

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

Departamento de Biologia Animal



Reproductive behaviour and the evolutionary history of
the hybridogenetic complex *Squalius alburnoides*
(Pisces, Cyprinidae)

CARLA PATRÍCIA CÂNDIDO DE SOUSA SANTOS

Doutoramento em Biologia

(Biologia Evolutiva)

2007

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

Departamento de Biologia Animal



Reproductive behaviour and the evolutionary history of
the hybridogenetic complex *Squalius alburnoides*
(Pisces, Cyprinidae)

CARLA PATRÍCIA CÂNDIDO DE SOUSA SANTOS

Tese co-orientada por:

Professor Doutor Vítor Almada (ISPA)

Professora Doutora Maria João Collares-Pereira (FCUL)

Doutoramento em Biologia

(Biologia Evolutiva)

2007

Os trabalhos de investigação realizados durante esta tese tiveram o apoio financeiro da Fundação para a Ciência e a Tecnologia (Bolsa de Doutoramento SFRH/BD/8320/2002 e Programas Plurianuais U I&D 331/94 e U I&D 329/94).

The research conducted in this thesis was funded by the FCT Pluriannual Program U I&D 331/94 and U I&D 329/94 (FEDER participation) and by a PhD grant from FCT (SFRH/BD/8320/2002).

Nota prévia

Na elaboração da presente dissertação foram integralmente utilizados artigos já publicados em revistas internacionais indexadas, à excepção de apenas um artigo que se encontra submetido para publicação (ver “List of Papers”). Uma vez que os referidos trabalhos foram realizados em colaboração com outros investigadores, e nos termos do Artigo 15º, Capítulo I, do Regulamento de Doutoramento da Universidade de Lisboa publicado no Diário da República – II série Nº 194 de 19-08-1993, a autora esclarece que participou em todas as fases de elaboração dos artigos: planeamento, obtenção e análise de dados, interpretação dos resultados e redacção do manuscrito final. Relativamente ao único artigo em que não é primeira autora, participou activamente na obtenção e análise de dados e na interpretação dos resultados.

Preliminary note

This thesis comprises papers that were published or submitted for publication in indexed international journals (see “List of Papers”). Since these papers were done in collaboration with other investigators, and according to specific regulations of the University of Lisbon, the author states that she has participated in all stages of the elaboration of the papers: planning, data collection, data analysis, interpretation of the results and in the writing of the final manuscripts. In the only paper in which she was not the first author, she actively participated in data collection, data analysis and in the interpretation of the results.

À minha família,
porque são as raízes que mantêm as árvores de pé.

Contents

Acknowledgements	iv
Resumo	v
Summary	x
List of papers	xi

PART I - GENERAL INTRODUCTION

Chapter 1 - Introduction	1
1.1 Hybridization: a historical perspective	2
1.2 Outcomes of hybridization	3
1.2.1 Polyploidy	3
1.2.2 Asexuality	5
1.2.3 Reticulate evolution	6
1.2.4 Introgression	7
1.3 Hybridization and allopolyploidy in vertebrates	7
1.4 Reproduction in hybrid fish complexes	9
1.5 The case-study of <i>S. alburnoides</i>	12
1.5.1 General description	12
1.5.2 Origin and dispersal	14
1.5.3 Reproductive dynamics	15
1.5.4 Biology and ecology	17
1.6 Aims and structure of the thesis	18
1.7 References	21

PART II – METHODOLOGICAL TOOLS FOR GENOTYPE ASSESSMENT

Chapter 2

Heterozygous indels as useful tools in haplotype reconstruction of DNA nuclear genes of hybrid organisms – an example using cyprinid fishes. *DNA Sequence* 16(6): 462-467 (2005) 31

Chapter 3

Phenotypic expression of different genome constitutions in a diploid-polyploid hybridogenetic complex. *Journal of Fish Biology*(submitted) 45

PART III – REPRODUCTION

Chapter 4

May a hybridogenetic complex regenerate the nuclear genome of both sexes of a missing ancestor? - First evidence on the occurrence of a nuclear nonhybrid *Squalius alburnoides* (Cyprinidae) female based on DNA sequencing. *Journal of Natural History* 40(23-24): 1443-1448 (2006) 71

Chapter 5

Fertile triploid males – an uncommon case among hybrid vertebrates. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* 307A(4): 220-225 (2007) 83

Chapter 6

Reproductive success of nuclear nonhybrid males of *Squalius alburnoides* hybridogenetic complex (Teleostei, Cyprinidae): an example of interplay between female choice and ecological pressures? *Acta Ethologica* 9(1): 31-36 (2006) 97

PART IV – ORIGIN, DISPERSAL AND INTROGRESSION

Chapter 7

Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. *Molecular Phylogenetics and Evolution* 39(1): 276- 113

Chapter 8

Reading the history of a hybrid fish complex from its molecular record. *Molecular Phylogenetics and Evolution* (in press) 127

Chapter 9

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. *Journal of Fish Biology* 68 (Supplement B): 292-301 (2006) 185

PART V – DISCUSSION AND CONCLUDING REMARKS

Chapter 10 – General Discussion 199

10.1 Identification of inherited genomes	200
10.2 Origin of the <i>S. alburnoides</i> complex	206
10.2.1 The identity of the paternal ancestor	206
10.2.2 Paleobiogeography of the ancestors	207
10.2.3 Single origin – where and why?	209
10.2.4 Unidirectionality of interspecific crosses – possible causes	213
10.2.5 Origin of the various forms of the complex	215
10.2.6 The extinction of the paternal ancestor	218
10.3 Foundation of the <i>S. alburnoides</i> Portuguese populations	222
10.3.1 Patterns of dispersal of the complex	222
10.3.2 Introgression of genes from other <i>Squalius</i> species	224
10.4 Maintenance and present day dynamics of the complex	227
10.4.1 Intraspecific reproductive behaviour	227
10.4.2 Interspecific reproductive behaviour	231
10.4.3 Variations among populations	232
10.4.4 Implications for conservation	241
10.5 Sexual parasitism or sexual autonomy in <i>S. alburnoides</i> ?	242
10.6 References	245

Chapter 11 – Concluding Remarks 255

Acknowledgements

This work would not have been possible without the help and support of many people. To some, for their crucial role, I would like to express, in my mother language, my sincere gratitude:

Ao Professor Doutor Vítor Almada (UIE-ISPA) e à Professora Doutora Maria João Collares-Pereira (FCUL) pelos conhecimentos transmitidos e pela disponibilidade, constante encorajamento, confiança e dedicação com que orientaram esta tese.

À Mestre Joana Robalo (UIE-ISPA) pela amizade e pelo entusiasmo na partilha deste percurso.

Aos colegas da UIE-ISPA Ana Pereira, André Levy, Clara Amorim, Cláudia Faria, Sara Francisco e Vera Domingues, pelas ajudas e trocas de ideias.

Aos meus grandes amigos de sempre.

Aos meus pais, mano e restante família, pela paciência, encorajamento e apoio incondicional.

Ao Gonçalo por boiar comigo, de cara ao sol, quando o mar está calmo, e por me trazer pela mão à superfície para me ajudar a respirar quando o mar está revolto.

Ao pequeno Gil, que veio mudar tudo, para (ainda) melhor.

Resumo

O complexo hibridogenético *Squalius alburnoides* (Pisces: Cyprinidae) constitui um exemplo bem sucedido no conjunto de vertebrados de origem híbrida da possibilidade de certas linhagens híbridas poderem compensar as desvantagens da assexualidade, nomeadamente a perda de variabilidade genética.

A origem deste complexo endémico da Península Ibérica deve-se a cruzamentos entre fêmeas de *Squalius pyrenaicus* (genoma P) e um ancestral paterno até agora desconhecido e virtualmente extinto (genoma A), dos quais terão resultado os primeiros híbridos (diplóides PA). Estes, cruzando-se entre si e com as espécies parentais, geraram mais formas híbridas que diferem na sua ploidia (triplóides e tetraplóides) e na proporção de genomas das espécies parentais. Do cruzamento entre as formas híbridas resultou ainda uma linhagem constituída quase exclusivamente por machos com genoma nuclear não-híbrido (AA). No entanto, as várias populações de *S. alburnoides* são distintas no que se refere ao tipo de formas representadas e à sua percentagem relativa de ocorrência, para além de existirem, fora da área de distribuição da espécie materna (*S. pyrenaicus*), em simpatria com outras espécies de *Squalius* (*S. carolitertii* e *S. aradensis*), o que pode contribuir para um acréscimo de diferenciação populacional ao nível mitocondrial e nuclear.

Esta possibilidade de cruzamentos interespecíficos depende, em larga medida, da dinâmica dos comportamentos reprodutores das espécies envolvidas. Do mesmo modo, as diversas formas de ploidias e constituições genómicas distintas podem diferir nos seus comportamentos reprodutores nomeadamente no que se refere às escolhas de parceiros, afectando as proporções dos diferentes tipos de cruzamentos e, portanto, a proporção das formas geradas. Partindo destas hipóteses, a presente tese teve como principal objectivo estudar o contributo do comportamento reprodutor intra e interespecífico não só para o actual sucesso ecológico como também para a história evolutiva do complexo desde a sua origem. Para tal, a investigação desenvolvida teve como objectivos específicos: 1) procurar marcadores nucleares adequados para a identificação das constituições genómicas individuais; 2) investigar a relação entre a constituição genómica e o fenótipo das várias formas do complexo; 3) descrever o comportamento reprodutor de *S. alburnoides* e investigar a existência de escolhas de parceiro; 4) contribuir para um melhor conhecimento da constituição genómica das formas do complexo e do seu ancestral paterno; 5) confirmar e aprofundar o conhecimento acerca dos modos de reprodução dos machos de *S. alburnoides*; 6) avaliar a ocorrência de hibridação natural entre machos de *S. alburnoides* e

fêmeas de outras espécies de *Squalius* e suas implicações para a dinâmica genética do complexo; 7) discutir as consequências da introgressão recíproca entre *S. alburnoides* e outras espécies de *Squalius*, e 8) analisar possíveis localizações para a origem do complexo e traçar rotas de dispersão plausíveis que expliquem a sua actual distribuição geográfica.

Com vista ao cumprimento destes objectivos, foi efectuada uma significativa amostragem de *S. alburnoides* ao longo da sua área de distribuição em Portugal, cobrindo um total de seis bacias hidrográficas. Foram igualmente capturados exemplares de outras espécies de *Squalius* em vários rios nacionais para que o estabelecimento de comparações nos estudos morfológico e filogeográfico de *S. alburnoides* fosse possível. Sequências adicionais do gene mitocondrial do citocromo *b* já publicadas e depositadas no GenBank pertencentes a *Squalius* de populações nacionais e espanholas foram também utilizadas de modo a determinar com maior precisão a diversidade global das populações ibéricas.

A combinação de haplótipos parentais envolvidos na formação dos genomas híbridos, a constituição genética e a ploidia de cada indivíduo (informações essenciais para a prossecução dos estudos morfológico, etológico e filogeográfico e para a reconstrução da história evolutiva do complexo) foram determinadas com o auxílio de um novo método baseado na análise dos picos duplos gerados pela presença de indels heterozigóticos nos cromatogramas resultantes da sequenciação do gene nuclear da beta-actina. A validação do método desenvolvido foi efectuada recorrendo a medições da quantidade de ADN nuclear por citometria de fluxo e à análise de híbridos criados artificialmente por mistura de quantidades conhecidas de ADN de espécies diferentes.

O estudo dos aspectos etológicos da reprodução de *S. alburnoides* foi efectuada com exemplares provenientes da Bacia do Rio Tejo usando uma abordagem experimental semi-natural, embora tenham sido igualmente realizadas experiências com exemplares de outras bacias (*S. alburnoides* dos Rios Douro e Guadiana e *Squalius torgalensis* do Rio Mira).

De forma resumida, as principais conclusões dos estudos que integram a presente dissertação foram as seguintes:

a) As formas de *S. alburnoides* cuja morfologia foi analisada (AA, PA, PAA e PPA) apresentam diferenças significativas entre si. Foi também demonstrado que as formas híbridas expressam ambos os genomas parentais (A e P) e que existe um contínuo morfológico, ocupando as forma parentais AA e PP os extremos e as formas híbridas uma posição intermédia (PAA e PPA são as formas mais próximas de AA e PP, respectivamente). Estas observações sugerem a existência de um efeito de dosagem no fenótipo que se traduz, por exemplo, no aumento da frequência de dentes biseriados (característicos do ancestral materno) com o aumento da proporção de genomas P nos híbridos. A existência

de diferenças morfológicas e de coloração poderá estar na origem de segregação espacial entre as várias formas de *S. alburnoides*, influenciando a probabilidade de ocorrência de certos cruzamentos em detrimento de outros e, conseqüentemente, a dinâmica evolutiva do complexo;

b) Em condições semi-naturais, as fêmeas de *S. alburnoides* tendem a preferir machos não-híbridos (que, apesar de mais pequenos, apresentam cortes mais vigorosas) em situações em que machos híbridos receptivos estão igualmente disponíveis. Esta preferência por parte das fêmeas poderá ter efeitos importantes na dinâmica genética do complexo: 1) tendo uma taxa de fertilizações superior à dos machos híbridos, os machos não-híbridos garantem a sua sobrevivência (sendo o resultado de um único tipo de cruzamento - entre fêmeas PAA e machos iguais a eles próprios, AA – uma vez extintos de uma população não voltam a ser regenerados); 2) os cruzamentos entre machos AA e fêmeas PAA são os que contribuem mais para aumentar a variabilidade genética do complexo dado que, para além dos tetraplóides balanceados que são extremamente raros, são as únicas formas que produzem gâmetas após recombinação;

c) Num contexto interespecífico, os machos não-híbridos revelaram-se bastante activos na corte de fêmeas de *S. torgalensis* e assumiram uma estratégia de *sneaking* na presença de pares de *S. pyrenaicus* em fase de pré-postura. Este comportamento apresenta igualmente implicações importantes para a dinâmica do complexo: a fertilização de ovos de outras espécies de *Squalius* resulta na introgressão de novos genes em *S. alburnoides*. O papel dos machos não-híbridos como motores de introgressão de novos genes no complexo foi corroborado pelo facto de terem sido detectados níveis de introgressão superiores em populações com machos não-híbridos, quando comparados com os níveis apresentados por populações sem este tipo de machos. Contrariando estudos anteriores que apontavam para uma quase uniformidade genética a nível mitocondrial, ficou igualmente demonstrado que a introgressão de genes de espécies diferentes de *S. pyrenaicus* é muito significativa em todas as populações estudadas;

d) A sequenciação do gene nuclear da beta-actina permitiu a descoberta de uma fêmea não híbrida (com o genoma parental em homozigotia – AA) cuja ocorrência deverá ser extremamente rara (uma vez que até ao momento só havia conhecimento de um outro caso isolado) mas que chama a atenção para a possibilidade de re-emergência de uma espécie extinta a partir dos seus descendentes híbridos ainda que possuindo genes mitocondriais de outra espécie (*S. pyrenaicus*);

e) A aplicação do método desenvolvido com base na análise dos picos duplos que surgem nos cromatogramas das sequências do gene da beta-actina de híbridos permitiu a

confirmação da produção de gametas haplóides e diplóides pelos machos diplóides AA e PA, respectivamente, e a primeira demonstração da produção de espermatozoides clonal viável e funcional num macho triplóide PAA;

f) O mesmo marcador nuclear permitiu confirmar que o ancestral paterno do complexo terá sido uma espécie próxima de *Anaocypris hispanica* (como havia sido anteriormente proposto), muito provavelmente já extinta e pertencente à linhagem Europeia de *Alburnus*;

g) O estudo filogeográfico revelou que a origem do complexo deverá ter sido posterior à diferenciação das outras espécies de *Squalius*, que terão tido uma radiação em estrela a partir de um clado central e amplamente distribuído de *S. pyrenaicus*. A radiação dos clados derivados periféricos (*S. aradensis*, *S. torgalensis*, *S. carolitertii*, *S. valentinus* and *S. malacitanus*) terá ocorrido entre o Miocénico e o Pliocénico, aspecto corroborado por dados relativos à paleohidrografia da Península Ibérica;

h) A hipótese de *S. alburnoides* ter tido origens múltiplas e independentes não foi confirmada, tendo sido proposta uma origem única, datada do Pleistocénico superior (há cerca de 700.000 anos). Muito provavelmente as hibridações iniciais terão ocorrido quando a Bacia Hidrográfica do Tejo-Guadiana (na altura ainda não isoladas entre si e transportando o ancestral materno do complexo já diferenciado - *S. pyrenaicus*) capturou uma bacia endorreica onde o ancestral paterno estaria isolado. As hibridações intergenéricas que deram origem ao complexo terão, portanto, surgido na sequência dos rearranjos paleohidrográficos despoletados pela basculação para oeste da Península Ibérica. O facto de nunca terem sido detectados genes mitocondriais da espécie paterna significa que, muito provavelmente, esses cruzamentos intergenéricos foram assimétricos desde a origem do complexo, ou seja, os cruzamentos entre fêmeas de *S. pyrenaicus* com machos da espécie paterna terão sido sistematicamente mais frequentes do que os cruzamentos inversos. Esta hipótese foi corroborada pela mencionada estratégia de *sneaking* que os machos não-híbridos adoptam na presença de pares de *S. pyrenaicus*. Esta poderá, assim, ser uma característica ancestral presente na espécie paterna que terá reemergido quando a forma não-híbrida AA foi reconstituída a partir das formas híbridas de *S. alburnoides*. Após a origem do complexo, o ancestral paterno terá sofrido declínios consideráveis (por escolhas de parceiro desfavoráveis, alterações drásticas de habitat, condições climáticas desfavoráveis e/ou competição directa com os híbridos) que terão conduzido muito provavelmente à sua extinção, eliminando irreversivelmente a incorporação de genes mitocondriais típicos do ancestral paterno no complexo;

i) Segundo o cenário mais parcimonioso, a dispersão de *S. alburnoides* terá ocorrido depois de se terem originado todas as formas do complexo e após a extinção do ancestral paterno. O complexo terá dispersado para norte para as Bacias dos Rios Mondego e Douro através de duas vias independentes, ambas com origem no Tejo; para sudoeste para o Sado, também via-Tejo; e para sul para o Guadalquivir e para a Ribeira da Quarteira, através das secções média e inferior da Bacia do Guadiana, respectivamente;

j) À medida que o complexo dispersou para bacias hidrográficas fora da área de distribuição do ancestral materno *S. pyrenaicus*, estabeleceu contacto com outras espécies de *Squalius* e, conseqüentemente, novos genes mitocondriais e nucleares foram incorporados na sequência de cruzamentos interespecíficos. Ao contrário da introgressão primordial de genes mitocondriais no complexo, que parece ter sido quase exclusivamente unidireccional, a introgressão actual de genes nucleares é bidireccional, o que tem importantes conseqüências relativamente ao estatuto taxonómico das outras espécies de *Squalius* que são simpátricas com o complexo. De facto, as fêmeas PPA, ao produzirem óvulos haplóides com um genoma P, podem ser responsáveis pela introgressão de genes de *S. pyrenaicus* em populações de *S. carolitertii*, contribuindo para a homogeneização fenotípica e genotípica das duas espécies.

Os resultados obtidos permitem concluir que o comportamento reprodutor terá tido um papel crucial na origem de *S. alburnoides* e que, ao longo da sua história evolutiva, terá sido determinante na estruturação das populações ao afectar toda a dinâmica de circulação de genes não apenas dentro do complexo mas também entre o complexo e as espécies com que interage.

Finalmente, apesar da existência de processos gametogénicos atípicos e dos rácios sexuais das populações serem extremamente enviesadas a favor das fêmeas, a reprodução de *S. alburnoides* envolve sempre a corte entre indivíduos férteis de sexos opostos e a união de óvulos e esperma, não sendo estritamente dependente de outras espécies de *Squalius*, para além dos gâmetas serem muitas vezes produzidos após recombinação. Assim sendo, não se poderá dizer que o complexo é verdadeiramente sexuado nem assexuado. Alternativamente, e na tentativa de evitar rótulos restritivos ou erróneos, propõe-se que *S. alburnoides* seja referido como sendo um “complexo hibridogenético” ou um “complexo híbrido intergenérico” que apresenta mecanismos gametogénicos alternativos à meiose normal.

Palavras-chave comportamento reprodutor, filogeografia, introgressão, hibridação intergenérica, origem única, poliploidia, *Squalius* ibéricos.

Summary

The *Squalius alburnoides* complex is an exception among hybrid vertebrates, comprising diploid and polyploid forms of both sexes which often produce gametes after recombination. This dissertation aimed to investigate the contribution of the reproductive behaviour to its remarkable evolutionary history. Related issues were also addressed, namely, the identification of the genomic constitutions of the forms of the complex and of its paternal ancestor, the genotype-phenotype relationship, the type of sperm and fertility of triploid males, the hybridization with other *Squalius*, and the phylogeography of the complex. The parental haplotypes found in hybrid genomes, genomic constitutions and ploidies were assessed with a validated method based on the analysis of beta-actin nuclear gene sequences.

Results showed that: 1) the studied *S. alburnoides* forms were morphologically distinct; 2) the phenotypic expression of both parental genomes was dosage related; 3) *S. alburnoides* females preferred to mate with nonhybrid males, increasing the genetic variability of the complex and ensuring the persistence of nonhybrids; 4) nonhybrid males act as sneakers of *S. pyrenaicus* pairs, which may explain some detected high introgression levels; 5) a sneaking male behaviour may have caused the original hybridization events, explaining the virtual absence of mtDNA of the paternal ancestor; 6) introgression of *S. alburnoides* genes in other *Squalius* also occur; and 7) the complex is likely to have had a single Pleistocenic origin, when paleohydrographical rearrangements allowed the contact between its ancestors. Also, the beta-actin gene sequencing revealed: that the paternal ancestor of *S. alburnoides* was closely related to *A. hispanica*, the occurrence of an extremely rare non-hybrid female and also the production of clonal sperm by PAA males.

Therefore, the reproductive behaviour seems to have been determinant to the origin of *S. alburnoides* and, subsequently, to population structuring by influencing the genetic dynamics not only inside the complex but also between the complex and the *Squalius* species with which it has been interacting.

Keywords reproductive behaviour, phylogeography, introgression, intergeneric hybridization, polyploidy, single origin, Iberian *Squalius*.

List of Papers

Each published or submitted paper listed below corresponds to a Chapter of the present thesis.

1. **Sousa-Santos C**, Robalo JI, Collares-Pereira MJ & Almada VC. 2005. Heterozygous indels as useful tools in haplotype reconstruction of DNA nuclear genes of hybrid organisms – an example using cyprinid fishes. *DNA Sequence* 16(6): 462-467.
2. **Sousa-Santos C**, Collares-Pereira MJ & Almada VC. submitted. Phenotypic expression of different genome constitutions in a diploid-polyploid hybridogenetic complex. *Journal of Fish Biology*.
3. **Sousa-Santos C**, Collares-Pereira MJ & Almada VC. 2006. May a hybridogenetic complex regenerate the nuclear genome of both sexes of a missing ancestor? - First evidence on the occurrence of a nuclear nonhybrid *Squalius alburnoides* (Cyprinidae) female based on DNA sequencing. *Journal of Natural History* 40(23-24): 1443-1448.
4. **Sousa-Santos C**, Collares-Pereira MJ & Almada VC. 2007. Fertile triploid males – an uncommon case among hybrid vertebrates. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* 307A(4): 220-225.
5. **Sousa-Santos C**, Collares-Pereira MJ & Almada VC. 2006. Reproductive success of nuclear nonhybrid males of *Squalius alburnoides* hybridogenetic complex (Teleostei, Cyprinidae): an example of interplay between female choice and ecological pressures? *Acta Ethologica* 9(1): 31-36.
6. Robalo JI, **Sousa-Santos C**, Levy A & Almada VC. 2006. Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. *Molecular Phylogenetics and Evolution* 39(1): 276-281.
7. **Sousa-Santos C**, Collares-Pereira MJ & Almada VC. in press. Reading the history of a hybrid fish complex from its molecular record. *Molecular Phylogenetics and Evolution*.
8. **Sousa-Santos C**, Collares-Pereira MJ & Almada VC. 2006. 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. *Journal of Fish Biology* 68 (Supplement B): 292-301.

PART I

GENERAL INTRODUCTION

Chapter 1

Introduction

"The view generally entertained by naturalists is that species, when intercrossed, have been specially endowed with the quality of sterility, in order to prevent the confusion of all organic forms. [...] Looking to all the ascertained facts on the intercrossing of plants and animals, it may be concluded that some degree of sterility, both in first crosses and in hybrids, is an extremely general result; but that it cannot, under our present state of knowledge, be considered as absolutely universal."

Darwin (1859)

Almost 150 years after "On the Origin of Species", the accumulated knowledge does not contradict the view that animal hybrids are generally sterile although the same is not applied to plants, for which there is a high number of known fertile cases (Tate *et al.* 2005). However, curious examples of fertile hybrid animals have also been described even though in many cases their hybrid condition is coupled, among others, with asexuality¹, polyploidization of the genome, short evolutionary histories, dependency on

¹ The formerly used term "unisexuality" has been progressively replaced by "asexuality" by most authors and, thus, it will also be adopted in the present dissertation to designate the mode of reproduction exhibited by an organism or group of organisms that reproduce without sex, with or without syngamy and karyogamy, and without recombination (or occurring only rarely), as described by Beukeboom & Vrijenhoek (1998).

bisexual species for fertilization, inability to perform normal meiosis and/or impaired capacity of recombination (*e.g.* Dawley 1989; Avise *et al.* 1992).

One of the exceptions to sterility in hybrid animals is the extremely well succeeded *Squalius alburnoides* hybridogenetic complex, which comprises fish of both sexes with varying ploidy levels and proportions of the parental genomes. This complex is also independent from bisexual hosts (although crosses with bisexual species may occur) and recombination is known to occur in some of the forms.

The present dissertation intended to clarify the reproductive dynamics of the *S. alburnoides* complex, not only with conspecifics but also with members of other *Squalius* species, aiming to trace back its evolutionary history and to draw hypotheses to explain the high success of the complex, an uncommon feature in the context of polyploid complexes of hybrid origin. The presentation of the results and conclusions of the research conducted to achieve these major goals will be preceded by a description of the present knowledge on hybridization and of its outcomes as well as by a detailed characterization of the *S. alburnoides* complex. As studies on this complex have been abundant, there was the need to include also references to the papers presented ahead in this dissertation in order to chronologically characterize the state of the art.

1.1 Hybridization: a historical perspective

As hybridization has been considered to be a rare event in animal evolution, the importance of viable fertile hybrids has been frequently neglected and the creative role of hybridization in evolution has been a controversial matter (*e.g.* Arnold 1997; Vrijenhoek 2006). Since the 30s, when the process of natural hybridization began to receive more attention from the researchers (Arnold 1997), two divergent approaches emerged: botanists highlighted its potential for generating diversity (hybrids could occupy new habitats and originate new clades) and zoologists tended to see it as a reproductive mistake that limits diversification and retards evolution (reviewed by Barton 2001, Seehausen 2004 and Mallet 2005). According to the later view, hybridization is the

converse of reproductive isolation and challenges the biological species concept, thus the study of hybrids would only be relevant as a tool to understand the development of reproductive isolation (Arnold 1997; Mallet 2005).

Moreover, hybrids were seen as less fit organisms, being in general sterile or possessing low viability. Although some sexually reproducing polyploid species have been identified, Schultz (1969) proposed the existence of a close link between polyploidy, hybridization and asexuality. Indeed, all naturally occurring asexual vertebrates are of hybrid origin (Avisé *et al.* 1992; Schmidt 1996; Dowling & Secor 1997; Adams *et al.* 2003) and they were presumed to display a lack of genetic variability that due to their asexual (*i.e.*, non-recombinant or clonal, as defined by Beukeboom & Vrijenhoek 1998) modes of reproduction produced fewer and/or less fertile and less diverse progeny when compared to their parental species. Hence, hybrids were considered to be “evolutionary dead-ends”, “blind alleys” and “no-hopers” (reviewed in Wetherington *et al.* 1987). However, this viewpoint has progressively changed as new molecular markers exposed the important role of hybridization as a source of genetic variation, a way to generate functional novelty and adaptability, and an important mechanism in the formation of new species (*e.g.* Rieseberg *et al.* 1999; Seehausen 2004; Kearney 2005; Mallet 2005).

1.2 Outcomes of hybridization

1.2.1 Polyploidy

Polyoids are organisms that have multiple complete sets of chromosomes, generally resulting from the multiplication of one chromosome set within a species (autopolyploids) or from the merging of different chromosome sets from two more or less related hybridizing species (allopolyploids). Additionally, a third conceptual category was created to designate hybrids that, in contrast to what occurs in allopolyploids, exhibit the same number of chromosomes of their parental species: homoploids (Seehausen 2004). As it was the case specifically analysed in the present dissertation, allopolyploidy will be addressed in more detail.

Some reasons have been pointed out to explain the lower occurrence of polyploidy in animals (in the case of vertebrates, mostly fishes and amphibians) when compared to that of plants (reviewed in Schultz 1969 and in Gregory & Mable 2005). The first one is related to the disruption of sex determination: with the exception of hermaphrodites, animals do not have both sexes in the same individual like many plants and, thus, the formation of stable populations of autopolyploids would require an unlikely simultaneous occurrence of genome duplication in both gametes. As regards allopolyploidy, as sex is often chromosomally determined in vertebrates, the hybridization may well result in the disruption of the development and/or fertility of one of the sexes (Gregory & Mable 2005). Indeed, the combination of distinct genomes alters the gametogenesis so that unreduced eggs and sperm are produced and the sex ratio is commonly skewed towards females. Also, according to Muller (1925), a certain proportion of genes on sex chromosomes in relation to those on autosomes seems to be an essential condition for sex determination in animals. The disruption of this ratio by polyploidization may lead to populations constituted by only one fertile - but mateless - sex (Gregory & Mable 2005).

Second, it is difficult for a newly formed animal polyploid to become reproductively isolated from its diploid parents (in plants this may be overcome, for instance, by a different flowering time or ecological habitat of the polyploids) (Gregory & Mable 2005).

Regarding the process of polyploid formation, a first step is generally the production of unreduced, rather than haploid, gametes as a result of several possible meiotic errors (Dawley 1989). The union of two unreduced gametes would result in a tetraploid individual with an even number of chromosome sets, which would likely perform normal meiosis. However, it is very unlikely that meiotic errors would stochastically emerge simultaneously in both sexes, leading to the production of unreduced eggs and sperm. Instead, the most common situation is that the unreduced eggs or sperm are fertilized by haploid gametes, originating triploids.

Unlike what happens in autopolyploids, in the case of allopolyploids and homoploids the existence of chromosome sets derived from distinct species may disrupt meiosis and lead to partial or to total infertility of the descendants (Vrijenhoek 1989, 1994). This problem is particularly relevant in triploids, since the lack of homologous pairs for one

set of chromosomes precludes the formation of bivalents necessary for normal segregation, a situation that often leads to sterility (Van de Peer & Meyer 2005).

However, if triploids produce fertile unreduced gametes, they might configure a “triploid bridge” – *i.e.*, an intermediate stage between diploidy and tetraploidy as firstly hypothesized by Schultz (1969) – since, if backcrosses with the diploid species occur, viable tetraploid offspring may be generated (Dawley 1989). Alternatively, triploid hybrids may also reproduce gynogenetically (*Poeciliopsis* complex - Schultz 1967), hybridogenetically (*Rana esculenta* - Vinogradov *et al.* 1990; *Squalius alburnoides* - Alves *et al.* 2001; *Cobitis hankugensis-longicorpus* - Saitoh *et al.* 2004) or, more surprisingly, bisexually (*Bufo pseudoraddei baturae* - Stöck *et al.* 2002).

1.2.2 Asexuality

Some hybrids overcome their inability to perform normal meiosis by reproducing by altered gametogenic mechanisms, transmitting at least part of their genome without or with limited recombination to the offspring. These special hybrids have been named as “unisexuals” or “asexuals” and their alternative modes to sexual reproduction are known to have independently evolved many times as a result of hybridization events between distinct species (reviewed in Dawley & Bogart 1989). Although asexuals are generally constituted by all-female populations that reproduce without recombination, exceptions to this general definition are known: lineages that incorporate sperm from related males, that show limited recombination and/or that also include males (Dawley 1989). The application of the term “asexual” to these exceptions, especially when they involve the presence of both sexes and the occurrence of normal meiotic recombination is, however, problematic and will be discussed in Chapter 10.

The establishment and survival of these organisms depends on clonal reproductive modes such as parthenogenesis or gynogenesis; on the incorporation of unusual meiotic mechanisms (like meiotic hybridogenesis) that allow recombination to occur in asexual females, often with the exclusion of the heterospecific genetic complement; and/or on the ploidy elevation by the incorporation of additional genomic material (processes detailed in Dawley 1989; Vrijenhoek 1989; Avise *et al.* 1992; Vrijenhoek

1994; Alves *et al.* 1998; Beukeboom & Vrijenhoek 1998; Alves *et al.* 2001; Scali *et al.* 2003; Bi & Bogart 2006). A brief description of the known asexual modes of reproduction is presented below (Table 1).

1.2.3 Reticulate evolution

The production of viable and fertile hybrids may lead to at least three possible evolutionary scenarios: 1) if hybridization occurs repeatedly, the extensive gene flow may lead to the extinction of one of the hybridizing species through genetic assimilation (*e.g.* Rosenfield *et al.* 2004) or even to the merging of the hybridizing species (*e.g.* Taylor *et al.* 2006) – see section 1.2.4; 2) if hybrids show reduced fitness a hybrid zone may be established, where gene exchange may occur but merging of the parental taxa is prevented (Barton & Hewitt 1985); and 3) if hybrids are at least partially reproductively isolated from the hybridizing species, the formation of a new, allopolyploid or homoploid, hybrid species may occur - reticulate evolution (*e.g.* Buerkle *et al.* 2000; Comai *et al.* 2003; Smith *et al.* 2003; Chapman & Burke 2007).

Amongst animals, there is a low number of well-documented cases of hybrid speciation (Coyne & Orr 2004) perhaps due to the difficulty to produce a hybrid lineage that overcomes the direct competition with its parents and that is reproductively isolated in a way that merging with the parental species is prevented (Buerkle *et al.* 2000; Rieseberg 2001; Chapman & Burke 2007).

The reproductive isolation from the parental species may be achieved immediately if the hybrid offspring exhibit a duplicated chromosome set compared to the parental species (allopolyploids – Chapman & Burke 2007). Another possible isolation mechanism involves chromosomal rearrangements that suppress recombination in backcrosses and thereby reduce gene flow with the parental species (like happens in homoploid hybrid species - Rieseberg 2001). Also, the colonization of a novel habitat by hybrids could, in theory, contribute to the reproductive isolation from the parental species, as already described for plants (Buerkle *et al.* 2000).

1.2.4 Introgression

Once formed, fertile F1 hybrids may backcross with one or both the parental species, allowing the occurrence of gene flow and eventually leading, respectively, to the incorporation of the genes of one species into the genome of the other species (introgressive hybridization) or to the complete merging of the previously isolated hybridizing species ("hybrid swarm") (Scribner *et al.* 2001).

If two hybridizing species are common in their habitats, even low rates of hybridization may have important evolutionary consequences. Indeed, the introgression of genes through hybridization enhances the genetic variability and thus may ultimately contribute to adaptation and diversification (Dowling & Secor 1997; Mallet 2005).

As already mentioned above for the rarity of hybridization in animals, evidence for the occurrence of introgression of genes has also increased in the last decades, as more molecular techniques became available to the researchers. These techniques inclusively detect unsuspected cases of introgression, namely when parental species and their hybrids are morphologically similar. Thus, it seems reasonable to admit that in a near future the currently held biological species concept might be reformulated.

1.3 Hybridization and allopolyploidy in vertebrates

At a macroevolutionary level, recent studies of eukaryotic genomes indicated that polyploidization seems to have played a crucial role in the evolution of vertebrates by creating new genes by duplicating old ones, as first suggested by S. Ohno in 1970 (for an historical review see Wolfe 2001). This author also hypothesized two rounds (2R) of entire genome duplications in ancient vertebrate history, one when the lineage of craniates split from cephalochordates and the other at the origin of the gnathostomes lineage. Several studies (*e.g.* Sidow 1996; Spring 1997; Gibson & Spring 2000; McLysaght *et al.* 2002; Dehal & Boore 2005) have been producing evidence in support of the 2R hypothesis (for an historical review see Furlong & Holland 2004).

The occurrence of a third whole-genome duplication (3R) in the ray-finned fish (actinopterygians) lineage was confirmed by Jaillon *et al.* (2004) after the analysis of the genome of the freshwater fish *Tetraodon nigroviridis* and supported by other studies (*e.g.* Meyer & Schartl 1999; Christofells *et al.* 2004; Meyer & Van De Peer 2005; Woods *et al.* 2005; Crow *et al.* 2006).

However, and in contrast to the generally accepted whole-genome duplication in the actinopterygians lineage, there is still some scepticism towards the 2R hypothesis since it is hard to test such antique events and the detection of their traces may be difficult due to rediploidizations and degeneration of duplicated genes (*e.g.* Skrabanek & Wolfe 1998; Friedman & Hughes 2001a, 2001b; Martin 2001; Panopoulou & Poustka 2005; Van de Peer & Meyer 2005). Thus, some authors suggest that instead of whole-genome duplications, multiple independent gene duplications would have occurred in the evolution of vertebrates (Hughes *et al.* 2001).

Notwithstanding, and regardless of the outcome of this debate, the important role of gene duplication in the evolution of new functions, promoting diversification and evolutionary success, seems to be unquestionable (*e.g.* Hurles 2004; Comai 2005). According to the “duplication-degeneration-complementation model” (DDC) proposed by Force *et al.* (1999), gene duplications have three functional consequences: 1) one copy is lost due to the occurrence of a null mutation in the coding region that drifts to fixation (nonfunctionalization); 2) one copy acquires a mutation conferring a new function while the other retains the ancestral function (neofunctionalization); and 3) each copy loses part of the ancestral function by degenerative mutations and both copies are required for the full function (subfunctionalization).

The neofunctionalization of entire pathways rather than of individual gene functions would have been crucial for eukaryote genome evolution (Jaillon *et al.* 2004). However, the type of duplications involved (either by auto- or allopolyploidy) is another unclear issue (Panopoulou & Poustka 2005).

More recent polyploidization events have been reported for several vertebrates (many of which exhibit asexual modes of reproduction), belonging to 13 fish families, 10 families of amphibians, seven reptile families and two bird families (Otto & Whitton 2000).

In fishes, the predominant mode of polyploidization seems to be allopolyploidization, as suggested by the congruence between the orders including

polyploids and the list of 56 families for which hybrid fish are known (Collares-Pereira 1987/8; Le Comber & Smith 2004). Remarkable examples among fish are the entirely tetraploid Salmonidae and Catostomidae families (that might have resulted from an ancient hybridization event although it was not yet demonstrated) and the existence of individuals with distinct ploidy levels in other families like Acipenseridae, Atherinidae, Balitoridae, Percidae, Claridae, Heteropneustidae, Cobitidae and Cyprinidae (Otto & Whitton 2000; Leggatt & Iwama 2003; Le Comber & Smith 2004). In the two later families polyploidization is in fact recurrent (Gromicho & Collares-Pereira 2007).

The high incidence of hybridization in fish taxa seems to be a result of several contributing factors: external fertilization, weak behavioural isolating mechanisms, unequal abundance of the two parental species, competition for limited spawning habitat, decreasing habitat complexity, and susceptibility to secondary contact between recently evolved forms (reviewed by Scribner *et al.* 2001 and by Chávez & Turgeon 2007). Additionally, hybrids between distantly related fish species are frequently viable suggesting that fish appear to be less susceptible to the severe developmental incompatibilities that affect interspecific hybrids in other vertebrates (Scribner *et al.* 2001).

1.4 Reproduction in hybrid fish complexes

Asexual vertebrates show reproductive modes that differ drastically from typical sexual reproduction. While reptiles such as asexual lizards (*e.g.* *Cnemidophorus* sp., *Gymnophthalmus underwoodi*, *Lacerta* sp., *Menetia greyii* and Komodo dragons) reproduce sperm independently by strict or facultative parthenogenesis, asexual fish and amphibians (*e.g.* *Ambystoma* complex and *Rana esculenta*) usually reproduce by gynogenesis and hybridogenesis, involving the need to parasitize sexual species as sperm donors (Vrijenhoek *et al.* 1989; Adams *et al.* 2003; Watts *et al.* 2006). Table 1 summarizes all the reproductive modes described for asexual vertebrates and presents examples of asexual fish complexes for which some of the reproductive modes were described.

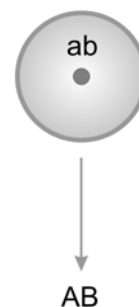
Table 1 — Sperm independent and sperm dependent asexual reproductive modes known for asexual vertebrates. Some examples of fish complexes exhibiting each reproductive mode are presented (adapted from Dawley 1989; Vrijenhoek *et al.* 1989; Avise *et al.* 1992; Dowling & Secor 1997; Alves *et al.* 2001; Saitoh *et al.* 2004; Gregory & Mable 2005; Schlupp 2005; Chávez & Turgeon 2007). In some cases, as happens in the *S. alburnoides* complex, the reproductive mode depends on the ploidy form considered.

Sperm independent

Parthenogenesis

Eggs are produced without syngamy, genetic recombination or reduction in ploidy and develop into clonal offspring.

(undescribed in fishes but reported in several reptiles).

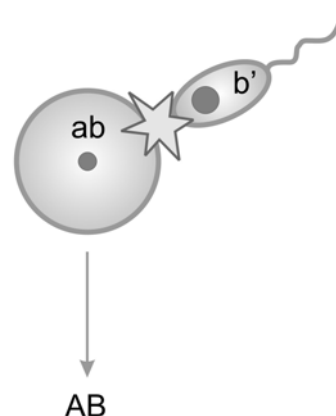


Sperm dependent

Gynogenesis

Similar to parthenogenesis but sperm is necessary to trigger embryogenesis (without karyogamy). The paternal genome is eliminated or degenerates and, thus, clonal offspring is produced.

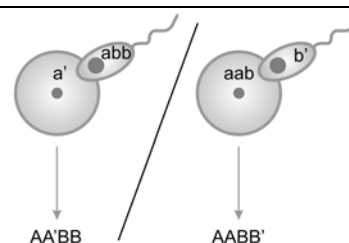
- *Misgurnus anguillicaudatus* (Cobitidae)
- *Cobitis* complexes (Cobitidae)
- *Carassius auratus* spp. (Cyprinidae)
- *Phoxinus eos-neogaeus* (Cyprinidae)
- *Poeciliopsis* complex (Poeciliidae)
- *Poecilia* complex (Poeciliidae)
- *Menidia clarkhubbsi* complex (Atherinidae)
- *Fundulus diaphanous* x *F. heteroclitus* hybrids (Fundulidae)
- *S. alburnoides* complex (Cyprinidae)¹



Production of unreduced clonal eggs

Females and males produce unreduced clonal gametes which, after fertilization, contribute to the elevation of the ploidy level of the offspring. The non-reduction of gametes may be achieved by apomixes (by mitosis) or by automixis (via premeiotic endoduplication).

- *S. alburnoides* complex (Cyprinidae)
-

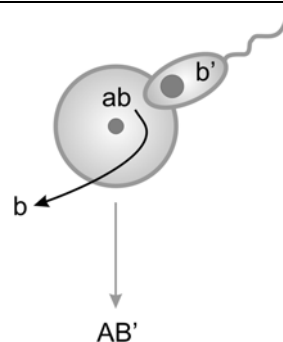


(Table 1 - cont.)

Hybridogenesis

During gametogenesis, the paternal genome is excluded and the maternal genome is clonally transmitted to the haploid egg. Eggs are then fertilized by sperm of the paternal species, restoring the hybrid condition. Thus, the paternal genome is replaced in each generation (hemiclonal inheritance). Requires syngamy and karyogamy of egg and sperm.

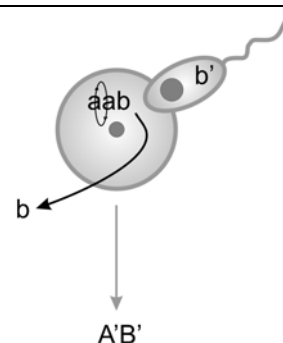
- *Poeciliopsis* complex (Poeciliidae)
- *Cobitis hankugensis-longicorpus* (Cobitidae)



Meiotic hybridogenesis

Similar to hybridogenesis but after the exclusion of the heterospecific genome the remaining homospecific genomes undergo meiosis with random segregation and recombination.

- *S. alburnoides* complex (Cyprinidae)



¹ confirmed only in laboratorial experiments

Concerning the reproduction of asexual fish, several studies have been focusing oogenesis (*e.g.* Zhang *et al.* 1998; Zhang & Arai 1999; Alves *et al.* 2004; Oshima *et al.* 2005; Itono *et al.* 2006), spermatogenesis (*e.g.* Monaco *et al.* 1981; Fan & Liu 1990; Gui *et al.* 1992; Alves *et al.* 1999; Lamatsch *et al.* 2000; Vasil'ev *et al.* 2003; Morishima *et al.* 2004; Nam & Kim 2004; Oshima *et al.* 2005), and inheritance of the parental genomes (*e.g.* Goddard & Schultz 1993; Alves *et al.* 1996, 1998; Carmona *et al.* 1997; Arai & Mukaino 1998; Arai & Inamori 1999; Momotani *et al.* 2002; Morishima *et al.* 2002; Zhang *et al.* 2002; Nam & Kim 2004; Pala & Coelho 2005; Crespo-López *et al.* 2006; Janko *et al.* 2007).

Contrastingly, investigations on the ethological aspects of the reproduction of these complexes are scarce albeit the recognized importance of mate choice in preventing hybridization between sympatric species (Mallet 2005). Studies on mate choice and mating success have been almost exclusively restricted to the *Poecilia* complex (*e.g.* Balsano *et al.* 1981, 1985; Farr *et al.* 1986; Schlupp *et al.* 1992, 1994, 1999, 2001; Schlupp & Ryan 1996, 1997; Körner *et al.* 1999; Landmann *et al.* 1999; Brooks & Endler 2001; Gabor &

Ryan 2001; Ptacek 2002), with few exceptions concerning the asexual fish complexes of *Carassius auratus* (Hakoyama & Iguchi 2001) and *Poeciliopsis monacha-lucida* (Lima & Bizerril 2002).

1.5 The case-study of *S. alburnoides*

1.5.1 General description

The endemic Iberian cyprinid *Squalius alburnoides* (initially described in the genus *Leuciscus* by Steindachner 1866, later transferred to *Rutilus* and to *Tropidophoxinellus* genera and finally returned to the original classification – Collares-Pereira *et al.* 1999) has a wide distribution range, being sympatric with at least three other *Squalius* species – *S. pyrenaicus*, *S. aradensis* and *S. carolitertii*. All the Iberian endemic *Squalius* (in a total of at least seven species) form a strongly supported monophyletic group (Zardoya & Doadrio 1999; Sanjur *et al.* 2003).

In the first detailed studies on *S. alburnoides* (Collares-Pereira 1983, 1984), the analysis of the biometric variability of Portuguese populations revealed the existence of two morphological forms, a discovery that led to the postulation of an “alburnoides complex”. The morphological and cytogenetic analyses (Collares-Pereira 1984, 1985) allowed the definition of a more common, widely distributed and almost all-female triploid form with some rare males and a less common diploid form that included both sexes. As the sex-ratio of the triploid form was extremely female biased, an asexual mode of reproduction was postulated for these females, instead of the bisexual reproduction that was expected to occur in the diploid form (Collares-Pereira 1985, 1989). Subsequent studies, using allozyme markers (Alves *et al.* 1997a; Carmona *et al.* 1997), highlighted the hybrid origin of the complex based on the detection of fixed heterozygosity at numerous allozyme loci in most specimens, with one set of alleles being identical to that of *Squalius pyrenaicus* (the maternal ancestor of the complex – P genome) whereas the other (paternal) set (coded as A-genome by Alves *et al.* 1998) could not be attributed to any

known living species. However, recent studies using 1) sequences of introns from aldolase-B and triosephosphate isomerase-B nuclear genes (Gilles *et al.* unpublished in Alves *et al.* 2001); 2) sequences of the beta-actin gene (Robalo *et al.* 2006 – **Chapter 7**); 3) microsatellites (Crespo-López *et al.* 2006); and 4) chromosome markers (Gromicho *et al.* 2006) revealed that the missing ancestor belonged to an extinct *Anaecypris*-like species (A genome), not identical to the extant *Anaecypris hispanica*. As the original mitochondrial DNA (mtDNA) of the *Anaecypris*-like ancestor was never detected, the *S. alburnoides* complex must have had its origin on an interspecific hybridization process that was, if not completely, at least predominantly, unidirectional (involving *S. pyrenaicus* females and *Anaecypris*-like males).

Aside from the diploid ($2n=50$), triploid ($3n=75$) and tetraploid ($4n=100$) hybrid forms, the complex also comprises nonhybrid individuals presenting the nuclear genome of the *Anaecypris*-like ancestor in homozygosity (AA genome). The presence of a *S. pyrenaicus*-like mtDNA indicates that these nonhybrids were reconstituted from the hybrids, discarding the hypothesis of being true descendants of the putative paternal lineage (Alves *et al.* 2002). All the specimens belonging to this nonhybrid form so far analyzed using morphological, genetic and cytogenetic markers (Collares-Pereira 1984; Alves *et al.* 1998, 2001, 2002; Gromicho & Collares-Pereira 2004, 2007; Gromicho *et al.* 2006) were males except one female reported by Carmona (1997), in the river Estena (Guadiana River Basin), and another female reported by Sousa-Santos *et al.* (2006a) -

Chapter 4.

Although firstly considered to be an asexual complex due to its reproductive modes, female preponderance, hybrid origin and various ploidy forms, *S. alburnoides* exhibits some characteristics that are uncommon among asexual vertebrates - the presence of around 20% of fertile diploid males, the incorporation of sperm into the majority of the eggs and the presence of recombination in fish of both sexes (Alves *et al.* 2001) – the implications of these exceptions will be further discussed in section 10.4.

This complex has also the peculiarity of having a distinct constitution depending on the river basin considered: some ploidies and genomic constitutions are lacking or are present in considerably distinct relative frequencies in some populations (Alves *et al.* 2001; Pala & Coelho 2005).

Additionally, as *S. alburnoides* is sympatric with other *Squalius*, interspecific crosses resulted in the introgression of non-*S. pyrenaicus* genes in the complex (Alves *et al.* 1997b), elevating the number of known genomic constitutions. Indeed, besides PA, PAA, PPA and PPAA genomes in the southern basins where *S. pyrenaicus* occur, one can also find CA, CAA, CCA and CCAA genomes in the northern drainages of Mondego and Douro (the letter “C” stands for the sympatric *S. carolitertii*), and QA, QAA, QQA and QQAA genomes in the southern Quarteira drainage (the letter “Q” stands for the sympatric *S. aradensis*). Theoretically, asymmetrical tetraploids may also occur but only one case was reported so far: a PAAA tetraploid from the Tagus drainage (Sousa-Santos *et al.* in press – **Chapter 8**).

These and other features that emerged from the research conducted on this complex during the last two decades will be detailed below.

1.5.2 Origin and dispersal

Although the hybrid and unidirectional origin seems to be consensual, the number of original hybridization events has been a matter of controversy. Indeed, the analysis of the mtDNA diversity of *S. alburnoides* populations led some authors to postulate multiple independent origins for the complex. Alves *et al.* (1997b), suggested that the complex had at least two independent origins (in the Sado drainage and in the Tagus-Guadiana region), and later Alves *et al.* (2002) postulated three original hybridization episodes, one in each of the southern drainages of Tagus, Sado and Guadiana. Recently, Cunha *et al.* (2004) extended this “multiple origins” hypothesis by suggesting five independent hybridizations episodes: in the Sado, in the Alagon-Douro region, in the Tagus-Mondego, in the Guadiana-Guadalquivir, and in the southwestern drainage of Quarteira. The formulation of this hypothesis was based on the fact that *S. alburnoides* showed a stronger affinity with *S. pyrenaicus* from the same region than with conspecifics from other regions. Independently of the number of hybridization episodes considered, all the hypotheses suggesting a polyphyletic origin for the complex presume the co-existence of both ancestors in all drainages and the independent extinction of the paternal ancestor in each of those drainages, after the formation of the complex.

The foundation of *de novo* hybrid lineages is virtually precluded by the present-day virtual absence of the paternal ancestor species. However, nonhybrid *S. alburnoides* males may be considered to be representatives of the *Anaecypris*-like ancestor since they preserve its nuclear genome (Alves *et al.* 2002; Gromicho & Collares-Pereira 2004, 2007).

1.5.3 Reproductive dynamics

The study of the peculiar reproductive modes of the *S. alburnoides* complex and the description of the pathways leading to the formation of its various genomic forms have received much attention from the researchers (Alves *et al.* 1996, 1997a,b, 1998, 1999, 2001, 2002, 2004; Carmona *et al.* 1997; Gromicho & Collares-Pereira 2004; Pala & Coelho 2005; Crespo-López *et al.* 2006). The determination of the ploidy level has been achieved by karyological analysis (reviewed in Gromicho & Collares-Pereira 2007) and, by an easier and faster way, using flow cytometry measurements of erythrocyte DNA content (Próspero & Collares-Pereira 2000). The subsequent identification of genomic constitutions has been accomplished with the controversial use of allozymes, through the detection of dosage-related staining intensities of electromorphs (Carmona *et al.* 1997); more effectively, with the analysis of some variable microsatellite loci with diagnostic alleles for both parental genomes (Pala & Coelho 2005; Crespo-Lopez *et al.* 2006); and with a new method based on beta-actin nuclear gene sequences (Sousa-Santos *et al.* 2005 – **Chapter 2**).

Summarizing the available information it can be stated that the reproductive modes exhibited by females include the production of unreduced eggs by diploid and triploid females (with subsequent incorporation of sperm); gynogenetic development of the eggs in a very low proportion (3%) of the diploid females (only observed in experimental crosses); “meiotic hybridogenesis”, in which triploid PAA females exclude the heterospecific genome after which they undergo normal or unreduced meiosis leading to A or AA gametes, respectively; and probably normal meiosis in balanced tetraploids (reviewed in Alves *et al.* 2001). It was also demonstrated that the same triploid PAA female is able to produce simultaneously haploid and triploid eggs (Alves *et al.* 2004).

Concerning the males, diploid hybrids produce fertile unreduced sperm, while nonhybrids and balanced tetraploids perform normal meiosis (reviewed in Alves *et al.*

2001). Although few individuals had been analyzed and despite their genomic constitutions were unknown, triploid males also appear to produce unreduced sperm (Alves *et al.* 1999; Sousa-Santos *et al.* 2007 – **Chapter 5**).

The reproductive modes and gametogenic processes known for females and males are determinant not only to the ploidy level of the progenies but also to the proportion of the P and A genomes found in the hybrids, contributing to the diverse array of *S. alburnoides* forms (summarized in Table 2). However, as mentioned above, the diversity and abundance of forms depends on the river basin considered, a situation that is still lacking a formal explanation.

Table 2 – Expected ploidies and genomic constitutions of the offspring resultant from crosses between *S. alburnoides* forms in southern populations (gametogenic processes described in Alves *et al.* 1999, 2001, 2004; Crespo-López *et al.* 2006; Sousa-Santos *et al.* 2007 – **Chapter 5**). The spermatogenesis of the PPA males is unknown but it was assumed that these males, like the PAA ones, produce unreduced sperm. As regards the northern populations, the gametogenic processes are likely identical, with the exception of the CA females, which seem to produce C-eggs (Carmona *et al.* 1997). *virtually unviable progeny (the maximum number of chromosome sets known for the complex is exceeded); **form never observed in nature but obtained in experimental crosses (Alves *et al.* 2001). Gynogenesis of the diploid females was not depicted since a very small fraction of descendants are produced by this reproductive mode (3% - Alves *et al.* 2001). Legend of the reproductive modes: H – Hybridogenesis; U – Production of unreduced gametes; M – normal meiosis; MH – meiotic hybridogenesis.

			MALES	AA	PA	PPAA	PAA	PPA
			Reproductive modes	M	U	M	U	U ?
FEMALES	Reproductive modes	gametes						?
PA	U		PAA	PPAA	PPAA	PPAAA*	PPPPAA*	
PAA	MH		AA	PAA	PAA	PAAA	PPAA	
	H		AAA**	PAAA	PAAA	PAAAA*	PPAAA*	
	U		PAAA	PPAAA*	PPAAA*	PPAAAA*	PPPPAA*	
PPA	MH		PA	PPA	PPA	PPAA	PPPA	
PPAA	M		PAA	PPAA	PPAA	PPAAA*	PPPPAA*	

Regarding interspecific crosses, there is evidence of introgressions of mtDNA of other species in the complex: specimens with *S. carolitertii*-like mtDNA were detected by Alves *et al.* (1997b) and Sousa-Santos *et al.* (in press – **Chapter 8**) in Douro and Mondego populations, and others with *S. aradensis*-like mtDNA were detected in the Quarteira River (Sousa-Santos *et al.* 2006b – **Chapter 9**). Despite these introgressions, the *S. pyrenaicus*-like mtDNA is prevalent in the complex (Cunha *et al.* 2004; Pala & Coelho 2005; Sousa-Santos *et al.* in press – **Chapter 8**), even outside the distribution range of *S. pyrenaicus*, which has been interpreted as indicating that crosses between *S. alburnoides* males and females of other species are uncommon, except at the origin of the complex (Alves *et al.* 2001).

Finally, the mechanism of sex determination of *S. alburnoides* is still an unclear issue. Indeed, although the eventual ZW/ZZ sex-determining system proposed for *S. pyrenaicus* (Collares-Pereira *et al.* 1998) would explain the data obtained in some crossing experiments, for other experimentally obtained results the putative involvement of non-W-linked genes was postulated (Alves *et al.* 1998). More recently, Gromicho & Collares-Pereira (2004, 2007) found no evidence of clearly heteromorphic sex chromosomes either in *S. pyrenaicus* or in the AA lineage.

1.5.4 Biology and ecology

Studies focusing aspects of the biology of the *S. alburnoides* complex (Fernández-Delgado & Herrera 1994; Peris *et al.* 1994; Ribeiro *et al.* 2003) revealed that females are multiple spawners; all fish attain sexual maturity early, until their second year of life; the onset of maturation is synchronous for diploid and triploid females and appears to be associated with the increase of water temperature in the spring; females show higher longevity than males and triploid females apparently live longer than the diploid ones.

Concerning the ecology of the complex, Martins *et al.* (1998) demonstrated the existence, at least in the Guadiana basin, of spatial segregation between different *S. alburnoides* forms: nonhybrids tend to prefer habitats characterized by shallow waters, higher temperatures and silt/sandy substrata, while diploid females show a preference for deeper waters, steeper gradients and coarse substrata and triploid females are mostly

found in areas with higher current velocities and instream cover. A spatial segregation of feeding niches between the three forms, especially during dry periods with lower prey availability, was also documented by Gomes-Ferreira *et al.* (2005): nonhybrid males exhibited a higher specialization for food, feeding mostly near the surface; diploid females preferred feeding grounds near the bottom and submerged vegetation; and triploid females, although showing an intermediate foraging behaviour, tended to have more affinities with the feeding habits of nonhybrid males.

Additionally to these differential ecological requirements, the *S. alburnoides* diploid and triploid forms also differ morphologically, namely in body size, position of the mouth, body profiles, pharyngeal tooth formula, number of scales in the lateral line and of gillrakers (Collares-Pereira 1984, 1985). However, at that time, the distinction only involved the triploid and diploid forms (named as A and B forms respectively) and, thus, the existence of morphological differences between all types of hybrids and nonhybrids could not be tested (this thematic was approached in the present thesis - Sousa-Santos *et al.* submitted - **Chapter 3**).

1.6 Aims and structure of the thesis

The main goal of this thesis was to investigate the reproductive behaviour of the *S. alburnoides* complex, a crucial issue to understand not only the present day success of the complex but also its remarkable evolutionary history. To do this, the research followed distinct but complementary domains, and the following specific objectives were formulated:

- 1) to search for nuclear markers suitable for a clear assessment of the genomic constitutions of the *S. alburnoides* individuals;
- 2) to investigate the relationships between the genomic constitution and the phenotype of the distinct forms of the complex;

- 3) to describe the spawning behaviour of the *S. alburnoides* complex and to search for the existence of mate preferences;
- 4) to contribute to a better understanding of the known genomic constitutions of the complex and of its missing paternal ancestor;
- 5) to confirm and deepen the previous knowledge on the reproductive modes of the *S. alburnoides* males;
- 6) to evaluate the natural occurrence of hybridization between *S. alburnoides* males and females of other *Squalius* species and both their relationships and implications to the genetic dynamics of the complex;
- 7) to discuss hypothetical consequences of the reciprocal introgression between *S. alburnoides* and other *Squalius* species;
- 8) to discuss possible locations for the origin of the complex and to draw plausible dispersal routes to explain the present day distribution range of the complex.

To achieve these goals, extensive fish sampling was made almost throughout the entire Portuguese distribution range of the complex (Douro, Mondego, Tagus, Sado, Guadiana and Quarteira drainages). The parental sequences involved in the formation of hybrid genomes and the individual genomic constitutions and ploidies (data that were essential to subsequent morphological, ethological and phylogeographical studies) were assessed with a new method based on the analysis of sequences of the beta-actin nuclear gene, validated by flow cytometry measurements of erythrocyte DNA content and by the analysis of artificial hybrids made with known volumes of DNA extracted from distinct species. In addition, other *Squalius* species were collected in several Portuguese rivers (including small basins where *S. alburnoides* does not occur) to allow comparisons in the morphological and phylogeographical studies. Additional cytochrome *b* gene sequences of specimens belonging to Portuguese and Spanish *Squalius* populations already published and deposited in GenBank were also used to get a better picture of the diversity of Iberian populations.

The study of the ethological aspects of the reproduction of the complex was conducted with specimens from the Tagus basin using a semi-natural approach, but

controlled experiments using fish from other basins (*S. alburnoides* from Guadiana and Douro and *S. torgalensis* from Mira) were also conducted.

The research performed to address the above mentioned specific objectives was already published (with the exception of one recently submitted manuscript) in indexed international journals. The resultant papers constitute the bulk of this thesis (Chapters 2 to 9) and are presented in their published version (or in the version submitted for publication in one case) but, for uniformity reasons, the formatting styles of the distinct journals were not maintained.

The thesis is organized in six Parts and eleven Chapters and obeys to the following structure:

Part I corresponds to the General Introduction of the dissertation and includes **Chapter 1**, focusing aspects related to hybridization: polyploidy, speciation, asexuality, introgression and their occurrence in fish. Descriptions of the asexual modes of reproduction known in hybrid fish complexes and of the case-study *S. alburnoides* complex are also addressed in this Chapter.

In **Part II**, including Chapters 2 and 3, methodological tools developed for the assessment of the genomic constitution of the *S. alburnoides* forms are presented. **Chapter 2** describes a new method based on the pattern of double peaks (generated due to the presence of heterozygous indels) that appear in the chromatograms after beta-actin nuclear gene sequencing. This method also proved to be useful in the determination of ploidies and in the reconstitution of the genetic complements involved in hybrids. In **Chapter 3** several morphological variables that can be used to differentiate the hybrid forms of the complex amongst themselves and from the representatives of both parental species (*S. alburnoides* nonhybrids and *S. pyrenaicus*) are presented. The effects of the parental genomes in the morphology of the hybrids and the possible influence of the morphology to the mating patterns exhibited by the distinct forms are discussed.

Part III includes studies related to the reproduction of the *S. alburnoides* complex: **Chapter 4** describes the very rare occurrence of a nonhybrid female (for which spawning behaviour and egg laying was observed) and discuss the implications of the finding to the evolutionary dynamics of the complex; **Chapter 5** describes the production of clonal

sperm by triploid males, giving rise to tetraploid offspring upon fertilization of *S. torgalensis* eggs, and confirms the already described types of gametes produced by nonhybrid and diploid hybrid males; and in **Chapter 6** the reproductive behaviour of *S. alburnoides* maintained in semi-natural conditions is described, with emphasis on the female mate preferences.

Chapters 7, 8 and 9 were grouped in **Part IV**, aiming to discuss the role of reproduction in the origin and dispersal (coupled with the subsequent introgression of genes) of *S. alburnoides*. In **Chapter 7**, the taxonomic identity of the paternal ancestor of the complex is highlighted using the nuclear beta-actin gene as molecular marker. **Chapter 8** presents a wide phylogeographical study of the *S. alburnoides* complex, framed by the reconstruction of the phylogeny and phylogeography of other sympatric *Squalius* species, and suggests a hypothetical location for the original hybridization event and the probable dispersal routes followed to attain the present day distribution area of the complex. **Chapter 9** describes an example on how the dispersal process allowed the contact between *S. alburnoides* and other *Squalius* species, whose genes were introgressed in the complex almost completely replacing the typical *S. pyrenaicus* mitochondrial DNA.

Finally, **Part V** comprises **Chapters 10 and 11** in which are presented, respectively, a general discussion aiming to integrate the results obtained to address the delineated specific objectives and the concluding remarks of this research with some proposals for future research.

1.7 References

- Adams M, Foster R, Hutchinson MN & Donnellan SC. 2003. The Australian scincid lizard *Menetia greyii*: a new instance of widespread vertebrate parthenogenesis. *Evolution* 57: 2619-2627.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 1996. Evidence for nonclonal reproduction in triploid *Rutilus alburnoides*. *Isozyme Bulletin* 29: 23.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 1997a. The *Rutilus alburnoides* complex (Cyprinidae): evidence for hybrid origin. *Journal of Zoological Systematics and Evolutionary Research* 35: 1-10.

- Alves MJ, Coelho MM & Collares-Pereira MJ. 1998. Diversity in the reproductive modes of females of the *Rutilus alburnoides* complex (Teleostei, Cyprinidae): a way to avoid the genetic constraints of uniparentalism. *Molecular Biology and Evolution* 15: 1233-1242.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 2001. Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: a genetic review. *Genetica* 111: 375-385.
- Alves MJ, Coelho MM, Collares-Pereira MJ & Dowling TE. 1997b. Maternal ancestry of the *Rutilus alburnoides* complex (Teleostei, Cyprinidae) as determined by analysis of cytochrome *b* sequences. *Evolution* 51: 1584-1592.
- Alves MJ, Coelho MM, Próspero MI & Collares-Pereira MJ. 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 MJ, Collares-Pereira MJ, Dowling TE & Coelho MM. 2002. The genetics of maintenance of an all-male lineage in the *Squalius alburnoides* complex. *Journal of Fish Biology* 60: 649-662.
- Alves MJ, Gromicho M, Collares-Pereira MJ, Crespo-López E & Coelho MM. 2004. Simultaneous production of triploid and haploid eggs by triploid *Squalius alburnoides* (Teleostei: Cyprinidae). *Journal of Experimental Zoology* 301A: 552-558.
- Arai K & Inamori Y. 1999. Viable hyperdiploid progeny between diploid female and induced triploid male in the loach, *Misgurnus anguillicaudatus*. *Suisanzoshoku* 47: 489-495.
- Arai K & Mukaino M. 1998. Electrophoretic analysis of the diploid progenies from triploid × diploid crosses in the loach *Misgurnus anguillicaudatus* (Pisces: Cobitidae). *Journal of Experimental Zoology* 280: 368-374.
- Arnold ML. 1997. Natural hybridization and evolution. Oxford University Press, Oxford, 215pp.
- Avice JC, Quattro JM & Vrijenhoek RC. 1992. Molecular clones within organismal clones: mitochondrial DNA phylogenies and the evolutionary histories of unisexual vertebrates. *Evolutionary Biology* 26: 225-246.
- Balsano JS, Kucharski K, Randle EJ, Rasch EM & Monaco PJ. 1981. Reduction of competition between bisexual and unisexual females of *Poecilia* in northeastern Mexico. *Environmental Biology of Fishes* 6: 39-48.
- Balsano JS, Randle EJ, Rasch EM & Monaco PJ. 1985. Reproductive behavior and the maintenance of all-female *Poecilia*. *Environmental Biology of Fishes* 12: 251-263.
- Barton NH. 2001. The role of hybridization in evolution. *Molecular Ecology* 10: 551-568.
- Barton NH & Hewitt GM. 1985. Analysis of hybrid zones. *Annual Review of Ecology and Systematics* 16: 113-148.
- Beukeboom LW & Vrijenhoek RC. 1998. Evolutionary genetics and ecology of sperm-dependent parthenogenesis. *Journal of Evolutionary Biology* 11: 755-782.
- Bi K & Bogart JP. 2006. Identification of intergenomic recombinations in unisexual salamanders of the genus *Ambystoma* by genomic in situ hybridization (GISH). *Cytogenetic and Genome Research* 112: 307-312.
- Brooks R & Endler J. 2001. Female guppies agree to differ: phenotypic and genetic variation in mate-choice behavior and the consequences for sexual selection. *Evolution* 55: 1644-1655.

- Buerkle CA, Morris RJ, Asmussen MA & Rieseberg LH. 2000. The likelihood of homoploid hybrid speciation. *Heredity* 84: 441-451.
- Carmona JA. 1997. Evolución y modelos de reproducción en un organismo unisexual de la Península Ibérica: el complejo *Tropidophoxinellus alburnoides*. PhD Thesis. Universidad Complutense, Madrid, 325 p.
- Carmona JA, Sanjur OI, Doadrio I, Machordom A & Vrijenhoek VC. 1997. Hybridogenetic reproduction and maternal ancestry of polyploid Iberian fish: the *Tropidophoxinellus alburnoides* complex. *Genetics* 146: 983-993.
- Chapman MA & Burke JM. 2007. Genetic divergence and hybrid speciation. *Evolution*. doi: 10.1111/j.1558-5646.2007.00134.x
- Chávez CH & Turgeon J. 2007. Asexual and sexual hybrids between *Fundulus diaphanous* and *F. heteroclitus* in the Canadian Atlantic region. *Molecular Ecology* 16: 1467-1480.
- Christofells A, Koh EGL, Chia J, Brenner S, Aparicio S & Venkatesh B. 2004. Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. *Molecular Biology and Evolution* 21: 1146-1151.
- Collares-Pereira MJ. 1983. *Estudo sistemático e citogenético dos pequenos ciprinídeos ibéricos pertencentes aos géneros Chondrostoma Agassiz, 1835, Rutilus Rafinesque, 1820 e Anaecypris Collares-Pereira, 1983*. PhD Thesis, Universidade de Lisboa, Lisboa, 511 pp.
- Collares-Pereira MJ. 1984. The "*Rutilus alburnoides* (Steindachner, 1866) complex" (Pisces, Cyprinidae). I. Biometrical analysis of some portuguese populations. *Arquivos do Museu Bocage (Série A)* III(8): 111-143.
- Collares-Pereira MJ. 1985. The "*Rutilus alburnoides* (Steindachner, 1866) complex". II. First data on the karyology of a well-established diploid-triploid group. *Arquivos do Museu Bocage (Série A)* III(5): 69-90.
- Collares-Pereira MJ. 1987/1988. O papel da poliploidia e da hibridação na evolução dos vertebrados. *Boletim da Sociedade Portuguesa das Ciências Naturais* (2ª série) 24: 95-114.
- Collares-Pereira MJ. 1989. Hybridization in European Cyprinids: evolutionary potential of unisexual populations. In Dawley RM & Bogart JP (Eds.) *Evolution and Ecology of Unisexual Vertebrates*. New York State Museum Bulletin 466, New York, pp 281-288.
- Collares-Pereira MJ, Alves MJ & Coelho MM. 1999. Reassessment of the generic position of the Iberian cyprinid "*alburnoides* complex": its return to the genus *Leuciscus*. *Journal of Fish Biology* 54: 464-468.
- Collares-Pereira MJ, Próspero MI, Biléu RI & Rodrigues EM. 1998. *Leuciscus* (Pisces, Cyprinidae) karyotypes: transect of portuguese populations. *Genetics and Molecular Biology* 21: 63-69.
- Comai L. 2005. The advantages and disadvantages of being polyploid. *Nature* 6: 836-846.
- Comai L, Madlung A, Josefsson C & Tyagi A. 2003. Do the different parental "heteromes" cause genomic shock in the newly formed allopolyploids? *Proceedings of the Royal Society of London Series B* 358: 1149-1155.
- Coyne JA & Orr HA. 2004. *Speciation*. Sinauer, Sunderland, 545pp.

- Crespo-López ME, Duarte T, Dowling T & Coelho MM. 2006. Modes of reproduction of the hybridogenetic fish *Squalius alburnoides* in the Tejo and Guadiana rivers: an approach using microsatellites. *Zoology* 109: 277-286.
- Crow KD, Stadler PF, Lynch VJ, Amemiya C & Wagner GP. 2006. The "fish-specific" Hox cluster duplication is coincident with the origin of Teleosts. *Molecular Biology and Evolution* 23: 121-136.
- Cunha C, Coelho MM, Carmona JA & Doadrio I. 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Molecular Ecology* 13: 2807-2817.
- Darwin C. 1859. On the origin of species: a facsimile of the first edition. Harvard University Press, Cambridge, 513p.
- Dawley R. 1989. An introduction to unisexual vertebrates. In Dawley RM & Bogart JP (Eds.) *Evolution and Ecology of Unisexual Vertebrates*. New York State Museum Bulletin 466, New York, pp 1-18.
- Dawley RM & Bogart JP (Eds.) 1989. *Evolution and Ecology of Unisexual Vertebrates*. New York State Museum Bulletin 466, New York.
- Dehal P & Boore JL. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biology* 3(10): 1700-1708.
- Dowling TE & Secor CL. 1997. The role of hybridization and introgression in the diversification of animals. *Annual Review of Ecology and Systematics* 28: 593-619.
- Fan Z & Liu G. 1990. The ploidy and reproductive mechanism of crucian carp, *Carassius auratus gibelio*. *Journal of Fish Biology* 36: 415-419.
- Farr JA, Travis J & Trexler JC. 1986. Behavioural allometry and interdemec variation in sexual behaviour of the sailfin molly, *Poecilia latipinna* (Pisces: Poeciliidae). *Animal Behaviour* 34: 497-509.
- Fernández-Delgado C & Herrera M. 1994. Population structure, growth and reproduction of *Tropidophoxinellus alburnoides* (Stendachner, 1866) in an intermittent stream of the Guadalquivir river basin (southern Spain). *Archiv für Hydrobiologie* 130: 359-369.
- Force A, Lynch M, Pickett FB, Amores A, Yan Y & Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151: 1531-1545.
- Friedman R & Hughes AL. 2001a. Gene duplication and the structure of eukaryotic genomes. *Genome Research* 11: 373-381.
- Friedman R & Hughes AL. 2001b. Pattern and timing of gene duplication in animal genomes. *Genome Research* 11: 1842-1847.
- Furlong RF & Holland PWH. 2004. Polyploidy in vertebrate ancestry: Ohno and beyond. *Biological Journal of the Linnean Society* 82: 425-430.
- Gabor CR & Ryan MJ. 2001. Geographical variation in reproductive character displacement in mate choice by male sailfin mollies. *Proceedings of the Royal Society of London B* 268: 1063-1070.

