S. alburnoides* by cannibalism of smaller fish, as found in the research presented here.

In many aspects, the habitat preferences of diploid male *S. alburnoides* are similar to those of *A. hispanica*, especially relating to the preference for low flow and shallow depth (Collares-Pereira et al., 2000; Ribeiro et al., 2000; Blanco-Garrido et al., 2009). Body size is likely an influencing factor for these preferences, as older and larger diploid males were also found in areas of greater depth. However, due to the highly endangered nature of *A. hispanica* (Collares-Pereira et al., 1999), there have not been any opportunities to carry out detailed dietary studies. However, it can be assumed, by examining habitat preferences, that some similarities will be found. The data presented here showed that dipteran adults are more likely to be consumed by individuals found in shallow depths, many of which are diploid males. The preference of *A. hispanica* to inhabit such habitats leads to the conclusion that dietary preferences between these species may overlap significantly. In addition, both the morphology of diploid male *S. alburnoides* and *A. hispanica* is adapted to consume terrestrial insects from the surface, due to the upturned position of the mouth (Karpouzi and Stergiou, 2003; Collares-Pereira et al., 2007; see Chapter 4.3.2).

Other polyploid fish species display similar patterns to those found for *S. alburnoides*, where differences between ploidy levels or between the hybrid and their parental species are found at a number of ecological levels. The hybrid *Phoxinus eos-neogaeus* lives in sympatry with both of its parental ancestors and presents mostly clonally-

reproducing diploids and some triploids that arise due to further hybridisation of the hybrid with one of the parental ancestors (Doeringsfeld et al., 2004). *Phoxinus eos-neogaeus* also displays habitat segregation between ploidy levels, similar to *S. alburnoides,* with diploids occupying a much broader niche than triploids, which have similar habitat preferences to the parental species (Doeringsfeld et al., 2004). The dietary preferences of polyploid species tend to indicate that differences are found between the hybrid species and their sympatric parental ancestors, reducing interspecies competition, such as for fish of the genus *Poeciliopsis* (Weeks et al., 1992; Gray and Weeks, 2001). Differences in the growth of *S. alburnoides,* however, do not corroborate the results from growth patterns of the polyploid spined-loach complex (genus *Cobitis*), where no differences were found between the available ploidy levels (Ritterbusch and Bohlen, 2000).

Previous work on the species, (e.g. Martins et al., 1998; Gomes-Ferreira et al., 2005) suggest that for the major portion of the year there are differences in resource use by different ploidy forms of *S. alburnoides*, however, results here show that interactions between ploidy levels also frequently occur outside of the reproductive period, including those relating to habitat selection. A number of life history differences also exist between diploid males and triploids, visible in growth rates and longevity. These differences may be affected by a number of variables that may vary between populations of the same species, such as resource use in early life stages. With the data available it was not possible to determine if growth and ecological differences were also present between other ploidy forms, such as triploids and diploid females. However, the literature suggests that no differences would be found, as these groups have similar behavioural and dietary preferences (Ribeiro et al., 2003). By necessity,

both ploidy levels of *S. alburnoides* studied (diploid male and triploid) are well adapted to cope with variability in prey and habitat availability of the Guadiana river basin, influenced by the ever changing hydrological regime. This does not however mean that their future is assured, as the pressure exerted by the increase in dams and water abstraction facilities may reduce habitat availability in summer even further and increase intra-species competition. This chapter has demonstrated that triploid *S. alburnoides* occupy a wide range of habitats within a river, which has been linked to improved chances of survival and adaptation in other polyploid species (Gray and Weeks, 2001). However, the reproductive mechanisms present within the species complex currently indicate that triploids rely on the presence of diploid males to activate fertilisation (Sousa-Santos, Collares-Pereira, and Almada, 2006c), especially if other potential sperm donors, such as *S. pyrenaicus*, are not present, as is the case for this study.

Chapter 6: General discussion

6.1 Morphological conclusions

This thesis has mostly focused on morphological characteristics of *S. alburnoides* and attempted to relate these to genetic composition, habitat and dietary preferences. A number of more specific questions were posed in individual chapters, with the aim of improving knowledge about the *S. alburnoides* species complex, not only for conservation purposes but to aid future studies in terms of ploidy identification within the species complex. Both Chapters 3 and 4 successfully demonstrated that differences in genetic composition of *S. alburnoides* individuals could be seen in scale, head and body morphology of individuals from the same river basin. Using scale morphology of fish of known ploidy, a prediction function was created that could identify individuals of unknown ploidy within the Guadiana river basin. While scales of three distinct ploidy groups were available, the prediction function was successful (with sufficient confidence) in predicting the ploidy of diploid male and triploid individuals. Diploid females were not identified to the required level of precision, in part due to differences in morphology between diploid female individuals and triploid individuals that were not as great as between non-hybrid diploid males and all other ploidy groups. The number of diploid females available to provide morphological scale data was also lower than the number of triploids samples available, meaning that less information on diploid female shape could be used to refine the prediction function.

In addition to identifying differences in ploidy in the Guadiana river, scale morphology was also used to identify differences between fish in the Tejo and the Sado river systems. Data from both of these rivers were able to provide some distinctions

between ploidy levels, however neither were capable of differentiating tetraploid individuals from other ploidy levels due to low numbers of tetraploid individuals available for morphological analysis. The low numbers of fish available from the Tejo and Sado rivers did not provide sufficient accuracy for the creation of prediction functions as for the Guadiana river system. However, the results found are encouraging, demonstrating that if more morphological data were to be collected, an accurate prediction function for both rivers could be created, with the ability to distinguish between most, if not all, ploidy levels present in the river basin.

The morphometric technique developed successfully allowed differentiation within the species and also revealed morphological differences not previously identified. The geometric methods involving body shape succeeded in morphological identification of *S. alburnoides* in the Guadiana river basin of Portugal where all previous ploidy identifications involved the use of flow cytometry or the counting of meristic characteristics (Prospero and Collares-Pereira, 2000; Collares-Pereira, 1984; Ribeiro et al., 2003). The work performed here confirmed findings that differentiation could be made (Cunha et al., 2010) using such techniques and added information about the morphology of non-hybrid diploid males. The knowledge gained here has shown that morphological differences between ploidy levels are not confined to a specific area of the body, such as the head, but are distributed across the whole of the body, such as in the caudal peduncle, body depth and position of the mouth. A comparison of morphological characteristics between river systems was performed, demonstrating that differences do exist, although the exact reasons for this difference are still not clear, as this may be an influence of ecological adaptation or due to genetic divergence within the species. The creation of these techniques has provided a novel framework

for the identification of *S. alburnoides* ploidy levels using geometric morphometrics, involving the use of scale, head and body morphology. The data has however shown that seasonal differences are also displayed in the morphology of the species, especially with regards to female individuals, and as such seasonality must be taken into account when using body morphology to identify *S. alburnoides,* whether traditional measurement based techniques or geometric morphometric methods be used.

The data available from the three rivers studies provided results that demonstrated that morphological differences not only existed between ploidy levels in a system, but also between the riverine systems. Differences in the average shape of each ploidy group could be seen between rivers, indicating that the varying population compositions between rivers were affecting morphological factors for each ploidy group. These differences in shape could also be explained by differing environmental conditions between rivers.

6.2. Ecological conclusions

Chapter 5 aimed in part to corroborate previous research that indicated that growth patterns between ploidy groups varied within the species, with diploid males having smaller length-at-age than all other ploidy groups. Due to only two groups within the species, diploid males and triploids, having been successfully identified in Chapter 3, comparisons in Chapter 5 could only be made between these two groups. Results successfully demonstrated that diploid males and triploids have different growth patterns, with triploids growing to longer lengths and living to greater ages than diploid males. The research showed that these differences usually became present after the first year of growth and that lengths before that were similar, irrespective of ploidy

group. Due to the sample sizes from cohorts being small, it is possible that some growth patterns, especially for later years, are not fully represented in the data available here. However, the growth patterns calculated within this thesis were performed on all individuals available, irrespective of cohort, and provide an insight into expected growth patterns of *S. alburnoides*.

The segregation of ploidy groups using ecological variables was also examined in Chapter 5. Between the two ploidy groups available, the results indicated that some dietary segregation occurred, identifying prey that were consumed exclusively by one ploidy or the other. Prey items consumed by diploid males included dipteran adults, which are associated with shallow habitat areas, generally the preferred habitat of younger diploid males. The consumption of these items is also related to morphological adaptations found amongst individuals of that ploidy group, such as the more upturned position of the mouth when compared to triploid individuals. Such adaptions have been found in other species of fish, from a wide range of genera (Czerwinski et al., 2008; Winemiller, 1991) and can be related to feeding at specific trophic levels (Karpouzi and Stergiou, 2003). Individuals of triploid ploidy were shown to be more generalist feeders, which could also be related to the morphology of the mouth, as the prey items consumed could be found in all levels of the water column with preferences for a variety of substrates (Longo S et al., 2010; Bae et al., 2011). Habitat segregation within *S. alburnoides* was successfully demonstrated and was found to be associated with size as well as ploidy. Numerous species of freshwater fish have been shown to modify their habitat preferences between juvenile and adult stages, such as the closely related *Squalius pyrenaicus* (Magalhaes, 1993a) and the polyploid hybrid *Phoxinus eos-neogaeus* (Schlosser, 1982). Older diploid male

individuals were more likely to be found sharing a habitat with triploid individuals than young diploid male individuals were.

The results achieved here, from both habitat and dietary aspects, confirmed previously published work, with *S. alburnoides* displaying a degree of separation for these ecological factors between non-hybrid diploid males and other groups present (Martins et al., 1998; Gomes-Ferreira et al., 2005). A certain amount of seasonal differences in diet consumption was found. The seasonal diet variation hinted at the possibility of further dietary variation by season. Unfortunately, logistic constraints, relating to time and availability of collaborators, made sampling on more than two occasions impossible. The highly variable nature of the river system and the extreme weather conditions present in winter and summer (Morais et al., 2009) would have no doubt also displayed an impact on the dietary preferences and intra-species competition. This would be especially true in summer, where interspecies competition is known to increase, due to limitation of resources caused by the drought (Morán-López et al., 2005).

The differences in habitat and diet within the Guadiana river basin are mostly also found in other river systems, such as the Tejo and Sado drainages. Sampling of these other rivers could have revealed alternative resource use for *S. alburnoides* individuals that may be highly dependant on the resources present within the system. This thesis did not focus on the comparison of how ploidy levels from different river basins selected their preferred habitats and whether this had an influence on prey choice. In addition, other rivers and stretches of the Guadiana have exotic introduced predators, such as *Micropterus salmoides*, that may play a role in habitat choice for predator avoidance, as demonstrated in numerous species of freshwater fish (Huntingford and

Wright, 1989; C. Brown, 2003; Reebs, 1999). No *M. salmoides* were found in the area of the Guadiana sampled, suggesting a population mostly undisturbed by introduced exotics.