- Gibson TJ & Spring J. 2000. Evolution of sequences, structures and genomes. *Biochemical Society Transactions* 28: 259-264.
- Goddard K & Schultz R. 1993. Aclonal reproduction by polyploid members of the clonal hybrid species *Phoxinus eos-neogaeus* (Cyprinidae). *Copeia* 1993(3): 650-660.
- Gomes-Ferreira A, Ribeiro F, Moreira da Costa L, Cowx IG & Collares-Pereira MJ. 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. *Journal of Fish Biology* 66: 454-467.
- Gregory TR & Mable BK. 2005. Polyploidy in animals. In Gregory TR (Ed.) The evolution of the genome. Elsevier Academic Press, Burlington, pp 428-517.
- Gromicho M & Collares-Pereira MJ. 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.
- Gromicho M & Collares-Pereira MJ. 2007. The evolutionary role of hybridization and polyploidy in an Iberian cyprinid fish – a cytogenetic review. In Pisano E, Ozouf-Costaz C, Foresti F & Kapoor BG (Eds.) Fish Cytogenetics. Science Publishers, Enfield, pp 41-67.
- Gromicho M, Coelho MM, Alves MJ & Collares-Pereira MJ. 2006. Cytogenetic analysis of *Anaecypris hispanica* and its relationship with the paternal ancestor of the diploid-polyploid *Squalius alburnoides* complex. *Genome* 49: 1621-1628.
- Gui JF, Jia J, Liang SC & Jiang YG. 1992. Meiotic chromosome behaviour in male triploid transparent coloured crucian carp, *Carassius auratus* L. *Journal of Fish Biology* 41: 317-326.
- Hakoyama H & Iguchi K. 2002. Male mate choice in the gynogenetic-sexual complex of crucian carp, *Carassius auratus*. *Acta Ethologica* 4: 85-90.
- Hughes AL, Silva J & Friedman R. 2001. Ancient genome duplications did not structure the Human Hox-bearing chromosomes. *Genome Research* 11: 771-780.
- Hurles M. 2004. Gene duplication: the genomic trade in spare parts. *PLoS Biology* 2(7): 900-904.
- Itono M, Morishima K, Fujimoto T, Bando E, Yamaha E & Arai K. 2006. Premeiotic endomitosis produces diploid eggs in the natural clone loach, *Misgurnus anguillicaudatus* (Teleostei: Cobitidae). *Journal of Experimental Zoology* 305A: 513-523.
- Jaillon O, Aury J-M, Brunet F, Petit J-L, Stange-Thomann N, Mauceli E, Bouneau L, Fischer C, Ozouf-Costaz C, Bernot A, Nicaud S, Jaffe D, Fisher S, Lutfalla G, Dossat C, Segurens B, Da Silva C, Salanoubat M, Levy M, Boudet N, Castellano S, Anthouard V, Jubin C, Castelli V, Katinka M, Vacherie B, Biémont C, Skalli Z, Cattolico L, Poulain J, Berardinis V, Cruaud C, Duprat S, Brottier P, Coutanceau J-P, Gouzy J, Parra G, Lardier G, Chapple C, McKernan KJ, McEwan P, Bosak S, Kellis M, Volff J-N, Guigó R, Zody MC, Mesirov J, Lindblad-Toh K, Birren B, Nusbaum C, Kahn D, Robinson-Rechavi M, Laudet V, Schachter V, Quétier F, Saurin W, Scarpelli C, Wincker P, Lander ES, Weissenbach J & Crollius HS. 2004. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431: 946-955.

- Janko K, Bohlen J, Lamatsch D, Flajshans M, Epplen JT, Ráb P, Kotlík P & Slechtová 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*. DOI 10.1007/s10709-006-9130-5.
- Kearney M. 2005. Hybridization, glaciation and geographical parthenogenesis. *Trends in Ecology and Evolution* 20: 495-502.
- Körner K, Lütjens O, Parzefall J & Schlupp I. 1999. The role of experience in mating preferences of the unisexual Amazon molly. *Behaviour* 136: 257-268.
- Lamatsch DK, Steinlein C, Schmid M & Scharl M. 2000. Noninvasive determination of genome size and ploidy level in fishes by flow cytometry: detection of triploid *Poecilia formosa*. *Cytometry* 39: 91-95.
- Landmann K, Parzefall J & Schlupp I. 1999. A sexual preference in the Amazon molly, *Poecilia formosa*. *Environmental Biology of Fishes* 56: 325-331.
- Le Comber SCL & Smith C. 2004. Polyploidy in fishes: patterns and processes. *Biological Journal of the Linnean Society* 82: 431-442.
- Leggatt RA & Iwama GK. 2003. Occurrence of polyploidy in the fishes. *Reviews in Fish Biology and Fisheries* 13: 237-246.
- Lima N & Bizerril C. 2002. Frequency-dependence of mating success in *Poeciliopsis monacha* (Pisces, Cyprinodontiformes) reproductive complex, Sonora, Mexico. *Brazilian Archives of Biology and Technology* 45(2): 161-169.
- Mallet J. 2005. Hybridization as an invasion of the genome. *Trends in Ecology and Evolution* 20: 229-237.
- Martin A. 2001. Is tetralogy true? Lack of support for the "one-to-four rule". *Molecular Biology and Evolution* 18: 89-93.
- Martins MJ, Collares-Pereira MJ, Cowx IG & Coelho MM. 1998. Diploids v. triploids of *Rutilus alburnoides*: spatial segregation and morphological differences. *Journal of Fish Biology* 52: 817-828.
- McLysaght A, Hokamp K & Wolfe KH. 2002. Extensive genomic duplication during early chordate evolution. *Nature Genetics* 31: 200-204.
- Meyer A & Scharl M. 1999. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Current Opinion in Cell Biology* 11: 699-704.
- Meyer A & Van de Peer Y. 2005. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *BioEssays* 27: 937-945.
- Momotani S, Morishima K, Zhang Q & Arai K. 2002. Genetic analyses of the progeny of triploid gynogens induced from unreduced eggs of triploid (diploid female x tetraploid male) loach. *Aquaculture* 204: 311-322.
- Monaco PJ, Rasch EM & Balsano JS. 1981. Sperm availability in naturally occurring bisexual-unisexual breeding complexes involving *Poecilia mexicana* and the gynogenetic teleost, *Poecilia formosa*. *Environmental Biology of Fishes* 6: 159-166.

- Morishima K, Horie S, Yamaha E & Arai K. 2002. A cryptic clonal line of the loach *Misgurnus anguillicaudatus* (Teleostei: Cobitidae) evidenced by induced gynogenesis, interspecific hybridization, microsatellite genotyping and multilocus DNA fingerprinting. *Zoological Science* 19: 565-575.
- Morishima K, Oshima K, Horie S, Fujimoto T, Yamaha E & Arai K. 2004. Clonal diploid sperm of the diploid-triploid mosaic loach, *Misgurnus anguillicaudatus* (Teleostei: Cobitidae). *Journal of Experimental Zoology* 301A: 502-511.
- Muller HJ. 1925. Why polyploidy is rarer in animals than in plants. *The American Naturalist* 59: 346-353.
- Nam YK & Kim DS. 2004. Ploidy status of progeny from the crosses between tetraploid males and diploid females in mud loach (*Misgurnus mizolepis*). *Aquaculture* 236: 575-582.
- Oshima K, Morishima K, Yamaha E & Arai K. 2005. Reproductive capacity of triploid loaches obtained from Hokkaido Island, Japan. *Ichthyological Research* 52: 1-8.
- Otto SP & Whitton J. 2000. Polyploid incidence and evolution. *Annual Review of Genetics* 34: 401-437.
- Pala I & Coelho MM. 2005. Contrasting views over a hybrid complex: between speciation and evolutionary "dead-end". *Gene* 347: 283-294.
- Panopoulou G & Poustka AJ. 2005. Timing and mechanism of ancient vertebrate genome duplications – the adventure of a hypothesis. *Trends in Genetics* 21(10): 559-567.
- Peris SJ, Marín A & Velasco JC. 1994. Age, growth and reproduction of the cyprinid *Tropidophoxinellus alburnoides* (Stephanidis, 1974) in River Agueda, West-Central Spain. *Archiv für Hydrobiologie* 131: 341-352.
- Próspero MI & Collares-Pereira MJ. 2000. Nuclear DNA content variation in the diploid-polyploid *Leuciscus alburnoides* complex (Teleostei, Cyprinidae) assessed by flow cytometry. *Folia Zoologica* 49: 53-58.
- Plácek MB. 2002. Patterns of inheritance of mating signals in interspecific hybrids between sailfin and shortfin mollies (Poeciliidae: *Poecilia Mollienesia*). *Genetica* 116: 329-342.
- Ribeiro F, Cowx IG, Tiago P, Filipe AF, Moreira da Costa L & Collares-Pereira MJ. 2003. Growth and reproductive traits of diploid and triploid forms of the *Squalius alburnoides* cyprinid complex in a tributary of the Guadiana River, Portugal. *Archiv für Hydrobiologie* 156: 471-484.
- Rieseberg LH. 2001. Chromosomal rearrangements and speciation. *Trends in Ecology and Evolution* 16: 351-358.
- Rieseberg LH, Archer MA & Wayne RK. 1999. Transgressive segregation, adaptation and speciation. *Heredity* 83: 363-372.
- Robalo JI, Sousa-Santos C, Levy A & Almada VC. 2006. Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. *Molecular Phylogenetics and Evolution* 39: 276-281.
- Rosenfield JA, Nolasco S, Lindauer S, Sandoval C & Kodric-Brown A. 2004. The role of hybrid vigor in the replacement of Pecos pupfish by its hybrids with sheepshead minnow. *Conservation Biology* 18: 1589-1598.

- Saitoh K, Kim I-S & Lee E-H. 2004. Mitochondrial gene introgression between spined loaches via hybridogenesis. *Zoological Science* 21: 795-798.
- Sanjurjo OI, Carmona JA & Doadrio I. 2003. Evolutionary and biogeographical patterns within Iberian populations of the genus *Squalius* inferred from molecular data. *Molecular Phylogenetics and Evolution* 29: 20-30.
- Scali V, Passamonti M, Marescalchi O & Mantovani B. 2003. Linkage between sexual and asexual lineages: genome evolution in *Bacillus* stick insects. *Biological Journal of the Linnean Society* 79: 137-150.
- Schlupp I. 2005. The evolutionary ecology of gynogenesis. *Annual Review of Ecology and Systematics* 36: 399-417.
- Schlupp I, Knab RM & Ryan MJ. 2001. Sexual harassment as a cost for molly females: bigger males cost less. *Behaviour* 138: 277-286.
- Schlupp I, Marler C & Ryan MJ. 1994. Benefit to male sailfin mollies of mating with heterospecific females. *Science New Series* 263(5145): 373-374.
- Schlupp I, Parzefall J, Epplen J, Nanda I, Schmid M & Scharf M. 1992. Pseudomale behaviour and spontaneous masculinization in the all-female teleost *Poecilia formosa* (Teleostei: Poeciliidae). *Behaviour* 122(1-2): 88-104.
- Schlupp I & Ryan MJ. 1996. Mixed-species shoals and the maintenance of a sexual-asexual mating system in mollies. *Animal Behaviour* 52: 885-890.
- Schlupp I & Ryan MJ. 1997. Male sailfin mollies (*Poecilia latipinna*) copy the mate choice of other males. *Behavioral Ecology* 8: 104-107.
- Schlupp I, Waschulewski M & Ryan MJ. 1999. Female preferences for naturally-occurring novel male traits. *Behaviour* 136: 519-527.
- Schmidt BR. 1996. Sexual and asexual reproduction in vertebrates. *Trends in Ecology and Evolution* 11: 253-254.
- Schultz RJ. 1967. Gynogenesis and triploidy in the viviparous fish *Poeciliopsis*. *Science* 157: 1564-1567.
- Schultz RJ. 1969. Hybridization, unisexuality, and polyploidy in the teleost *Poeciliopsis* (Poeciliidae) and other vertebrates. *The American Naturalist* 103: 605-619.
- Scribner KT, Page KS & Bartron ML. 2001. Hybridization in freshwater fishes: a review of case studies and cytonuclear methods of biological inference. *Reviews in Fish Biology and Fisheries* 10: 293-323.
- Seehausen O. 2004. Hybridization and adaptive radiation. *Trends in Ecology and Evolution* 19: 198-209.
- Sidow A (1996). Gen(om)e duplications in the evolution of early vertebrates. *Current Opinions in Genetics and Development* 6: 715-722.
- Skrabaneck L & Wolfe KH. 1998. Eukaryote genome duplication - where's the evidence? *Current Opinion in Genetics and Development* 8: 694-700.

- Smith PF, Konings A & Kornfield I. 2003. Hybrid origin of a cichlid population in Lake Malawi: implications for genetic variation and species diversity. *Molecular Ecology* 12: 2497-2504.
- Sousa-Santos C, Collares-Pereira MJ & Almada V. in press. Reading the history of a hybrid fish complex from its molecular record. *Molecular Phylogenetics and Evolution*
- Sousa-Santos C, Collares-Pereira MJ & Almada V. submitted. Phenotypic expression of different genome constitutions in a diploid-polyploid hybridogenetic fish complex. *Journal of Fish Biology*
- Sousa-Santos C, Collares-Pereira MJ & Almada V. 2006a. May a hybridogenetic complex regenerate the nuclear genome of both sexes of a missing ancestor? - First evidence on the occurrence of a nuclear nonhybrid *Squalius alburnoides* (Cyprinidae) female based on DNA sequencing. *Journal of Natural History* 40(23-24): 1443-1448.
- Sousa-Santos C, Collares-Pereira MJ & Almada V. 2006b. 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. *Journal of Fish Biology* 68 (Supplement B): 292-301.
- Sousa-Santos C, Collares-Pereira MJ & V Almada V. 2007. Fertile triploid males – an uncommon case among hybrid vertebrates. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* 307A(4): 220-225
- Sousa-Santos C, Robalo JI, Collares-Pereira MJ & Almada V. 2005. Heterozygous indels as useful tools in haplotype reconstruction of DNA nuclear genes of hybrid organisms – an example using cyprinid fishes. *DNA Sequence* 16(6): 462-467.
- Spring J. 1997. Vertebrate evolution by interspecific hybridisation – are we polyploid? *FEBS Letters* 400: 2-8.
- Stöck M, Lamatsch DK, Steinlein C, Epplen JT, Grosse W-R, Hock R, Klapperstück T, Lampert KP, Scheer U, Schmid M & Scharf M. 2002. A bisexually reproducing all-triploid vertebrate. *Nature Genetics* 30: 325-328.
- Tate JA, Soltis DE & Soltis PS. 2005. Polyploidy in plants. In Gregory TR (Ed.) The evolution of the genome. Elsevier Academic Press, Burlington, pp 372-426.
- Taylor EB, Boughman JW, Groenenboom M, Sniatynski M, Schluter D & Gow JL. 2006. Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology* 15:343-355.
- Van de Peer Y & Meyer A. 2005. Large-scale gene and ancient genome duplications. In Gregory TR (Ed.) The Evolution of the genome. Elsevier Academic Press, Burlington, pp 330-368.
- Vasil'ev VP, Akimova N, Emel'yanova NG, Pavlov DA & Vasil'eva ED. 2003. Reproductive capacities in the polyploid spined loaches from the unisexual-bisexual complex, occurred in the Moscow River. *Folia Zoologica* 51: 67-73.
- Vinogradov AE, Borkin LJ, Günther R & Rosanov JM. 1990. Genome elimination in diploid and triploid *Rana esculenta* males: cytological evidence from DNA flow cytometry. *Genome* 33, 619-627.
- Vrijenhoek R. 1989. Genetic and ecological constraints on the origins and establishment of unisexual vertebrates. In Dawley R & Bogart J (Eds.) Evolution and Ecology of Unisexual Vertebrates. New York State

Museum Bulletin 466, New York, pp 24-31.

Vrijenhoek RC. 1994. Unisexual fish: model systems for studying ecology and evolution. *Annual Review of Ecology and Systematics* 25: 71-96.

Vrijenhoek RC. 2006. Polyploid hybrids: multiple origins of a treefrog species. *Current Biology* 16: R245-R247.

Vrijenhoek RC, Dawley RM, Cole CJ & Bogart JP. 1989. A list of the known unisexual vertebrates. In Dawley R & Bogart J (Eds.) *Evolution and Ecology of Unisexual Vertebrates*. New York State Museum Bulletin 466, New York, pp 19-23.

Watts PC, Buley KR, Sanderson S, Boardman W, Ciofi C & Gibson R. 2006. Parthenogenesis in Komodo dragons. *Nature* 444: 1021-1022.

Wetherington JD, Kotora KE & Vrijenhoek RC. 1987. A test of the spontaneous heterosis hypothesis for unisexual vertebrates. *Evolution* 41: 721-731.

Wolfe KH. 2001. Yesterday's polyploids and the mystery of diploidization. *Nature Reviews Genetics* 2: 333-341.

Woods IG, Wilson C, Friedlander B, Chang P, Reyes DK, Nix R, Kelly PD, Chu F, Postlethwait JH & Talbot WS. 2005. The zebrafish gene map defines ancestral vertebrate chromosomes. *Genome Research* 15: 1307-1314.

Zardoya R & Doadrio I. 1999. Molecular evidence on the evolutionary and biogeographical patterns of European Cyprinids. *Journal of Molecular Evolution* 49: 227-237.

Zhang Q & Arai K. 1999. Distribution and reproductive capacity of natural triploid individuals and occurrence of unreduced eggs as a cause of polyploidization in the loach, *Misgurnus anguillicaudatus*. *Ichthyological Research* 46:153-161.

Zhang Q, Arai K & Yamashita M. 1998. Cytogenetic mechanisms for triploid and haploid egg formation in the triploid loach *Misgurnus anguillicaudatus*. *The Journal of Experimental Zoology* 281: 608-619.

Zhang Q, Hanada K & Arai K. 2002. Aberrant meioses in diploid and triploid progeny of gynogenetic diploids produced from eggs of natural tetraploid loach, *Misgurnus anguillicaudatus*. *Folia Zoologica* 51(2): 165-176.

PART II

METHODOLOGICAL TOOLS FOR GENOTYPE ASSESSMENT

Chapter 2

Heterozygous indels as useful tools in haplotype reconstruction of DNA nuclear genes of hybrid organisms – an example using cyprinid fishes

Sousa-Santos C, Robalo JI, Collares-Pereira MJ & Almada V (2005)

DNA Sequence 16(6): 462-467

Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genomic constitution of hybrid organisms

Sousa-Santos C, Robalo JI, Collares-Pereira MJ & Almada V

Abstract In this paper we describe a simple approach using double peaks in chromatograms generated as artefacts in the vicinity of heterozygous indels, to identify the specific sequences present in individual strands of a given DNA fragment. This method is useful to assign bases in individuals that are heterozygous at multiple sites. In addition, the relative sizes of the double peaks help to determine the ploidy level and the relative contribution of the parental genomes in hybrids. Our interpretation was confirmed with the analysis of artificial mixtures of DNA of two different species. Results were robust with varying PCR and sequencing conditions. The applicability of this method was demonstrated in hybrids of the *Squalius alburnoides* complex and in heterozygotes of *Chondrostoma oligolepis*. Far from being limited to these fish models and the gene where it was tested (*beta-actin*), this sequence reconstruction methodology is expected to have a broader application.

Keywords Double peaks, interspecific hybrids, haplotype reconstruction, beta-actin gene, cyprinid fish

Introduction

When an organism is heterozygous for several linked sites in a given DNA fragment that is being sequenced, one major problem is to assign the individual bases in each heterozygous position to each of the parental genomes. This problem is found both in intraspecific studies and especially in interspecific hybrids whose parental sequences often differ in many nucleotide positions.

The presence of a heterozygous indel in a fragment of nuclear DNA generates a disturbance in the sequencing process characterized by a succession of false double peaks. Bhangale *et al.* (2005) used the double peaks generated to identify indels in a set

of 330 human genes. Once identified, the length of the indel was inferred from the pattern of peaks by performing a pairwise alignment of bases corresponding to the two allelic sequences, obviating the need of having a previous knowledge of both of the sequences involved (Bhangale *et al.* 2005).

In this paper we explore further potentialities of this approach. First, we show with intra and interspecific data how the two parental haplotypes can be read from the chromatogram if we previously know its characteristic indels. In addition, we evidence how this method can help to determine the ploidy of each hybrid and to access the relative contributions of the parental genomes, constituting a useful alternative tool to quantitative PCR methods.

Methods

To test this method in interspecific hybrids we used individuals of the Iberian minnow *Squalius alburnoides* complex. Its hybrid origin resulted from interspecific crosses between *S. pyrenaicus* females (PP) and males from an unknown species (AA), generating $2n=50$, $3n=75$ and $4n=100$ hybrid forms (reviewed in Alves *et al.* 2001) and reconstituted diploid nonhybrids with the nuclear AA genome of the missing paternal ancestor (Alves *et al.* 2002; Robalo *et al.* in press ¹), which are morphologically distinct from the diploid hybrid form of the complex.

In order to be able to analyse the hybrid nuclear genomes, a total of 31 individuals of the parental species were analysed: 11 *S. pyrenaicus*, nine *S. carolitertii* and 11 diploid nonhybrid *S. alburnoides* (GenBank: AY943863 to AY943896). Samples of *S. carolitertii* were used since in the river basins where *S. pyrenaicus* is absent, although the mtDNA found in *S. alburnoides* fish is also *S. pyrenaicus*-like, the complex seems to be maintained by crosses with males of *S. carolitertii* (CC) and by diploid hybrid males (CA) (Cunha *et al.* 2004; Pala & Coelho 2005). Samples from 19 hybrids of *S. alburnoides* (eight diploids, 10 triploids and one tetraploid) were used. The ploidy of the hybrids was previously determined by flow cytometry using fresh fin clips, following an adaptation of the method proposed by Lamatsch *et al.* (2000) (Collares-Pereira *et al.* unpublished).

¹ Updated citation: Robalo *et al.* (2006) *Molecular Phylogenetics and Evolution* 39: 276-281.

To illustrate the applicability of the method in intraspecific studies we sampled 13 individuals of *Chondrostoma oligolepis*² (formerly known as *Chondrostoma macrolepidotum*) another Iberian minnow with 2n=50 (Collares-Pereira 1985).

Total genomic DNA was extracted from fin clips preserved in ethanol by an SDS/proteinase-k based protocol (adapted from Sambrook *et al.* 1989). A total of 927bp of the *beta-actin* gene was amplified using the primers For-5'-ATGGATGATGAAATTGCCGC-3' and Rev-5'-AGGATCTTCATGAGGTAGTC-3' (J. Robalo unpublished). The amplification process was conducted as follows: 35 cycles of [94°C(30sec.), 55°C(40sec.) and 72°C(1min30sec)]. Amplification and sequencing of DNA from six diploids and six triploids was repeated using different PCR conditions: 35 cycles of [94°C(30sec.), 42°C(40sec.), 72°C(1min30sec.)]. The amplified fragment is homologous to a region of the *beta-actin* gene of *Cyprinus carpio* (GenBank: M24113), between the positions 1622 and 2550, including introns B and C and three exons. Each sample was sequenced in both directions with the same primers used for PCR. Sequences were aligned with BioEdit® v.5.0.6.

Ploidy assessment methodology

To test the hypothesis that unbalanced proportions of parental genomes can be detected in nonquantitative PCR products, we produced six groups of artificial hybrids simulating hybrid forms of the *S. alburnoides* complex. Each group included PA, PAA and PPA forms made with mixtures of re-suspended DNA from the same genome donors (previously sequenced and all differentiated by specific point mutations): 9µl of each DNA suspension from the A- and P-genome donors to produce PA hybrids; 6µl of DNA suspension from the P-genome donor and 12µl from the A-genome donor for PAA hybrids; and the reversed quantities for PPA hybrids.

The contribution of each parental complement to the hybrid genome was quantified by the "measuring method": measuring both overlapping peaks in each position (using ImageTool 2.0 UTHSCSA® with a screen resolution of 1024x768 pixels) and calculating the ratio "height of peak from P/(height of peak from P + height of peak from A)" (P/P+A ratio). The provenience of the higher overlapping peak was also registered to calculate the percentage of P-peaks that were greater than A-peaks ("P-count") – "count method".

² *Achondrostoma oligolepis* according to the recent taxonomical revision of Robalo *et al.* (2007) *Mol. Phylogenet. Evol.* 42: 362-372.

Different DNA concentrations caused by variations in the extraction procedure could be responsible for excesses of one of the genomes. The value of the P/A ratio for each position measured in an artificial diploid was used as a "correction coefficient" for the peak heights measured in the chromatograms of the artificial triploids made with the same genome donors. These corrected P-peak values were then used to calculate the P/P+A ratio in each position. In natural hybrids the P/A correction was unnecessary.

Results

There was no clear distinction between *S. pyrenaicus* and *S. carolitertii* for the analysed gene segment (one to four mutations between pairs of haplotypes) thus they were here designated as "*S. pyrenaicus/S. carolitertii*" (PP genome). The chromatograms of the parental species showed single peaks, except for one to four single nucleotide polymorphisms for some fish.

All hybrids sequenced showed double peaks in segments that varied between 544 and 667bp (58.7% and 72.0% of the amplified fragment, respectively), involving the bases we expected to find if we overlapped the parental genomes. We assumed that analysing about a hundred double peaks should provide a sufficient number of distinct points to allow an adequate statistical analysis. Thus, a segment of 176bp was randomly selected in the double peaks region, containing 118 positions with overlapping peaks (the remaining were single peaks as a result of the addition of equal bases – see Figure 1).

Reconstruction of the parental sequences in a hybrid

In the presence of one heterozygous indel in a fragment of nuclear DNA, the sequencer starts to read two bases in the same position, a situation that generates a pattern of overlapping peaks in the chromatogram. In these regions of double peaks, the bases are out of phase as many positions as the number of bases of the indel (Figure 1a), a condition that will be maintained until a second indel of opposite direction counterbalances the first, which could be many dozens or hundreds of bases downstream from the initial indel. Thus, starting at the position where the first heterozygous indel occurs, it is possible to read the complements involved in the formation of a hybrid genome.

S. pyrenaicus/*S. carolitertii* and nuclear nonhybrid *S. alburnoides* differed by a total of seven indels. The reconstruction of the genomes of the hybrids was possible because each parental species had a characteristic number and location of indels (Figure 1b). It was also favoured by the existence of several point mutations characteristic of each parental species (in homozygous condition) that marked out highly conserved regions and made possible to ascribe unambiguously each peak in a double peaks region to the correct parental genome.

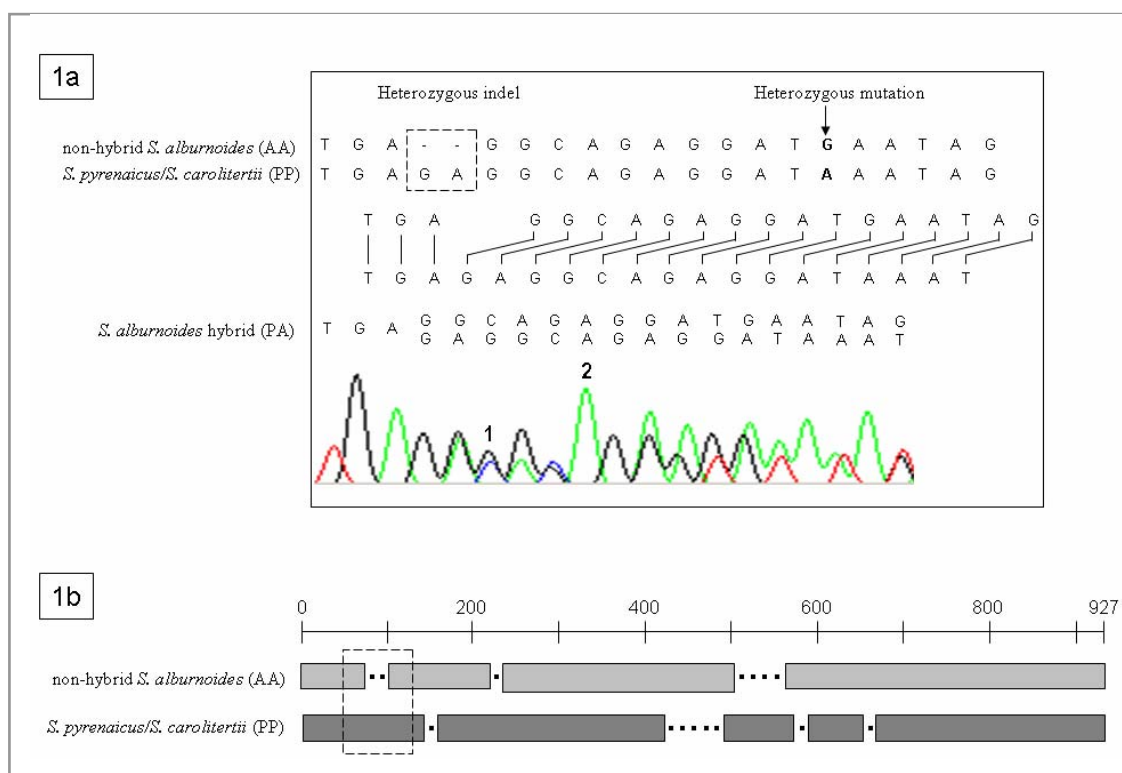


Figure 1 — 1a) Demonstration of the disturbance generated by the first heterozygous indel in the sequencing process of a *S. alburnoides* individual with P and A complements. When the bases in both alleles are the same the chromatogram shows a single higher peak reflecting the presence of double quantities of the same base, which means that the sequences can be read even in the homologous regions flanked by overlapping peaks. 1 — double peak; 2 — single peak resulting from the addition of equal bases. 1b) Schematic representation of the characteristic indels of the parental species in the analysed fragment of 927bp (deletions are represented by black dots). Indel size ranged from one to five bases (mean=1.88±1.68). The broken line box indicates the localization of the indel represented in Figure 1a.

Information on the ploidy and hybrid genome constitution

Since the height of a peak in a chromatogram reflects approximately the amount of a specific base in that position, one can expect that different genome constitutions exhibit chromatograms with different peak heights. Exception is made when there is a

suppression of signal at a given position because during the sequencing process the reading of a specific base may be affected by the constitution of the preceding one (for example, it is frequent to observe weak G's after C's or A's – see Hills *et al.* (1997) for more information on peak patterns). However, this “effect of the adjacent base” affects all the samples in the same way, as demonstrated by a strong correlation of the heights of the peaks between samples (for six samples, Spearman-R ranged between 0.88 and 0.97, for 70 analysed nucleotide positions of a homozygous segment).

After the reconstruction of the genomes involved, it was possible to assess the ploidy level and to determine the hybrid genome constitution. In artificial hybrids we obtained P/P+A mean values of 0.29 ± 0.08 for PAA and 0.70 ± 0.07 for PPA hybrids. For all of the six groups the P/P+A values of the PAA and PPA triploids did not overlap with those of the diploids (forced to be 0.50). All differences were highly significant even after Bonferroni correction for multiple comparisons ($N=118$, p values= 4.5×10^{-20} to 2.4×10^{-16} , Wilcoxon tests). The “count method” was also applied and the resulting P-count was significantly different: $0.71\% \pm 0.84$ for PAA and $96.32\% \pm 5.76$ for PPA hybrids ($p=0.0036$, $N_1=N_2=6$, Mann-Whitney test). The data for each group of artificial hybrids are summarized in Table I.

Table I – Results from the measuring and count methods applied to the six groups of artificial hybrids. Mean values and standard deviations for P/P+A ratios (both raw and uncorrected values) and P-counts calculated from the analysis of 118 sites are presented for each group. Total means and standard deviations of the P/P+A ratios and P-counts for the three types of artificial hybrids are also presented.

group	Measuring method						Count method	
	raw P/P+A ratio			corrected P/P+A ratio			P-count (%)	
	PA	PAA	PPA	cPA	cPAA	cPPA	PAA	PPA
1	0.45 $\pm$ 0.13	0.33 $\pm$ 0.11	0.63 $\pm$ 0.13	0.50	0.37 $\pm$ 0.06	0.69 $\pm$ 0.06	1.74	99.12
2	0.30 $\pm$ 0.18	0.15 $\pm$ 0.09	0.48 $\pm$ 0.21	0.50	0.29 $\pm$ 0.07	0.70 $\pm$ 0.11	0.85	97.46
3	0.31 $\pm$ 0.12	0.18 $\pm$ 0.08	0.58 $\pm$ 0.17	0.50	0.32 $\pm$ 0.04	0.77 $\pm$ 0.08	0.00	99.15
4	0.47 $\pm$ 0.12	0.34 $\pm$ 0.10	0.65 $\pm$ 0.10	0.50	0.36 $\pm$ 0.03	0.68 $\pm$ 0.02	0.00	100.00
5	0.27 $\pm$ 0.20	0.05 $\pm$ 0.04	0.58 $\pm$ 0.27	0.50	0.16 $\pm$ 0.10	0.79 $\pm$ 0.13	1.69	97.46
6	0.35 $\pm$ 0.17	0.16 $\pm$ 0.11	0.42 $\pm$ 0.21	0.50	0.26 $\pm$ 0.06	0.59 $\pm$ 0.10	0.00	84.75
total mean	0.36	0.20	0.56	0.50	0.29	0.70	0.71	96.32
total sd	0.08	0.11	0.09	0.00	0.08	0.07	0.84	5.76

The application of the “count method” to natural hybrids showed that diploids had higher peaks attributable to P-genome (55.77%±6.88; range 47.01%-65.25%), which were significantly different from the 8.13%±0.95 (range 6.84%-9.40%) calculated for the PAA triploids ($p=0.0005$, $N_1=8$, $N_2=9$, Mann-Whitney test). Among the triploids we found a morphologically distinct individual that had extremely high values of P-count (95.69%), suggesting a PPA genome. The P-count for the single tetraploid (1.69%) was lower than the smallest value for PAA triploids (6.84%), suggesting a PAAA constitution. Sections of chromatograms of the hybrids are illustrated in Figure 2.

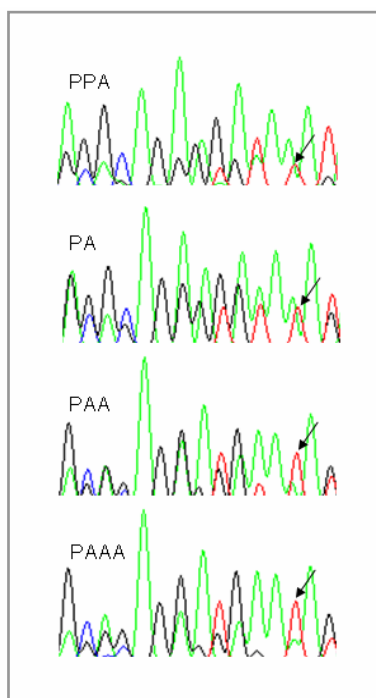


Figure 2 — Sections of the chromatograms of *S. alburnoides* natural hybrids where different dosages of A and P genomes are evident. The arrow indicates a base attributed to the non-hybrid *S. alburnoides* progenitor (A genome) that increases its relative height from PPA to PA to PAA to PAAA.

The “measuring method” also returned significant differences: mean values of P/P+A for PA and PAA hybrids were, respectively, 0.53±0.01 (range 0.51–0.55) and 0.34 ± 0.01 (range 0.33–0.36) ($p=0.0005$, $N_1=8$, $N_2=9$, Mann-Whitney test). Values of P/P+A ratio for the morphologically distinct triploid (0.70) and for the tetraploid (0.23) also differed markedly from the remaining fish, corroborating the genomic constitutions suggested by the “count method”.

The comparison between the measures taken for the same individuals in two different PCR and sequencing runs demonstrated a repeatable distinction between diploids and triploids. In the second PCR and sequencing of the same material, the mean P/P+A values were still significantly different between diploids and triploids: 0.58 ± 0.03 and 0.39 ± 0.01 , respectively ($p=0.0039$, $N_1=N_2=6$, Mann-Whitney test). The results showed a high correlation between the P/P+A values obtained for the two chromatograms from the same individual for diploids and triploids (Spearman-R ranged between 0.66 and 0.77 and between 0.76 and 0.83, respectively).

A comparison of the results obtained when lower numbers of double peaks are analysed showed that all the differences between ploidies were recovered and the values showed only slight deviations from those obtained with the entire data series (Figure 3). Indeed, even groups of 20 overlapping peaks provided estimates that were sufficiently accurate to assign the proportion of each genome in different ploidy groups. This means that even disturbances that are much shorter than the ones used in this study may recover basically the same ploidy information.

Application to intraespecific heterozygotes

Concerning intraespecific heterozygotes, four of the *Chondrostoma oligolepis* samples sequenced revealed one heterozygous indel of seven bases in the same fragment of the *beta-actin* gene, generating a double peaks region of 823bp. The reconstructed parental haplotypes were recovered in the remaining individuals (GenBank: AY943897 to AY943905): the strand with the deletion was present in homozygosity in seven individuals (CO1 genome), and the other strand was present in two individuals homozygous for the insertion (CO2 genome). Measures of the peaks in the chromatograms of the heterozygotes yielded a CO1/CO1+CO2 ratio of 0.58 ± 0.03 , a value that is comparable to that obtained for the PA diploids of *S. alburnoides* from the same PCR and sequencing run ($p=1.000$, $N_1=6$, $N_2=4$, Mann-Whitney test).

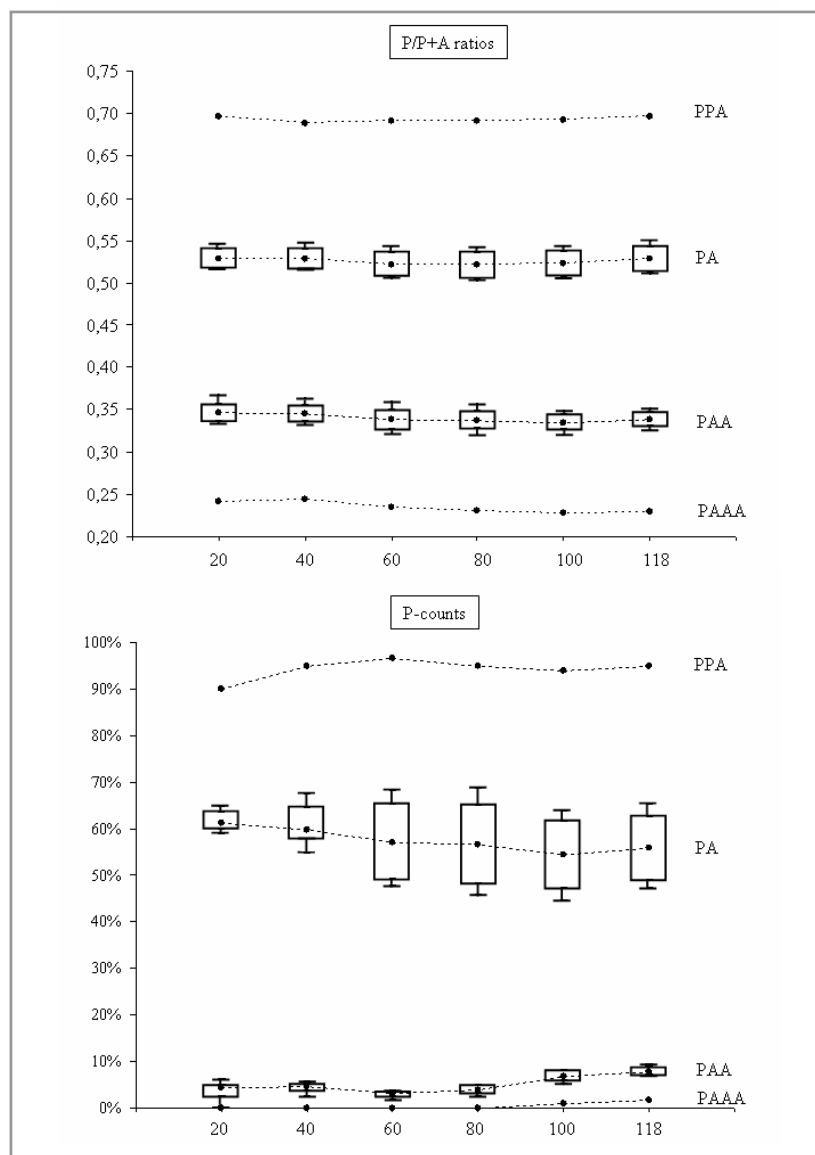


Figure 3 — Plot of means, standard deviations, minimums and maximums of the P/P+A ratios and P-counts in PA diploids (N=8), PAA triploids (N=9), PPA triploid (N=1) and PAAA tetraploid (N=1), as a function of the number of double peaks analysed (20, 40, 60, 80, 100 and 118).

Discussion

The results demonstrated that the pattern of double peaks generated by heterozygous indels proved to be useful: i) to reconstruct the parental sequences involved in large DNA segments of individuals that are heterozygotes for several linked sites and of interspecific hybrids; ii) to determine the ploidy of the hybrids; and iii) to identify the relative contributions of the parental sequences in non-diploid hybrids. The method was repeatable in different PCR and sequencing conditions and the interpretation of ploidy and

genomic proportions was always consistent. To control for possible variations we recommend that: in all PCR and sequencing runs at least one diploid hybrid should be included to serve as a standard control to that run; and results should be confirmed with forward and reverse sequencing.

One may ask which of the two methods described in this paper is the more accurate. The "count method" is much less time consuming and in the present study discriminated all the genomic constitutions that had been identified with the measuring method. However, we suggest that the "measuring method" should be preferred over the "count method" since it presents the great advantage of controlling for varying PCR artefacts and PCR conditions and to obviate the effects of neighbouring bases on the height of each peak. In addition, the "measuring method" is more reliable and powerful than the "count method" to discriminate different hybrid forms in which one of the parental complements is predominant – for instance, when we want to discriminate between a PAA and a PAAA individual.

When compared with allozyme electrophoresis, our method avoids killing the specimens and problems of regulation of gene expression. Although crude when compared with quantitative PCR procedures, it is sufficiently precise and inexpensive to be considered a useful tool in the study of interspecific hybrids and in population genetics.

DNA segments that harbour several and closely located indels provide ideal material for the application of this method as they provide several reference points that allow recovery of the specific sequences starting from both directions and minimize the risk of error caused by recombination. The process of sequence reconstruction is also favoured by the existence of several characteristic point mutations fixed for each parental species. These mutations mark out highly conserved regions and make it possible to ascribe unambiguously each peak in a double peaks region to the correct parental genome. Although advantageous, these conditions are not essential and their absence would not necessarily make the method impracticable. In fact, as referred by Bhangale *et al.* (2005), since the length of the indel is inferred from the pattern of double peaks, the process does not require the presence of homozygotes for both of the alleles in the surveyed sample, a situation that widens its applications.

As flow cytometry methodology can only determine the ploidy level of the samples and not their exact genome composition, mainly when parental species have similar DNA

contents, our approach also constitutes a valuable complementary tool for the analysis of hybrids with either balanced or non balanced parental genomes.

Acknowledgements We thank the help of G. Lemos, Sousa-Santos family, T. Bento and A. Levy. We are also grateful to A. Gomes-Ferreira, M. Gromicho and L. Moreira da Costa for technical assistance with flow cytometry. DGF provided authorizations for field work. Study funded by the FCT Pluriannual Program (UI&D 331/94 and UI&D 329/94) (FEDER participation). C. Sousa-Santos was supported by a PhD grant from FCT (SFRH/BD/8320/2002).

References

- Alves MJ, Coelho MM & Collares-Pereira MJ. 2001. Evolution in action through hybridization and polyploidy in an Iberian freshwater fish: a genetic review. *Genetica* 111: 375–385.
- Alves MJ, Collares-Pereira MJ, Dowling TE & Coelho MM. 2002. The genetics of maintenance of an all-male lineage in the *Squalius alburnoides* complex. *Journal of Fish Biology* 60: 649–662.
- Bhangale TR, Rieder MJ, Livingston RJ & Nickerson DA. 2005. Comprehensive identification and characterization of diallelic insertion-deletion polymorphisms in 330 human candidate genes. *Human Molecular Genetics* 14: 59–69.
- Collares-Pereira MJ. 1985. Cytotaxonomic studies in Iberian Cyprinids. II. Karyology of *Anaecypris hispanica* (Steindachner, 1866), *Chondrostoma lemmingi* (Steindachner, 1866), *Rutilus arcasi* (Steindachner, 1866) and *R. macrolepidotus* (Steindachner, 1866). *Cytologia* 50: 879–890.
- Cunha C, Coelho MM, Carmona JA & Doadrio I. 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Molecular Ecology* 13: 2807–2817.
- Hills HG, Elliot MD, Grills G & Spicer EK. 1997. ABRF'97: Techniques at the genome/proteome interface - DNA sequence analysis. Available from: <http://www.biotech.iastate.edu/facilities/DSSF/ABRF/default.html> via the INTERNET. Accessed 2005 Aug 31.
- Lamatsch DK, Steinlein C, Schmid M & Scharl M. 2000. Noninvasive determination of genome size and ploidy level in fishes by flow cytometry: detection of triploid *Poecilia formosa*. *Cytometry* 39: 91–95.
- Pala I & Coelho MM. 2005. Contrasting views over a hybrid complex: between speciation and evolutionary “dead-end”. *Gene* 347: 283–294.
- Robalo J, Sousa-Santos C, Levy A & Almada V (in press) Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. *Molecular Phylogenetics and Evolution*.
- Sambrook J, Fritsch EF & Maniatis T. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, New York.

Chapter 3

Phenotypic expression of different genome constitutions in a
diploid-polyploid hybridogenetic fish complex

Sousa-Santos C, Collares-Pereira MJ & Almada V (submitted)

Journal of Fish Biology

Phenotypic expression of different genome constitutions in a diploid-polyploid hybridogenetic fish complex

Sousa-Santos C, Collares-Pereira MJ & Almada V

Abstract In this paper a nuclear marker was used to differentiate the genome constitutions of individuals of the hybridogenetic Iberian fish complex *Squalius alburnoides*, aiming to analyse the phenotypes of the most common forms and to compare them with the phenotype of specimens belonging to both the maternal species and to representatives of the extinct paternal ancestor. This study showed that: 1) there are a number of diagnostic features that can be directly used in the field to differentiate *S. alburnoides* from other *Squalius* and *S. alburnoides* hybrid and nonhybrid specimens; 2) the analysed *S. alburnoides* genome constitutions may be clearly distinguished by several morphological variables; and 3) globally, *S. alburnoides* hybrid forms are intermediate between the groups representing their paternal and maternal ancestors. Moreover, an apparent gene-dosage effect on the phenotype was detected for several of the analysed morphological variables, although the inheritance of some traits seemed to be only partially additive. It is argued that such hybrid complexes may provide excellent model systems to study the relationship between genotype and phenotype expression. As the different forms of *S. alburnoides* tend to be ecologically segregated, it is also suggested that differences in morphology, colouration pattern and habitat preferences may influence the likelihood of distinct mating patterns, and thus affect the evolutionary dynamics of the complex.

Keywords *Squalius alburnoides*, hybridogenetic complex, intermediate morphology, beta-actin, gene dosage

Introduction

The external mode of fertilization of cyprinid fishes is certainly one of the factors favouring hybridization even among distantly related species which use the same spawning grounds (Scribner *et al.* 2001). Hybrids resulting from these occasional matings may share

a number of characters with each one of the parental species and display others that are intermediate between those of the parents (*e.g.* Albertson & Kocher 2001; Mateos & Vrijenhoek 2002; Gante *et al.* 2004). Hybridization episodes that lead to the foundation of a new, stable and ecologically successful species were traditionally considered uncommon, although an increasing number of instances highlighting the role of hybridization in speciation have been already reported (Coyne & Orr 2004).

The *Squalius alburnoides* (Steindachner) complex, a hybrid complex of cyprinid fishes endemic to the Iberian Peninsula, configures one of these instances. Its origin involved intergeneric crosses between *Squalius pyrenaicus* (Günther) females and a presumably extinct *Anaecypris*-like paternal ancestor (Alves *et al.* 2001; Crespo-López *et al.* 2006; Gromicho *et al.* 2006; Robalo *et al.* 2006). Peculiar reproductive modes and gametogenic processes, namely hybridogenesis, production of unreduced gametes and meiotic hybridogenesis (Carmona *et al.* 1997; Alves *et al.* 1997a, 2001, 2004; Gromicho & Collares-Pereira 2004; Crespo-López *et al.* 2006; Sousa-Santos *et al.* 2007), were responsible for the production of several forms within the complex, differing in their ploidy level and genome constitution: SA, SSA, SAA, SSAA, SAAA, SSSA, and also a non-hybrid diploid AA form ("S" and "A" denotes the *Squalius* and the *Anaecypris*-like complements, respectively). This non-hybrid form exhibit *S. pyrenaicus*-like mtDNA, indicating that it has resulted from crosses between the hybrid forms (Alves *et al.* 2002). For instance, SAA females produce an important fraction of A gametes because during egg formation they exclude the S genome, after which the two A chromosome sets undergo normal meiosis. These A eggs fertilized by an A-sperm regenerate the AA genome, giving rise almost entirely to males - only two females were reported so far, both from the Guadiana river basin (Carmona 1997; Sousa-Santos *et al.* 2006c). All the other *S. alburnoides* hybrid forms – diploids, triploids and tetraploids – include fish of both sexes, although triploid females are usually the dominant form. Triploid males and tetraploids of both sexes are rare in natural populations (Alves *et al.* 2001; Pala & Coelho 2005). Although the reasons are yet unclear, the relative frequency with which the *S. alburnoides* forms occur may differ among drainages and some forms do not exist in some populations (*e.g.* non-hybrid males occur only in the south, being absent from the northern populations of Douro and Mondego; and tetraploids and diploid hybrid males were never captured in the Guadiana drainage - Alves *et al.* 2001; Pala & Coelho 2005).

Although of relatively recent Pleistocenic origin in the bulk of Iberia (Cunha *et al.* 2004; Sousa-Santos *et al.* in press), the complex shows a wide distribution range, occupying at least nine distinct river basins. As it dispersed to the peripheral areas of the Iberian Peninsula, it became sympatric with already established populations of other *Squalius* species: *Squalius carolitertii* (Doadrio) in the north, *S. pyrenaicus* in the south and *Squalius aradensis* (Coelho, Bogutskaya, Rodrigues & Collares-Pereira) in the southwest (see Figure 1). It is known that *S. alburnoides* may interbreed with them and nuclear and mitochondrial gene transfers have already been evidenced in several studies (Alves *et al.* 1997b; Cunha *et al.* 2004; Sousa-Santos *et al.* 2006b).

Contrasting with the considerable amount of research involving genetic aspects of its reproductive biology, the phenotypic variation within the complex is poorly known and no attempts have been made to search for concordance between *S. alburnoides* genome constitutions and the distinct phenotypes. Indeed, Collares-Pereira (1984) described the complex as comprising two morphological forms, mainly differentiated by the number of gillrakers and pharyngeal teeth, that were latter identified as being hybrid and non-hybrid forms (Alves *et al.* 1997a). More recently, Martins *et al.* (1998) failed to distinguish *S. alburnoides* of different ploidy levels on a morphological basis although they found sex-related differences. Careful research on this topic is however of great relevance to a proper understanding of the origin and maintenance of this hybridogenetic complex. Indeed, if fish with different genome constitutions differ in their morphology and other phenotypic traits, they may have distinct life history and ecological traits (as apparently happens in *S. alburnoides*, as demonstrated by Martins *et al.* 1998; Ribeiro *et al.* 2003 and Gomes-Ferreira *et al.* 2005) which, in turn, may drastically influence the probability of crosses between the different forms.

Thus, the aim of the present study was to search for morphological characters that might be helpful in sorting the distinct *S. alburnoides* forms. This kind of investigations may also help to search for the mechanisms of gene expression and interaction in hybrid organisms.

Materials and methods

A total of 70 specimens from *S. alburnoides* were analysed, covering six of the seven Portuguese river basins where the complex has been found (see Figure 1 for sampling

locations): Douro (N=7), Mondego (N=9), Tagus (N=19), Sado (N=13), Guadiana (N=4) and Quarteira (N=20).

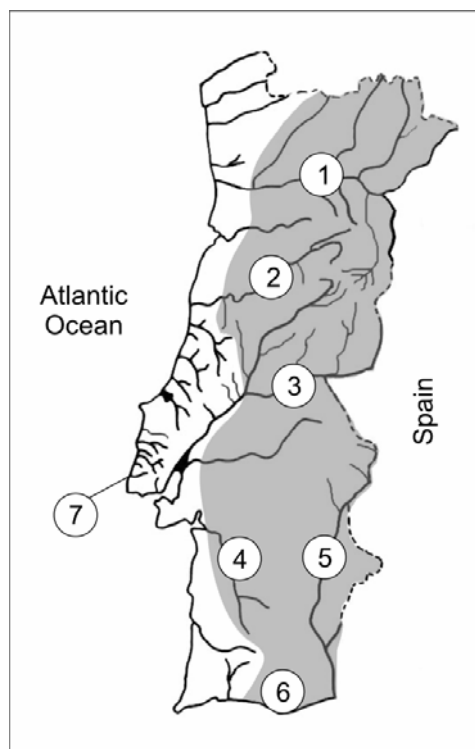


Figure 1 – Distribution area of *S. alburnoides* in Portugal (in grey) and sampling locations: 1 – Douro drainage (River Sabor); 2 – Mondego drainage (River Ceira); 3 – Tagus drainage (Rivers Sorraia, Zêzere, Erges and Ocreza); 4– River Sado; 5 – Guadiana drainage (Rivers Caia and Ardila); 6 – River Quarteira; 7 – River Colares.

Total genome DNA was extracted from fin clips preserved in ethanol by an SDS/proteinase-k based protocol, precipitated with isopropanol and washed with ethanol before re-suspension in water (adapted from Sambrook *et al.* 1989). The genome constitution of the *S. alburnoides* individuals was assessed by sequencing the beta-actin nuclear gene (935bp) – primers and PCR conditions may be found in Sousa-Santos *et al.* (2005). Sequences of the homozygous individuals were aligned with BioEdit v.5.0.6. For heterozygous diploid and polyploid individuals, the genome complements had to be recovered following the procedures described in Sousa-Santos *et al.* (2005) before the alignment process. As mentioned above, the nuclear haplotypes derived from the paternal ancestor of the *S. alburnoides* complex were designated as A-haplotypes and the ones derived from the maternal ancestor were designated as S-haplotypes. However, when the geographical location of the specimens was worth mention, the letter “S” was replaced by the letters “P”, “C” or “Q”, which stand for the sympatric *Squalius* complements, respectively, *S. pyrenaicus* (in the southern populations), *S. carolitertii* (in the northern populations) or *S. aradensis* (in the southwestern populations of Quarteira river

basin). Genbank accession numbers of all the analysed *S. alburnoides* (already published elsewhere) are presented in Table I. A sample of *S. pyrenaicus* (N=26) from River Colares (a small western drainage where *S. alburnoides* is absent) was also included in the analysis and data concerning the typical number of gillrakers of *S. pyrenaicus* were obtained from Rodrigues (1999).

Table I – Genome constitution and GenBank accession numbers of the *S. alburnoides* specimens analysed. The nuclear haplotypes derived from the maternal ancestor of the *S. alburnoides* complex were designated as A-haplotypes and the ones derived from the paternal ancestor were designated as P-, C- or Q-haplotypes according to the sympatric *Squalius*. *S. pyrenaicus* (in the southern populations), *S. carolitertii* (in the northern populations) or *S. aradensis* (in the southwestern populations of Quarteira river basin), respectively. In the presence of distinct P, C, Q or A haplotypes in the same individual the designations P', C', Q' or A' were used. GenBank accession numbers were already published by: 1 – Sousa-Santos *et al.* (2006c), 2 – Sousa-Santos *et al.* (2006b), 3 – Sousa-Santos *et al.* (2005), and 4 – Sousa-Santos *et al.* (in press).

Label	Drainage	Genome	GenBank Accession Number	Public.
ALBG2N153	Guadiana	AA	DQ010337	1
ALBQ12	Quarteira	AA'	strand A1: EF459360; strand A2: EF459361	2
ALBQ13	Quarteira	AA	AY943867	3
ALBQ14	Quarteira	AA	DQ128102	1
ALBQ15	Quarteira	AA	EF458692	4
ALBQ17	Quarteira	AA	EF458693	4
ALBQ11	Quarteira	AA'	strand A1: EF459358; strand A2: EF459359	4
ALBQ16	Quarteira	AA'	strand A1: DQ150335; strand A2: DQ150336	2
ALBQ19	Quarteira	AA'	strand A1: EF459382; strand A2: EF459383	4
ALBQ21	Quarteira	AA'	strand A1: EF459384; strand A2: EF459385	4
ALBQ8	Quarteira	AA'	strand A1: DQ150341; strand A2: DQ150342	2
ALBAA13	Tagus	AA	AY943896	3
ALBAA159	Tagus	AA	EF458695	4
ALBAA160	Tagus	AA	EF458696	4
ALBAA164	Tagus	AA	AY943895	3
ALBAA44	Tagus	AA	AY943863	3
ALBTJSE28	Tagus	AA	AY943894	3
ALBQ18	Quarteira	AA'	strand A1: EF459380; strand A2: EF459381	4
ALBG2N155	Guadiana	PA	strand P: DQ335474; strand A: DQ335475	1
ALBQ55	Quarteira	QA	strand A: DQ150355; strand P: DQ150356	2
ALBLSAD13	Sado	PA	strand P: EF459390; strand A: EF459391	4
ALBLSAD2	Sado	PA	strand P: EF459388; strand A: EF459389	4
ALBSAD17	Sado	PA	strand P: EF459394; strand A: EF459395	4
ALBLSAD4	Sado	PA	strand P: EF459392; strand A: EF459393	4
ALB2N128	Tagus	PA	strand P: EF459400; strand A: EF459401	4
ALB2N31	Tagus	PA	strand P: EF459405; strand A: EF459406	4
ALB2N68	Tagus	PA	strand P: EF459396; strand A: EF459397	4
ALBD3N145	Douro	CAA	strand P: EF459199; strand A1: EF459200; strand A2: EF459201	4
ALBDSA1	Douro	CAA	strand P: EF459193; strand A1: EF459194; strand A2: EF459195	4

(Table 1 - cont.)

Label	Drainage	Genome	GenBank Accession Number	Public.
ALBDSA2	Douro	CAA	strand P: EF459178; strand A1: EF459179; strand A2: EF459180	4
ALBDSA3	Douro	CAA	strand P: EF459202; strand A1: EF459203; strand A2: EF459204	4
ALBDSA5	Douro	CAA	strand P: EF459196; strand A1: EF459197; strand A2: EF459198	4
ALBDSA8	Douro	CAA	strand P: EF459184; strand A1: EF459185; strand A2: EF459186	4
ALBG43B	Guadiana	PAA	strand P: DQ335480; strand A: DQ335481	1
ALBG3N154	Guadiana	PAA	strand P: DQ335472; strand A: DQ335473	1
ALBMDC29	Mondego	CAA	strand P: EF459235; strand A1: EF459236; strand A2: EF459237	4
ALBMGC14	Mondego	CAA	strand P: EF459238; strand A1: EF459239; strand A2: EF459240	4
ALBMGC18	Mondego	CAA	strand P: EF459253; strand A1: EF459254; strand A2: EF459255	4
ALBMGC19	Mondego	CAA	strand P: EF459256; strand A1: EF459257; strand A2: EF459258	4
ALBMGC21	Mondego	CAA	strand P: EF459241; strand A1: EF459242; strand A2: EF459243	4
ALBMGC22	Mondego	CAA	strand P: EF459244; strand A1: EF459245; strand A2: EF459246	4
ALBQ1	Quarteira	QAA	strand A: DQ150345; strand P: DQ150346	2
ALBQ10	Quarteira	QAA	strand A: DQ150350; strand P: DQ150351	2
ALBQ6	Quarteira	QAA	strand A: DQ150337; strand P: DQ150338	2
ALBQ9	Quarteira	QAA	strand A: DQ150343; strand P: DQ150344	2
ALBQ2	Quarteira	QAA'	strand A1: DQ150347; strand A2: DQ150348; strand P: DQ150349	2
ALBQ58	Quarteira	QAA'	strand A1: DQ150357; strand A2: DQ150358; strand P: DQ150359	2
ALBSAD11	Sado	PAA	strand P: EF459286; strand A1: EF459287; strand A2: EF459288	4
ALBSAD14	Sado	PAA	strand P: EF459298; strand A1: EF459299; strand A2: EF459300	4
ALBSAD2	Sado	PAA	strand P: EF459292; strand A1: EF459293; strand A2: EF459294	4
ALBSAD22	Sado	PAA	strand P: EF459310; strand A1: EF459311; strand A2: EF459312	4
ALBSAD38	Sado	PAA	strand P: EF459319; strand A1: EF459320; strand A2: EF459321	4
ALBSAD25	Sado	PAA	strand P: EF459313; strand A1: EF459314; strand A2: EF459315	4
ALB3N10	Tagus	PAA	strand P: EF459337; strand A1: EF459338; strand A2: EF459339	4
ALB3N100	Tagus	PAA	strand P: EF459340; strand A1: EF459341; strand A2: EF459342	4
ALB3N53	Tagus	PAA	strand P: EF459352; strand A1: EF459353; strand A2: EF459354	4
ALB3N96	Tagus	PAA	strand P: EF459349; strand A1: EF459350; strand A2: EF459351	4
ALBTJO1	Tagus	PAA'	strand P: EF459334; strand A1: EF459335; strand A2: EF459336	4
ALBLDCA1R	Douro	CCA	strand P1: EF459148; strand P2: EF459149; strand A1: EF459150	4
ALBMGC10	Mondego	CC'A	strand P1: EF459274; strand P2: EF459275; strand A1: EF459276	4
ALBMGC9	Mondego	CC'A	strand P1: EF459229; strand P2: EF459230; strand A1: EF459231	4
ALBMGC33	Mondego	CCA	strand P1: EF459226; strand P2: EF459227; strand A1: EF459228	4
ALBQ28	Quarteira	QQ'A	strand A: DQ150352; strand P1: DQ150353; strand P2: DQ150354	2
ALBSAD27	Sado	PPA	strand P1: EF459289; strand P2: EF459290; strand A1: EF459291	4
ALB3N20	Tagus	PPA	strand P1: EF459346; strand P2: EF459347; strand A1: EF459348	4
ALBTJZ1	Tagus	PP'A	strand P1: EF459355; strand P2: EF459356; strand A1: EF459357	4

All fish were captured by electrofishing, preserved in ethanol after being killed with an overdose of anesthetic MS-222 and deposited in the collection of UIE/ISPA (Lisbon). The

morphological study involved the acquisition of 25 morphometric and 9 meristic variables (most of them described in Figure 2): standard length (S_L); head length (H_L); head height (H_H); pre-orbital head length (P_{ROL}); post-orbital head length (P_{SOL}); eye diameter (E_D); distance between the eye and the body margin (D_{EB}); pre-dorsal fin length (P_{rD}); pre-ventral fin length (P_{rV}); pre-anal fin length (P_{rA}); distance between the insertions of the ventral and anal fins (V_A); dorsal fin length (D_{FL}); dorsal fin base width (D_{FW}); pectoral fin length (P_{FL}); pectoral fin base width (P_{FW}); ventral fin length (V_{FL}); ventral fin base width (V_{FW}); anal fin length (A_{FL}); anal fin base width (A_{FW}); maximum width of the third circum-orbital bone (W_{COB}); axillary scale length (A_S); interorbital width (I_{OW}); snout angle (S_{NA}); upper snout profile perimeter (U_{SP}); lower snout profile perimeter (L_{SP}); scales in the lateral line (S_{LL}); number of transverse rows above lateral line (S_{UTR}); number of transverse rows under lateral line (S_{LTR}); branched dorsal fin rays (D_R); branched anal fin rays (A_R); gillrakers on the left first gill arch (G_L); gillrakers on the right first gill arch (G_R); left bone pharyngeal teeth (P_L) and right bone pharyngeal teeth (P_R). Measurements were made with a digital display calliper (0.01mm accuracy). For a detailed description of each morphometric and meristic variable see Collares-Pereira (1983). Five morphometric indices were also used to describe particular aspects of the morphology: H_L/H_H , E_D/H_L , A_S/V_{FW} , P_{rD}/P_{rV} and U_{SP}/L_{SP} (for descriptions of the abbreviations see the text above).

To reduce the impact of size, the standardized residuals of the regression of the logarithm of each morphometric variable on the logarithm of standard length (S_L) were calculated and used in subsequent analyses (Reist 1985). Discriminant analyses were used to determine the variables that most contribute to the distinction of previously defined groups and to evaluate whether each individual was correctly assigned to its group according to the generated model.

To allow the inclusion of the pharyngeal teeth formula and the intensity of colouration of the longitudinal bands in the discriminant analysis, these discrete variables were coded as follows. For the pharyngeal teeth formula the following numerical codes were used: 1 – uniseriated teeth only; 2 – uniseriated and biseriated teeth; and 3 – biseriated teeth only. As regards the colouration of the longitudinal band, since it was absent in *S. pyrenaicus*, almost imperceptible in SSA individuals, more visible in SA and SAA forms and darker in AA individuals (see Results), the numerical codes adopted to designate the colouration of these four forms were, respectively: 0, 1, 2, 3 and 4.

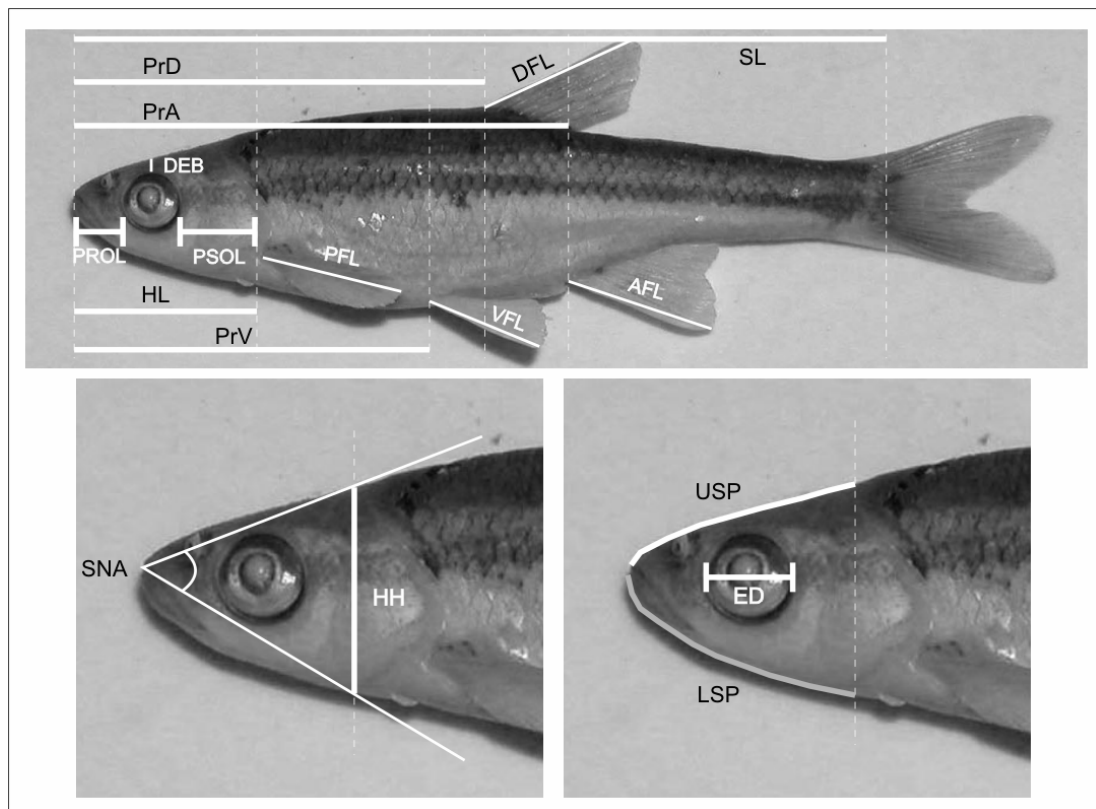


Figure 2 — Description of the morphological variables analysed (some variables - AS, WCOB, IOL, VA and the base widths of all the fins - could not be represented since the reference points were not visible in the image). See legends in the “Materials and Methods” section.

The existence of significant differences between the mean values of the variables studied was assessed using ANOVA followed by post-hoc pairwise comparisons with the appropriate Bonferroni corrections. Statistica 7.0 software (StatSoft, Inc.) was used for all the statistical analyses performed.

Results

The beta-actin gene sequencing enabled the identification of four distinct genome constitutions in *S. alburnoides*: non-hybrids (AA), diploid hybrids (PA and QA), triploid hybrids with duplicated *Anaocypris*-like complements (PAA, CAA and QAA) and triploid hybrids with duplicated *Squalius* complements (PPA, CCA and QQA) – see Table II. Tetraploids were also detected but they were discarded due to their low number (N=2).

Table II – Genome constitution and geographical provenience of the *S. alburnoides* individuals analysed. The northern populations of *S. alburnoides* are sympatric with *S. caroliteirii* (C-genome), the southern ones with *S. pyrenaicus* (P-genome) and the southwestern population of Quarteira with *S. aradensis* (Q-genome). The mean, standard deviations (sd), minimum (min), and maximum (max) standard lengths of the global samples (containing all the specimens of each specific genome constitution) are also presented.

Drainage		Nuclear genome			
		AA	CA	CAA	CCA
North	Douro	-	-	6	1
	Mondego	-	-	6	3
		AA	PA	PAA	PPA
South	Tagus	6	5	5	2
	Sado	-	6	6	1
	Guadiana	1	1	2	-
		AA	QA	QAA	QQA
Southwest	Quarteira	11	1	6	1
		AA	SA	SAA	SSA
TOTAL		18	13	31	8
Standard length (mm)		45.3±7.8	68.2±18.6	65.1±15.1	65.9±19.5
mean±sd (min – max)		(38.0 – 66.7)	(39.4 – 104.8)	(39.3 – 102.5)	(33.9 – 98.1)

Colour pattern

Live specimens of the *S. alburnoides* complex show a dark line along the dorsal fin base (extremely well visible when the fish is seen from above, while swimming), that is absent from all the other Iberian *Squalius*. Another external difference between *S. alburnoides* and other *Squalius* species is the presence of a longitudinal dark band along the flanks of the fish (Collares-Pereira 1984), instead of the typical silvery coloration of the flanks of the other *Squalius* species. This line was darker in AA and SAA forms, less conspicuous in SA individuals and, in most cases, almost undetectable in live SSA individuals (although after death the line became more visible) – Figure 3.

Non-hybrids exhibited at least two external features that differ markedly from the hybrid forms: presence of a typical straight forward descending anterior portion of the lateral line (hybrids showed a curvilinear lateral line); and a longitudinal band very dark coloured, extremely well marked and extending forward to the posterior edge of the eye

(in hybrids the contrast was not so intense, the limits of the band were attenuated and it extended only to the operculum).

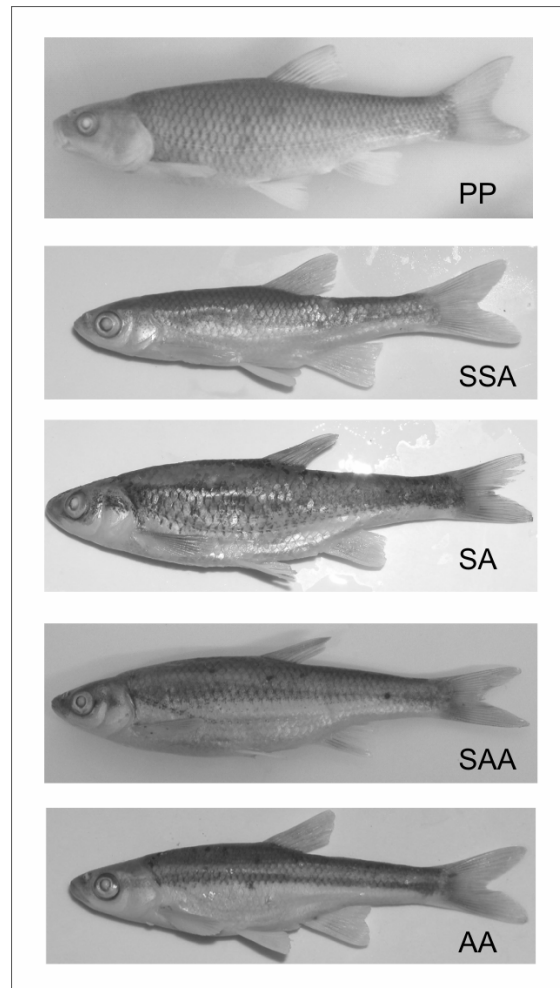


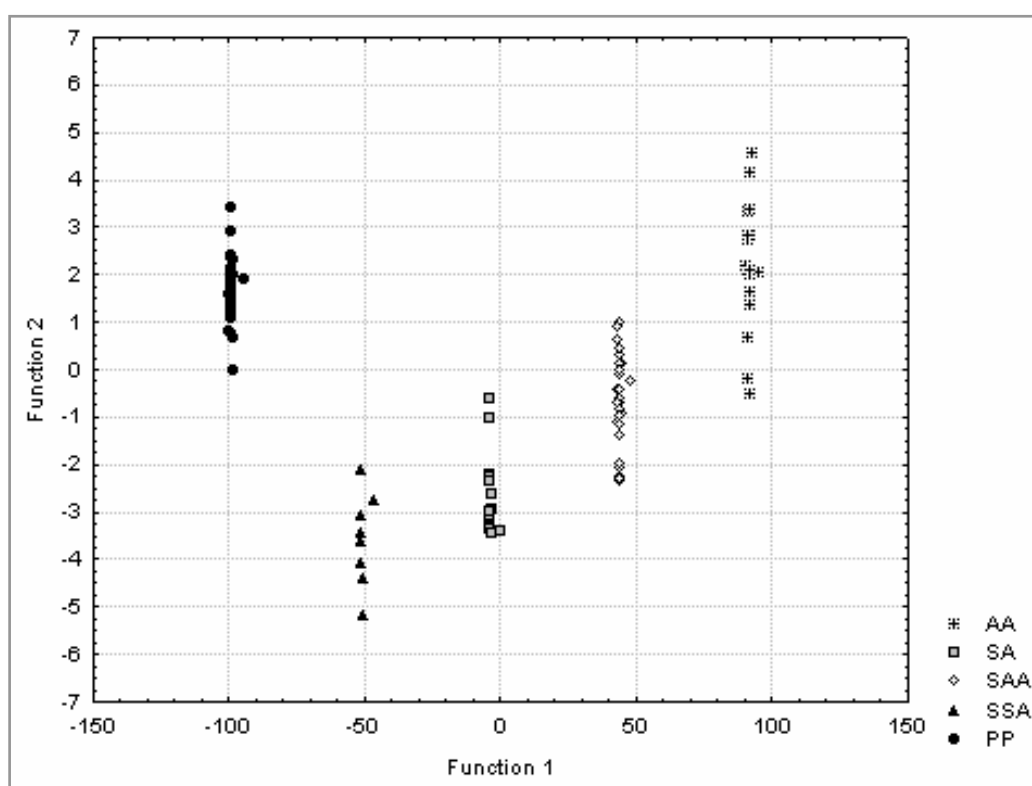
Figure 3 — Morphological continuum of *S. alburnoides* hybrid forms, *S. pyrenaicus* and non-hybrid *S. alburnoides*. From top to bottom: PP, SSA, SA, SAA and AA forms.

Morphological pattern of the distinct S. alburnoides genome constitutions

S. alburnoides were grouped according to their genome constitution to test for the existence of morphological differences, regardless of their geographical provenance. A fifth group containing pure *S. pyrenaicus* (PP), the maternal ancestor of the complex, from Colares (N=26) was also included in the discriminant analysis. The forward stepwise discriminant analysis (21 variables included in the model) was significant ($p \approx 0.000$; Wilk's

$\Lambda = 0.00001$; $F(84.282)=61.032$). The classification scores obtained were 100% for all five groups. The p-values associated with the calculated Mahalanobis distances showed that all groups were significantly differentiated ($p \approx 0.000$ between all pairs). Canonical analysis showed that the hybrid forms of *S. alburnoides* were intermediate between the AA and PP groups, the SSA group being more closely related to PP, and SAA the group closest to AA (Figure 4).

Figure 4 — Projection of the centroids of the groups SA, SAA, SSA, AA and PP in the space defined by the first two canonical discriminant functions.



The discriminant power of each variable included in the model is shown in Table III. The variables G_L , G_R , W_{COB} and I_{OW} were *a priori* discarded from this discriminant analysis since we had no data on these variables for the PP group.

According to the results of ANOVA, the variables that showed significantly distinct mean values for each pair of genotypes are listed in Table IV.

These results, together with the calculated mean values for each variable (Table V), showed that AA non-hybrids may be distinguished from all the studied hybrid forms by the significant higher numbers of gillrakers in both arches, of scales in the lateral line and

of transverse rows above lateral line, by a significantly more upturned mouth (lower U_{SP}/L_{SP}) and by significantly smaller axillary scales (A_S/V_{FW}).

Table III — Wilks' Lambda, F-values and associated p-values for all the variables included in the model (ordered according to decreasing F-values). Significant p-values are marked with *.