6.3 Combining morphological and ecological data

The *S. alburnoides* species complex is part of a small group of vertebrate animals that present polyploidy due to hybridisation, which is often accompanied by modified reproductive tactics and unisexual lineages (Dawley, 1989; Vrijenhoek and Dawley, 1989). It has been proposed that these complexes are evolutionary dead-ends due to their reproductive methods and interspecies competition with their parental species for limited ecological resources (AA Echelle and AF Echelle, 1997; Vrijenhoek, 1989). Other research suggests that this is not the case, with polyploid complexes living separate from their parental ancestors (Adams et al., 2003) and evidence to show that polyploids can occupy a wider range of niches than their ancestors (Gray and Weeks, 2001; Weeks et al., 1992), potentially gaining an advantage over these. This appears to be the case with *S. alburnoides*, which not only shows an ability to occupy a number of habitats and thrive, but also shows some habitat segregation within the species, thus potentially reducing intraspecific competition. Extensive evidence of niche partitioning has been demonstrated for fish of the genus *Poeciliopsis* (Weeks et al., 1992; Gray and Weeks, 2001) and those of the *P. eos-neogaeus* complex (Schlosser et al., 1998; Doeringsfeld et al., 2004). The differences in resource use of *S. alburnoides* are also evident in morphology, where differences in a number of morphological characteristics are visible, some of which, such as position of the mouth, can be correlated directly to ecological preferences. The morphological differences were expected, as work on a similar polyploid complex of the genus *Cobitis* shows a number of morphological

variations on fin rays, spines and scales (Kotusz, 2008). One aspect of morphology that the geometric morphometrics conducted in this thesis did not take into account was the inherent difference in size between the two ploidy levels available for study. Analyses of age and growth succeeded in confirming previous findings that indicated that non-hybrid diploid males systematically had a smaller length at age than other ploidy groups found within the same river system (Ribeiro et al., 2003).

The *S. alburnoides* species complex displays a lot of evidence of on-going evolution, which may play a vital understanding in our overall view of evolutionary processes. The study here has shown that different ploidy levels, with their differing morphological features, have varied ecological traits, which can often be directly related to their morphological features. An example of this would be for the apparent preference for adult diptera by non-hybrid AA males, that display a superior mouth position. The abundance of *S. alburnoides* in a variety of microhabitats confirms research that postulates that polyploid hybrids are adaptable to a wider range of conditions than their ancestors. In addition to this, *S. alburnoides* individuals are capable of producing a variety of viable offspring without *S. pyrenaicus*, the maternal ancestor, ensuring their continued survival should *S. pyrenaicus* become extinct. *S. alburnoides* overall shows great potential for adaptability, not only at a reproductive level, but at an ecological level. Both morphology and dietary consumption indicate that this polyploid species is capable of successfully inhabiting a variety of habitats throughout its range. These conclusions can also be applied to a number of other polyploid species, showing that these species are an important step in fish evolutionary processes and should not be considered to have no evolutionary merit.

6.4 Limitations and recommendations for future work

Future work on *S. alburnoides* scale morphology should look to improve sample sizes, especially diploid females, and add to the existing database of scales, with the hope of further improving the prediction function and creating a clear distinction between diploid females and other ploidy groups. Accurate prediction functions can be obtained with as few as 30 fish for each ploidy level, however a sample size of 75 for each would allow greater discrimination between ploidy levels with similar morphology.

One of the main limitations of the present work was that the only prediction function created used scale morphology from fish caught in the Guadiana river. Currently, fish caught in any other river basin may only be identified using more traditional methods rather than morphology. Despite this, the framework laid out here has shown that the creation of prediction functions is possible and that they have a high degree of accuracy. If a thorough comparison of scale and body shape between rivers were to be carried out, it would not only provide an extremely useful tool for identification of *S. alburnoides* ploidy levels but may also shed light on how the different populations have evolved since their original hybridisation event (Sousa-Santos, Collares-Pereira, and Almada, 2007b). In comparisons between the Tejo river and the northern basins of the Douro and the Mondego, analysis of morphology may provide insight on how these populations have diverged since the original colonisation of the northern basins, and what influence the sperm donor *S. carolitertii* has had on the morphology (Sousa-Santos, Collares-Pereira, and Almada, 2007b).

The differences in population composition between river basins also highlights the problem that in the scales available for the Tejo river basin, there were no examples from non-hybrid diploid males, a form that is found in this river (Gromicho and

Collares-Pereira, 2004). In both the Sado and Tejo drainages, examples of scales from tetraploid individuals were discarded from the final analysis, due to insufficient sample sizes. This missing information means that it is not possible to provide a complete picture of the morphological variation that exists within a river basin. On the same note, although female diploids were present in the available sample, they were not sufficient in number to accurately predict individuals. This may have also led to some misidentification, whereby diploid females were classified as females. It was deemed that this would not cause a problem for the subsequent ecological analyses as previous work showed little habitat and dietary differentiation between diploid females and triploid individuals (Martins et al., 1998).

Further work should be done on all rivers where *S. alburnoides* is present, highlighting existing differences between rivers and providing explanations as to which environmental variables could cause these differences. It may also be possible to discover similarities between rivers that are related to the evolutionary history of the species, identifying populations that remain closely related and confirming proposed hybridisation events in the river systems studied (Sousa-Santos, Collares-Pereira, and Almada, 2007b).

Despite the success achieved using body morphology as an indicator of ploidy, a prediction function was not created. This was due to the fact that the ploidy of individuals used for body shape analysis was inferred using the scale shape prediction function, rather than through traditional methods using flow cytometry (Sousa-Santos, Collares-Pereira, and Almada, 2007a). Such a function should be created by gathering morphological information from individuals genotyped by flow cytometry in future work. As with scale shape, at least 30 individuals for each ploidy level, organised by

river basin, should allow the creation of a detailed prediction function relating to body and head morphology. This prediction function could be further improved by gathering scale shape information and combining all morphological aspects. By verifying morphology on two separate features of an individual fish, confidence in the accuracy of the prediction function would increase.

This thesis has added considerable knowledge to the ecology of *S. alburnoides*, specifically relating to their differences according to ploidy level. Previous research on this was not accurate enough in the description of ecological data collected. This thesis has attempted to remedy this, for example, by identifying dietary items consumed to the lowest taxonomic level possible. Future work should continue this, with particular focus on habitat characteristics, such as accurately measuring depth and flow.

This should also be applied to dietary analyses. Due to logistical reasons this study was not able to sampled the macroinvertebrate fauna present in the rivers when sampling occurred. Detailed information on macroinvertebrate presence and its seasonal variability would allow future studies to draw detailed conclusions about dietary preferences of *S. alburnoides* and whether competition or resource partitioning exists between ploidy levels. This dietary information could also be applied to determine whether differences exist in diet at different life stages. Molecular studies can be used to identify diet composition in larval stages.

The recommendations established for *S. alburnoides* can also be applied to a number of other polyploid species as well as non-polyploid species of fish. The morphometric techniques developed within this thesis can be used on a number of species, allowing the identification of fish within and between species. The aim is that this tool can be

used by field operatives, provided a digital camera and laptop are available. This would allow for sample sizes to be more evenly distributed, by ensuring researchers know beforehand the characteristics of the fish being sampled. It is evident that this technique is not for sole use in the field, and that where possible, more detailed confirmation, continuing using morphometric techniques, should be performed in the laboratory.

In summary, the aims of the thesis were successfully fulfilled, adding knowledge to the field of study of *S. alburnoides* and to the identification of individuals within a polyploid species using geometric morphometric methods. The methodology used to identify ploidy levels within a species using scale morphology had never been performed previously. Geometric morphometric methods are very reliant on finding points that can be repeatably identified throughout the data sample provided. The addition of semilandmarks to describe scale shape has allowed the development of a method that can be used for scales that do not display clearly defined lateral tips (Ibañez et al., 2007), as was the case with scales of *S. alburnoides*. The development of this method, and its success on *S. alburnoides*, has allowed the creation of an identification method that is rapid, easily performed in the field, relatively inexpensive and does not harm the fish in any discernible way.

References

Abbas,K., Li,M.Y., Wang,W.M. and Zhou,X.Y. (2009) First record of the natural occurrence of hexaploid loach Misgurnus anguillicaudatus in Hubei Province, China. J FIsh Biol, 75, 435–441.

Adams,M., Foster,R., Hutchinson,M.N., Hutchinson,R.G. and Donnellan,S.C. (2003) The australian scincid lizard Menetia greyii: A new instance of widespread vertebrate parthenogenesis. Evolution, 57, 2619–2627.

Aguirre,W.E. (2009) Microgeographical diversification of threespine stickleback: body shape-habitat correlations in a small, ecologically diverse Alaskan drainage. Biol J Linn Soc, 98, 139–151.

Alarcón,C., Cubillos,L.A. and Acuña,E. (2011) Length-based growth, maturity and natural mortality of the cockfish Callorhinchus callorhynchus (Linnaeus, 1758) off Coquimbo, Chile. Environ Biol Fish, 92, 65–78.

Alves,M. and Coelho,M.M (2001) Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: a genetic review. Genetica.

Alves,M., Coelho,M.M. and Collares-Pereira,M.J. (1998) Diversity in the reproductive modes of females of the Rutilus alburnoides complex (Teleostei, Cyprinidae): A way to avoid the genetic constraints of uniparentalism. Mol Biol Evol, 15, 1233–1242.

Alves,M., Coelho,M.M., Collares-Pereira,M.J. and Dowling,T.E. (1997) Maternal ancestry of the Rutilus alburnoides complex (Teleostei, Cyprinidae) as determined by analysis of cytochrome b sequences. Evolution, 51, 1584–1592.

Alves,M., Coelho,M.M., Prospero,M.I. and Collares-Pereira,M.J. (1999) Production of fertile unreduced sperm by hybrid males of the Rutilus alburnoides complex (Teleostei, Cyprinidae): An alternative route to genome tetraploidization in unisexuals. Genetics, 151, 277–283.

Alves,M., Gromicho,M., Collares-Pereira,M.J., Crespo-Lopez,M.E. and Coelho,M.M. (2004) Simultaneous production of triploid and haploid eggs by triploid Squalius alburnoides (Teleostei : Cyprinidae). J Exp Zool Part A, 301A, 552–558.