Variable	Wilks' Lambda	F	p-level
Longitudinal band	0,0047	9626,3920	0,0000 *
E_D/H_L	0,0000	8,1535	0,0000 *
P_{FL}	0,0000	7,2632	0,0001 *
P_{rd}/P_{rv}	0,0000	7,1080	0,0001 *
S_{UTR}	0,0000	6,6463	0,0001 *
D_R	0,0000	6,0833	0,0003 *
S_{NA}	0,0000	4,6582	0,0021 *
D_{EB}	0,0000	4,3706	0,0032 *
Pharyngeal teeth	0,0000	4,0512	0,0051 *
H_H	0,0000	3,0262	0,0231 *
P_{rA}	0,0000	2,9821	0,0246 *
P_{rv}	0,0000	2,9281	0,0267 *
D_{FL}	0,0000	2,6807	0,0384 *
V_A	0,0000	2,1565	0,0827
S_{LTR}	0,0000	2,0574	0,0955
A_{FL}	0,0000	1,6966	0,1603
S_{LL}	0,0000	1,6043	0,1827
P_{FW}	0,0000	1,5480	0,1977
U_{SP}/L_{SP}	0,0000	1,3609	0,2562
P_{SOL}	0,0000	1,0186	0,4037
A_R	0,0000	0,9916	0,4179

Additionally, after correcting for size, non-hybrids had significantly larger eyes (E_D/H_L) than SAA and SA hybrids; significantly less dorsal fin rays and smaller interorbital distances than SA and SSA hybrids; higher heads (higher H_H) than SAA hybrids; and, when compared to SA hybrids, their ventral fins were significantly more posterior (lower P_{rd}/P_{rv}) and they had significantly longer pectoral and ventral fins and more anal fin rays. Among the hybrids, the SA form was distinct from the SAA form by significantly shorter pectoral fins and by exhibiting significantly lower numbers of upper transverse rows, gillrakers in both arches and anal fin rays. SA hybrids were also differentiated from SSA hybrids by significantly shorter pectoral and ventral fins. Finally, SAA showed significantly higher numbers of gillrakers in both arches than SSA hybrids.

The results of ANOVA also corroborated the above mentioned distances between the PP group and the four forms of *S. alburnoides*: a higher number of variables with

significantly distinct mean values was detected between the PP and AA and between PP and SAA groups (Table IV).

Table IV — ANOVA results showing which variables (Var) had significant differences at $p < 0.05$ when the mean values for each *S. alburnoides* group (SAA, SSA, SA and AA) were compared with each other and with that of the PP group, using Bonferroni post-hoc tests (see Methods section for codes of the variables).

Genome constitution	AA	SAA	SA	SSA
	Var p	Var p	Var p	Var p
SAA	S_{LL} 0.000 S_{UTR} 0.000 U_{SP}/L_{SP} 0.007 G_L <0.001 G_R <0.001 H_H 0.012 E_D/H_L 0.011 A_S/V_{FW} <0.001			
SA	S_{LL} 0.000 S_{UTR} 0.000 D_R 0.024 U_{SP}/L_{SP} <0.001 G_L 0.000 G_R 0.000 P_{rD}/P_{rV} 0.018 E_D/H_L 0.001 A_R 0.047 I_{OW} 0.035 V_{FL} 0.007 P_{FL} <0.001 A_S/V_{FW} <0.001	S_{UTR} 0.007 A_R 0.013 G_L <0.001 G_R 0.001 P_{FL} <0.001		
SSA	S_{LL} <0.001 S_{UTR} 0.000 D_R 0.014 U_{SP}/L_{SP} <0.001 I_{OW} 0.005 G_L 0.000 G_R 0.000 A_S/V_{FW} <0.001	G_L 0.008 G_R 0.040	P_{FL} 0.005 V_{FL} 0.011	
PP	S_{LL} <0.001 S_{UTR} 0.000 S_{LTR} <0.001 D_R 0.000 S_{NA} 0.006 U_{SP}/L_{SP} 0.000 H_L/H_H 0.000 E_D/H_L 0.000 A_S/V_{FW} 0.000 P_{rD}/P_{rV} 0.000 H_L 0.010 P_{SOL} <0.001 P_{rV} 0.003 P_{rA} 0.011	S_{UTR} 0.000 S_{LTR} 0.029 D_R 0.000 S_{NA} 0.000 U_{SP}/L_{SP} <0.001 H_L/H_H <0.001 E_D/H_L 0.000 A_S/V_{FW} 0.000 P_{rD}/P_{rV} 0.000 H_L <0.001 H_H 0.020 P_{SOL} 0.000 D_{EB} <0.001 P_{rV} 0.001 P_{rA} 0.006 A_{FW} 0.011	S_{LTR} 0.027 D_R <0.001 S_{NA} 0.000 H_L/H_H 0.007 E_D/H_L 0.000 A_S/V_{FW} <0.001 P_{rD}/P_{rV} <0.001 D_{EB} <0.001 P_{FL} <0.001 V_{FL} 0.006	S_{LTR} 0.048 D_R 0.010 S_{NA} 0.001 H_L/H_H 0.025 E_D/H_L 0.000 P_{rD}/P_{rV} <0.001 H_L 0.011 P_{SOL} 0.002 P_{rA} 0.012

Six variables showed significantly different mean values between the PP group and all the *S. alburnoides* forms (Table IV), indicating that *S. alburnoides* may be clearly differentiated from *S. pyrenaicus* by having significantly lower number of transverse rows and of dorsal fin rays and by deeper heads (lower H_L/H_H and higher S_{NA}), smaller axillary scales (A_S/V_{FW}) and more posterior dorsal fins (P_{rD}/P_{rV}) – see Table V for mean values.

Inspection of Table V also showed that for eight variables (S_{LTR} , D_R , H_L/H_H , A_S/V_{FW} , U_{SP}/L_{SP} , P_{SOL} , P_{rV} and P_{rA}) the mean values increased with increasing proportions of the *Squalius*-like genome; for five variables (S_{UTR} , E_D/H_L , P_{rD}/P_{rV} , G_L and G_R) the tendency was the reverse; and for two additional variables (S_{LL} and S_{NA}) the mean values were lower or higher than those of both non-hybrid forms (AA and PP).

Table V – Means (X), standard deviations (SD), and minimum (Min) and maximum (Max) values for all the morphological variables (except the pharyngeal teeth formula) and morphometric indices studied for the groups AA, SA, SAA, SSA and PP. To facilitate comparisons among groups, mean values for the morphometric variables were calculated after a standardization procedure that consisted in dividing each raw measure by the standard body length of the considered individual. Concerning the PP group, the values presented for G_L and G_R (marked with **) were retrieved from Rodrigues (1999) who used a sample of *S. pyrenaicus* from the Tagus drainage. Variables that showed a tendency to increase or decrease from AA to PP are marked with “+” and “-”, respectively. Variables from which the mean values of the hybrid forms were higher or lower than both the non-hybrid AA and PP forms are marked with a “°”.

	AA	SAA	SA	SSA	PP
	X±SD (min-max) N	X±SD (min-max) N	X±SD (min-max) N	X±SD (min-max) N	X±SD (min-max) N
° S_{LL}	44.56±1.54 18 (42-47)	41.65±1.84 31 (38-46)	40.77±1.17 13 (39-43)	41.14±1.57 7 (38-43)	42.12±1.18 26 (40-45)
+ S_{LTR}	2.56±0.24 18 (2-3)	2.79±0.42 31 (2-3.5)	2.71±0.37 13 (2.5-3.5)	2.64±0.48 7 (2-3.5)	3.09±0.27 28 (2.5-3.5)
+ D_R	6.89±0.32 18 (6-7)	7.03±0.18 31 (7-8)	7.31±0.63 13 (7-9)	7.43±0.53 7 (7-8)	8.00±0.38 28 (7-9)
A_R	8.11±0.47 18 (7-9)	8.13±0.56 31 (7-9)	7.62±0.51 13 (7-8)	8.14±0.38 7 (8-9)	7.89±0.31 28 (7-8)
+ H_L/H_H	1.26±0.09 18 (1.11-1.44)	1.32±0.10 31 (1.10-1.54)	1.33±0.08 13 (1.21-1.53)	1.32±0.05 8 (1.24-1.40)	1.44±0.07 28 (1.30-1.60)
- E_D/H_L	0.32±0.03 18 (0.28-0.38)	0.29±0.02 31 (0.25-0.34)	0.28±0.02 13 (0.24-0.31)	0.30±0.03 8 (0.25-0.33)	0.23±0.02 28 (0.20-0.28)
+ A_S/V_{FW}	0.73±0.14 16 (0.46-0.96)	1.02±0.20 29 (0.77-1.68)	1.12±0.26 13 (0.70-1.63)	1.25±0.24 6 (0.97-1.62)	1.45±0.22 28 (1.00-2.00)
- P_{rD}/P_{rV}	1.16±0.03 18 (1.10-1.23)	1.14±0.03 31 (1.08-1.21)	1.12±0.03 13 (1.09-1.19)	1.14±0.03 8 (1.10-1.20)	1.08±0.03 28 (0.98-1.13)
° S_{NA}	50.50±1.64 16 (47.99-53.27)	52.68±3.53 30 (44.91-57.80)	53.30±3.07 13 (48.89-58.70)	52.06±2.92 7 (49.07-57.22)	47.19±1.85 28 (43.86-51.74)
+ U_{SP}/L_{SP}	0.86±0.07 16 (0.77-0.98)	0.93±0.06 30 (0.77-1.07)	0.96±0.05 13 (0.87-1.04)	0.98±0.09 7 (0.84-1.11)	1.02±0.03 27 (0.96-1.08)
- G_L	18.24±1.56 17 (16-21)	15.94±1.65 31 (13-19)	13.82±1.25 13 (12-16)	13.88±1.64 8 (11-16)	8.81±0.92 276** (6-12)**
- G_R	18.06±1.39 17 (15-20)	15.68±1.60 31 (13-19)	13.69±1.11 13 (12-16)	14.00±1.93 8 (11-17)	8.76±0.91 276** (6-12)**
W_{COB}	0.01±0.00 17 (0.01-0.02)	0.01±0.00 30 (0.00-0.02)	0.02±0.01 9 (0.01-0.06)	0.01±0.00 7 (0.01-0.02)	- -

(Table V - cont.)

	AA		SAA		SA		SSA		PP	
	X±SD (min-max)	N	X±SD (min-max)	N	X±SD (min-max)	N	X±SD (min-max)	N	X±SD (min-max)	N
I _{OW}	0.08±0.02 (0.06-0.14)	18	0.08±0.00 (0.07-0.09)	31	0.09±0.02 (0.07-0.13)	13	0.09±0.01 (0.08-0.09)	8	-	-
H _L	0.23±0.01 (0.21-0.26)	18	0.23±0.02 (0.19-0.25)	31	0.24±0.02 (0.21-0.27)	13	0.23±0.01 (0.22-0.24)	8	0.25±0.01 (0.23-0.27)	28
H _H	0.18±0.01 (0.17-0.20)	18	0.17±0.01 (0.15-0.19)	31	0.18±0.01 (0.17-0.19)	13	0.17±0.01 (0.16-0.18)	8	0.18±0.01 (0.16-0.20)	28
P _{ROL}	0.05±0.01 (0.04-0.08)	18	0.05±0.00 (0.04-0.06)	31	0.05±0.01 (0.05-0.07)	13	0.05±0.00 (0.04-0.06)	8	0.06±0.01 (0.05-0.07)	28
E _D	0.07±0.01 (0.06-0.08)	18	0.07±0.01 (0.05-0.08)	31	0.07±0.01 (0.05-0.08)	13	0.07±0.01 (0.06-0.07)	8	0.06±0.01 (0.05-0.07)	28
+ P _{SOL}	0.10±0.01 (0.09-0.12)	18	0.11±0.01 (0.09-0.13)	31	0.11±0.01 (0.10-0.12)	13	0.11±0.01 (0.10-0.13)	8	0.14±0.01 (0.11-0.15)	28
D _{EB}	0.01±0.00 (0.01-0.02)	18	0.01±0.00 (0.01-0.02)	31	0.01±0.00 (0.01-0.02)	13	0.01±0.00 (0.01-0.02)	8	0.02±0.00 (0.01-0.03)	28
P _{FD}	0.52±0.02 (0.49-0.57)	18	0.53±0.01 (0.50-0.55)	31	0.52±0.01 (0.50-0.55)	13	0.53±0.01 (0.51-0.54)	8	0.53±0.01 (0.50-0.55)	28
+ P _{rv}	0.45±0.02 (0.42-0.48)	18	0.46±0.01 (0.44-0.49)	31	0.47±0.01 (0.44-0.49)	13	0.46±0.01 (0.44-0.48)	8	0.49±0.01 (0.47-0.51)	28
+ P _{ra}	0.64±0.02 (0.60-0.68)	18	0.65±0.01 (0.63-0.69)	31	0.67±0.01 (0.65-0.69)	13	0.65±0.01 (0.63-0.66)	8	0.69±0.01 (0.66-0.71)	28
V _A	0.16±0.03 (0.12-0.28)	18	0.17±0.02 (0.15-0.24)	31	0.17±0.01 (0.15-0.20)	12	0.15±0.01 (0.14-0.17)	8	0.17±0.02 (0.14-0.20)	28
D _{FL}	0.20±0.02 (0.15-0.22)	18	0.20±0.01 (0.17-0.23)	31	0.19±0.01 (0.16-0.20)	13	0.20±0.01 (0.19-0.22)	8	0.19±0.01 (0.16-0.21)	28
D _{FW}	0.11±0.03 (0.09-0.23)	18	0.11±0.01 (0.10-0.12)	31	0.11±0.02 (0.09-0.16)	13	0.11±0.01 (0.10-0.12)	8	0.11±0.01 (0.09-0.15)	28
P _{FL}	0.20±0.03 (0.16-0.24)	12	0.19±0.02 (0.14-0.22)	23	0.17±0.02 (0.13-0.19)	13	0.20±0.02 (0.18-0.23)	5	0.19±0.01 (0.17-0.21)	28
P _{FW}	0.05±0.03 (0.03-0.17)	18	0.04±0.01 (0.03-0.06)	31	0.04±0.02 (0.03-0.10)	13	0.04±0.00 (0.04-0.05)	8	0.04±0.01 (0.03-0.05)	28
A _{FL}	0.16±0.02 (0.13-0.18)	18	0.17±0.02 (0.13-0.20)	30	0.15±0.02 (0.11-0.18)	13	0.17±0.03 (0.11-0.19)	8	0.15±0.01 (0.14-0.18)	27
A _{FW}	0.12±0.03 (0.10-0.24)	18	0.12±0.01 (0.10-0.13)	31	0.11±0.02 (0.10-0.17)	13	0.12±0.01 (0.10-0.14)	8	0.10±0.01 (0.08-0.11)	28
V _{FL}	0.16±0.01 (0.12-0.17)	17	0.15±0.01 (0.12-0.18)	29	0.14±0.01 (0.11-0.16)	13	0.16±0.01 (0.14-0.17)	8	0.15±0.01 (0.14-0.17)	28
V _{FW}	0.05±0.03 (0.03-0.17)	18	0.04±0.00 (0.03-0.05)	31	0.04±0.02 (0.03-0.10)	13	0.04±0.00 (0.04-0.05)	8	0.04±0.00 (0.03-0.05)	28

Concerning the pharyngeal teeth formula (Table VI), while all individuals with AA and SAA genome constitutions exhibited uniserial teeth, SA and SSA genome constitutions included also individuals with two rows of pharyngeal teeth, in one or in both bones, like in *S. pyrenaicus* and in other *Squalius*. Interestingly, in these two types of hybrids, the individual variation in the pharyngeal teeth formula was surprisingly high: while AA and SAA individuals showed only three types of formulas, the remaining SSA and SA individuals showed eight distinct formulas. Moreover, the presence of biserial teeth in both bones, as typically found in other *Squalius* species, was observed in 37.5% of the

SSA individuals, in 15.4% of the SA individuals and in none of the SAA and AA analysed specimens. Conversely, 100% of AA and SAA, 46.16% of SA and 12.5% of SSA individuals showed uniseriated teeth in both bones.

Table VI – Types of pharyngeal teeth formulas (and respective percentages of occurrence) identified for each group of individuals analysed. Values for the PP group were retrieved from Rodrigues (1999).

	N	Left bone	Right bone	Most common formula	Other formulas
AA	18	5 (83.33%)	5 (94.44%)	5/4 (83.33%)	4/4 (11.11%)
		4 (16.67%)	4 (5.56%)		4/5 (5.56%)
SAA	31	5 (90.32%)	5 (58.06%)	5/5 (45.16%)	5/4 (32.26%)
		4 (9.68%)	4 (41.94%)		4/4 (9.68%)
SA	13	5 (53.85%)	5 (38.46%)	5/4 (23.08%)	5+1/5+1 (15.38%)
		5+1 (30.77%)	4 (38.46%)	5/5 (23.08%)	5/5+1 (7.69%)
		4+1 (7.69%)	5+1 (23.08%)		4+1/5 (7.69%)
		5+2 (7.69%)			5+1/4 (7.69%)
					5+1/5 (7.69%)
					5+2/4 (7.69%)

Discussion

A visual inspection of *S. alburnoides* specimens clearly showed a gradient of morphotypes that progressively acquire a more fusiform body and, in the extreme, almost lose the dark longitudinal band along the flanks that is characteristic of the complex in a way that may lead to misidentifications with other *Squalius* species. However, the general shape of the fishes that make their appearance look more or less fusiform was not fully captured in the present analysis since the geometric methodology was not applied. This was due to the fact that, especially in females, the abdomen varies markedly with the reproductive condition, thus a statistically valid comparison would require the analysis of fish of the same sex and breeding stage, which was not possible.

Concerning the dark band, it seems reasonable to conclude that it is determined by the A-genome and that the intensity of its expression reflects the proportion of the A-complements in the genome of the specimen. The variation that follows a gradient from AA to PP genome also holds for the pharyngeal teeth from exclusively uniseriated to

exclusively biseriated, the position of the mouth from upturned to almost terminal (increasing U_{sp}/L_{sp}), the relative size of the axillary scale relative to the ventral fin base (increasing A_s/V_{FW}) and the dorsal fin length.

Despite this continuum, it is possible to distinguish *S. alburnoides* from other *Squalius* when observing live specimens. Indeed, the results here presented showed that, in the field, *S. alburnoides* may be unambiguously identified by the presence of a dark line along the dorsal fin base that is absent from all the other *Squalius* species and also by the presence of a longitudinal dark band along the flanks, instead of the typical silvery coloration of the flanks of the other *Squalius* species. However, since this later aspect may lead to misidentifications, mainly between SSA and SS individuals due to the almost imperceptible dark bands in the former, the above mentioned presence of a dark line along the dorsal fin base should be preferred as a diagnostic feature for live specimens of all the forms of the complex.

The distinction between *S. alburnoides* hybrids and non-hybrids is possible without using meristic and morphometric variables, which may be helpful in the field. In fact, non-hybrids (AA) may be differentiated from hybrids by-eye since they show a typical straight forward descending anterior portion of the lateral line and extremely well marked longitudinal bands that extend to the posterior edge of the eye. Using preserved specimens, regardless of the drainage to which they belong, the analysed forms of the *S. alburnoides* complex may be morphologically differentiated from each other (synthesized in Figure 5).

This study clearly showed that both S and A genomes are expressed in the *S. alburnoides* hybrids, unlike what happens in other hybridogenetic complexes such as *Rana esculenta* (Linnaeus) (Tunner 2000) where only one of the parental genomes is in general expressed.

The existence of a morphological continuum that has the AA and PP non-hybrid forms as its extreme representatives and the hybrid forms (SA, SAA and SSA) in an intermediate position was already reported in other studies involving hybrid fish and their parental species (e.g. Albertson & Kocher 2001; Kittell *et al.* 2005). In this particular case with distinct ploidies, such a continuum may reflect a gene dosage effect on the phenotype exhibited by the *S. alburnoides* hybrids, since SAA and SSA genome constitutions were more closely related to AA and PP, respectively. However, since the

pattern of variation was not identical for all the variables analysed (see Table V), the general expression of both parental genomes may not be fully additive.

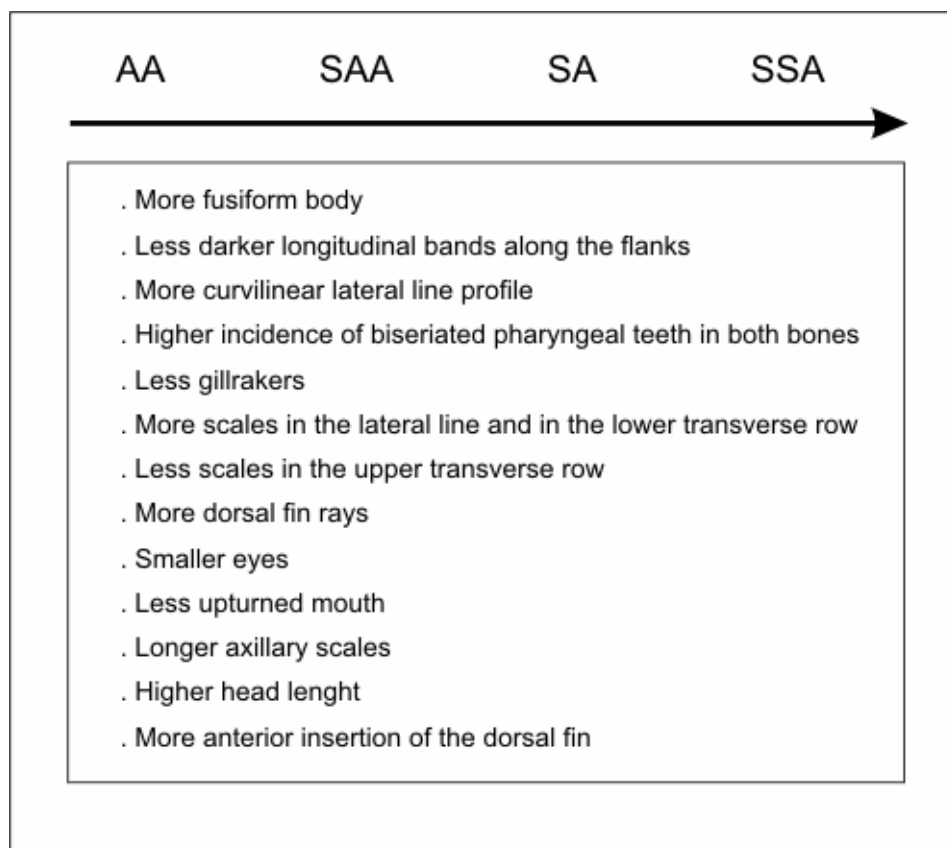


Figure 5 — Diagram synthesizing the phenotypic traits that vary in a continuum from AA to *S. alburnoides* forms and to PP.

The number of pharyngeal teeth and the intensity of the longitudinal dark band on the flanks may be seen as illustrative examples of the additive model of gene expression. The representatives of the maternal ancestor of the complex (PP genome) always exhibit two rows of pharyngeal teeth (the most common formula being 5+2/5+2) (Rodrigues 1999) and no dark band on the flanks, while the non-hybrid form (AA genome) shows only one teeth row like the extant species more closely related to the paternal ancestor of the complex, *Anaocypris hispanica* Collares-Pereira (Collares-Pereira 1983) and a dark longitudinal band on the flanks that is more intense and more extended than in all hybrid forms. Thus, as expected, some SA hybrids showed uniseriated and others biseriated pharyngeal teeth; all SAA hybrids had uniseriated teeth; and the majority of the SSA

hybrids showed biseriated teeth. Accordingly, as stated above, the intensity of the longitudinal dark band decreases with a decreasing proportion of the A to S genome.

The reported variation of the pharyngeal teeth formula in the hybrids highlights another interesting feature of these fishes: an apparent disturbance of the regulatory process involved in development such that the phenotypic expression of each genome constitution is substantially more variable than in non-hybrid forms.

With the modern techniques available to study gene expression, hybrids (especially fertile ones like *S. alburnoides*) have the potential to become excellent models to study the genetic control of the development of morphometric characters and the regulation of gene expression.

The fact that *S. alburnoides* currently hybridizes with at least three distinct *Squalius* species (Alves *et al.* 2001; Cunha *et al.* 2004; Sousa-Santos *et al.* 2006b), which in turn differ in their morphology (Coelho *et al.* 1998; Rodrigues 1999; Sousa-Santos 2001), provides an additional source of genetic variation in the form of distinct “natural experiments” that may be explored in the future to test genotype/phenotype relationships.

Aside from its interest as a model system in quantitative genetics, this hybrid complex may be of great relevance to solve problems of taxonomy and biogeography of the Iberian *Squalius*. Indeed, as PPA females produce P-gametes (Crespo-López *et al.* 2006), crosses involving these females and males of other *Squalius* species may generate non-*alburnoides* individuals. Thus, the introgression of genes can be bidirectional: not only genes from other *Squalius* may be introgressed into the complex, but also the *S. alburnoides* complex may be responsible for the introgression of genes in other *Squalius* species, likely contributing to their geno- and phenotypic de-characterization. The introgression of *S. pyrenaicus* genes mediated by *S. alburnoides* in *S. carolitertii*, for instance, may raise questions about the taxonomical boundaries of the species.

As previously mentioned, the abundance of the different forms of the *S. alburnoides* complex differs significantly among the drainages. One hypothetical explanation for this phenomenon is the existence of a female preference to mate with non-hybrid males (Sousa-Santos *et al.* 2006a) combined with different ecological requirements of hybrid females and non-hybrid males. As demonstrated by Martins *et al.* (1998), at least in the Guadiana drainage, outside the reproductive period there are differences in habitat use: non-hybrid AA males prefer shallow and warmer waters and silt/sand substrates; diploid

hybrid females (PA) are most commonly found in deeper waters, steeper gradients and coarse substrata; and triploid hybrid females (PAA) prefer areas with higher current velocities and instream cover. Thus, since some degree of spatial segregation seems to exist, it is very unlikely that all possible combinations of spawning pairs are equally probable in this or in other river basins (although similar studies performed in other drainages are lacking). Instead, the males find in the spawning areas of the diploid *S. alburnoides* females might be distinct from the males that triploid females find in their spawning areas. Our suggestion is that distinct ecological requirements and mate preferences that seem to play a crucial role in the genetic dynamics of the complex may be at least in part related to the morphological differentiation between the forms of the *S. alburnoides* complex, found in the present study. Indeed, the combination of morphological, ecological and ethological traits may have contributed to reduce intraspecific competition between the ploidy forms and be one of the explanations for the high success of this fish complex.

This diversity of genome constitutions, body shapes and related ecological preferences may allow *S. alburnoides* hybrid and nonhybrid specimens to take advantage of a more diverse array of habitats than the other sympatric *Squalius* species, raising fascinating questions about the evolutionary potential of this hybridogenetic complex.

Acknowledgments We thank Sousa-Santos family and J. Robalo for help in sample collection. This study was funded by the FCT Pluriannual Program (U I&D 331/94 and U I&D 329/94) (FEDER participation). C. Sousa-Santos was supported by a PhD grant from FCT (SFRH/BD/8320/2002).

References

- Albertson RC & Kocher TD. 2001. Assessing morphological differences in an adaptive trait: a landmark-based morphometric approach. *Journal of Experimental Zoology* 289: 385–403.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 1997a. The *Rutilus alburnoides* complex (Cyprinidae): evidence for hybrid origin *Journal of Zoological Systematics & Evolutionary Research* 35: 1–10.
- Alves MJ, Coelho MM, Collares-Pereira MJ & Dowling TE. 1997b. Maternal ancestry of the *Rutilus alburnoides* complex (Teleostei, Cyprinidae) as determined by analysis of cytochrome *b* sequences. *Evolution* 51: 1584–1592.

- Alves MJ, Coelho MM & Collares-Pereira MJ. 2001. Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: a genetic review. *Genetica* 111: 375-385.
- Alves MJ, Collares-Pereira MJ, Dowling MJ & Coelho MM. 2002. The genetics of maintenance of an all-male lineage in the *Squalius alburnoides* complex. *Journal of Fish Biology* 60: 649-662.
- Alves MJ, Gromicho M, Collares-Pereira MJ, Crespo-López E & Coelho MM. 2004. Simultaneous production of triploid and haploid eggs by triploid *Squalius alburnoides* (Teleostei: Cyprinidae). *Journal of Experimental Zoology* 301A: 552-558.
- Carmona JA. 1997. Evolución y modelos de reproducción en un organismo unisexual de la Península Ibérica: el complejo *Tropidophoxinellus alburnoides*. PhD Thesis. Universidad Complutense, Madrid, 325 p.
- Carmona JA, Sanjur OI, Doadrio I, Machordom A & Vrijenhoek VC. 1997. Hybridogenetic reproduction and maternal ancestry of polyploid Iberian fish: the *Tropidophoxinellus alburnoides* complex. *Genetics* 146: 983-993.
- Coelho MM, Bogutskaya N, Rodrigues JA & Collares-Pereira MJ. 1998. *Leuciscus torgalensis* and *L. aradensis*, two new cyprinids for Portuguese fresh waters. *Journal of Fish Biology* 52: 937-950.
- Collares-Pereira MJ. 1983. Estudo sistemático e citogenético dos pequenos ciprinídeos ibéricos pertencentes aos géneros *Chondrostoma* Agassiz 1835, *Rutilus* Rafinesque, 1820 e *Anaecypris* Collares-Pereira, 1983. PhD Thesis. Universidade de Lisboa, Lisboa, 511p.
- Collares-Pereira MJ. 1984. The "*Rutilus alburnoides* (Steindachner, 1866) complex" (Pisces, Cyprinidae). I. Biometrical analysis of some Portuguese populations. *Arquivos Museu Bocage (Série A)* 11: 111-143.
- Coyne JA & Orr HA. 2004. Speciation. Sinauer, Massachusetts, 545p.
- Crespo-López ME, Duarte T, Dowling T & Coelho MM. 2006. Modes of reproduction of the hybridogenetic fish *Squalius alburnoides* in the Tagus and Guadiana rivers: an approach using microsatellites. *Zoology* 109: 277-286.
- Cunha C, Coelho MM, Carmona JA & Doadrio I. 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Molecular Ecology* 13: 2807-2817.
- Gante H, Collares-Pereira MJ & Coelho MM. 2004. Introgressive hybridisation between two Iberian *Chondrostoma* species (Teleostei, Cyprinidae) revisited: new evidence from morphology, mitochondrial DNA, allozymes and NOR-phenotypes. *Folia Zoologica* 53: 423-432.
- Gomes-Ferreira A, Ribeiro F, Moreira da Costa L, Cowx IG & Collares-Pereira MJ. 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. *Journal of Fish Biology* 66: 454-467.
- Gromicho M & Collares-Pereira MJ. 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.

- Gromicho M, Coelho MM, Alves MJ & Collares-Pereira MJ. 2006. Cytogenetic analysis of *Anaecypris hispanica* and its relationship with the paternal ancestor of the diploid-polyploid *Squalius alburnoides* complex. *Genome* 49: 1621-1627.
- Kittell MM, Harvey MN, Contreras-Balderas S & Ptacek MB. 2005. Wild-caught hybrids between sailfin and shortfin mollies (Poeciliidae, *Poecilia*): Morphological and Molecular Verification. *Hidrobiologica* 15: 131-137.
- Martins MJ, Collares-Pereira MJ, Cowx IG & Coelho MM. 1998. Diploids v. triploids of *Rutilus alburnoides*: spatial segregation and morphological differences. *Journal of Fish Biology* 52: 817-828.
- Mateos M & Vrijenhoek RC. 2002. Ancient versus reticulate origin of a hemiclinal lineage. *Evolution* 56: 985-992.
- Pala I & Coelho MM. 2005. Contrasting views over a hybrid complex: between speciation and evolutionary "dead-end". *Gene* 347: 283-294.
- Reist J. 1985. An empirical evaluation of several univariate methods that adjust for size variation in morphometric data. *Canadian Journal of Zoology* 63: 1429-1439.
- Ribeiro F, Cowx IG, Tiago P, Filipe AF, Moreira da Costa L & Collares-Pereira MJ. 2003. Growth and reproductive traits of diploid and triploid forms of the *Squalius alburnoides* cyprinid complex in a tributary of the Guadiana River, Portugal. *Archiv für Hydrobiologie* 156: 471-484.
- Robalo JI, Sousa-Santos C, Levy A & Almada VC. 2006. Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. *Molecular Phylogenetics and Evolution* 39: 276-281.
- Rodrigues JA. 1999. Aspectos da bio-ecologia das populações de *Leuciscus pyrenaicus* Günther, 1868 (Pisces, Cyprinidae) na Bacia Hidrográfica do rio Tagus. PhD Thesis. Universidade de Lisboa, Lisboa, 302p.
- Sambrook J, Fritsch EF & Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor, New York, 999 p.
- Scribner KT, Page KS & Bartron ML. 2001. Hybridization in freshwater fishes: a review of case studies and cytonuclear methods of biological inference. *Reviews in Fish Biology and Fisheries* 10: 293-323.
- Sousa-Santos C. 2001. O género *Leuciscus* Cuvier, 1816 (Pisces: Cyprinidae) na região Litoral Oeste e suas afinidades biogeográficas. Análise morfológica, genética e do comportamento reprodutor. Undergraduate thesis. Universidade de Lisboa, Lisboa, 56p.
- Sousa-Santos C, Collares-Pereira MJ & Almada VC. 2006a. Reproductive success of nuclear non-hybrid males of *Squalius alburnoides* hybridogenetic complex (Teleostei, Cyprinidae): an example of interplay between female choice and ecological pressures? *Acta Ethologica* 9: 31-36.
- Sousa-Santos C, Collares-Pereira MJ & Almada VC. 2006b. 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. *Journal of Fish Biology* 68 (Supplement B): 292-301.
- Sousa-Santos C, Collares-Pereira MJ & Almada VC. 2006c. 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. *Journal of Natural History* 40: 1443-1448.

Sousa-Santos C, Collares-Pereira MJ & Almada VC. in press. Reading the history of a hybrid fish complex from its molecular record. *Molecular Phylogenetics and Evolution*

Sousa-Santos C, Collares-Pereira MJ & Almada VC. 2007. Fertile triploid males – an uncommon case among hybrid vertebrates. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* 307A: 220-225.

Sousa-Santos C, Robalo JI, Collares-Pereira MJ & Almada VC. 2005. Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genome composition of hybrid organisms. *DNA Sequence* 16: 462-467.

Tunner HG. 2000. Evidence for genome imprinting in unisexual triploid hybrid frogs. *Amphibia-Reptilia* 21: 135-141.

PART III

REPRODUCTION

Chapter 4

May a hybridogenetic complex regenerate the nuclear genome of both sexes of a missing ancestor? - First evidence on the occurrence of a nuclear nonhybrid *Squalius alburnoides* (Cyprinidae) female based on DNA sequencing

Sousa-Santos C, Collares-Pereira MJ & Almada V (2006)

Journal of Natural History 40(23-24): 1443-1448

May a hybridogenetic complex regenerate the nuclear genome of both sexes of a missing ancestor? - First evidence on the occurrence of a nuclear nonhybrid *Squalius alburnoides* (Cyprinidae) female based on DNA sequencing

Sousa-Santos C, Collares-Pereira MJ & Almada V

Abstract Based on molecular evidence and on direct observation of gonads and morphology, we describe the occurrence of a female of the hybridogenetic minnow *Squalius alburnoides* bearing the nuclear genome of the paternal ancestor of the complex and the mtDNA of *S. pyrenaicus* (the maternal species). The paternal ancestor is believed to be extinct and the available molecular evidence indicates that it was a species distant from the maternal ancestor and closer to a very different genus (*Anaocypris*). Its nuclear genes were perpetuated through hybrids and through diploid males originated from the hybrids and containing two copies of the paternal genome. The discovery of a diploid female with the pure nuclear genome of the paternal ancestor, even if it represents a very rare occurrence, illustrates a very interesting biological phenomenon: the possibility of re-emergence of an extinct species from its descendent hybrids, although carrying the mtDNA of another species.

Keywords hybrid, non-sexual, extinct ancestor, beta-actin, Cyprinidae, *Squalius alburnoides*

Introduction

The Iberian *Squalius alburnoides* (Steindachner 1866) complex is an exceptionally interesting hybrid system because apart from diploid, triploid and tetraploid fishes that reproduce mainly by meiotic hybridogenesis, it also includes diploid males whose nuclear genome is nonhybrid and undergoes normal meiosis (reviewed in Alves *et al.* (2001).

Robalo *et al.* (2006) showed that the *S. alburnoides* complex resulted from an intergeneric hybridization process, since the nuclear DNA of its "missing paternal ancestor" was phylogenetically close to the species *Anaocypris hispanica* Steindachner

1866 – as previously suggested in Alves *et al.* (2001), while the maternally inherited mitochondrial genome is typical of the genus *Squalius* (Alves *et al.* 2001).

Depending on the species of *Squalius* available in the different drainages where *S. alburnoides* lives, it can form hybrids with *S. pyrenaicus*, *S. carolitertii* or *S. aradensis* (Alves *et al.* 1997, 2001; Carmona *et al.* 1997; Cunha *et al.* 2004; Pala & Coelho 2005; Sousa-Santos *et al.* 2006a), a finding confirmed by protein electrophoresis and DNA analysis.

The diploid males carrying two copies of the nuclear genome of the paternal ancestor are morphologically very distinct from the hybrids. They have a narrow body with pointed head, straight dorsal profile and convex ventral profile, terminal mouth with a greater lower jaw so that the large buccal opening is turned slightly upward, deeply forked caudal lobes, (39)40 to 45 scales in the lateral line, 17 to 25(26) gillrakers, and a typical non-symmetrical pharyngeal teeth formula of (4)5/4(5) (Collares-Pereira 1984).

All the nuclear nonhybrid specimens from natural populations and experimental crosses so far analysed using genetic and cytogenetic markers were males (see Alves *et al.* 1998, 2001, 2002; Gromicho & Collares-Pereira 2004). One nuclear nonhybrid female was reported by Carmona (1997), in the river Estena (Guadiana River Basin). In this paper we report the first occurrence of such a nonhybrid female explicitly confirmed with a nuclear marker. This diploid nonhybrid female, also captured in the Guadiana River basin, was morphologically similar to the nonhybrid diploid males. It carried the mtDNA of *S. pyrenaicus* and the nuclear genome of the paternal ancestor.

Methods

The *S. alburnoides* nonhybrid female was electrofished in 2000 in river Caia (Guadiana River basin) and maintained in a tank with other *S. alburnoides* individuals until 2004. The sex of the fish was confirmed by direct observation of the spawning behaviour, extrusion of gametes by applying a mild pressure on the abdomen and *post mortem* examination. To evaluate its readiness to spawn and attractiveness to males, the female was maintained in an outdoor aquarium with four diploid nonhybrid males (confirmed by flow cytometry and *beta-actin* gene sequencing), under natural conditions of light and

temperature.

Assignment to a morphological type was made *post mortem* by gill-raker and lateral scale counts, inspection of pharyngeal teeth, general body shape and position of the mouth.

Small fin clips were taken from 12 *S. alburnoides* (including the nonhybrid female) and from six *S. pyrenaicus*. Total genomic DNA was extracted by an SDS/proteinase-k based protocol, precipitated with isopropanol and washed with ethanol before re-suspension in water (adapted from Sambrook *et al.* 1989). DNA samples of the nonhybrid female and of six nonhybrid males from the Guadiana (N=1), Tejo (N=1) and Quarteira (N=4) river basins were genetically analysed for a segment of 927bp of the nuclear *beta-actin* gene. The amplification process was conducted using the primers BACTFOR and BACTREV (Sousa-Santos *et al.* 2005). In addition, a total of 1123bp of the cytochrome *b* (*cytb*) gene was also amplified using the primers LCB1 (Brito *et al.* 1997) and HA (Schmidt & Gold 1993) from samples of the nonhybrid female, of six *S. alburnoides* with distinct nuclear genomic constitutions previously assessed by beta-actin sequencing (one nonhybrid AA, two diploid PA and three triploid PAA hybrids) and of six *S. pyrenaicus* (PP), all from the Guadiana drainage. PCR conditions for both genes followed the procedures described in Sousa-Santos *et al.* (2005). Each sample was sequenced in both directions with the same primers used for PCR. Sequences were aligned with BioEdit® v.5.0.6 and deposited in GenBank under the accession numbers: AY943863-65, AY943867, DQ128102, DQ010337, AY943891, DQ335472-481 and DQ350252 (for *beta-actin*); and DQ263227-39 (for *cytb*). Note that, in the case of the *beta-actin* gene sequences, there are more accession numbers than individuals since some fishes were heterozygous for the analysed fragment and each strand had its own accession number. For a description of the procedure that allows the identification of the two different sequences involved in hybrids see Sousa-Santos *et al.* (2005). A nonhybrid nuclear constitution was identified when the chromatograms showed only clear single peaks, a situation that contrasted with the presence of series of double peaks in the chromatograms of *S. alburnoides* that presented a hybrid nuclear genome, as described in Sousa-Santos *et al.* (2005).

PAUP® 4.0 software (Swofford 1998) was used to calculate the mean percentage of divergence between the *cytb* gene sequences - defined as the average number of pairwise differences among sequences, expressed as a percentage of the total length of

the gene fragment analysed. A parsimony network was constructed with TCS[®] v1.21 (Clement *et al.* 2000) using a 95% confidence limit.

Ploidy assessments were made by flow cytometry using fresh fin clips, following an adaptation of Lamatsch *et al.* (2000) method developed by MJ Collares-Pereira *et al.* (*unpublished*).

Results

Morphologically, the nonhybrid diploid female presented the same characteristics as all reconstituted nonhybrid males so far described: 45 scales in the lateral line, 18 gillrakers, and a 5/4 pharyngeal teeth formula. Up to now, there are no available estimates of the frequency of these nonhybrid diploid females in natural populations. In our survey, this was the only nuclear nonhybrid female collected, while more than 50 such males were found.

Results of the flow cytometry analysis confirmed the diploidy of this female. The analysis of the fragment of the *beta-actin* gene revealed an unambiguous nonhybrid constitution of the nuclear genome, as shown by the existence of clear single peaks in the chromatogram retrieved from the sequencing process. This female and all the six other nuclear nonhybrid males sequenced shared the same *beta-actin* haplotype (GenBank accession numbers: AY943863-65, AY943867, DQ128102, DQ010337 and DQ350252).

The analysis of the *cytb* gene showed that the nonhybrid female presented a *S. pyrenaicus*-like mitochondrial genome that was similar to the ones found in the *S. pyrenaicus* and *S. alburnoides* individuals sequenced, differing only by few mutations (Figure 1). The mean percentage of divergence values between the *cytb* haplotypes of the nonhybrid female, of the *S. pyrenaicus* and of the remaining *S. alburnoides* individuals is presented in Table I. This female underwent normal maturation and ovulation, which was demonstrated by the very easy release of mature eggs when a mild pressure was applied to the abdomen of the fish (about two weeks before the spawning observations) and by inspection of the gonads and oviducts upon dissection. In addition, it was fully attractive to males. Indeed, the courtship behaviours displayed by the males towards the nonhybrid female were similar to the ones performed towards the much more common hybrid

females (Sousa-Santos *et al.* 2006b) and involved female attraction to the spawning site with quivering, fin shuffling and manoeuvrings. The female interacted with the courting males and spawned with them several times (the females of *S. alburnoides* do not spawn all the eggs at one time but instead make repeated spawnings that may extend for an entire day - Sousa-Santos *et al.* 2006b).

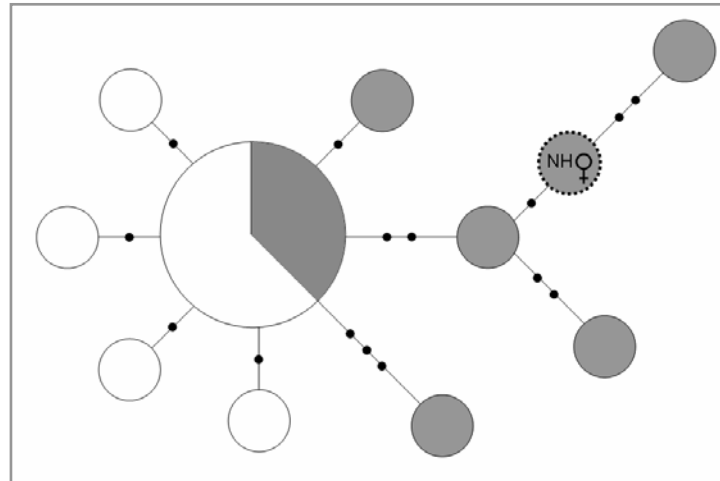


Figure 1 — Minimum spanning network among *cytb* haplotypes. The majority of the haplotypes (represented by circles) are exclusive of *S. alburnoides* (in grey) and of *S. pyrenaicus* (in white), except for the central one which is a haplotype shared between one *S. alburnoides* and two *S. pyrenaicus* individuals. NH indicates the haplotype of the nonhybrid female. The number of mutations between haplotypes is represented by small black dots.

Table 1 — Mean percentage of divergence between *cytb* gene haplotypes from the nonhybrid female, *S. alburnoides* and *S. pyrenaicus* from Guadiana. The intraspecific mean percentage of divergence values are presented in the diagonal.

		Nonhybrid female	<i>S. pyrenaicus</i>	<i>S. alburnoides</i>
Nonhybrid female	N=1	-		
<i>S. pyrenaicus</i>	N=6	0.33 ± 0.05	0.13 ± 0.05	
<i>S. alburnoides</i>	N=6	0.25 ± 0.10	0.28 ± 0.16	0.34 ± 0.13

Eggs had a normal appearance and the development reached the stage where embryos with already fully pigmented eyes were visible moving actively in the egg capsule. Unfortunately, all the progeny was lost due to a fungal infection.

Discussion

Although the possibility of occurrence of females with a nonhybrid nuclear genome within the *S. alburnoides* complex has been confirmed, the mechanism responsible for the production of these apparently rare reconstituted nonhybrid females is as yet unclear. The fertilization of haploid A oocytes (produced by PAA hybridogenetic females) by A sperm of nuclear nonhybrid AA males has always generated male offspring in all experimental crosses analysed to date (Alves *et al.* 1998; Gromicho & Collares-Pereira 2004). Studies based on such crosses have shown that around 3% of hybrid females may reproduce by gynogenesis (Alves *et al.* 2001), and that the most common female's biotype in southern drainages - the triploids PAA - generally produces haploid gametes (A) by a meiotic hybridogenetic process. However, there are also evidences that PAA females may, in addition to these reduced eggs, produce gametes with partially (AA) or totally (PAA) unreduced genomes (Alves *et al.* 2004).

In addition to the fertilization of a eggs by A sperm, two other possibilities ending up in the production of these rare nuclear nonhybrid females may be hypothetically considered: a gynogenetic development process of diploid AA oocytes produced by PAA females and eventually, the fertilization by A sperm of reduced A oocytes after P-genome extrusion in PA females, although they generally transmit their genome clonally to offspring (see Alves *et al.* 2001; Gromicho & Collares-Pereira 2004).

From a general evolutionary perspective, and regardless of the mechanism involved, the most relevant aspect of the finding here reported is that it highlights the possibility of the reconstitution of the nuclear genome of an extinct species, whose genes had been perpetuated in hybrids for many generations. The only mark of the past hybridization history that remains is apparently the mtDNA inherited from the maternal species that originated the hybrids. However, before we can achieve a full assessment of the significance of this finding, it will be necessary to investigate if these nuclear nonhybrid diploid females are genetically distinct from the corresponding males or are mere developmental deviations of what are in fact genetically males. In addition, it will be of paramount importance for the study of this fish complex to investigate what kind of gametes these reconstituted females originate. If they prove to display normal meiosis, if their gametes combined with the corresponding males produce fish of both sexes, and if they reach sufficient frequency in nature to have a reasonable chance of mating with

reconstituted males of the same genetic makeup, they could originate a viable lineage fully independent from the hybrids. Indeed, the observation that the AA progeny of nuclear nonhybrid diploid males is composed only by males, as mentioned above, may be affected by the fact that they mate with fish of hybrid origin. It may well be that if they mate with AA females they may produce progeny of both sexes, as their ancestors had to do. The rarity of these females may simply be due to the fact that the hybrids are so more abundant that the AA males will mate with hybrids in most occasions.

There is still another feature that makes *S. alburnoides* a very remarkable case among the hybridogenetic vertebrates so far investigated: in several of its forms there are males and/or females which include meiosis and recombination in their gametogenic processes (Alves *et al.* 2001), giving the complex the capability of maintaining a high level of genetic variability which is unavailable to most other vertebrate hybrids so far described.

Acknowledgments The authors wish to thank the field team of CBA (University of Lisbon) and Sousa-Santos family for fish captures, and J. Robalo and T. Bento for help in maintaining the fish tank. We appreciate the skilful technical assistance provided by A. Gomes-Ferreira, M. Gromicho e L. Moreira da Costa during ploidy determinations by flow cytometry. The study was funded by the FCT Pluriannual Program (UI&D 331/94 and UI&D 329/94) (FEDER participation). C. Sousa-Santos was supported by a PhD grant from FCT (SFRH/BD/8320/2002).

References

- Alves MJ, Coelho MM & Collares-Pereira MJ. 1997. The *Rutilus alburnoides* complex (Cyprinidae): evidence for hybrid origin. *Journal of Zoological Systematics and Evolutionary Research* 35: 1–10.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 1998. Diversity in the reproductive modes of females of the *Rutilus alburnoides* complex (Teleostei, Cyprinidae): a way to avoid the genetic constraints of uniparentalism. *Molecular Biology and Evolution* 15: 1233–1242.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 2001. Evolution in action through hybridization and polyploidy in an Iberian freshwater fish: a genetic review. *Genetica* 111: 375–385.
- Alves MJ, Collares-Pereira MJ, Dowling TE & Coelho MM. 2002. The genetics of maintenance of an all-male lineage in the *Squalius alburnoides* complex. *Journal of Fish Biology* 60: 649–662.
- Alves MJ, Gromicho M, Collares-Pereira MJ, Crespo-López E & Coelho MM. 2004. Simultaneous production of triploid and haploid eggs by triploid *Squalius alburnoides* (Teleostei: Cyprinidae). *Journal of Experimental Zoology* 301A: 552–558.
- Brito RM, Briolay J, Galtier N, Bouvet Y & Coelho MM. 1997. Phylogenetic relationships within genus *Leuciscus*

(Pisces, Cyprinidae) in Portuguese freshwaters, based on mitochondrial cytochrome *b* sequences. *Molecular Phylogenetics and Evolution* 8: 435-442.

Carmona JA. 1997. Evolución y modelos de reproducción en un organismo unisexual de la Península Ibérica: el complejo *Tropidophoxinellus alburnoides*. PhD Thesis. Universidad Complutense, Madrid, 325p.

Carmona JA, Sanjur OI, Doadrio I, Machordom A & Vrijenhoek VC. 1997. Hybridogenetic reproduction and maternal ancestry of polyploid Iberian fish: the *Tropidophoxinellus alburnoides* complex. *Genetics* 146: 983-993.

Collares-Pereira MJ. 1984. The "*Rutilus alburnoides* (Steindachner, 1866) complex" (Pisces, Cyprinidae). I. Biometrical analysis of some Portuguese populations. *Arquivos do Museu Bocage Série A II*: 111-143.

Clement M, Posada D & Crandall K. 2000. TCS®: a computer program to estimate gene genealogies. *Molecular Ecology* 9: 1657-1660.

Cunha C, Coelho MM, Carmona JA & Doadrio I. 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Molecular Ecology* 13: 2807-2817.

Gromicho M & Collares-Pereira MJ. 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.

Lamatsch DK, Steinlein C, Schmid M & Scharf M. 2000. Noninvasive determination of genome size and ploidy level in fishes by flow cytometry: detection of triploid *Poecilia formosa*. *Cytometry* 39: 91-95.

Pala I & Coelho MM. 2005. Contrasting views over a hybrid complex: between speciation and evolutionary "dead-end". *Gene* 347: 283-294.

Robalo JI, Sousa-Santos C, Levy A & Almada VC. 2006. Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. *Molecular Phylogenetics and Evolution* 39: 276-281.

Sambrook J, Fritsch EF & Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor, New York, 999 p.

Schmidt TR & Gold JR. 1993. Complete sequence of the mitochondrial cytochrome *b* gene in the Cherryfin Shiner, *Litrus roseipinnis* (Teleostei: Cyprinidae). *Copeia* 3: 880-883.

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

Sousa-Santos C, Collares-Pereira MJ & Almada VC. 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. *Journal of Fish Biology* 68 (Supplement B): 292-301.

Sousa-Santos C, Collares-Pereira MJ & Almada VC. 2006b. Reproductive success of nuclear nonhybrid males of *Squalius alburnoides* hybridogenetic complex (Teleostei, Cyprinidae): an example of interplay between female choice and ecological pressures? *Acta Ethologica* 9: 31-36.

Swofford DL. 1998. PAUP® - phylogenetic analysis using parsimony (and other methods) version 4.0. Sinauer Associates, Sunderland.

Chapter 5

Fertile triploid males – an uncommon case among hybrid vertebrates

Sousa-Santos C, Collares-Pereira MJ & Almada V (2007)

Journal of Experimental Zoology Part A 307A(4): 220-225

Fertile triploid males – an uncommon case among hybrid vertebrates

Sousa-Santos C, Collares-Pereira MJ & Almada V

Abstract The endemic Iberian minnow *Squalius alburnoides* is a complex of fishes of hybrid origin including both males and females with distinct ploidy levels and varying proportions of the parental genomes. In this paper we demonstrated that in contrast to many vertebrate hybrid lineages the sperm of triploid hybrid males of *S. alburnoides* is viable and fully functional. Flow cytometry and analysis of sequences of a fragment of the beta-actin nuclear gene applied to progenitors and offspring evidenced that these males produced their sperm clonally, as already described for diploid hybrids. The presence of different types of fertile males (non-hybrid diploids with normal meiosis, and both diploid and triploid hybrids) coupled with hybridogenetic meiosis in females endows this vertebrate complex with a high level of independence from other species and contributes to maintain its genetic variability.

Keywords double peaks, interspecific hybrids, haplotype reconstruction, heterozygotes, beta-actin gene, Cyprinid fish

Introduction

Often when two distinct species hybridize there is a disruption of normal meiosis that leads to partial or total infertility of the descendants (Vrijenhoek 1989, 1994). Some hybrids, however, overcome this problem by reproducing by diverse gametogenic mechanisms. These alternative modes of reproduction among hybrid lineages of vertebrates are known to have evolved independently in many groups and may involve asexual or modified forms of sexual reproduction (reviewed in Dawley & Bogart 1989).

Within the fish family Cyprinidae, the small Iberian minnow *Squalius alburnoides* is an example of an ecologically extremely well succeeded hybrid lineage that was originated by intergeneric crosses between *S. pyrenaicus* females - P genome - and males of a probably extinct paternal species - A genome (Alves *et al.* 2001),

phylogenetically close to the extant *Anaocypris hispanica* (Robalo *et al.* 2006). This interesting female sex biased complex comprises diploid, triploid and tetraploid hybrid forms and a diploid nonhybrid form (AA) constituted almost exclusively by males (reviewed in Alves *et al.* 2001). This diploid AA male lineage contains two genomic complements inherited from the paternal ancestor but a *pyrenaicus*-like mtDNA (Alves *et al.* 2002), a clear indication that they resulted from the hybrids and are not “survivors” of the paternal progenitor of the complex. Instead of being a typical asexual all-female lineage that reproduces clonally, this complex includes also fertile males and has varied reproductive modes and gametogenic processes.

Concerning the spermatogenetic processes, it is known that normal meiosis occurs in diploid nonhybrids (AA) and in balanced tetraploids (PPAA) and that on the contrary, diploid hybrids (PA) produce fertile unreduced sperm, transmitting the intact genome to the offspring (Alves *et al.* 1998, 1999). A study of two triploid males, which seem to be very rare in natural populations, by flow cytometry of the DNA content of erythrocytes and spermatozoa, showed that they also produced unreduced sperm (Alves *et al.* 1999) but their genomic constitution (PAA or PPA) remained unknown and their sperm functionality was not tested.

Therefore, the aims of the present study were: 1) to clarify the reproductive mode and fertility of triploid males by accurately determining the genomic constitution of their sperm and of progeny using a suitable nuclear marker (beta-actin gene) combined with ploidy confirmation; and 2) to evaluate the viability of the progeny produced by triploid males. To validate the applied methodology crosses with diploid PA and AA males were also performed.

Materials and Methods

Five *S. alburnoides* males (one triploid, two hybrid diploids and two nonhybrid diploids) were selected after the determination of their genomic constitutions by beta-actin gene sequencing and the confirmation of their ploidy by flow cytometry. These males were then used in crossing experiments with *S. alburnoides* and *S. torgalensis* females and their live progeny (N=31) was subsequently analysed, also by beta-actin gene sequencing and flow

cytometry. Although DNA nuclear measurements were impracticable in very small dead juveniles, the DNA content of 33 additional preserved juveniles was analysed since ploidy assessment was also possible through the analysis of the beta-actin gene sequences (see below). The ploidy of the *S. torgalensis* females was not determined since it is expected to be a diploid species like *S. aradensis*, *S. carolitertii* and *S. pyrenaicus* (Collares-Pereira & Moreira da Costa 1999). This was confirmed by the fact that of more than 25 *S. torgalensis* sequenced for the beta-actin gene none gave indications of being non-diploid (*unpublished data*). Additional and more detailed methodological information is presented below.

Assessment of the genomic constitutions

When a nuclear gene of a hybrid between closely related species is sequenced, single nucleotide polymorphisms appear in chromatograms as single double peaks. However, in the case of a hybrid between species that differ by one or more indels in a particular gene fragment, the sequencing process generates artefacts in the resultant chromatograms that appear as series of false double peaks and it is possible to reconstruct the complements involved since the bases are out of phase for as many positions as the size of the indel – for more details see Sousa-Santos *et al.* (2005). This process of reconstruction of the genomes present in hybrids is favoured by a previous knowledge about some fixed mutations that are specific of the parental genomes as these act as reference points that make possible to ascribe unambiguously each peak to the correct parental genome (Sousa-Santos *et al.* 2005). Thus, to evaluate the result of the spermatogenetic process of a *S. alburnoides* male through the observation of the genomic constitution of its progeny it was considered ideal to cross the male with a female of a species having characteristic marker mutations absent in the male but sufficiently related to prevent reproductive isolation. This important role could be performed by *S. torgalensis* females from Mira River, where *S. alburnoides* complex is absent. Indeed, concerning the beta-actin gene of the Portuguese *Squalius* species, three groups of haplotypes were found so far, clearly distinct from each other by characteristic patterns of fixed mutations and indels (Sousa-Santos *et al.* 2005, 2006): P and C haplotypes in *S. pyrenaicus* and *S. carolitertii* (respectively); Q and T haplotypes in *S. aradensis* and *S. torgalensis* (respectively); and A haplotypes in *S. alburnoides*. At the

beta-actin gene level, the P and C haplotypes are not distinguishable from each other and the same occurs between the Q and T haplotypes (Sousa-Santos *et al.* 2006). *S. alburnoides* hybrids exhibit a mixture of A and P, C or Q haplotypes depending on whether they are, respectively, descendants of *S. pyrenaicus*, *S. carolitertii* or of *S. aradensis*, the three *Squalius* species with which *S. alburnoides* occur in sympatry.

Since all *S. alburnoides* males selected to be used in the crossing experiments were from Tagus and Douro river basins (where *S. pyrenaicus* and *S. carolitertii*, respectively, occur in sympatry with *S. alburnoides*) they presented similar beta-actin haplotypes (P and C). Crossings with *S. torgalensis* bearing beta-actin T-haplotypes that are clearly distinct from P and C haplotypes would, therefore, allow an easy discrimination of the genomes inherited by the progeny, a situation that would not occur if the females used in the experiments were *S. pyrenaicus* or *S. carolitertii* since they would bear P and C haplotypes like the *S. alburnoides* male progenitors. The count method described by Sousa-Santos *et al.* (2005) was applied to a segment of 176bp and the resultant P-count values were used to assess the ploidy and genome constitution of the *S. alburnoides* hybrids.

Experimental crosses

Each *S. alburnoides* male (one CAA triploid from the Douro River and two PA diploids from the Tagus River basin) was isolated with one *S. torgalensis* female (TT) from Mira River (crosses 1-3) in a tank (30 litres) with vegetation and coarse substratum, under natural conditions of light and temperature, for two months (May and June 2005). The spawning was not artificially induced and the experimental pairs spawned naturally, after which they were removed from the tank to minimize the loss of eggs and juveniles by predation. The majority of the eggs was lost due to fungal infections but a total of 31 juveniles was reared to the age of eight months to attain a sufficiently large body length (2-4 cm) that allowed their survival after blood and fin clip sampling. A total of 33 juveniles that died during this period were immediately preserved in 96% ethanol.

In addition to these controlled experiments, two juveniles (nine months old) resultant from group spawnings between seven *S. alburnoides* AA males from Tagus River and eight *S. torgalensis* females from Mira River and eight juveniles (21 months old) resultant from crosses between two AA males and two PA females of *S. alburnoides* from

the Tagus river were also reared for comparative analyses (crosses 4 and 5). Due to their low number, only the beta-actin gene sequencing was performed since blood sampling of small fish for flow cytometry procedures involves a high risk of mortality.

In total, considering the controlled experiments with only one pair of spawners and the experiments involving groups with more than one pair of spawners, 74 juveniles were sequenced which resulted from five crosses (two PAXTT, PAAxTT, AAxTT and AAxPA).

All the experiments were made in conformity to the ethical guidelines for animal care approved by the authors' institutions.

Laboratory methods

Ploidy determination by flow cytometry was made using small blood samples drawn from the caudal vein of live juveniles or using fresh fin clips in the case of the adults used in the crossing experiments (since these measurements were made prior to the spawnings and the removal of fin clips involved less risks than blood sampling to the survival of the adults). The procedures used for blood samples were the ones described in Alves *et al.* (1999) and for the analysis of fin samples an adaptation of the method proposed by Lamatsch *et al.* (2000a) developed by MJ Collares-Pereira *et al.* (*unpublished*) was used. The DNA content estimation for individual fishes followed the procedure described in Alves *et al.* (1999). Total genomic DNA was extracted from fin clips preserved in ethanol by an SDS/proteinase-k based protocol, precipitated with isopropanol and washed with ethanol before re-suspension in water (adapted from Sambrook *et al.* 1989). A total of 927bp of the *beta-actin* gene was amplified using the primer BACTFOR and conditions described in Robalo *et al.* (2006). Sequences were aligned with BioEdit® v.5.0.6.

Results

The analysis of the beta-actin gene sequences of the *S. alburnoides* males used in the crossing experiments that involved only one pair of spawners showed that they were all hybrids with both P and A complements. P-count values showed that males used in crosses 1 and 2 were diploids (PA genome) and that male of cross 3 was a triploid with a

CAA genomic constitution. The flow cytometry analysis confirmed the ploidies of these three males (Table 1).

Table 1 — Genomic constitution and ploidy of the males used in the controlled crossing experiments and their respective progenies inferred by flow cytometry (DNA content of erythrocytes and/or fin cells) and beta-actin nuclear gene sequencing (P-counts) analyses. All crosses were mothered by *S. torgalensis* females (TT genome). P, C and A beta-actin haplotypes were inherited from the male progenitors. Flow cytometry reference values: 2.44 ± 0.08 pg/cell for diploids, 3.64 ± 0.16 pg/cell for triploids and 4.83 ± 0.08 for tetraploids (Alves *et al.* 1999). P-count reference values: $55.77\% \pm 6.88$ for diploids and balanced tetraploids and 95.69% for PPA triploids (Sousa-Santos *et al.* 2005).

		Cross 1	Cross 2	Cross 3
<i>S. alburnoides</i> male	P-count	53.39%	61.02%	14.41%
	Erythrocytes (pg / cell)	2.68	2.73	3.85
	Inferred ploidy/genome	2n / PA	2n / PA	3n / CAA
Juveniles	N (live/preserved)	19 (2/17)	24 (16/8)	21 (13/8)
	P-count	$89.70\% \pm 1.90$ (N=19)	$94.55\% \pm 2.36$ (N=24)	$63.83\% \pm 7.78$ (N=21)
	Erythrocytes (pg/cell)	3.80 ± 0.25	3.68 ± 0.12	5.23 ± 0.22
	Mean $\pm$ SD	(N=2)	(N=16)	(N=13)
	Inferred genome	TPA	TPA	TCAA

As shown in Table 1, all juveniles fathered by diploid hybrid males (crosses 1 and 2) were triploids with a TPA genomic constitution (N=43). The descendants of the triploid male (cross 3) were all tetraploids with a TCAA genome (N=21). This means that the sperm produced by the diploid and triploid males was also diploid and triploid, respectively. For the 31 offspring for which we had data from beta-actin sequencing and flow cytometry the results showed, without exception, full concordance between the two approaches. Live progeny that have resulted from these three crosses reached one year of age and are still currently exhibiting an apparently normal development. These juveniles are being kept alive in order to try to evaluate their fertility and gametogenic mechanisms.

Concerning the experimental crosses that involved more than one pair of spawners (crosses 4 and 5), both groups involved *S. alburnoides* males with a AA genome (mean 2.65 ± 0.10 pg/cell, N=7; and mean 2.67 ± 0.21 pg/cell, N=2, respectively). In cross 4 the nonhybrid males fertilized the eggs of *S. torgalensis* females and the resulting progeny was constituted by diploid TA hybrids (P-count= $64.41\% \pm 3.60$, N=2). In the cross 5, the nonhybrid males mated with *S. alburnoides* females with a PA genome (mean

2.54±0.01pg/cell and P-count=58.47%±1.20, N=2) and the resulting progeny was triploid and exhibited a PAA genome (P-count=5.93%±1.81, N=8). This means that in both crosses, the AA males produced haploid sperm which fertilized, respectively, haploid eggs from the *S. torgalensis* females and diploid eggs from *S. alburnoides* PA females (as expected since meiosis does not occur in this type of females, as described by Alves *et al.* 1998).

Discussion

The occurrence of triploid vertebrates is considered to be a rare phenomenon and is generally coupled with infertility or, more rarely, with asexuality in allotriploids, since normal meiosis is likely to be disrupted by cytological mechanisms that preclude synapsis between heterospecific chromosomes (Dawley 1989; Stöck *et al.* 2002). Thus, the establishment and survival of these organisms depends on clonal reproductive modes such as parthenogenesis or gynogenesis; on the incorporation of unusual meiotic mechanisms (like hybridogenesis) that allow recombination to occur in triploid females, often with the exclusion of the genetic complement inherited from the male parent; and/or on the ploidy elevation by the incorporation of additional genomic material (processes detailed in Dawley 1989; Avise *et al.* 1992; Vrijenhoek 1994; Beukeboom & Vrijenhoek 1998; Alves *et al.* 2001; Scali *et al.* 2003; Bi & Bogart 2006).

Triploid males of several species often develop testes and successfully complete spermatogenesis but are in general sterile due to the production of aneuploid and/or abnormally shaped spermatozoa, although some may produce functional sperm and generate embryos with a very low survival rate. The sterility of triploid male fishes has been documented for both artificially induced triploids [*e.g.* *Rhodeus ocellatus ocellatus* (Kawamura *et al.* 1999), *Carassius auratus* (Gui *et al.* 1992), *Verasper moseri* (Mori *et al.* 2006), *Salvelinus fontinalis* (Allen & Stanley 1978) and *Oncorhynchus mykiss* (Thorgaard & Gall 1979)], and for naturally occurring individuals, whose triploidy resulted from the incorporation of additional genomic material by an unreduced egg [*e.g.* *Misgurnus anguillicaudatus* (Matsubara *et al.* 1995; Zhang & Arai 1999; Oshima *et al.* 2005), *Poecilia formosa* (Lamatsch *et al.* 2000b) and *Tinca tinca* (Flajshans *et al.* 1993)].

In contrast, the results presented in this study showed that a triploid CAA *S. alburnoides* male produced unreduced sperm (like the diploid PA males) which, after fertilization of haploid eggs from *S. torgalensis*, gave rise to viable tetraploid offspring. The fertilization by the unreduced sperm of these triploid males leads to ploidy elevation of the offspring, as described by Bogart *et al.* (1989) for hybrid salamanders of the genus *Ambystoma*. Interestingly, however, this process does not seem to increase indefinitely the maximum number of chromosomes of the complex since females seem to act as regulators of the ploidy level. Indeed, triploid *S. alburnoides* females, the most abundant form in natural populations, exhibit a mechanism of gamete production (meiotic hybridogenesis) that constitutes a limitation to an uncontrolled growth of the ploidy level: the ability to exclude the unmatched set of chromosomes is likely to enable, in general, a subsequent normal meiosis and the production of haploid eggs. The same type of reduction to a haploid stage was observed in fertile triploid males of the Batura toads all-triploid species *Bufo pseudoraddei baturae* (Stöck *et al.* 2002), in the water frog *Rana esculenta* triploid males (Graf & Pelaz 1989) and in *Rana nigromaculata-lessona* triploid males (Ohtani 1992).

Triploid males of *S. alburnoides*, however, seem to produce only clonal sperm as do *S. alburnoides* hybrid diploids (Alves *et al.* 1999). Normal meiosis occurs only in *S. alburnoides* males that are nonhybrids and in tetraploids with balanced doses of both genomes (Alves *et al.* 2001). It must be bared in mind, however, that the number of males used in this study was very small and although likely representative of the more typical situation does not allow to entirely exclude other gametogenic processes. The example of PAA females illustrates well the need for further investigation of the gamete formation in *S. alburnoides*: although they frequently produce haploid A eggs after exclusion of the P-genome followed by normal meiosis, the same individual may also produce diploid AA and triploid PAA gametes (Alves *et al.* 2004), stressing the instability of the process. The same type of instability cannot be ruled out in hybrid males with the available data. In addition, the gametogenic behaviour of other types of males like PPA and PAAA is still awaiting investigation.

The fertility of the *S. alburnoides* males coupled with the ability of triploid females to undergo hybridogenetic meiosis is certainly an important factor contributing to the ecological success of the complex since it expands the number of possible routes to recombination and genetic variation. This scenario corroborates the point of view of

several authors that highlighted the evolutionary potential of hybridization (*e.g.* Vrijenhoek, 2006). Indeed, instead of being necessarily “evolutionary dead-ends” (Dawley 1989) due to difficulties to escape the accumulation of deleterious mutations and to cope with the changing environment, hybrid complexes may experience genetic diversification that enhances their hypotheses of success (*e.g.* Dawley 1989; Vrijenhoek 1994; Dowling & Secor 1997; Alves *et al.* 2001; Stöck *et al.* 2002; Scali *et al.* 2003; Janko *et al.* 2005; Bi & Bogart 2006; Gromicho *et al.* 2006).

In contrast with the majority of the vertebrate hybrid complexes already described, the *S. alburnoides* complex is neither a typical asexual lineage (as described by Dawley 1989) nor is it dependent on a bisexually reproducing host. Instead, this complex of minnows is constituted by both fertile males and females involved in varied reproductive modes that enhance its genetic variation, being an amazing example of the creative role that interspecific hybridization may play in evolution.

Acknowledgements The authors would like to thank G. Lemos, Sousa-Santos family, J. Robalo and T. Bento for help in the maintenance of the fish tanks and rearing juveniles. We are also grateful to C. Carvalho, M. Gromicho and L. Moreira da Costa for technical assistance with flow cytometry. DGF provided permissions for field work. Study funded by the FCT Pluriannual Program (UI&D 331/94 and UI&D 329/94) (FEDER participation). C. Sousa-Santos was supported by a PhD grant from FCT (SFRH/BD/8320/2002).

Literature cited

- Allen SK & Stanley JG. 1978. Reproductive sterility in polyploidy brook trout, *Salvelinus fontinalis*. *Transactions of the American Fisheries Society* 107: 473-478.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 1998. Diversity in the reproductive modes of females of the *Rutilus alburnoides* complex (Teleostei, Cyprinidae) as determined by analysis of cytochrome *b* sequences. *Evolution* 51: 1584-1592.
- Alves MJ, Coelho MM, Próspero MI & Collares-Pereira MJ. 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 MJ, Coelho MM & Collares-Pereira MJ. 2001. Evolution in action through hybridization and polyploidy in an Iberian freshwater fish: a genetic review. *Genetica* 111: 375-385.
- Alves MJ, Collares-Pereira MJ, Dowling TE & Coelho MM. 2002. The genetics of maintenance of an all-male lineage in the *Squalius alburnoides* complex. *Journal of Fish Biology* 60: 649-662.

- Alves MJ, Gromicho M, Collares-Pereira MJ, Crespo-Lopez E & Coelho MM. 2004. Simultaneous production of triploid and haploid eggs by triploid *Squalius alburnoides* (Teleostei: Cyprinidae). *Journal of Experimental Zoology* 301A: 552–558.
- Avisé JC, Quattro JM & Vrijenhoek RC. 1992. Molecular clones within organismal clones: Mitochondrial DNA phylogenies and the evolutionary history of unisexual vertebrates. *Evolutionary Biology* 26: 225-246.
- Beukeboom LW & Vrijenhoek RC. 1998. Evolutionary genetics and ecology of sperm-dependent parthenogenesis. *Journal of Evolutionary Biology* 11: 755-782.
- Bi K & Bogart JP. 2006. Identification of intergenomic recombinations in unisexual salamanders of the genus *Ambystoma* by genomic in situ hybridization (GISH). *Cytogenetics and Genome Research* 112: 307-312.
- Bogart JP, Elinson R & Licht LE. 1989. Temperature and sperm incorporation in polyploidy salamanders. *Science* 246: 1032-1034.
- Collares-Pereira MJ & Moreira da Costa L. 1999. Intraspecific and interspecific genome size variation in Iberian Cyprinidae and the problem of diploidy and polyploidy, with review of genome sizes within the family. *Folia Zoologica* 48(1): 61-76.
- Dawley RM. 1989. An introduction to unisexual vertebrates. In Dawley RM & Bogart JP (Eds.) *Evolution and ecology of unisexual vertebrates*. New York State Museum Bulletin 466, New York, pp 1-28.
- Dawley RM & Bogart JP (Eds.) 1989. *Evolution and ecology of unisexual vertebrates*. New York State Museum Bulletin 466, New York.
- Dowling TE & Secor CL. 1997. The role of hybridization and introgression in the diversification of animals. *Annual Review of Ecology and Systematics* 28: 593-619.
- Graf J-D & Pelaz MP. 1989. Evolutionary genetics of the *Rana esculenta* complex. In Dawley RM & Bogart JP (Eds.) *Evolution and ecology of unisexual vertebrates*. New York State Museum Bulletin 466, New York, pp 289-301.
- Gromicho M, Countanceau J-P, Ozouf-Costaz C & Collares-Pereira MJ. 2006. Contrast between extensive variation of 28S rDNA and stability of 5S rDNA and telomeric repeats in the diploid-polyploid *Squalius alburnoides* complex and in its maternal ancestor *Squalius pyrenaicus* (Teleostei, Cyprinidae). *Chromosome Research* 14: 297-306.
- Gui JF, Jia J, Liang SC & Jiang YG. 1992. Meiotic chromosome behaviour in male triploid transparent coloured crucian carp *Carassius auratus*. *Journal of Fish Biology* 41: 317-326.
- Flajshans M, Kvasnicka P & Ráb P. 1993. Genetic studies in tench (*Tinca tinca* L.): high incidence of spontaneous triploid. *Aquaculture* 110: 243-248.
- Janko K, Culling MA, Ráb P & Kotlík P. 2005. Ice age cloning – comparison of the Quaternary evolutionary histories of sexual and clonal forms of spiny loaches (*Cobitis*; Teleostei) using the analysis of mitochondrial DNA variation. *Molecular Ecology* 14: 2991-3004.
- Kawamura K, Ueda T, Aoki K & Hosoya K. 1999. Spermatozoa in triploids of the rosy bitterling *Rhodeus ocellatus ocellatus*. *Journal of Fish Biology* 55: 420-432.

- Lamatsch DK, Steinlein C, Schmid M & Scharfl M. 2000a. Noninvasive determination of genome size and ploidy level in fishes by flow cytometry: detection of triploid *Poecilia formosa*. *Cytometry* 39: 91-95.
- Lamatsch DK, Nanda I, Epplen JT, Schmid M & Scharfl M. 2000b. Unusual triploid males in a micromosome-carrying clone of the amazon molly, *Poecilia formosa*. *Cytogenetic and Genome Research* 91: 148-156.
- Matsubara K, Arai K & Suzuki R. 1995. Survival potential and chromosomes of progeny of triploid and pentaploid females in the loach, *Misgurnus anguillicaudatus*. *Aquaculture* 131: 37-48.
- Mori T, Saito S, Kishioka C & Arai K. 2006. Aquaculture performance of triploid barfin flounder *Verasper moseri*. *Fish Science* 72: 270-277.
- Oshima K, Morishima K, Yamaha E & Arai K. 2005. Reproductive capacity of triploid loaches obtained from Hokkaido Island. *Ichthyological Research* 52: 1-8.
- Ohtani H. 1992. Mechanism of chromosome elimination in the hybridogenetic spermatogenesis of allotriploid males between Japanese and European water frogs. *Chromosoma* 102: 158-162.
- Robalo J, Sousa-Santos C, Levy A & Almada VC. 2006. Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. *Molecular Phylogenetics and Evolution* 39: 276-281.
- Sambrook J, Fritsch EF & Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor, New York.
- Scali V, Passamonti M, Marescalchi O & Mantovani B. 2003. Linkage between sexual and asexual lineages: genome evolution in *Bacillus* stick insects. *Biological Journal of the Linnean Society* 79: 137-150.
- Sousa-Santos C, Collares-Pereira MJ & Almada VC. 2006. 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. *Journal of Fish Biology* 68(Supplement B): 292-301.
- Sousa-Santos C, Robalo JI, Collares-Pereira MJ & Almada VC. 2005. Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genomic constitution of hybrid organisms. *DNA Sequence* 16: 462-467.
- Stöck M, Lamatsch DK, Steinlein C, Epplen JT, Grosse W-R, Hock R, Klapperstück T, Lampert KP, Scheer U, Schmid M & Scharfl M. 2002. A bisexually reproducing all-triploid vertebrate. *Nature Genetics* 30: 325-328.
- Thorgaard GH & Gall GAE. 1979. Adult triploids in a rainbow trout family. *Genetics* 93: 961-973.
- Vrijenhoek RC. 1989. Genetic and ecological constraints on the origins and establishment of unisexual vertebrates. In Dawley RM & Bogart JP (Eds.) Evolution and ecology of unisexual vertebrates. New York State Museum Bulletin 466, New York, pp 24-31.
- Vrijenhoek RC. 1994. Unisexual fish: model systems for studying ecology and evolution. *Annual Reviews in Ecology and Systematics* 25: 71-96.
- Vrijenhoek RC. 2006. Polyploid hybrids: multiple origins of a treefrog species. *Current Biology* 16: R245-R247.

Zhang Q & Arai K. 1999. Distribution and reproductive capacity of natural triploid individuals and occurrence of unreduced eggs as a cause of polyploidization in the loach, *Misgurnus anguillicaudatus*. *Ichthyological Research* 46: 153-161.

Chapter 6

Reproductive success of nuclear non-hybrid males of *Squalius alburnoides* hybridogenetic complex (Teleostei, Cyprinidae): an example of interplay between female choice and ecological pressures?

Sousa-Santos C, Collares-Pereira MJ & Almada V (2006)

Acta Ethologica 9(1): 31-36

Reproductive success of nuclear non-hybrid males of *Squalius alburnoides*
hybridogenetic complex (Teleostei, Cyprinidae): an example of interplay between
female choice and ecological pressures?

Sousa-Santos C, Collares-Pereira MJ & Almada V

Abstract The hybridogenetic fish complex *Squalius alburnoides* comprises diploid males with non-hybrid nuclear genomes and several hybrid forms varying in ploidy and relative proportions of the parental genomes. In this paper, we present evidence that in captivity females prefer to mate with non-hybrid males. We suggest that female choice combined with different ecological requirements of hybrid and non-hybrid males may explain the extreme variation in the relative abundance of male types among drainages.

Keywords Cyprinidae, female choice, sexual selection, reproductive behaviour, courting displays

Introduction

The *Squalius alburnoides* minnows are endemic to the Iberian Peninsula and constitute a complex of various ploidy forms, originated by interspecific hybridization between *S. pyrenaicus* females (P genome) and males from an extinct *Anaecypris*-like ancestor (A genome) (Alves *et al.* 2001; Robalo *et al.* 2006).

Aside from variation in the ploidy level, the *S. alburnoides* individuals also differ in the relative proportion of the P and A genomes (Alves *et al.* 2001). It is important to stress that both male and female diploid PA hybrids (the form that probably was at the origin of the complex) do not undergo meiosis, producing gametes clonally (reviewed in Alves *et al.* 2001). Most triploids are PAA females that reproduce by meiotic hybridogenesis once they exclude the P-genome, after which the AA genomes undergo meiosis and recombination, generating haploid A gametes, although they can produce other gametes, namely, AA and even PAA eggs (Alves *et al.* 2004). Meiosis is also apparently normal in PPAA

tetraploids, a rare form that produces PA gametes (Alves *et al.* 2001). There is also a AA form, which is much smaller and morphologically distinct from the hybrid forms (Collares-Pereira 1984) – see details below. These diploid fish with nuclear non-hybrid genomes are males that undergo normal meiosis and generate A sperm. When the A sperm of these males meets A eggs from PAA females, new diploid males carrying the nuclear genome of the paternal ancestor are found again. This brief description outlines the high complexity and diversity of this fish complex but we believe it is helpful to give some background on this unusual diploid-polyploid complex.

Although triploid females predominate in all populations, the frequency at which the other forms occur differs significantly among drainages – Table 1.

Table 1 – Minimum and maximum relative frequencies of the different forms of both genders of the *S. alburnoides* found in the northern and southern river basins (adapted from Alves *et al.* 2001 and Pala & Coelho 2005). *River basins subjected to extreme summer droughts.

	Ploidy	Northern river basins			Genome	Southern river basins		
		Genome	Douro	Mondego		Tejo	Sado*	Guadiana*
females	2n	CA	0-4%	0-10%	PA	0-15%	36-77%	0-35%
	3n	CAA/CCA	36-90%	80-90%	PAA/PPA	50-100%	19-70%	11-88%
	4n	CCAA	0-1%	-	PPAA	0-10%	0-2%	-
males	2n	CA	10-14%	5-14.9%	PA	0-23%	-	-
	3n	CAA/CCA	0-1%	5.3%	PAA/PPA	0-22%	0-4%	0-5%
	4n	CCAA	0-1%	0-5%	PPAA	0-14%	0-2%	-
	2n	AA	-	-	AA	0-16%	0-48%	8-89%

In contrast with the situation in many other complexes of hybrid origin that are almost exclusively composed of females (Vrijenhoek *et al.* 1989), fertile males of different ploidies and genome constitutions are also found (Alves *et al.* 2001). Thus, females of *S. alburnoides* mate not only with males of other sympatric *Squalius* species – *S. carolitertii* in the northern rivers, *S. pyrenaicus* in the south and *S. aradensis* in the southwest – but also with males of *S. alburnoides* (Alves *et al.* 2001; Cunha *et al.* 2004; Pala & Coelho 2005; Sousa-Santos *et al.* 2005, in press¹). The later include not only individuals with hybrid nuclear genomes, but also non-hybrid diploids reconstituted from the hybrids and

¹ Updated citation: Sousa-Santos et al. (2006) *Journal of Fish Biology* 68 (Supplement B): 292-301.

presenting the nuclear genome of the paternal ancestor in homozygosity (AA genome) (Alves *et al.* 2002). Their mitochondrial DNA is that of other *Squalius* species with which the complex hybridizes, the original mtDNA of the missing ancestor having been lost in evolution (Alves *et al.* 1997; Carmona *et al.* 1997).

The non-hybrid and hybrid males constitute two very distinct morphotypes, differing in size, morphology and ecological requirements (Collares-Pereira 1984; Martins *et al.* 1998; Ribeiro *et al.* 2003): non-hybrids are smaller than hybrids and the males of other sympatric *Squalius* species, and tend to prefer habitats characterized by shallow waters and higher temperatures, conditions that are typical of the rivers of south Iberia, when summer droughts reduce many streams to pools. In contrast, *S. alburnoides* females and other *Squalius* species show a preference for deeper waters and higher current velocities (Martins *et al.* 1998; Pires *et al.* 1999). The non-hybrid males are the most common male type in most of the southern populations but are apparently absent from the populations of the northern rivers (Table 1). The causes of these discrepancies among drainages are still unclear.

The studies of the processes of gamete formation in this fish complex (see details in Alves *et al.* 2001) support the conclusion that the persistence of these nuclear non-hybrid males is self dependent, as they can only be generated from crosses involving males with their own genomic constitution and the most abundant form of triploid females (PAA). Their persistence in many populations constitutes a very interesting evolutionary problem, as the frequency of this type of males is inherently unstable. Indeed, if they are lost from a population, they are unlikely to be produced again. Furthermore, females do not depend on this male type, because they can mate with hybrid males and with the males of other *Squalius* species. Hence, some advantage must exist that explains the evolutionary persistence of this male type in the southern populations.

Female choice may play an important role in determining the persistence and abundance of the non-hybrid males in southern drainages: if, despite being smaller, the non-hybrid males are favoured by female choice this may explain their abundance in the southern drainages, where the ecological conditions tend also to favour them. To test this hypothesis, we investigated the spawning behaviour of *S. alburnoides* in captivity in a setting where *S. alburnoides* females had free access to hybrid and non-hybrid conspecific males.

Materials and methods

A captive population of 56 *Squalius alburnoides* was established with individuals electrofished in the Tagus River basin (R. Sorraia), in the central part of Portugal. Fish were maintained in an outdoor aquarium (150x50x70cm) under natural conditions of light and temperature, and were fed with commercial flake food. This density of fish that, at first glance, may look much higher than that found in nature, is likely not very different from the conditions that the complex meets in its natural habitats. Indeed, the complex lives for the most part in Mediterranean rivers and streams, which in the spring and summer get almost entirely reduced to small pools where fish are forced to congregate, being retained there for many months. This population showed breeding cycles that coincided with the reproduction in the field (see Ribeiro *et al.* 2003) and spawned spontaneously in 3 consecutive years.

In the absence of any externally visible sexual dimorphism, sex identification was made during the breeding season by applying a mild pressure on the abdomen of the fish. Mature males were easily identified by the release of milt (N=11), while in mature females a few eggs were extruded (N=45).

Three categories of fish were visually distinguished during the breeding season: females, non-hybrid males and hybrid males. Mature females were identifiable in the video recordings due to their larger size and swollen abdomens. The distinction between hybrid (mean body size of 83.35 ± 13.24 mm, N=7) and non-hybrid males (mean body size of 53.18 ± 9.56 mm, N=4) was based on morphological differences: non-hybrid males (previously described as "form B" of the complex by Collares-Pereira 1984) are small fish with a narrow body and pointed head, straight dorsal profile and convex ventral profile, terminal mouth with a greater lower jaw so that the opening is turned slightly upward, and deeply forked caudal lobes. The presence of a straightforward anterior portion of the lateral line was also an unambiguous diagnostic character used to discriminate between non-hybrid and hybrid males (which presented a typical curvilinear lateral line) (C. Sousa-Santos, unpublished data). This distinction was made after the observation of 33 specimens, whose homozygous genomic constitution was confirmed with the sequencing of the beta-actin nuclear gene (unpublished data and Sousa-Santos *et al.* 2005).

Ad libitum observations (sensu Martin & Bateson 1993) were conducted 5 days a week between March 18 and July 4, 2003. The breeding season extended from April 30 to

July 3 during which 8 days of spawning activity were observed, comprising a total of 211 spawning events (with egg-laying). Spawning behaviour started early in the morning and continued until sundown. Maximum water temperatures, recorded daily during the observation period, ranged from 18 to 28°C (mean $\pm$ SE=21.6 $\pm$ 2.73°C). A total of 1,190 min of videotape recordings (Canon MV3 camera) were used to allow subsequent analysis of courtship and spawning. Agonistic behaviours, when present, were recorded. From all the spawning events recorded, 31 were chosen for the analysis due to optimal image requisites, such as the possibility to track the courting behaviours before the spawning and a clear distinction of the individuals involved. The number of times the focal females interrupted the spawning sequence when courted in different courtship contexts was compared with a Kruskal-Wallis test as implemented in Statistica 5.0 (Statsoft). Dunn's tests were subsequently performed to identify differences between specific contexts after the implementation of Siegel & Castellan (1988). To test if there was any bias favouring a given male type, we compared the frequencies of successful spawnings involving pairs with a hybrid and pairs with a non-hybrid male, against the frequencies that would be expected if females paired randomly. This null hypothesis assumed that the spawning pairs involving a hybrid or a non-hybrid would reflect the proportions of the two male types in the observation tank. This comparison was performed with a goodness of fit Chi²-square (χ^2), the significance of which was accessed by a simulation procedure with the program ADERSIM (Vitor Almada). This program generates 1,000 simulations in which values are randomly assigned to the different classes with probabilities reflecting their expected frequencies. The number of times out of 1,000 that for each class-observed values are equal or greater and equal or smaller than the simulated values allows the assessment of the significance of the results. In addition, the number of times that the χ^2 was equal or exceeded by the χ^2 computed for each simulation is also provided. The procedure is described in detail in Almada & Oliveira (1997) and has the advantage over the conventional χ^2 tests of being free of the assumptions of the χ^2 distributions, at the same time allowing the assessment of the significance of the deviations of individual classes.

Ethical note

Fishes were captured in a non-imperilled population. The sample size was chosen to maximize the likelihood of collecting all the forms in the complex but without risking

depletion of the natural stock. Electrofishing was performed in low duration pulses to avoid killing juveniles (400-500 V and 2-4 A). All fish survived the transport to the laboratory, using aerated containers. The captivity conditions were enriched to simulate natural conditions with respect to vegetation and substrate. The sex determination procedure was performed while the fish were captured briefly in a hand net, and this took only a few seconds. Fish were immediately returned to the water and no casualties or behavioural abnormalities were detected.

Results

During the breeding season, fish aggregated close to the substrate, in the most aerated zone of the tank. Indeed, when a blue dye was added to the water close to the outlet of the filter pump it became apparent that the spawning area was the part of the substrate more directly exposed to the current. All spawning sequences were initiated when a female entered the spawning area, where one or more males were displaying courtship behaviours: quivering, shuffling the dorsal and pectoral fins, and assuming "head-down" postures with the abdomen oriented towards the female. These male displays could evoke two different responses in females: 1) the female left the spawning site (and was generally chased by one or two courting males), and 2) the female also exhibited courting behaviours that culminated in the deposition of eggs.

More details on the interactions between the two sexes during reproduction are provided in the "Electronic supplementary material".

The spawning sequence was performed in synchrony by the female and by only one male in 61.29% ($N=19$) of the events, while the remaining spawnings involved trios (38.71%, $N=12$). When the female spawned in a trio, each male assumed a side by side position on one of the flanks of the female. Trios comprised two non-hybrid males ($N=5$) or one non-hybrid and one hybrid male ($N=7$).

In pairs, when the female spawned with a single male, the male was a hybrid in only two out of 19 occasions and a non-hybrid in the remaining 17 cases. This difference was significant ($\chi^2_1=5.48$, $N=2$, $P<0.05$), allowing us to reject the null hypothesis that all males had equal chances to participate in spawnings. In all cases, it was the female that

approached the spawning site, reaching to a very close proximity of the displaying male(s).

Concerning courting trios with two non-hybrid males ($N=5$), the first male manoeuvred to position himself side by side with the female, followed by a second male that was in the vicinity and occupied the vacant flank. Immediately thereafter, the typical spawning sequence was initiated.

In trios with hybrid and non-hybrid males ($N=7$), the mate that first took the side-by-side position against the female was the non-hybrid male in three cases and the hybrid male in the remaining events. However, it is interesting to note that when the hybrid male was the first to contact the flank of the female, the female initiated a series of manoeuvres and changes of position which ceased only when the non-hybrid male occupied the vacant flank. At this stage, the female initiated a new spawning sequence, culminating with egg deposition. Before that, when only the hybrid male was present, the female always interrupted the sequence at an initial stage, manoeuvring to another place in the spawning site where the spawning sequence was initiated again.

We assumed that the number of times that the female interrupts a spawning sequence (IS value) is inversely related to its readiness to spawn. The medians of the IS value (and the interquartile ranges - IQR) for pairs with one non-hybrid, pairs with one hybrid, trios with non-hybrids and trios with one hybrid and one non-hybrid were, 1 (IQR between 1 and 1), 1.5 (IQR between 1.25 and 1.75), 2 (IQR between 2 and 2) and 5 (IQR between 3 and 5.5), respectively. The medians of the four groups differed significantly (Kruskal-Wallis ANOVA on ranks: $H_{3, 31}=9.619$, $P=0.02$) - and all post hoc comparisons using a Dunn's test were significant at $P<0.05$ (two-tailed test). The smallest values were found when the females were mating with a single non-hybrid male and were largest when a hybrid male was present in a trio.

During the spawning events, agonistic male-male interactions were negligible: two isolated agonistic behaviours were performed in the same spawning event by a hybrid towards a non-hybrid male ("pushing" and "blocking path") and in two other different events, a hybrid male tried to position himself side by side with the female, laterally displacing the other male (a non-hybrid male in one of the events and a hybrid in the other).

Discussion

During the breeding season, *S. alburnoides* males gather in the spawning area and perform a series of courtship displays that attract females, which ultimately have the choice of the mate(s) and of the location and timing of the spawning. In fact, as described by Katano & Hakoyama (1997), the presence of courtship behaviour in fishes classified as open substratum lithophil spawners (Balon 1975), such as *S. alburnoides*, suggests that female mate choice likely plays an important role in determining male mating success.

Our data point to the existence of a female preference to spawn with nuclear non-hybrid males: 1) the majority of the spawning events involved at least one non-hybrid male, although there were more hybrid than non-hybrid males available; 2) there were significantly fewer interruptions by the female after the spawning sequence had been initiated when the female spawned with one non-hybrid male, than when the spawning event involved a trio with one non-hybrid and one hybrid male; and 3) in trios with one hybrid and one non-hybrid male, the female, when approached first by the hybrid male, manoeuvred to avoid him and only performed the complete spawning sequence when the non-hybrid male joined the pair.

As the fish were not individually marked, some degree of pseudoreplication that could limit the validity of our conclusions cannot be ruled out. Even if the fish were marked, with the large number of individuals involved and the speed of spawning movements, it would be very difficult to identify the participants reliably. We believe, however, that the results presented are not artefacts due to mere idiosyncratic peculiarities of some individuals. In females, the abdomen shrinks after spawning and only gets swollen again after many days or weeks, when a female is ready to release a new batch of eggs (see Ribeiro *et al.* 2003). This difference between females that are ready to spawn and those that have just spawned made us almost sure that successive observations very likely represented spawning events involving different females. The males in our study differed in size and morphology and, although we cannot demonstrate that the same male was not observed more than once, we are confident that the observations used here involved different individuals.

As non-hybrid males are much smaller than hybrid males and males of other *Squalius* species with which *S. alburnoides* females mate, it is not possible to test experimentally if the females are choosing based on size or on other attributes of males.

Also, there are no interspecific studies on the choice by females in this group of *Squalius* and it is unknown how the differences in size may affect male success and female preferences.

Regardless of the mechanism involved in female choice, the important point is that non-hybrid males are favoured to the detriment of hybrid males, playing an important role in the reproduction of the *S. alburnoides* complex. This is surprising when we consider their small size and the fact that females may also mate with other males, both hybrid and members of other *Squalius* species. In fact, the smaller size of these non-hybrid males, which at first sight could be seen as a disadvantage, probably enables them to perform a more vigorous body quivering and faster manoeuvres. Their displays in a situation of high density of males can, therefore, represent more intense stimuli than the ones performed by other males, which could be a possible explanation for the preference of the females. A similar situation was described for *Rhodeus sericeus*: the vigour of male courtship has the stronger effect on female mate choice decisions, followed by male body size, in experiments with no male-male competition (Reichard *et al.* 2005), despite the fact that in this species males actively defend territories (Smith *et al.* 2004), a condition that usually favours bigger males. This is in agreement with Andersson (1994) who compared several taxa and showed that large body size in males is usually selected via male-male competition and less often by female choice. There are no interspecific studies on mate choice by females in this group of *Squalius*.

As stated above, genetic evidence showed that the non-hybrid males are always sons of males of their own type, thus, their substantial frequency in southern drainages is in itself a direct proof that they achieve a substantial number of fertilizations.

Given the preference of females for the small non-hybrid males, how can the difference in their frequencies among drainages be explained? In the first place, as noted above, the mechanism of their production implies that if they became extinct from a population, they are unlikely to be regenerated. If ecological conditions are adverse to the non-hybrid males, as is probably the case in the north, where rivers have no summer droughts and tend to have stronger currents, they may be at a disadvantage when compared with hybrid males and males of other sympatric *Squalius*. Indeed, non-hybrids are so much smaller that it is unlikely that they are in a disadvantage when compared to larger males in the northern rivers, where currents are often strong. This may have led to a

gradual decrease in their number even if the females from the northern populations had the same mating preferences for smaller non-hybrid males.

By contrast, in the south, their small size and tolerance to shallow, warm and still waters (Martins *et al.* 1998) combined with the female preference may help to maintain them in populations.

In addition, the fact that the non-hybrids undergo normal meiosis, that is absent in hybrids, which are mostly diploids that produce sperm clonally (Alves *et al.* 1999; Pala & Coelho 2005), will also contribute to maintain a higher level of genetic variability in the progeny. This fact may be important in the harsh and highly unpredictable Mediterranean environments of the south (Mesquita *et al.* 2005). Thus, the preference of the females for the small AA males may be evolutionarily advantageous due to the fact that these small males undergo normal meiosis and recombination, and females mating preferentially with them could benefit their offspring.

If our hypothesis proves to be correct, the persistence of non-hybrid males may depend upon a balance between the behavioural preferences of the females and differential ecological requirements of both male types, with distinct outcomes in different environments. Experimental tests in tanks simulating various habitats differing in water flow intensity, combined with DNA fingerprinting of adults and progenies, may provide a way to test these hypotheses.

All these considerations assumed that the AA males, like all *S. alburnoides*, originated in the southern drainages, having subsequently migrated to the northern rivers (Alves *et al.* 1997). In this scenario, AA males should be present in *S. alburnoides* populations, as their origin and their absence in the north should be viewed as a secondary loss once it is difficult to explain why, during the process of colonization of northern rivers, the small AA males so common in several southern populations would be excluded from the northward dispersal event.

Acknowledgements The authors wish to thank R. Sousa-Santos for help in sample collection, J. Robalo and T. Bento for help in maintaining the fish tank and two anonymous referees that helped to improve the text with interesting suggestions. The study was funded by the FCT Pluriannual Program (UI&D 331/94 and UI&D

329/94) (FEDER participation). C. Sousa-Santos was supported by a PhD grant from FCT (SFRH/BD/8320/2002). A permit for field work was provided by "Direcção Geral dos Recursos Florestais". The procedures involved in fish capture, observations and maintenance are in full agreement with the Portuguese and European laws and regulations, namely those pertaining to animal welfare and wildlife conservation.

References

- Almada VC & Oliveira RF. 1997. Sobre o Uso de Estatística de Simulação em Estudos de Comportamento. *Análise Psicológica* 15: 97-109.
- Alves MJ, Coelho MM, Collares-Pereira MM & Dowling TE. 1997. Maternal ancestry of the *Rutilus alburnoides* complex (Teleostei, Cyprinidae) as determined by analysis of cytochrome *b* sequences. *Evolution* 51: 1584–1592.
- Alves MJ, Coelho MM, Próspero MI & Collares-Pereira MJ. 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 MJ, Coelho MM & Collares-Pereira MJ. 2001. Evolution in action through hybridization and polyploidy in an Iberian freshwater fish: a genetic review. *Genetica* 111: 375–385.
- Alves MJ, Collares-Pereira MJ, Dowling TE & Coelho MM. 2002. The genetics of maintenance of an all-male lineage in the *Squalius alburnoides* complex. *Journal of Fish Biology* 60: 649-662.
- Andersson M. 1994. *Sexual selection*. Princeton University Press, Princeton.
- Balon EK. 1975. Reproductive guilds of fishes: a proposal and definition. *Journal of the Fisheries Research Board of Canada* 32: 821-864.
- Collares-Pereira MJ. 1984. The "*Rutilus alburnoides* (Steindachner, 1866) complex" (Pisces, Cyprinidae). I. Biometrical analysis of some Portuguese populations. *Arquivos do Museu Bocage Série AII*: 111-143.
- Cunha C, Coelho MM, Carmona JA & Doadrio I. 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Molecular Ecology* 13: 2807-2817.
- Katano O & Hakoyama H. 1997. Spawning behaviour of *Hemibarbus barbus* (Cyprinidae). *Copeia* 3: 620-622.
- Martin P & Bateson P. 1993. *Measuring behaviour – An introductory guide*. Cambridge University Press, Cambridge.
- Martins MJ, Collares-Pereira MJ, Cowx IG & Coelho MM. 1998. Diploids *v.* triploids of *Rutilus alburnoides*: spatial segregation and morphological differences. *Journal of Fish Biology* 52: 817-828.
- Mesquita N, Hänfling B, Carvalho GR & Coelho MM. 2005. Phylogeography of the cyprinid *Squalius aradensis* and implications for conservation of the endemic freshwater fauna of southern Portugal. *Molecular Ecology* 14: 1939-1954.
- Pala I & Coelho MM. 2005. Contrasting views over a hybrid complex: between speciation and evolutionary "dead-end". *Gene* 347: 283-294.

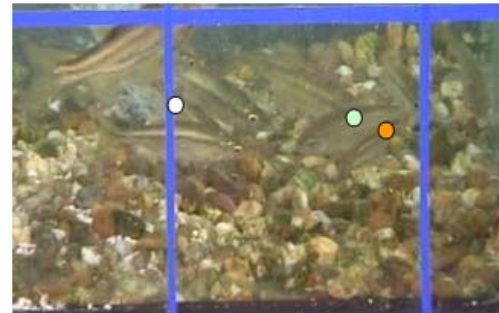
- Pires AM, Cowx IG & Coelho MM. 1999. Seasonal changes in fish community structure of intermittent streams in the middle reaches of the Guadiana basin, Portugal. *Journal of Fish Biology* 54: 235-249.
- Reichard M, Bryja J, Ondracková M, Dávidová M, Kaniewska P & Smith C. 2005. Sexual selection for male dominance reduces opportunities for female mate choice in the European bitterling (*Rhodeus sericeus*). *Molecular Ecology* 14: 1533-1542.
- Ribeiro F, Cowx IG, Tiago P., Filipe AF & Collares-Pereira MJ. 2003. Growth and reproductive traits of diploid and triploid forms of *Squalius alburnoides* cyprinid complex in a tributary of the Guadiana River, Portugal. *Archiv für Hydrobiologie* 156: 471-484.
- Robalo JI, Sousa-Santos C, Levy A & Almada VC. 2006. Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. *Molecular Phylogenetics and Evolution* 39: 276-281.
- Siegel S & Castellan NJ. 1988. *Nonparametric statistics for the behavioural sciences*. McGraw-Hill Book Company, Singapore.
- Smith C, Reichard M, Jurajda P & Przybylski M. 2004. The reproductive ecology of the European bitterling (*Rhodeus sericeus*). *Journal of Zoology* 262: 107-124.
- Sousa-Santos C, Collares-Pereira MJ & Almada V. in press. 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. *Journal of Fish Biology* 68 (Supplement B): 292-301.
- Sousa-Santos C, Robalo J, Collares-Pereira MJ & Almada V. 2005. Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genomic composition of hybrid organisms. *DNA Sequence* 16: 462-467.
- Vrijenhoek RC, Dawley RM, Cole CJ & Bogart JP. 1989. A list of the known unisexual vertebrates. In Dawley RM & Bogart JP (Eds.) *Evolution and Ecology of Unisexual Vertebrates*. New York State Museum Bulletin 466, New York, pp 19-23.

Electronic Supplementary Material

Legend: ○ female ● nonhybrid male ● hybrid male



1 – female enters the spawning site; both males assume a head-down position



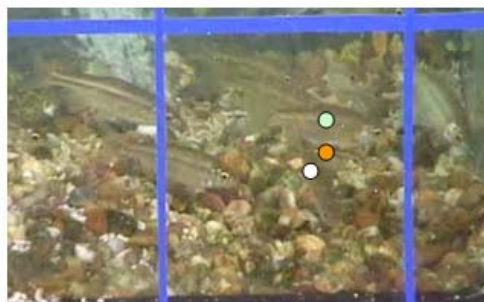
2 – males swim towards the female



3 – both males manoeuvre to be side-by-side with the female; the hybrid male is the first to contact the flank of the female

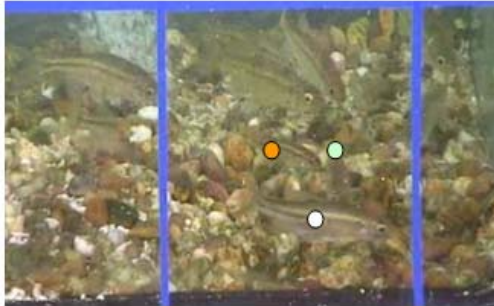


4 – female swims up in the water column (first demobilization)

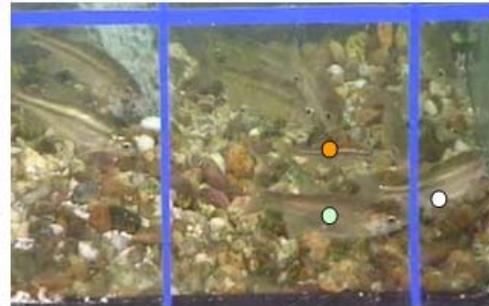


5 – female manoeuvres to other location

Legend: ○ female ● nonhybrid male ● hybrid male



6 – the hybrid male swims towards the female



7 – the hybrid male tries to be side-by-side with the female but she swims away (second demobilization)



8 – female swims towards the nonhybrid male, followed by the hybrid male



9 – female manoeuvres to be side-by-side with the nonhybrid male and they both assume a head-down position in synchrony; the hybrid male occupies the vacant flank of the female also in a head-down position



10 – at the same time, the female and the two males hit the substratum with the ventral part of the body (spawning in a trio)

PART **IV**

ORIGIN, DISPERSAL AND INTROGRESSION

Chapter 7

Molecular insights on the taxonomic position of the paternal ancestor
of the *Squalius alburnoides* hybridogenetic complex

Robalo J, Sousa-Santos C, Levy A & Almada V (2006)

Molecular Phylogenetics and Evolution 39(1): 276-281

Molecular insights on the taxonomic position of the paternal ancestor of the
Squalius alburnoides hybridogenetic complex

Robalo J, Sousa-Santos C, Levy A & Almada V

Introduction

Many hybrid fish lineages result from intrageneric crosses and are often asexual, being considered as “evolutionary dead-ends” (Vrijenhoek 1998). This seems not to be true in the *Squalius alburnoides* (Iberian minnow) Steindachner 1866 complex. This complex, that seems to have originated in a cross between distant species, has a remarkable variability in ploidy level and genomic composition, includes fertile fish of both sexes and some of its forms retain meiosis and recombination (Alves *et al.* 2001).

The *S. alburnoides* complex probably originated from crosses between *S. pyrenaicus* females and males from an unknown species. The complex includes $2n=50$, $3n=75$ and $4n=100$ hybrid forms (reviewed in Alves *et al.* 2001) with varying proportions of the two parental genomes (denoted by P and A, corresponding to the *S. pyrenaicus* and paternal ancestor genomes, respectively). Reconstituted diploid non-hybrids (AA genome) are also produced and are morphologically distinct from the diploid hybrid form of the complex (PA). These non-hybrid fish, normally males (females seem to be extremely rare), exhibit the nuclear genome of the paternal ancestor (AA) and *S. pyrenaicus*-like mtDNA. In the absence of an identified paternal species this suggests that they are reconstituted from the hybrids (Alves *et al.* 2002). The oogenesis of triploid females with PAA genomes frequently involves discarding the *Squalius* (P genome), followed by normal meiosis and recombination, generating A gametes. When these gametes fuse with the sperm of AA males, which also undergo normal meioses, new AA male progeny is generated. In the absence of *S. pyrenaicus*, such as in the northern basins of Portugal, this complex seems to be maintained by crosses with males of *S. carolitertii* (CC) and by diploid hybrid males (CA), although the mtDNA found in *S. alburnoides* fish is *S. pyrenaicus*-like (Cunha *et al.* 2004; Pala & Coelho 2005).

The Leuciscini presently found in the Iberian Peninsula include mainly species of the genera *Squalius*, *Chondrostoma* and *Anaecypris*. Studies based on allozymes showed

that *S. alburnoides* did not originate from intrageneric crosses between *Squalius* species and also ruled out members of the genera *Chondrostoma* and *Anaecypris* as paternal ancestors (Alves *et al.* 1997; Carmona *et al.* 1997). Recent studies still unpublished (Gilles, Dowling, Alves, Coelho and Collares-Pereira) using introns from two nuclear genes, seem to suggest that perhaps *A. hispanica*-like individuals may represent the paternal ancestor of this complex.

The aim of the present work was to investigate which of the genera considered is phylogenetically closest to the paternal species that originated this complex. This approach is based on the amplification of a segment of the beta-actin nuclear gene from a number of genera closely related to *S. pyrenaicus* and from the non-hybrid males of *S. alburnoides*. The topology obtained with beta-actin was compared to relatively complete phylogenies of the European cyprinids based on *cytb* (*e.g.* Briolay *et al.* 1998; Gilles *et al.* 1998; Zardoya & Doadrio 1999) and with our own reconstruction using a set of species comparable to that used with beta-actin.

Methods

Total genomic DNA was extracted from fin clips preserved in ethanol by an SDS/proteinase-k based protocol (adapted from Sambrook *et al.* 1989). For the *beta-actin* gene a total of 1062bp was amplified using the primers For-5'-ATGGATGATGAAATGCCGC-3' and Rev-5'-AGGATCTTCATGAGGTAGTC-3' (J. Robalo, *unpublished*). The amplification process was conducted as follows: 35 cycles of 94°C (30s), 55°C (40s), and 72°C (1min 30s). Further details may be requested from the corresponding author. The amplified fragment is homologous to a region of the beta-actin gene of *Cyprinus carpio* (GenBank Accession No. M24113), including introns B and C and three exons.

For the *cytb* gene a total of 1044 bp was amplified using the primers LCB1-5'-AATGACTTGAAGAACCACCGT-3' (Brito *et al.* 1997) and HA-5'-CAACGATCTCCGGTTACAAGAC-3' (Schmidt & Gold 1993). PCR conditions followed those in Cunha *et al.* (2004).

For both genes, each sample was sequenced in both directions with the same primers used for PCR. Sequencing reactions were performed by MacroGen Inc. in a MJ Research PTC-225 Peltier Thermal Cycler using a ABI PRISM BigDye™ Terminator Cycle Sequencing Kits with AmpliTaq DNA polymerase (FS enzyme) (Applied Biosystems), following the protocols supplied by the manufacturer.

All Accession Numbers from the present and previous studies are listed in Table I.

Sequences were aligned with BioEdit v.5.0.6.

The aligned sequences were analysed using distance (minimum evolution, ME), maximum likelihood (ML), maximum parsimony (MP) and Bayesian methods. For ME and ML analyses, we performed a hierarchical likelihood ratio test (LRT), using the program ModelTest 3.6 (Posada & Crandall 1998) to find the model of evolution that best fitted our data. These phylogenetic analyses were performed using PAUP 4.0 (Swofford 1998) and MrBayes 3.1 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). Bootstrap analysis were used to assess the relative robustness of branches of the MP (1000 replicates), the ME (1000 replicates) and the ML (100 replicates) trees. For the Bayesian analysis, *cytb* data were partitioned by codon base position, and a GTR+I+ Γ model was used for third base positions and a HKY model for first and second base positions. For beta-actin, gaps were coded as separate characters (Simmons & Ochoterena, 2000), using GapCoder (Young & Healy 2003). Data were partitioned and separate models used for each region, thus: GTR+I+ Γ model used for third base positions of exon and intron region, HKY model for first and second base positions of exons, and single rate for gap characters. For both genes, four separate analyses were performed (one with four million generations and three with one million generations) with four chains per analysis. The first 50,000 generations were discarded as "burn-in", the remaining generations were sampled every 100 generations, and majority-rule consensus trees were calculated from samples at stationarity.

To compare topologies recovered with both genes, the same species were used in both analyses, except for the genera *Phoxinus* and *Misgurnus* where, for lack of corresponding sequences for both fragments, we used different congeneric species.

We decided to include a phylogeny for the *cytb* gene because we could not get beta-actin sequences for all the species used in previous *cytb* studies. Thus, there was a

risk that our reduced data set for beta-actin was not directly comparable to *cyt b* studies with more taxa and different outgroups.

Species from the genus *Misgurnus* (Cobitidae) were used as outgroups.

Table 1 - Species considered in this study and their GenBank accession numbers (*cyt b* and beta-actin gene). Note: DQ061937 and AY943882 are haplotypes found for the beta-actin gene that are shared by *S. aradensis*- *S. torgalensis* and *S. pyrenaicus*- *S. carolitertii*, respectively. Legend: * - result from this work.

Species Name	GenBank accession number Cyt b gene	GenBank accession number Beta-actin gene
<i>Phoxinus lagowskii steindachneri</i>	AB162650	
<i>Misgurnus anguillicaudatus</i>	AF051868	
<i>Misgurnus mizolepis</i>		AF270649
<i>Leuciscus idus</i>	AY026397	DQ061947*
<i>Scardinius erythrophthalmus</i>	AY509848	DQ061949*
<i>Phoxinus oxycephalus</i>		AF200957
<i>Anaocypris hispanica</i>	AJ427814	DQ061936*
<i>Chondrostoma genei</i> ¹	AF533766	DQ061938*
<i>Chondrostoma lemmingii</i> ²	DQ089654*	DQ061940*
<i>Chondrostoma lusitanicum</i> ³	AY254584	DQ061941*
<i>Chondrostoma oligolepis</i> ⁴	DQ061932*	DQ061942*
<i>Chondrostoma polylepis</i> ⁵	AF045982	DQ061945*
<i>Chondrostoma prespense</i>	AF090747	DQ061944*
<i>Chondrostoma soetta</i>	AY568623	DQ061939*
<i>Chondrostoma turiense</i> ⁶	AY568619	DQ061946*
<i>Telestes souffia</i>	AY509862	DQ061950*
<i>Rutilus rutilus</i>	DQ061933*	DQ061948*
<i>Squalius aradensis</i>	AF421825	DQ061937*
<i>Squalius carolitertii</i>	AF045994	AY943882*
<i>Squalius pyrenaicus</i>	AF421826	AY943882*
<i>Squalius torgalensis</i>	DQ061934*	DQ061937*
<i>Squalius alburnoides</i>		AY943863*
<i>Cyprinus carpio</i>	NC001606	M24113
<i>Gobio gobio</i>	AY426589	DQ061935*

The sequence of *S. alburnoides* chosen to integrate this paper is the most common haplotype found through out the species area of distribution, both in hybrids and in reconstituted non-hybrid males, and was recovered from 51 fishes out of 103 (unpublished data). All other haplotypes differ from this most common one by few

According to the recent taxonomical revision of Robalo *et al.* (2007) *Mol. Phylogenet. Evol.* 42: 362-372:

¹ *Protochondrostoma genei*

² *Iberochondrostoma lemmingii*

³ *Iberochondrostoma lusitanicum*

⁴ *Achondrostoma oligolepis*

⁵ *Pseudochondrostoma polylepis*

⁶ *Parachondrostoma turiense*

mutational steps that rarely exceed one and are unlikely to affect the results presented. This sequence was found in homozygous condition in many reconstituted non-hybrid males which allowed us to reconstruct it from the chromatograms without ambiguities. As we also obtained homozygous sequences for the maternal ancestor of *S. alburnoides*, *S. pyrenaicus*, we were in a position to identify the overlapping patterns of the two species when they were present in hybrids. In this respect the beta-actin gene proved a very useful marker because it possesses a combination of three very convenient features: a) well-preserved exons that provide good landmarks to align the sequences and identify homologous regions; b) introns that are sufficiently variable to accumulate information useful in phylogenetic reconstructions; and c) the presence of frequent indels. These indels, when in heterozygous condition make possible the reconstruction of the two sequences that overlap in the chromatogram of a hybrid fish. Indeed, was recently published a method that permits the recognition and reconstruction of different nuclear DNA sequences in the same chromatogram through the detection of indels (Bhangale *et al.* 2005) that with some alterations (Sousa Santos *et al.*, unpublished) makes possible to distinguish copies of the gene of maternal origin (similar to those found in *S. pyrenaicus*) and reconstruct the remaining ones, of paternal origin.

For both genes, the saturation of transitions and transversions was checked by plotting the absolute number of changes of each codon position against uncorrected sequence divergence values (p). There was no evidence of saturation in the ingroup (graph not shown).

Results

Beta-actin gene

Among all the sequences studied, 283 sites were variable, and 77 were parsimony informative. The general time reversible model with among-site rate heterogeneity HKY+G (HKY, Hasegawa *et al.* 1985) was selected by ModelTest as the best fit to the data. Base frequencies were A=0.22, C=0.24, G=0.20, and T=0.34. Among-site rate variation was approximated with the gamma distribution shape parameter $\alpha=0.27$. The proportion of

invariable sites (I) was 0.7. MP analysis resulted in a consensus tree of 392 steps (Consistency Index, CI=0.87; Homoplasy Index, HI=0.12; Retention Index, RI=0.70). The results of the four phylogenetic inference methods are summarized in Fig. 1.

Cytb gene

Among all the sequences studied, 445 sites were variable and 363 were parsimony informative. The general time reversible model with among-site rate heterogeneity GTR+I+G (Lanave *et al.* 1984; Yang 1994) was selected by ModelTest as the best fit to the data. The rate matrix parameters estimated were: $R(a)=1.10$, $R(b)=38.6$, $R(c)=0.63$, $R(d)=1.81$, and $R(e)=1.0$. Base frequencies were A=0.29, C=0.31, G=0.11, and T=0.28. Among-site rate variation was approximated with the gamma distribution shape parameter $\alpha=1.02$. The proportion of invariable sites (I) was 0.52. MP analysis resulted in a consensus tree of 1596 steps (CI=0.43; Homoplasy Index, HI=0.57; Retention Index, RI=0.41). The results of the four phylogenetic inference methods are summarized in Fig. 2.

Phylogenetic reconstructions with all inference methods show some important congruent features for both genes (Figs. 1 and 2) and for the most part agreed with the findings of Briolay *et al.* (1998), Gilles *et al.* (1998) and Zardoya & Doadrio (1998, 1999). The division between the subfamilies Cyprininae and Leuciscinae was recovered with Cyprininae in a basal position. *Gobio gobio* is more closely related to Leuciscinae than to Cyprininae. Phoxinins and leuciscins emerge in a single clade, although in the present work the position of the *Phoxinus* species varied according with the reconstruction method. The beta-actin gene also confirms the polyphyly of the old genus *Leuciscus*, *Telestes* (= *Leuciscus*) *souffia* being undoubtedly related with the genus *Chondrostoma*. *Leuciscus idus* remains in an unresolved situation, at least with the taxa included in this analysis. The Iberian *Squalius* remain divided in the two groups: *S. aradensis*-*S. torgalensis* and *S. pyrenaicus*-*S. carolitertii*. In the monophyletic genus *Chondrostoma* the beta-actin gene did not resolve the polytomy found in the previous *cytb* studies of Doadrio & Carmona (2004), Durand *et al.* (2003) and Zardoya & Doadrio (1999). Trees derived from the two genes disagree in the placement of *Scardinius erythrophthalmus*. In the beta-actin tree, *S. alburnoides* is, according with all methods, strongly associated with *Anaecypris hispanica* (in the *cytb* tree it is not present because its mitochondrial DNA is *Squalius*-like).

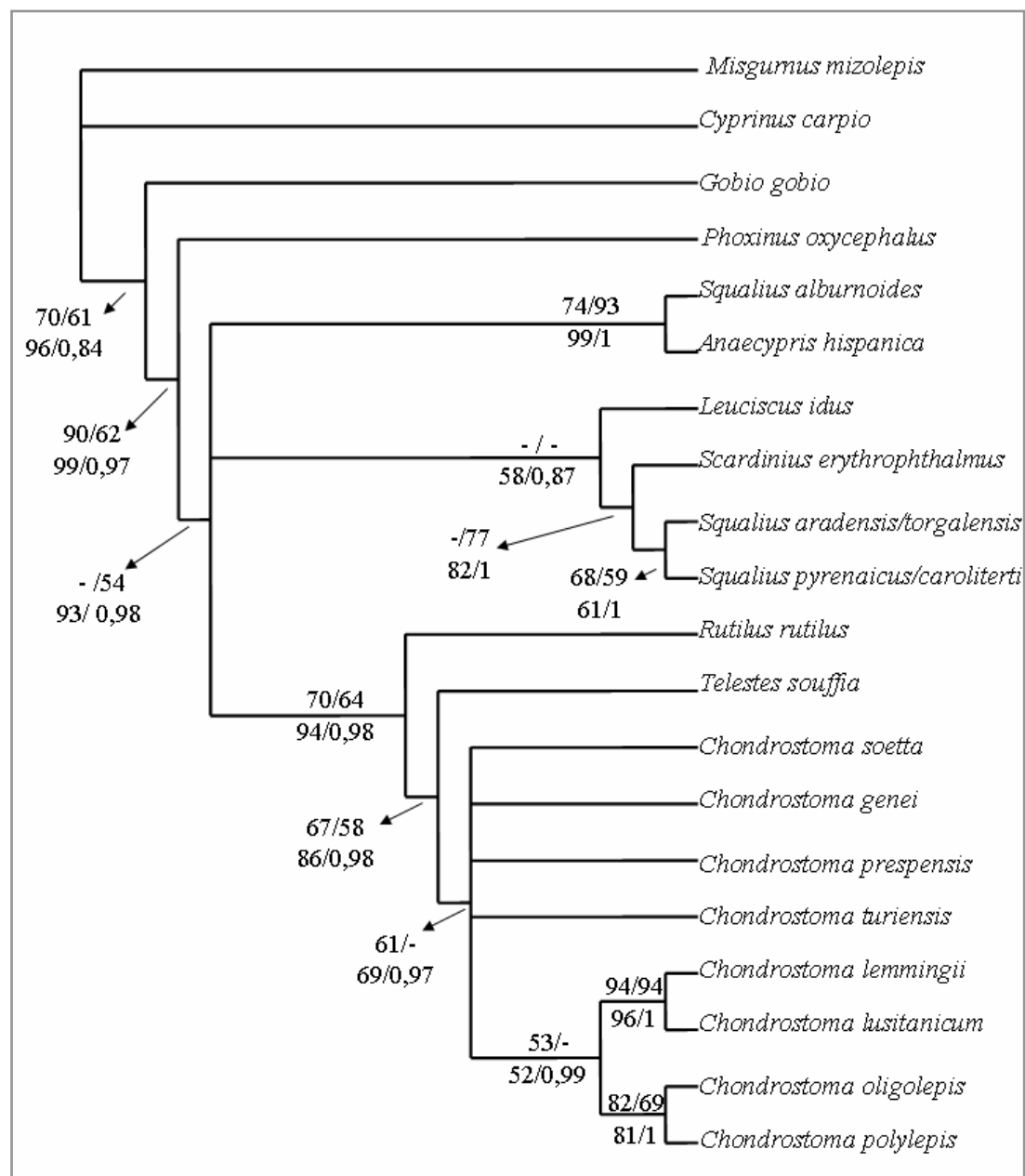


Figure 1 — Figure based on the MP tree. Phylogenetic relationships among the species based on *cytb* sequences. For each branch, numbers above represent the bootstrap values obtained for ME and ML; numbers below indicate those for MP and the posterior probabilities for Bayesian inference.

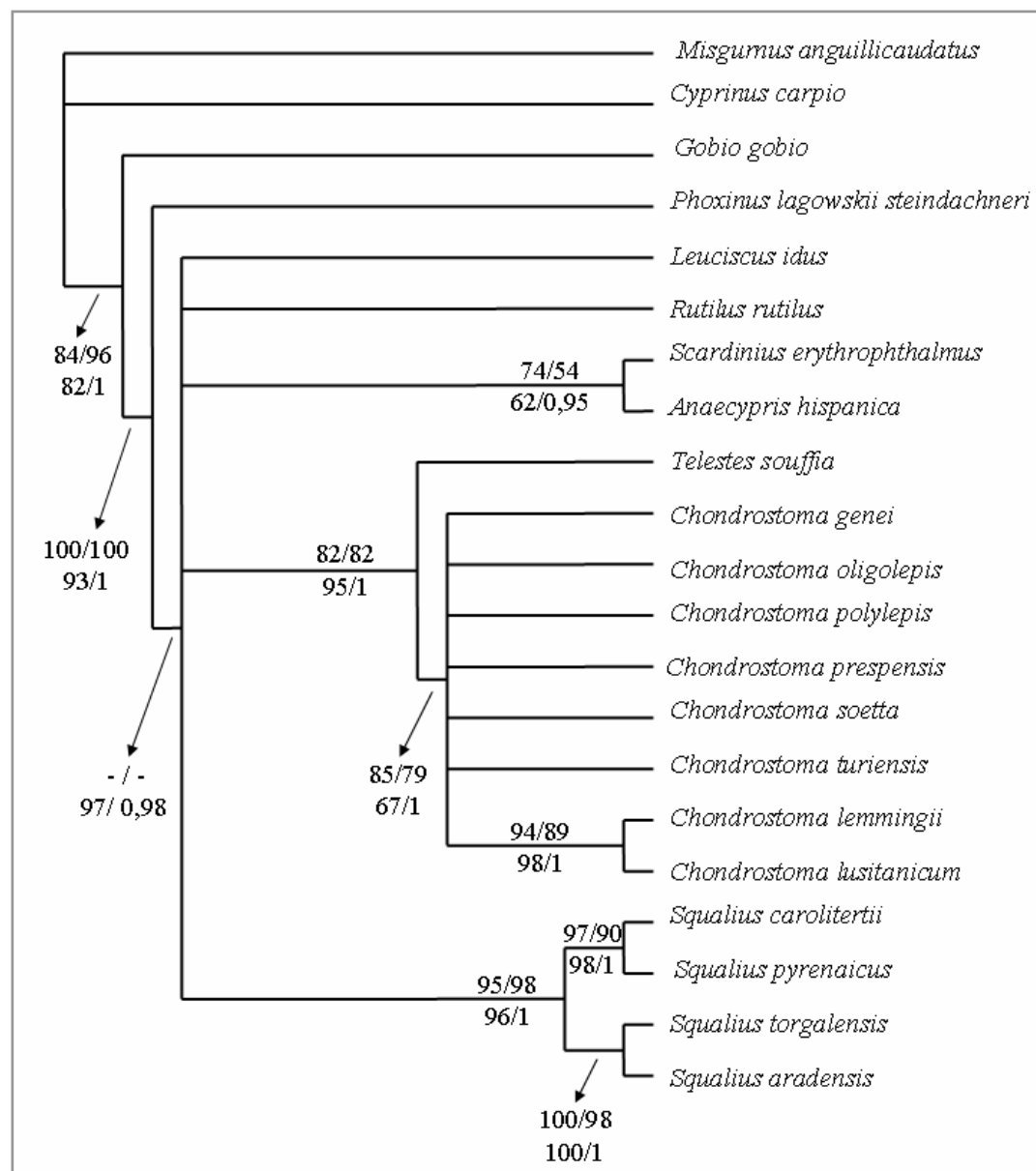


Figure 2 — Figure based on the MP tree. Phylogenetic relationships among the species based on beta-actin sequences. For each branch, numbers above represent the bootstrap values obtained for ME and ML; numbers below indicate those for MP and the posterior probabilities for Bayesian inference.

Discussion

The results of the present study clearly point to two main conclusions: 1) the paternal ancestor of *S. alburnoides* did not belong to the genera *Squalius* or *Chondrostoma* and 2) of the species included in this study, *A. hispanica* is the closest to the paternal ancestor (as already suggested by Gilles, Dowling, Alves, Coelho and Collares-Pereira, unpublished).

During the history of *S. alburnoides* some recombination may have taken place between the *Squalius* and *alburnoides* genomes and thus the beta-actin sequences of the reconstituted non-hybrid males may not be identical to the sequence of the paternal ancestor. The patristic distance between *S. alburnoides* and *Squalius pyrenaicus*/*S. carolitertii* beta-actin sequences is 1.48% while that between *A. hispanica* and *S. pyrenaicus*/*S. carolitertii* is 1.80%. The *p* distance between *S. alburnoides* and *A. hispanica* is 1.48%. Although the hypothesis of some recombination cannot be excluded, the fact that all inference methods recovered the *S. alburnoides* beta-actin sequence with that of *A. hispanica*, with very high bootstrap support, indicates that the signal still present in the beta-actin gene is sufficiently strong to trace its phylogenetic relationships.

Our suggestion is that, in the past, one species of the same clade of *A. hispanica* may have hybridized with *S. pyrenaicus* and originated *S. alburnoides*. The morphology of the reconstituted AA males supports this conclusion: they are similar to *A. hispanica* in size, general body shape and coloration, but they also differ from it in several important characters (*e.g.* structure of the lateral line, lateral scale counts and number of gill rakers) (Collares-Pereira 1983). The recent discovery of frequent natural intergeneric hybrids between *Squalius cephalus* and a species of *Chalcalburnus* (Ünver & Erk'Äkan 2005), a member of the same clade as *A. hispanica* (Zardoya & Doadrio 1999), supports our hypothesis of an intergeneric hybrid origin.

In spite of its hybrid origin *S. alburnoides* seems to be, from an evolutionary perspective, a very successful fish. It is often much more abundant than other sympatric *Squalius*. Alves *et al.* (2001) argue that the continuous shifting between forms, with P and A nuclear genomes being cyclically lost, gained or replaced by new genomes, allows the introduction of new genetic material. The evolutionary potential of this species, in terms of recombination and maintenance of genetic variability, may be even enhanced by the

presence of tetraploid fishes of both sexes in some natural populations. Indeed, PPAA tetraploids seem to undergo normal meiosis. Crosses between them mean that sexual reproduction is restored, while their crosses with other forms of the complex are a way to introduce recombination in the whole system (Alves *et al.* 2001).

The beta-actin gene evolves at a much slower rate than the *cytb* gene. For example, the difference between *S. pyrenaicus* and *S. carolitertii* was 6.13 % for the *cytb* gene, while the majority of fish from both species share the same haplotype for the beta-actin gene. The distance between species of *Squalius* and *Chondrostoma* averaged 14.05% for the *cytb* and 1.77% for beta-actin gene. Because of its slow rate of evolution, the beta-actin may prove potentially very useful in studies of cyprinid phylogeny, particularly in resolving intrageneric and intertribe relationships.

Acknowledgements We thank G. Lemos, Sousa-Santos family and T. Bento for all the help in field work and fish maintenance. We also thank I. Doadrio for the samples of European cyprinids. DGF provided all authorizations for field work. This study was funded by the Pluriannual Program (FCT, UI&D 331/94 with FEDER participation). C. Sousa-Santos and A. Levy were supported by FCT grants (SFRH/BD/8320/2002 and SFRH/BPD/18067/2004, respectively).

References

- Alves MJ, Coelho MM & Collares-Pereira MJ. 1997. The *Rutilus alburnoides* complex (Cyprinidae): evidence for hybrid origin. *Journal of Zoological Systematics and Evolutionary Research* 35: 1–10.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 2001. Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: a genetic review. *Genetica* 111: 375–385.
- Alves MJ, Collares-Pereira MJ, Dowling MJ & Coelho MM. 2002. The genetics of maintenance of an all-male lineage in the *Squalius alburnoides* complex. *Journal of Fish Biology* 60: 649–662.
- Bhangale TR, Rieder MJ, Livingston RJ & Nickerson DA. 2005. Comprehensive identification and characterization of diallelic insertion-deletion polymorphisms in 330 human candidate genes. *Human Molecular Genetics* 14: 59–69.
- Briolay J, Galtier N, Brito MR & Bouvet Y. 1998. Molecular phylogeny of Cyprinidae inferred from cytochrome *b* DNA sequences. *Molecular Phylogenetics and Evolution* 9: 100–108.

- Brito RM, Briolay J, Galtier N, Bouvet Y & Coelho MM. 1997. Phylogenetic relationships within genus *Leuciscus* (Pisces, Cyprinidae) in Portuguese freshwaters, based on mitochondrial cytochrome *b* sequences. *Molecular Phylogenetics and Evolution* 8: 435-442.
- Carmona JA, Sanjur OI, Doadrio I, Machordom A & Vrijenhoek VC. 1997. Hybridogenetic reproduction and maternal ancestry of polyploid Iberian fish: the *Tropidophoxinellus alburnoides* complex. *Genetics* 146: 983-993.
- Collares-Pereira MJ. 1983. Estudo sistemático e citogenético dos pequenos ciprinídeos ibéricos pertencentes aos géneros *Chondrostoma* Agassiz, 1835, *Rutilus* Rafinesque, 1820 and *Anaocypris* Collares-Pereira, 1983. PhD Thesis. University of Lisbon, Lisbon, 511p.
- Cunha C, Coelho MM, Carmona JA & Doadrio I. 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Molecular Ecology* 13: 2807-2817.
- Doadrio I & Carmona JA. 2004. Phylogenetic relationships and biogeography of the genus *Chondrostoma* inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* 33: 802-815.
- Durand JD, Bianco PG, Laroche J & Gilles A. 2003. Insight into the origin of endemic Mediterranean ichthyofauna - Phylogeography of *Chondrostoma* genus (Teleostean, Cyprinidae). *Journal of Heredity* 94: 315-328.
- Gilles A, Lecointre G, Faure E, Chappaz R & Brun G. 1998. Mitochondrial phylogeny of the European cyprinids: implications for their systematics, reticulate evolution, and colonization time. *Molecular Phylogenetics and Evolution* 10: 132-143.
- Hasegawa M, Kishino H & Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 21: 160-174.
- Huelsenbeck JP & Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17(8): 754-755.
- Lanave C, Preparata G, Saccone C & Serio G. 1984. A new method for calculating evolutionary substitution rates. *Journal of Molecular Evolution* 20: 86-93.
- Pala I & Coelho MM. 2005. Contrasting views over a hybrid complex: between speciation and evolutionary "dead-end". *Gene* 347: 283-294.
- Posada D & Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14(9): 817-818.
- Ronquist F & Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(12): 1572-1574.
- Sambrook J, Fritsch EF & Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor, New York, 999 pp.
- Schmidt TR & Gold JR. 1993. Complete sequence of the mitochondrial cytochrome *b* gene in the Cherryfin Shinner, *Liturus roseipinnis* (Teleostei: Cyprinidae). *Copeia* 3: 880-883.

- Simmons MP & Ochoterena H. 2000. Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* 49: 369-381.
- Swofford DL. 1998. PAUP® - phylogenetic analysis using parsimony (and other methods) version 4.0. Sinauer Associates, Sunderland.
- Ünver B & Erk'akan F. 2005. A natural hybrid of *Leuciscus cephalus* (L.) and *Chalcalburnus chalcoides* (Güldenstädt) (Osteichthyes-Cyprinidae) from Lake Tödürge (Sivas, Turkey). *Journal of Fish Biology* 66: 899-910.
- Vrijenhoek RC. 1998. Clonal organisms and the benefits of sex. In Carvalho GR (Ed.) *Advances in Molecular Ecology*. IOS Press, Amsterdam, pp 151-172.
- Yang Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution*. 39: 306-314.
- Young ND & Healy J. 2003. GapCoder automates the use of indel characters in phylogenetic analysis. *BMC Bioinformatics* 4: 6.
- Zardoya R & Doadrio I. 1998. Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. *Proceedings of the Royal Society B* 265: 1365-1372.
- Zardoya R & Doadrio I. 1999. Molecular evidence on the evolutionary and biogeographical patterns of European Cyprinids. *Journal of Molecular Evolution* 49: 227-237.

Chapter 8

Reading the history of a hybrid fish complex from its molecular record

Sousa-Santos C, Collares-Pereira MJ & Almada V (in press)

Molecular Phylogenetics and Evolution

Reading the history of a hybrid fish complex from its molecular record

Sousa-Santos C, Collares-Pereira MJ & Almada V

Abstract *Squalius alburnoides* is a widely distributed intergeneric hybrid complex with fish of both sexes, varying ploidy levels and proportions of the parental genomes. Its dispersal routes were here delineated and framed by the reconstruction of the phylogeny and phylogeography of other *Squalius* with which it hybridizes, based on the available data on the paleohydrographical history of the Iberian Peninsula. Results based on sequences of cytochrome *b* and beta-actin genes showed that: proto-*Squalius pyrenaicus* originated at least five species as it dispersed throughout the Iberian Peninsula in the Mio-Pliocene; the *S. alburnoides* complex likely had a single origin in the bulk of Iberia, in the Upper Tagus/Guadiana area, when hydrographical rearrangements allowed the contact between its ancestors (around 700,000 years ago); interspecific crosses allowed the introgression of mitochondrial and nuclear genes of *S. alburnoides* in allopatric species/populations of other *Squalius* and vice-versa; and reconstituted *S. alburnoides* non-hybrid males may contribute to the replacement of the typical mtDNA of the complex (in the populations where they occur, crosses with females of other *Squalius* seem to have been especially frequent). A number of dispersal events and colonization routes are proposed.

Keywords *Squalius alburnoides*; Hybrid complex; Introgression; Paleobiogeography; Iberia

Introduction

S. alburnoides is a curious example of a very successful intergeneric hybrid complex of cyprinid fishes endemic to the Iberian Peninsula. Allozymes (Alves *et al.* 1997a; Carmona *et al.* 1997), microsatellites (Crespo-López *et al.* 2006), beta-actin gene sequencing (Robalo *et al.* 2006) and cytogenetic studies (Gromicho *et al.* 2006) revealed that the crosses that originated *S. alburnoides* involved *S. pyrenaicus* females (P-haplotype) and a presumably extinct *Anaecypris*-like paternal ancestor (A-haplotype). As a result of this hybridization process, the bisexual reproductive mode was disturbed, originating new patterns of

gamete production which generated a diverse array of ploidy levels and genomic constitutions (see Table S1 - electronic supplementary material).

As regards primary freshwater fish, the Iberian Peninsula is almost an island which had very limited historical connections with the rest of Europe (Banarescu 1973; Almaça 1978). The first known Iberian cyprinid fossil, *Rutilus antiquus*, was detected in the eastern margin of the Peninsula, in the region of the Ebro river basin, and was dated from the Upper Oligocene (Cabrera & Gaudant 1985; De la Peña 1995). At this time, the uplift of the Pyrenean Mountains was still an active process that culminated only in the Late Pliocene (Andeweg 2002), thus, the colonization of the Peninsula in the Oligocene by cyprinids (including the ancestors of *S. alburnoides*) likely occurred through a freshwater passage from south France to the Ebro river basin. For an unlikely alternative (Doadrio & Carmona 2003) involving wide migrations across the Mediterranean during the Messinian salinity crisis see Bianco (1990). Also, although Iberia and North Africa were in contact for a short period in the Upper Miocene (Andeweg 2002), the extreme scarcity of Leuciscinae in Africa (Froese & Pauly 2007) also makes this African alternative very unlikely. For studies on the phylogeny of Iberian cyprinids see Zardoya & Doadrio (1998), Zardoya & Doadrio (1999) and Cunha *et al.* (2002).

Although the identity of the ancestors of the complex appears to be a clarified issue, the number of original hybridization events has been a matter of controversy. Indeed, the analysis of the mitochondrial DNA (mtDNA) diversity of *S. alburnoides* populations led some authors to postulate multiple independent origins for the complex: Alves *et al.* (1997b) postulated two independent origins in the Tagus-Guadiana and in the Sado basins and Cunha *et al.* (2004) postulated three additional ones: in Guadiana-Guadalquivir basins, in Douro and in the Quarteira river (see Figure 1 for river locations).

Along its wide distribution range, *S. alburnoides* is currently sympatric with at least three other *Squalius* species, *S. pyrenaicus*, *S. carolitertii* and *S. aradensis* (Figure 1), whose mitochondrial and nuclear genes are found in the complex (Alves *et al.* 1997b; Cunha *et al.* 2004; Sousa-Santos *et al.* 2006a). To discuss whether the origin of the *S. alburnoides* complex was a unique event or if it occurred independently in more than one river basin, it is crucial to evaluate the existence of past and present interspecific gene flow between *S. alburnoides* and the other three sympatric *Squalius* species. Moreover, the interpretation of the phylogeographic patterns of *S. alburnoides* has to be framed by the

patterns exhibited by the other *Squalius* species with which the complex hybridizes and corroborated by the paleohydrographical history of Iberia.

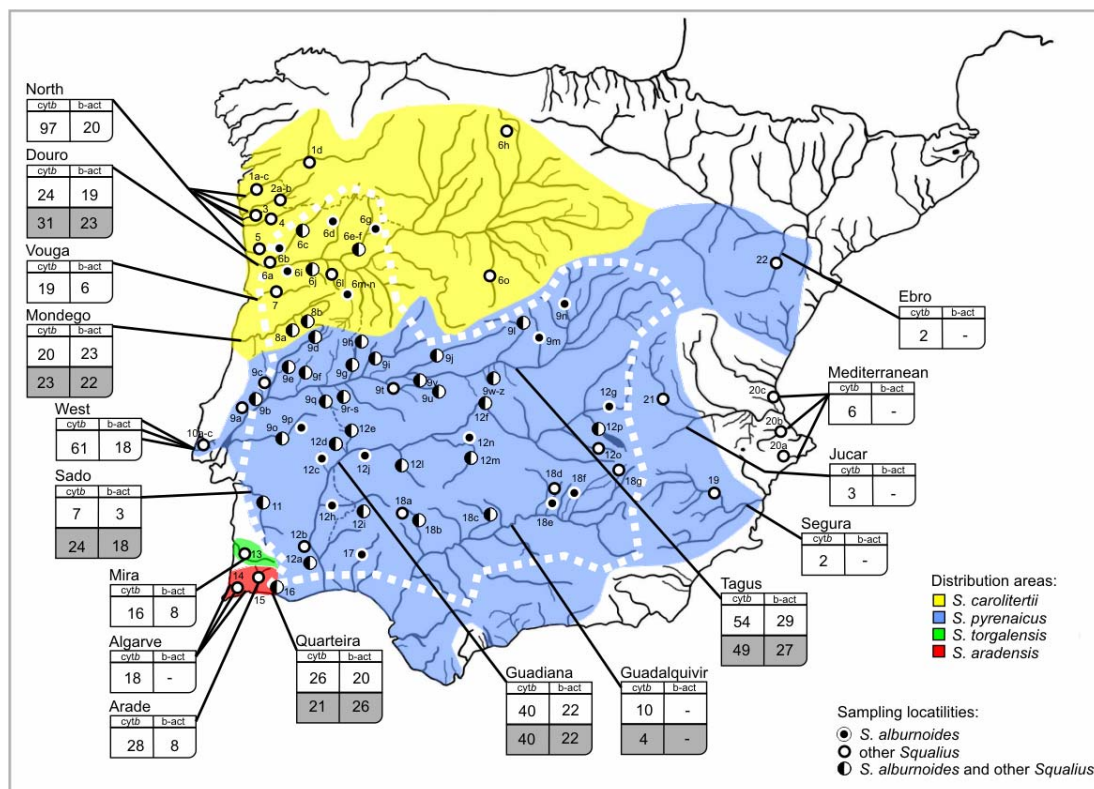


Figure 1 — Distribution areas of the *Squalius* species studied, sampling locations and number of individuals, from each river basin, sequenced for *cytb* and beta-actin genes. The distribution area of *S. alburnoides* is represented by a white broken line, overlapping the distribution areas of *S. carolitertii* (in yellow), *S. pyrenaicus* (in blue). *S. torgalensis* (in green) and *S. aradensis* (in red). Sampling locations are represented by different signs (legended in the figure) according to the species captured in each location. Each river basin has a box associated in which the number of *S. alburnoides* (grey half of the box) and the number of other *Squalius* (white half of the box) sequenced for *cytb* and beta-actin genes are summarized. Legend: 1 — Minho (1a-Coura, 1b-Mouro, 1c-Tea, 1d-Vivey); 2 — Lima (2a-Salas, 2b-Vez); 3 — Neiva; 4 — Cávado; 5 — Ave; 6 — Douro (6a-Minas, 6b-Sousa, 6c-Tâmega; 6d-Calvo, 6e-Sabor, 6f-Azibo, 6g-Maçãs, 6h-Boedo, 6i-Arda, 6j-Paiva, 6l-Távora, 6m-Coa, 6n-Águeda, 6o-Adaja); 7 — Vouga; 8 — Mondego (8a-Alva, 8b-Ceira); 9 — Tagus [9a-Maior, 9b-Alviela, 9c-Nabão, 9d-Zêzere, 9e-Sertã, 9f-Ocreza, 9g-Erges, 9h-Trejejana, 9i-Alagon tributaries (Acebo, Arrago, Jerte, Caparro), 9j-Tiétar, 9l-Cofio, 9m-Guadarrama, 9n-Jarama, 9o-Sorraia, 9p-Seda, 9q-Sever, 9r-Alburrel, 9s-Vid, 9t-Pesquero, 9u-Almonte, 9v-Aurela, 9w-Huso, 9x-Gebalo, 9z-Cedena]; 10 — Western rivers (10a-Lizandro, 10b-Samarra, 10c-Colares); 11 — Sado; 12 — Guadiana (12a-Vascão, 12b-Oeiras, 12c-Degebe, 12d-Caia, 12e-Xévora, 12f-Estena, 12g-Zancara, 12h-Ardila, 12i-Sillo, 12j-Albuera, 12l-Matachel, 12m-Zujar, 12n-Quejigares, 12o-Azuer, 12p-Ruidera); 13 — Mira; 14 — Algarve rivers (Aljezur, Seixe and Alvor); 15 — Arade; 16 — Quarteira; 17 — Odiel; 18 — Guadalquivir (18a-Mollinos, 18b-Montemayor, 18c-Manzano, 18d-Robledillo, 18e-Jandula, 18f-Guadiel, 18g-Guadalupe); 19 — Segura; 20 — Mediterranean rivers (20a-Algar, 20b-Serpis, 20c-Valencia Lagoon); 21 — Jucar; 22 — Ebro.

During the Miocene most river systems, instead of flowing to the sea, drained to a large number of inland lakes, some of which persisted well into the Pliocene (Friend & Dabrio 1996; Andeweg 2002). The current exorheic river network is very recent, being of Plio-Pleistocene origin (Andeweg 2002). Thus, hypotheses drawn to explain the dispersal

and foundation of new *Squalius* populations have to be concordant with the available geological data concerning the definition of the respective fluvial networks.

The aims of this study were: 1) to draw plausible dispersal routes for the Iberian *Squalius*; 2) to discuss possible locations for the origin of the *S. alburnoides* complex; 3) to postulate the putative dispersal routes of the *S. alburnoides* complex that are compatible with geological data and with the distribution of other *Squalius* species; and 4) to evaluate the relationships between *S. alburnoides* and other *Squalius* species and their implications to the genetic dynamics of the complex.

The opportunity to reconstruct the history of this complex was made possible because although recombination has been demonstrated to occur among similar genetic complements (*e.g.* among two A or two P chromosome sets – Alves *et al.* 2004; Crespo-López *et al.* 2006), no evidence of recombination among dissimilar genetic complements (A and P) was yet found. This absence of recombination between the A and P genomes means that by the combined use of a mitochondrial and a nuclear marker it is possible to infer the parentage of each form of the complex. In this paper, fragments of the cytochrome *b* and the beta-actin genes were used to achieve this goal.

As it was possible to obtain samples covering almost all the known distribution area of *S. alburnoides*, this study allowed a deep insight into the evolutionary dynamics of an intergeneric hybrid complex. Indeed, the reconstruction of paleogeographic scenarios may be helpful for the evaluation of its evolutionary potential and capacity to adapt to distinct environments and interact with sympatric species.

Materials and methods

Brief presentation of the S. alburnoides complex

The *S. alburnoides* complex is known to occur presently in nine Iberian river basins (Douro, Vouga, Mondego, Tagus, Sado, Guadiana, Quarteira, Odiel and Guadalquivir) (Cabral *et al.* 2005; Ribeiro *et al.* 2007) (Figure 1). This complex comprises distinct forms whose frequency may differ significantly according to the river basin: the southern populations include diploids (PA), triploids (PAA and PPA), tetraploids (PPAA, PAAA and PPPA), and also

a non-hybrid form (AA), reconstituted from the hybrids and almost constituted by males (females appear to be extremely rare – Sousa-Santos *et al.* 2006c), that is absent from the northern populations where all the other hybrid forms may be found (reviewed in Alves *et al.* 2001). The designation of the nuclear genomes of the *S. alburnoides* hybrid forms includes a capital “A” (referring to the paternal ancestor of the complex), while the letter “P” used above to denote the genome of the still existing maternal ancestor *S. pyrenaicus*, must be replaced by “C” or “Q” when the sympatric *Squalius* are, respectively, *S. carolitertii* or *S. aradensis*. A synthesis on the distribution and frequency of the *S. alburnoides* forms, including the northern ones, may be found in Table S2 (electronic supplementary material).

The mtDNA found in *S. alburnoides* is usually *S. pyrenaicus*-like (the maternal ancestor of the complex) but some introgressions were already reported: one individual with *S. carolitertii*-like mtDNA in the Douro drainage (Alves *et al.* 1997b), and several with *S. aradensis*-like mtDNA in the Quarteira River (Sousa-Santos *et al.* 2006a).

Field work and laboratorial procedures

Sequences of specimens from the *S. alburnoides* complex and from five other *Squalius* species (*S. aradensis*, *S. carolitertii*, *S. pyrenaicus*, *S. torgalensis* and *S. valentinus*), covering a total of 27 river basins, were analysed for the cytochrome *b* (*cytb*) and beta-actin genes. Sampling locations and the respective number of individuals sequenced, in a total of 149 *S. alburnoides*, 143 *S. pyrenaicus*, 164 *S. carolitertii*, 42 *S. aradensis* and 18 *S. torgalensis* are depicted in Figure 1. In order to get the most complete coverage of the populations, some additional sequences of the *cytb* gene were retrieved from GenBank which, added to the sequences that resulted from this study, amounted to a total of 217 *S. alburnoides*, 214 *S. pyrenaicus*, 175 *S. carolitertii*, 75 *S. aradensis*, 18 *S. torgalensis* and 6 *S. valentinus*. GenBank accession numbers may be found in Table S3 (electronic supplementary material). Samples of *S. alburnoides* from River Vouga should have been included, however, despite some attempts in the main river and in the River Sul, no *S. alburnoides* specimens were captured.

Fish were electrofished, morphologically identified, and in general returned to the river after the removal of small fin clips. Total genomic DNA was extracted from fin clips preserved in ethanol by an SDS/proteinase-k based protocol, precipitated with isopropanol and washed with ethanol before re-suspension in water (adapted from

Sambrook *et al.* 1989). A total of 799bp of the *cytb* gene and of 935bp of the beta-actin gene were amplified. The primers and PCR conditions may be found in Sousa-Santos *et al.* (2005).

Data analysis

Sequences of the *cytb* gene were aligned with BioEdit v.5.0.6 and their phylogenetic relationships reconstructed with a maximum parsimony method using PAUP 4.0 (Swofford, 1998). The resultant phylogenetic tree was analysed to assess the existence of present and ancient crosses between *S. alburnoides* and other sympatric *Squalius* species in a given river basin. It was assumed that if *S. alburnoides* crossed only with conspecifics and/or with males of other *Squalius* species no haplotypes will be shared between *S. alburnoides* and other *Squalius* species. In contrast, the existence of derived haplotypes shared between *S. alburnoides* and other *Squalius* species would reflect the existence of interspecific crosses. Thus, when analysing the phylogenetic tree, a) terminal haplotypes shared between *S. alburnoides* and other *Squalius* species were assumed to be representatives of recent crosses; and b) missing common ancestors between *S. alburnoides* and other *Squalius* species located in the internal nodes of the tree were interpreted as indicative of ancient crosses. The mean number of mutational steps linking each pair of haplotypes to their common ancestor was used to construct a histogram reflecting the contacts through time between *S. alburnoides* and other *Squalius* species.