Alves,M., Collares-Pereira,M.J., Dowling,T.E. and Coelho,M.M. (2002) The genetics of maintenance of an all-male lineage in the Squalius alburnoides complex. J FIsh Biol, 60, 649–662.

Bae,M.-J., Kwon,Y., Hwang,S.-J., Chon,T.-S., Yang,H.-J., Kwak,I.-S., Park,J.-H., Ham,S.-A. and Park,Y.-S. (2011) Relationships between three major stream assemblages and their environmental factors in multiple spatial scales. Ann Limnol-Int J Lim, 47, 91–105.

Bagenal,T.B. (1978) Methods for assessment of fish production in fresh waters 3rd ed. Blackwell, Oxford.

Beamish,F.W.H., Sa-ardrit,P. and Tongnunui,S. (2006) Habitat characteristics of the cyprinidae in small rivers in central thailand. Environ Biol Fish, 76, 237–253.

Bertalanffy,von,L. (1938) A Quantitative Theory Of Organic Growth (Inquiries On Growth Laws II). Human biology, 10, 181–213.

Blanco-Garrido,F., Clavero,M. and Prenda,J. (2009) Jarabugo (Anaecypris hispanica) and freshwater blenny (Salaria fluviatilis): habitat preferences and relationship with exotic fish species in the middle Guadiana basin. Limnetica, 28, 139–148.

Bolger,T., Bracken,J.J. and Daoud,H.A. (1990) The Feeding Relationships of Brown Trout, Minnow and 3-Spined Stickleback in an Upland Reservoir System. Hydrobiologia, 208, 169–185.

Bookstein,F. (1997) Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Med Image Anal.

Braak,ter,C. and Verdonschot,P. (1995) Canonical Correspondence-Analysis and Related Multivariate Methods in Aquatic Ecology. Aquat Sci, 57, 255–289.

Britton,J.R. (2007) Reference data for evaluating the growth of common riverine fishes in the UK. J Appl Ichthyol, 23, 555–560.

Brosse,S., Gabas,S., Cever,K. and Lek,S. (1999) Comparison of two point abundance sampling methods to assess young roach (Rutilus rutilus L.) microhabitat in the littoral zone of lake Pareloup (France). Ann Limnol-Int J Lim, 35, 199–204.

Brown,C. (2003) Habitat-predator association and avoidance in rainbowfish (Melanotaenia spp.). Ecol Freshw Fish, 12, 118–126.

Burns,J.G., Di Nardo,P. and Rodd,F.H. (2009) The role of predation in variation in body shape in guppies Poecilia reticulata: a comparison of field and common garden phenotypes. J FIsh Biol, 75, 1144–1157.

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

Cadrin,S. (2000) Advances in morphometric identification of fishery stocks. Rev Fish Biol Fisher, 10, 91–112.

Cardinale,M., Doering-Arjes,P., Kastowsky,M., Mosegaard,H. (2004) Effects of sex, stock, and environment on the shape of known-age Atlantic cod (Gadus morhua) otoliths. Can J Fish Aquat Sci, 61, 158–167.

Carlson,R.L. and Wainwright,P.C. (2010) The ecological morphology of darter fishes (Percidae: Etheostomatinae). Biol J Linn Soc, 100, 30–45.

Carroll,S.P., Hendry,A.P., Reznick,D.N. and Fox,C.W. (2007) Evolution on ecological time-scales. Funct Ecol, 21, 387–393.

Cavalcanti,M., Monteiro,L.R. and Lopes,P.R.D. (1999) Landmark-based morphometric analysis in selected species of serranid fishes (Perciformes : Teleostei). Zool Stud, 38, 287–294.

Chícharo,M.A., Chícharo,L. and Morais,P. (2006) Inter-annual differences of ichthyofauna structure of the Guadiana estuary and adjacent coastal area (SE Portugal/ SW Spain): Before and after Alqueva dam construction. In, Estuarine Coastal and Shelf Science. pp. 39–51.

Christiansen,D.G. and Reyer,H.-U. (2009) From Clonal to Sexual Hybrids: Genetic Recombination via Triploids in All-Hybrid Populations of Water Frogs. Evolution, 63, 1754–1768.

Collares-Pereira,M.J. (1984) The Rutilus alburnoides (Steindachner, 1866) complex (Pisces, Cyprinidae). I. Biometrical analysis of some Portuguese populations. Arquivos do Museu Bocage (Ser. A), (Ser. A), 111–143.

Collares-Pereira,M.J. and Coelho,M. (2010) Reconfirming the hybrid origin and generic status of the Iberian cyprinid complex Squalius alburnoides. J FIsh Biol, 76, 707–715.

Collares-Pereira,M.J. and Moreira da Costa,L. (1999) Intraspecific and inter-specific genome size variation in Iberian Cyprinidae and the problem of diploidy and polyploidy, with review of genome sizes within the family. Folia Zool, 48, 61–76.

Collares-Pereira,M.J., Filipe,A.F. and Moreira da Costa,L. (2007) Os peixes do Guadiana que futuro? 1st ed. Edicoes Cosmos.

Collares-Pereira,M.J., Cowx,I.G., Rodrigues,J.A., Rogado,L. and Moreira da Costa,L. (1999) The status of Anaecypris hispanica in Portugal: Problems of conserving a highly endangered Iberian fish. Biol Conserv, 88, 207–212.

Collares-Pereira,M.J., Cowx,I.G., Ribeiro,F., Rodrigues,J.A. and Rodago,L. (2000) Threats imposed by water resource development schemes on the conservation of endangered fish species in the Guadiana River Basin in Portugal. Fisheries Manag Ecol, 7, 167–178.

Copp,G. and Garner,P. (1995) Evaluating the Microhabitat Use of Fresh-Water Fish Larvae and Juveniles with Point Abundance Sampling by Electrofishing. Folia Zool, 44, 145–158.

Copp,G. and Penaz,M. (1988) Ecology of Fish Spawning and Nursery Zones in the Flood-Plain, Using a New Sampling Approach. Hydrobiologia, 169, 209–224.

Cordeiro-Estrela,P., Baylac,M., Denys,C. and Polop,J. (2008) Combining geometric morphometrics and pattern recognition to identify interspecific patterns of skull variation: case study in sympatric Argentinian species of the genus Calomys (Rodentia : Cricetidae : Sigmodontinae). Biol J Linn Soc, 94, 365–378.

Costa,C., Tibaldi,E., Pasqualetto,L. and Loy,A. (2006) Morphometric comparison of the cephalic region of cultured Acipenser baerii (Brandt, 1869), Acipenser naccarii (Bonaparte, 1836) and their hybrid. J Appl Ichthyol, 22, 8–14.

Coyne,J.A. and Orr,H.A. (2004) Speciation 1st ed. Sinauer Associates, Inc.

Crespo-Lopez,M.E., Pala,I., Duarte,T.L., Dowling,T.E. and Coelho,M.M. (2007) Genetic structure of the diploid-polyploid fish Squalius alburnoides in southern Iberian basins Tejo and Guadiana, based on microsatellites. J FIsh Biol, 71, 423–436.

Crespo-Lopez,M.E., Duarte,T., Dowling,T. and Coelho,M.M. (2006) Modes of reproduction of the hybridogenetic fish Squalius alburnoides in the Tejo and Guadiana rivers: An approach with microsatellites. Zoology, 109, 277–286.

Crivelli,A.J. (2011) Iberocypris alburnoides. In: IUCN 2011. IUCN Red List of Threatened Species. Version 2011.2. http://www.iucnredlist.org.

Cunha,C., Bastir,M., Coelho,M.M. and Doadrio,I. (2010) Body shape evolution among ploidy levels of the Squalius alburnoides hybrid complex (Teleostei, Cyprinidae). J Evolution Biol, 22.

Cunha,C., Coelho,M.M., Carmona,J.A. and Doadrio,I. (2004) Phylogeographical insights into the origins of the Squalius alburnoides complex via multiple hybridization events. Mol Ecol, 13, 2807–2817.

Cunha,C., Doadrio,I., Abrantes,J. and Coelho,M.M. (2011) The evolutionary history of the allopolyploid Squalius alburnoides (Cyprinidae) complex in the northern Iberian Peninsula. Heredity, 106, 100–112.

Czerwinski,I.A., Gutiérrez-Estrada,J.C., Soriguer,M.C. and Hernando,J.A. (2008) Morphometric Relations for Body Size and Mouth Dimensions for Four Fish Species in the Strait of Gibraltar. Acta Ichthyol Piscat, 38, 81–90.

Dawley,R. (1989) An introduction to unisexual vertebrates Evolution and ecology of unisexual vertebrates.

Dieterman,D. and Galat,D. (2005) Variation in body form, taste buds, and brain patterns of the sicklefin chub, Macrhybopsis meeki, in the Missouri River and lower Yellowstone River, USA. J Freshwater Ecol, 20, 561–573.

Doadrio,I. (2002) Atlas y Libro Rojo de Los Peces Continentals de España 2nd ed. Doadrio,I. (ed) Dirección General de Conservación de la Naturaleza, Museo Nacional de Ciencias Naturales, Madrid.

Doering-Arjes,P., Carnidale,M. and Mosegaard,H. (2008) Estimating population age structure using otolith morphometrics: a test with known-age Atlantic cod (Gadus morhua) individuals. Can J Fish Aquat Sci, 65, 2342–2350.

Doeringsfeld,M., Schlosser,I.J., Elder,J.F. and Evenson,D.P. (2004) Phenotypic consequences of genetic variation in a gynogenetic complex of Phoxinus eos-neogaeus clonal fish (Pisces : Cyprinidae) inhabiting a heterogeneous environment. Evolution, 58, 1261–1273.

Echelle,AA and Echelle,AF (1997) Patterns of abundance and distribution among members of a unisexual-bisexual complex of fishes (Atherinidae: Menidia). Copeia, 249–259.

Esteves,E., Pina,T., Chicharo,M.A. and Andrade,J.P. (2000) The distribution of estuarine fish larvae: Nutritional condition and co-occurrence with predators and prey. Acta Oecol, 21, 161–173.

Etheridge,E.C., Bean,C.W., Maitland,P.S. and Adams,C.E. (2010) Morphological and ecological responses to a conservation translocation of powan (Coregonus lavaretus) in Scotland. Aquat Conserv, 20, 274–281.

Fitter,R. and Manuel,R. (1986) A Field Guide to the Freshwater Life of Britain and North-west Europe Harper-Collins Publishers, London.

Fox,M.G., Vila-Gispert,A. and Copp,G.H. (2007) Life-history traits of introduced Iberian pumpkinseed Lepomis gibbosus relative to native populations. Can differences explain colonization success? J FIsh Biol, 71, 56–69.

Francis,R. (1990) Back-Calculation of Fish Length - a Critical-Review. J FIsh Biol, 36, 883–902.

Friess,M. and Baylac,M. (2003) Exploring artificial cranial deformation using elliptic Fourier analysis of Procrustes aligned outlines. Am J Phys Anthropol, 122, 11–22.

Garcia-Berthou,E. and Moreno-Amich,R. (2000) Food of introduced pumpkinseed sunfish: ontogenetic diet shift and seasonal variation. J FIsh Biol, 57, 29–40.

Gavrilets,S. (2003) Perspective: Models of speciation: What have we learned in 40 years? Evolution, 57, 2197–2215.

Godinho,F., Ferreira,M.T. and Cortes,R.V. (1997) The environmental basis of diet variation in pumpkinseed sunfish, Lepomis gibbosus, and largemouth bass, Micropterus salmoides, along an Iberian river basin. Environ Biol Fish, 50, 105–115.

Godinho,F., Ferreira,M.T. and Santos,J.M. (2000) Variation in fish community composition along an Iberian river basin from low to high discharge: relative contributions of environmental and temporal variables. In, Ecology of Freshwater Fish. pp. 22–29.

Gomes,J.L.J. and Monteiro,L. (2008) Morphological divergence patterns among populations of Poecilia vivipara (Teleostei Poeciliidae): test of an ecomorphological paradigm. Biol J Linn Soc, 93, 799–812.

Gomes-Ferreira,A., Ribeiro,F., Moreira da Costa,L., Cowx,I.G. and Collares-Pereira,M.J. (2005) Variability in diet and foraging behaviour between sexes and ploidy forms of the hybridogenetic Squalius alburnoides complex (Cyprinidae) in the Guadiana River basin, Portugal. J FIsh Biol, 66, 454–467.

Gray,M. and Weeks,S. (2001) Niche breadth in clonal and sexual fish (Poeciliopsis): a test of the frozen niche variation model. Can J Fish Aquat Sci, 58, 1313–1318.

Griffith,O.L., Moodie,G.E.E. and Civetta,A. (2003) Genome size and longevity in fish. Exp Gerontol, 38, 333–337.

Gromicho,M. and Collares-Pereira,M.J. (2004) Polymorphism of major ribosomal gene chromosomal sites (NOR-phenotypes) in the hybridogenetic fish Squalius alburnoides complex (Cyprinidae) assessed through crossing experiments. Genetica, 122, 291–302.

Hale,L.F. and Lowe,C.G. (2008) Age and growth of the round stingray Urobatis halleri at Seal Beach, California. J FIsh Biol, 73, 510–523.

Heery,E.C. and Berkson,J. (2009) Systematic Errors in Length Frequency Data and their Effect on Age-Structured Stock Assessment Models and Management. T Am Fish Soc, 138, 218–232.

Helland,I.P., Vøllestad,L.A., Freyhof,J. and Mehner,T. (2009) Morphological differences between two ecologically similar sympatric fishes. J FIsh Biol, 75, 2756–2767.

Hermida,M., Fernandez,J., Amaro,R. and San Miguel,E. (2005) Morphometric and meristic variation in Galician threespine stickleback populations, northwest Spain. Environ Biol Fish, 73, 189–200.

Hickley,P. and Dexter,K.F. (1979) A Comparative Index for Quantifying Growth in Length of Fish. Fisheries Management, 10, 147–151.

Huntingford,F. and Wright,P. (1989) How Sticklebacks Learn to Avoid Dangerous Feeding Patches. Behav Process, 19, 181–189.

Hurvitz,A., Jackson,K., Yom-Din,S. Degani,G. and Levavi-Sivan,B. (2008) Sexual development in Russian sturgeon (Acipenser gueldenstaedtii) grown in aquaculture. In, Cybium. pp. 283–285.

Hynes,H.B.N. (1950) The Food of Fresh-Water Sticklebacks (Gasterosteus aculeatus and Pygosteus pungitius), with a Review of Methods Used in Studies of the Food of Fishes. J Anim Ecol, 19, 36–58.

Ibáñez,A.L., Britton,J.R. and Cowx,I.G. (2008) Relationship between scale growth checks, circuli formation rate and somatic growth in Rutilus rutilus (L.) a fish farm-reared cyprinid. J FIsh Biol, 72, 1023–1034.

Ibañez,A.L., Cowx,I.G. and O'Higgins,P. (2007) Geometric morphometric analysis of fish scales for identifying genera, species, and local populations within the Mugilidae. Can J Fish Aquat Sci, 64, 1091–1100.

Ibañez,A.L., Cowx,I.G. and O'Higgins,P. (2009) Variation in elasmoid fish scale patterns is informative with regard to taxon and swimming mode. Zool J Linn Soc-Lond, 155, 834–844.

Ihssen,P., Booke,H.E., Casselman,J.M, Mcglade,J.M., Payne,N.R. and Utter,F.M. (1981) Stock Identification - Materials and Methods. Can J Fish Aquat Sci, 38, 1838–1855.

Janko,K., Bohlen,J., Lamatsch,D., Flajšhans,M., Epplen,J.T., Ráb,P., Kotlík,P. and Šlechtová,V. (2007) The gynogenetic reproduction of diploid and triploid hybrid spined loaches (Cobitis : Teleostei), and their ability to establish successful clonal lineages-on the evolution of polyploidy in asexual vertebrates. Genetica, 131, 185–194.

Jiangfeng,Z., Xiaojie,D. and Yong,C. (2011) Species composition and diversity of pelagic fishes based on a longline fishery catch in the North Pacific Ocean. Chin J Oceanol Limn, 29, 261–269.

Joly,S. and Bruneau,A. (2004) Evolution of triploidy in Apios americana (Leguminosae) revealed by genealogical analysis of the histone H3-D gene. Evolution, 58, 284–295.

Karpouzi,V. and Stergiou,K. (2003) The relationships between mouth size and shape and body length for 18 species of marine fishes and their trophic implications. J FIsh Biol, 62, 1353–1365.

Kendall,N.W. and Quinn,T.P. (2009) Effects of population-specific variation in age and length on fishery selection and exploitation rates of sockeye salmon (Oncorhynchus nerka). Can J Fish Aquat Sci, 66, 896–908.

Klingenberg,C. (2011) MorphoJ: an integrated software package for geometric morphometrics. Mol Ecol Resour.

Klingenberg,C., Barluenga,M. and Meyer,A. (2002) Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry. Evolution, 56, 1909–1920.

Kotusz,J. (2008) Morphological Relationships Between Polyploid Hybrid Spined Loaches of the Genus Cobitis (Teleostei: Cobitidae) and Their Parental Species. Ann Zool, 58, 891–905.

Lampert,K.P. (2009) Clonal reproduction in freshwater fish: mechanisms, systematic overview, genetic and ecological consequences. Fund Appl Limnol, 174, 245–260.

Lea,E. (1910) On the methods used in the herring investigations. ICES Publication de Circonstance, 53, 5–174.

Leica Microsystems (2008) Leica Application Suite, Wetzlar, Germany.

Longo S,M., Zamora, H.G., Guisande, C.G. and Jairo Ramirez,J.R. (2010) Macroinvertebrate community dynamics in the Potrerillos stream (Colombia): Response to seasonal flow changes. Limnetica, 29, 195–210.

Madurell,T. and Cartes,J. (2005) Temporal changes in feeding habits and daily rations of Hoplostethus mediterraneus in the bathyal Ionian Sea (eastern Mediterranean). Mar Biol, 146, 951–962.

Magalhaes,M. (1993a) Effects of Season and Body-Size on the Distribution and Diet of the Iberian Chub Leuciscus-Pyrenaicus in a Lowland Catchment. J FIsh Biol, 42, 875–888.

Magalhaes,M. (1993b) Feeding of an Iberian Stream Cyprinid Assemblage: Seasonality of Resource Use in a Highly Variable Environment. Oeologia, 96, 253–260.

Maitland,P.S. and Campbell,R.N. (1992) Freshwater fishes of the British Isles Harper Collins.

Mallet,J. (2007) Hybrid speciation. Nature, 446, 279–283.

Martínez-Capel,F., García de Jalón,D., Werenitzky,D., Baeza,D. and Rodilla-Alamá,M. (2009) Microhabitat use by three endemic Iberian cyprinids in Mediterranean rivers (Tagus River Basin, Spain). Fisheries Manag Ecol, 16, 52–60.

Martins,M., Collares-Pereira,M.J., Cowx,I.G. and Coelho,M.M. (1998) Diploids v. triploids of Rutilus alburnoides: spatial segregation and morphological differences. J FIsh Biol, 52, 817–828.

Mercier,L. Panfili,J., Paillon,C., N'diaye,A., Nouillot,D. and Darnaude,A.M. (2011) Otolith reading and multi-model inference for improved estimation of age and growth in the gilthead seabream Sparus aurata (L.). Estuar Coast Shelf S, 92, 534–545.

Merkel,F.R., Jamieson,S.E., Falk,K. and Mosbech,A. (2007) The diet of common eiders wintering in Nuuk, Southwest Greenland. Polar Biol, 30, 227–234.

Mesquita,N., Cunha,C., Carvalho,G.R. and Coelho,M.M. (2007) Comparative phylogeography of endemic cyprinids in the south-west Iberian Peninsula: evidence for a new ichthyogeographic area. J FIsh Biol, 71, 45–75.

Mesquita,N., Hanfling,B., Carvalho,G.R. and Coelho,M.M. (2005) Phylogeography of the cyprinid Squalius aradensis and implications for conservation of the endemic freshwater fauna of southern Portugal. Mol Ecol, 14, 1939–1954.

Mitteroecker,P. and Bookstein,F. (2011) Linear Discrimination, Ordination, and the Visualization of Selection Gradients in Modern Morphometrics. Evol Biol, 38, 100–114.

Monteiro,L. and Gomes,J.L.J. (2005) Morphological divergence rate tests for natural selection: Uncertainty of parameter estimation and robustness of results. Genet Mol Biol, 28, 345–355.

Morais,P., Chicharo,M.A. and Chicharo,L. (2009) Changes in a temperate estuary during the filling of the biggest European dam. Sci Total Environ, 407, 2245–2259.

Morán-López,R., Da Silva,E., Perez-Bote,J.L. and Corbacho Amado,C. (2006) Associations between fish assemblages and environmental factors for Mediterranean-type rivers during summer. J FIsh Biol, 69, 1552–1569.

Morán-López,R., Perez-Bote,J.L., Rubio,E.D. and Amado,C.C. (2005) Summer habitat relationships of barbels in south-west Spain. J FIsh Biol, 67, 66–82.