Concerning the beta-actin gene, the sequences of homozygous individuals were also aligned with BioEdit v.5.0.6. However, for heterozygous diploid and polyploid individuals the genome complements had to be recovered following the procedures described in Sousa-Santos *et al.* (2005) before the alignment process. Since the low differentiation of the beta-actin *Squalius* haplotypes does not allow the distinction of all the species validated by the mtDNA analysis, the nuclear genomes of *S. pyrenaicus*, *S. carolitertii*, *S. torgalensis* and *S. aradensis* were herein designated as P-haplotypes for simplicity reasons. The nuclear haplotypes derived from the paternal ancestor of the *S. alburnoides* complex were designated as A-haplotypes. In the presence of distinct P or A haplotypes in the same individual the designation P' or A' was used.

In previous studies with this marker in cyprinids, involving more than ten species and some hundreds of individuals (Robalo *et al.* 2006, 2007; Sousa-Santos *et al.* 2006a,

2006c, 2007), only a single locus was detected in diploids for the beta-actin gene, assuring that it is a single copy gene, thus excluding the risk of using paralogous sequences when analysing polyploids of hybrid origin.

Networks of mitochondrial and nuclear haplotypes were performed with Network 4.201 (www.fluxus-engineering.com), using a median-joining algorithm (Bandelt *et al.* 1999). Mean number of pairwise differences, diversity indices, AMOVA and pairwise comparisons of haplotype frequencies among populations were calculated with Arlequin 3.01 (Excoffier *et al.* 2005). When computing mean divergence among populations, the values were corrected to remove within population variation. To perform this correction, the mean divergence between all possible pairs of sequences having a member in each population was first computed and then the average divergence of all pairs of sequences involving members of the same population was subtracted, as implemented in Arlequin.

Estimations of the divergence time based on the *cytb* gene were calculated using an evolutionary rate of 1.05% sequence divergence per million years (MY), as suggested by Dowling *et al.* (2002) for North American cyprinids.

Results

MtDNA variation

From a total of 625 *cytb* gene sequences, 211 distinct haplotypes were identified: 39.81% belonging to *S. alburnoides*, 53.08% belonging to other *Squalius* species and 7.11% shared haplotypes between *S. alburnoides* and other *Squalius* species. A summary of the genetic structure of each population is summarized in Table 1.

The haplotype network showed a clear distinction of the populations belonging to the five *Squalius* species studied (Figure 2).

Table I – Sample size (N), number of haplotypes (N hap), haplotype diversity (HD), gene diversity (GD), nucleotide diversity (ND) and mean number of pairwise differences (MNPd) for *S. alburnoides* and other *Squalius* species from distinct drainages sequenced for the *cytb* gene.

River basins	N	N hap	HD (%)	GD $\pm$ sd	MNPd $\pm$ sd	ND $\pm$ sd
<i>S. alburnoides</i>						
Douro	31	10	32.26%	0.626 $\pm$ 0.100	11.755 $\pm$ 5.470	0.015 $\pm$ 0.008
Mondego	24	9	37.50%	0.775 $\pm$ 0.079	16.670 $\pm$ 7.690	0.021 $\pm$ 0.011
Tagus	49	40	81.63%	0.990 $\pm$ 0.007	7.051 $\pm$ 3.368	0.009 $\pm$ 0.005
Sado	24	10	41.67%	0.620 $\pm$ 0.117	3.967 $\pm$ 2.057	0.005 $\pm$ 0.003
Guadiana	40	27	67.50%	0.951 $\pm$ 0.022	5.362 $\pm$ 2.641	0.007 $\pm$ 0.004
Guadalquivir	4	4	100.00%	1.000 $\pm$ 0.177	7.667 $\pm$ 4.533	0.010 $\pm$ 0.007
Quarteira	21	6	28.57%	0.552 $\pm$ 0.122	14.219 $\pm$ 6.641	0.018 $\pm$ 0.009
<i>Other Squalius</i>						
Minho	22	1	4.55%	0.000 $\pm$ 0.000	0.000 $\pm$ 0.000	0.000 $\pm$ 0.000
Lima	22	4	18.18%	0.398 $\pm$ 0.122	0.429 $\pm$ 0.402	0.001 $\pm$ 0.001
Neiva	13	1	7.69%	0.000 $\pm$ 0.000	0.000 $\pm$ 0.000	0.000 $\pm$ 0.000
Cávado	20	2	10.00%	0.521 $\pm$ 0.042	0.521 $\pm$ 0.456	0.001 $\pm$ 0.001
Ave	20	2	10.00%	0.100 $\pm$ 0.088	0.100 $\pm$ 0.178	0.000 $\pm$ 0.000
Douro	24	7	29.17%	0.779 $\pm$ 0.057	1.667 $\pm$ 1.014	0.002 $\pm$ 0.001
Vouga	19	2	10.53%	0.105 $\pm$ 0.092	0.105 $\pm$ 0.183	0.000 $\pm$ 0.000
Mondego	20	8	40.00%	0.847 $\pm$ 0.051	17.584 $\pm$ 8.156	0.022 $\pm$ 0.011
Lizandro	21	8	38.10%	0.791 $\pm$ 0.076	2.238 $\pm$ 1.284	0.003 $\pm$ 0.002
Samarra	22	2	9.09%	0.091 $\pm$ 0.081	0.182 $\pm$ 0.245	0.000 $\pm$ 0.000
Colares	18	4	22.22%	0.529 $\pm$ 0.117	0.699 $\pm$ 0.553	0.001 $\pm$ 0.001
Tagus	54	36	66.67%	0.980 $\pm$ 0.008	15.881 $\pm$ 7.195	0.020 $\pm$ 0.010
Guadiana	40	24	60.00%	0.942 $\pm$ 0.027	4.165 $\pm$ 2.115	0.005 $\pm$ 0.003
Mira	16	3	18.75%	0.242 $\pm$ 0.135	0.250 $\pm$ 0.297	0.000 $\pm$ 0.000
Arade	28	9	32.14%	0.833 $\pm$ 0.050	2.611 $\pm$ 1.439	0.003 $\pm$ 0.002
Quarteira	26	3	11.54%	0.151 $\pm$ 0.093	0.2310 $\pm$ 7.779	0.000 $\pm$ 0.011
Sado	7	4	57.14%	0.810 $\pm$ 0.130	1.238 $\pm$ 0.885	0.002 $\pm$ 0.001
Guadalquivir	10	9	90.00%	0.978 $\pm$ 0.054	7.089 $\pm$ 3.636	0.009 $\pm$ 0.005
Alvor	6	1	16.67%	0.000 $\pm$ 0.000	0.000 $\pm$ 0.000	0.000 $\pm$ 0.000
Aljezur	6	2	33.33%	0.533 $\pm$ 0.172	0.533 $\pm$ 0.508	0.001 $\pm$ 0.001
Seixe	6	2	33.33%	0.333 $\pm$ 0.215	0.333 $\pm$ 0.380	0.000 $\pm$ 0.001
Segura	2	2	100.00%	1.000 $\pm$ 0.500	2.000 $\pm$ 1.732	0.003 $\pm$ 0.003
Algar	2	2	100.00%	1.000 $\pm$ 0.500	2.000 $\pm$ 1.732	0.003 $\pm$ 0.003
Serpis	2	1	50.00%	0.000 $\pm$ 0.000	0.000 $\pm$ 0.000	0.000 $\pm$ 0.000
Valencia Lagoon	2	2	100.00%	1.000 $\pm$ 0.500	1.000 $\pm$ 1.000	0.001 $\pm$ 0.002
Júcar	3	2	66.67%	0.667 $\pm$ 0.314	0.667 $\pm$ 0.667	0.001 $\pm$ 0.001
Ebro	2	1	50.00%	0.000 $\pm$ 0.000	0.000 $\pm$ 0.000	0.000 $\pm$ 0.000

The *S. pyrenaicus* sub-network was the most diverse, with 155 different haplotypes that ranged in their level of divergence from one to 22 mutations. The haplotypes from the River Tagus occupied a central position within this sub-network, from where other *S. pyrenaicus* populations from Guadiana, Guadalquivir and Sado river basins, and *S. valentinus* branched.

Within the *S. carolitertii* sub-network, three groups of haplotypes were clearly identified, differing from one to 31 mutations: haplotypes that were found only in River Zêzere; haplotypes that were found only in River Mondego; and a third group of haplotypes belonging to northern rivers from Minho to Mondego. *S. aradensis* and *S. torgalensis* haplotypes also formed well defined sub-networks. Globally, eight phylogroups of *Squalius* were identified: *S. pyrenaicus*-Tagus/Guadiana, *S. pyrenaicus*-Sado, *S. valentinus*, *S. carolitertii*-North, *S. carolitertii*-Mondego, *S. carolitertii*-Zêzere, *S. aradensis* and *S. torgalensis*. The mean percentage of divergence between the described phylogroups is presented in Figure 2.

In general, the *Squalius* individuals that were not *S. alburnoides* generally presented the expected mtDNA considering their geographical provenience, with only few exceptions: four *S. pyrenaicus* from Guadiana with Tagus-like mtDNA; five *S. pyrenaicus* in Mondego (where the expected species should be *S. carolitertii*); and one *S. pyrenaicus* from Lizandro with Guadiana-like mtDNA.

To allow the analysis of the gene flow between distinct river basins, only the populations of *S. alburnoides* bearing *S. pyrenaicus*-mtDNA were considered. Other *Squalius* harbouring mtDNA of other sympatric *Squalius* were excluded because they represent introgressions that would introduce artefacts in the estimation of divergence times.

The calculated divergence values between pairs of populations ranged between 0% and 1.15% (Table 2). Also shown in Table 2 are the divergence times between populations of other *Squalius* species analysed in this study.

Nuclear DNA variation

The genomic constitution of 138 *S. alburnoides* and of 176 individuals belonging to the other *Squalius* species is presented in Table S3 (electronic supplementary material). Concerning the later, although the majority of the individuals were homozygous for the beta-actin gene, considerably high numbers of heterozygous individuals (PP') were found (40.34%).

Figure 2 — Network of mtDNA haplotypes. The global network representing eight phylogroups (SCN—*S. caroliretii*-North, SCZ—*S. caroliretii*-Zêzere, SCM—*S. caroliretii*-Mondego, SPS—*S. pyrenaicus*-Sado, SPT/G—*S. pyrenaicus*-Tagus/Guadiana, SV—*S. valentinus*, SA—*S. aradenis* and ST—*S. torgalensis*), with the respective mean percentage of pairwise differences to the closest phylogroups, is depicted in a central position. On both sides of this central diagram the network of haplotypes of each phylogroup is depicted, with information concerning the species from where the haplotypes were retrieved and their respective river basin: haplotypes that were exclusive of *S. alburnoides* are represented by triangles; haplotypes that were exclusive of other *Squalius* are represented by circles; and haplotypes shared between *S. alburnoides* and other *Squalius* species are represented by squares. Each haplotype is identified with a numerical code (the same that was used in Figure S1 and Table S3 — both in electronic supplementary material) and its geographic provenience is represented by different colour patterns (see legend in the Figure). The number of mutations between haplotypes is represented by the number of small black dots placed on the branches linking haplotypes.

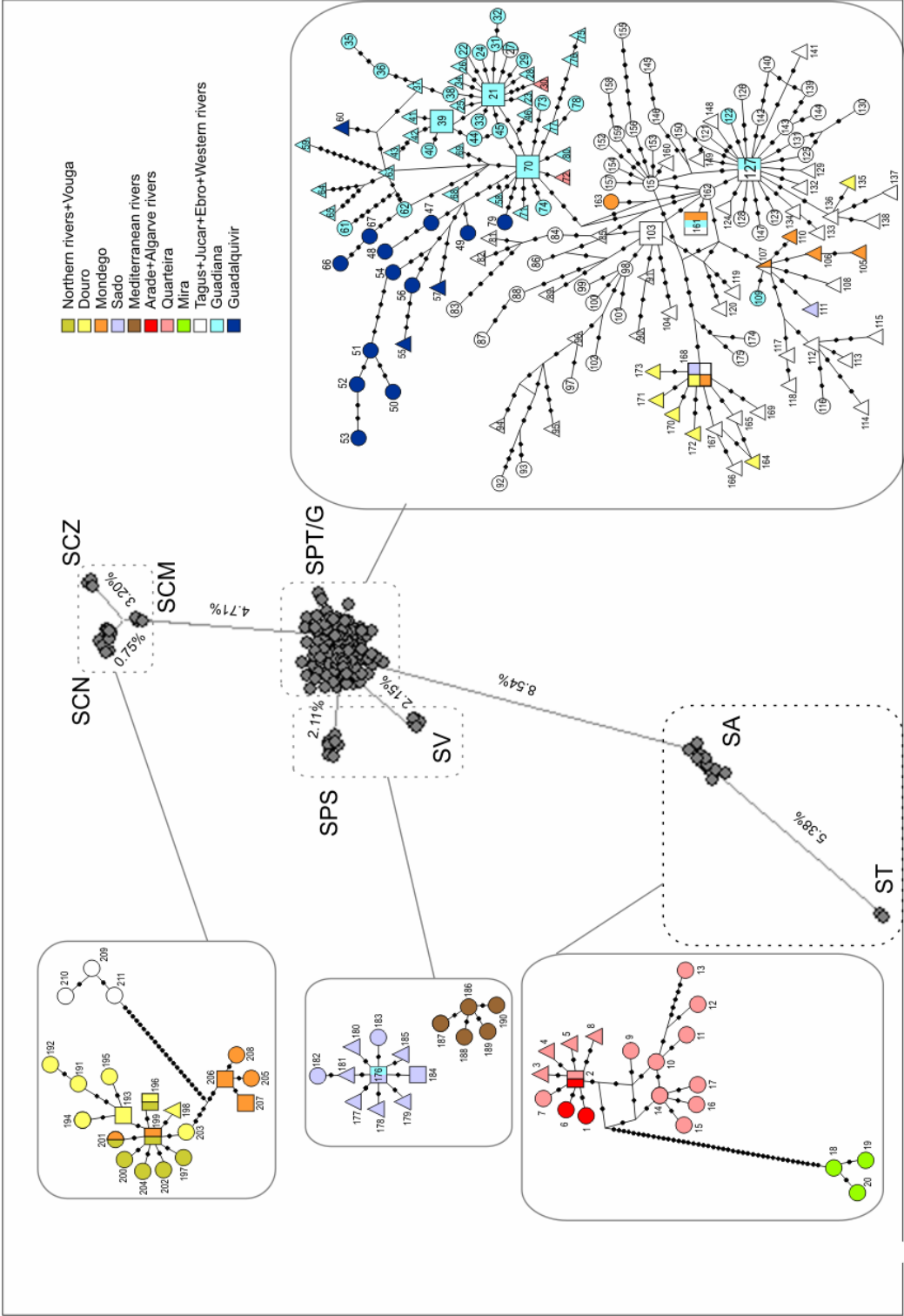


Table 2 — Mean percentage of divergence (below diagonal) between the *S. alburnoides* and other *Squalius* populations, intrapopulation divergence percentage values (in the diagonal) and estimated divergence times between populations in MY (above diagonal). To infer more accurately the colonization routes followed by the complex, samples of the major drainages (Tagus and Guadiana) were grouped according to their geographic provenience: tributaries of the western left margin (WLM), eastern left margin (ELM), western right margin (WRM) and eastern right margin (ERM). Additionally, a fifth sub-group of the Tagus River containing the samples of the Alagon River (ALAG) was created since special haplotypes were previously reported to this tributary (Cunha *et al.* 2004). Concerning the river Mondego, besides the population of *S. alburnoides*, calculations were made independently for the groups of samples belonging to three distinct phylogroups found in this drainage: *S. pyrenaicus* (Mondego a), *S. carolitertii*-North (Mondego b) and *S. carolitertii*-Mondego (Mondego c).

<i>S. alburnoides</i>	Tagus ALAG	Tagus ERM	Tagus ELM	Tagus WRM	Tagus WLM	Guadiana WRM	Guadiana WLM	Guadiana ERM	Guadiana ELM	Guadalquivir	Quarteira	Douro	Mondego	Sado
Tagus ALAG	0,28	0,77	0,77	0,72	0,76	0,85	0,89	0,80	0,99	1,09	0,85	0,01	0,69	0,01
Tagus ERM	0,80	0,60	< 0	0,16	0,21	0,78	0,81	0,75	0,93	1,01	0,74	0,70	0,30	0,17
Tagus ELM	0,81	-0,01	0,81	0,07	0,15	0,69	0,73	0,68	0,83	0,92	0,64	0,71	0,22	0,13
Tagus WRM	0,75	0,17	0,07	0,81	0,09	0,74	0,78	0,75	0,94	1,01	0,70	0,66	0,05	< 0
Tagus WLM	0,80	0,22	0,16	0,09	0,62	0,71	0,75	0,69	0,88	0,97	0,67	0,70	0,23	0,11
Guadiana WRM	0,89	0,82	0,73	0,78	0,75	0,43	0,00	0,02	0,23	0,28	< 0	0,81	0,84	0,50
Guadiana WLM	0,94	0,86	0,76	0,82	0,78	0,00	0,43	< 0	0,26	0,30	< 0	0,86	0,88	0,55
Guadiana ERM	0,84	0,79	0,71	0,78	0,73	0,02	-0,01	0,92	0,10	0,18	< 0	0,78	0,83	0,49
Guadiana ELM	1,04	0,98	0,87	0,99	0,93	0,24	0,27	0,10	1,75	0,05	0,24	1,03	1,01	0,72
Guadalquivir	1,15	1,06	0,96	1,06	1,01	0,29	0,32	0,19	0,05	0,96	0,29	1,11	1,11	0,79
Quarteira	0,89	0,78	0,67	0,74	0,70	-0,06	-0,06	-0,02	0,25	0,30	0,75	0,80	0,79	0,48
Douro	0,01	0,73	0,74	0,69	0,73	0,85	0,90	0,82	1,09	1,17	0,84	0,18	0,63	< 0
Mondego	0,72	0,31	0,23	0,05	0,24	0,88	0,92	0,87	1,06	1,17	0,83	0,66	0,30	< 0
Sado	0,01	0,18	0,13	-0,03	0,12	0,53	0,57	0,51	0,75	0,83	0,50	-0,02	-0,08	1,13

<i>S. pyrenaicus</i>	Tagus ALAG	Tagus ERM	Tagus ELM	Tagus WRM	Tagus WLM	Guadiana WRM	Guadiana WLM	Guadiana ERM	Guadiana ELM	Guadalquivir	Samarra	Lizandro	Colares	Mondego a
Tagus ALAG	0,68	0,43	0,16	0,11	0,15	0,66	1,10	0,81	0,85	1,15	0,35	0,20	0,22	0,20
Tagus ERM	0,45	0,67	0,16	0,34	0,18	0,81	1,23	0,95	1,02	1,15	0,52	0,35	0,39	0,33
Tagus ELM	0,17	0,17	0,87	0,12	0,03	0,57	0,98	0,66	0,74	0,98	0,33	0,16	0,20	0,13
Tagus WRM	0,12	0,36	0,87	0,61	0,03	0,53	0,95	0,66	0,73	0,99	0,19	0,04	0,06	0,06
Tagus WLM	0,16	0,19	0,03	0,03	0,63	0,53	0,97	0,68	0,75	0,99	0,20	0,05	0,08	0,02
Guadiana WRM	0,70	0,85	0,60	0,55	0,56	0,53	0,49	0,07	0,02	0,42	0,73	0,49	0,62	0,45
Guadiana WLM	1,16	1,29	1,03	1,00	1,01	0,51	0,00	0,43	0,46	0,76	1,18	0,95	1,06	0,86
Guadiana ERM	0,85	1,00	0,70	0,69	0,72	0,07	0,64	0,38	0,05	0,39	0,89	0,65	0,78	0,56
Guadiana ELM	0,90	1,07	0,78	0,77	0,79	0,02	0,70	0,06	0,45	0,41	0,96	0,70	0,85	0,64
Guadalquivir	1,20	1,20	1,03	1,04	1,04	0,44	0,80	0,41	0,43	0,89	1,22	0,97	1,11	0,90
Samarra	0,72	0,89	0,80	0,51	0,54	1,05	1,25	1,14	1,24	1,73	0,02	0,13	0,12	0,29
Lizandro	0,69	0,85	0,75	0,49	0,51	0,92	1,14	1,01	1,10	1,60	0,14	0,28	0,02	0,12
Colares	0,62	0,79	0,69	0,42	0,44	0,96	1,16	1,05	1,16	1,65	0,13	0,02	0,09	0,16
Mondego a	0,78	0,90	0,80	0,59	0,56	0,97	1,13	1,00	1,12	1,61	0,54	0,49	0,44	0,45
Sado	2,30	2,51	2,31	2,22	2,37	2,15	2,64	2,23	2,22	2,96	2,62	2,33	2,51	2,29
Jucar	0,41	0,21	0,14	0,32	0,11	0,81	1,25	0,95	1,03	1,30	0,50	0,33	0,37	0,30
Ebro	0,41	0,21	0,14	0,32	0,11	0,81	1,25	0,95	1,03	1,30	0,50	0,33	0,37	0,30
Segura	1,28	1,25	1,13	1,13	1,14	0,66	0,88	0,58	0,65	0,32	1,37	1,12	1,24	1,03

<i>S. valentinus</i>	Algar	2,54	2,67	2,48	2,29	2,39	2,32	2,63	2,29	2,36	2,45	2,62	2,35	2,49	2,23
Serpis	2,66	2,80	2,60	2,41	2,52	2,34	2,63	2,29	2,36	2,45	2,74	2,47	2,62	2,35	
Valencia lagoon	2,54	2,67	2,48	2,29	2,39	2,32	2,63	2,29	2,36	2,45	2,62	2,35	2,49	2,23	

<i>S. carolitertii</i>	Minho	5,84	5,92	5,69	5,71	5,76	5,83	5,88	5,78	5,81	6,10	5,89	5,73	5,79	5,56
Lima	5,86	5,95	5,71	5,73	5,78	5,85	5,91	5,80	5,83	6,12	5,92	5,76	5,81	5,58	
Neiva	5,84	5,92	5,69	5,71	5,76	5,83	5,88	5,78	5,81	6,10	5,89	5,73	5,79	5,56	
Cávado	5,78	5,87	5,63	5,66	5,70	5,77	5,83	5,72	5,75	6,04	5,84	5,68	5,74	5,50	
Ave	5,84	5,93	5,70	5,72	5,76	5,84	5,89	5,78	5,81	6,10	5,90	5,74	5,80	5,56	
Douro	5,93	6,01	5,78	5,80	5,85	5,86	5,97	5,83	5,84	6,12	5,98	5,81	5,88	5,65	
Vouga	5,85	5,93	5,70	5,72	5,76	5,84	5,89	5,78	5,81	6,10	5,90	5,74	5,80	5,56	
Mondego b	5,88	6,05	5,81	5,84	5,88	5,96	6,01	5,90	5,93	6,22	6,02	5,86	5,92	5,68	
Mondego c	5,27	5,36	5,12	5,19	5,19	5,26	5,31	5,21	5,24	5,58	5,33	5,17	5,22	4,71	
Tagus Zêzere	6,07	6,41	6,11	5,96	6,09	5,87	5,86	5,90	5,84	5,90	6,06	5,74	5,93	5,57	

<i>S. aradensis</i>	Quarteira	8,84	8,77	8,57	8,72	8,67	9,16	9,08	9,19	9,18	10,77	9,23	8,99	9,04	8,77
Arade	8,50	8,60	8,34	8,37	8,39	8,83	8,75	8,86	8,85	9,55	8,88	8,65	8,69	8,44	
Alvor	9,00	8,93	8,72	8,87	8,82	9,34	9,26	9,37	9,36	9,46	9,39	9,15	9,20	8,94	
Aljezur	9,01	8,94	8,73	8,88	8,83	9,35	9,27	9,38	9,37	9,46	9,40	9,16	9,21	8,94	
Seixe	8,46	8,64	8,36	8,33	8,38	8,80	8,72	8,83	8,82	8,91	8,84	8,61	8,65	8,39	

<i>S. torgalensis</i>	Mira	10,97	10,91	10,65	10,82	10,81	10,91	11,00	10,88	10,90	11,42	11,37	11,08	11,24	10,75
-----------------------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

(Table 2 – cont.)

	Sado	Jucar	Ebro	Segura	Algar	Serpis	Valencia lagoon	Minho	Lima	Neiva	Cavado	Ave	Douro	Vouga	Mondego b	Mondego c	Tagus Zézere	Quarteira	Arade	Alvor	Aljezur	Seixe	Mira
<i>S. pyrenaicus</i>																							
Tagus ALAG	2.19	0.39	0.39	1.22	2.42	2.53	2.42	5.24	5.23	5.24	5.15	5.23	5.22	5.24	5.28	4.65	5.41	8.42	8.10	8.57	8.58	8.06	10.45
Tagus ERM	2.39	0.20	0.20	1.19	2.54	2.66	2.54	5.32	5.32	5.32	5.24	5.32	5.31	5.32	5.44	4.73	5.74	8.35	8.19	8.50	8.51	8.22	10.39
Tagus ELM	2.20	0.14	0.14	1.08	2.36	2.48	2.36	5.00	5.00	5.00	4.92	5.00	4.99	5.00	5.12	4.41	5.36	8.16	7.95	8.31	8.31	7.96	10.14
Tagus WRM	2.11	0.30	0.30	1.07	2.18	2.30	2.18	5.15	5.14	5.15	5.06	5.15	5.13	5.15	5.27	4.60	5.33	8.30	7.97	8.45	8.46	7.93	10.31
Tagus WLM	2.26	0.11	0.11	1.08	2.28	2.40	2.28	5.19	5.18	5.19	5.10	5.19	5.17	5.19	5.30	4.60	5.46	8.26	7.99	8.40	8.40	7.98	10.30
Guadiana WRM	2.05	0.77	0.77	0.63	2.21	2.23	2.21	5.30	5.29	5.30	5.21	5.30	5.23	5.30	5.42	4.71	5.29	8.73	8.41	8.90	8.90	8.38	10.39
Guadiana WLM	2.51	1.19	1.19	0.83	2.50	2.50	2.50	5.60	5.60	5.60	5.52	5.60	5.59	5.60	5.72	5.01	5.54	8.65	8.33	8.82	8.83	8.30	10.47
Guadiana ERM	2.13	0.91	0.91	0.55	2.19	2.19	2.19	5.32	5.31	5.32	5.23	5.32	5.27	5.32	5.44	4.73	5.39	8.76	8.44	8.93	8.94	8.41	10.36
Guadiana ELM	2.11	0.98	0.98	0.62	2.24	2.24	2.24	5.32	5.31	5.32	5.23	5.32	5.25	5.32	5.44	4.73	5.30	8.75	8.43	8.92	8.93	8.40	10.38
Guadaluquivir	2.32	1.23	1.23	0.30	2.33	2.33	2.33	5.38	5.38	5.38	5.30	5.38	5.31	5.38	5.50	4.84	5.15	8.83	8.52	9.01	9.01	8.49	10.44
Samarra	2.49	0.48	0.48	1.30	2.49	2.61	2.49	5.60	5.60	5.60	5.52	5.60	5.58	5.60	5.72	5.01	5.77	8.79	8.46	8.94	8.95	8.42	10.83
Lizandro	2.22	0.32	0.32	1.07	2.24	2.35	2.24	5.33	5.32	5.33	5.24	5.33	5.30	5.33	5.45	4.74	5.47	8.56	8.24	8.72	8.72	8.20	10.55
Colares	2.39	0.35	0.35	1.18	2.38	2.49	2.38	5.47	5.47	5.47	5.47	5.39	5.47	5.46	5.47	5.59	4.88	8.61	8.28	8.76	8.77	8.24	10.70
Mondego a	2.18	0.29	0.29	0.98	2.12	2.24	2.12	5.08	5.07	5.08	4.99	5.08	5.06	5.08	5.20	4.49	5.30	8.36	8.04	8.51	8.52	7.99	10.24
Sado	0.15	2.51	2.51	2.63	3.11	3.23	3.11	5.94	5.93	5.94	5.96	5.94	5.84	5.94	6.06	5.38	6.62	9.35	9.25	9.53	9.54	9.25	11.17
Jucar	2.64	0.08	0.00	1.31	2.50	2.62	2.50	5.13	5.12	5.13	5.04	5.13	5.11	5.13	5.24	4.54	5.54	8.32	8.15	8.46	8.47	8.18	10.36
Ebro	2.64	0.00	0.00	1.31	2.50	2.62	2.50	5.13	5.12	5.13	5.04	5.13	5.11	5.13	5.24	4.54	5.54	8.32	8.15	8.46	8.47	8.18	10.36
Segura	2.77	1.38	1.38	0.25	2.26	2.38	2.26	5.48	5.48	5.48	5.48	5.48	5.47	5.48	5.60	4.89	5.66	8.76	8.45	8.94	8.95	8.42	10.36
<i>S. valentinus</i>																							
Algar	3.27	2.63	2.63	2.38	0.25	0.12	0.00	5.72	5.72	5.72	5.74	5.72	5.71	5.72	5.84	5.13	6.37	9.60	9.29	9.77	9.78	9.26	11.31
Serpis	3.39	2.75	2.75	2.50	0.13	0.00	0.12	5.84	5.84	5.84	5.86	5.84	5.83	5.84	5.96	5.25	6.49	9.72	9.41	9.89	9.90	9.38	11.31
Valencia lagoon	3.27	2.63	2.63	2.50	0.00	0.13	0.13	5.72	5.71	5.72	5.74	5.72	5.71	5.72	5.84	5.13	6.37	9.49	9.17	9.65	9.66	9.14	11.19
<i>S. carolitei</i>																							
Minho	6.31	5.38	5.38	5.76	6.01	6.13	6.01	0.00	0.00	0.00	0.02	0.00	0.03	0.00	0.12	0.73	3.28	8.22	8.25	8.22	8.23	8.30	9.40
Lima	6.33	5.38	5.38	5.75	6.00	6.13	6.00	0.00	0.05	0.00	0.02	0.00	0.03	0.00	0.12	0.73	3.24	8.20	8.23	8.20	8.21	8.28	9.38
Neiva	6.31	5.38	5.38	5.76	6.01	6.13	6.01	0.00	0.00	0.00	0.02	0.00	0.03	0.00	0.12	0.73	3.28	8.22	8.25	8.22	8.23	8.30	9.40
Cavado	6.37	5.29	5.29	5.67	6.03	6.16	6.03	0.02	0.03	0.02	0.07	0.02	0.05	0.02	0.14	0.75	3.19	8.13	8.16	8.14	8.15	8.22	9.32
Ave	6.32	5.38	5.38	5.76	6.01	6.13	6.01	0.00	0.00	0.00	0.02	0.01	0.03	0.00	0.11	0.73	3.28	8.21	8.25	8.22	8.23	8.30	9.40
Douro	6.32	5.37	5.37	5.74	5.99	6.12	5.99	0.03	0.03	0.03	0.06	0.03	0.21	0.03	0.15	0.67	3.14	8.24	8.28	8.25	8.26	8.33	9.34
Vouga	6.32	5.38	5.38	5.76	6.01	6.13	6.01	0.00	0.00	0.00	0.02	0.00	0.03	0.01	0.12	0.73	3.28	8.22	8.22	8.22	8.23	8.30	9.40
Mondego b	6.44	5.51	5.51	5.88	6.13	6.26	6.13	0.13	0.13	0.13	0.15	0.11	0.16	0.13	0.00	0.85	3.39	8.32	8.37	8.34	8.35	8.42	9.52
Mondego c	5.78	4.76	4.76	5.14	5.39	5.51	5.39	0.77	0.77	0.77	0.79	0.77	0.70	0.77	0.89	0.10	3.05	8.21	8.24	8.22	8.23	8.30	8.90
Tagus Zézere	7.08	5.82	5.82	5.94	6.69	6.82	6.69	3.44	3.40	3.44	3.35	3.44	3.29	3.44	3.56	3.20	0.10	9.36	9.00	9.47	9.48	8.96	10.05
<i>S. aradensis</i>																							
Quarteira	10.95	8.74	8.74	9.20	10.08	10.21	9.96	8.63	8.61	8.63	8.54	8.63	8.66	8.63	8.74	8.62	9.83	2.12	0.27	0.12	0.13	0.48	5.18
Arade	9.95	8.56	8.56	8.87	9.75	9.88	9.63	8.66	8.64	8.66	8.57	8.66	8.69	8.66	8.79	8.66	9.45	1.50	0.33	0.39	0.39	0.04	5.17
Alvor	10.01	8.89	8.89	9.39	10.26	10.39	10.14	8.64	8.61	8.64	8.55	8.64	8.67	8.64	8.76	8.63	9.95	0.13	0.40	0.00	0.01	0.60	5.35
Aljezur	10.01	8.89	8.89	9.40	10.27	10.40	10.15	8.64	8.62	8.64	8.56	8.64	8.68	8.64	8.77	8.64	9.96	0.13	0.41	0.01	0.07	0.60	5.36
Seixe	9.72	8.59	8.59	8.84	9.72	9.85	9.60	8.72	8.70	8.72	8.63	8.72	8.75	8.72	8.84	8.72	9.40	0.50	0.05	0.63	0.63	0.04	5.23
<i>S. torquensis</i>																							
Mira	11.82	10.87	10.87	10.87	11.87	11.87	11.75	9.87	9.85	9.87	9.78	9.87	9.81	9.87	10.00	9.35	10.56	5.44	5.61	5.62	5.62	5.49	0.03

From all the *S. alburnoides* individuals sampled, 23.19% were nuclear non-hybrids (AA or AA'), proceeding from the Tagus, Guadiana and Quarteira rivers. The remaining 76.81% were diploid or polyploid hybrids (13.04% PA, 50.00% PAA, 11.59% PPA, 1.45% PPAA and 0.72% PAAA) – Table S2 (electronic supplementary material). From a total of 314 specimens of all the *Squalius* sampled, 37 P-haplotypes and 14 A-haplotypes were identified.

The network of the A-haplotypes (Figure 3) showed that the ancestral haplotype (A4) was dispersed throughout the sampling area with the exception of the Douro river basin. In this drainage only one haplotype (A9) was found, shared with the Tagus and Mondego Rivers.

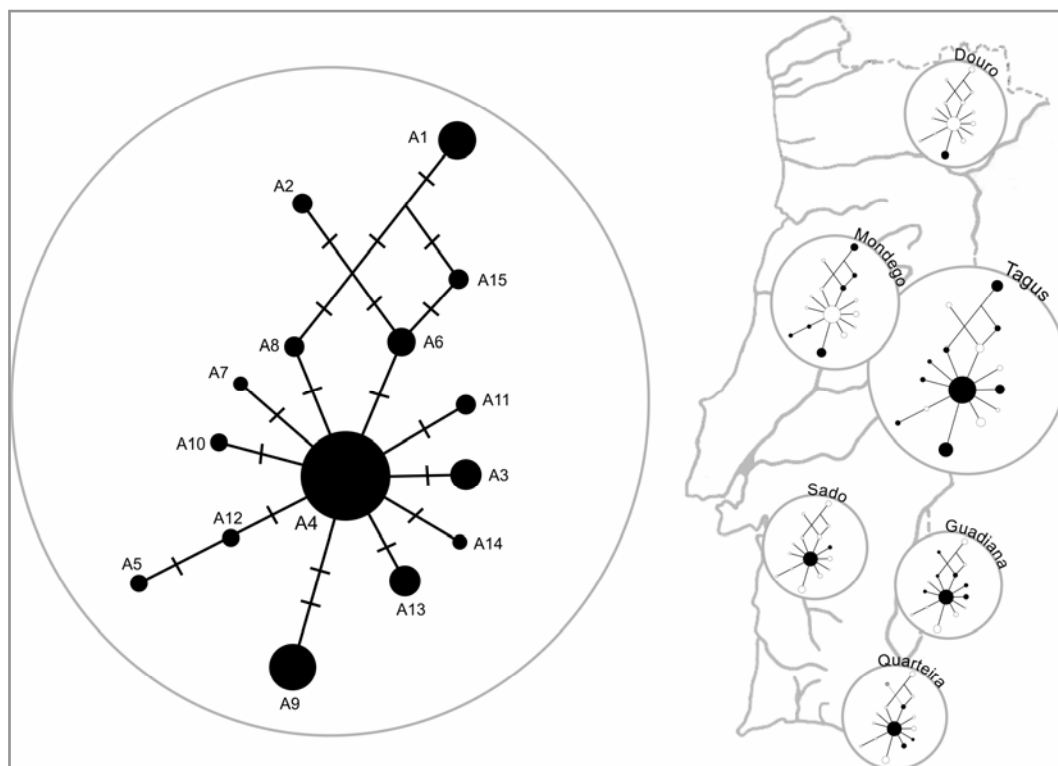


Figure 3 — Network of nuclear A-haplotypes and their respective occurrences in each of the populations/drainages. Haplotypes are represented by circles with diameters that are proportional to the number of individuals that shared each haplotype. Mutations between haplotypes are represented by the small lines perpendicular to the branch linking haplotypes. The haplotypes found in each population (map on the right) are depicted in black.

Mondego and Tagus specimens also shared three other haplotypes (A1, A15 and A5) and showed, respectively, one and three unique haplotypes. In the south, the Guadiana population showed one exclusive haplotype (A2) and four haplotypes shared with other river basins: Sado (A4 and A11), Quarteira (A4 and A6), Tagus (A4 and A3) and Mondego

(A6). The Quarteira population, besides the two haplotypes shared with Guadiana, also presented two exclusive haplotypes (A13 and A14).

Concerning the P-haplotypes (Figure 4), the ancestral haplotype (P3) was found in all the northern rivers from Minho to Mondego, in the Tagus and in the Guadiana. These last three river basins were the most diverse, with nine different haplotypes being found in Mondego and Guadiana, and ten in the Tagus.

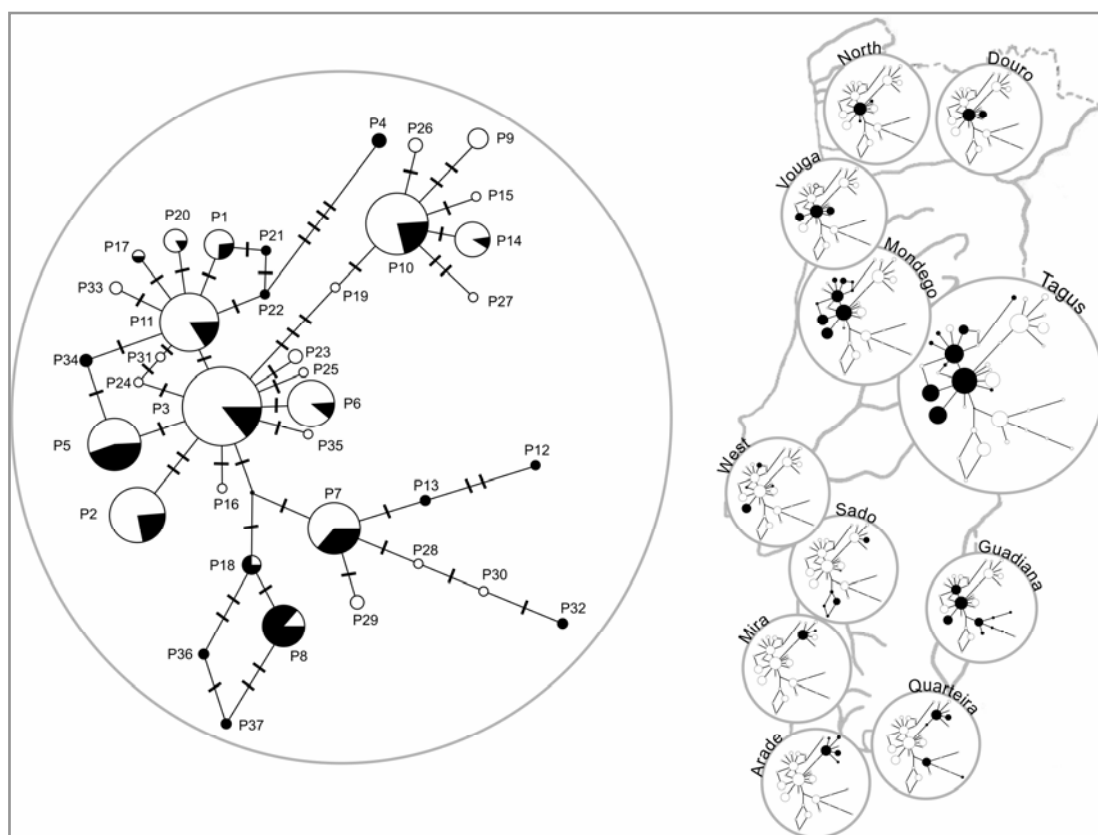


Figure 4 — Network of nuclear P-haplotypes, represented by circles with diameters that are proportional to the number of individuals that shared each haplotype. Each mutation between haplotypes is represented by a small line perpendicular to the branch linking haplotypes. Black and white circles were used to distinguish the haplotypes found, respectively, in *S. alburnoides* and in other *Squalius* species. When the haplotypes were shared between *S. alburnoides* and other *Squalius* species, black and white slices proportional to the respective number of individuals were depicted. The distribution of the haplotypes in the distinct populations is also depicted (haplotypes found in each drainage are black coloured).

The network depicted in Figure 4 showed that starting from the ancestral haplotype, three distinct southern lineages seem to have differentiated: one that includes haplotypes found in the southwestern rivers of Mira and Arade (P9, P10, P14, P15, P26 and P27); and two lineages linked by a missing common ancestor, one from the Sado (P8, P18,

P36 and P37) and the other from the Guadiana (P7, P12, P13, P28, P29 and P30). In the Sado, in addition to the above mentioned four exclusive haplotypes, one haplotype from the Arade lineage was detected in one individual (P14). The Quarteira population also presented a mixture of haplotypes: from the Arade (P10, P14 and P19) and from the Guadiana (P7 and P32) lineages.

In the north, the majority of the populations exhibited the ancient haplotypes P3 and P6 (only found in this region). The population of Vouga also presented haplotype P5 that was shared with Mondego and with the Erges tributary of Tagus. In the Mondego, four out of nine haplotypes were only found in this population (P20, P21, P22 and P34), the remaining being shared with the neighbour drainages of Vouga (P3 and P5) and Tagus (P1, P2, P3, P5 and P11).

The Tagus populations presented four exclusive haplotypes (P17, P31, P33 and P35) and haplotypes shared with northern rivers (P3 and P5), with Mondego (P1, P2, P3, P5 and P11), with the western rivers (P1 and P2) and with Guadiana (P2, P3 and P11). The haplotypes found in the western rivers of Lizandro, Samarra and Colares were either shared with the Tagus populations (P1 and P2) or derived by one mutational step from the ancestral haplotype P3 also present in the Tagus (P24 and P25).

Interspecific gene flow

The nuclear P-haplotypes found in *S. alburnoides* were the ones found in the sympatric *Squalius* of the considered river basin, with a few exceptions (Figure 4). These exceptions may represent new mutations, *Squalius* genomes brought by *S. alburnoides* that dispersed from other rivers, or simply insufficient sampling causing a failure to detect less common haplotypes in all *Squalius* of a given river basin.

In the case of Tagus, Guadiana and Guadalquivir drainages, the mtDNA haplotypes presented by *S. alburnoides* merged into the *S. pyrenaicus* sub-network. Some *S. alburnoides* haplotypes from Sado, Mondego, Douro and Quarteira river basins were included into the sub-network of the *Squalius* species that is sympatric with the complex in the same river basin (Figure 2). The extent of the introgressions of mtDNA of different phylogroups in the *S. alburnoides* complex is presented in Table 3.

Table 3 — Distribution of the mtDNA of the different phylogroups of *S. alburnoides*, as an indicator of the distinct levels of introgression by other species in the complex. For each river basin, the number (and percentage) of individuals with a given mtDNA type is indicated.

	Douro	Mondego	Tagus	Sado	Guadiana	Quarteira	Guadalquivir
Phylogroup	N=31	N=23	N=49	N=24	N=40	N=21	N=4
<i>S. pyrenaicus</i> Tagus/Guadiana	27 (87.1%)	18 (78.3%)	49 (100%)	2 (8.3%)	40 (100%)	2 (9.5%)	4 (100%)
<i>S. pyrenaicus</i> Sado	-	-	-	22 (91.7%)	-	-	-
<i>S. carolitertii</i> North	4 (12.9%)	1 (4.3%)	-	-	-	-	-
<i>S. carolitertii</i> Mondego	-	4 (17.4%)	-	-	-	-	-
<i>S. aradensis</i>	-	-	-	-	-	19 (90.5%)	-

The parsimony phylogenetic tree exhibited a topology that was similar to trees already published by other authors (Sanjur *et al.* 2003; Cunha *et al.* 2004; Doadrio & Carmona 2006) and, due to its extension it was herein presented as electronic supplementary material (Figure S1). However, in contrast to previous papers, the inclusion in this study of *S. alburnoides* from almost all its distribution area made possible the identification of crosses with members of other *Squalius* species that occurred at different time scales. As shown in Figure 5, although the majority of the interspecific crosses were recent (15 shared haplotypes), there was also a high frequency of past interspecific crosses. Indeed, the average branch lengths measured from the missing common ancestor to the terminal nodes of each branch, ranged from 0.62 to 10.39 mutational steps. Note that, as stated above, only the branches of the tree that contained *S. alburnoides* and other *Squalius* species as terminal nodes were included in this analysis.

These findings clearly showed that *S. alburnoides* did not exchange mtDNA with other *Squalius* species only one to five times, as suggested by previous authors. Indeed, accepting a different origin of the complex for each major basin like Cunha *et al.* (2004), we would expect one or a few old common ancestors at the basal nodes of the subtree

corresponding to that basin, with distinct evolutionary lines in *S. alburnoides* and other sympatric *Squalius* without shared haplotypes along the branches of the tree. On the contrary, the pattern presented in Figures 5 and S1 clearly showed that sharing of haplotypes took place at multiple occasions within each basin, although the crosses were not so massive as to blur the distinctiveness of *S. alburnoides* and the other *Squalius*.

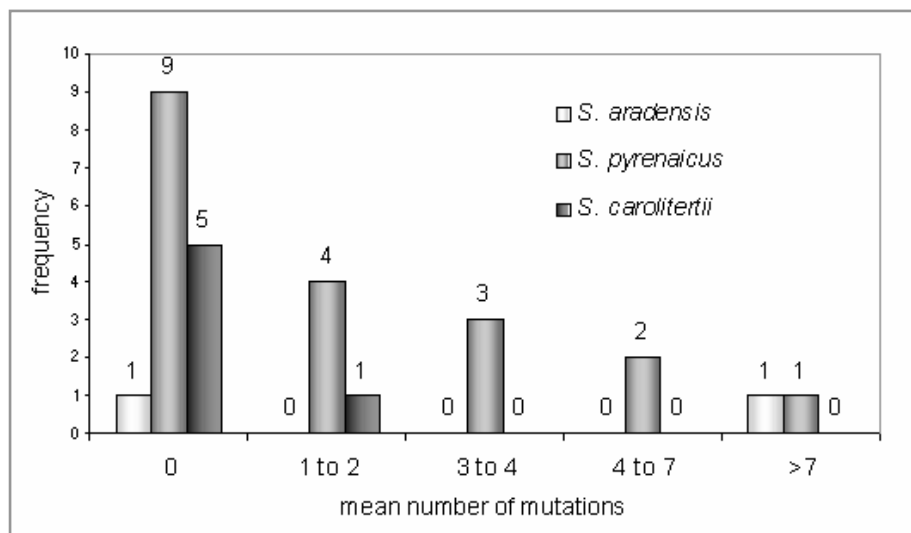


Figure 5 — Frequency of interspecific crosses, represented in the phylogenetic tree by shared haplotypes or missing common ancestors between *S. alburnoides* and other *Squalius* species, against a temporal line expressed by the mean number of mutations leading from the common ancestor to its terminal nodes.

For the river basins Tagus and Guadiana, there were a sufficient number of samples of both *S. alburnoides* and *S. pyrenaicus* to perform a statistical comparison of the populations. In Table 4 we present the results of an AMOVA in which two groups were considered (Tagus and Guadiana) with two populations per group (*S. alburnoides* and *S. pyrenaicus* of the same drainage). Inspection of Table 4 shows that the variation among groups is much greater than that among populations, supporting the view that much of the history of *S. pyrenaicus* and *S. alburnoides* in each basin was shared. The populations of *S. alburnoides* and *S. pyrenaicus* from Tagus are significantly distinct ($p=0.00$) and those from Guadiana approach significance ($p=0.07$). At the same time, the corrected mean number of pairwise differences between *S. alburnoides* and *S. pyrenaicus* of the same basin were very low (Table 4). This could be explained in one of two ways: either there was a single origin of *S. alburnoides* in each basin and the populations were so recent that they had little time to diverge; or many instances of haplotype transfers between

S. pyrenaicus and *S. alburnoides* took place on a much longer history. The within population variation is higher than the variation among populations and even greater than the variation among groups, indicating that each population had a considerably long history of mutation accumulation and haplotype diversification (Table 4) – the very low mean numbers of pairwise differences contrasted with the much higher levels of intrapopulation mean number of pairwise differences, as shown in Table 1. These results are in agreement with the structure of the tree (Figure S1) and of the networks presented in Figure 2.

Table 4 — AMOVA using pairwise differences with two groups (Tagus and Guadiana), each with two populations: *S. alburnoides* and *S. pyrenaicus* from Tagus (albT and pyrT); and *S. alburnoides* and *S. pyrenaicus* from Guadiana (albG and pyrG). The corrected mean number of pairwise differences within (diagonal) and between (below diagonal) populations is also presented. Significance p values for the exact test of population differentiation are indicated above diagonal.

	pyrT	albT	pyrG	albG		DF	Variance components	% of variation
pyrT	15.881	0.000	0.000	0.000	Among groups	1	2.234	32.03%
albT	1.119	7.051	0.000	0.000	Among populations within groups	2	0.358	4.90%
pyrG	4.873	4.904	4.165	0.090	Within populations	179	4.450	63.06%
albG	5.218	5.268	0.139	5.362	Total	182	7.042	

Discussion

The information retrieved from the molecular analyses led to the postulation of a single origin for the *S. alburnoides* complex (outlined below) and allowed the reconstitution of a hypothetical dispersal scenario that aimed to explain the foundation of populations in distinct river basins. However, since the dispersal of primary freshwater fish is only possible through fluvial connections, the routes followed by *S. alburnoides* had to be corroborated by the phylogeography of other *Squalius*.

Phylogeographical patterns of the Squalius species

The general pattern that emerged from the phylogenetic and phylogeographic analyses was that *S. pyrenaicus* is a highly diversified species with a wide distribution range whose ancestral population originated at least five new species as it dispersed towards the peripheral areas of the Iberian Peninsula: *S. carolitertii* in the North, *S. aradensis* and *S. torgalensis* in the Southwest, *S. valentinus* in the Southeast and *S. malacitanus* in the South (not studied; see Doadrio & Carmona 2006).

This picture involving a central, very widely distributed clade which originated a star-like pattern of peripheral derived clades resembles the peripatric speciation model proposed by Mayr (1982) and is congruent with our understanding of the Miocenic hydrography of Iberia, with a number of large endorheic lakes which subsequently branched, underwent fragmentation and became connected with rivers.

A detailed chronological description of those phylogeographical events, with estimated dates based on the estimated mtDNA divergence times (Table 2) and supported by geological data, is presented below - for a synthesis see Figure 6.

Miocenic pathways (Figure 6a)

Before it became an exorheic river in the Pliocene (Cunha *et al.* 1993; Andeweg 2002), the River Tagus was a system of at least five endorheic lakes, likely connected at some time since Middle Miocene fossils morphologically similar to extant *Squalius* species (Doadrio & Carmona 2006), were found in the Lower Tagus basin (Gaudant 1977).

The Lower Tagus basin (Figure 6a) may have acted as a littoral corridor that allowed the colonization of the southern Rivers Mira and Arade by a *Squalius* ancestor. Indeed, since *S. aradensis* and *S. torgalensis* are sister-species (Brito *et al.* 1997; Coelho *et al.* 1998), their differentiation depended on the arrival of a common ancestor to southwest Portugal at least in the Upper Miocene, the estimated age of the common ancestor of these species with the common ancestor of *S. pyrenaicus*/*S. carolitertii* (Doadrio & Carmona 2003; Sanjurjo *et al.* 2003).

Hypothetically, two routes could bring primary freshwater fish to southwest Portugal at that time: one from the East (from the Guadiana) or one from North (involving the Lower Tagus and the primitive basin of the Sado, which was intermittently connected

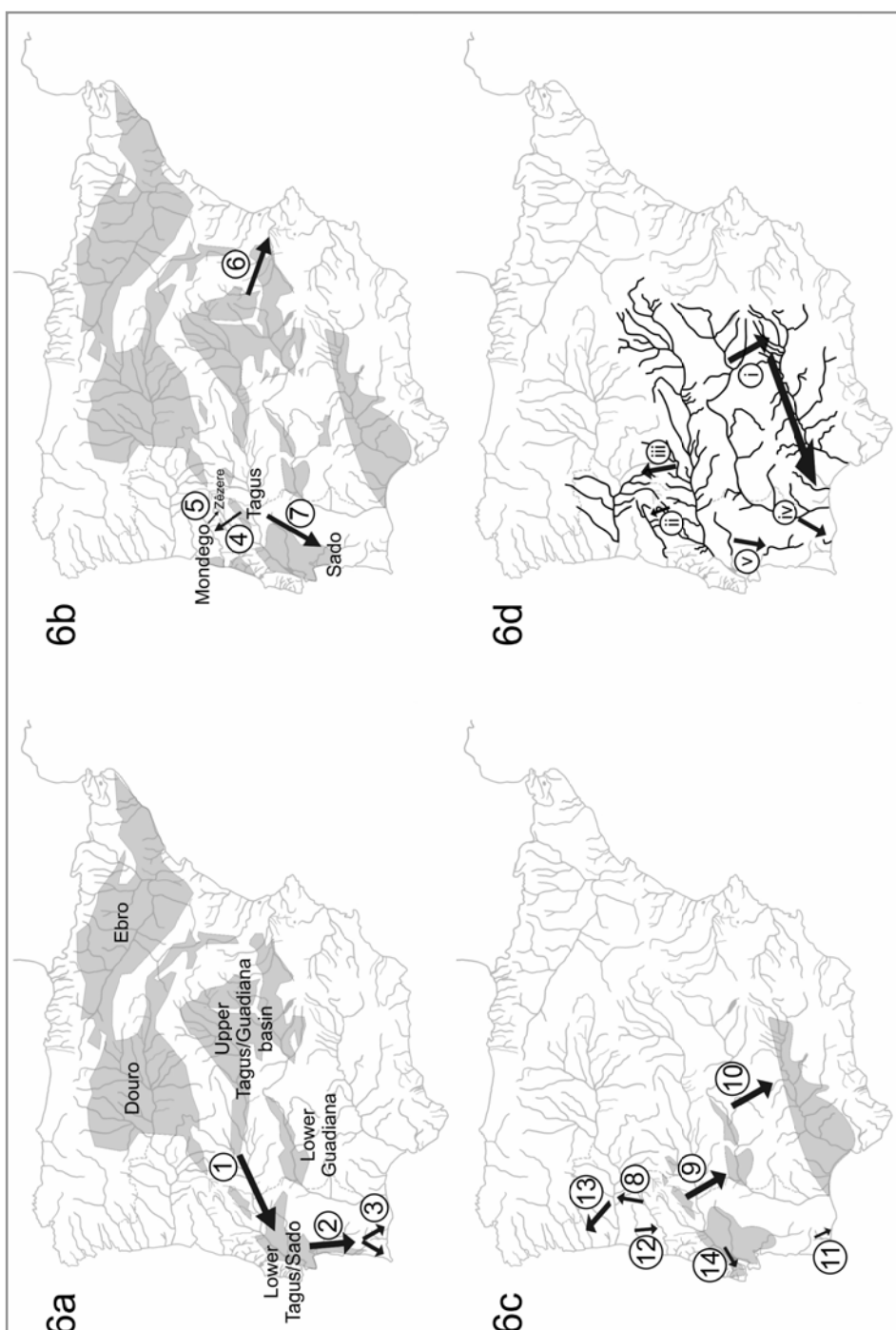


Figure 6 — Hypothesized pathways for *Squalius* species in the colonization of Iberian drainages. The numbered arrows represent the hypothetical pathways occurring in the Miocene (6a), Pliocene (6b), Pleistocene (6c) and Pleisto-Holocene (6d). Dates for each colonization event were estimated from the mtDNA divergence times (Table 2). Representations of the endorheic lakes were adapted from Andeweg (2002). Legend: 1 - Arrival of a proto-*S. pyrenaicus* to western Iberia (Middle Miocene); 2 - Colonization of the Southwest (Upper Miocene); 3 - Differentiation of *S. aradensis* and *S. torgalensis* (5.13 MY); 4 - Differentiation of the *S. carolierii*-Mondego phylogroup (4.49 MY); 5 - Differentiation of the *S. carolierii*-Tézere phylogroup (3.05 MY); 6 - Differentiation of the *S. valentinus* phylogroup (2.05 MY); 7 - Differentiation of the *S. pyrenaicus*-Sado phylogroup (2.01 MY); 8 - Differentiation of the *S. carolierii*-North phylogroup (0.72 MY); 9 - Renewed contacts Tagus-Guadiana (0.53 MY); 10 - Dispersal Upper Guadiana-Upper Guadalquivir (0.39 MY); 11 - Dispersal Arade-Quarteira (0.27 MY); 12 - Dispersal Vouga-Mondego (0.12 MY); 13 - Northward radiation of *S. carolierii* (0.05-0.00 MY); 14 - Dispersal Tagus-West (0.04 MY); i to v - hypothesized colonization pathways followed by *S. alburnoides* (see also text).

with the Tagus – T. Azevedo pers.com.). This last scenario seems more likely since 1) the southwestern endemism *Iberochondrostoma almakai* diverged from a common ancestor with *I. lusitanicum* (Robalo *et al.* 2007) that occurs in Sado but is absent from Guadiana; 2) *Squalius* from Arade and Sado shared a beta-actin haplotype; 3) the Guadiana River only drained to the south in the Pleistocene (Rodríguez-Vidal *et al.* 1991, 1993), which is posterior to the estimated age of these species; and 4) cyprinid species that inhabit the Guadiana are absent from the southwestern area (*Anaecypris hispanica*, *Pseudochondrostoma willkommii* and *Barbus microcephalus*).

The elevation of the Caldeirão Mountain, between the Guadiana and the southwestern rivers of Arade and Mira, 5.3 to 3.4 MY ago (Dias 2001), must have isolated the recently arrived *Squalius* ancestor. Moreover, the geomorphological changes that took place may have isolated a subpopulation in the River Mira and another in the River Arade, allowing the differentiation of *S. torgalensis* and *S. aradensis*, respectively. The estimated age of about 5.13 MY for the divergence between these two species is congruent with the timing of the elevation of the Caldeirão Mountain and with the divergence values obtained by Doadrio & Carmona (2003), Sanjur *et al.* (2003) and Mesquita *et al.* (2005). The haplotype P10 of the beta-actin gene found in specimens from the Mira and Arade Rivers may be one of the last vestiges of the connection between the two populations.

Pliocenic pathways (Figure 6b)

The northward migration to the Mondego of a common ancestor to the *S. pyrenaicus*-Tagus/Guadiana phylogroup was corroborated by the positioning of the haplotypes of *S. carolitertii* from Mondego in both mtDNA and beta-actin networks. From the Middle Miocene to the Early Pliocene, at least three small Tagus endorheic lagoons were present in the vicinity of the spring of the Mondego River (Andeweg 2002). Between 3.6 and 2.6 MY ago the Tagus River was acquiring its longitudinal profile, as the Upper Tagus basin drained towards west and the formerly isolated endorheic lakes were being united (Cunha *et al.* 1993). Thus, a connection with the adjacent Mondego basin was plausible, allowing the passage of the *S. carolitertii* ancestor.

According to the mtDNA analysis, *S. carolitertii* from the Zêzere River, a tributary of the right bank of the River Tagus located in the southeast side of the Estrela Mountain, diverged from that of Mondego in the Middle Pliocene. This suggests that after an initial dispersal of its ancestor from Tagus to Mondego, the derived phylogroup *S. carolitertii*-

Mondego reinvaded a Tagus tributary (Zêzere) at a later date, leading to the foundation of the *S. caroliterti*-Zêzere phylogroup. Our hypothesis is that a tributary of Mondego (whose headwaters are on the northwestern side of the Estrela Mountain but only about 15 km away from the sampling location of the Zêzere) must have drained to one of the endorheic lakes that existed in the area (Cunha *et al.* 1993).

The foundation of the Zêzere population by migrants proceeding from the Douro seems unlikely since their divergence was higher (3.39%) than with the Mondego population (3.20%); the most common nuclear haplotype in Douro (P6) is absent from Zêzere; and there is no geological support for contacts between Zêzere and Douro basins.

In the Pliocene, *S. pyrenaicus* from the Tagus also seem to have dispersed southwards, reinvading the Sado (2.11 MY ago) and allowing the foundation of the *S. pyrenaicus*-Sado phylogroup (represented by the nuclear haplotypes P8, P18, P36 and P37). This dispersal path is supported by the mtDNA networks (Figure 2) and by the nuclear P-haplotypes (Figure 4). To corroborate the Tagus-Sado pathway in detriment of a hypothetical Guadiana-Sado pathway is the fact that the nase and barbel species present in Sado are *Pseudochondrostoma polylepis* and *Barbus bocagei* that are present in Tagus but not in Guadiana. The Albufeira Lagoon, located between the mouths of the Tagus and Sado rivers, and inhabited by *I. lusitanicum* (Collares-Pereira 1983; Robalo *et al.* 2007), may be considered a last vestige of the connection between the two rivers. The differentiation of the *S. pyrenaicus*-Sado phylogroup could have been favoured by the isolation of subpopulations in more elevated areas where freshwaters persisted during frequent transgression episodes that caused an intermittent regime of contact with the adjacent Tagus drainage (Pimentel 1997; Andeweg 2002). Additionally, in the Upper Pliocene, a climatic crisis impeded the persistence of fluvial canals and was responsible for the disorganization of the Sado drainage network (Pimentel 1997). Consequently, the original *Squalius* populations may have suffered declines, accelerating the process of lineage sorting and causing the loss of the proto-*S. aradensis* mtDNA (the presence of the haplotype P14 in a fish from Sado is probably a last vestige of the Miocenic spread of *Squalius* to Arade through the Tagus-Sado corridor).

As it presents divergence values from *S. pyrenaicus* that are similar to the ones that allowed the description of *S. valentinus* as a distinct species (Doadrio & Carmona

2006), we suggest that the *S. pyrenaicus* population from Sado River basin should be eventually considered a new species.

In the Pliocene, the Upper Guadiana (which was isolated from its Lower section) was a tributary of the Upper Tagus (Moya-Palomares 2002), so that both rivers must have shared the same *S. pyrenaicus* populations at that period. The existence of two distinct sets of species distributed in the Upper Tagus (*Iberochondrostoma lemmingii* and *Achondrostoma arcasii*) and in the Lower Tagus (*I. lusitanicum*), with similarities with the species present, respectively, in the Guadiana and Sado basins, led us to postulate that the Tagus may have established independent connections with these two basins, at a time when the communication between its Lower and Upper sections was interrupted. The fact that *B. comizo* is shared by Tagus and Guadiana but not Sado and *P. polylepis* is shared by Tagus and Sado but not Guadiana, also supports this view.

The geographic proximity between the Upper Tagus and tributaries of some Mediterranean rivers may have also allowed the migration of the *S. valentinus* ancestors in the Pliocene.

Pleistocenic pathways (Figure 6c)

In the Pleistocene the Iberian hydrographical network acquired its current profile but some connections between drainages were still possible. The following Pleistocenic colonizations were postulated according to the topology of the network of mtDNA haplotypes and to the respective values of divergence between haplotypes:

1) Tagus-Guadiana (0.53 MY ago) – a contact between the lower sections of both drainages would explain the occurrence of four *S. pyrenaicus* from Guadiana carrying Tagus-like mtDNA and seems plausible according to geological data: one or more Portuguese tributaries of the Guadiana, in the region of Mora-Pavia, were tributaries of the Tagus (T. Azevedo, pers. com.); and contacts between the two basins occurred in the Badajoz area (Moya-Palomares 2002);

2) Tagus-Western rivers (0.04 MY ago) – migration of *S. pyrenaicus* to Rivers Lizandro, Samarra and Colares prior to the arrival of *S. alburnoides* to the lower section of the Tagus (since the complex is absent from the western rivers);

3) Guadiana-Guadalquivir (0.39 MY ago) - migrants of *S. pyrenaicus* proceeding from the Upper Guadiana must have reached the adjacent Guadalquivir drainage;

4) Arade-Quarteira (0.27 MY ago) - the geographic proximity between tributaries in the lowlands south of the Caldeirão Mountain might have allowed the migration of *S. aradensis* from Arade to Quarteira;

5) Mondego-Douro (0.35 MY ago) - the low mitochondrial and nuclear diversity of the *S. carolitertii* populations of the Douro corroborates a very recent colonization and/or reflects major depletions of the original fauna caused by glaciations. A similar pattern of higher levels of genetic diversity in Mondego populations when compared to the Douro and other northern populations was also detected in the golden-striped salamanders *Chioglossa lusitanica* (Alexandrino *et al.* 2002), for which a recent colonization by a small number of founders was suggested to explain the almost genetically uniform populations located north of the Douro;

6) Douro-Northern rivers (0.06 to 0.03 MY ago) – the very recent radiation of *S. carolitertii* to the other northern rivers is supported by the star-like mtDNA network (with a highly abundant root haplotype and many closely associated haplotypes) and may have been favoured by the major regression that took place 0.018 MY ago, during which almost all of the Portuguese continental shelf was above sea level, allowing the confluence of the mouths of the northern rivers (Dias *et al.* 2000).

7) Vouga-Mondego (0.13 MY ago) - postulated to explain the sharing of the P5 nuclear haplotype and the existence of mtDNA haplotypes belonging to the *S. carolitertii*-North phylogroup in the Mondego drainage.

Origin of the S. alburnoides complex

In previous studies (Alves *et al.* 1997b; Cunha *et al.* 2004) the similarities between the mtDNA haplotypes of *S. alburnoides* and of other *Squalius* from the same river basin were interpreted as evidence of an independent origin of the complex in that particular river basin. In our view, however, they may reflect the occurrence of recent interspecific crosses involving females of the sympatric *Squalius* species. As hypothesized by Sousa-Santos *et al.* (2006a) and corroborated by the results from the present work, the available data are

consistent with a single origin for the *S. alburnoides* complex, when both the maternal and paternal ancestors became sympatric, due to the historical rearrangements of the Iberian hydrographical network. This hypothesis seems more parsimonious than admitting the prolonged coexistence of the maternal and paternal ancestors in multiple river basins, the independent synthesis of the complex in each of those basins, and the subsequent extinction of one or both the ancestors depending on the river basin considered (Sousa-Santos *et al.* 2006a). The much higher levels of intrapopulation mean number of pairwise differences when compared to the very low mean numbers of pairwise differences between populations (Table 2) also contradict the hypothesis of a single origin per basin and supports the view that, from time to time, haplotypes of one population passed on to the other.

According to our findings, the origin of the complex must have occurred in the bulk of Iberia, in the Middle-Upper Pleistocene (less than 0.7 MY ago), more recently than the Upper Pliocene age proposed by Cunha *et al.* (2004). Our hypothesis is that the differentiation of the *Anaocypris*-like paternal ancestor of the complex occurred in a southern endorheic lake that remained isolated until it was captured by an ancient river carrying the *S. pyrenaicus* maternal ancestor. We suggest that this refuge was located in the area of what is now the River Guadiana since the distribution of the extant *A. hispanica* is restricted to this basin and the paternal ancestor, belonging to a derived clade, was also presumably favoured by the southern ecological conditions (namely, higher temperatures and intermittent conditions).

The capture of the endorheic lake must have been possible since with the tilting of the Peninsula towards the Atlantic in the Pleistocene, the Tagus and Guadiana rivers (which, at the time, were connected in the area where are now the headwaters of the Guadiana) began to drain towards west, acquiring their present longitudinal profiles and capturing the isolated endorheic lakes located on the way (Moya-Palomares 2002). As a result, the paternal ancestor of the complex must have become sympatric with *S. pyrenaicus* (at the time already differentiated in Tagus and Guadiana) and interspecific crosses gave rise to the complex. Afterwards, with the ongoing basculation process, the Tagus and Guadiana became completely isolated from each other but continued their path towards west (Moya-Palomares 2002), already carrying their respective *S. alburnoides* and *S. pyrenaicus* populations.

Dispersal of the S. alburnoides complex

Once originated, the complex must have dispersed throughout the connections between river basins that were still available in the Upper Pleistocene-Holocene, which explains why it has a wide distribution in the main drainages and is absent from the smaller and peripheral river basins of the Peninsula, already isolated at that time. These colonizations allowed the contact, not only with different *S. pyrenaicus* populations, but also with other *Squalius* species with which the complex interbreeds.

The dispersal route of the *S. alburnoides* complex, based on the estimated mtDNA divergence times (Table 2), likely included at least five colonization paths (represented by the same Arabic numbers in the text below and in Figure 6d):

i) from Upper Guadiana to Upper Guadalquivir (0.05 MY ago)

Path corroborated by the lowest divergence values between the *S. alburnoides* populations from Guadalquivir and from the left bank tributaries of the Upper Guadiana. Stream captures may also explain the migration of *S. alburnoides* from Guadalquivir to the adjacent Odiel drainage;

ii) from Tagus to Mondego (0.05 MY ago)

Through fluvial captures involving adjacent tributaries of the right bank of the Tagus basin, as corroborated by the lower divergence values involving *S. alburnoides* from the Zêzere-Erges area. These contacts may have also allowed the migration of *S. pyrenaicus* whose genes were probably diluted in the more abundant populations of its sister-species *S. carolitertii*. The five presumably *S. carolitertii* individuals from Mondego with *S. pyrenaicus* mtDNA may either be true *S. pyrenaicus* proceeding from Tagus or, alternatively, may be reconstituted from crosses between PPA females (carrying the *S. pyrenaicus* mtDNA) and *S. carolitertii* males. The later hypothesis would be discarded if the nuclear genomes of these five individuals showed P-haplotypes that were exclusive of the Tagus basin. However, two individuals were homozygous for a haplotype that was shared between Mondego and Tagus (P5) and the remaining three were heterozygous with one or both complements shared between the two basins. Brito *et al.* (1997) also found one *S. pyrenaicus* individual in the Mondego but interpreted it as a result of anthropogenic introductions;

iii) from Tagus to Douro (0.01 MY ago)

The *S. alburnoides* population of Douro showed a very low divergence value from the population of the Alagon river (tributary of Tagus, in the vicinity of the Portuguese border), which suggested that this may have been the corridor used in the colonization of the Douro basin. This hypothesis is corroborated by the fact that, in contrast to the wide distribution in the Portuguese Douro basin, the distribution of the complex in the Spanish Douro basin is restricted to a few tributaries of the left bank that are located in the vicinity of the Alagon area. Thus, after the colonization of those tributaries, the *S. alburnoides* complex may have reached the main course of Douro and, from there, dispersed virtually to all Portuguese tributaries. An upstream migration may have been impeded by the existence of a geological barrier of about 400 meters near the Portuguese border (Ribeiro *et al.* 1987). As in the case of the colonization of Mondego, the contact between Tagus and Douro may also have allowed the passage of *S. pyrenaicus*. This introgression was not yet detected but increased sampling effort and the use of new nuclear markers (the beta-actin does not differ between *S. carolitertii* and *S. pyrenaicus*) will very likely solve this issue;

iv) from lower Guadiana to Quarteira (0 MY ago)

After the colonization of Quarteira by *S. aradensis* from Arade River, a second colonization might have occurred: *S. alburnoides* proceeding from Guadiana seem to have colonized this river basin very recently, when the contact with the Arade had already ceased (since the complex is absent from Arade). The Guadiana acquired its present configuration and a southward draining pattern (that must have allowed the connections with Quarteira) very recently, in the Upper Pleistocene (Rodriguez-Vidal *et al.* 1993), which is in accordance with our results. The presence of *I. lemmingii*, which is present in Guadiana and Quarteira but not in Arade, also supports this route;

v) from Tagus to Sado (0 MY ago)

The colonization of Sado probably occurred when the upper section of the Tagus, carrying the *S. alburnoides* complex, merged with its lower section, yet connected with the Sado River basin. The *S. alburnoides* from Sado, in addition to mtDNA that is typical of the *S. pyrenaicus* from this basin, also exhibited Tagus-like mtDNA, which corroborates the postulation of a third dispersal wave from Tagus towards Sado (see the other two

postulated episodes above), that according to the geomorphological history of both drainages is not unlikely.

*Relationships between *S. alburnoides* and other *Squalius* species*

The mtDNA analysis showed a low number of haplotypes shared between *S. alburnoides* and other *Squalius* species, indicating that present crosses involving *S. alburnoides* males and females of other *Squalius* species are scarce. Scarcity is not, however, synonymous of absence and some proofs of the occurrence of interspecific crosses were found: 1) the complete replacement of the typical *S. pyrenaicus* mtDNA of the complex by *S. aradensis* mtDNA in Quarteira; and 2) the introgression of *S. carolitertii* mtDNA in some *S. alburnoides* individuals from Mondego and Douro. Thus, the reproduction of the *S. alburnoides* complex seems to involve mating with conspecifics, with males of other *Squalius* species, and, at least occasionally, with females of all the three sympatric *Squalius* species (*S. pyrenaicus*, *S. carolitertii* and *S. aradensis*), allowing the introgression of their mtDNA in the complex.

The presence of non-hybrid *S. alburnoides* males seems to be of extreme relevance to the process of replacement of the typical mtDNA of the complex. These males are probably more efficient in the diffusion of the mtDNA of other *Squalius* species, as corroborated by the finding that in the river basins where AA males are abundant (Sado, Guadiana, Tagus, Guadalquivir and Quarteira), the mtDNA of the *S. alburnoides* is identical to the mtDNA of the sympatric *Squalius* species. This is probably a result of higher attractiveness and fertilization success of these small males (Sousa-Santos *et al.* 2006b).

In contrast, in the river basins where non-hybrid males are absent (Mondego and Douro) the detected levels of introgression were much inferior. According to the proposed dispersal scenarios, the Mondego and Douro rivers were apparently colonised by *S. alburnoides* proceeding from tributaries of the right bank of the Tagus, where non-hybrid *S. alburnoides* males have not been found. This virtual absence of non-hybrids may be explained by unfavourable ecological conditions since they seem to prefer shallow waters with higher temperatures (Martins *et al.* 1998). Indeed, the tributaries of the right bank of the Tagus have higher discharges and lower water temperatures when compared to the tributaries of the left bank, whose ecological regimes resemble more the ones from the southern Mediterranean rivers. Moreover, the calculated age of the

colonization of Mondego (0.05 MY) predated the last glacial maximum (0.018 MY), that may have been responsible for severe bottlenecks, as river discharges were extremely higher due to a longer pluvial season and to the effect of spring ice melting (Dias *et al.* 2000), combined with a cooling that was probably unfavourable to AA males. Moreover, the persistence of non-hybrid males in populations is self-dependent, as they can only be originated by crosses between males of their own type and PAA females producing A gametes (Alves *et al.* 2002; Sousa-Santos *et al.* 2006b). Thus, if a secondary loss of this kind of males occurred in the northern populations, it is unlikely that they could be originated *de novo*.