Morgado-Santos,M., Matos,I., Vicente,L. and Collares-Pereira,M.J. (2010) Scaleprinting: individual identification based on scale patterns. J FIsh Biol, 76, 1228–1232.

Morinville,G.R. and Rasmussen,J.B. (2008) Distinguishing between juvenile anadromous and resident brook trout (Salvelinus fontinalis) using morphology. Environ Biol Fish, 81, 171–184.

Morozov-Leonov,S.I., Mezhzherin,S.V., Nekrasova,O.D., Shabanov,D.A., Korshunov,A.V. and Kurtiak,F.F. (2009) Inheritance of parental genomes by a hybrid form Rana “esculenta” (Amphibia, Ranidae). Genetika, 45, 488–495.

Mullen,J.A., Bramblett,R.G., Guy,C.S., Zale,A.V. and Roberts,D.W. (2011) Determinants of Fish Assemblage Structure in Northwestern Great Plains Streams. T Am Fish Soc, 140, 271–281.

Murta,A. (2000) Morphological variation of horse mackerel (Trachurus trachurus) in the Iberian and North African Atlantic: implications for stock identification. ICES J Mar Sci, 57, 1240–1248.

Neves,F. and Monteiro,L. (2003) Body shape and size divergence among populations of Poecilia vivipara in coastal lagoons of south-eastern Brazil. J FIsh Biol, 63, 928–941.

Nunn,A.D., Cowx,I.G. and Harvey,J.P. (2002) Recruitment patterns of six species of cyprinid fishes in the lower River Trent, England. Ecol Freshw Fish, 11, 74–84.

Ohira,Y., Shimizu,M., Ura,K. and Takagi,Y. (2007) Scale regeneration and calcification in goldfish Carassius auratus: quantitative and morphological processes. Fisheries Sci, 73, 46–54.

Pala,I. and Coelho,M. (2005) Contrasting views over a hybrid complex: Between speciation and evolutionary. Gene.

Pires,A., Cowx,I.G. and Coelho,M.M. (2001) Diet and growth of two sympatric Iberian barbel, Barbus steindachneri and Barbus microcephalus, in the middle reaches of the Guadiana Basin (Portugal). Folia Zool, 50, 291–304.

Pires,A., Cowx,I.G. and Coelho,M.M. (1999) Seasonal changes in fish community structure of intermittent streams in the middle reaches of the Guadiana basin, Portugal. J FIsh Biol, 54, 235–249.

Pires,A., Cowx,I.G., and Coelho,M.M. (2000a) Benthic macroinvertebrate communities of intermittent streams in the middle reaches of the Guadiana Basin (Portugal). Hydrobiologia, 435, 167–175.

Pires,A., Cowx,I.G., and Coelho,M. (2000b) Life history strategy of Leuciscus pyrenaicus (Cyprinidae) in intermittent streams of the Guadiana basin (Portugal). Cybium, 24, 287–297.

Poulet,N., Reyjol,Y., Collier,H. and Lek,S. (2005) Does fish scale morphology allow the identification of populations at a local scale? A case study for rostrum dace Leuciscus leuciscus burdigalensis in River Viaur (SW France). Aquat Sci, 67, 122–127.

Prospero,M. and Collares-Pereira,M.J. (2000) Nuclear DNA content variation in the diploid-polyploid Leuciscus alburnoides complex (Teleostei, Cyprinidae) assessed by flow cytometry. Folia Zool, 49, 53–58.

Ptacek,M., Gerhardt,H.C. and Sage,R.D. (1994) Speciation by Polyploidy in Treefrogs - Multiple Origins of the Tetraploid, Hyla Versicolor. Evolution, 48, 898–908.

Pusey,B.J., Arthington,A.H., Stewart-Koster,B., Kennard,M.J. and Read,M.G. (2010) Widespread omnivory and low temporal and spatial variation in the diet of fishes in a hydrologically variable northern Australian river. J FIsh Biol, 77, 731–753.

R Development Core Team (2011) R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria.

Reebs,S. (1999) Time-place learning based on food but not on predation risk in a fish, the inanga (Galaxias maculatus). Ethology, 105, 361–371.

Restrepo,V.R., Diaz,G.A., Walter,J.F., Neilson,J.D., Campana,S.E., Secor,D. and Wingate,R.L. (2010) Updated estimate of the growth curve of Western Atlantic bluefin tuna. Aquat Living Resour, 23, 335–342.

Ribeiro,F., Cowx,I.G., Tiago,P., Filipe,A.F., Moreira da Costa,L. and Collares-Pereira,M.J. (2003) Growth and reproductive traits of diploid and triploid forms of the Squalius alburnoides cyprinid complex in a tributary of the Guadiana River, Portugal. Arch Hyrdobiol, 156, 471–484.

Ribeiro,F., Cowx,I.G. and Collares-Pereira,M.J. (2000) Life history traits of the endangered Iberian cyprinid Anaecypris hispanica and their implications for conservation. Arch Hydrobiol, 149, 569–586.

Ritterbusch,D. and Bohlen,J. (2000) On the ecology of spined loach in Lake Muggelsee. In, Folia Zoologica. pp. 187–192.

Robalo,J.I., Sousa-Santos,C., Levy,A. and Almada,V.C. (2006) Molecular insights on the taxonomic position of the paternal ancestor of the Squalius alburnoides hybridogenetic complex. Mol Phylogenet Evol, 39, 276–281.

Robinson,J.M., Jirka,K.J. and Chiotti,J.A. (2010) Age and growth analysis of the central mudminnow, Umbra limi (Kirtland, 1840). J Appl Ichthyol, 26, 89–94.

Rohlf,F.J. (2010a) TPSDig2, Version 2.16. Department of Ecology and Evolution, State University of New York, Stony Brook, NY. http://life.bio.sunysb.edu/morph

Rohlf,F.J. (2010b) TPSRelw, Version 1.49. Department of Ecology and Evolution, State University of New York, Stony Brook, NY. http://life.bio.sunysb.edu/morph

Salini,J., Milton,D.A., Rahman,M.J. and Hussain,M.G. (2004) Allozyme and morphological variation throughout the geographic range of the tropical shad, hilsa Tenualosa ilisha. Fish Res, 66, 53–69.

Sanjur,O., Carmona,J.A. and Doadrio,I. (2003) Evolutionary and biogeographical patterns within Iberian populations of the genus Squalius inferred from molecular data. Mol Phylogenet Evol, 29, 20–30.

Schaefer,J., Kreiser,B.R., Champagne,C., Mickle,P.M. and Duvernell,D.D. (2009) Patterns of co-existence and hybridisation between narrowly endemic (Fundulus euryzonus) and broadly distributed (F-olivaceus) topminnows in a riverine contact zone. Ecol Freshw Fish, 18, 360–368.

Schlosser,I. (1982) Fish Community Structure and Function along Two Habitat Gradients in a Headwater Stream. Ecol Monogr, 52, 395–414.

Schlosser,I., Doeringsfeld,M.R., Elder,J.F. and Arzayus,L.F. (1998) Niche relationships of clonal and sexual fish in a heterogeneous landscape. Ecology, 79, 953–968.

Schlupp,I. (2005) The evolutionary ecology of gynogenesis. Annu Rev Ecol Evol S, 36, 399–417.

Schoener,T. (1970) Nonsynchronous Spatial Overlap of Lizards in Patchy Habitats. Ecology, 51, 408–&.

Schultz,R. (1977) Evolution and ecology of unisexual fishes. Evol Biol, 10, 277–331.

Schultz,R. (1971) Special Adaptive Problems Associated with Unisexual Fishes. Am Zool, 11, 1–10.

Silva,A. (2003) Morphometric variation among sardine (Sardina pilchardus) populations from the northeastern Atlantic and the western Mediterranean. ICES J Mar Sci, 60, 1352–1360.

Sire,J. and Akimenko,M. (2004) Scale development in fish: a review, with description of sonic hedgehog (shh) expression in the zebrafish (Danio rerio). Int J Dev Biol, 48, 233–247.

Sire,J. and Arnulf,I. (1990) The development of squamation in four teleostean fishes with a survey of the literature. Jpn J Ichthyol, 37, 133-143.

Sobel,J.M., Chen,G.F., Watt,L.R. and Schemske,D.W. (2010) The biology of speciation. Evolution, 64, 295–315.

Sousa-Santos,C., Collares-Pereira,M.J., and Almada,V.C. (2006a) Evidence of extensive mitochondrial introgression with nearly complete substitution of the typical Squalius pyrenaicus-like mtDNA of the Squalius alburnoides complex (Cyprinidae) in an independent Iberian drainage. J FIsh Biol, 68, 292–301.

Sousa-Santos,C., Collares-Pereira,M.J., and Almada,V.C. (2006b) May a hybridogenetic complex regenerate the nuclear genome of both sexes of a missing ancestor? First evidence on the occurrence of a nuclear non-hybrid Squalius alburnoides (Cyprinidae) female based on DNA sequencing. J Nat Hist, 40, 1443–1448.

Sousa-Santos,C., Collares-Pereira,M.J., and Almada,V.C. (2006c) Reproductive success of nuclear nonhybrid males of Squalius alburnoides hybridogenetic complex (Teleostei,

Cyprinidae): an example of interplay between female choice and ecological pressures? Acta Ethol, 9, 31–36.

Sousa-Santos,C., Collares-Pereira,M.J., and Almada,V.C. (2007a) Fertile triploid males: an uncommon case among hybrid vertebrates. J Exp Zool Part A, 307, 220–225.

Sousa-Santos,C., Collares-Pereira,M.J., and Almada,V.C. (2007b) Reading the history of a hybrid fish complex from its molecular record. Mol Phylogenet Evol, 45, 981–996.

Spiegel,J.R., Quist,M.C. and Morris,J.E. (2010) Precision of Scales and Pectoral Fin Rays for Estimating Age of Highfin Carpsucker, Quillback Carpsucker, and River Carpsucker. J Freshwater Ecol, 25, 271–278.

Stransky,C., Murta,A.G., Schlickeisen,J. and Zimmermann,C. (2008) Otolith shape analysis as a tool for stock separation of horse mackerel (Trachurus trachurus) in the Northeast Atlantic and Mediterranean. Fish Res, 89, 159–166.

Thanukos,A. (2008) Views from Understanding Evolution: Parsimonious Explanations for Punctuated Patterns. Evo Edu Outreach, 1, 138–146.

Ungerer,M.C., Baird,S.J., Pan,J. and Rieseberg,L.H. (1998) Rapid hybrid speciation in wild sunflowers. P Natl Acad Sci Usa, 95, 11757–11762.