Conversely, while non-hybrids may contribute to the introgression of distinct mtDNA in the complex, triploid PPA females might play an extremely important role in the introgression of nuclear and mtDNA in other *Squalius* species. As these females discard the uneven genome and perform normal meiosis (Crespo-López *et al.* 2006), the generated eggs carry a single P-haplotype. Thus, populations with abundant PPA females reflect a certain degree of autonomy from the sympatric *Squalius* species as P-donors. Additionally, the PPA females that colonized the Douro and Mondego drainages and crossed with *S. carolitertii* males transmitted nuclear genes of *S. pyrenaicus* to the offspring. However, if only mtDNA sequencing was performed, this transference of nuclear genes would be undetectable since the resultant offspring would be classified as *S. pyrenaicus*, when they should instead be classified as hybrids between *S. pyrenaicus* and *S. carolitertii*. Thus, the mtDNA analysis, when considered alone, may underestimate the extent of gene introgression between *Squalius* species.

The *S. alburnoides* complex is, therefore, besides being introgressed with sympatric *Squalius* genes, also responsible for the transference of mitochondrial and nuclear genes to different *Squalius* species, contributing to a homogenization of the *Squalius* genomes. This situation is particularly relevant at the nuclear DNA level since recombination between nuclear genes belonging to distinct species may occur, raising taxonomical problems related to the definition of the species.

To conclude, after reading the history of this hybrid complex from its molecular record, our results may be summarized as follows: 1) its origin may be traced back to the Pleistocene; 2) it is likely to have had a single origin, from hybridizations between an extinct *Anaecypris*-like species and *S. pyrenaicus* in the centre of Iberia, in the area of the

Upper Tagus/Upper Guadiana; 3) it apparently dispersed afterwards along several routes, namely: Guadiana-Guadalquivir-Odiel; Guadiana-Quarteira; Tagus-Sado; Tagus-Douro and Tagus-Mondego; and 4) it may have played a major role in bidirectional nuclear and mtDNA gene transfer with allopatric species and populations of *Squalius*.

Many fish hybrid lineages are mere sinks for the genes of sexual species they parasitize for reproducing. However, *S. alburnoides*, in its history of about 700,000 years, interacted with several other *Squalius* species, promoting bidirectional gene transfers. In this respect, the peculiar modes of reproduction of this hybrid complex emphasized by Alves *et al.* (2001), place it in a unique position not only in terms of its own evolution but in the evolutionary dynamics of other fish species. If the scenarios here reconstructed are correct, the dispersion/colonization paths of many other primary fish species might have occurred using the same fluvial connections. Thus, it is essential to delineate a wider research program with a much more intense sampling of other sympatric species and molecular markers to test for the signature of the events now postulated.

In this study we combined conventional phylogenetic inference procedures with phylogeographical tools and geological information, and our results suggest that phylogeographical analysis of slowly evolving nuclear genes like the beta-actin gene, may help to get a better picture of the past of a clade because these genes will be also more slowly affected by the processes leading to lineage sorting. Thus, ancestral haplotypes and historic relationships that left no equivalent signature in the mtDNA may be recovered, providing ways to get a more accurate phylogenetic reconstruction.

This research also illustrated the advantage of analysing phylogroups of mtDNA haplotypes instead of simply taking each species as a single collection of samples, as a way of identifying the relevant clades. Thus, the concerted use of phylogenetic and phylogeographical methods designed for studies at various time scales may be considered a promising combination of tools in paleobiogeography. In the future, the use of nuclear genes varying in their rates of evolution, combined with mtDNA analysis, may provide the necessary tools for validating the history of groups of organisms at the multiple time scales now advocated.

Acknowledgements We thank Sousa-Santos family and J. Robalo for help in sample collection and T. Azevedo for the helpful informations on the geological history of the Guadiana and Tagus drainages. Samples of *S. alburnoides* from River Almonte (Tagus) were kindly supplied by I. Doadrio. The study was funded by the FCT Pluriannual Program (U I&D 331/94 and U I&D 329/94) (FEDER participation). C. Sousa-Santos was supported by a PhD grant from FCT (SFRH/BD/8320/2002).

References

- Alexandrino J, Arntzen JW & Ferrand N. 2002. Nested clade analysis and the genetic evidence for population expansion in the phylogeography of the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Heredity* 88: 66-74.
- Almaça C. 1978. Répartition géographique des Cyprinidae Ibériques et secteurs ichthyogéographiques de la Péninsule Ibérique. *Vestník Československé Společnosti Zoologické* 42: 241-248.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 1997a. The *Rutilus alburnoides* complex (Cyprinidae): evidence for hybrid origin. *Journal of Zoological Systematics and Evolutionary Research* 35: 1-10.
- Alves MJ, Coelho MM, Collares-Pereira MJ & Dowling TE. 1997b. Maternal ancestry of the *Rutilus alburnoides* complex (Teleostei, Cyprinidae) as determined by analysis of cytochrome *b* sequences. *Evolution* 51: 1584-1592.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 2001. Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: a genetic review. *Genetica* 111: 375-385.
- Alves MJ, Collares-Pereira MJ, Dowling MJ & Coelho MM. 2002. The genetics of maintenance of an all-male lineage in the *Squalius alburnoides* complex. *Journal of Fish Biology* 60: 649-662.
- Alves MJ, Gromicho M, Collares-Pereira MJ, Crespo-López E & Coelho MM. 2004. Simultaneous production of triploid and haploid eggs by triploid *Squalius alburnoides* (Teleostei: Cyprinidae). *Journal of Experimental Zoology* 301A: 552-558.
- Andeweg B. 2002. Cenozoic tectonic evolution of the Iberian Peninsula, causes and effects of changing stress fields. PhD Thesis. Vrije Universiteit Amsterdam, Amsterdam, 178p.
- Banareescu P. 1973. Some reconsiderations on the zoogeography of the Euro-Mediterranean freshwater fish fauna. *Revue Romaine de Biologie (Zoologie)* 18: 257-264.
- Bandelt HJ, Forster P & Röhl A. 1999. Median-Joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 1681: 37-48.
- Bianco PG. 1990. Potential role of the paleohistory of the Mediterranean and Paratethys basins on the early dispersal of Euro-Mediterranean freshwater fishes. *Ichthyological Exploration of Freshwaters* 1: 167-184.
- Brito RM, Briolay J, Galtier N, Bouvet Y & Coelho MM. 1997. Phylogenetic relationships within genus *Leuciscus* (Pisces, Cyprinidae) in Portuguese freshwaters, based on mitochondrial cytochrome *b* sequences. *Molecular Phylogenetics and Evolution* 8: 435-442.

- Cabral MJ, Almeida J, Almeida PR, Dellinger T, Ferrand de Almeida N, Oliveira ME, Palmeirim JM, Queiroz AI, Rogado L & Santos-Reis M (Eds.) 2005. Livro Vermelho dos Vertebrados de Portugal. Instituto de Conservação da Natureza, Lisboa, 660p.
- Cabrera L & Gaudant J. 1985. Los Ciprínidos (Pisces) del sistema lacustre Oligocénico-Miocénico de los Monegros (sector SE de la Cuenca del Ebro, provincias de Lleida, Tarragona, Hoesca y Zaragoza). *Acta Geologica Hispanica* 20: 219-226.
- Carmona JA, Sanjur OI, Doadrio I, Machordom A & Vrijenhoek VC. 1997. Hybridogenetic reproduction and maternal ancestry of polyploid Iberian fish: the *Tropidophoxinellus alburnoides* complex. *Genetics* 146: 983-993.
- Coelho MM, Bogutskaya N, Rodrigues JA & Collares-Pereira MJ. 1998. *Leuciscus torgalensis* and *L. aradensis*, two new cyprinids for Portuguese fresh waters. *Journal of Fish Biology* 52: 937-950.
- Collares-Pereira MJ. 1983. Estudo sistemático e citogenético dos pequenos ciprinídeos ibéricos pertencentes aos géneros *Chondrostoma* Agassiz, 1835, *Rutilus* Rafinesque, 1820 and *Anaecypris* Collares-Pereira, 1983. PhD Thesis. University of Lisbon, Lisbon, 511p.
- Crespo-López ME, Duarte T, Dowling T & Coelho MM. 2006. Modes of reproduction of the hybridogenetic fish *Squalius alburnoides* in the Tejo and Guadiana rivers: an approach using microsatellites. *Zoology* 109: 277-286.
- Cunha P, Barbosa B & Reis R. 1993. Synthesis of the Piacenzian onshore record between the Aveiro and Setúbal parallels (Western Portuguese margin). *Ciências da Terra* (UNL) 12: 35-43.
- Cunha C, Coelho MM, Carmona JA & Doadrio I. 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Molecular Ecology* 13: 2807-2817.
- Cunha C, Mesquita N, Dowling TE, Gilles A & Coelho MM. 2002. Phylogenetic relationships of Eurasian and American cyprinids using cytochrome *b* sequences. *Journal of Fish Biology* 61: 929-944.
- De la Peña A. 1995. Tertiary fishes from the Iberia continental basins: History and fossil record. *Colóquios de Paleontologia* 47, Editorial Complutense, Madrid.
- Dias JMA, Boski T, Rodrigues A & Magalhães F. 2000. Coast line evolution in Portugal since the Last Glacial Maximum until present – a synthesis. *Marine Geology* 170: 177-186.
- Dias RP. 2001. Neotectónica da região do Algarve. PhD Thesis. University of Lisbon, Lisbon, 369p.
- Doadrio I & Carmona JA. 2003. Testing freshwater Lago Mare dispersal theory on the phylogeny relationships of Iberian cyprinid genera *Chondrostoma* and *Squalius* (Cypriniformes, Cyprinidae). *Graellsia* 59: 457-473.
- Doadrio I & Carmona JA. 2006. Phylogenetic overview of the genus *Squalius* (Actinopterygii, Cyprinidae) in the Iberian Peninsula, with description of two new species. *Cybium* 30: 199-214.
- Dowling T, Tibbets CA, Minckley WL & Smith GR. 2002. Evolutionary relationships of the Plagopterygins (Teleostei: Cyprinidae) from cytochrome *b* sequences. *Copeia* 2002(3): 665-678.

- Excoffier L, Laval G & Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47-50.
- Friend PF & Dabrio CJ (Eds.) 1996. Tertiary basins of Spain: the stratigraphic record of crustal kinematics. Cambridge University Press, Cambridge, 400p.
- Froese R & Pauly D (Eds.) 2007. FishBase. World Wide Web electronic publication, www.fishbase.org, version 01/2007.
- Gaudant J. 1977. Contributions à la paléontologie du Miocène moyen continental du bassin du Tage. II. Observations sur les dents pharyngiennes de poissons cyprinidés-Póvoa de Santarém. *Ciências da Terra (UNL)* 3: 129-141.
- Gromicho M, Coelho MM, Alves MJ & Collares-Pereira MJ. 2006. Cytogenetic analysis of *Anaocypris hispanica* and its relationship with the paternal ancestor of the diploid-polyploid *Squalius alburnoides* complex. *Genome* 49: 1621-1628.
- Gromicho M. & 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.
- Martins MJ, Collares-Pereira MJ, Cowx IG & Coelho MM. 1998. Diploids *v.* triploids of *Rutilus alburnoides*: spatial segregation and morphological differences. *Journal of Fish Biology* 52: 817-828.
- Mayr E. 1982. Processes of speciation in animals. In Barigozzi C (Ed.) Mechanisms of Speciation. Alan R. Liss, New York, pp 1-19.
- Mesquita N, Hänfling B, Carvalho GR & Coelho MM. 2005. Phylogeography of the cyprinid *Squalius aradensis* and implications for conservation of the endemic freshwater fauna of southern Portugal. *Molecular Ecology* 14: 1939-1954.
- Moya-Palomares ME. 2002. Evolución sedimentológica y geomorfológica de las Vegas Bajas del Guadiana entre Mérida y Badajoz (España). PhD Thesis. Universidad Complutense de Madrid, Madrid, 297p.
- Pala I & Coelho MM. 2005. Contrasting views over a hybrid complex: between speciation and evolutionary "dead-end". *Gene* 347: 283-294.
- Pimentel NL. 1997. O terciário da bacia do Sado. Sedimentologia e análise tectono-sedimentar. PhD Thesis. University of Lisbon, Lisbon, 381p.
- Ribeiro F, Beldade R, Dix M & Bochechas J. 2007. Carta Piscícola Nacional Direcção Geral dos Recursos Florestais - Fluviatilis, Lda. World Wide Web electronic publication, www.fluviatilil.com, version 01/2007.
- Ribeiro O, Lautensach H & Daveau S. 1987. Geografia de Portugal - I. A posição geográfica e o território. Edições João Sá da Costa, Lisboa, 334p.
- Robalo JI, Sousa-Santos C, Levy A & Almada VC. 2006. Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. *Molecular Phylogenetics and Evolution* 39: 276-281.

- Robalo JI, Almada VC, Levy A & Doadrio I. 2007. Re-examination and phylogeny of the genus *Chondrostoma* based on mitochondrial and nuclear data and the definition of 5 new genera. *Molecular Phylogenetics and Evolution* 42: 362–372.
- Rodríguez-Vidal J, Cáceres L & Ramirez A. 1991. La red fluvial cuaternaria en el piedemonte de Sierra Morena occidental. *Cuadernos de Investigación Geográfica* 17: 37–46.
- Rodríguez-Vidal J, Cáceres L & Ramirez A. 1993. Modelo evolutivo de la red fluvial cuaternaria en el suroeste de la Península Ibérica. Actas da 3ª Reunião do Quaternário Ibérico, Coimbra, pp 93–96.
- Sambrook J, Fritsch EF & Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor, New York, 999p.
- Sanjur OI, Carmona JA & Doadrio I. 2003. Evolutionary and biogeographical patterns within Iberian populations of the genus *Squalius* inferred from molecular data. *Molecular Phylogenetics and Evolution* 29: 20–30.
- Sousa-Santos C, Collares-Pereira MJ & Almada VC. 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. *Journal of Fish Biology* 68 (Supplement B): 292–301.
- Sousa-Santos C, Collares-Pereira MJ & Almada VC. 2006b. Reproductive success of nuclear non-hybrid males of *Squalius alburnoides* hybridogenetic complex (Teleostei, Cyprinidae): an example of interplay between female choice and ecological pressures? *Acta Ethologica* 9: 31–36.
- Sousa-Santos C, Collares-Pereira MJ & Almada VC. 2006c. 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. *Journal of Natural History* 40: 1443–1448.
- Sousa-Santos C, Collares-Pereira MJ & Almada VC. 2007. Fertile triploid males – an uncommon case among hybrid vertebrates. *Journal of Experimental Zoology* 307A: 220–225.
- Sousa-Santos C, Robalo J, Collares-Pereira MJ & Almada VC. 2005. Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genomic composition of hybrid organisms. *DNA Sequence* 16: 462–467.
- Swofford DL. 1998. PAUP® - phylogenetic analysis using parsimony (and other methods) version 4.0. Sinauer Associates, Sunderland.
- Zardoya R & Doadrio I. 1998. Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. *Proceedings of the Royal Society B* 265: 1365–1372.
- Zardoya R & Doadrio I. 1999. Molecular evidence on the evolutionary and biogeographical patterns of European Cyprinids. *Journal of Molecular Evolution* 49: 227–237.

Electronic Supplementary Material

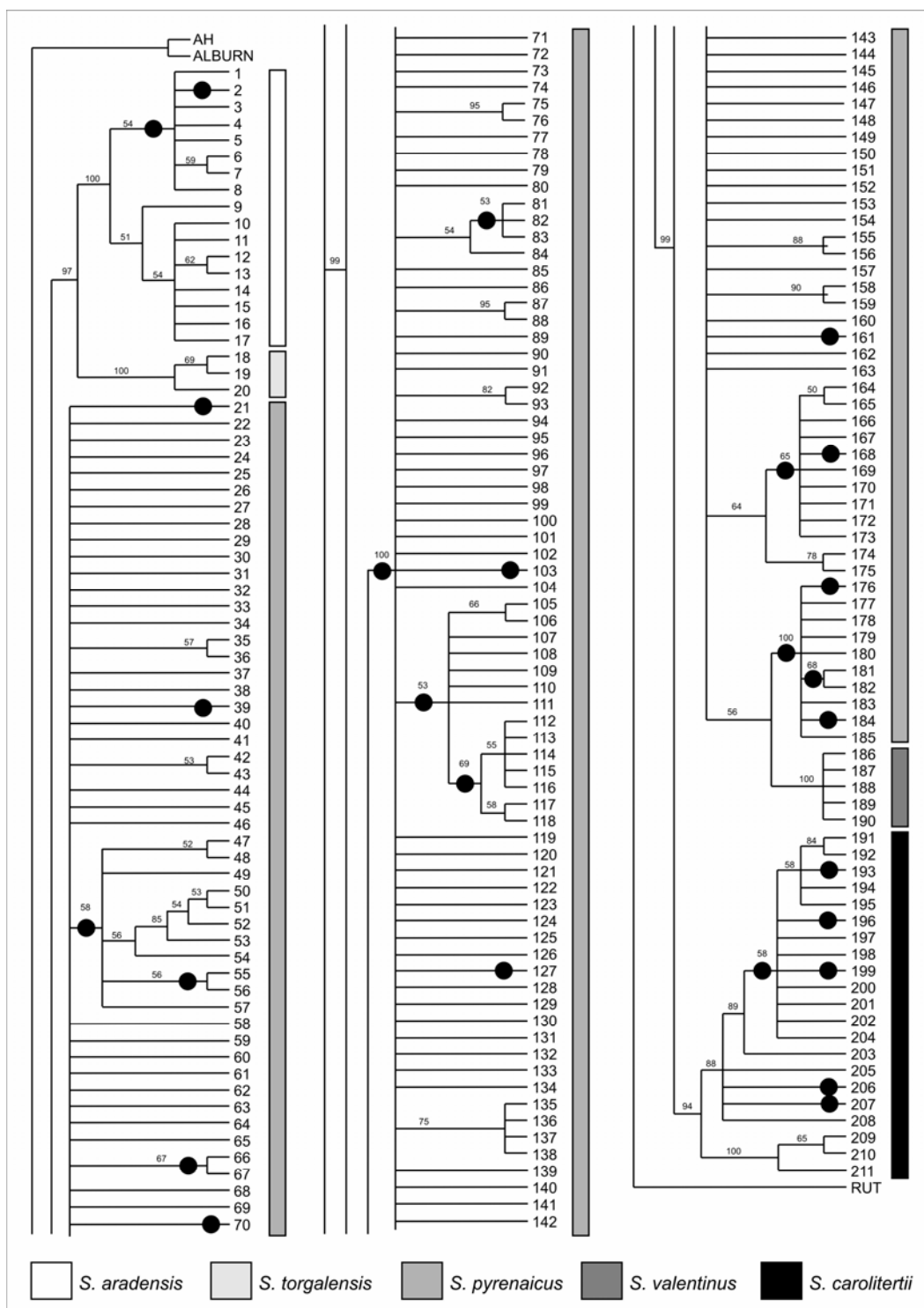


Figure S1 — Phylogenetic relationships between *Squalius* species, based on *cytb* gene sequences (haplotypes are numbered from 1 to 211 - the same numerical codes that were used in Figure 2 and Table S2). Numbers associated with the branches represent the bootstrap values obtained for 1000 replications for maximum-parsimony (nodes with bootstrap values lower than 50% were forced to collapse). Black dots refer to haplotypes that were shared between *S. alburnoides* and other *Squalius* species, representing recent and ancient interspecific crosses. *Anaecypris hispanica* (AH), *Alburnus alburnus* (ALBURN) and *Rutilus rutilus* (RUT) were used as outgroups (GenBank accession nos. [AJ427814](#), [DQ350253](#) and [DQ061933](#), respectively).

Table S1 — Known reproductive modes of *S. alburnoides* from southern populations (AA, PA, PAA, PPA and PPAA), according to Alves *et al.* (2001, 2004), Gromicho & Collares-Pereira (2004), Crespo-López *et al.* (2006) and Sousa-Santos *et al.* (2007), and the expected progeny from all possible crosses. Legend: H — Hybridogenesis; U — Production of unreduced gametes; M — normal meiosis; MH — meiotic hybridogenesis; X — virtually unviable progeny (the maximum number of chromosome sets known for the complex is exceeded); ? — unknown fertility.

			males					
			AA	PA	PAA	PPA	PPAA	
			M	U	U	?	M	
			a	pa	paa	?	pa	
females	PA	U	pa	PAA	PPAA	X	?	PPAA
	PAA	MH	a	AA	PAA	PAAA	?	PAA
		H	aa	AAA	PAAA	X	?	PAA
		U	paa	PAAA	X	X	?	X
	PPA	MH	p	PA	PPA	PPAA	?	PPA
	PPAA	M	pa	PAA	PPAA	X	?	PPAA

Table S2 — Frequencies of the different forms of the *S. alburnoides* males and females found in the populations of the northern and southern river basins (minimum and maximum values adapted from Alves *et al.* 2001; Pala & Coelho 2005).

		Northern river basins			Southern river basins			
Ploidy		Genome	Douro	Mondego	Genome	Tagus	Sado	Guadiana
females	2n	CA	0-4%	0-10%	PA	0-15%	36-77%	0-35%
	3n	CAA/CCA	36-90%	80-90%	PAA/PPA	50-100%	19-70%	11-88%
	4n	CCAA	0-1%	-	PPAA	0-10%	0-2%	-
males	2n	CA	10-14%	5-14.9%	PA	0-23%	-	-
	3n	CAA/CCA	0-1%	5.3%	PAA/PPA	0-22%	0-4%	0-5%
	4n	CCAA	0-1%	0-5%	PPAA	0-14%	0-2%	-
	2n	AA	-	-	AA	0-16%	0-48%	8-89%

Table S3 — GenBank accession numbers of all the individuals studied (ALB — *S. alburnoides*, CAR — *S. caroliterti*, PIR — *S. pyrenaicus*, ARD — *S. aradensis*, TOR — *S. torgalensis*, VAL — *S. valentinus*). Since some fishes were heterozygous for the analysed fragment of the beta-actin gene, each strand has its own accession number. Sequences resulted from the present study are identified with *. The remaining sequences were retrieved from GenBank and resulted from previous studies: 1 — Cunha *et al.* (2004), 2 — Alves *et al.* (1997b); 3 — Sanjurjo *et al.* (1997b); 4 — Zardoya & Doadrio (1998); 5 — Sousa-Santos *et al.* (2006c); 6 — Sousa-Santos *et al.* (2006a); 7 — Alves *et al.* (2001b); 8 — Mesquita *et al.* (2005); 9 — Durand *et al.* (2000); 10 — Sousa-Santos *et al.* (2005). The numbers used to designate the haplotypes of the *cytb* gene (1 to 211) are the same than the ones used in Figures 2 and S1.

Specimen	Drainage	River	Cytb haplotype	GenBank <i>cytb</i>	Source	Genomic constitution	Beta-actin haplotypes	GenBank beta-actin	Source
ALB 1	Douro	Águeda	164	<u>AJ627330</u>	1	-	-	-	-
ALB 2	Douro	Ardã	168	<u>EF458756</u>	*	-	-	-	-
ALB 3	Douro	Azibo	-	-	*	CAA	P3A9A9	strand P: <u>EF459142</u> , strand A1: <u>EF459143</u> , strand A2: <u>EF459144</u>	*
ALB 4	Douro	Azibo	-	-	*	CAA	P3A9A9	strand P: <u>EF459145</u> , strand A1: <u>EF459146</u> , strand A2: <u>EF459147</u>	*
ALB 5	Douro	Calvo	173	<u>EF458758</u>	*	-	-	-	-
ALB 6	Douro	Calvo	196	<u>EF458757</u>	*	CCA	P3P3A9	strand P1: <u>EF459148</u> , strand P2: <u>EF459149</u> , strand A1: <u>EF459150</u>	*
ALB 7	Douro	Coa	168	<u>EF458759</u>	*	CAA	P3A9A9	strand P: <u>EF459151</u> , strand A1: <u>EF459152</u> , strand A2: <u>EF459153</u>	*
ALB 8	Douro	Coa	168	<u>EF458760</u>	*	CAA	P3A9A9	strand P: <u>EF459154</u> , strand A1: <u>EF459155</u> , strand A2: <u>EF459156</u>	*
ALB 9	Douro	Maçãs	-	-	*	CA	P3A9	strand P: <u>EF459362</u> , strand A: <u>EF459363</u>	*
ALB 10	Douro	Maçãs	168	<u>AJ627335</u>	1	-	-	-	-
ALB 11	Douro	Maçãs	168	<u>X99431</u>	2	-	-	-	-
ALB 12	Douro	Paiva	168	<u>AJ627331</u>	1	-	-	-	-
ALB 13	Douro	Paiva	168	<u>AJ627332</u>	1	-	-	-	-
ALB 14	Douro	Paiva	168	<u>AJ627333</u>	1	-	-	-	-
ALB 15	Douro	Pega	173	<u>EF458761</u>	*	-	-	-	-
ALB 16	Douro	Sabor	-	-	*	CAA	P6A9A9	strand P: <u>EF459199</u> , strand A1: <u>EF459200</u> , strand A2: <u>EF459201</u>	*
ALB 17	Douro	Sabor	-	-	*	CAA	P3A9A9	strand P: <u>EF459202</u> , strand A1: <u>EF459203</u> , strand A2: <u>EF459204</u>	*
ALB 18	Douro	Sabor	135	<u>EF458775</u>	*	CAA	P3A9A9	strand P: <u>EF459196</u> , strand A1: <u>EF459197</u> , strand A2: <u>EF459198</u>	*
ALB 19	Douro	Sabor	168	<u>EF458762</u>	*	CAA	P3A9A9	strand P: <u>EF459157</u> , strand A1: <u>EF459158</u> , strand A2: <u>EF459159</u>	*
ALB 20	Douro	Sabor	168	<u>EF458766</u>	*	CAA	P6A9A9	strand P: <u>EF459169</u> , strand A1: <u>EF459170</u> , strand A2: <u>EF459171</u>	*
ALB 21	Douro	Sabor	168	<u>EF458767</u>	*	CAA	P3A9	strand P: <u>EF459172</u> , strand A1: <u>EF459173</u> , strand A2: <u>EF459174</u>	*
ALB 22	Douro	Sabor	168	<u>EF458768</u>	*	CAA	P3A9A9	strand P: <u>EF459175</u> , strand A1: <u>EF459176</u> , strand A2: <u>EF459177</u>	*
ALB 23	Douro	Sabor	168	<u>EF458769</u>	*	CAA	P3A9A9	strand P: <u>EF459178</u> , strand A1: <u>EF459179</u> , strand A2: <u>EF459180</u>	*
ALB 24	Douro	Sabor	168	<u>EF458770</u>	*	CAA	P3A9A9	strand P: <u>EF459181</u> , strand A1: <u>EF459182</u> , strand A2: <u>EF459183</u>	*
ALB 25	Douro	Sabor	168	<u>EF458771</u>	*	CAA	P3A9A9	strand P: <u>EF459184</u> , strand A1: <u>EF459185</u> , strand A2: <u>EF459186</u>	*
ALB 26	Douro	Sabor	168	<u>EF458772</u>	*	CAA	P6A9A9	strand P: <u>EF459187</u> , strand A1: <u>EF459188</u> , strand A2: <u>EF459189</u>	*
ALB 27	Douro	Sabor	170	<u>EF458773</u>	*	CAA	P3A9	strand P: <u>EF459190</u> , strand A1: <u>EF459191</u> , strand A2: <u>EF459192</u>	*
ALB 28	Douro	Sabor	171	<u>EF458774</u>	*	CAA	P3A9A9	strand P: <u>EF459193</u> , strand A1: <u>EF459194</u> , strand A2: <u>EF459195</u>	*
ALB 29	Douro	Sabor	193	<u>EF458763</u>	*	CCA	P3P3A9	strand P1: <u>EF459160</u> , strand P2: <u>EF459161</u> , strand A1: <u>EF459162</u>	*
ALB 30	Douro	Sabor	196	<u>EF458765</u>	*	CAA	P3A9A9	strand P: <u>EF459166</u> , strand A1: <u>EF459167</u> , strand A2: <u>EF459168</u>	*

ALB 31	Douro	Sabor	198	EF458764	*	CAA	P3A9A9	strand P: EF459163 ; strand A1: EF459164 ; strand A2: EF459165	*
ALB 32	Douro	Sousa	168	AJ627334	1	-	-	-	-
ALB 33	Douro	Tâmega	168	AJ627336	1	-	-	-	-
ALB 34	Douro	Tâmega	168	EF458776	*	-	-	-	-
ALB 35	Douro	Tâmega	172	EF458777	*	CAA	P3A9A9	strand P: EF459205 ; strand A1: EF459206 ; strand A2: EF459207	*
ALB 36	Douro	Tâmega	172	EF458778	*	-	-	-	-
ALB 37	Guadalquivir	Guaditel	66	AJ627337	1	-	-	-	-
ALB 38	Guadalquivir	Jandula	57	AJ627338	1	-	-	-	-
ALB 39	Guadalquivir	Manzano	55	AJ627339	1	-	-	-	-
ALB 40	Guadalquivir	Montemayor	60	AJ627340	1	-	-	-	-
ALB 41	Guadiana	Albuera	61	AJ627341	1	-	-	-	-
ALB 42	Guadiana	Ardila	21	EF458789	*	AA'	A3A6	strand A1: EF459374 ; strand A2: EF459375	*
ALB 43	Guadiana	Ardila	21	EF458790	*	-	-	-	-
ALB 44	Guadiana	Ardila	21	EF458791	*	PAA	P7A4A4	strand P: EF459220 ; strand A1: EF459221 ; strand A2: EF459222	*
ALB 45	Guadiana	Ardila	21	EF458792	*	AA'	A2A4	strand A1: EF459376 ; strand A2: EF459377	*
ALB 46	Guadiana	Ardila	23	EF458787	*	PAA	P7A4A4	strand P: EF459217 ; strand A1: EF459218 ; strand A2: EF459219	*
ALB 47	Guadiana	Ardila	25	EF458786	*	PAA	P7A4A4	strand P: EF459214 ; strand A1: EF459215 ; strand A2: EF459216	*
ALB 48	Guadiana	Ardila	26	EF458785	*	PA	P7A4	strand P: EF459370 ; strand A: EF459371	*
ALB 49	Guadiana	Ardila	39	EF458780	*	AA'	A4A11	strand A1: EF459364 ; strand A2: EF459365	*
ALB 50	Guadiana	Ardila	41	EF458782	*	PAA	P7A4A4	strand P: EF459208 ; strand A1: EF459209 ; strand A2: EF459210	*
ALB 51	Guadiana	Ardila	58	EF458788	*	AA'	A4A11	strand A1: EF459372 ; strand A2: EF459373	*
ALB 52	Guadiana	Ardila	70	EF458783	*	AA'	A2A4	strand A1: EF459368 ; strand A2: EF459369	*
ALB 53	Guadiana	Ardila	70	EF458784	*	PAA	P7A4A4	strand P: EF459211 ; A1: EF459212 ; strand A2: EF459213	*
ALB 54	Guadiana	Ardila	77	EF458781	*	AA'	A3A4	strand A1: EF459366 ; strand A2: EF459367	*
ALB 55	Guadiana	Arronches	70	AJ627342	1	-	-	-	-
ALB 56	Guadiana	Caia	70	X99426	2	-	-	-	-
ALB 57	Guadiana	Degebe	75	AJ627343	1	-	-	-	-
ALB 58	Guadiana	Degebe	76	X99425	2	-	-	-	-
ALB 59	Guadiana	Estena	65	AJ627344	1	-	-	-	-
ALB 60	Guadiana	Estena	77	AF045992	4	-	-	-	-
ALB 61	Guadiana	Guadiana	-	-	*	AA'	A4A6	AY943893	10
ALB 62	Guadiana	Guadiana	-	-	*	AA	A4A4	AY943864	10
ALB 63	Guadiana	Guadiana	176	EF458779	*	AA'	A4A6	AY943892	10
ALB 64	Guadiana	Guadiana	28	DQ263231	5	PA	P7A3	strand P: DQ335476 ; strand A: DQ335477	5
ALB 65	Guadiana	Guadiana	46	DQ263230	5	PA	PI3A4	strand P: DQ335474 ; strand A: DQ335475	5
ALB 66	Guadiana	Guadiana	70	DQ263227	5	AA	A4A4	DQ010337	5
ALB 67	Guadiana	Guadiana	70	DQ263228	5	PAA	PI2A4A4	strand P: DQ335472 ; strand A: DQ335473	5
ALB 68	Guadiana	Guadiana	71	DQ263232	5	PAA	P7A4A4	strand P: DQ335478 ; strand A: DQ335479	5
ALB 69	Guadiana	Guadiana	80	DQ263229	5	AA'	A2A6	AY943891	10

[illegible]

ALB 109	Quarteira	Quarteira	-	-	*	AA'	A4A13	strand A1: EF459380 ; strand A2: EF459381	*
ALB 110	Quarteira	Quarteira	-	-	*	AA'	A4A13	strand A1: EF459382 ; strand A2: EF459383	*
ALB 111	Quarteira	Quarteira	-	-	*	AA	A4A4	DQ350252	5
ALB 112	Quarteira	Quarteira	-	-	*	AA'	A4A13	strand A1: EF459384 ; strand A2: EF459385	*
ALB 113	Quarteira	Quarteira	-	-	*	AA	A4A4	EF458694	*
ALB 114	Quarteira	Quarteira	-	-	*	AA'	A4A6	strand A1: EF459386 ; strand A2: EF459387	*
ALB 115	Quarteira	Quarteira	2	DQ003239	6	QAA	P10A4A4	strand A: DQ150345 ; strand P: DQ150346	6
ALB 117	Quarteira	Quarteira	2	DQ003240	6	QAA	P10A4A13	strand A1: DQ150347 ; strand A2: DQ150348 ; strand P: DQ150349	6
ALB 118	Quarteira	Quarteira	2	DQ003241	6	AA'	A4A13	strand A1: EF459358 ; strand A2: EF459359	*
ALB 119	Quarteira	Quarteira	2	DQ003242	6	AA	A4A4	AY943865	10
ALB 119	Quarteira	Quarteira	2	DQ003253	6	QAA	P14A4A4	strand A: DQ150343 ; strand P: DQ150344	6
ALB 120	Quarteira	Quarteira	2	DQ003254	6	QAA	P10A4A4	strand A: DQ150350 ; strand P: DQ150351	6
ALB 121	Quarteira	Quarteira	2	DQ003259	6	AA	A4A4	AY943867	10
ALB 122	Quarteira	Quarteira	2	DQ141679	6	QQ'AA	P7P10A4A4	strand P1: EF458677 ; strand P2: EF458678	*
ALB 123	Quarteira	Quarteira	2	DQ141680	6	QAA	P10A4A4	strand A1: EF458679 ; strand A2: EF458680	6
ALB 124	Quarteira	Quarteira	2	DQ141681	6	QAA	P10A4A4	strand A: DQ150337 ; strand P: DQ150338	6
ALB 125	Quarteira	Quarteira	2	DQ141682	6	AA'	A4A13	strand A: DQ150339 ; strand P: DQ150340	6
ALB 126	Quarteira	Quarteira	2	DQ141683	6	AA	A4A4	strand A1: DQ150341 ; strand A2: DQ150342	6
ALB 127	Quarteira	Quarteira	2	DQ141684	6	AA	A4A4	DQ128102	5
ALB 128	Quarteira	Quarteira	2	DQ141687	6	QQ'A	P7P32A14	EF458692	*
ALB 129	Quarteira	Quarteira	3	DQ003246	6	AA'	A4A13	strand A: DQ150352 ; strand P1: DQ150353 ; strand P2: DQ150354	6
ALB 130	Quarteira	Quarteira	30	DQ003238	6	QAA	P10A4A4	strand A1: EF459360 ; strand A2: EF459361	*
ALB 131	Quarteira	Quarteira	4	DQ003255	6	QA	P10A4	strand A: AY943866 ; strand P: DQ150360	6
ALB 132	Quarteira	Quarteira	4	DQ141686	6	AA	A4A4	strand A: DQ150355 ; strand P: DQ150356	6
ALB 133	Quarteira	Quarteira	4	DQ141688	6	QAA'	A4A4	EF458693	*
ALB 134	Quarteira	Quarteira	5	DQ141685	6	AA'	P10A4A13	strand A1: DQ150357 ; strand A2: DQ150358 ; strand P: DQ150359	6
ALB 135	Quarteira	Quarteira	72	AJ627358	6	AA'	A6A13	strand A1: DQ150335 ; strand A2: DQ150336	6
ALB 136	Sado	Odivelas	176	AJ427838	7	-	-	-	-
ALB 137	Sado	Odivelas	177	X99432	1	-	-	-	-
ALB 138	Sado	Odivelas	181	AJ627357	1	-	-	-	-
ALB 139	Sado	Sado	111	EF458814	*	PA	P8A11	strand P: EF459388 ; strand A: EF459389	*
ALB 140	Sado	Sado	168	EF458813	*	PAA	P8A4	strand P: EF459286 ; strand A1: EF459287 ; strand A2: EF459288	*
ALB 141	Sado	Sado	176	EF458819	*	PA	P18A4	strand P: EF459390 ; strand A: EF459391	*
ALB 142	Sado	Sado	176	EF458815	*	PA	P8A4	strand P: EF459392 ; strand A: EF459393	*
ALB 143	Sado	Sado	176	EF458821	*	PAA	P8A4	strand P: EF459298 ; strand A1: EF459299 ; strand A2: EF459300	*
ALB 144	Sado	Sado	176	EF458822	*	PAA	P8A4A4	strand P: EF459301 ; strand A1: EF459302 ; strand A2: EF459303	*
ALB 145	Sado	Sado	176	EF458823	*	PA	P8A4	strand P: EF459394 ; strand A: EF459395	*
ALB 146	Sado	Sado	176	EF458824	*	PAA	P8A4A4	strand P: EF459304 ; strand A1: EF459305 ; strand A2: EF459306	*

ALB 147	Sado	176	<u>EF458825</u>	*	PAA	P8A4A4	strand P: <u>EF459307</u> ; strand A1: <u>EF459308</u> ; strand A2: <u>EF459309</u>	
ALB 148	Sado	176	<u>EF458826</u>	*	PAA	P8A4A4	strand P: <u>EF459310</u> ; strand A1: <u>EF459311</u> ; strand A2: <u>EF459312</u>	
ALB 149	Sado	176	<u>EF458827</u>	*	PAA	P8A4A4	strand P: <u>EF459313</u> ; strand A1: <u>EF459314</u> ; strand A2: <u>EF459315</u>	
ALB 150	Sado	176	<u>EF458828</u>	*	PAA	P8A4A4	strand P: <u>EF459316</u> ; strand A1: <u>EF459317</u> ; strand A2: <u>EF459318</u>	
ALB 151	Sado	176	<u>EF458829</u>	*	PAA	P8A4A4	strand P: <u>EF459319</u> ; strand A1: <u>EF459320</u> ; strand A2: <u>EF459321</u>	
ALB 152	Sado	176	<u>EF458830</u>	*	PAA	P8A4A4	strand P: <u>EF459322</u> ; strand A1: <u>EF459323</u> ; strand A2: <u>EF459324</u>	
ALB 153	Sado	176	<u>EF458831</u>	*	PAA	P8A4A4	strand P: <u>EF459325</u> ; strand A1: <u>EF459326</u> ; strand A2: <u>EF459327</u>	
ALB 154	Sado	176	<u>EF458832</u>	*	-	-	-	
ALB 155	Sado	178	<u>EF458817</u>	*	PAA	P8A4A4	strand P: <u>EF459292</u> ; strand A1: <u>EF459293</u> ; strand A2: <u>EF459294</u>	
ALB 156	Sado	179	<u>EF458816</u>	*	PPA	P8P8A4	strand P1: <u>EF459289</u> ; strand P2: <u>EF459290</u> ; strand A1: <u>EF459291</u>	
ALB 157	Sado	180	<u>EF458834</u>	*	-	-	-	
ALB 158	Sado	184	<u>EF458818</u>	*	PAA	P8A4A4	strand P: <u>EF459295</u> ; strand A1: <u>EF459296</u> ; strand A2: <u>EF459297</u>	
ALB 159	Sado	185	<u>EF458833</u>	*	-	-	-	
ALB 160	Tagus	169	<u>AJ627359</u>	1	-	-	-	
ALB 161	Tagus	118	<u>AJ627364</u>	1	-	-	-	
ALB 162	Tagus	103	<u>EF458836</u>	*	PAA	P11A15A15	strand P: <u>EF459331</u> ; strand A1: <u>EF459332</u> ; strand A2: <u>EF459333</u>	
ALB 163	Tagus	104	<u>AJ627365</u>	1	-	-	-	
ALB 164	Tagus	81	<u>EF458854</u>	*	-	-	-	
ALB 165	Tagus	85	<u>EF458835</u>	*	PAA	P11A15A15	strand P: <u>EF459328</u> ; strand A1: <u>EF459329</u> ; strand A2: <u>EF459330</u>	
ALB 166	Tagus	127	<u>EF458837</u>	*	PPA	P2P4A4	strand P1: <u>EF458681</u> ; strand P2: <u>EF458682</u> ; strand A1: <u>EF458683</u>	
							strand P1: <u>EF458684</u> ; strand P2: <u>EF458685</u> ;	
							strand A1: <u>EF458686</u> ; strand A2: <u>EF458687</u>	
ALB 167	Tagus	127	<u>EF458838</u>	*	PPAA'	P2P2A4A10		
ALB 168	Tagus	165	<u>AJ627360</u>	1	-	-	-	
ALB 169	Tagus	168	<u>AJ627361</u>	1	-	-	-	
ALB 170	Tagus	112	<u>AJ627367</u>	1	-	-	-	
ALB 171	Tagus	113	<u>AJ627366</u>	1	-	-	-	
ALB 172	Tagus	167	<u>AJ627362</u>	1	-	-	-	
ALB 173	Tagus	95	<u>AJ627368</u>	1	-	-	-	
ALB 174	Tagus	91	<u>AJ627369</u>	1	-	-	-	
ALB 175	Tagus	-	-	-	PA	P1A4		
ALB 176	Tagus	107	<u>EF458840</u>	*	PA	P3A9	strand P: <u>EF459398</u> ; strand A: <u>EF459399</u>	
ALB 177	Tagus	89	<u>EF458839</u>	*	PA	P11A1	strand P: <u>EF459396</u> ; strand A: <u>EF459397</u>	
ALB 178	Tagus	91	<u>AJ627370</u>	1	-	-	-	
ALB 179	Tagus	91	<u>AJ627371</u>	1	-	-	-	
ALB 180	Tagus	96	<u>AJ627372</u>	1	-	-	-	
ALB 181	Tagus	94	<u>AJ627373</u>	1	-	-	-	
ALB 182	Tagus	90	<u>AJ627374</u>	1	-	-	-	
ALB 183	Tagus	166	<u>AJ627363</u>	1	-	-	-	
ALB 184	Tagus	-	-	-	PA	P3A9	strand P: <u>EF459400</u> ; strand A: <u>EF459401</u>	

ALB 185	Tagus	Ocreza	-	-	-	PAA'	P11A5A9	strand P: EF459334 ; strand A1: EF459335 ; strand A2: EF459336	*
ALB 186	Tagus	Seda	-	-	-	AA	A4A4	EF458695	*
ALB 187	Tagus	Seda	-	-	-	AA	A4A4	EF458696	*
ALB 188	Tagus	Seda	-	-	-	AA'	A4A8	AY943895	10
ALB 189	Tagus	Seda	-	-	-	AA'	A4A10	AY943894	10
ALB 190	Tagus	Serfă	138	AJ627375	*	-	-	-	-
ALB 191	Tagus	Serfă	108	X99429	2	-	-	-	-
ALB 192	Tagus	Sever	108	AJ627376	1	-	-	-	-
ALB 193	Tagus	Sever	114	X99427	2	-	-	-	-
ALB 194	Tagus	Sorraia	160	-	-	-	-	-	-
ALB 195	Tagus	Sorraia	-	-	-	AA	A4A4	AY943863	10
ALB 196	Tagus	Sorraia	112	EF458847	*	PAA	P3A9A9	strand P: EF459340 ; strand A1: EF459341 ; strand A2: EF459342	*
ALB 197	Tagus	Sorraia	112	EF458848	*	PAA	P3A9A9	strand P: EF459343 ; strand A1: EF459344 ; strand A2: EF459345	*
ALB 198	Tagus	Sorraia	115	EF458850	*	PAA	P3A9A9	strand P: EF459349 ; strand A1: EF459350 ; strand A2: EF459351	*
ALB 199	Tagus	Sorraia	124	EF458842	*	PA	P1A3	strand P: EF459405 ; strand A: EF459406	*
		Sorraia	127	EF458853	*	PAAA	P4A4A4A4	strand P: EF458688 ; strand A1: EF458689 ; strand A2: EF458690 ; strand A3: EF458691	*
ALB 200	Tagus	Sorraia	129	EF458846	*	PAA	P2A3A3	strand P: EF459337 ; strand A1: EF459338 ; strand A2: EF459339	*
ALB 201	Tagus	Sorraia	132	AJ427839	7	-	-	-	-
ALB 202	Tagus	Sorraia	132	AJ627377	1	-	-	-	-
ALB 203	Tagus	Sorraia	133	EF458844	*	PA	P2A4	strand P: EF459409 ; strand A: EF459410	*
ALB 204	Tagus	Sorraia	134	EF458845	*	PA	P2A3	strand P: EF459411 ; strand A: EF459412	*
ALB 205	Tagus	Sorraia	136	EF458843	*	PA	P2A4	strand P: EF459407 ; strand A: EF459408	*
ALB 206	Tagus	Sorraia	137	EF458851	*	PAA	P2A3A3	strand P: EF459352 ; strand A1: EF459353 ; strand A2: EF459354	*
ALB 207	Tagus	Sorraia	141	EF458852	*	AA'	A7A8	AY943896	10
ALB 208	Tagus	Sorraia	148	AJ627378	1	-	-	-	-
ALB 209	Tagus	Sorraia	148	X99428	2	-	-	-	-
ALB 210	Tagus	Sorraia	149	EF458849	*	PPA	P3P3A3	strand P1: EF459346 ; strand P2: EF459347 ; strand A: EF459348	*
ALB 211	Tagus	Tietar	119	AJ627381	1	-	-	-	-
ALB 212	Tagus	Tietar	120	AJ627380	1	-	-	-	-
ALB 213	Tagus	Trevijana	82	AJ627382	1	-	-	-	-
ALB 214	Tagus	Vid	117	AJ627383	1	-	-	-	-
ALB 215	Tagus	Zêzere	-	-	-	PPA	P2P17A4	strand P1: EF459355 ; strand P2: EF459356 ; strand A: EF459357	*
ALB 216	Tagus	Zêzere	160	AJ627379	1	-	-	-	-
ARD 1	Aljezur	Aljezur	6	AJ583062	8	-	-	-	-
ARD 2	Aljezur	Aljezur	6	AJ583063	8	-	-	-	-
ARD 3	Aljezur	Aljezur	6	AJ583064	8	-	-	-	-
ARD 4	Aljezur	Aljezur	6	AJ583065	8	-	-	-	-
ARD 5	Aljezur	Aljezur	7	AJ583066	8	-	-	-	-
ARD 6	Aljezur	Aljezur	7	AJ583067	8	-	-	-	-

ARD 7	Alvor	6	AJ583068	8	-	-
ARD 8	Alvor	6	AJ583069	8	-	-
ARD 9	Alvor	6	AJ583070	8	-	-
ARD 10	Alvor	6	AJ583071	8	-	-
ARD 11	Alvor	6	AJ583072	8	-	-
ARD 12	Alvor	6	AJ583073	8	-	-
ARD 13	Arade	-	-	QQ'	P26P27	strand P1: <u>DQI50305</u> ; strand P2: <u>DQI50306</u>
ARD 14	Arade	10	AJ583077	8	-	-
ARD 15	Arade	10	AJ583078	8	-	-
ARD 16	Arade	10	AJ583079	8	-	-
ARD 17	Arade	12	AF421825	3	-	-
ARD 18	Arade	13	AF421824	3	-	-
ARD 19	Arade	14	AJ583074	8	-	-
ARD 20	Arade	14	AJ583075	8	-	-
ARD 21	Arade	2	AJ583076	8	-	-
ARD 22	Azilheira	14	EF458856	*	QQ'	strand P1: <u>DQI50303</u> ; strand P2: <u>DQI50304</u>
ARD 23	Arade	2	EF458855	*	QQ'	strand P1: <u>DQI50301</u> ; strand P2: <u>DQI50302</u>
ARD 24	Arade	-	-	QQ'	P9P10	strand P1: <u>DQI50293</u> ; strand P2: <u>DQI50294</u>
ARD 25	Arade	11	EF458874	*	-	-
ARD 26	Arade	14	EF458860	*	QQ'	strand P1: <u>DQI50295</u> ; strand P2: <u>DQI50296</u>
ARD 27	Odelouca	14	EF458861	*	-	-
ARD 28	Arade	14	EF458862	*	-	-
ARD 29	Arade	14	EF458863	*	-	-
ARD 30	Arade	14	EF458864	*	-	-
ARD 31	Arade	14	EF458865	*	-	-
ARD 32	Odelouca	14	EF458866	*	-	-
ARD 33	Arade	15	EF458871	*	-	-
ARD 34	Arade	15	EF458872	*	-	-
ARD 35	Arade	15	EF458873	*	-	-
ARD 36	Arade	16	EF458857	*	-	-
ARD 37	Arade	2	EF458858	*	-	-
ARD 38	Arade	2	EF458859	*	-	-
ARD 39	Arade	9	EF458867	*	QQ'	strand P1: <u>DQI50299</u> ; strand P2: <u>DQI50300</u>
ARD 40	Arade	9	EF458868	*	QQ'	strand P1: <u>DQI50297</u> ; strand P2: <u>DQI50298</u>
ARD 41	Arade	9	EF458869	*	QQ'	strand P1: <u>DQI50291</u> ; strand P2: <u>DQI50292</u>
ARD 42	Arade	9	EF458870	*	-	-
ARD 43	Quarteira	-	-	QQ	P10P10	<u>DQI50267</u>
ARD 44	Quarteira	1	DQ003250	6	QQ	<u>DQI50261</u>
ARD 45	Quarteira	2	DQI45177	6	QQ	<u>DQI50271</u>

ARD 46	Quarteira	2	<u>AJ583080</u>	8	-	-	-	strand P1: <u>DQ150321</u> , strand P2: <u>DQ150322</u>	6	-
ARD 47	Quarteira	2	<u>AJ583081</u>	8	-	-	-	<u>DQ150262</u>	6	-
ARD 48	Quarteira	2	<u>AJ583082</u>	8	-	-	-	strand P1: <u>DQ150311</u> , strand P2: <u>DQ150312</u>	6	-
ARD 49	Quarteira	2	<u>AJ583083</u>	8	-	-	-	<u>DQ150269</u>	6	-
ARD 50	Quarteira	2	<u>AJ583085</u>	8	-	-	-	strand P1: <u>DQ150317</u> , strand P2: <u>DQ150318</u>	6	-
ARD 51	Quarteira	2	<u>AJ583086</u>	8	-	-	-	strand P1: <u>DQ150315</u> , strand P2: <u>DQ150316</u>	6	-
ARD 52	Quarteira	2	<u>DQ003243</u>	6	QQ'	-	-	<u>DQ150268</u>	6	-
ARD 53	Quarteira	2	<u>DQ003244</u>	6	QQ	P14P19	P10P10	strand P1: <u>DQ150309</u> , strand P2: <u>DQ150310</u>	6	-
ARD 54	Quarteira	2	<u>DQ003245</u>	6	QQ'	P10P10	P10P14	<u>DQ150270</u>	6	-
ARD 55	Quarteira	2	<u>DQ003247</u>	6	QQ	P14P14	P14P14	strand P1: <u>DQ150319</u> , strand P2: <u>DQ150320</u>	6	-
ARD 56	Quarteira	2	<u>DQ003248</u>	6	QQ'	P10P14	P10P14	<u>DQ150260</u>	6	-
ARD 57	Quarteira	2	<u>DQ003249</u>	6	QQ'	P10P14	P10P14	strand P1: <u>DQ150307</u> , strand P2: <u>DQ150308</u>	6	-
ARD 58	Quarteira	2	<u>DQ003251</u>	6	QQ	P10P10	P10P10	strand P1: <u>DQ150313</u> , strand P2: <u>DQ150314</u>	6	-
ARD 59	Quarteira	2	<u>DQ003252</u>	6	QQ'	P10P14	P10P14	<u>DQ150266</u>	6	-
ARD 60	Quarteira	2	<u>DQ003256</u>	6	QQ	P14P14	P14P14	<u>DQ150265</u>	6	-
ARD 61	Quarteira	2	<u>DQ003257</u>	6	QQ'	P10P14	P10P14	<u>DQ150264</u>	6	-
ARD 62	Quarteira	2	<u>DQ003258</u>	6	QQ	P10P10	P10P10	<u>DQ150263</u>	6	-
ARD 63	Quarteira	2	<u>DQ141689</u>	6	QQ'	P10P14	P10P14	-	-	-
ARD 64	Quarteira	2	<u>DQ141690</u>	6	QQ	P10P10	P10P10	-	-	-
ARD 65	Quarteira	2	<u>DQ141691</u>	6	QQ	P10P10	P10P10	-	-	-
ARD 66	Quarteira	2	<u>DQ141692</u>	6	QQ	P10P10	P10P10	-	-	-
ARD 67	Quarteira	2	<u>DQ141693</u>	6	QQ	P10P10	P10P10	-	-	-
ARD 68	Quarteira	2	<u>DQ141694</u>	6	QQ	P10P10	P10P10	-	-	-
ARD 69	Quarteira	8	<u>AJ583084</u>	8	-	-	-	-	-	-
ARD 70	Seixe	14	<u>AJ583056</u>	8	-	-	-	-	-	-
ARD 71	Seixe	14	<u>AJ583058</u>	8	-	-	-	-	-	-
ARD 72	Seixe	14	<u>AJ583059</u>	8	-	-	-	-	-	-
ARD 73	Seixe	14	<u>AJ583060</u>	8	-	-	-	-	-	-
ARD 74	Seixe	14	<u>AJ583061</u>	8	-	-	-	-	-	-
ARD 75	Seixe	17	<u>AJ583057</u>	8	-	-	-	-	-	-
CAR 1	Ave	196	<u>EF458876</u>	*	CC'	-	P3P6	strand P1: <u>EF459413</u> , strand P2: <u>EF459414</u>	*	-
CAR 2	Ave	196	<u>EF458877</u>	*	-	-	-	-	-	-
CAR 3	Ave	196	<u>EF458878</u>	*	-	-	-	-	-	-
CAR 4	Ave	196	<u>EF458879</u>	*	CC'	-	P3P6	strand P1: <u>EF459415</u> , strand P2: <u>EF459416</u>	*	-
CAR 5	Ave	196	<u>EF458880</u>	*	-	-	-	-	-	-
CAR 6	Ave	196	<u>EF458881</u>	*	CC	-	P6P6	<u>EF458697</u>	*	-
CAR 7	Ave	196	<u>EF458882</u>	*	-	-	-	-	-	-
CAR 8	Ave	196	<u>EF458883</u>	*	-	-	-	-	-	-
CAR 9	Ave	196	<u>EF458884</u>	*	-	-	-	-	-	-

[illegible]

CAR 49	Douro	Boedo	195	<u>AF421797</u>	3	-	-	P3P3	-	
CAR 50	Douro	Minas	194	<u>EF458930</u>	*	CC	CC	P3P3	<u>EF458703</u>	*
CAR 51	Douro	Minas	194	<u>EF458931</u>	*	CC	CC	P3P3	<u>EF458704</u>	*
CAR 52	Douro	Minas	194	<u>EF458932</u>	*	CC	CC	P3P3	<u>EF458705</u>	*
CAR 53	Douro	Paiva	-	=	*	CC	CC	P6P6	<u>EF458706</u>	*
CAR 54	Douro	Paiva	196	<u>Y10135</u>	2	-	-	-	-	-
CAR 55	Douro	Sabor	196	<u>EF458917</u>	*	CC	CC	P3P3	<u>EF458707</u>	*
CAR 56	Douro	Tâmega	196	<u>EF458927</u>	*	CC	CC	P3P6	strand P1: <u>EF459423</u> ; strand P2: <u>EF459424</u>	*
CAR 57	Douro	Tâmega	196	<u>EF458928</u>	*	CC	CC	P6P6	<u>EF458714</u>	*
CAR 58	Douro	Tâmega	203	<u>EF458918</u>	*	CC	CC	P3P3	<u>EF458708</u>	*
CAR 59	Douro	Tâmega	203	<u>EF458919</u>	*	CC	CC	P3P6	strand P1: <u>EF459417</u> ; strand P2: <u>EF459418</u>	*
CAR 60	Douro	Tâmega	203	<u>EF458920</u>	*	CC	CC	P6P6	<u>EF458709</u>	*
CAR 61	Douro	Tâmega	203	<u>EF458921</u>	*	CC	CC	P6P6	<u>EF458710</u>	*
CAR 62	Douro	Tâmega	203	<u>EF458922</u>	*	CC	CC	P6P6	<u>EF458711</u>	*
CAR 63	Douro	Tâmega	203	<u>EF458923</u>	*	CC	CC	P6P6	<u>EF458712</u>	*
CAR 64	Douro	Tâmega	203	<u>EF458924</u>	*	CC	CC	P3P6	strand P1: <u>EF459419</u> ; strand P2: <u>EF459420</u>	*
CAR 65	Douro	Tâmega	203	<u>EF458925</u>	*	CC	CC	P3P6	strand P1: <u>EF459421</u> ; strand P2: <u>EF459422</u>	*
CAR 66	Douro	Tâmega	203	<u>EF458926</u>	*	CC	CC	P3P3	<u>EF458713</u>	*
CAR 67	Douro	Távora	196	<u>EF458929</u>	*	CC	CC	P3P3	<u>AY943890</u>	10
CAR 68	Lima	Lima	-	-	*	CC	CC	P3P3	<u>AY943883</u>	10
CAR 69	Lima	Lima	-	-	*	CC	CC	P3P3	<u>EF458718</u>	*
CAR 70	Lima	Lima	196	<u>EF458937</u>	*	-	-	-	-	*
CAR 71	Lima	Lima	196	<u>EF458938</u>	*	-	-	-	-	*
CAR 72	Lima	Lima	196	<u>EF458939</u>	*	CC	CC	P3P3	<u>EF458715</u>	*
CAR 73	Lima	Lima	196	<u>EF458940</u>	*	CC	CC	P3P3	<u>EF458716</u>	*
CAR 74	Lima	Lima	196	<u>EF458941</u>	*	CC	CC	P6P6	<u>EF458717</u>	+
CAR 75	Lima	Lima	196	<u>EF458942</u>	*	-	-	-	-	*
CAR 76	Lima	Lima	199	<u>EF458933</u>	*	-	-	-	-	-
CAR 77	Lima	Lima	199	<u>EF458934</u>	*	-	-	-	-	-
CAR 78	Lima	Lima	199	<u>EF458935</u>	*	-	-	-	-	-
CAR 79	Lima	Lima	200	<u>EF458943</u>	*	CC	CC	P3P6	strand P1: <u>EF459425</u> ; strand P2: <u>EF459426</u>	*
CAR 80	Lima	Lima	202	<u>EF458936</u>	*	-	-	-	-	*
CAR 81	lima	Salas	196	<u>AF421795</u>	3	-	-	-	-	-
CAR 82	lima	Salas	196	<u>AF421796</u>	3	-	-	-	-	-
CAR 83	Lima	Vez	196	<u>EF458944</u>	*	-	-	-	-	-
CAR 84	Lima	Vez	196	<u>EF458945</u>	*	-	-	-	-	-
CAR 85	Lima	Vez	196	<u>EF458946</u>	*	-	-	-	-	-
CAR 86	Lima	Vez	196	<u>EF458947</u>	*	-	-	-	-	-
CAR 87	Lima	Vez	196	<u>EF458948</u>	*	-	-	-	-	-

CAR 88	Lima	Vez	196	<u>EF458949</u>	*	-	-	-	strand P1: EF459427; strand P2: EF459428 AY943886	10
CAR 89	Lima	Vez	196	<u>EF458950</u>	*	-	-	-		
CAR 90	Lima	Vez	196	<u>EF458951</u>	*	-	-	-		
CAR 91	Lima	Vez	196	<u>EF458952</u>	*	-	-	-		
CAR 92	Minho	Coura	196	<u>EF458953</u>	*	CC	P3P6			
CAR 93	Minho	Coura	196	<u>EF458954</u>	*	CC	P3P3			
CAR 94	Minho	Mouro	196	<u>EF458955</u>	*	-	-	-		
CAR 95	Minho	Mouro	196	<u>EF458956</u>	*	-	-	-		
CAR 96	Minho	Mouro	196	<u>EF458957</u>	*	-	-	-		
CAR 97	Minho	Mouro	196	<u>EF458958</u>	*	-	-	-		
CAR 98	Minho	Mouro	196	<u>EF458959</u>	*	-	-	-		
CAR 99	Minho	Mouro	196	<u>EF458960</u>	*	-	-	-		
CAR 100	Minho	Mouro	196	<u>EF458961</u>	*	-	-	-		
CAR 101	Minho	Mouro	196	<u>EF458962</u>	*	-	-	-		
CAR 102	Minho	Mouro	196	<u>EF458963</u>	*	-	-	-		
CAR 103	Minho	Mouro	196	<u>EF458964</u>	*	-	-	-		
CAR 104	Minho	Mouro	196	<u>EF458965</u>	*	-	-	-		
CAR 105	Minho	Mouro	196	<u>EF458966</u>	*	-	-	-		
CAR 106	Minho	Mouro	196	<u>EF458967</u>	*	-	-	-		
CAR 107	Minho	Mouro	196	<u>EF458968</u>	*	-	-	-		
CAR 108	Minho	Mouro	196	<u>EF458969</u>	*	-	-	-		
CAR 109	Minho	Mouro	196	<u>EF458970</u>	*	-	-	-		
CAR 110	Minho	Tea	196	<u>EF458971</u>	*	CC	P6P6	EF458719		
CAR 111	Minho	Tea	196	<u>EF458972</u>	*	CC	P6P6	<u>EF458720</u>		
CAR 112	minho	Viverey	196	<u>AF421793</u>	3	-	-	-		
CAR 113	minho	Viverey	196	<u>AF421794</u>	3	-	-	-		
CAR 114	mondengo	Alva	205	<u>YI0136</u>	2	-	-	-		
CAR 115	Mondogo	Ceira	-	=	*	CC	P5P20	strand P1: EF459453; strand P2: EF459454		
CAR 116	Mondogo	Ceira	-	-	*	CC	P3PI1	strand P1: EF459455; strand P2: EF459456 AY943887		10
CAR 117	Mondogo	Ceira	-	-	*	CC	P5P5			*
CAR 118	Mondogo	Ceira	-	-	*	CC	P3P20	strand P1: EF459457; strand P2: EF459458		*
CAR 119	Mondogo	Ceira	-	-	*	CC	P5PI1	strand P1: EF459459; strand P2: EF459460		*
CAR 120	Mondogo	Ceira	-	-	*	CC	P5PI1	strand P1: EF459449; strand P2: EF459450 EF458724		*
CAR 121	Mondogo	Ceira	161	<u>EF458988</u>	*	CC	P5P5			*
CAR 122	Mondogo	Ceira	161	<u>EF458989</u>	*	CC	P5PI1	strand P1: EF459451; strand P2: EF459452		*
CAR 123	Mondogo	Ceira	163	<u>EF458991</u>	*	CC	P5P5	<u>EF458725</u>		*
CAR 124	Mondogo	Ceira	168	<u>EF458987</u>	*	CC	PI1P20	strand P1: EF459447; strand P2: EF459448		*
CAR 125	Mondogo	Ceira	201	<u>EF458985</u>	*	CC	P3P3	<u>EF458723</u>		*
CAR 126	Mondogo	Ceira	201	<u>EF458986</u>	*	-	-	-		*

CAR 127	Mondego	Ceira	206	<u>EF458973</u>	*	CC'	P5P11	strand P1: <u>EF459422</u> , strand P2: <u>EF459430</u>	*
CAR 128	Mondego	Ceira	206	<u>EF458974</u>	*	CC	P11P11	<u>EF458721</u>	*
CAR 129	Mondego	Ceira	206	<u>EF458975</u>	*	CC'	P5P11	strand P1: <u>EF459431</u> , strand P2: <u>EF459432</u>	*
CAR 130	Mondego	Ceira	206	<u>EF458976</u>	*	CC'	P2P5	strand P1: <u>EF459433</u> , strand P2: <u>EF459434</u>	*
CAR 131	Mondego	Ceira	206	<u>EF458977</u>	*	CC	P5P5	<u>AY943888</u>	10
CAR 132	Mondego	Ceira	206	<u>EF458978</u>	*	CC'	P2P5	strand P1: <u>EF459435</u> , strand P2: <u>EF459436</u>	*
CAR 133	Mondego	Ceira	207	<u>EF458979</u>	*	CC'	P3P11	strand P1: <u>EF459437</u> , strand P2: <u>EF459438</u>	*
CAR 134	Mondego	Ceira	207	<u>EF458980</u>	*	CC'	P5P11	strand P1: <u>EF459439</u> , strand P2: <u>EF459440</u>	*
CAR 135	Mondego	Ceira	207	<u>EF458981</u>	*	CC'	P5P11	strand P1: <u>EF459441</u> , strand P2: <u>EF459442</u>	*
CAR 136	Mondego	Ceira	207	<u>EF458982</u>	*	CC'	P5P20	strand P1: <u>EF459443</u> , strand P2: <u>EF459444</u>	*
CAR 137	Mondego	Ceira	207	<u>EF458983</u>	*	CC'	P3P20	strand P1: <u>EF459445</u> , strand P2: <u>EF459446</u>	*
CAR 138	Mondego	Ceira	208	<u>EF458984</u>	*	CC	P5P5	<u>EF458722</u>	*
CAR 139	Neiva	Neiva	-	-	*	CC'	P3P23	strand P1: <u>EF459461</u> , strand P2: <u>EF459462</u>	*
CAR 140	Neiva	Neiva	-	-	*	CC'	P3P23	strand P1: <u>EF459463</u> , strand P2: <u>EF459464</u>	*
CAR 141	Neiva	Neiva	196	<u>EF458992</u>	*	-	-	<u>AY943884</u>	-
CAR 142	Neiva	Neiva	196	<u>EF458993</u>	*	CC'	P3P16		10
CAR 143	Neiva	Neiva	196	<u>EF458994</u>	*	-	-		-
CAR 144	Neiva	Neiva	196	<u>EF458995</u>	*	-	-		-
CAR 145	Neiva	Neiva	196	<u>EF458996</u>	*	-	-		-
CAR 146	Neiva	Neiva	196	<u>EF458997</u>	*	-	-		-
CAR 147	Neiva	Neiva	196	<u>EF458998</u>	*	-	-		-
CAR 148	Neiva	Neiva	196	<u>EF458999</u>	*	-	-		-
CAR 149	Neiva	Neiva	196	<u>EF459000</u>	*	-	-		-
CAR 150	Neiva	Neiva	196	<u>EF459001</u>	*	-	-		-
CAR 151	Neiva	Neiva	196	<u>EF459002</u>	*	-	-		-
CAR 152	Neiva	Neiva	196	<u>EF459003</u>	*	-	-		-
CAR 153	Neiva	Neiva	196	<u>EF459004</u>	*	-	-		-
CAR 154	Vouga	Sul	196	<u>EF459005</u>	*	CC	P3P3	<u>AY943885</u>	10
CAR 155	Vouga	Sul	197	<u>EF459006</u>	*	-	-		-
CAR 156	Vouga	Vouga	-	-	*	CC'	-	strand P1: <u>EF459467</u> , strand P2: <u>EF459468</u>	-
CAR 157	Vouga	Vouga	-	-	*	CC	P3P6	<u>EF458727</u>	*
CAR 158	Vouga	Vouga	-	-	*	CC	P3P3	<u>AY943882</u>	*
CAR 159	Vouga	Vouga	196	<u>EF459007</u>	*	-	-		10
CAR 160	Vouga	Vouga	196	<u>EF459008</u>	*	-	-		-
CAR 161	Vouga	Vouga	196	<u>EF459009</u>	*	-	-		-
CAR 162	Vouga	Vouga	196	<u>EF459010</u>	*	-	-		-
CAR 163	Vouga	Vouga	196	<u>EF459011</u>	*	CC'	P5P6	strand P1: <u>EF459465</u> , strand P2: <u>EF459466</u>	*
CAR 164	Vouga	Vouga	196	<u>EF459012</u>	*	CC	P6P6	<u>EF458726</u>	*
CAR 165	Vouga	Vouga	196	<u>EF459013</u>	*	-	-		-

CAR 166	Vouga	196	EF459014	*	-	pP25	strand P1: EF459469 , strand P2: EF459470
CAR 167	Vouga	196	EF459015	*	-	P2P2	EF458728
CAR 168	Vouga	196	EF459016	*	-	P2P2	EF458729
CAR 169	Vouga	196	EF459017	*	-	P2P2	EF458731
CAR 170	Vouga	196	EF459018	*	-	-	-
CAR 171	Vouga	196	EF459019	*	-	-	-
CAR 172	Vouga	196	EF459020	*	-	-	-
CAR 173	Vouga	196	EF459021	*	-	-	-
CAR 174	Vouga	196	EF459022	*	-	-	-
CAR 175	Vouga	196	EF459023	*	-	-	-
PXR 1	Colares	-	-	*	pp'	P2P25	-
PXR 2	Colares	-	-	*	pp	P2P2	-
PXR 3	Colares	-	-	*	pp	P2P2	-
PXR 4	Colares	-	-	*	pp	P2P2	-
PXR 5	Colares	121	EF459026	*	-	-	-
PXR 6	Colares	121	EF459027	*	-	-	-
PXR 7	Colares	121	EF459028	*	-	-	-
PXR 8	Colares	121	EF459041	*	pp	P2P2	EF458730
PXR 9	Colares	126	EF459025	*	-	-	-
PXR 10	Colares	127	EF459029	*	-	-	-
PXR 11	Colares	127	EF459030	*	-	-	-
PXR 12	Colares	127	EF459031	*	-	-	-
PXR 13	Colares	127	EF459032	*	-	-	-
PXR 14	Colares	127	EF459033	*	-	-	-
PXR 15	Colares	127	EF459034	*	-	-	-
PXR 16	Colares	127	EF459035	*	-	-	-
PXR 17	Colares	127	EF459036	*	-	-	-
PXR 18	Colares	127	EF459037	*	-	-	-
PXR 19	Colares	127	EF459038	*	-	-	-
PXR 20	Colares	127	EF459039	*	-	-	-
PXR 21	Colares	127	EF459040	*	-	-	-
PXR 22	Colares	162	EF459024	*	-	-	-
PXR 23	Ebro	98	AJ252821	9	-	-	-
PXR 24	Ebro	98	AF421802	3	-	-	-
PXR 25	Guadalquivir	49	AF421816	3	-	-	-
PXR 26	Guadalquivir	79	AF421817	3	-	-	-
PXR 27	Manzano	53	AJ627287	1	-	-	-
PXR 28	Manzano	53	AJ627288	1	-	-	-
PXR 29	Mollinos	67	AJ627289	1	-	-	-

PYR 30	Guadalquivir	Montemayor	50	<u>AF42I790</u>	3	-	-	-
PYR 31	Guadalquivir	Montemayor	51	<u>AJ627290</u>	1	-	-	-
PYR 32	Guadalquivir	Montemayor	52	<u>AJ627291</u>	1	-	-	-
PYR 33	Guadalquivir	Montemayor	54	<u>AJ627292</u>	1	-	-	-
PYR 34	Guadalquivir	Robledillo	56	<u>AJ627293</u>	1	-	-	-
PYR 35	Guadiana	Arronches	31	<u>AJ627294</u>	1	-	-	-
PYR 36	Guadiana	Azuér	32	<u>AF42I805</u>	3	-	-	-
PYR 37	Guadiana	Azuér	32	<u>AJ627295</u>	1	-	-	-
PYR 38	Guadiana	Azuér	45	<u>AJ627296</u>	1	-	-	-
PYR 39	Guadiana	Azuér	70	<u>AF42I804</u>	3	-	-	-
PYR 40	Guadiana	Caia	31	<u>YI0I34</u>	2	-	-	-
PYR 41	Guadiana	Estena	33	<u>AF42I814</u>	3	-	-	-
PYR 42	Guadiana	Estena	33	<u>AJ627298</u>	1	-	-	-
PYR 43	Guadiana	Estena	38	<u>AJ627297</u>	1	-	-	-
PYR 44	Guadiana	Estena	70	<u>AJ627300</u>	1	-	-	-
PYR 45	Guadiana	Estena	74	<u>AF045991</u>	4	-	-	-
PYR 46	Guadiana	Estena	78	<u>AF42I813</u>	3	-	-	-
PYR 47	Guadiana	Estena	78	<u>AJ627299</u>	1	-	-	-
PYR 48	Guadiana	Guadiana	-	-	*	PP	P2P2	<u>AY943877</u>
PYR 49	Guadiana	Guadiana	-	-	*	PP'	P7P29	strand P1; <u>DQI50331</u> ; strand P2; <u>DQI50332</u>
PYR 50	Guadiana	Guadiana	109	<u>EF459048</u>	*	PP	P7P7	<u>AY943878</u>
PYR 51	Guadiana	Guadiana	122	<u>EF459049</u>	*	-	-	-
PYR 52	Guadiana	Guadiana	127	<u>EF459050</u>	*	PP	P7P7	<u>DQI50281</u>
PYR 53	Guadiana	Guadiana	161	<u>EF459051</u>	*	-	-	-
PYR 54	Guadiana	Guadiana	21	<u>DQ263234</u>	5	PP	P7P7	<u>DQI50283</u>
PYR 55	Guadiana	Guadiana	21	<u>EF459044</u>	*	PP'	P3P11	strand P1; <u>EF459471</u> ; strand P2; <u>EF459472</u>
PYR 56	Guadiana	Guadiana	21	<u>EF459045</u>	*	-	-	-
PYR 57	Guadiana	Guadiana	21	<u>EF459046</u>	*	-	-	-
PYR 58	Guadiana	Guadiana	21	<u>EF459047</u>	*	-	-	-
PYR 59	Guadiana	Guadiana	39	<u>EF459042</u>	*	-	-	-
PYR 60	Guadiana	Guadiana	70	<u>EF459043</u>	*	-	-	-
PYR 61	Guadiana	Matachel	62	<u>AJ627301</u>	1	-	-	-
PYR 62	Guadiana	Oeiras	-	-	*	PP	P7P7	<u>DQI50282</u>
PYR 63	Guadiana	Oeiras	-	-	*	PP'	P7P28	strand P1; <u>DQI50329</u> ; strand P2; <u>DQI50330</u>
PYR 64	Guadiana	Oeiras	-	-	*	PP	P7P7	<u>DQI50279</u>
PYR 65	Guadiana	Oeiras	-	-	*	PP'	P7P30	strand P1; <u>DQI50327</u> ; strand P2; <u>DQI50328</u>
PYR 66	Guadiana	Oeiras	-	-	*	PP'	P7P29	strand P1; <u>DQI50325</u> ; strand P2; <u>DQI50326</u>
PYR 67	Guadiana	Oeiras	-	-	*	PP'	P3P11	strand P1; <u>DQI50323</u> ; strand P2; <u>DQI50324</u>
PYR 68	Guadiana	Oeiras	-	-	*	PP	P7P7	<u>DQI50274</u>