Vargas,J., Real,R. and Guerrero,J.C. (1998) Biogeographical regions of the Iberian peninsula based on freshwater fish and amphibian distributions. Ecography, 21, 371–382.

Vilizzi,L. and Copp,G. (2008) Diel densities of young-of-the-year fishes in the River Lee (Hertfordshire) over a 24-hour period in late summer. The London Naturalist, 87, 113–117.

Vrijenhoek,R. (1989) Genetic and ecological constraints on the origins and establishment of unisexual vertebrates Evolution and ecology of unisexual vertebrates.

Vrijenhoek,R. and Dawley,R. (1989) A list of the known unisexual vertebrates Evolution and ecology of unisexual vertebrates.

Wallace,R.,Jr (1981) An assessment of diet-overlap indexes. T Am Fish Soc.

Washington,H. (1984) Diversity, Biotic and Similarity Indexes - a Review with Special Relevance to Aquatic Ecosystems. Water Res, 18, 653–694.

Weber,M.J. and Brown,M.L. (2011) Comparison of common carp (Cyprinus carpio) age estimates derived from dorsal fin spines and pectoral fin rays. J Freshwater Ecol, 26, 195–202.

Weeks,S., Gaggiotti,O.E., Schenck,R.A., Spindler,K.P. and Vrijenhoek,R.C. (1992) Feeding-Behavior in Sexual and Clonal Strains of Poeciliopsis. Behav Ecol Sociobiol, 30, 1–6.

Wesner,J.S., Billman,E.J., Meier,A. and Belk,M.C. (2011) Morphological convergence during pregnancy among predator and nonpredator populations of the livebearing fish Brachyrhaphis rhabdophora (Teleostei: Poeciliidae). Biol J Linn Soc, 104, 386–392.

Winemiller,K. (1991) Ecomorphological Diversification in Lowland Fresh-Water Fish Assemblages From 5 Biotic Regions. Ecol Monogr, 61, 343–365.

Zar,J.H. (1999) Biostatistical analysis 4th ed. Prentice-Hall, Upper Saddle River, NJ.

Zardoya,R. and Doadrio,I. (1999) Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. J Mol Evol, 49, 227–237.

Zelditch,M. (2004) Geometric morphometrics for biologists: a primer.

Zymonas,N.D. and McMahon,T.E. (2009) Comparison of pelvic fin rays, scales and otoliths for estimating age and growth of bull trout, Salvelinus confluentus. Fisheries Manag Ecol, 16, 155–164.

Appendix

Appendix 1: Discriminant analysis in two steps

Guadiana

This methodology describes another possible way of classifying the ploidy levels of *S. alburnoides*. The results presented here are comparable to the results found in the main body of the thesis. This methodology is included for information purposes. Predictions may be carried out in the same manner as previously described, also yielding similar results. It was however deemed that the approach presented previously was a simpler and more effective way of discriminating and classifying individuals.

This method initially separated fish by diploid and triploid individuals. Following this, all diploid individuals were reclassified to discriminate them between diploid males (AA) and other diploids (PA).

The first discriminant analysis contained 132 diploid individuals, all for which sex was known, and 199 triploid individuals. Diploid individuals of unknown sex present in previous analyses were removed in order to avoid diploid males (AA) being incorrectly labeled as diploids (PA) and influencing the accuracy of the linear discriminant function. Table 1A.1 shows that 66.7% of diploid individuals and 87% of triploid individuals were classified correctly by cross-validation. The CVA plot distinguishes between individuals of different ploidy with minimal overlap (Figure 1A.1).

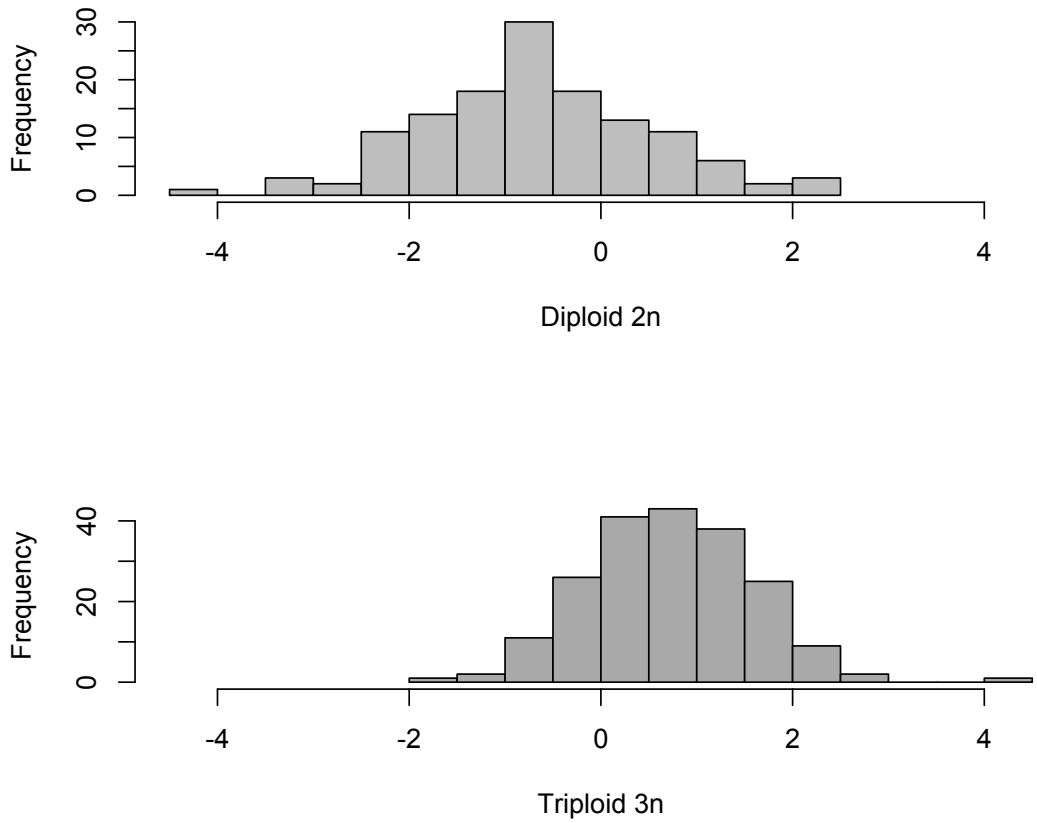


Figure 1A.1: CVA plot of Guadiana river, showing shape differences between diploid (2) and triploid (3) individuals.

Table 1A.1: Cross-validation of diploid individuals of known sex (2n) and triploid individuals from the Guadiana river basin.

	2n	3n
2n	88	44
3n	24	175

After this discriminant analysis was performed, triploid individuals were removed from the dataset so that only diploid individuals remained. 84 individuals were classed as diploid males (AA) and 48 individuals were known to be female diploids, thus were classed as PA individuals. The canonical variates analysis, discriminant analysis and cross-validation were able to effectively separate the diploid individuals into two

distinct groups (Figure 1A.2). 86% of diploid males were correctly identified and 64% of diploid females were also found to be correct (Table 1A.2).

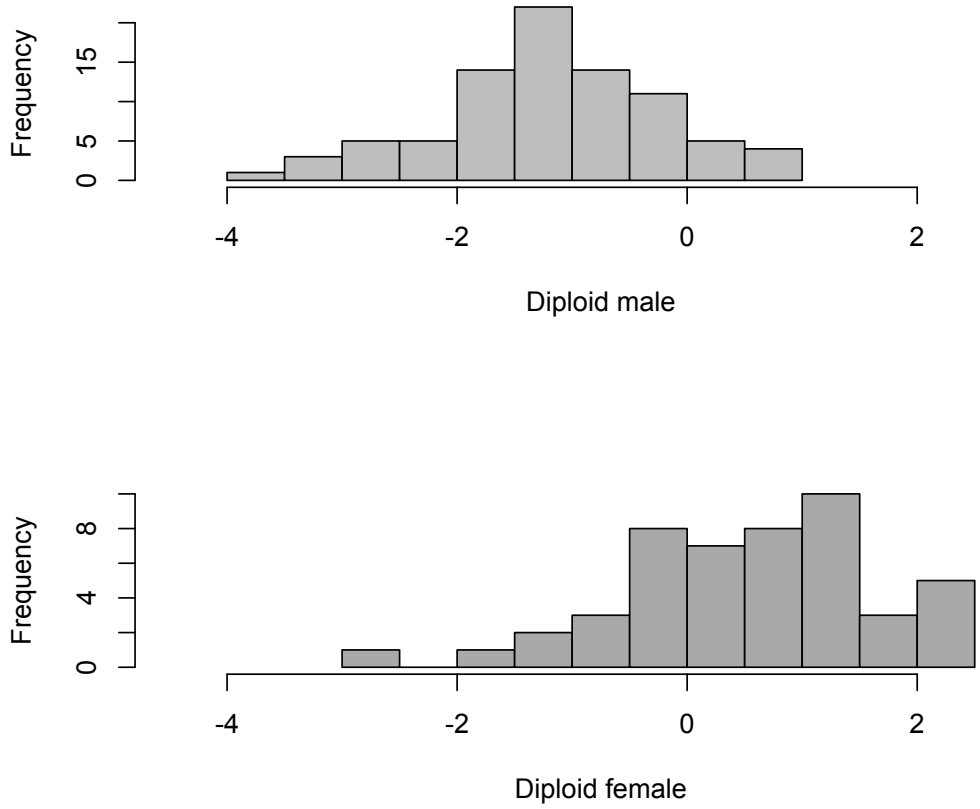


Figure 1A.2: CVA plot of Guadiana river, showing shape differences between diploid females (2n) and diploid males (AA) individuals.

Table 1A.2: Cross-validation of diploid females (2n) and diploid males (AA) from the Guadiana river basin.

	AA	2n
AA	73	11
2n	17	31

The two discriminant functions combined accurately classified the majority of individuals. The linear discriminant functions created using this dataset could then be used to predict the ploidy of unknown fish by analysing the scale shape.

All datasets

As with the Guadiana dataset, a two-step discriminant analysis was performed, to see if this approach could potentially improve the classification between triploids and diploid females. The cross-validation for the first step showed good accuracy in discriminating between diploid males and the other groups (Table 1A.3 - A and Figure 1A.3 - A). However, the second step showed very little accuracy in discriminating diploid females from triploids (Table 1A.3 - B and Figure 1A.3 - B).

Table 1A.3: A) Cross-validation table for all river basins, between diploid male (2AA) and triploid and diploid female (2n - 3n) individuals. B) Cross-validation table for all river basins, between female diploid (2n) and triploid (3n) individuals.

A)

	2AA	2n - 3n
2AA	58	28
2n - 3n	18	260

B)

	2n	3n
2n	7	43
3n	5	223

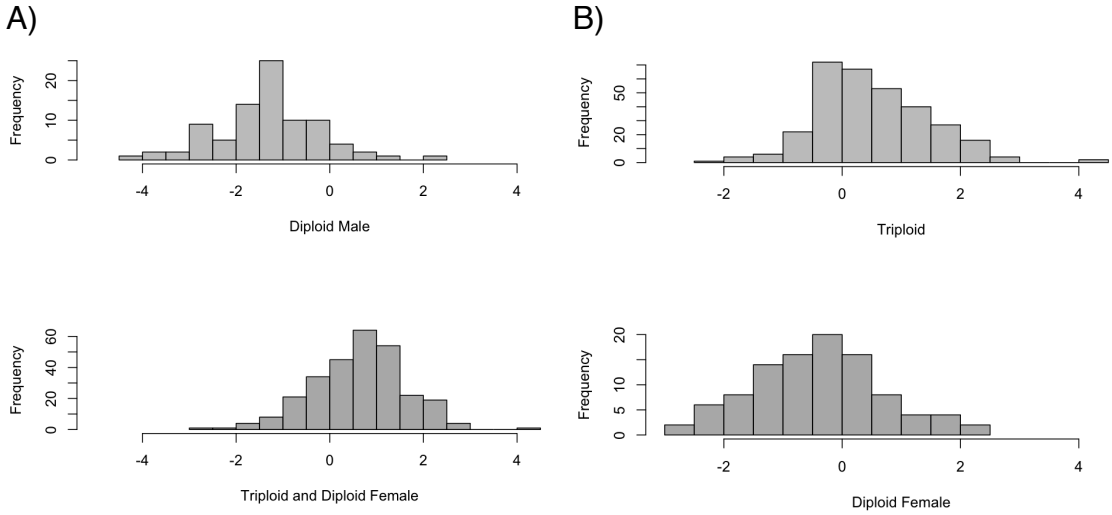


Figure 1A.3: A) CVA plot for all river basins, between diploid males and grouped triploid and diploid female individuals. B) CVA plot for all river basins, between triploid and female diploid individuals.

Appendix 2: Prediction values and probabilities

Guadiana by scale shape

The following table shows the individuals for which ploidy was predicted using the linear discriminant function in Chapter 3, using scale shape, as well as the probability of the prediction being accurate.

Table 2A.1: Prediction values and probabilities derived from analysis of scale shape. Individuals for which prediction probability is over 60% are in bold.

Specimen	Predicted ploidy	Diploid male probability	Diploid female probability	Triploid probability
V1	**Diploid male**	**0.908256551**	**0.021931489**	**0.069811961**
V2	**Diploid male**	**0.984499072**	**0.004652165**	**0.010848763**
V3	**Diploid male**	**0.996706532**	**0.001924477**	**0.001368991**
V4	**Triploid**	**0.015811088**	**0.289536099**	**0.694652813**
V5	**Diploid male**	**0.971904132**	**0.005895084**	**0.022200784**
V7	Triploid	0.185820286	0.280900014	0.5332797
V8	**Diploid male**	**0.844642945**	**0.038641306**	**0.116715749**
V9	Triploid	0.105218358	0.313106235	0.581675406
V10	**Triploid**	**0.037869136**	**0.305316891**	**0.656813973**
V11	**Triploid**	**0.143770504**	**0.196920427**	**0.659309069**
V12	**Triploid**	**0.089851124**	**0.207472652**	**0.702676225**
V13	**Diploid male**	**0.964429594**	**0.003836505**	**0.031733902**
V14	**Diploid male**	**0.956348293**	**0.009823013**	**0.033828693**
V15	**Diploid male**	**0.665637162**	**0.140064116**	**0.194298722**
V16	**Triploid**	**0.106762934**	**0.143784811**	**0.749452255**
V17	Diploid male	0.543697183	0.105493922	0.350808895
V18	**Triploid**	**0.048521101**	**0.287535882**	**0.663943017**
V19	**Triploid**	**0.010288473**	**0.129478222**	**0.860233304**
V20	Triploid	0.283120579	0.191223663	0.525655758
V21	**Triploid**	**0.015878905**	**0.181076446**	**0.803044648**
V22	Triploid	0.010045957	0.470531515	0.519422528
V23	**Triploid**	**0.113418088**	**0.191281447**	**0.695300465**
V24	Triploid	0.032387187	0.397329956	0.570282857
V25	Triploid	0.099132183	0.317041518	0.583826299

Specimen	Predicted ploidy	Diploid male probability	Diploid female probability	Triploid probability
V26	**Diploid male**	**0.945186719**	**0.027605558**	**0.027207723**
V27	**Triploid**	**0.092123766**	**0.21839086**	**0.689485374**
V28	**Diploid male**	**0.996984813**	**0.001325625**	**0.001689562**
V29	Triploid	0.191524848	0.210529694	0.597945458
V30	**Triploid**	**0.126301841**	**0.221278894**	**0.652419265**
V31	**Triploid**	**0.022701581**	**0.275648888**	**0.701649532**
V32	**Diploid male**	**0.947109296**	**0.030373639**	**0.022517065**
V33	**Diploid male**	**0.982837659**	**0.012406232**	**0.004756109**
V34	**Diploid male**	**0.871180034**	**0.036633191**	**0.092186775**
V35	**Triploid**	**0.039535578**	**0.116650253**	**0.843814168**
V36	**Diploid male**	**0.74463043**	**0.076387659**	**0.178981911**
V37	**Triploid**	**0.041511339**	**0.233327631**	**0.72516103**
V38	**Triploid**	**0.13961603**	**0.173211466**	**0.687172504**
V39	**Triploid**	**0.003446657**	**0.131349589**	**0.865203754**
V40	**Triploid**	**0.013802846**	**0.198475917**	**0.787721237**
V41	**Triploid**	**0.094152601**	**0.300890671**	**0.604956727**
V42	**Triploid**	**0.01195309**	**0.205251191**	**0.782795718**
V43	**Diploid male**	**0.949190083**	**0.018836177**	**0.03197374**
V44	**Diploid male**	**0.612792542**	**0.124993325**	**0.262214133**
V45	Triploid	0.142404089	0.358374421	0.49922149
V46	Triploid	0.227885942	0.339050827	0.433063231
V47	Diploid male	0.506541494	0.154773753	0.338684754
V48	**Triploid**	**0.004157939**	**0.080338953**	**0.915503109**
V49	**Triploid**	**0.004293678**	**0.156201113**	**0.839505209**
V50	**Triploid**	**0.057754231**	**0.109883947**	**0.832361822**
V51	Diploid female	0.252226338	0.457447692	0.29032597
V52	**Triploid**	**0.025527218**	**0.367546557**	**0.606926225**
V53	**Diploid male**	**0.985264422**	**0.003536734**	**0.011198845**
V54	Triploid	0.268800772	0.176485303	0.554713925
V55	**Triploid**	**0.00679469**	**0.323194442**	**0.670010867**
V56	**Triploid**	**0.178505472**	**0.106236499**	**0.715258029**
V57	Triploid	0.041801398	0.373609919	0.584588682
V58	Triploid	0.265327634	0.155350153	0.579322213
V59	**Triploid**	**0.052439141**	**0.267894718**	**0.679666141**

Specimen	Predicted ploidy	Diploid male probability	Diploid female probability	Triploid probability
V60	**Triploid**	**0.017521123**	**0.194923858**	**0.787555019**
V61	**Triploid**	**0.008351729**	**0.158147551**	**0.83350072**
V62	**Triploid**	**0.029808361**	**0.084997373**	**0.885194266**
V63	**Triploid**	**0.171380804**	**0.219913411**	**0.608705785**
V64	**Triploid**	**0.159317056**	**0.233189261**	**0.607493684**
V65	**Triploid**	**0.021999766**	**0.145613902**	**0.832386333**
V66	Triploid	0.114160518	0.359793898	0.526045584
V67	**Triploid**	**0.00592769**	**0.267110989**	**0.726961321**
V68	**Triploid**	**0.015942589**	**0.313991214**	**0.670066197**
V69	**Diploid male**	**0.960130374**	**0.020471665**	**0.019397961**
V70	**Triploid**	**0.075108724**	**0.100907038**	**0.823984238**
V71	**Diploid male**	**0.941028526**	**0.017832637**	**0.041138837**
V72	**Triploid**	**0.011505839**	**0.267344379**	**0.721149782**
V73	**Diploid male**	**0.015720209**	**0.239625328**	**0.744654463**
V74	Diploid male	0.587327511	0.096100206	0.316572283
V75	**Triploid**	**0.038122179**	**0.090226294**	**0.871651527**
V76	**Triploid**	**0.098219206**	**0.207460184**	**0.69432061**
V77	**Triploid**	**0.024297536**	**0.171020828**	**0.804681636**
V78	**Triploid**	**0.18968362**	**0.168808121**	**0.641508259**
V79	**Triploid**	**0.041293109**	**0.249184906**	**0.709521985**
V80	**Triploid**	**0.008779467**	**0.265917908**	**0.725302626**
V81	**Triploid**	**0.082062321**	**0.166284688**	**0.751652992**
V82	**Triploid**	**0.121493252**	**0.155606672**	**0.722900077**
V83	**Triploid**	**0.131828551**	**0.236447131**	**0.631724317**
V84	**Triploid**	**0.069553213**	**0.281039226**	**0.649407562**
V85	Diploid male	0.515362681	0.183325889	0.30131143
V86	**Triploid**	**0.008155525**	**0.114653383**	**0.877191091**
V87	**Triploid**	**0.012967291**	**0.170067182**	**0.816965527**
V88	Diploid female	0.104500082	0.514908873	0.380591045
V89	**Triploid**	**0.098506751**	**0.165368726**	**0.736124523**
V90	**Diploid male**	**0.995510321**	**0.003050974**	**0.001438705**
V91	Diploid male	0.562433245	0.193371419	0.244195336
V92	**Diploid male**	**0.997895619**	**0.000475469**	**0.001628912**
V93	**Diploid male**	**0.885152544**	**0.009157992**	**0.105689465**