PYR 69	Guadiana	Oeiras	-	-	*	pp	P7P7	<u>DQ150273</u>	6
PYR 70	Guadiana	Oeiras	-	-	*	pp	P7P7	<u>DQ150272</u>	6
PYR 71	Guadiana	Oeiras	21	<u>DQ263235</u>	5	pp	P7P7	<u>DQ150280</u>	6
PYR 72	Guadiana	Oeiras	21	<u>EF459053</u>	*	pp	P7P7	<u>AY943879</u>	10
PYR 73	Guadiana	Oeiras	22	<u>DQ263236</u>	5	pp	P7P7	<u>DQ150278</u>	6
PYR 74	Guadiana	Oeiras	24	<u>DQ263237</u>	5	pp	P7P7	<u>DQ150277</u>	6
PYR 75	Guadiana	Oeiras	27	<u>DQ263239</u>	5	pp	P7P7	<u>DQ150275</u>	6
PYR 76	Guadiana	Oeiras	39	<u>DQ263238</u>	5	pp	P7P7	<u>DQ150276</u>	6
PYR 77	Guadiana	Oeiras	44	<u>EF459054</u>	*	pp	P7P7	<u>EF458732</u>	*
PYR 78	Guadiana	Ruidera	21	<u>AF421823</u>	3	-	-	-	-
PYR 79	Guadiana	Ruidera	21	<u>AJ627303</u>	1	-	-	-	-
PYR 80	Guadiana	Ruidera	29	<u>AF421822</u>	3	-	-	-	-
PYR 81	Guadiana	Ruidera	29	<u>AJ627302</u>	1	-	-	-	-
PYR 82	Guadiana	Sillo	73	<u>AJ627304</u>	1	-	-	-	-
PYR 83	Guadiana	Vascão	40	<u>EF459052</u>	*	-	-	-	-
PYR 84	Guadiana	Xévara	36	<u>AJ627305</u>	1	-	-	-	-
PYR 85	Guadiana	Zujar	35	<u>AJ627306</u>	1	-	-	-	-
PYR 86	Jucar	Jucar	98	<u>AF421807</u>	3	-	-	-	-
PYR 87	Jucar	Jucar	98	<u>EF459055</u>	*	-	-	-	-
PYR 88	Jucar	Jucar	99	<u>AF421806</u>	3	-	-	-	-
PYR 89	Lizandro	Cheleiros	127	<u>AJ427835</u>	1	-	-	<u>EF458733</u>	*
PYR 90	Lizandro	Lizandro	-	-	*	pp	P2P2	strand P1: <u>EF459475</u> , strand P2: <u>EF459476</u>	*
PYR 91	Lizandro	Lizandro	-	-	*	pp'	P2P24	<u>EF458734</u>	*
PYR 92	Lizandro	Lizandro	-	-	*	pp	P1P1	<u>EF458735</u>	*
PYR 93	Lizandro	Lizandro	-	-	*	pp	P2P2	-	*
PYR 94	Lizandro	Lizandro	127	<u>EF459063</u>	*	-	-	-	-
PYR 95	Lizandro	Lizandro	127	<u>EF459064</u>	*	-	-	-	-
PYR 96	Lizandro	Lizandro	127	<u>EF459065</u>	*	-	-	-	-
PYR 97	Lizandro	Lizandro	127	<u>EF459066</u>	*	-	-	-	-
PYR 98	Lizandro	Lizandro	127	<u>EF459067</u>	*	-	-	-	-
PYR 99	Lizandro	Lizandro	127	<u>EF459068</u>	*	-	-	-	-
PYR 100	Lizandro	Lizandro	127	<u>EF459069</u>	*	-	-	-	-
PYR 101	Lizandro	Lizandro	127	<u>EF459070</u>	*	-	-	-	-
PYR 102	Lizandro	Lizandro	130	<u>EF459074</u>	*	-	-	-	-
PYR 103	Lizandro	Lizandro	131	<u>EF459071</u>	*	-	-	-	-
PYR 104	Lizandro	Lizandro	131	<u>EF459072</u>	*	-	-	-	-
PYR 105	Lizandro	Lizandro	139	<u>EF459060</u>	*	-	-	-	-
PYR 106	Lizandro	Lizandro	139	<u>EF459061</u>	*	-	-	-	-
PYR 107	Lizandro	Lizandro	143	<u>EF459062</u>	*	-	-	-	-

181

PYR 147	Segura	Bogardá	48	AF421820	3	-	-	-	-	-
PYR 148	Tagus	Acebo	158	AF421827	3	-	-	-	-	-
PYR 149	Tagus	Acebo	159	AJ627310	1	-	-	-	-	-
PYR 150	Tagus	Alburrel	103	AJ627314	1	-	-	-	-	-
PYR 151	Tagus	Almonte	83	AF421791	3	-	-	-	-	-
PYR 152	Tagus	Almonte	86	AJ627315	1	-	-	-	-	-
PYR 153	Tagus	Alviela	-	-	*	PP	P2P2	EF458742	-	*
PYR 154	Tagus	Alviela	127	EF459100	*	PP'	PIP2	strand P1: AY943868 , strand P2: AY943869	10	
PYR 155	Tagus	Arrago	155	AF421826	3	-	-	-	-	-
PYR 156	Tagus	Arrago	156	AJ627311	1	-	-	-	-	-
PYR 157	Tagus	Aurela	84	AJ627316	1	-	-	-	-	-
PYR 158	Tagus	Caparro	153	AJ627312	1	-	-	-	-	-
PYR 159	Tagus	Cedena	92	AJ627318	1	-	-	-	-	-
PYR 160	Tagus	Cedena	93	AJ627317	1	-	-	-	-	-
PYR 161	Tagus	Cofio	100	AJ627319	1	-	-	-	-	-
PYR 162	Tagus	Vid	101	AJ627320	1	-	-	-	-	-
PYR 163	Tagus	Erges	-	-	*	PP	P1IP11	EF458743	-	-
PYR 164	Tagus	Erges	-	-	*	PP'	P3P11	strand P1: EF459485 , strand P2: EF459486	*	
PYR 165	Tagus	Erges	-	-	*	PP'	P5P11	strand P1: EF459487 , strand P2: EF459488	*	
PYR 166	Tagus	Erges	151	EF459101	*	-	-	-	-	-
PYR 167	Tagus	Erges	151	EF459102	*	-	-	-	-	-
PYR 168	Tagus	Erges	151	EF459103	*	PP'	P3P11	strand P1: AY943870 , strand P2: AY943871	10	
PYR 169	Tagus	Erges	151	EF459104	*	-	-	-	-	-
PYR 170	Tagus	Gevaldo	102	AJ627321	1	-	-	-	-	-
PYR 171	Tagus	Huso	97	AJ627322	1	-	-	-	-	-
PYR 172	Tagus	Jefe	150	AJ627313	1	-	-	-	-	-
PYR 173	Tagus	Maior	-	-	*	PP	P1IP11	EF458744	-	-
PYR 174	Tagus	Maior	-	-	*	PP'	P2P33	strand P1: EF459489 , strand P2: EF459490	*	
PYR 175	Tagus	Maior	-	-	*	PP	PIP1	EF458745	*	
PYR 176	Tagus	Maior	-	-	*	PP	P1IP11	EF458746	*	
PYR 177	Tagus	Maior	-	-	*	PP	PIP1	AY943875	10	
PYR 178	Tagus	Maior	145	EF459109	*	-	-	-	-	-
PYR 179	Tagus	Maior	146	EF459107	*	-	-	-	-	-
PYR 180	Tagus	Maior	146	EF459108	*	-	-	-	-	-
PYR 181	Tagus	Maior	147	EF459110	*	-	-	-	-	-
PYR 182	Tagus	Maior	174	EF459111	*	-	-	-	-	-
PYR 183	Tagus	Maior	175	EF459105	*	PP	PIP1	AY943874	10	
PYR 184	Tagus	Maior	175	EF459106	*	-	-	-	-	-
PYR 185	Tagus	Nabão	123	EF459113	*	PP	P2P2	EF458747	*	

PYR 186	Tagus	Nabão	125	<u>EF459112</u>	*	pp'	P2P11	strand P1: <u>EF459491</u> , strand P2: <u>EF459492</u>	*
PYR 187	Tagus	Nabão	127	<u>EF459114</u>	*	pp'	P3P35	strand P1: <u>EF459493</u> , strand P2: <u>EF459494</u>	*
PYR 188	Tagus	Ocreza	-	-	*	pp	P11P11	<u>AY943880</u>	10
PYR 189	Tagus	Ocreza	-	-	*	pp	P11P11	<u>EF458748</u>	*
PYR 190	Tagus	Ocreza	-	-	*	pp'	P2P33	strand P1: <u>EF459495</u> , strand P2: <u>EF459496</u>	*
PYR 191	Tagus	Ocreza	146	<u>EF459115</u>	*	-	-	-	-
PYR 192	Tagus	Pesqueiro	157	<u>AF421812</u>	3	-	-	-	-
PYR 193	Tagus	Pesqueiro	154	<u>AF421811</u>	3	-	-	-	-
PYR 194	Tagus	Pesqueiro	154	<u>AJ627323</u>	1	-	-	-	-
PYR 195	Tagus	Pesqueiro	157	<u>AJ627324</u>	1	-	-	-	-
PYR 196	Tagus	Serfã	116	<u>Y10131</u>	2	-	-	-	-
PYR 197	Tagus	Serfã	161	<u>AJ627325</u>	1	-	-	-	-
PYR 198	Tagus	Sever	116	<u>AJ627326</u>	1	-	-	-	-
PYR 199	Tagus	Sever	161	<u>Y10132</u>	2	-	-	-	-
PYR 200	Tagus	Sorraia	-	-	*	pp'	P2P31	strand P1: <u>EF459497</u> , strand P2: <u>EF459498</u>	*
PYR 201	Tagus	Sorraia	128	<u>AJ627327</u>	1	-	-	-	-
PYR 202	Tagus	Tietar	87	<u>AF045993</u>	4	-	-	-	-
PYR 203	Tagus	Tietar	88	<u>AJ627328</u>	1	-	-	-	-
PYR 204	Tagus	Trevijana	152	<u>AJ627329</u>	1	-	-	-	-
PYR 205	Tagus	Zêzere	127	<u>EF459123</u>	*	pp	P11P11	<u>AY943881</u>	10
PYR 206	Tagus	Zêzere	127	<u>EF459124</u>	*	pp	P11P11	<u>EF458755</u>	*
PYR 207	Tagus	Zêzere	161	<u>EF459125</u>	*	pp'	P11P17	strand P1: <u>AY943872</u> , strand P2: <u>AY943873</u>	10
PYR 208	Tagus	Zêzere	209	<u>EF459119</u>	*	pp	P11P11	<u>EF458752</u>	*
PYR 209	Tagus	Zêzere	209	<u>EF459120</u>	*	pp	P11P11	<u>EF458753</u>	*
PYR 210	Tagus	Zêzere	209	<u>EF459121</u>	*	pp	P11P11	<u>EF458754</u>	*
PYR 211	Tagus	Zêzere	209	<u>EF459122</u>	*	pp'	P2P11	strand P1: <u>EF459499</u> , strand P2: <u>EF459500</u>	*
PYR 212	Tagus	Zêzere	210	<u>EF459116</u>	*	pp	P11P11	<u>EF458749</u>	*
PYR 213	Tagus	Zêzere	211	<u>EF459117</u>	*	pp	P3P3	<u>EF458750</u>	*
PYR 214	Tagus	Zêzere	211	<u>EF459118</u>	*	pp	P11P11	<u>EF458751</u>	*
TOR 1	Mira	Mira	-	-	*	TT	P10P10	<u>DQ150284</u>	6
TOR 2	Mira	Mira	-	-	*	TT	P10P10	<u>DQ150285</u>	6
TOR 3	Mira	Mira	18	<u>EF459128</u>	*	TT	P10P10	<u>DQ150290</u>	6
TOR 4	Mira	Mira	18	<u>EF459129</u>	*	TT	P10P10	<u>DQ150289</u>	6
TOR 5	Mira	Mira	18	<u>EF459130</u>	*	TT	P10P10	<u>DQ150287</u>	6
TOR 6	Mira	Mira	18	<u>EF459131</u>	*	-	-	-	-
TOR 7	Mira	Mira	18	<u>EF459132</u>	*	-	-	-	-
TOR 8	Mira	Mira	18	<u>EF459133</u>	*	-	-	-	-
TOR 9	Mira	Mira	18	<u>EF459134</u>	*	TT'	P10P15	strand P1: <u>DQ150333</u> , strand P2: <u>DQ150334</u>	6
TOR 10	Mira	Mira	18	<u>EF459135</u>	*	-	-	-	-

TOR 11	Mira	18	<u>EF459136</u>	*	-	-	-	-
TOR 12	Mira	18	<u>EF459137</u>	*	-	-	-	-
TOR 13	Mira	18	<u>EF459138</u>	*	-	-	-	-
TOR 14	Mira	18	<u>EF459139</u>	*	-	-	-	-
TOR 15	Mira	18	<u>EF459140</u>	*	-	-	-	-
TOR 16	Mira	18	<u>EF459141</u>	*	-	-	-	-
TOR 17	Mira	19	<u>EF459126</u>	*	TT	P10P10	<u>DQ150288</u>	6
TOR 18	Mira	20	<u>EF459127</u>	*	TT	P10P10	<u>DQ150286</u>	6
VAL 1	Algar	188	<u>AF421818</u>	3	-	-	-	-
VAL 2	Algar	189	<u>AF421819</u>	3	-	-	-	-
VAL 3	Serpis	190	<u>AF421810</u>	3	-	-	-	-
VAL 4	Serpis	190	<u>AF421815</u>	3	-	-	-	-
VAL 5	Valencia Lagoon	186	<u>AF421809</u>	3	-	-	-	-
VAL 6	Valencia Lagoon	187	<u>AF421808</u>	3	-	-	-	-

Chapter 9

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

Sousa-Santos C, Collares-Pereira MJ & Almada V (2006)

Journal of Fish Biology 68 (Supplement B): 292-301

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

Sousa-Santos C, Collares-Pereira MJ & Almada V

Abstract We report the first occurrence of massive mitochondrial introgression of *Squalius aradensis* genes in *S. alburnoides*, a hybridogenetic complex that usually carries mtDNA of its maternal ancestor (*S. pyrenaicus*), and discuss the possible implications of such introgressions for the history of the complex.

Keywords beta-actin; cytochrome *b*; hybridogenetic complex; mitochondrial introgression; *Squalius alburnoides*, *Squalius aradensis*

Although still lacking a formal generic recognition, the cyprinid species of *Squalius*, previously included in the genus *Leuciscus*, form a very well defined monophyletic clade only distantly related to *Leuciscus leuciscus* (Briolay *et al.* 1998), within which the Iberian species form a strongly supported monophyletic group (Zardoya & Doadrio 1999; Sanjurjo *et al.* 2003). The species of *Squalius* have attracted considerable attention from researchers studying fish hybridization since the group provided several instances of formation of hybridogenetic lineages [*e.g.* *Squalius alburnoides* (Alves *et al.* 1997a) and possibly *Squalius palaciosi* (Zardoya & Doadrio 1998)], extensive interspecific introgression (*e.g.* Durand *et al.* 2000) and intergeneric hybridization (*e.g.* Ünver & Erk'akan 2005; Freyhof *et al.* 2005; Robalo *et al.* 2006). Since its first description in the 80's (Collares-Pereira 1984, 1985), great advances have been made to understand the origin and maintenance of the *S. alburnoides* (Steindachner) hybridogenetic complex, which comprises diploid (2n=50), triploid (3n=75) and tetraploid (4n=100) hybrid forms (reviewed in Alves *et al.* 2001). This complex of widely distributed minnows seems to have resulted from unidirectional interspecific crosses between *S. pyrenaicus* (Günther) females (P genome) and males from an unknown (and possibly extinct) species - A genome (reviewed in Alves *et al.* 2001). The peculiar mechanisms of gamete formation in the

complex also lead to the reconstitution of diploid “nuclear nonhybrids” exhibiting *S. pyrenaicus*-like mtDNA and the nuclear genome of the missing paternal ancestor (Alves *et al.* 2002). This nonhybrid form, constituted almost entirely by males (females are inexistent or extremely rare), has only been found in Tagus, Sado and Guadiana drainages (see also Cunha *et al.* 2004; Pala & Coelho 2005) – numbered 3, 4 and 5, respectively, in Figure 1.

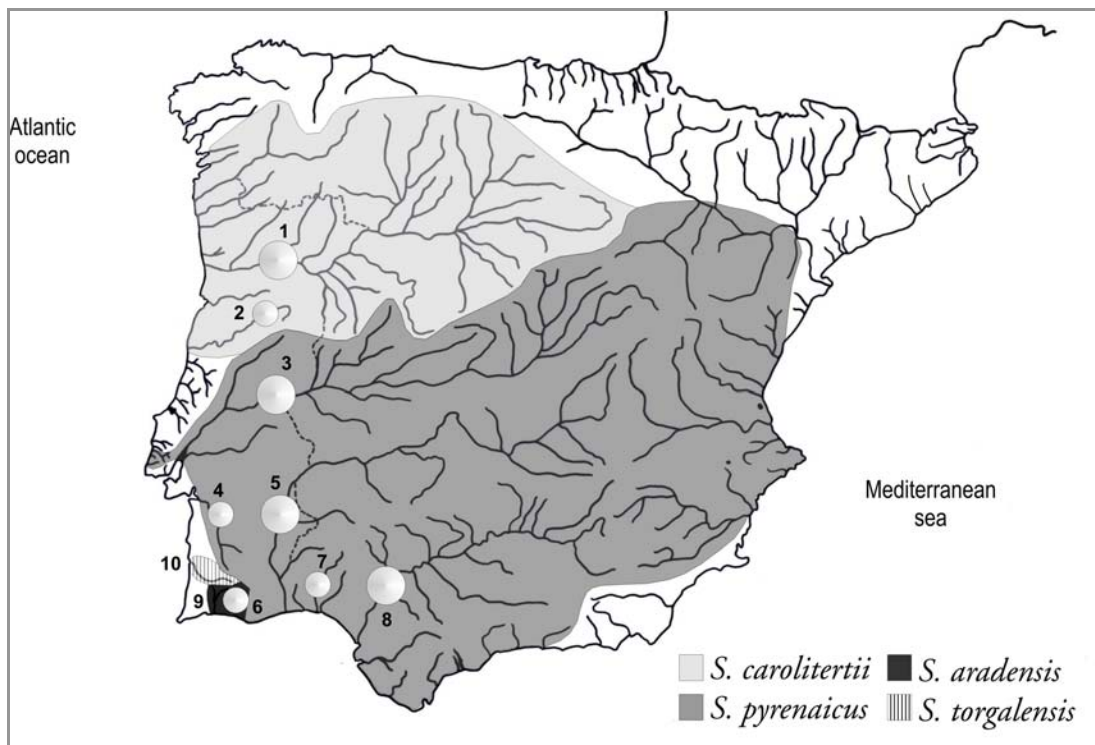


Figure 1 – Distribution areas of the Iberian endemic minnows *S. carolitertii*, *S. pyrenaicus*, *S. torgalensis* and *S. aradensis*. The numbered circles represent the already analysed river basins for *S. alburnoides*. 1 – Douro; 2 – Mondego; 3 – Tagus; 4 – Sado; 5 – Guadiana; 6 – Quarteira; 7 – Odier; 8 – Guadalquivir. River basins numbered 9 and 10 are, respectively, Arade and Mira.

In the drainages to the north of the Tagus River, *S. pyrenaicus* is absent but *S. alburnoides* is sympatric with another *Squalius* species – *S. carolitertii* (Doadrio) (Figure 1). In such river basins, the mtDNA of *S. alburnoides* is *S. pyrenaicus*-like (Alves *et al.* 1997b, 2002; Cunha *et al.* 2004; Pala & Coelho 2005), with only one exception reported by Alves *et al.* (1997b): a specimen with *S. carolitertii*-like mtDNA. As *S. alburnoides* males are much less common than females, it has been assumed that the maintenance of the complex depends mainly on unidirectional crosses between *S. alburnoides* females and *S. pyrenaicus*/*S. carolitertii* males (Alves *et al.* 2001; Cunha *et al.* 2004; Pala & Coelho

2005). The almost exclusive presence of *S. pyrenaicus* mtDNA in *S. alburnoides* has been interpreted as indicating that crosses between *S. alburnoides* males and females of other species are absent or extremely uncommon, except at the origin of the complex (Alves *et al.* 2001). Based on the pattern of variation of the *cyt b* gene, Alves *et al.* (1997b) postulated that the complex had at least two distinct hybridogenetic origins - one in the Sado drainage and the other in the Guadiana and Tagus drainages - and that it had dispersed from the Tagus into the northern drainages. More recently Cunha *et al.* (2004), using a broader data set, postulated five independent hybridization origins based on the fact that the mtDNA of *S. alburnoides* had a stronger affinity with that of *S. pyrenaicus* from the same river than with conspecifics from other drainages. In Quarteira, a small independent drainage in southwest Portugal, *S. alburnoides* is sympatric with another *Squalius* species, *S. aradensis* (Coelho, Bogutskaya, Rodrigues & Collares-Pereira), an endemic fish with a very restricted range (Figure 1). Apart from the Quarteira drainage, it only occurs in a few distinct independent drainages, mainly in that of the Arade River (Figure 1) which was recently considered as the evolutionary centre of origin of the species (Mesquita *et al.* 2005). The Quarteira drainage represents the western limit of the geographical distribution of *S. alburnoides* and the eastern limit of *S. aradensis* (Mesquita & Coelho 2002).

In the present paper we report for the first time the occurrence of natural hybridization between *S. alburnoides* and *S. aradensis* and demonstrate the occurrence of an extensive introgression of *S. aradensis* genes into *S. alburnoides* in the Quarteira drainage, specially marked at the mitochondrial level. This observation contradicts the current views on the unidirectionality of interspecific crosses involving *S. alburnoides*. We also show that the crosses between *S. alburnoides* and *S. aradensis* reconstitute nonhybrid *S. alburnoides* males with the nuclear genome of the missing ancestor of the complex. Finally, the occurrence of fish bearing *S. aradensis* and *S. pyrenaicus* nuclear genomes is also reported.

Live *S. alburnoides* individuals are easily distinguished in the field by the presence of a black line surrounding the base of the dorsal fin, which is absent in all other Iberian *Squalius* species (C. Sousa-Santos *unpublished*). This distinction was confirmed in *S. alburnoides* from all the drainages where it occurs, after comparing hundreds of fish with specimens of all other *Squalius* species of Portuguese freshwaters. In addition,

S. alburnoides fish show a conspicuous longitudinal dark band above the lateral line (Collares-Pereira 1984) that is absent in the remaining Iberian *Squalius*.

A small fin clip was taken from *S. alburnoides* and *S. aradensis* from Quarteira and the fish were safely returned to the water. Previous experience with all the cyprinid fish that underwent this procedure and were temporally kept in aquaria for other purposes demonstrated that the procedure described above does cause neither mortality nor diseases. Total genomic DNA was extracted from fin clips preserved in ethanol by a SDS/proteinase-k based protocol, precipitated with isopropanol and washed with ethanol before re-suspension in water (adapted from Sambrook *et al.* 1989).

The amplification process was conducted as follows: 35 cycles of [94°C (30sec.), 55°C (40sec.) and 72°C (1min30sec)]. Each sample was sequenced in both directions with the same primers used for PCR. Sequences were aligned with BioEdit® v.5.0.6.

A total of 1123bp of the *cytb* gene of samples of 20 *S. alburnoides* (GenBank: DQ003238-42, DQ003246, DQ003253-55, DQ003259, DQ14179-88) and of 19 *S. aradensis* (GenBank: DQ003243-45, DQ003247-52, DQ003256-58, DQ14189-94, DQ145177) from Quarteira were amplified using the primers LCB1 (Brito *et al.* 1997) and HA (Schmidt & Gold 1993).

For comparisons of the *cytb* gene with that of other *Squalius* species we used samples available in Genbank: *S. pyrenaicus* from Sado (Y10133), from Guadiana (Y10134, AF421822-23, AF421813-14, AF421804-05, AF045991), from Guadalquivir (AF421816-17, AF421790) and from Tagus (Y10131-132, AF421811-12, AF421826-27, AF421791, AF045993). Additional samples of *S. aradensis* from Arade (AF421824-25) and *S. torgalensis* (Coelho, Bogutskaya, Rodrigues & Collares-Pereira) (Z75929) from Mira were also included - *S. alburnoides* is absent from both these drainages. A segment of 927bp of the nuclear beta-actin gene was also amplified from samples of 20 *S. aradensis* (GenBank: DQ150307-22 and DQ150260-71) and 15 *S. alburnoides* (GenBank: DQ150335-61, DQ128102, AY943867 and AY943865) from Quarteira (for primers see Sousa-Santos *et al.* in press¹). This fragment is homologous to a region of the beta-actin gene of *Cyprinus carpio* (GenBank: M24113) between the positions 1622 and 2550, including introns B and C and three exons. The methods used for recovering the parental sequences of nuclear genes of hybrids were described elsewhere (see Sousa-Santos *et al.* in press). As there

¹ Updated citation: Sousa-Santos *et al.* (2005). *DNA Sequence* 16:462-467.

were no sequences of the beta-actin gene available in GenBank, we also sequenced DNA from samples of 20 *S. pyrenaicus* (GenBank: AY943877-79, DQ150272-83 and DQ150323-332) and of four nuclear nonhybrid *S. alburnoides* from Guadiana (GenBank: AY943892-93, AY943864, DQ010337); of four nuclear nonhybrid *S. alburnoides* from Tagus (GenBank: AY943863, AY943894-96); of eight *S. aradensis* from Arade (DQ150291-306); and of eight *S. torgalensis* from Mira (GenBank: DQ150284-290 and DQ150333-34) - see Figure 1 for river locations. Note that there are more accession numbers than individuals since some fishes are heterozygous for the analysed fragment of the beta-actin gene and each strand has its own accession number. For a description of a procedure that allows the identification of the two different sequences involved in heterozygotes and hybrids when analysing nuclear gene sequences see Bhangale *et al.* (2005).

The sequencing of beta-actin gene yielded three groups of haplotypes that showed characteristic patterns of mutations and indels that were sufficiently distinct to be identifiable when hybrid genomes were analysed. Based on these characteristic patterns, each individual fish could be classified as showing A, Q or P genome or other possible combinations: Q+A, P+A or P+Q (this notation does not denote the ploidy of the fish but merely what types of genomes were present) (Table I). "Q haplotypes" were found in the fish morphologically diagnosed as *S. aradensis* from Quarteira, in the specimens of *S. aradensis* from Arade and in the specimens of *S. torgalensis* from Mira; "P haplotypes" were observed in the samples of *S. pyrenaicus* from Guadiana and in one *S. alburnoides* individual from Quarteira; and "A haplotypes" were found in homozygosity in specimens morphologically diagnosed as *S. alburnoides* nonhybrids (five from Quarteira, four from Tagus and four from Guadiana). All other fish from Quarteira morphologically diagnosed as *S. alburnoides* hybrids showed hybrid genomes with a combination of A and P/Q haplotypes.

The average percentage of divergence between all six Q haplotypes found in fishes from Quarteira (N=28), Arade (N=8) and Mira (N=8) was 0.21 ± 0.08 . Of these six Q haplotypes, three were identified in fish from Quarteira, one being specific to that drainage and the other two being shared with Arade. The average percentage of divergence between the seven distinct A haplotypes that were found in fishes from Quarteira (N=15), Guadiana (N=4) and Tejo (N=4) was 0.26 ± 0.13 . *S. alburnoides* from Quarteira exhibited four of these six A haplotypes (two of them being specific to that

drainage, one shared with Guadiana and the other shared with Guadiana and Tejo). Nine different P haplotypes were recovered from fishes from Quarteira (N=1) and Guadiana (N=20), with an average percentage of divergence of 0.35 ± 0.16 between them. One triploid *S. alburnoides* from Quarteira showed two distinct P haplotypes, one of them exclusive of that river and the other shared with Guadiana. The average percentage of divergence between the three types of nuclear haplotypes was: 0.73 ± 0.16 between P and Q, 2.81 ± 0.13 between Q and A and 3.03 ± 0.77 between P and A.

Inspection of Table I clearly shows that all fish morphologically diagnosed as *S. alburnoides* exhibited either pure A or hybrid beta-actin genomes. All fish morphologically diagnosed as *S. aradensis* showed Q genomes, except one with a P+Q genome, indicative of a history of hybridization between *S. aradensis* and *S. pyrenaicus*. The *S. aradensis* from Arade showed Q genomes that helped to confirm the identification of *S. aradensis* from Quarteira. Thus, the morphological distinction between *S. alburnoides* and *S. aradensis* was fully supported by the beta-actin data. Of the ten hybrid *S. alburnoides*, nine showed Q+A combinations indicating that they had a history of hybridization between *S. alburnoides* and *S. aradensis*, while one showed a P+A constitution indicative of the incorporation of a *S. pyrenaicus* haplotype in its ancestry. These results clearly showed that in the Quarteira drainage the DNA of *S. aradensis* introgressed massively into *S. alburnoides*.

The sequencing of the *cytb* gene of *S. alburnoides* (both hybrids and nonhybrids) and of *S. aradensis* from Quarteira yielded nine *S. aradensis*-like mitochondrial haplotypes distinguished by one to seven mutations (group I), and one *S. pyrenaicus* mtDNA haplotype found in one *S. alburnoides* which differed from the others by 111 to 116 mutational steps (haplotype PYR) – see Figure 2 (haplotype network performed using TCS 1.21[®], Clement *et al.* 2000).

When compared with the *cytb* sequences of other *Squalius* species (Figure 2), the haplotypes included in group I were phylogenetically closer to *S. aradensis*, while the haplotype PYR had more affinities with *S. pyrenaicus* being closer to haplotypes from Guadiana (mean percentage of divergence between pairs of haplotypes of 0.83 ± 0.19 for Guadiana, 1.48 ± 0.45 for Guadalquivir, 1.77 ± 0.29 for Tagus, and 2.67 for Sado specimens).

Table 1 – Morphological identification, mitochondrial and nuclear haplotypes of the specimens analysed from River Quarteira.

Morphological identification	Mitochondrial haplotype	Nuclear haplotype
<i>S. alburnoides</i> hybrid 1	<i>S. aradensis</i>	Q+A
<i>S. alburnoides</i> hybrid 2	<i>S. aradensis</i>	Q+A
<i>S. alburnoides</i> hybrid 3	<i>S. aradensis</i>	(not sequenced)
<i>S. alburnoides</i> hybrid 4	<i>S. aradensis</i>	(not sequenced)
<i>S. alburnoides</i> hybrid 5	<i>S. aradensis</i>	(not sequenced)
<i>S. alburnoides</i> hybrid 6	<i>S. aradensis</i>	(not sequenced)
<i>S. alburnoides</i> hybrid 7	<i>S. aradensis</i>	Q+A
<i>S. alburnoides</i> hybrid 8	<i>S. aradensis</i>	P+A
<i>S. alburnoides</i> hybrid 9	<i>S. pyrenaicus</i>	Q+A
<i>S. alburnoides</i> hybrid 10	<i>S. aradensis</i>	(not sequenced)
<i>S. alburnoides</i> hybrid 11	<i>S. aradensis</i>	Q+A
<i>S. alburnoides</i> hybrid 12	<i>S. aradensis</i>	Q+A
<i>S. alburnoides</i> hybrid 13	<i>S. aradensis</i>	Q+A
<i>S. alburnoides</i> hybrid 14	<i>S. aradensis</i>	Q+A
<i>S. alburnoides</i> hybrid 15	<i>S. aradensis</i>	Q+A
<i>S. alburnoides</i> nonhybrid 16	<i>S. aradensis</i>	A
<i>S. alburnoides</i> nonhybrid 17	<i>S. aradensis</i>	A
<i>S. alburnoides</i> nonhybrid 18	<i>S. aradensis</i>	A
<i>S. alburnoides</i> nonhybrid 19	<i>S. aradensis</i>	A
<i>S. alburnoides</i> nonhybrid 20	<i>S. aradensis</i>	A
<i>S. aradensis</i> 1	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 2	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 3	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 4	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 5	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 6	(not sequenced)	Q
<i>S. aradensis</i> 7	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 8	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 9	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 10	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 11	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 12	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 13	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 14	<i>S. aradensis</i>	P+Q
<i>S. aradensis</i> 15	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 16	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 17	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 18	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 19	<i>S. aradensis</i>	Q
<i>S. aradensis</i> 20	<i>S. aradensis</i>	Q

The haplotype network (Figure 2) strongly suggests that in the history of the *S. alburnoides* population of Quarteira several hybridization events involving females of *S. aradensis* took place. Indeed, if a single cross had taken place at the origin of the complex, with no subsequent involvement of *S. aradensis* females, a very different

haplotype network was to be expected, with the haplotypes of *S. alburnoides* and *S. aradensis* clustering in two well differentiated clades.

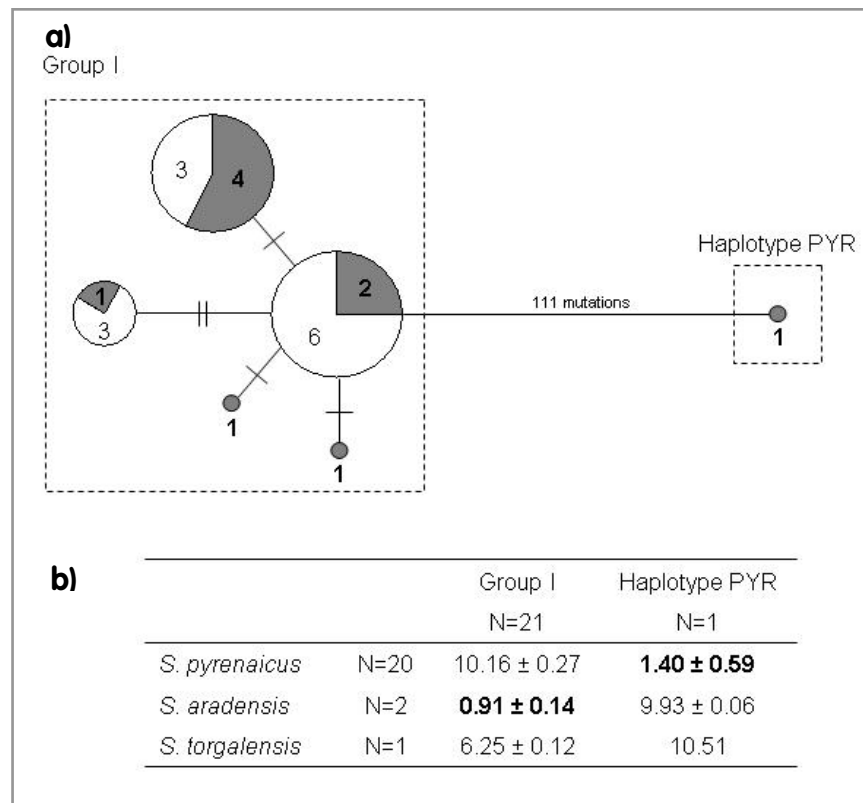


Figure 2 — (a) Haplotype network of *S. alburnoides* and *S. aradensis* (N=39) from the river Quarteira. Each haplotype is represented by a circle, proportional to the number of individuals of *S. alburnoides* (in grey) and *S. aradensis* (in white) that share that haplotype (number of individuals indicated in arabic). The number of mutations between haplotypes is represented by the number of small lines perpendicular to the branches linking haplotypes. (b) Mean percentage of divergence (“p” distance) and respective standard deviations between pairs of haplotypes from Quarteira and from *S. pyrenaicus*, *S. torgalensis* and *S. aradensis*.

The analysis of molecular variance (AMOVA, Excofier *et al.* 1992), performed using Arlequin 2.0[®] (Schneider *et al.* 1999), showed no clear distinction between the *cytb* gene sequences of all *S. aradensis* and *S. alburnoides* from Quarteira: 1.09% of variation among species and 98.91% of variation within species. Neither AMOVA nor the exact test of population differentiation based on haplotype frequencies yielded significant results (F_{ST} value=0.0109 with $p=0.1896\pm0.0116$, and $p=0.2003\pm0.0064$, respectively). The AMOVA results were also consistent with the hypothesis of a history with multiple hybridization events involving *S. aradensis* females and *S. alburnoides* males. Finally, the finding that the five *S. alburnoides* with pure A nuclear genomes have *S. aradensis* mtDNA suggests

that they were reconstituted from hybrids and that crosses involving *S. aradensis* females occurred in their ancestry. The beta-actin sequences of these fishes were identical to the sequences found in other nuclear nonhybrid males of *S. alburnoides*: mean percentage of divergence between pairs of haplotypes of 0.19 ± 0.16 and 0.29 ± 0.15 for fishes with pure A nuclear genomes from Guadiana and from Tagus drainages, respectively.

The results of the present study demonstrate that in the Quarteira drainage *S. alburnoides* suffered massive introgression from both nuclear and mitochondrial genes of *S. aradensis* and suggest that probably many crossings involving *S. aradensis* females occurred in the history of this population.

An ancient connection between the Quarteira and Guadiana river basins would have allowed the passage of *S. pyrenaicus*-like mtDNA carriers and their subsequent crossing with *S. aradensis*. This massive introgression could have been promoted by a disproportionately low number of colonizers and/or by behavioural/ecological preferences that favoured crosses with *S. aradensis* females. This does not exclude the possibility of other crosses, namely between *S. alburnoides* females (which tend to be the most abundant sex in the populations – Alves *et al.* 2001) and males of other *Squalius* species. The maternal inheritance of mtDNA means, however, that a massive introgression of mtDNA of *S. aradensis* in *S. alburnoides* of Quarteira must have been achieved through crosses between males of *S. alburnoides* and females of *S. aradensis*. As in the Guadiana the great majority of *S. alburnoides* males are nuclear nonhybrids (Alves *et al.* 2001) it is probable that these were the *S. alburnoides* males most frequently involved in the crosses with *S. aradensis* females.

It is also interesting to note that the reconstituted nuclear nonhybrids seem to be inexistent or nearly absent in the northern rivers (Pala & Coelho 2005) but were detected in Quarteira with a considerable frequency (five out of 15 individuals). These nonhybrid individuals, almost always males, may be the key to understand the synthesis of new hybrid lineages through crosses with *S. aradensis* females.

Observations on the reproductive behaviour in captivity (C. Sousa-Santos *unpublished*) showed that i) these nonhybrid males assume a sneaking behaviour in the presence of courting pairs of *S. pyrenaicus* and ii) actively spawned and fertilized the eggs of females of *S. torgalensis* (a species closely related to *S. aradensis* – Brito *et al.* 1997; Sanjur *et al.* 2003; Mesquita *et al.* 2005) in the absence of conspecific males. The

presence of nonhybrid males in Quarteira might have, therefore, promoted the expansion of the complex, contributing to a rapid replacement of the typical *S. pyrenaicus* mtDNA lineage.

An independent origin of the complex in Quarteira (assuming that the paternal ancestor existed in this drainage and went extinct), as postulated by Cunha *et al.* (2004), although not disproved by this study seems unlikely since we detected the presence of one individual with a *S. pyrenaicus*-like mtDNA and one individual with *S. pyrenaicus*-like *beta-actin* sequences although possessing *S. aradensis* mtDNA. Cunha *et al.* (2004) found that in several drainages the mtDNA of *S. alburnoides* is more similar to the mtDNA of other sympatric *Squalius* species than to the mtDNA of *S. alburnoides* from other drainages. It is interesting to note that this phenomenon is only apparent in the drainages where nuclear nonhybrid males have been reported – Tagus, Sado and Guadiana. Conversely to their interpretation and postulation of multiple independent origins for *S. alburnoides*, these similarities in mtDNA between *S. alburnoides* and *S. pyrenaicus* may have been promoted by frequent crosses between *S. alburnoides* nuclear nonhybrid males and local *S. pyrenaicus* females. Indeed, when there are no such males in the populations, as apparently happens in Douro and Mondego (Alves *et al.* 2001; Pala & Coelho 2005), it is expected that the most frequent crosses involve triploid *S. alburnoides* females and the most abundant males (either *S. carolitertii* or diploid hybrid *S. alburnoides*), explaining the high frequency of individuals with *S. pyrenaicus*-like mtDNA in these drainages (Pala & Coelho 2005).

The available evidence is consistent with a single origin for the *S. alburnoides* complex, whose dispersal caused by historical changes of the Iberian hydrographical network would have allowed multiple crosses with the different *Squalius* species found in the newly colonized drainages. This scenario is more parsimonious than admitting the coexistence of the ancestors of the complex in multiple river basins, the independent synthesis of the complex in each of those basins, and the subsequent extinction of one or both the ancestors depending on the river basin considered. However, plausibility is not equivalent to definitive proof and further studies are needed to a more accurate evaluation of the different hypotheses proposed to explain the history of this interesting complex.

Acknowledgements The authors wish to thank G. Lemos and Sousa-Santos family for help in sample collection, J. Robalo for laboratory help, and an anonymous referee for valuable comments and suggestions. Direcção Geral das Florestas provided permissions for field work. The study was funded by the FCT Pluriannual Program (U I&D 331/94 and U I&D 329/94) (FEDER participation). C. Sousa-Santos was supported by a PhD grant from FCT (SFRH/BD/8320/2002).

References

- Alves MJ, Collares-Pereira MJ, Dowling TE & Coelho MM. 2002. The genetics of maintenance of an all-male lineage in the *Squalius alburnoides* complex. *Journal of Fish Biology* 60: 649-662.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 1997a. The *Rutilus alburnoides* complex (Cyprinidae): evidence for hybrid origin. *Journal of Zoological Systematics and Evolutionary Research* 35: 1-10.
- Alves MJ, Coelho MM, Collares-Pereira MM & Dowling TE. 1997b. Maternal ancestry of the *Rutilus alburnoides* complex (Teleostei, Cyprinidae) as determined by analysis of cytochrome *b* sequences. *Evolution* 51: 1584-1592.
- Bhangale TR, Rieder MJ, Livingston RJ & Nickerson DA. 2005. Comprehensive identification and characterization of diallelic insertion-deletion polymorphisms in 330 human candidate genes. *Human Molecular Genetics* 14: 59-69.
- Brito RM, Briolay J, Galtier N, Bouvet Y & Coelho MM. 1997. Phylogenetic relationships within genus *Leuciscus* (Pisces, Cyprinidae) in Portuguese freshwaters, based on mitochondrial cytochrome *b* sequences. *Molecular Phylogenetics and Evolution* 8: 435-442.
- Clement M, Posada D & Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9: 1657-1660.
- Collares-Pereira MJ. 1984. The "*Rutilus alburnoides* (Steindachner, 1866) complex" (Pisces, Cyprinidae). I. Biometrical analysis of some Portuguese populations. *Arquivos do Museu Bocage (Série A)* II: 111-143.
- Collares-Pereira MJ. 1985. The "*Rutilus alburnoides* (Steindachner, 1866) complex". II. First data on the karyology of a well-established diploid-triploid group. *Arquivos do Museu Bocage (Série A)* III: 69-90.
- Cunha C, Coelho MM, Carmona JA & Doadrio I. 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Molecular Ecology* 13: 2807-2817.
- Durand JD, Ünlü E, Doadrio I & Pipoyan S. 2000. Origin, radiation, dispersion and allopatric hybridization in the chub *Leuciscus cephalus*. *Proceedings of the Royal Society of London Series B* 267: 1687-1697.
- Excoffier L, Smouse PE & Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.
- Freyhof J, Lieckfeldt D, Pitra C & Ludwig A. 2005. Molecules and morphology: evidence for introgression of mitochondrial DNA in Dalmatian cyprinids. *Molecular Phylogenetics and Evolution* 37: 347-354.

Mesquita N & Coelho MM. 2002. The ichthyofauna of the small Mediterranean-type drainages of Portugal: its importance for conservation. *In* Collares-Pereira MJ, Cowx IG & Coelho MM (Eds.) Conservation of Freshwater Fishes: Options for the Future. Fishing News Books, Blackwell Science, Oxford, pp 65-71.

Mesquita N, Hänfling B, Carvalho GR & Coelho MM. 2005. Phylogeography of the cyprinid *Squalius aradensis* and implications for conservation of the endemic freshwater fauna of southern Portugal. *Molecular Ecology* 14: 1939-1954.

Pala I & Coelho MM. 2005. Contrasting views over a hybrid complex: between speciation and evolutionary "dead-end". *Gene* 347: 283-294.

Robalo JI, Sousa-Santos C, Levy A & Almada VC. 2006. Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. *Molecular Phylogenetics and Evolution* 39(1): 276-281.

Sambrook J, Fritsch EF & Maniatis T. 1989. *Molecular cloning: a laboratory manual*. Cold Spring Harbor, New York.

Sanjur OI, Carmona JA & Doadrio I. 2003. Evolutionary and biogeographical patterns within Iberian populations of the genus *Squalius* inferred from molecular data. *Molecular Phylogenetics and Evolution* 29: 20-30.

Schmidt TR & Gold JR. 1993. Complete sequence of the mitochondrial cytochrome *b* gene in the Cherryfin Shinner, *Liturus roseipinnis* (Teleostei: Cyprinidae). *Copeia* 3: 880-883.

Schneider S, Roessli D & Excofier L. 1999. Arlequin version 2.0: a software for genetic data analysis. University of Geneva, Genetics and Biometry Laboratory.

Sousa-Santos C, Robalo J, Collares-Pereira MJ & Almada V. in press. Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genomic composition of hybrid organisms. *DNA Sequence*

Ünver B & Erk'akan F. 2005. A natural hybrid of *Leuciscus cephalus* (L.) and *Chalcalburnus chalcoides* (Güldenstädt) (Osteichthyes: Cyprinidae) from Lake Tödürge (Sivas, Turkey). *Journal of Fish Biology* 66: 899-910.

Zardoya R & Doadrio I. 1998. Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. *Proceedings of the Royal Society of London Series B* 265: 1365-1372.

Zardoya R & Doadrio, I. 1999. Molecular evidence on the evolutionary and biogeographical patterns of European Cyprinids. *Journal of Molecular Evolution* 49: 227-237.

PART **V**

DISCUSSION AND CONCLUDING REMARKS

Chapter 10

General Discussion

The results presented in the Chapters 2 to 9 were already subjected to specific discussions, concerning various aspects related to the reproduction of *S. alburnoides* that followed the objectives drawn for this thesis. In this Chapter, these main results and conclusions will be integrated in an attempt to achieve a holistic discussion on the established goals: the reproductive behaviour and evolutionary history of the *S. alburnoides* complex.

This general discussion will thus be structured as an overview of the evolutionary history of *S. alburnoides*, from its origin to its dispersal and, finally, to its present day dynamics. For each one of those stages, the contribution of the reproductive behaviour to the dynamics of the complex will be discussed, based on the main achievements. However, this overview will be preceded by the discussion of the developed method for the assessment of genomic constitutions (**Chapter 2**), as its application was fundamental for the ethological, morphological and phylogeographical researches conducted.

10.1 Identification of inherited genomes

Although the analysis of the mtDNA of *S. alburnoides* and the use of allozymes had been crucial for the understanding of several aspects of the biology of the complex, revealing the gametogenesis of several of its forms and contributing to the comprehension of the formation of the various genotypes (Alves *et al.* 1996, 1997a, 1997b, 1998, 1999, 2001, 2002, 2004; Carmona *et al.* 1997; Crespo-López *et al.* 2006), the inexistence of suitable nuclear markers precluded the determination of individual genotypes, of paternally inherited genomes or of male fertilization rates.

A tentative identification of *S. alburnoides* genotypes using allozymes was first conducted by Carmona *et al.* (1997), who proposed that triploid and tetraploid genome constitutions could be inferred from gene dosage patterns when analysing the staining intensities of electromorphs. With this methodology, these authors described the presence of two types of triploids in Douro and Guadiana river basins. However, Alves *et al.* (1997a), using the same methodology and analysing seven instead of three loci with virtually fixed heterozygosity, failed to detect dosage-related patterns.

More recently, Pala & Coelho (2005) were successful in finding microsatellite loci with diagnostic alleles for both *Anaocypris*-like (A) and *S. carolitertii* (C) nuclear genomes, thus enabling the identification of *S. alburnoides* genotypes but only in the northern populations.

The description of a new method based on the presence of double peaks in chromatograms generated by heterozygous indels after beta-actin gene sequencing (**Chapter 2**) configures the molecular tool that was missing and it may be applied to all *S. alburnoides* populations.

In recent years, the amplification of nuclear genes combined with modern automated sequencers and more reliable software for chromatogram analysis made possible the detection of single nucleotide polymorphisms (SNP's): positions within a DNA sequence that vary by a single nucleotide substitution, deletion or insertion. SNP's have enormous potential for studies of evolutionary and population genetics, including the characterization of recombination rates, population origins and gene flow (Belfiore *et al.* 2003). Currently, the analysis of SNP's is a popular method for the investigation of human

genetic variation (*e.g.* Takatani *et al.* 2001; Takeuchi *et al.* 2001; Liguori *et al.* 2003; Toone *et al.* 2003; Sykiotis *et al.* 2003; Campo *et al.* 2004; Linch *et al.* 2004; Malvagias *et al.* 2004), aiding research into areas such as population genetics, genetic disease and drug development (French *et al.* 2001).

In spite of having lower variation compared with microsatellites, SNP's should make the comparison of genomic diversity and history of different species (the core goal of "comparative biogeography") more straightforward than has been possible with microsatellites (Brumfield *et al.* 2003). The use of SNP's is advantageous, when compared to that of microsatellites, since they are more prevalent in the genome and some may directly affect protein structure or expression levels (which is useful for the determination of the genetic mechanisms of diseases), among others (Landegren *et al.* 1998). In addition, and contrary to microsatellites for which models of evolution are still unclear, SNP's can usually be amenable to the identification of ancestral and derived states, making them much more informative in phylogenetic and phylogeographical studies (Schlötterer 2004).

In the presence of many heterozygous loci, several algorithms may be applied to ascribe each base to the corresponding sequence, in order to reconstruct the two haplotypes involved in the heterozygote (*e.g.* Clark 1990; Excoffier & Slatkin 1995; Stephens *et al.* 2001; Niu *et al.* 2002; Zhang *et al.* 2003). However, these haplotype inference methods can be replaced by a simpler approach when sequences having indels in heterozygous condition are analyzed (Bhangale *et al.* 2005). This situation is responsible for a disturbance in the sequence process that generates a series of double peaks that can be confused with contamination or "garbage" in the chromatograms. However, they have great potential for haplotype reconstruction because, if properly analysed, a single indel generates a series of double peaks that allow the correct assignment of each base to the corresponding sequence. Indeed, the presence of heterozygous indels and the overlap of bases out of phase for as many positions as the size of the indel proved to be useful to attain two objectives: 1) the isolation and the recovering of the parental sequences involved in the formation of a heterozygote or interspecific hybrid; and 2) to complement the determination of ploidy obtained by other methods and identify the relative contributions of the parental sequences in non-diploid hybrids (**Chapter 2**).

In the case of *S. alburnoides*, the process of reconstruction of the parental haplotypes was favoured by the existence of several specific point mutations characteristic

of each parental species. These mutations marked out highly conserved regions and made possible to ascribe unambiguously each peak in a double peaks region to the correct parental genome.

In diploids, the identification was simple since there were only two genomes involved. If one of the two bases is characteristic of one of the genomes for that site and the other is unknown it is reasonable to ascribe the mutated base to the other genome (Figure 1). This criterion will fail only in the unlikely situation of having two genomes harbouring mutations at the same site.

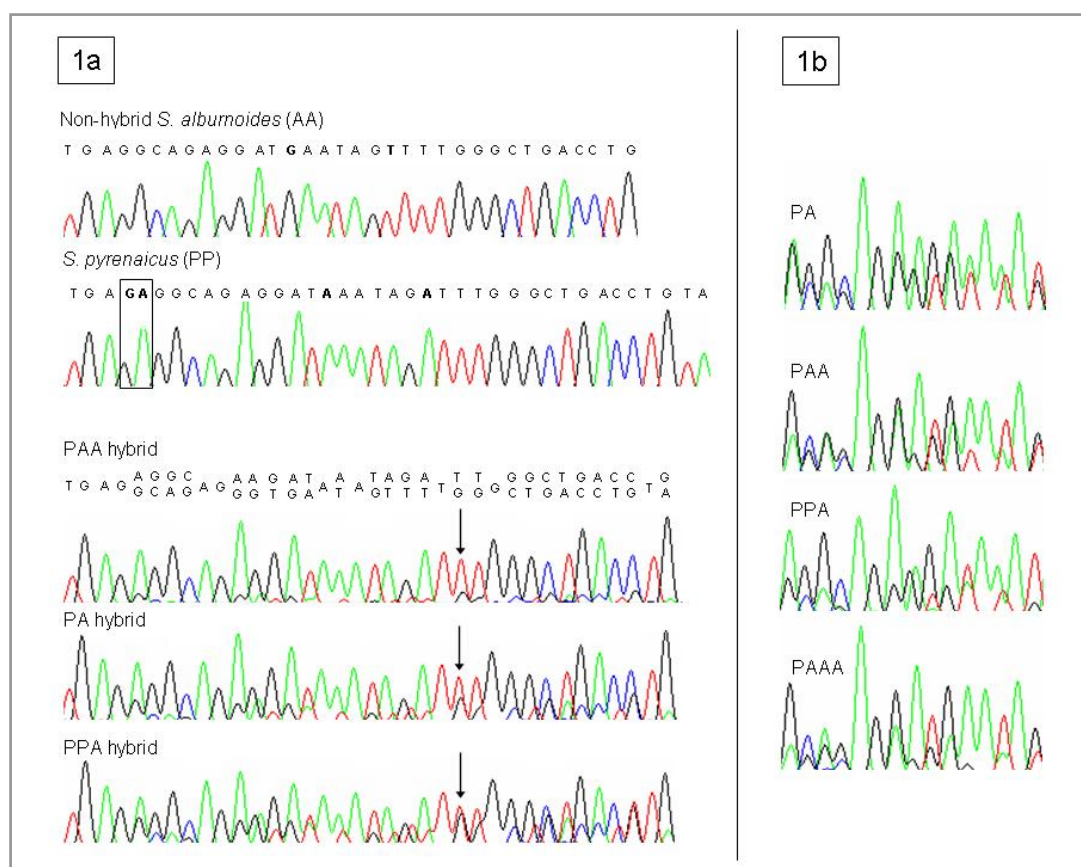


Figure 1 - 1a) Sections of the chromatograms of the parental species and of the three types of artificial hybrids (made as part of the validation of the method based on the analysis of double peaks - **Chapter 2**). The rectangle in the chromatogram of *S. pyrenaicus* shows the insertion of two bases that generate the subsequent peak overlap in the hybrids. The arrow indicates a position where it is notorious the decreasing amount of the base attributed to the nonhybrid *S. alburnoides* progenitor from PAA to PA to PPA, as it was expected. 1b) Sections of the chromatograms of *S. alburnoides* natural hybrids (diploid, triploids and tetraploid), where different dosages of A and P genomes are evident.

In triploids, by comparing the relative heights of the peaks ("measuring method") and by counting the number of times the higher overlapping peak corresponds to each progenitor genome ("count method"), it was possible to evaluate the constitution of the hybrid genomes (whether they were PAA or PPA) – Figure 1. Additionally, as in triploids one of the genomes is in double dose, if those two haplotypes are different, it is possible to identify specific point mutations which allow us to reconstruct the three distinct genomes involved in the formation of the triploid. For example, in the case of a PPA triploid of *S. alburnoides*, if we have two known bases characteristic of the P- and A-genomes, respectively, the third peak would be ascribed to the other P strand. This attribution of the unknown base to a P-genome would fail only in a situation in which both A and P suffered mutations by base substitutions at the same site which, although possible, is unlikely. The same reasoning applies to the interpretation of triple peaks in PAA genomes.

Another example in triploids is the occurrence of a double peak of known bases but in which the base ascribed to the non-duplicated genome appears disproportionably high (when compared with the normal height of that base in that particular position, regardless of the "effect of the adjacent base"). In this situation, one of the duplicated genomes might have mutated to the same base as the one from the non-duplicated genome, resulting in the addition of the peaks in a single higher peak.

If even so these situations could not be resolved, the unknown base (or the segment of unknown bases) should be called point (or zone) of uncertainty and the reconstruction of the genomes should be recovered ahead, if possible, starting at a known characteristic mutation of one of the parental species. Another example of a point of uncertainty occurs when a heterozygote has more than one isolated double peak. In these circumstances, however, methods have been proposed that take advantage of the knowledge of the frequencies of different mutations in the populations to estimate the more likely combinations (Clark 1990; Excoffier & Slatkin 1995; Stephens *et al.* 2001; Niu *et al.* 2002).

Tetraploids in *S. alburnoides* are relatively rare and in theory cannot be significantly distinct from diploid hybrids if they show two copies of each genome. However, one could expect that a tetraploid with a PPPA genome would present higher values of P-count and of P/P+A ratio than the ones for PPA triploids. On the contrary, a tetraploid with a PAAA genome should present lower values of these ratios than a PAA

triploid, an expectation that was corroborated by our data. Additionally, the method presented in this paper may be useful even in the above mentioned case of symmetrical (PPAA) tetraploids: if the genomes involved are different (for instance two A- or two P-genomes distinct by one or more mutations) the method will help to differentiate them from diploids.

Besides the confirmation of the ploidy level of the sequenced individuals by flow cytometry measurements of the erythrocyte DNA content, the methodology used for the determination of the genomic constitutions was also validated by the analysis of artificial hybrids. Indeed, the mixture of known volumes of DNA in order to produce artificial PA, PAA and PPA hybrids that were subsequently sequenced for the beta-actin gene was a crucial test to ensure that the relative size of the peaks attributable to both DNA donors was dosage-dependent.

The comparison of chromatograms of the same fish obtained in different PCR and sequencing runs showed that the absolute sizes of peaks can vary radically from chromatogram to chromatogram and are, thus, sensitive to the PCR and sequencing processes. However, as the method is not dependent on absolute peak sizes, but on the ratio of relative sizes of peaks of different genomes, these variations of amplification and sequencing conditions do not compromise its use. Moreover, to control for these variations: 1) in all PCR and sequencing runs at least one diploid hybrid should be included to serve as a standard control to that run; 2) the “measuring method”, although more time-consuming, should be preferred over the “count method”; and 3) results should be always confirmed with forward and reverse sequencing (only applicable for the reconstitution of the sequences involved in a hybrid).

It should be noted, however, that the pattern of double peaks was not identical in forward and reverse sequences. This had to do with the presence of seven heterospecific indels in *S. alburnoides* hybrids (see Figure 1 in **Chapter 2**). Thus, in forward readings, the series of double peaks started at the first indel (around base pair number 85) whereas reverse sequences showed double peaks from the 5' end until the position of the last indel, around base pair number 650 (which was the first to be read since the sequencer starts reading from the 3' to the 5' end). As a consequence, the reconstitution of the sequences involved in a hybrid has to be done 5' to 3' in forward sequences and 3' to 5' in the reverse ones.

Using this artefact of the sequencing process, caused by the existence of indels, one can access the haplotype diversity of heterozygotes and interspecific hybrids. Therefore, in the case of interspecific studies it is possible to establish phylogeographical inferences related to the origin of the hybridization events, ancient contacts between populations and migratory routes. As flow cytometry methodology can only determine the ploidy level of the samples and not their genome composition (Próspero & Collares-Pereira 2000), this method constitutes a complementary tool for the analysis of hybrids with non balanced parental genomes.

Additionally, when compared with allozyme electrophoresis, this method presents the great advantage of avoiding the problems of regulation of gene expression that make often the dosages of different allozymes impossible to determine (Alves *et al.* 1997a; Letting *et al.* 1999). It is also more advantageous than the use of microsatellites, which is a more expensive and time-consuming method since several loci are needed to obtain sufficient statistical power.

Far from being restricted to the study of interspecific hybrids, this method may find a broad spectrum of applications in population genetics. When several SNP's are found in a given DNA segment it is often difficult to assign each base of each polymorphic site to the strand inherited from one of the parents. The application of the procedure described in this paper makes possible to assign multiple SNP's to each of the two parental strands. This is especially so if 1) these SNP's are located in the vicinity of an indel, 2) some individuals that are heterozygous for the indel were sequenced, and 3) the sequences for individuals that are homozygous for the indel are known.

The more distant a SNP is from the onset of the pattern of double peaks, the less reliable is the attribution of specific bases to the DNA strands because the likelihood of recombination will tend to increase with the length of the segment considered. Thus, DNA segments that harbour several indels that are relatively close to each other provide ideal material for the application of this method, as they provide several reference points that allow recovery of the specific sequences starting from both directions and minimize the risk of error caused by recombination.

When using this method it is essential to avoid the risk of amplifying, in the same PCR, sequences from paralogous copies of the same gene. Thus, care is needed to design primers that are specific to a particular copy of a gene or, preferably, targeting

single copy genes. Introns, as non-codifying regions inside nuclear genes, have high rates of substitution and indel formation (Forsdyke 1981) so one can expect to find many genes with heterozygous indels with a considerable frequency.

10.2 Origin of the *S. alburnoides* complex

10.2.1 The identity of the paternal ancestor

Despite the early clarification of the hybrid origin of the *S. alburnoides* complex and of the involvement of *S. pyrenaicus* as the maternal ancestor (Alves *et al.* 1997a; Carmona *et al.* 1997), the identity of the paternal ancestor of the complex was unknown until very recently. Although some unpublished data (cited by Alves *et al.* 2001) based on sequences of introns from two nuclear genes suggested that the complex was fathered by an *Anaocypris*-like species, the final evidence came from the analysis of the beta-actin nuclear gene (**Chapter 7**) and, subsequently, from microsatellites (Crespo-López *et al.* 2006) and cytogenetic analysis (Gromicho *et al.* 2007). Since some divergence (patristic distance of 1.48%) was detected between the beta-actin sequences of *A. hispanica* and *S. alburnoides* nonhybrids (presenting the nuclear paternal genome in homozygosity - AA), the paternal ancestor was not *A. hispanica* but instead a presumably extinct species belonging to the clade of *A. hispanica*, whose nuclear genome was perpetuated by *S. alburnoides*. To corroborate this hypothesis is the fact that although *S. alburnoides* nonhybrids share some traits with *A. hispanica* (small size, high number of gillrakers and uniseriated pharyngeal teeth – Collares-Pereira 1983), the divergence between them at the beta-actin gene level is considerable and identical to the patristic distance that separates distinct species like *S. pyrenaicus* and *S. valentinus* (2.15% - **Chapter 8**).

This *Anaocypris*-like ancestor must have radiated from a proto-*Anaocypris*, belonging to the European *Alburnus* lineage (**Chapter 7**), that entered the Iberian Peninsula and dispersed towards west, at least to the area of what is now the River Guadiana basin.

10.2.2 Paleobiogeography of the ancestors

Cyprinids reached Europe in the Oligocene, from Asia, when the Turgai Sea that separated the two continents dried out (Briggs 1987). The migrations of European cyprinid lineages towards the Iberian Peninsula must have taken place, in most cases, in the Oligo-Miocene (Almaça 1976). After that, the uplift of the Pyrenean Mountains irreversibly prevented further colonisations of Iberia by freshwater fish, possibly except through temporary connections with Africa in the Late Miocene (Andeweg 2002) and through dispersals along adjacent rivers during marine regressions. Concerning the connection with Africa, apparently it was only relevant for some barbels (Machordom & Doadrio 2001; Doadrio *et al.* 2002) and cobitids (Doadrio & Perdices 2005) but not for Leuciscinae (Zardoya & Doadrio 1998).

An alternative colonisation route also of Miocene age, involving wide migrations across the Mediterranean during the Messinian Salinity Crisis was proposed by Bianco (1990). However, this Mediterranean route, often called “Lago Mare”, was recently discarded based on *Chondrostoma* and *Squalius* phylogenies (Doadrio & Carmona 2003). Thus, the colonisation of the Iberian Peninsula by Leuciscinae, including both the ancestors of the *S. alburnoides* complex, must have predated the complete isolation by the Pyrenees, likely through freshwater passages from southern France to the Ebro river basin – Figure 2.

Indeed, cyprinid fossils of Miocene age were detected throughout the Iberian Peninsula (De la Peña 1995), indicating that, since its colonisation in the Oligocene, the dispersal of freshwater fish from the eastern to the western margin occurred in only 10-15 million years. This dispersal process was likely favoured by the fact that during the Miocene most Iberian river systems, instead of flowing to the sea, drained to a large number of inland lakes some of which persisted well into the Pliocene (Friend & Dabrio 1996; Andeweg 2002).

As shown in **Chapter 8**, the paleohydrographical history of Iberia seems to have been determinant to the dispersal and radiation of *Squalius* and *Alburnus* lineages. Regarding the *Squalius* lineage, a proto-*S. pyrenaicus* clade originated at least five *Squalius* species as is dispersed from the centre towards the peripheral areas of the Peninsula, since the Miocene to the Pliocene: *S. torgalensis* and *S. aradensis* in the

southwest, *S. carolitertii* in the North, *S. valentinus* in the Southeast and *S. malacitanus* in the South – Figure 3.



Figure 2 — Representation of the arrival of the *Squalius* and *Alburnus* European lineages to the Iberian Peninsula. The localization of the Quaternary endorheic lakes was adapted from Andeweg (2002). Image credits: Jacques Descloitres, MODIS Rapid Response Team, NASA/GSFC.

The *Alburnus* lineage must have followed at least some of the freshwater connections used by *Squalius* but, as evidenced by its present day restricted distribution (limited to the southern part of Iberia – Figure 3), must have experienced major depletions and extinctions. Indeed, the last three representatives of an *Alburnus* lineage in the Iberian Peninsula seem to be *A. hispanica*; the paternal ancestor of the *S. alburnoides* complex (presumably extinct but whose nuclear genome was preserved in *S. alburnoides*); and the paternal ancestor of *S. palaciosi* (likely another case of a hybridogenetic complex between *S. pyrenaicus* and a presumably extinct *Anaocypris*-like ancestor, in the Guadalquivir River). The hybrid origin of *S. palaciosi*, although not demonstrated genetically, is strongly

suggested by the morphological resemblance of this fish with *S. alburnoides* (Doadrio 2001) and by a *S. pyrenaicus*-like mtDNA (Zardoya and Doadrio 1998).



Figure 3 — Radiation of the proto-*S. pyrenaicus* clade, generating: 1 — *S. aradensis* (in pink), 2 — *S. torgalensis* (in green), 3 — *S. carolitertii* (in yellow), 4 — *S. malacitanus* (in purple) and 5 — *S. valentinus* (in blue). The present day distribution area of *S. pyrenaicus* is depicted in white. Red circles represent the geographical areas to where the last three representatives of the *Alburnus* lineage are restricted: 6 — paternal ancestor of the *S. alburnoides* complex (in the Guadiana basin - virtually extinct?); 7 — *Anaecypris hispanica* (in the Guadiana basin); 8 — paternal ancestor of the *S. palaciosi* complex (in the Guadalquivir basin - virtually extinct?). Image credits: Jacques Descloitres, MODIS Rapid Response Team, NASA/GSFC.

10.2.3 Single origin – where and why?

Results now obtained (**Chapter 8**) showed that the origin of the complex is of Middle-Upper Pleistocene age (around 700.000 years ago). Previously, Cunha *et al.* (2004) had postulated a more ancient origin (in the Upper Pliocene). These discrepancies may

eventually result from a less broad sampling effort and/or from the apparent use of uncorrected patristic distances by those authors.

When compared with the estimated age of less than 500.000 years of several asexual vertebrates (Avice *et al.* 1992; Scharl *et al.* 1995; Mateos & Vrijenhoek 2005; Robertson *et al.* 2006), *S. alburnoides* seems to have achieved a higher evolutionary longevity.

The hybridization between two species, conducive to the formation of a new species, requires the sympatry between the parental species and the absence of reproductive isolation between them. Thus, the parental species had to be isolated until the formation of the *S. alburnoides* complex – Figure 4a. One possible explanation relies on the recent establishment of the current exorheic Iberian river network, in the Plio-Pleistocene (Andeweg 2002). As mentioned above, at this time the *Squalius* lineage had already dispersed throughout the Peninsula thus, the isolation of the *Anaecypris*-like ancestor in a southern endorheic refuge seems to be a plausible hypothesis to explain the absence of sympatry of the ancestors for such a long time. Under this scenario, both ancestors first contacted when, as a result of the hydrographical rearrangements, the Tagus-Guadiana drainage carrying *S. pyrenaicus* captured the endorheic refuge where the *Anaecypris*-like ancestor remained isolated – Figure 4b.

This hypothesis is corroborated by geological data indicating that the medium section of the Guadiana was an endorheic lake until it was captured, in the Pleistocene, by the Tagus and Guadiana upper basins that were still connected at the time (Moya-Palomares 2002). Indeed, according to Dowling & Secor (1997), the onset of hybridization is often associated with habitat disturbances, either by anthropogenic activities or as a result of dramatic environmental changes. Thus, contrary to other authors that postulated multiple origins for the complex (Alves *et al.* 1997b; Cunha *et al.* 2004), it is proposed that *S. alburnoides* had a single origin in the Pleistocene, when the upper Tagus and Guadiana river basins were still connected (Moya-Palomares 2002). Moreover, this central location in the Peninsula of the endorheic refuge of the paternal ancestor would explain the wide distribution of the complex that does not include the smaller and peripheral river basins of the Peninsula: the Tagus and Guadiana Rivers, with the basculation of the Peninsula towards the Atlantic, became separated and progressively acquire their current exorheic longitudinal profiles, thus, acting as dispersal corridors

towards north, south and west, through the still available fluvial connections between the main drainages, during the formation of the modern Iberian hydrographical network.

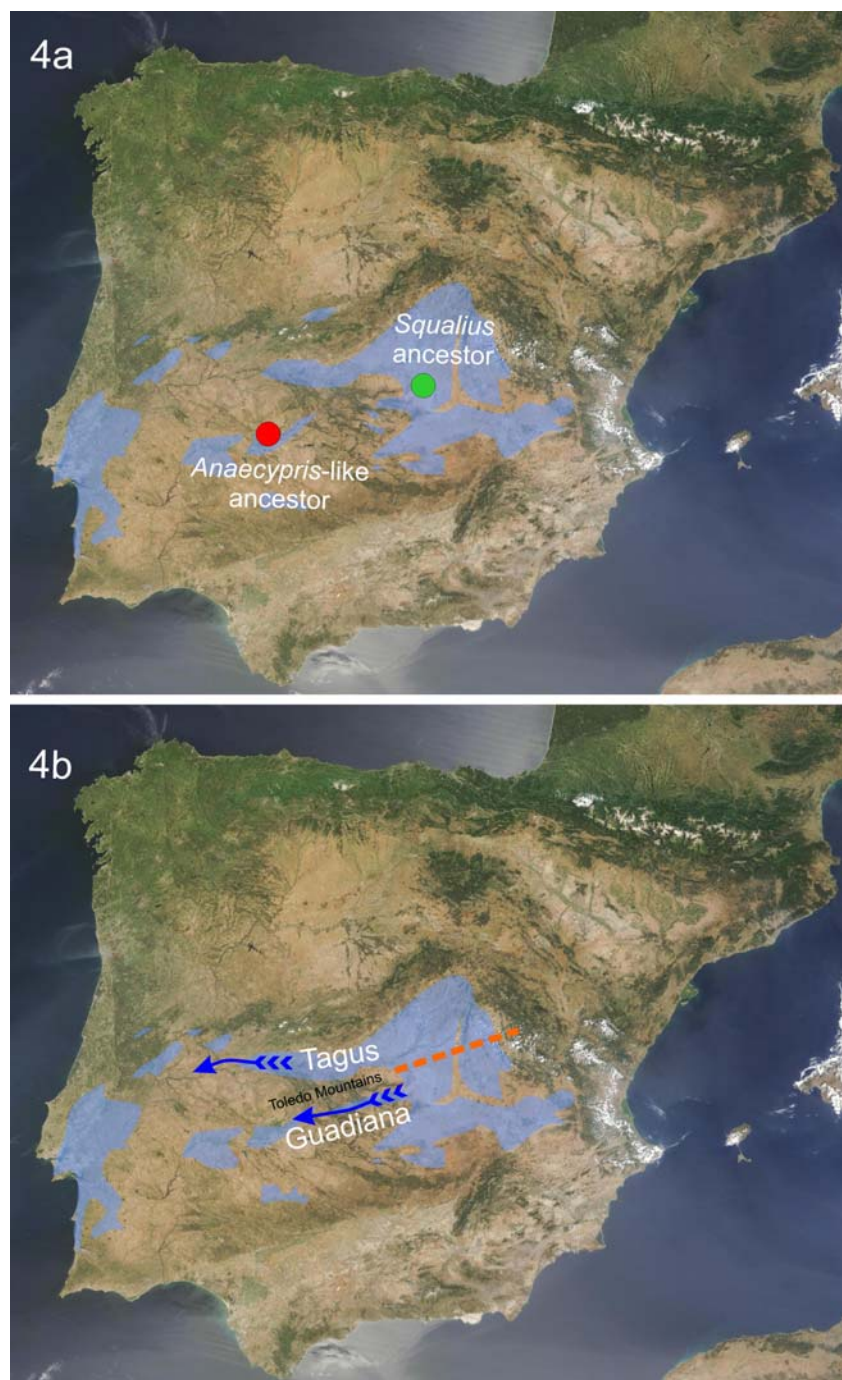


Figure 4 — Diagram representing the hypothesized origin of the *S. alburnoides* complex in the Tagus-Guadiana area. 4a — Both ancestors were isolated in independent endorheic lakes: the *Squalius* ancestor in the Tagus-Guadiana basin (green dot) and the *Anaecypris*-like ancestor in a more southern refuge (red dot); 4b — With the basculation of the Peninsula, the endorheic basins started to drained towards west (represented by the blue arrows) and the rivers progressively acquire their longitudinal profiles. The rivers Tagus and Guadiana became independent from each other (red line), flowing north and south of the Toledo Mountains, respectively (Moya-Palomares 2002). The localization of the Quaternary endorheic lakes was adapted from Andeweg (2002). Image credits: Jacques Descloitres, MODIS Rapid Response Team, NASA/GSFC.

If instead of a single origin one assumes multiple independent ones, several unparsimonious assumptions have to be considered, namely that the *S. pyrenaicus* and the *Anaecypris*-like ancestors had dispersed to different basins (preserving their identity during the process) and, subsequently, for unknown reasons, began to hybridize independently in those basins after which the paternal ancestor became extinct independently in each of those basins (see section 10.2.6).

In contrast to the wide distribution of *S. alburnoides*, the extant distribution areas of *A. hispanica* and *S. palaciosi* are very restricted (Figure 3), suggesting that their endorheic refuges became isolated for a longer time.

Concerning *S. palaciosi*, the exorheism of the Guadalquivir River might have allowed the contact between its paternal ancestor and *S. pyrenaicus* (that most likely reached the Guadalquivir through connections with tributaries of the Upper Guadiana).

Finally, the third representative of the *Alburnus* lineage in Iberia, *A. hispanica*, must have been isolated in a more southern endorheic lake, in the region of the lower Guadiana – the Moura-Marmelar endorheic basin is one of the hypotheses (Moya-Palomares 2002). Since Guadiana River acquired its present configuration very recently, in the Upper Pleistocene (Rodríguez-Vidal *et al.* 1993), at a time when the connections with other river basins were no longer possible, the capture of a endorheic basin near the terminal part of the present course of the river may explain why the distribution of *A. hispanica* is now restricted only to the median and lower sections of the Guadiana drainage (Collares-Pereira *et al.* 2002). If, alternatively, the refuge of *A. hispanica*, was located in the bulk of Iberia, one should expect that the distribution range of this species would be similar to that of *S. alburnoides* and that the already reported hybrids between *A. hispanica* and *S. alburnoides* (Collares-Pereira 1983) were disseminated instead of being restricted to the Guadiana basin.

Besides sympatry, another premise for hybridization is the absence of reproductive isolation between the hybridizing species, which seems to have been the case in the formation of the *S. alburnoides* complex, even though the ancestors belonged to distinct lineages and were relatively distant species: mean patristic distances of 1.80% between P and A beta-actin sequences (**Chapter 7**) and of about 14% between *cytb* sequences from *S. pyrenaicus* and *A. hispanica* (C. Sousa-Santos unpublished data). Some intrageneric hybridizations have been described for European cyprinids (*e.g.* Crespin &

Berrebi 1994; Durand *et al.* 2000; Tsigenopoulos *et al.* 2002; Salzburger *et al.* 2003), but reported hybridizations between distantly related species are much more uncommon: *Rutilus rutilus* x *Abramis brama* (Pitts *et al.* 1997); *Leuciscus cephalus* x *Chalcalburnus chalcoides* (Ünver & Erk'Äkan 2005), *Abramis brama* x *Rutilus rutilus* (Yakovlev *et al.* 2000), and *Scardinius dergle* x *Squalius tenellus* (Freyhof *et al.* 2005).

10.2.4 Unidirectionality of interspecific crosses – possible causes

The mitochondrial DNA of the paternal ancestor of the complex was never detected, after numerous studies involving *S. alburnoides*. Assuming that it must have been lost in evolution, one possible explanation for this loss may be the occurrence of almost unidirectional hybridizing crosses, as happens in most of the extant hybrid complexes (Avice *et al.* 1992) – the Central European *Cobitis elongatoides-taenia* complex configures one of the exceptions as apparently reciprocal hybridizations occur (Janko *et al.* 2003).