Specimen	Predicted ploidy	Diploid male probability	Diploid female probability	Triploid probability
V94	**Diploid male**	**0.908810271**	**0.023820771**	**0.067368958**
V95	**Diploid male**	**0.961880831**	**0.027538663**	**0.010580507**
V97	**Diploid male**	**0.969301658**	**0.001664309**	**0.029034032**
V98	Diploid male	0.500464732	0.055655899	0.443879369
V99	**Diploid male**	**0.903758675**	**0.007460454**	**0.088780871**
V100	**Diploid male**	**0.782026241**	**0.03719785**	**0.180775909**
V101	Diploid male	0.551676791	0.105839049	0.342484161
V102	**Diploid male**	**0.983728495**	**0.00165072**	**0.014620786**
V103	**Triploid**	**0.116028271**	**0.273760301**	**0.610211429**
V104	**Diploid male**	**0.949668439**	**0.01007856**	**0.040253001**
V105	**Triploid**	**0.01207491**	**0.187990591**	**0.799934499**
V106	Triploid	0.359396208	0.203179532	0.43742426
V107	**Diploid male**	**0.983035587**	**0.007318705**	**0.009645708**
V108	Triploid	0.18156331	0.287585871	0.530850818
V109	**Triploid**	**0.0058972**	**0.184010559**	**0.810092241**
V110	**Diploid male**	**0.827608041**	**0.08376072**	**0.088631239**
V111	**Triploid**	**0.065736093**	**0.080345721**	**0.853918185**
V112	**Triploid**	**0.015017483**	**0.183957603**	**0.801024914**
V113	Triploid	0.240180794	0.191440117	0.568379089
V114	**Triploid**	**0.188865218**	**0.203249932**	**0.60788485**
V115	**Diploid male**	**0.9867029**	**0.002367226**	**0.010929874**
V116	**Diploid male**	**0.995191842**	**0.000230031**	**0.004578127**
V117	**Triploid**	**0.028146579**	**0.108957358**	**0.862896064**
V118	**Triploid**	**0.017203101**	**0.092412357**	**0.890384542**
V119	**Diploid male**	**0.984184344**	**0.005288923**	**0.010526733**
V120	Triploid	0.46578147	0.015140904	0.519077626
V121	**Triploid**	**0.006654951**	**0.095307509**	**0.89803754**
V122	**Diploid male**	**0.666925611**	**0.118903243**	**0.214171146**
V123	Diploid female	0.08539537	0.532229477	0.382375153
V124	**Triploid**	**0.072879536**	**0.105162279**	**0.821958185**
V125	**Diploid male**	**0.860754493**	**0.022203186**	**0.117042322**
V126	Diploid male	0.537072719	0.039843212	0.423084068
V127	Diploid male	0.553160536	0.011077269	0.435762196
V128	**Triploid**	**0.021012849**	**0.276034232**	**0.702952918**

Specimen	Predicted ploidy	Diploid male probability	Diploid female probability	Triploid probability
V129	**Triploid**	**0.044323003**	**0.118513945**	**0.837163052**
V130	**Triploid**	**0.022536074**	**0.194186679**	**0.783277248**
V131	Triploid	0.167158263	0.334612056	0.498229681
V132	Triploid	0.040135211	0.437397073	0.522467716
V133	**Triploid**	**0.05588864**	**0.202670527**	**0.741440834**
V134	Triploid	0.297105148	0.342526843	0.36036801
V135	**Triploid**	**0.001926447**	**0.118355124**	**0.879718429**
V136	Diploid male	0.533191548	0.194906923	0.271901529
V137	**Diploid male**	**0.651874899**	**0.024823773**	**0.323301328**
V138	**Diploid male**	**0.674848635**	**0.050472922**	**0.274678443**
V139	**Triploid**	**0.022945218**	**0.110418767**	**0.866636015**
V140	**Diploid male**	**0.94175411**	**0.016620248**	**0.041625642**
V141	**Diploid male**	**0.964589366**	**0.003805212**	**0.031605422**
V142	**Diploid male**	**0.615527053**	**0.055839763**	**0.328633185**
V143	**Triploid**	**0.008302891**	**0.062058678**	**0.929638431**
V144	**Diploid male**	**0.995348097**	**0.001676151**	**0.002975752**
V145	Diploid male	0.514112264	0.056777929	0.429109807
V146	**Diploid male**	**0.984663565**	**0.004767045**	**0.01056939**
V147	**Diploid male**	**0.949247891**	**0.016893552**	**0.033858557**
V148	**Diploid male**	**0.896337894**	**0.003522837**	**0.100139269**
V149	**Diploid male**	**0.91897525**	**0.016193354**	**0.064831396**
V150	Triploid	0.160448802	0.300856266	0.538694932
V151	**Triploid**	**0.009665379**	**0.071856317**	**0.918478304**
V152	**Diploid male**	**0.665804661**	**0.054538253**	**0.279657086**
V153	**Triploid**	**0.052286669**	**0.149643981**	**0.79806935**
V154	**Triploid**	**0.031028721**	**0.247491137**	**0.721480142**
V155	**Triploid**	**0.096445334**	**0.291030307**	**0.612524359**
V156	**Triploid**	**0.084568664**	**0.136856345**	**0.778574991**
V157	Diploid male	0.506989993	0.120681477	0.37232853
V158	**Diploid male**	**0.903203594**	**0.04945503**	**0.047341376**
V159	**Triploid**	**0.008789762**	**0.160542183**	**0.830668055**
V160	Triploid	0.242385136	0.164467662	0.593147202
V161	Diploid female	0.080048519	0.591340482	0.328610999
V162	**Triploid**	**0.095634046**	**0.153608809**	**0.750757145**

Specimen	Predicted ploidy	Diploid male probability	Diploid female probability	Triploid probability
V163	Diploid male	0.943443167	0.024217584	0.032339249

Guadiana by body shape

The following table shows the probability of prediction for fish that were predicted using the linear discriminant functions for body shape, as described in Chapter 4.3.1.

Table 2A.2: Prediction values and probabilities derived from analysis of body shape. Individuals for which prediction probability is less than 60% are in bold.

Specimen	Predicted ploidy	Diploid male probability	Triploid probability
V6	Triploid	2.325712^{-5}	9.999767^{-1}
V7	Triploid	1.186365^{-8}	1.000000
V9	Triploid	1.660314^{-5}	9.999834^{-1}
V17	Triploid	1.215145^{-2}	9.878485^{-1}
V20	Triploid	1.959865^{-3}	9.980401^{-1}
V22	Triploid	2.098933^{-8}	1.000000
V24	Triploid	8.806557^{-5}	9.999119^{-1}
V25	Triploid	1.444322^{-8}	1.000000
V29	Triploid	4.149547^{-4}	9.995850^{-1}
V45	Triploid	7.968583^{-3}	9.920314^{-1}
V46	Triploid	2.077433^{-5}	9.999792^{-1}
V47	Triploid	3.058727^{-10}	1.000000
V51	Triploid	5.895374^{-6}	9.999941^{-1}
V54	Triploid	1.672338^{-2}	9.832766^{-1}
V57	Triploid	2.715970^{-3}	9.972840^{-1}
V58	Diploid male	9.999991^{-1}	9.371069^{-7}
V66	Triploid	9.392420^{-6}	9.999906^{-1}
V74	Diploid male	9.996891^{-1}	3.109034^{-4}
V85	Triploid	0.004332876	0.99566712
V88	Triploid	0.021403234	0.97859677
V91	Triploid	0.279861989	0.72013801
V96	Diploid male	0.938199796	0.06180020

Specimen	Predicted ploidy	Diploid male probability	Triploid probability
V98	Triploid	0.003789328	0.99621067
V101	**Diploid male**	**0.556503125**	**0.44349687**
V106	Triploid	0.006590474	0.99340953
V108	Triploid	0.018274874	0.98172513
V113	Triploid	0.007153055	0.99284694
V120	Diploid male	0.903695382	0.09630462
V123	Triploid	0.004657150	0.99534285
V126	Diploid male	0.989770431	0.01022957
V127	Triploid	0.011546045	0.98845395
V131	Triploid	0.048924626	0.95107537
V132	Triploid	0.021017622	0.97898238
V134	Triploid	0.001043562	0.99895644
V136	Triploid	0.006449394	0.99355061
V145	Diploid male	0.989103948	0.01089605
V150	Triploid	0.001125741	0.99887426
V157	Triploid	0.139701484	0.86029852
V160	Diploid male	0.849280324	0.15071968
V161	Triploid	0.048878001	0.95112200

Guadiana by head shape

The following table shows the probability of prediction for fish that were predicted using the linear discriminant functions for head shape, as described in Chapter 4.3.2.

Table 2A.3: Prediction values and probabilities derived from analysis of head shape. Individuals for which prediction probability is less than 60% are in bold.

Specimen	Predicted ploidy	Diploid male probability	Triploid probability
V6	Diploid male	0.85959759	0.1404014
V7	Triploid	0.04800345	0.9519965
V9	Triploid	0.08853453	0.9114655
V17	Triploid	0.03576680	0.9642332
V20	Triploid	0.23254752	0.7674525
V22	Triploid	0.04228175	0.9577182
V24	Triploid	0.05579993	0.9442001
V25	Triploid	0.03370719	0.9662982
V29	Triploid	0.11007008	0.8899299
V45	Triploid	0.17324554	0.8267545
V46	Triploid	0.04286455	0.9571354
V47	Triploid	0.01167013	0.9883299
V51	**Triploid**	**0.49972248**	**0.5002775**
V54	Triploid	0.27523561	0.7247644
V57	Triploid	0.03820763	0.9617924
V58	Triploid	0.13165762	0.8683424
V66	Triploid	0.15043015	0.8495698
V74	Triploid	0.01215810	0.9878419
V85	Triploid	0.30576174	0.694238262
V88	Triploid	0.01903969	0.980960315
V91	Diploid male	0.92942335	0.070576648
V96	Diploid male	0.97822356	0.0212776442
V98	Triploid	0.01378669	0.986213315

Specimen	Predicted ploidy	Diploid male probability	Triploid probability
V101	Triploid	0.34182462	0.6558175375
V106	Triploid	0.03308192	0.966918085
V108	Triploid	0.07915749	0.920842507
V113	Triploid	0.05741096	0.942589038
V120	Triploid	0.04193356	0.958066439
V123	Triploid	0.08773773	0.912262271
V126	Diploid male	0.98892033	0.011079670
V127	Diploid male	0.95457108	0.045428916
V131	Triploid	0.01272256	0.987277440
V132	Triploid	0.34686633	0.653133673
V134	Triploid	0.02333340	0.976666598
V136	Triploid	0.07635730	0.923642697
V145	Diploid male	0.99728647	0.002713528
V150	Triploid	0.0664763	0.933526367
V157	**Diploid male**	**0.53928520**	**0.460714799**
V160	Diploid male	0.98078656	0.019213445
V161	Triploid	0.10397766	0.896022336