In the particular case of fishes, several pre- and postzygotic causes have been postulated to explain cases of unidirectional hybridizations (reviewed by Wirtz 1999). Prezygotic causes include size differences between species, sneaking behaviour (furtive fertilizations of the eggs of females that are being courted by other males), unidirectional forced copulations or differences in the discrimination intensity against allospecific males by the females of the two species, among others (Wirtz 1999). On the other side, postzygotic causes may include the production of unviable offspring in the reciprocal crosses, random extinction of one of the two types of hybrids, and extinction of one type of hybrids due to comparatively lower fitness (Wirtz 1999).

Concerning *S. alburnoides*, if interspecific crosses between *S. pyrenaicus* females and *Anaecypris*-like males were systematically much more frequent than the reverse ones, this could explain the loss of the paternal species mtDNA after many generations.

From an ethological point of view this hypothesis seems to be plausible. In non-territorial fish like many cyprinids (Johnston 1999), males frequently have to chase and follow females to achieve fertilizations. The larger a female, the more eggs it can carry (confirmed by Ribeiro *et al.* 2003 for *S. alburnoides* females) and probably the stronger it is as a stimulus for the males. Thus, it may be a very simple process that leads males to

chase and follow the larger females it can find, while the reverse is probably not so common (*i.e.* small females being chased by larger males). Specimens of *S. pyrenaicus* may reach a maximum body length of 18.1cm (Ribeiro *et al.* 2007), whilst *A. hispanica* are typically very small (maximum body length of 7.5cm - Collares-Pereira 1980), as probably was the case of the *Anaecypris*-like ancestor. Thus, it is plausible to hypothesize that the large *S. pyrenaicus* males would likely prefer to mate with the also large *S. pyrenaicus* females to the detriment of the smaller *Anaecypris*-like ones. The remaining males, belonging to the *Anaecypris*-like species, would either mate with conspecific females of similar size or attempt to mate with *S. pyrenaicus* females, following the usual criterion of male mate choice by size. If this hypothesis proves to be correct, the reproductive behaviour of both hybridizing species would have been determinant for the evolutionary history of *S. alburnoides* since an early stage.

In the absence of the paternal ancestor of the complex, a possible way to test this hypothesis would be to analyze the reproductive behaviour of *S. pyrenaicus* and *A. hispanica* of both sexes maintained in the same experimental setting or, alternatively, *S. pyrenaicus* of both sexes and nonhybrid *S. alburnoides* males, the living representatives of the paternal ancestor of the complex. As shown in this thesis (**Chapter 5**), these nonhybrid males, despite their much smaller size, were successful in stimulating the spawning and fertilizing the eggs of *S. torgalensis* females. Moreover, when a nonhybrid male was placed in a tank during active courtship episodes of *S. pyrenaicus* spawning pairs, the nonhybrid male tried to place himself between the spawners (C. Sousa-Santos unpublished data). Thus, although more observations are needed, the occurrence of the original interspecific crosses that gave rise to the *S. alburnoides* complex, rather than being stochastic events consequent of a common usage of the same spawning grounds by distinct species, may be explained by ethological reasons: eventual adoption of an interspecific sneaking behaviour by the *Anaecypris*-like paternal ancestor (see section 10.4.2), male mate choice based on size and/or presence of female choice based on more vigorous displays performed by the smaller nonhybrid males. Behavioural causes seem also to explain the directionality of hybridization between *Fundulus heteroclitus* and *F. diaphanous* (Chávez & Turgeon 2007).

10.2.5 Origin of the various forms of the complex

As already described in section 1.5.3, various reproductive modes have been reported for *S. alburnoides* males and females (Alves *et al.* 1999, 2001, 2004; Pala & Coelho 2005; Crespo-López *et al.* 2006; **Chapter 5**). The confirmation of the production of clonal sperm by PAA males (as proposed by Alves *et al.* 1999) was made by the analysis of the offspring of only one specimen (**Chapter 5**), which although probably typifying the most usual reproductive mode, does not allow the exclusion of the occurrence of other spermatogenic processes.

From all the known *S. alburnoides* forms, only the fertility and gametogenic processes of PPA males and unbalanced tetraploids remain unknown. Concerning the remaining forms, and in contrast to what happens in many hybrid complexes (Dawley 1989), means to escape infertility have evolved not only in diploids, but also in triploids and tetraploids. However, the reproductive modes exhibited by females are more diverse and the hypothesis of recombination is much more frequent. Indeed, with the exception of nonhybrids and balanced tetraploids, the other forms of *S. alburnoides* males (diploid and probably both types of triploid hybrids) produce only clonal sperm, contributing much less to the genetic diversity of the complex than females. Moreover, as apparently spermatogenesis of hybrid males involves no reduction in ploidy, the offspring sired by these males will possibly be, in many cases, unviable (ploidy levels higher than $4n$).

Knowing the types of gametes produced by almost all the *S. alburnoides* forms, it is possible to predict the genome constitutions of the offspring that resulted from the first generations after the original crosses that gave rise to the complex. These simulations are presented in Table 1 and their results, namely the proportion of each genome constitution expected to have resulted from all possible crosses in F_0 , F_1 and F_2 generations, are summarized in Table 2.

The next generation (F_3) was not considered since the only difference from the F_2 generation was the inclusion of two additional progenitors, PPPA and PAAA tetraploids (the later already reported, in River Tagus - **Chapter 8**), for which the fertility and gametogenesis are unknown.

Table I — Simulation of the genome constitutions produced as a result of all the possible crosses at the formation of the *S. alburnoides* complex (F_0) and at the two subsequent generations (F_1 and F_2). For simplicity reasons, the following assumptions were considered: PPA males produce clonal sperm as do PAA males; only the most common type of eggs produced by PAA females (according to Alves *et al.* 2001) was considered in the simulations; and in each cross females and males were expected to have been originated. *virtually unviable progeny (the maximum number of chromosome sets known for the complex is exceeded); **form never observed in nature but obtained in experimental crosses (Crespo-López *et al.* 2006).

F ₀	MALES	AA	PP
FEMALES	gametes	a	P
AA	a	AA	PA
PP	p	PA	PP

F _I	MALES	AA	PP	PA
FEMALES	gametes	a	p	pa
AA	a	AA	PA	PAA
PP	p	PA	PP	PPA
PA	pa	PAA	PPA	PPAA

F ₂	MALES	AA	PP	PA	PAA	PPA	PPAA
FEMALES	gametes	a	p	pa	paa	ppa	pa
AA	a	AA	PA	PAA	PAAA	PPAA	PAA
PP	p	PA	PP	PPA	PPAA	PPPA	PPA
PA	pa	PAA	PPA	PPAA	*	*	PPAA
PAA	a	AA	PA	PAA	PAAA	PPAA	PAA
	aa	AAA*	PAA	PAAA	*	*	PAAA
	paa	PAAA	PPAA	*	*	*	*
PPA	p	PA	PP	PPA	PPAA	PPPA	PPA
PPAA	pa	PAA	PPA	PPAA	*	*	PPAA

Obviously, *S. pyrenaicus* males and females were also included in these simulations since they occurred in sympatry with the paternal ancestor of the complex and there are clear evidence that they cross with *S. alburnoides* (**Chapters 5, 8 and 9**). A similar reasoning was also applied to the inclusion of the females of the paternal species of the complex. Indeed, although the reconstituted nonhybrid form is almost exclusively constituted by males (Alves *et al.* 2001; **Chapter 4**), there is no reason to suspect that the virtually extinct paternal species was not bisexual.

Table 2 — Summarized results from the simulations presented in Table 1 regarding the genome constitutions of the offspring produced at the origin of the complex (F_0) and at the two subsequent generations (F_1 and F_2). For each generation, the number of possible crosses that result in a giving genome constitution is depicted as a proportion of the total number of possible crosses. * form never observed in nature but obtained in experimental crosses (Crespo-López *et al.* 2006).

Offspring	Generations		
	F_0	F_1	F_2
AA	1/4	1/9	2/48
PP	1/4	1/9	2/48
PA	2/4	2/9	4/48
PAA	-	2/9	7/48
PPA	-	2/9	6/48
PPAA	-	1/9	9/48
PAAA	-	-	5/48
PPPA	-	-	2/48
AAA*			1/48
Unviable (>4n)	-	-	10/48

As shown in Tables 1 and 2, if all crosses were equally probable, all the known forms of the complex could have been produced in only two generations after the original hybridizations between the ancestors of *S. alburnoides*. Also theoretically, the parental genomes would become rarer with the increasing number of crosses resulting in the formation of *S. alburnoides* hybrids. However, as will be detailed in section 10.4.3, the composition of the complex in natural populations is dramatically distinct from what theory predicts, not only in terms of the relative frequency with which the different forms occur but also regarding the presence/absence of some forms. In addition, in all the populations the sex is extremely biased towards females (Alves *et al.* 2001) but the sex determination mechanism remains unclear.

Phenotypically, it was demonstrated that both P and A genomes were expressed in the phenotype of the *S. alburnoides* hybrids and that the most common forms (AA, PAA,

PA and PPA) could be clearly differentiated (**Chapter 3**), which, according to Kearney & Shine (2004), may cause a reduction in niche overlap, allowing varying degrees of coexistence between them.

Moreover, a morphological continuum was detected for several traits, having AA and PP as extreme forms and the hybrids in an intermediate position: triploids with double A-genomes were more closely related to nonhybrids (AA), followed by diploid hybrids and finally by triploid hybrids with double P-genomes (the hybrid form more closely related to PP). This variation suggested the existence of a gene dosage effect on the phenotype exhibited by *S. alburnoides* hybrids, extremely well marked in the almost disappearance of the longitudinal dark band or in the increasing number of biseriated pharyngeal teeth as hybrids with increasing dosages of the P-genome were considered (**Chapter 3**). Amongst hybrid lineages with distinct ploidies, several cases of dosage-related expression of morphological (Albertson & Kocher 2001; Placek 2002; Kittell *et al.* 2005; Mulcahy *et al.* 2006) and also of behavioural traits (Lima & Bizerril 2002) were already reported.

10.2.6 The extinction of the paternal ancestor

As mentioned in section 10.2.4, the occurrence of almost unidirectional crosses between the ancestors of *S. alburnoides* may be explained by similar mate preferences exhibited by the males of both hybridizing species: *S. pyrenaicus* and *Anaocypris*-like males would likely prefer to mate with the larger *S. pyrenaicus* females.

If this view is confirmed it leads to the surprising conclusion that, as hybridizing events progressed, the reproductive success of the *Anaocypris*-like females must have decreased to a point where they became extinct, a situation that raises a question of who was the winner and who was the loser in this game. In a sense, the presence of *S. pyrenaicus* promoted the formation of a hybridogenetic fish but at the cost of the extinction of the *Anaocypris*-like species.

If the hybrid *S. alburnoides* males retained the same preference to mate with larger females as their ancestors, they could also have contributed to the complete elimination of the *Anaocypris*-like ancestral species.

An effect that apparently has been overlooked in previous discussions is the fact that when a small male spawns with a much larger female it will be able to fertilize more eggs than if it spawns with a small size female, because of the typical increase in fecundity with female size (Bagenal 1978), also observed in this complex (Ribeiro *et al.* 2003). Thus, the genes of the ancestral species would tend to be propagated much more efficiently by crosses with *S. pyrenaicus* females or with the already formed *S. alburnoides* hybrid females, especially the triploids, than through crosses with the females of the ancestral species. If these considerations proved to be correct even when ecological conditions tend to be less favourable, the fitness of the small AA males will tend to be higher than that of their female counterparts.

Apart from the ethologically-driven disadvantage of the small AA females, at least two ecological conditions may have also favoured the extinction of the ancestral AA females.

First, the *Anaecypris*-like paternal species, perhaps due to its small dimensions (inferred from the morphology of the extant *A. hispanica* and *S. alburnoides* nuclear nonhybrids - Collares-Pereira 1983) could have been favoured by the ecological conditions of the endorheic southern basins, namely higher water temperatures, intermittent conditions and waters with reduced flow. However, with the exorheism events and the consequent change to less favourable ecological conditions associated with the cooling of the climate in the Quaternary and more torrential regimes (Ribeiro *et al.* 1987; Hewitt 1999, 2000), the paternal ancestor of *S. alburnoides*, as well as the other two representatives of the *Alburnus* lineage in Iberia (see section 10.2.3), could have suffered considerable declines and local extinctions. Indeed, with the Pleistocene glaciations the freshwater habitats were altered on a dramatic scale and aquatic species were particularly affected since the opportunities for dispersal were very limited due to the scarce water connections still available despite the advancing glacial fronts (Bernatchez & Wilson 1998). Thus, the successive cycles of expansion-contraction of glacial ice sheets must have played an extremely important role in shaping the distribution of the taxa (Rowe *et al.* 2004). After the last glaciation, the previously affected rivers returned to typical Mediterranean conditions, which may explain why the survival of the also small *S. alburnoides* nonhybrids (AA), contrary to that of their paternal ancestors, was not threatened by climatic conditions, at least in southern populations (see section 10.4.3).

Secondly, the extinction of the paternal species might have been a consequence of direct competition with the successful *S. alburnoides* complex. Indeed, several cases of displacement or elimination of one of the hybridizing species by the resultant hybrids are known (*e.g.* Levin *et al.* 1996; Ayres *et al.* 2004; Konishi & Takata 2004; Rosenfield *et al.* 2004; Rhymer 2006). Also, although not sufficiently studied yet, in river basins where *S. alburnoides* occurs, the populations of *S. pyrenaicus* or *S. carolitertii* seem to be experiencing depletions. This apparent tendency was detected during the collection of samples for the present research (Table 3) and may be simply related to differential habitat occupation (fish sampling was mainly performed in low velocity and shallow stretches of rivers, which may be a factor biasing the results) or may, alternatively, be a result of the high ecological success of *S. alburnoides* (likely constituting an example of a hybrid species that has outnumbered its progenitors).

Table 3 — Number of captured specimens belonging to *S. alburnoides* and other *Squalius* species. For each sampling location the percentage of *S. alburnoides* from the total of specimens captured is depicted.

Sampling location	Captured specimens					% of <i>S. alburnoides</i>
	Drainage	River	<i>S. alburnoides</i>	Other <i>Squalius</i>	Total	
1	Douro	Sabor	30	0	30	100,00%
2	Douro	Tâmega	23	21	44	52,27%
3	Douro	Paiva	0	7	7	0,00%
4	Guadiana	Caia	60	0	60	100,00%
5	Mondego	Dão	0	9	9	0,00%
6	Mondego	Ceira	2	3	5	40,00%
7	Mondego	Ceira	99	26	125	79,20%
8	Sado	Sado	67	0	67	100,00%
9	Sado	Sado	26	3	29	89,66%
10	Tagus	Sorraia	15	0	15	100,00%
11	Tagus	Sorraia	54	2	56	96,43%
12	Tagus	Sorraia	4	0	4	100,00%
13	Tagus	Ocreza	10	0	10	100,00%
14	Tagus	Ocreza	5	0	5	100,00%
15	Tagus	Ocreza	79	56	135	58,52%
16	Tagus	Zêzere	0	38	38	0,00%
17	Tagus	Pônsul	11	6	17	64,71%
18	Tagus	Erges	37	2	39	94,87%
19	Tagus	Seda	53	0	53	100,00%

Regardless of the hypothesized cause for the extinction of the *Anaocypris*-like ancestor - by ethological causes, unfavourable climatic conditions, drastic habitat changes and/or competition with *S. alburnoides* - it seems reasonable to admit that it must have preceded the dispersal of the *S. alburnoides* complex, a scenario that is corroborated by the fact that the mtDNA of the paternal species was never detected in the analyzed specimens from all of the drainages where the complex occurs. The alternative hypothesis of extinction posterior to the dispersal of *S. alburnoides* would require admitting the concomitant dispersal of the complex and of both the parental species (as mentioned in section 10.2.4) and the independent extinction of the *Anaocypris*-like ancestor in each of the basins, which seems to be a much less plausible scenario.

The nonhybrid female described (**Chapter 4**) presented a *S. pyrenaicus*-like mtDNA, as all the nonhybrid males so far analyzed using genetic and cytogenetic markers (Alves *et al.* 1998, 2001, 2002; Gromicho & Collares-Pereira 2004; **Chapters 3, 8 and 9**). As postulated for males, this nonhybrid female was certainly reconstituted from the hybrids - the alternative hypothesis of being a survivor of the paternal ancestor is, thus, discarded.

Inspection of Table 1 shows that offspring with an AA genome constitution can only be originated from crosses between progenitors with that genome constitution, or between an AA male and a PAA female producing haploid eggs bearing one A-genome. Thus, since nonhybrid females are extremely rare (with only two cases reported - Carmona 1997; **Chapter 4**), it is unlikely that AA individuals have been originated by the first type of crosses (AAxAA) but rather from PAAxAA crosses. As a consequence, and as the sex determination mechanism seems to dictate that the offspring of PAAxAA crosses would be exclusively constituted by males (Alves *et al.* 1998; Gromicho & Collares-Pereira 2004; Crespo-López *et al.* 2006), the reported cases of nonhybrid females should be seen as a result of a hypothetical fault in the mechanism. Hypothetically, crosses between AA males and females could yield offspring of both sexes but, as the AA females are extremely rare, AA males will mate almost always, if not exclusively, with hybrid females or females of other *Squalius* species.

Both the above mentioned pathways to the origin of nonhybrids require the existence of one or both progenitors of their own type. Thus, the survival of the AA form is self-dependent and, once lost, cannot be originated *de novo*.

The preceding arguments were all based in the assumption that the paternal ancestor of *S. alburnoides* is completely extinct, which seems to be true according to the available data. However, this assumption should not be taken for granted since neither all of the captured nonhybrids were sequenced for the *cyt b* gene, nor all of the rivers along the distribution range of the complex were subjected to exhaustive collections of samples, despite of the many years of research on this complex.

10.3 Foundation of the *S. alburnoides* Portuguese populations

10.3.1 Patterns of dispersal of the complex

As detailed in section 10.2.2, the origin of the *S. alburnoides* complex seems to have been related to the Pleistocenic capture of an endorheic lake (where the paternal ancestor was likely isolated) in the present day middle section of the Guadiana drainage by the, at the time interconnected, Tagus and Guadiana upper basins (which likely carried the maternal ancestor of the complex). Thus, under this scenario, the Tagus-Guadiana drainage was the first dispersal corridor for the newly formed hybrid complex, as it drained towards west due to the tilting of the Iberian Peninsula (Moya-Palomares 2002). With the ongoing of the basculation process, the Guadiana and Tagus became two independent river basins with their presently known longitudinal profiles (Moya-Palomares 2002) and this event likely configured the first split of the *S. alburnoides* complex in two distinct populations. After that, three dispersal routes were apparently followed by the complex, using the fluvial connections still available in the Upper Pleistocene-Holocene: 1) a northward radiation proceeding from the Tagus; 2) a southward radiation via Guadiana; and 3) a south-westward radiation from Tagus – Figure 5.

According to the now estimated mtDNA divergence times (**Chapter 8**), the northern rivers of Mondego and Douro were colonised, respectively, around 0.05 and 0.01 MY ago, through two independent routes of migration, both proceeding from the Tagus – Figure 5. As corroborated by the geographical proximity and by the lowest of all the divergence values between all possible groups of specimens belonging to the sampled

tributaries of Mondego and Tagus, it is likely that *S. alburnoides* reached Mondego through fluvial captures involving adjacent tributaries of the right bank of the Tagus, in the Zêzere-Erges area (**Chapter 8**).



Figure 5 — Colonisation routes followed by *S. alburnoides* proceeding from the Tagus (1 — Tagus to Mondego, 2 — Tagus to Douro via the Alagon and Águeda tributaries, 3 — Tagus to Sado) and Guadiana (4 — Guadiana to Guadalquivir, and from this basin to Odiel, 5 — Guadiana to Quarteira) drainages. The geological barrier of 400m that likely impeded the radiation of *S. alburnoides* to the Spanish section of the Douro is represented by a red rectangle. Image credits: Jacques Descloitres, MODIS Rapid Response Team, NASA/GSFC.

On the other hand, the colonisation of the Douro by *S. alburnoides* from Tagus seems to have been possible by fluvial captures between tributaries of both rivers that are in close proximity in the Alagon area, which is corroborated by the lowest divergence values between the specimens belonging to these tributaries, when compared with that between specimens from other tributaries of both rivers (**Chapter 8**). This hypothesis is also corroborated by the restricted dispersal of the complex in the Spanish section of the Douro: *S. alburnoides* is found in several localities along the Águeda River (a natural

border separating Spain and Portugal) (Doadrio 2001) and in scattered sampling points in the upper section of the Douro. Indeed, considering the hypothesized corridor between Tagus and Douro through the Alagon and Águeda tributaries, respectively, the entrance of the complex in the Douro occurred downstream of a geological barrier of 400m located at the Portuguese border – Figure 5. Thus, the upstream migration of *S. alburnoides* was likely impeded by this insurmountable barrier, which explains the widespread distribution of the complex in the Portuguese tributaries and its virtual absence in the Spanish ones with the exception of the Águeda River. The above mentioned scattered occurrences in the upper section of the Douro (Doadrio 2002) might be a result of contacts with Tagus tributaries, allowing localized migrations of the complex.

Concerning the dispersal route that allowed the colonisation of the southern drainage of Guadalquivir, it seems to have proceeded from the left bank tributaries of the Guadiana about 0.05 MY ago (**Chapter 8**) and, once there, the complex may have reached the adjacent Odiel basin also by stream captures – Figure 5.

More recently, in the Holocene (0 MY ago), the southern Quarteira drainage was also colonised by *S. alburnoides* from Guadiana (Figure 5), likely through fluvial captures between tributaries of both basins that became possible as a result of the acquisition of the southward draining pattern of the Guadiana basin, which occurred only in the Upper-Pleistocene (Rodríguez-Vidal *et al.* 1993).

Finally, the colonisation of Sado also occurred very recently, in the Holocene (0 MY ago), probably with the final acquisition of the present day longitudinal profile of the Tagus basin, that must have implied the capture of the previously isolated Lower Tagus basin, with which the Sado was connected (**Chapter 8**).

10.3.2 Introgression of genes from other *Squalius* species

The results based of mtDNA analysis (**Chapters 8** and **9**) proved that *S. alburnoides* may interbreed with all the sympatric *Squalius* species (*S. pyrenaicus*, *S. carolitertii* and *S. aradensis*). This potential of *S. alburnoides* to interbreed with other *Squalius* species was also verified in crosses involving a *Squalius* species that has no contact with *S. alburnoides* in the wild (**Chapter 5**): in captivity, *S. alburnoides* males spawned and produced viable

hybrids with the allopatric *S. torgalensis* and the spontaneous reverse cross between *S. alburnoides* females and *S. torgalensis* males was also demonstrated.

This means that the absence of reproductive isolation between the ancestors of the complex (which allowed interspecific hybridization), was maintained between the hybrids and at least four other *Squalius* species in the course of evolution.

The present day introgression of *S. pyrenaicus* mitochondrial genes in the complex has small effects on the gene pool of the complex since they are similar to the ones belonging to the maternal lineage of the complex. A visible effect of these introgressions is the detection, in the same population, of a mixture of ancient and more recent nuclear *S. pyrenaicus* haplotypes that were introgressed as the complex crossed with distinct *S. pyrenaicus* populations during its dispersal pathways (**Chapter 8**).

On the contrary, considerable changes in the typical gene pool of the complex (**Chapters 8** and **9**) resulted from the colonisation of river basins outside the distribution range of *S. pyrenaicus*, a process that allowed the contact with other *Squalius* species and the subsequent introgression of non-*S. pyrenaicus* genes in the complex. Indeed, until now, it was assumed that all *S. alburnoides* bore *S. pyrenaicus*-like mtDNA, since only one exception was reported: one *S. alburnoides* from Mondego River with a *S. carolitertii*-like mtDNA (Alves *et al.* 1997b). This scenario had at least three possible explanations: 1) the *S. alburnoides* populations had limited contacts with the sympatric *Squalius* species due to habitat segregation; 2) although coexistent, crosses between *S. alburnoides* and other *Squalius* species were uncommon; or 3) interspecific crosses occurred but mainly between *S. alburnoides* females and males of the other *Squalius* species, the reverse crosses being absent or extremely rare.

However, using samples with higher numbers of individuals (**Chapters 8** and **9**), the above described scenario of almost genetic uniformity of the complex at the mtDNA level was substantially modified. It was demonstrated that in river basins where the complex is sympatric with *S. carolitertii* and *S. aradensis*, the levels of introgression were surprisingly high: 12.9%, 21.7% and 90.5% in Douro, Mondego and Quarteira river basins, respectively. Although not formally described, considering the *S. pyrenaicus* population of Sado River as a new species (based on the divergence values calculated for the *cytb* sequences – **Chapter 8**) the estimated level of introgression of non-*S. pyrenaicus* mtDNA in the *S. alburnoides* population of this drainage was also extremely high (91.7%).

As a consequence, we have to admit that crosses with the sympatric *Squalius* females are frequent or would have had to be frequent at least in an initial stage of the colonisation of a new river basin. Considering this later hypothesis, once the non-*S. pyrenaicus* mtDNA was disseminated in the complex, the *S. alburnoides* females could ensure its perpetuation throughout generations, without further interventions of the sympatric *Squalius*. However, this hypothesis is not corroborated by the analysis of the nuclear and mtDNA haplotype diversity since *S. alburnoides* from Douro, Mondego, Quarteira and Sado exhibited not only ancient but also more recently derived haplotypes that were also found in the sympatric *Squalius* (**Chapter 8**). Thus, it seems reasonable to conclude that crosses between *S. alburnoides* males and females of the sympatric *Squalius* species are frequent.

The introgression of *S. alburnoides* genes in other *Squalius* populations, as likely evidenced by the five *S. carolitertii* specimens from Mondego with *S. pyrenaicus*-like mtDNA (**Chapter 8**) and also by the finding of four individuals with *S. carolitertii*-like mtDNA but with allozymes that were typical of *S. pyrenaicus* by Brito *et al.* (1997) in the Mondego River, may raise taxonomical problems that will be addressed in section 10.4.4. A similar case of introgression between two allopatric sexual species mediated by a hybrid complex was reported in the genus *Poeciliopsis* (Mateos & Vrijenhoek 2005), in the loach complex *Cobitis hankugensis-longicarpus* (Saitoh *et al.* 2004) and in salamanders (Bogart 1989).

The reverse situation of introgression of non-*S. pyrenaicus* genes in *S. alburnoides*, may result in the arising of novel traits in the complex, contributing to its phenotypic and genotypic reshaping. Indeed, discriminant analyses performed on the morphometric data compiled (**Chapter 3**) but unpublished due to a low number of samples, revealed that when PAA specimens were grouped according to the respective sympatric species (*i.e.* the ones that were sympatric with *S. pyrenaicus* – N=13, *S. carolitertii* – N=12, and *S. aradensis* – N=6) significant differences were detected amongst all the three groups at $p < 0.05$, without misclassifications. Also, AA specimens from Quarteira (N=11) were also significantly discriminated from AA specimens from Tagus (N=6) at $p < 0.05$, without misclassifications. Thus, although it cannot be stated firmly, there seems to be a tendency for geographical differentiation within, at least, the PAA and AA *S. alburnoides* forms.

Introgressive hybridization may have considerable impacts in fishes being responsible, for instance, for the speciation of *Gila seminuda* (DeMarais *et al.* 1992); for the complete substitution of the mtDNA of *Salvelinus fontinalis* (Bernatchez *et al.* 1995) and of *Gila robusta* (Gerber *et al.* 2001); and for the imperilment of wild gene pools by the introgression of genes from stocked species in wild populations, as happened in brown trout (Machordom *et al.* 2000; Almodóvar *et al.* 2001).

10.4 Maintenance and present day dynamics of the complex

10.4.1 Intraspecific reproductive behaviour

The reproductive behaviour of cyprinids is quite diverse, ranging from male territoriality, nest building and even male parental care to simple promiscuous mass spawning (Johnston 1999). In a somewhat intermediate position are the males of *S. alburnoides*, that congregate in the spawning site to court females and even engage in brief episodes of defence, although they frequently leave the area (**Chapter 6**) so that the situation is not fully comparable to a typical male territoriality as found in other fish species (Turner 1986). Thus, the spawning mode of *S. alburnoides* configures neither a typical situation of establishment of territories nor leks, but instead a situation of male communal occupation of spawning sites to which females are attracted. In fact, by definition, lek polygyny occurs when each male defends a small territory with no resources located within a display site ("lek") that females visit only to spawn or to choose a mate (Turner 1986; Goodenough *et al.* 1993). In addition, some authors suggest that in the case of animals with external fertilization the term "lek" should be avoided (Wedekind 1996 and references therein). On the other hand, true territoriality occurs when males defend well defined areas from where intruders are expelled (Goodenough *et al.* 1993; Wilson 2000).

The observed spawning mode of *S. alburnoides* seems to configure a different situation: males do not defend exclusive display sites in the spawning area like in lek polygyny, neither alternate between bourgeois and parasitic tactics as occurs in lek-like mating systems (Wedekind 1996), not even defend the spawning site as do isolated

territory owners. Indeed, the reproduction of this complex configures an intermediate situation with lek-like and spawning site signalling behaviours alternating with behaviours that are typical of mass spawning reproduction when receptive females are in the spawning site.

In fact, *S. alburnoides* males gather in spawning sites, probably well suited for the development of the eggs, and perform a series of courtship displaying that attract the female to their vicinity (**Chapter 6**). Like what happens in other broadcast spawners (Taborsky 1999), *S. alburnoides* males try to be closer to the female than the competitors in space and time. Their displays are more or less stimulating to females and, consequently, females apparently have the ultimate choice of the mate and localization of the spawning in space and time.

Often females assume an exploratory behaviour, entering the spawning area without descending to the spawning site and generally these inspections are followed by chasing and leading behaviours performed by the courting male(s) (**Chapter 6**). Such a pattern was already reported for other cyprinid species like *Cyprinella whipplii* (Pflieger 1965), *C. leedsii* (Rabito & Heins 1985), *Pimephales promelas* (Cole & Smith 1987), *Nocomis biguttatus* (Vives 1990) *Rutilus rutilus* (Wedekind 1996) and *Rhodeus ocellatus* (Kano 2000). When a *S. alburnoides* female effectively entered the spawning site, her touching of the substratum with the snout in a head-down position seemed to be the trigger to defuse the typical pre-spawning behavioural sequence. Other female behaviours that signal the readiness to initiate spawning and seem to be strong stimuli for males were already described for cyprinids, like the approach and swimming beneath the male one or several times in *Nocomis biguttatus* (Vives 1990) or the circling movements performed up and down along a vertical wall by *Iberochondrostoma lusitanicum* females (Carvalho *et al.* 2002).

According to the eco-ethological classification of fishes proposed by Balon (1975), *S. alburnoides* is an open substratum lithophil spawner. This broadcasting behaviour refers to the release and abandonment of eggs and sperm over an unprepared substratum and is considered the primitive spawning mode in minnows (Johnston 1999). However, the presence of courting behaviour in this category of fishes suggests that female mate choice may be occurring (Katano & Hayakama 1997), as was now confirmed

(**Chapter 6**): the AA males seemed to be favoured to the detriment of hybrid males (see below).

The experimental setting used (**Chapter 6**) intended to mimic the natural conditions not only in terms of temperature, light and habitat, but also in terms of interactions between individuals. As fish were able to reproduce virtually with all the available mates, it is assumed that the results of the ethological observations provide some degree of biological realism.

In this nearly-natural context, agonistic behaviours by an isolated hybrid male towards intruding hybrid and nonhybrid males were observed outside the spawning area (C. Sousa-Santos *unpublished data*), which seems to have been a consequence of the small density of males in that area. Such male-male interactions are common in several cyprinids – e.g. *Campostoma pullum* (Miller 1962), *Cyprinella whipplii* (Pflieger 1965), *Cyprinella leedsi* (Rabito & Heins 1985), *Nocomis biguttatus* (Vives 1990), *Notropis baileyi* and *Notropis chrosomus* (Johnston & Kleiner 1994), *Hemibarbus barbus* (Katano & Hakoyama 1997) and *Rhodeus sericeus* (Smith *et al.* 2004).

Previous studies with *R. sericeus* and *C. pullum* have demonstrated that in the presence of several males, due to the consequent higher male intrusion pressure, territorial males abandon the defence of their territories and male-male aggression becomes sporadic or inexistent (Miller 1962; Mills & Reynolds 2003; Reichard *et al.* 2004a, b).

One interesting situation for the dynamics of the complex is the possibility of existence of alternative mating tactics according to the density of males in the spawning area, as observed by Mills & Reynolds (2003) and Reichard *et al.* (2004a, b) in *R. sericeus*. In the presence of few males of *S. alburnoides*, as normally occurs in nature since in the majority of the river basins the sex ratio is extremely biased towards females, it is expected that males could establish territories. In such situations, it is known that larger and stronger males will be dominant over intruders (Candolin & Voigt 2001; Mills & Reynolds 2003), which lead us to predict that hybrid males would be more successful than nonhybrid males. This prediction is corroborated by the fact that, in the only observed situation of territoriality, a hybrid male defended an area in which AA males were not allowed to enter.

Conversely, in situations of high density of males, usually no territories are established and larger males are expected to chase other males significantly less often (Mills & Reynolds 2003). Moreover, the absence of control of spawning sites by dominant males offers more opportunities for smaller males to compete (Mills & Reynolds 2003), which may explain why hybrid and nonhybrid males were observed together in the same spawning site and even in the same spawning act (**Chapter 6**).

A possible explanation for the low density of fishes outside the spawning area and for the fact that a single spawning event was observed there may be its poor quality in what concerns aeration. Thus, females may be able to discriminate spawning areas that are suitable for egg deposition (as described for *R. sericeus* by Smith *et al.* 2000, 2001), which could explain the high density of fish in the spawning area where all but one spawnings observed during the breeding season took place (**Chapter 6**). Perhaps as a consequence of this high density of fish, no territory has been formed, situation that is similar to the one described for *Rutilus rutilus* and *Scardinius erythrophthalmus*. Svårdson (1952) postulated that the males of these species had no territories but instead were concentrated in groups of 20 to 50 specimens *per* m² of the spawning ground and females occasionally entered this "male belt" after which they were immediately chased by the males.

The hypothesis that the reproductive behaviour of *S. alburnoides* may be dependent on varying male densities could eventually explain major differences in the genetic composition of the complex in different drainages (as will be detailed in section 10.4.3).

The results presented in **Chapter 6** also highlighted the important active role of nonhybrid males in the reproduction of *S. alburnoides* females. Since females seem to prefer nonhybrids to the detriment of hybrid males (at least in a situation of high density of males) in the populations where these nonhybrid males exist their perpetuation through generations might be related to their ability to attract females to spawn with them. Indeed, by fertilizing the A-eggs of triploid females (the most common form of females in the populations where AA males exist – Alves *et al.* 2001), the nonhybrid males are guarantying the production of more males of their own kind.

In addition to the effect of the density of males in the spawning area, female choice may be also conditioned by the existence of significant morphological differences

between nonhybrid, diploid, triploid SAA and triploid SSA males (**Chapter 3**). The effect of the morphology may either be direct (*i.e.* females might show preference by specific morphological traits) or indirect, when a spatial segregation of the different forms of males motivated by differences in morphology would likely affect female choice by a conditioned availability of the distinct males.

10.4.2 Interspecific reproductive behaviour

The ability of *S. alburnoides* to cross with other *Squalius* species (at least with other *Squalius* females), generating viable offspring is extremely well expressed in the considerable extent of introgressed mitochondrial genes in the complex (**Chapters 8 and 9**). Interestingly, the levels of introgression in the northern populations of Douro and Mondego are much lower than those of the southern drainages of Sado and Quarteira, which may be related to the presence of AA males in the later drainages. Indeed, and although a detailed ethological analysis was not performed, it was observed that nonhybrid males displayed courtship behaviours towards *S. pyrenaicus* females that were similar to those performed towards *S. alburnoides* females. Moreover, these small nonhybrid males assumed an extremely active sneaking behaviour in the presence of courting pairs of *S. pyrenaicus* by placing themselves between the members of the pair when egg laying was about to occur. Thus, AA males, acting as sneakers, may be responsible for the foundation of new hybrid lineages through crosses with non-*S. pyrenaicus* females, as apparently occurred in the Quarteira and Sado drainages where the typical mtDNA of the complex was almost replaced (see section 10.3.2). This hypothesis is corroborated by the fact that, as mentioned above, populations without nonhybrids show considerably lower levels of introgression of other *Squalius* genes.

If this behavioural trait of nonhybrids was inherited, the origin of *S. alburnoides* complex might be explained by the existence of interspecific sneaking behaviour promoting the fertilization of *S. pyrenaicus* eggs by sperm of the sneakers belonging to the paternal species. This hypothesis would also explain the postulated unidirectionality of the original crosses and the consequent *Squalius*-like mtDNA of the complex (see also section 10.2.4).

Interspecific hybridizations due to male sneaking behaviour were observed in the Atlantic salmon (Garcia-Vazquez *et al.* 2002): *Salmo salar* males fertilized up to 65% of the eggs laid by brown trout (*S. trutta*) females spawning with males of its own species. However, in this case, low survival of the hybrids and agonistic behaviours performed by trout males towards sneakers resulted in low numbers of interspecific hybrids.

10.4.3 Variations among populations

Considering the hypothesis advocated in **Chapter 8** that *S. alburnoides* had a single origin in the bulk of Iberia and the results of the simulations presented in section 10.2.5 showing that two generations after the initial hybridization events all the forms of the complex could have already been originated, it should be expected that all these forms would have colonised the river basins of the present day distribution range of the complex. However, the Portuguese *S. alburnoides* populations exhibit differences in composition (*i.e.*, some forms are lacking in some populations - Table 4) in relation to what was expected from the theoretical model (Table 1).

Table 4 — Known occurrences of the *S. alburnoides* forms in the Portuguese river basins where the complex is found (data compiled from Alves *et al.* 2001 and Pala & Coelho 2005) — “x” indicates presence.

Drainage	Females			Males			
	2n	3n	4n	AA	2n	3n	4n
Douro	x	x	x		x	x	x
Mondego	x	x			x	x	x
Tagus	x	x	x	x	x	x	x
Sado	x	x	x	x		x	x
Guadiana	x	x		x		x	

As shown in Table 4, with the exception of the Tagus population, for all the other *S. alburnoides* populations at least one form was never reported.

In addition to differences in composition, the populations of *S. alburnoides* also exhibit differences in the relative abundances with which the same genome constitution occurs among drainages. Authors like Alves *et al.* (2001) and Pala & Coelho (2005) reported the relative abundances of the distinct forms but presented them as intervals of percentage values that illustrated their variation among collection episodes or among samplings made in different rivers of the same drainage – Table 5. Due to the difficulty in detecting both genders (outside the reproductive period) and distinct ploidies when sampling, these data were never compiled in absolute numbers to get a more accurate picture of the reality. Despite of that, the data presented in Table 5 are informative concerning the most abundant forms of males and females and also concerning the forms that were never reported for each population.

Table 5 – Minimum and maximum relative frequencies of the different forms of both genders of the *S. alburnoides* found in the northern and southern river basins (adapted from Alves *et al.* 2001 and Pala & Coelho 2005). For simplicity reasons, the mentioned genome constitution of tetraploids was CCAA or PPAA, however, asymmetrical tetraploids (CCCA, PPPA, CAAA and PAAA) may also occur.

		Northern river basins			Southern river basins			
	Ploidy	Genome	Douro	Mondego	Genome	Tagus	Sado	Guadiana
females	2n	CA	0-4%	0-10%	PA	0-15%	36-77%	0-35%
	3n	CAA/CCA	36-90%	80-90%	PAA/PPA	50-100%	19-70%	11-88%
	4n	CCAA	0-1%	-	PPAA	0-10%	0-2%	-
males	2n	CA	10-14%	5-14.9%	PA	0-23%	-	-
	3n	CAA/CCA	0-1%	5.3%	PAA/PPA	0-22%	0-4%	0-5%
	4n	CCAA	0-1%	0-5%	PPAA	0-14%	0-2%	-
	2n	AA	-	-	AA	0-16%	0-48%	8-89%

Although the sex determination mechanism in *S. alburnoides* remains unclear, experimental crosses performed by several authors (Alves *et al.* 1998, 1999, 2004;

Gromicho & Collares-Pereira 2004; Crespo-López *et al.* 2006) have yielded important information regarding the analysis of the sex of the offspring obtained from several crosses involving the *S. alburnoides* forms – Table 6.

Table 6 – Sex of the offspring resultant from several crosses involving distinct *S. alburnoides* forms (M-male, F-female, ?-undetermined). Data compiled from Alves *et al.* 1998, 1999, 2004; Gromicho & Collares-Pereira 2004; and Crespo-López *et al.* 2006.

FEMALES	MALES	AA	PA	PPAA	PP
	gametes	a	pa	pa	p
PA	pa	40 F (5 crosses)	9 M + 7 ? (2 crosses)	3 F + 6 M + 15? (1 cross)	
PAA	a	75 M + 26 ? (8 crosses)		11 F + 1 M (2 crosses)	136 F + 32 M + 43 ? (12 crosses)
	aa	6 M + 6 ? (2 crosses)			9 F + 1M + 9 ? (2 crosses)
	paa				11 F (1 cross)
PPAA	pa			1 F + 1 ? (2 crosses)	

Hypothetical explanations for the virtual absences of some *S. alburnoides* forms in some populations (Guadiana, Sado, Mondego and Douro), based on the mate preferences detailed (**Chapter 6**) and on the data compiled in Tables 5 and 6, are detailed below:

i) Guadiana populations

As shown in Table 7, all the known forms of *S. alburnoides* in the Guadiana River may be originated through intraspecific crosses with the exception of triploid males, whose origin depends on crosses between *S. alburnoides* females and *S. pyrenaicus* males. Crosses between PA females and AA males might be considered another possible route to triploid offspring however experimental crosses demonstrated that this route yielded all-female progeny (Table 6).

The extremely low abundance of triploid males may explain the virtual absence of tetraploids (that may be originated by the likely uncommon crosses between triploid males and females) and possibly indicates that crosses with *S. pyrenaicus* males are rare. It is likely that crosses involving females of this sympatric species also occur sporadically, as corroborated by the low number of *S. alburnoides* mtDNA haplotypes shared with *S. pyrenaicus* from Guadiana: three out of 27 of the mtDNA different haplotypes found in *S. alburnoides* from Guadiana (**Chapter 8**). This apparent low frequency of interspecific crosses may be a consequence of consistent mate preferences, of spatial segregation between the two species or of the depletion of the *S. pyrenaicus* populations in this river basin (thus making the crossings with *S. alburnoides* less probable).

Table 7 —*S. alburnoides* forms and their respective relative frequencies known for the Guadiana drainage (data from Alves *et al.* 2001).
*virtually unviable progeny (the maximum number of chromosome sets known for the complex is exceeded)

			8-89%	0-5%		
MALES			AA	PAA	PPA	PP
FEMALES gametes			a	paa	ppa	p
0-35%	PA	pa	PAA	*	*	PPA
11-88%	PAA	a	AA	PAAA	PPAA	PA
		aa	AAA	*	*	PAA
		paa	PAAA	*	*	PPAA
	PPA	p	PA	PPAA	PPPA	PP
	PP	p	PA	PPAA	PPPA	PP

On the other hand, the formation of PPA females requires the involvement of *S. pyrenaicus* males but their apparent rarity (one case reported by Crespo-López *et al.* 2006) corroborates the above mentioned scarcity of interspecific crosses.

PPA females may act as P-genome donors and contribute to the formation of PA females by crosses with AA males. Alternatively, and as PPA females are rare, the formation of the PA females may also result from crosses involving *S. pyrenaicus* - between PP females with AA males or between PP males with PAA females. Experimental

crosses demonstrated that this later cross yields progeny of both sexes, thus, if this was the route for the production of PA females it was expected that at least some PA males would have been detected in this drainage, which was not the case. Therefore, it is likely that the former cross, between PP females and AA males, would explain the formation of all-female PA *S. alburnoides*. The ethological observations detailed in section 10.4.2 regarding the vigorous displays performed by AA males towards *S. pyrenaicus* females also corroborate this hypothesis.

In these populations, the relative frequency of AA males is extremely high, which implies that PAA females would also have to occur in high frequencies since the cross PAAxAA is the only pathway to generate AA males in this basin. This hypothesis is strongly supported by the ethological observations (**Chapter 6**), which demonstrated that *S. alburnoides* females show a preference to spawn with nonhybrid males to the detriment of hybrid males.

On the other way, the production of a high number of PAA females requires frequent crosses between PA females and AA males or between *S. pyrenaicus* males and PAA females producing diploid eggs (a situation that seems rarer compared with the production of haploid A-eggs – Alves *et al.* 2004). The higher number of eggs produced by diploid females when compared to that produced by triploid ones in the Murtega River (Guadiana drainage) (Ribeiro *et al.* 2003), may also explain the observed high percentage of triploids in this drainage.

The presented arguments and hypotheses aiming to explain the genetic dynamics of the complex in Guadiana populations also applies to Sado populations, whose composition seems to be very similar to that of Guadiana (Table 5), except that in the Sado, tetraploids of both sexes were already reported (although in a very low relative frequency – less than 2%).

ii) Mondego populations

As hypothesized for Guadiana, the absence of tetraploid females and the apparently rare tetraploid males in Mondego might result from the low abundance of triploid males (Table 8).

However, in contrast to Guadiana, CCA triploids were formed in these populations in a considerable number (nine out of a total of 21 triploids – **Chapter 8**). The formation of this genome constitution requires interspecific crosses with the sympatric *Squalius* species, a situation that is corroborated (at least in what concerns crosses involving *S. carolितertii* females) by the 21.7% of introgression of *S. carolितertii* mitochondrial genomes in Mondego populations.

Table 8 - *S. alburnoides* forms and their respective relative frequencies known for the Mondego drainage (data from Alves *et al.* 2001 and Pala & Coelho 2005).

			5-14.9%	5.3%		0-5%	
			MALES	CAA	CCA	CCAA	CC
			FEMALES gametes	ca	caa	cca	c
0-10%	CA	c		CCA	CCAA	CCCA	CCA
80-90%	CAA	a		CAA	CAAA	CCAA	CAA
	CCA	c		CCA	CCAA	CCCA	CCA
	CC	c		CCA	CCAA	CCCA	CCA

Crosses between CCA or CA females and CC males reconstitute specimens that are homozygotic for the *S. carolितertii* genome complement (Table 8), which constitutes an example of how *S. alburnoides* may be responsible for the introgression of genes in other *Squalius* species.

Interestingly, all the possible crosses between the *S. alburnoides* forms found in Mondego originate only triploid or tetraploid offspring - the formation of CA males and females depends on crosses with *S. carolितertii*.

The situation in the Douro populations seems to be very similar to that of Mondego, except in the fact that rare tetraploid females (less than 1%) were already reported to date for Douro populations (Table 5), but more samplings are needed.

The absence of nonhybrid AA males in the northern populations (Alves *et al.* 2001; Pala & Coelho 2005; **Chapter 6**) may have resulted from the effect of ecological conditions that were adverse to their small size and to their preference to shallow, warm and still waters of the Mediterranean-type southern rivers (Martins *et al.* 1998). These adverse conditions, namely stronger currents, colder water and torrential regimes, might have led

to severe depletions and, ultimately, to the secondary loss of the small AA males, even if the northern *S. alburnoides* females preferred them as mates. Moreover, it is known that the northern drainages of Mondego and Douro were colonised, respectively, around 0.05 and 0.01 MY ago (**Chapter 8**). Thus, it is likely that the last glaciation, whose effects were more dramatic in the north due to the higher altitude and latitude (Hewitt 1996), had an important role in the depletion and/or extinction of the AA males. Once extinguished from a population the nonhybrid males cannot be originated *de novo* since their production is self-dependent, *i.e.*, they are sired exclusively by males with the same genome constitution, which has serious implications for conservation – see section 10.4.4.

When comparing the dynamics of the *S. alburnoides* populations from southern and northern drainages, it seems reasonable to admit that the composition and relative frequency of the forms of the complex seems to be extremely affected by female mate choices. Indeed, the female preference to spawn with nonhybrid males (**Chapter 6**) is apparently the main factor to explain, in the southern drainages, the perpetuation of PAA females and AA males (the most common forms) and the minimal intervention of *S. pyrenaicus* males in the genetic dynamics of the populations. The same behaviour of the nonhybrid males likely contributes to a renewed interbreeding with *Squalius* females.

Contrastingly, in the northern basins, from where nonhybrid males are apparently absent, if *S. alburnoides* females only crossed with conspecifics, the resultant progeny would be triploid or tetraploid, and the diploid hybrid form (CA) would be condemned to extinction. Thus, its maintenance must be ensured by crosses with the sympatric *S. carolitertii*.

As shown above, the reasons for the discrepancies in composition among drainages are unlikely related to sampling deficiencies but instead may rely on factors that increase the relative frequency of some forms and in the secondary loss of other forms (assuming, as mentioned above, that all drainages were colonised by migrants of all the forms of the complex). Thus, the abundance and survival of *S. alburnoides* forms may be influenced by ecological factors like unfavourable climatic conditions (in the last 700.000 years, precisely the estimated age of the complex, major ice ages occurred approximately every 100.000 years – Hewitt 1996), abnormally destructive winter river discharges, anoxia and food scarcity during summer droughts (when the populations are confined to pools), intense predation due to smaller size, degradation of the preferred habitat or more

intense angling due to larger size, amongst others. It is also plausible to admit the influence of stochastic factors at the foundation of a new population. Indeed, if, by chance, in the process of colonisation of a given river basin, some forms were not represented in the pool of migrants, the dynamics of the complex in that population would necessarily be distinct when compared to a population that received all the forms of the complex.

Regardless of the factor that is at the origin of the depletion/extinction of a given genome constitution, all the changes in the composition of the complex at population level, whether they occurred in the past or are a result of ongoing processes, will drastically affect the dynamics of the complex.

The effects of mate preferences in the frequency and survival of the forms of the complex may also play an extremely important role in the dynamics of the *S. alburnoides* populations. If the composition of the complex (in terms of the number of forms present and of their relative frequencies) was identical among populations, the existence of mate preferences (see section 10.4.1) would probably affect them in a similar way. On the contrary, since the complex is very likely to have a distinct composition according to the drainage considered, the mates available will be distinct and, consequently, the partners available for choice differ among populations. Moreover, female preference seems to be affected by the density of males (see section 10.4.1), with the hypothetical consequences described below using as an illustrative example the Tagus population.

The most abundant females in the Tagus basin have a PAA genome and produce, through hybridogenetic meiosis, oocytes with A genome (only rarely AA and PAA oocytes - Alves *et al.* 2001, 2004). In a natural population with a low density of males it is expected that the most frequent crosses would be with PA males (the most abundant hybrid form of male in most of Tagus populations) since, due to their size, they would more likely establish and maintain territories from where they expel other males (this was at least the situation during the behavioural observations when the male density was low - section 10.4.1). As these males produce fertile unreduced sperm (Alves *et al.* 1999), the most common progeny will have a PAA genome like the female progenitor. In situations where males are more abundant, however, crosses with AA males will be facilitated as no territories will be expected, generating progeny with the same genome as the male progenitor (AA) (since these males undergo a normal meiosis). Due to sex determination

mechanisms as yet unknown, individuals with PAA genome are more commonly females and the form AA is almost exclusively composed of males.

This differential density of males in the spawning areas may be caused by ecological factors. Indeed, in contrast to what happens in northern rivers, the Mediterranean-type southern rivers are subjected to extreme summer droughts, being frequently reduced to small pools where fish are retained in extremely high densities, which may also be one possible explanation for the success of AA males in these rivers and their virtual absence in the northern populations.

Also, if these expectations are correct, they would be corroborated by the relative proportions of the forms of the complex found in different river basins. Thus, if a given population has a more balanced sex ratio, then from generation to generation more AA males will be originated. In the opposite scenario, if a given population had an initial sex ratio extremely biased towards females, then the future of the males is threatened in that population since the most common crosses will generate mostly females (with PAA genome). In fact, the survival of the AA form seems to be self-dependent since the more AA males exist, the more AA males will be generated. In agreement with this prediction, it is known that the frequency of AA males tends to be bimodal: in some rivers (southern populations) they occur in considerable numbers while in others (northern populations) they are virtually absent - in these situations where *S. alburnoides* nonhybrid males are absent or rare, the populations seem to be maintained with diploid hybrid males or males of other *Squalius* species. Data collected in southern populations by Collares-Pereira (1983) corroborate this hypothesis: in the samples from Tagus River Basin, PA males were more abundant than AA males (59 against 16, respectively); and in the Guadiana River AA males were very common but PA males seem to be extremely rare if not totally absent.

In conclusion, the genetic dynamics of the *S. alburnoides* complex in a river basin seems to depend on at least three factors, regulated by stochastic, ecological and/or behavioural causes: the availability of the different forms of the complex; the proportions in which they occur; and the frequency of intra- and interspecific crosses.

10.4.4 Implications for conservation

Johnston (1999) suggested that spawning modes are important to the formulation of recovery plans and proactive conservation efforts. According to this author, species with simpler spawning modes reveal a tendency to be more imperilled than species exhibiting complex behaviours such as mound-building, pit-building and egg-clustering.

Although minnows comprise the largest family of freshwater fishes (Nelson 1994), reports on their reproductive behaviour are scarce. Exception is made to the North American cyprinids, whose behavioural strategies received more attention (Johnston 1999 and references therein), in part due to the need of establishing conservation measures.

In Europe, in spite of the large number of native species and the endangered status of many of them (Collares-Pereira *et al.* 2002), studies on the reproductive behaviour are limited to few species [e.g. *Rutilus rutilus* (Svärdson 1952; Diamond 1985; Wedekind 1996); *Scardinius erythrophthalmus* (Svärdson 1952); *Abramis brama* (Diamond 1985; Poncin *et al.* 1996); *Barbus barbus* and *Barbus meridionalis* (Poncin *et al.* 1994, 1997); *Gobio gobio* (Poncin *et al.* 1997); *Rhodeus sericeus* (Smith *et al.* 2004); *Iberochondrostoma lusitanicum* (Carvalho *et al.* 2002); *S. alburnoides* (**Chapter 6**)].

Besides being important to the understanding of the dynamics and evolution of the *S. alburnoides* complex, the knowledge about the reproductive behaviour is also relevant to conservation efforts. Fish that aggregate in large numbers to spawn over suitable substratum, may be seriously imperilled if alterations of the physical habitat reduce the size of spawning areas or cause the fragmentation of populations (Johnston 1999; Shumway 1999). This situation is particularly dramatic for *S. alburnoides* since, as demonstrated in section 10.4.3, even a small decrease in the relative frequency of a genome constitution may change the genetic dynamics of the population and the formation of one or more forms may eventually be imperilled.

A paradigmatic example of an irreversible secondary loss in the *S. alburnoides* complex is the absence of nonhybrid AA males in the northern populations of Douro and Mondego. Indeed, as explained in section 10.4.3, the survival of AA males is strictly dependent on the fertilization of haploid eggs from PAA females by sperm from males with the same nonhybrid AA genome. Thus, any impediment to this kind of crosses or any serious threat to the survival of AA males or PAA females would result in the depletion or in the extinction of the AA form of the complex.

Another threat to the integrity of the *S. alburnoides* populations may be the hybridization with exotic fish, as might eventually occur in River Guadiana after the recent introduction of *Alburnus alburnus* (Vinyoles *et al.* 2007), a species that is related to the paternal ancestor of the complex (**Chapter 7**).

As the populations of *S. alburnoides* have characteristic compositions and, consequently, each of them has an eventually exclusive genetic dynamics, they should be treated as unique entities when delineating conservation plans and strategies, as proposed by Bohlen & Rab (2001) for several locally unique complexes of spined loaches.

Moreover, situations that may have a predictable negative effect on a previously balanced *S. alburnoides* population, such as habitat destruction or the introduction of exotic fish, must be actively prevented.

Another issue relevant to conservation which in a sense is the reverse of the one discussed above is the need to prevent *S. alburnoides* to reach new drainages. With its ability to hybridize with all the endemic *Squalius* and of transferring *Squalius* genes from species to species, accidental or deliberate introduction of *S. alburnoides* in new drainages may irreversibly degrade the genetic constitution of the local *Squalius* populations. Thus, while we need to conserve the populations of *S. alburnoides* in their habitats, we must be equally effective in preventing their spread to areas where they are not native.

10.5 Sexual parasitism or sexual autonomy in *S. alburnoides*?

It has been generally assumed that animal hybrid lineages tend to be short-lived on an evolutionary time scale (see section 1.1). As in most cases they are mainly constituted by females and, in a way or another, are unable to perform normal meiosis or perform it rarely (especially in the case of intergeneric hybrids), their ability to achieve recombination is impaired (Vrijenhoek 1989). Thus, the potential to retain genetic diversity and to eliminate deleterious mutations is reduced, limiting their ability to respond adaptively to

environmental changes (reviewed in Arnold 1997). These limitations of non-sexual (*i.e.* non-recombinant) hybrid lineages are aggravated as many of them depend on one or more bisexual species to fertilize the gametes, thus becoming dependent on their availability (Dawley 1989). In summary, *sensu* Dawley (1989), asexuals are of interspecific origin, almost (or nearly) constituted by females, with aberrant gametogenic mechanisms that inhibit recombination (clonal inheritance) and often include polyploids. However, this author highlighted the fact that there are exceptions to this definition (some all-female lineages incorporate sperm, others show some level of recombination and in others males are also present) but, despite of that, all asexuals reproduce by altered gametogenic mechanisms and transmit at least part of their genome with limited or without recombination.

The *S. alburnoides* complex was initially described as an asexual complex due to the preponderance of females (Collares-Pereira 1985, 1989). However, as more studies have been conducted, several exceptions to the general pattern exhibited by asexuals were revealed: 1) there is, in comparative terms, a considerable number of males; 2) most males produce viable and fertile sperm; 3) although much of the gametogenesis is aberrant, the reproduction involves singamy and karyogamy; 4) recombination between homospecific genomes occurs in triploid females, in diploid nonhybrids and in balanced tetraploids (Alves *et al.* 2004; Pala & Coelho 2005; Crespo-López *et al.* 2006); and 5) there is no strict dependency on a bisexual species as a sexual host.

Concerning the relationship between *S. alburnoides* and sympatric bisexual species, the absence of interspecific crosses would likely result in populations with distinct compositions but the viability of the complex would not be compromised. Moreover, the introduction of P-genomes in the complex may be ensured by PPA females (which produce P-eggs - Crespo-López *et al.* 2006), contributing even more to the independency of the complex from other *Squalius* species. Thus, *S. alburnoides* may be considered an autonomous complex whose maintenance, in contrast to its origin, does not depend on crosses with other *Squalius* species.

The ability to perform recombination contradicts the point of view that *S. alburnoides* may constitute an evolutionary “dead-end”, as predicted long before (Collares-Pereira 1989). Indeed, authors like Pala & Coelho (2005) postulated a short lived history for the *S. alburnoides* population of Mondego, based on a lower genetic diversity

when compared to that of southern populations and on a more limited set of distinct forms of the complex (that would potentially decrease the probability of generating novel genetic material). However, this conclusion should not be drawn without taking into account the phylogeography of the complex and the effects of the glaciations on its populations. Indeed, the results presented in **Chapter 8** showed that the northern populations of Mondego and Douro exhibited haplotype diversities that were lower than those of Tagus, Sado, Guadiana and Guadalquivir but not higher than that of the Quarteira population. Thus, it is likely that the lower diversity highlighted by Pala & Coelho (2005) for the Mondego population is a reflection of the dynamics of a population that is still recovering from the probable dramatic consequences of the last glaciation. Indeed, areas that have been colonised after the Pleistocenic ice ages are expected to harbour lower levels of genetic diversity than more southern areas that remained occupied due to the existence of varied suitable habitats (refugia) – Hewitt (1996, 2000).

This hypothesis is also corroborated by the considerably lower mtDNA haplotype diversity of the *S. carolitertii* populations from the Minho to the Vouga (ranging from 4.55% to 29.17%) when compared to the values exhibited by the *Squalius* populations located in more southern drainages: *S. carolitertii* from Mondego (40.00%), *S. pyrenaicus* from Tagus and Guadiana (66.67% and 60.00%, respectively) and *S. aradensis* from Arade (32.14%) (C. Sousa-Santos *unpublished data*). *S. torgalensis* from Mira showed only 18.75% of haplotype diversity (C. Sousa-Santos *unpublished data*), which may be a result of a low number of migrants or of bottlenecks caused by transgressions.

In conclusion, the *S. alburnoides* complex shows a peculiar positioning in the context of hybrid complexes. Indeed, it is neither a true asexual species due to the above mentioned exceptions to the typical pattern, nor a true bisexual species with a balanced sex ratio and with both sexes producing gametes with half of the somatic ploidy level after a normal meiosis. Thus, to avoid misinterpretations and, more importantly, to avoid labels that have a narrower scope, references to the reproduction mode (asexual or sexual) should be omitted and the *S. alburnoides* complex should only be designated as a “hybridogenetic complex” or a “intergeneric hybrid complex”.

Moreover, the available evidence showed that the *S. alburnoides* complex is likely an intermediate step towards tetraploidy. The emergence of a new tetraploid species with normal meiosis may be, however, prevented since the scarce tetraploids already

originated in most of the populations coexist with all the other ploidies of the complex, not being able to stabilize as an independent species. This would only occur if, by chance, a group constituted mainly by tetraploids became confined and/or reproductively isolated from the remaining forms of the complex, again highlighting the importance of the relative proportions of the different forms in the complex dynamics, as well as of establishing specifically oriented conservation measures.

10.6 References

- Albertson RC & Kocher TD. 2001. Assessing morphological differences in an adaptive trait: a landmark-based morphometric approach. *Journal of Experimental Zoology* 289: 385–403.
- Almaça C. 1976. Zoogeografia e especiação dos ciprinídeos da Península Ibérica. *Revista da Sociedade Portuguesa das Ciências Naturais* 4: 1-28.
- Almodóvar A, Suárez J, Nicola GG & Nuevo M. 2001. Genetic introgression between wild and stocked brown trout in the Douro River basin, Spain. *Journal of Fish Biology* 59 (Supplement A): 68-74.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 1996. Evidence for nonclonal reproduction in triploid *Rutilus alburnoides*. *Isozyme Bulletin* 29: 23.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 1997a. The *Rutilus alburnoides* complex (Cyprinidae): evidence for hybrid origin. *Journal of Zoological Systematics and Evolutionary Research* 35: 1–10.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 1998. Diversity in the reproductive modes of females of the *Rutilus alburnoides* complex (Teleostei, Cyprinidae): a way to avoid the genetic constraints of uniparentalism. *Molecular Biology and Evolution* 15: 1233-1242.
- Alves MJ, Coelho MM & Collares-Pereira MJ. 2001. Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: a genetic review. *Genetica* 111: 375-385.
- Alves MJ, Collares-Pereira MJ, Dowling TE & Coelho MM. 2002. The genetics of maintenance of an all-male lineage in the *Squalius alburnoides* complex. *Journal of Fish Biology* 60: 649-662.
- Alves MJ, Coelho MM, Collares-Pereira MJ & Dowling TE. 1997b. Maternal ancestry of the *Rutilus alburnoides* complex (Teleostei, Cyprinidae) as determined by analysis of cytochrome *b* sequences. *Evolution* 51: 1584-1592.
- Alves MJ, Coelho MM, Próspero MI & Collares-Pereira MJ. 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 MJ, Gromicho M, Collares-Pereira MJ, Crespo-López E & Coelho MM. 2004. Simultaneous production of triploid and haploid eggs by triploid *Squalius alburnoides* (Teleostei: Cyprinidae). *Journal of Experimental Zoology* 301A: 552-558.

Andeweg B. 2002. Cenozoic tectonic evolution of the Iberian Peninsula, causes and effects of changing stress fields. PhD Thesis. Universiteit Amsterdam, Amsterdam, 178p.

Arnold ML. 1997. Natural hybridization and evolution. Oxford University Press, Oxford, 215p.

Avise JC, Quattro JM & Vrijenhoek RC. 1992. Molecular clones within organismal clones: mitochondrial DNA phylogenies and the evolutionary histories of unisexual vertebrates. *Evolutionary Biology* 26: 225-246.

Ayres DR, Zaremba K & Strong DR. 2004. Extinction of a common native species by hybridization with an invasive congener. *Weed Technology* 18: 1288-1291.

Bagenal TB. 1978. Aspects of fish fecundity. In Gerking SD (Ed.) Ecology of freshwater fish production. Wiley, New York, pp 75-101.

Balon EK. 1975. Reproductive guilds of fishes: a proposal and definition. *Journal of Fisheries Research Board of Canada* 32: 821-864.

Belfiore NM, Hoffman FG, Baker RJ & DeWoody JA. 2003. The use of nuclear and mitochondrial single nucleotide polymorphisms to identify cryptic species. *Molecular Ecology* 12: 2011-2017.

Bernatchez L & Wilson CC. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology* 7: 431-452.

Bernatchez L, Glémet H, Wilson CC & Danzmann RG. 1995. Introgression and fixation of Arctic char (*Salvelinus alpinus*) mitochondrial genome in an allopatric population of brook trout (*Salvelinus fontinalis*). *Canadian Journal of Fisheries and Aquatic Sciences* 52: 179-185.

Bhangale TR, Rieder MJ, Livingston RJ & Nickerson DA. 2005. Comprehensive identification and characterization of diallelic insertion-deletion polymorphisms in 330 human candidate genes. *Human Molecular Genetics* 14: 59-69.

Bianco PG. 1990. Potential role of the paleohistory of the Mediterranean and Paratethys basins on the early dispersal of Euro-Mediterranean freshwater fishes. *Ichthyological Exploration of Freshwaters* 1: 167-184.

Bogart JP. 1989. A mechanism for interspecific gene exchange via all-female salamander hybrids. In Dawley RM & Bogart JP (Eds.) Evolution and Ecology of Unisexual Vertebrates. New York State Museum Bulletin 466, New York, pp 170-179.

Bohlen J & Ráb P. 2001. Species and hybrid richness in spined loaches of the genus *Cobitis* (Teleostei: Cobitidae), with a checklist of European forms and suggestions for conservation. *Journal of Fish Biology* 59 (Supplement A): 75-89.

Briggs JC. 1987. Antitropicality and vicariance. *Systematic Zoology* 36: 206-207.

- Brito RM, Briolay J, Galtier N, Bouvet Y & Coelho MM. 1997. Phylogenetic relationships within genus *Leuciscus* (Pisces, Cyprinidae) in Portuguese freshwaters, based on mitochondrial cytochrome *b* sequences. *Molecular Phylogenetics and Evolution* 8: 435-442.
- Brumfield RT, Beerli P, Nickerson DA & Edwards SV. 2003. The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology and Evolution* 18: 249-256.
- Campo S, Sardo AM, Campo GM, Avenoso A, Castaldo M, D'Ascola A, Giunta E, Calatroni A & Saitta A. 2004. Identification of paraoxonase 3 gene (PON3) missense mutations in a population of southern Italy. *Mutation Research* 546: 75-80.
- Candolin U & Voigt H-R. 2001. Correlation between male size and territory quality: consequence of male competition or predation susceptibility? *Oikos* 95: 225-230.
- Carmona JA. 1997. Evolución y modelos de reproducción en un organismo unisexual de la Península Ibérica: el complejo *Tropidophoxinellus alburnoides*. PhD Thesis. Universidad Complutense, Madrid, 325p.
- Carmona JA, Sanjur OI, Doadrio I, Machordom A & Vrijenhoek VC. 1997. Hybridogenetic reproduction and maternal ancestry of polyploid Iberian fish: the *Tropidophoxinellus alburnoides* complex. *Genetics* 146: 983-993.
- Carvalho V, Robalo JI & Almada VC. 2002. A description of the reproductive behaviour of the endangered Iberian cyprinid *Chondrostoma lusitanicum* Collares-Pereira 1980 in captivity. *Etología* 10: 23-25.
- Chávez CH & Turgeon J. 2007. Asexual and sexual hybrids between *Fundulus diaphanous* and *F. heteroclitus* in the Canadian Atlantic region. *Molecular Ecology* 16: 1467-1480.
- Clark AG. 1990. Inference of haplotypes from PCR-amplified samples of diploid populations. *Molecular Biology and Evolution* 7: 111-122.
- Cole KS & Smith RJF. 1987. Male Courting Behaviour in the Fathead Minnow, *Pimephales promelas*. *Environmental Biology of Fishes* 18: 235-239.
- Collares-Pereira, M.J. 1980. Population variability of *Pseudophoxinus hispanicus* (Steindachner, 1866) (Pisces Cyprinidae). *Arquivos do Museu Bocage* 7: 363-388.
- Collares-Pereira MJ. 1983. Estudo sistemático e citogenético dos pequenos ciprinídeos ibéricos pertencentes aos géneros *Chondrostoma* Agassiz, 1835, *Rutilus* Rafinesque, 1820 and *Anaecypris* Collares-Pereira, 1983. PhD Thesis. University of Lisbon, Lisbon, 511p.
- Collares-Pereira MJ. 1985. The "*Rutilus alburnoides* (Steindachner, 1866) complex". II. First data on the karyology of a well-established diploid-triploid group. *Arquivos do Museu Bocage (Série A) III(5)*: 69-90.
- Collares-Pereira MJ. 1989. Hybridization in European Cyprinids: evolutionary potential of unisexual populations. In Dawley RM & Bogart JP (Eds.) *Evolution and Ecology of Unisexual Vertebrates*. New York State Museum Bulletin 466, New York, pp 281-288.
- Collares-Pereira MJ, Cowx IG, Rodrigues JA & Rogado L. 2002. The conservation strategy for *Anaecypris hispanica*: a picture of LIFE for a highly endangered Iberian fish. In Collares-Pereira MJ, Cowx IG & Coelho MM

- (Eds.) Conservation of Freshwater Fishes: Options for the Future. Fishing News Books, Blackwell Science, Oxford, pp 186-197.
- Crespin L & Berrebi P. 1994. L'hybridation naturelle entre le barbeau commun et le barbeau meridional en France: compte rendu de dix années de recherche. *Bulletin Français de la Pêche et de la Pisciculture* 334: 177-189.
- Crespo-López ME, Duarte T, Dowling T & Coelho MM. 2006. Modes of reproduction of the hybridogenetic fish *Squalius alburnoides* in the Tejo and Guadiana rivers: an approach using microsatellites. *Zoology* 109: 277-286.
- Cunha C, Coelho MM, Carmona JA & Doadrio I. 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Molecular Ecology* 13: 2807-2817.
- Dawley R. 1989. An introduction to unisexual vertebrates. In Dawley RM & Bogart JP (Eds.) Evolution and Ecology of Unisexual Vertebrates. New York State Museum Bulletin 466, New York, pp 1-18.
- De la Peña A. 1995. Tertiary fishes from the Iberia continental basins: History and fossil record. *Colóquios de Paleontologia* 47, Editorial Complutense, Madrid.
- DeMarais BD, Dowling TE, Douglas ME, Minckley WL & Marsh PC. 1992. Origin of *Gila seminuda* (Teleostei: Cyprinidae) through introgressive hybridization: implications for evolution and conservation. *Proceedings of the National Academy of Sciences USA* 89: 2747-2751.
- Diamond M. 1985. Some observations of spawning by roach, *Rutilus rutilus* L., and bream, *Abramis brama* L., and their implications for management. *Aquaculture and Fisheries Management* 16: 359-367.
- Doadrio I & Carmona JA. 2003. Testing freshwater Lago Mare dispersal theory on the phylogeny relationships of Iberian cyprinid genera *Chondrostoma* and *Squalius* (Cypriniformes, Cyprinidae). *Graellsia* 59: 457-473.
- Doadrio I & Perdices A. 2005. Phylogenetic relationships among the Ibero-African cobitids (*Cobitis*, Cobitidae) based on cytochrome *b* sequence data. *Molecular Phylogenetics and Evolution* 37: 484-493.
- Doadrio I. 2001. Atlas y libro rojo de los peces continentales de España. Dirección General de Conservación de la Naturaleza - Museo Nacional de Ciencias Naturales, Madrid, 364p.
- Doadrio I, Carmona JA & Machordom A. 2002 Haplotype diversity and phylogenetic relationships among the Iberian barbels (*Barbus*, Cyprinidae) reveal two evolutionary lineages. *Journal of Heredity* 93: 140-147.
- Dowling TE & Secor CL. 1997. The role of hybridization and introgression in the diversification of animals. *Annual Review of Ecology and Systematics* 28: 593-619.
- Durand JD, Ünlu E, Doadrio I, Pipoyan S & Templeton AR. 2000. Origin, radiation, dispersion and allopatric hybridization in the chub *Leuciscus cephalus*. *Proceedings of the Royal Society of London Series B* 267: 1687-1697.
- Excoffier L, Laval G & Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47-50.
- Forsdike DR. 1981. Are introns in-series error-detecting sequences? *Journal of Theoretical Biology* 93: 861-866.

- French DJ, Archard CL, Brown T & McDowell DG. 2001. HyBeaconTM probes: a new tool for DNA sequence detection and allele discrimination. *Molecular and Cellular Probes* 15: 363-374.
- Freyhof J, Lieckfeldt D, Pitra C & Ludwig A. 2005. Molecules and morphology: evidence for introgression of mitochondrial DNA in Dalmatian cyprinids. *Molecular Phylogenetics and Evolution* 37: 347-354.
- Friend PF & Dabrio CJ (Eds.) 1996. Tertiary basins of Spain: the stratigraphic record of crustal kinematics. Cambridge University Press, Cambridge, 400p.
- Garcia-Vazquez E, Moran P, Perez J, Martinez JL, Izquierdo JL, de Gaudemar B & Beall E. 2002. Interspecific barriers between salmonids when hybridisation is due to sneak mating. *Heredity* 89: 288-292.
- Gerber AS, Tibbets CA & Dowling TE. 2001. The role of introgressive hybridization in the evolution of the *Gila robusta* complex (Teleostei: Cyprinidae). *Evolution* 55: 2028-2039.
- Goodenough J, McGuire B & Wallace R. 1993. Perspectives on Animal Behaviour. John Wiley & Sons, New York.
- Gromicho M & Collares-Pereira MJ. 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.
- Gromicho M, Coelho MM, Alves MJ & Collares-Pereira MJ. 2007. Cytogenetic analysis of *Anaecypris hispanica* and its relationships with the paternal ancestor of the diploid-polyploid *Squalius alburnoides* complex. *Genome* 49: 1621-1627.
- Hewitt G. 1996. Some genetic consequences of ice ages and their role in divergence and speciation. *Biological Journal of the Linnean Society* 58: 217-276.
- Hewitt GM. 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* 68: 87-112.
- Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405: 907-913.
- Janko K, Kotlík P & Ráb P. 2003. Evolutionary history of asexual hybrid loaches (*Cobitis*: Teleostei) inferred from phylogenetic analysis of mitochondrial DNA. *Journal of Evolutionary Biology* 15: 1280-1287.
- Johnston CE & Kleiner KJ. 1994. Reproductive behaviour of the rainbow shiner (*Notropis chrosomus*) and the rough shiner (*Notropis baileyi*), nest associates of the bluehead chub (*Nocomis leptcephalus*) (Pisces: Cyprinidae) in the Alabama River drainage. *Journal of the Alabama Academy of Science* 65: 230-240.
- Johnston CE. 1999. The relationship of spawning mode to conservation of North American minnows (Cyprinidae). *Environmental Biology of Fishes* 55: 21-30.
- Kanoh Y. 2000. Reproductive success associated with territoriality, sneaking, and grouping in male rose bitterlings, *Rhodeus ocellatus* (Pisces: Cyprinidae). *Environmental Biology of Fishes* 2: 143-154.
- Katano O & Hayokama H. 1997. Spawning behaviour of *Hemibarbus barbus* (Cyprinidae). *Copeia* 3: 620-622.
- Kearney M & Shine R. 2004. Morphological and physiological correlates of hybrid parthenogenesis. *The American Naturalist* 164: 803-813.

- Kittell MM, Harvey MN, Contreras-Balderas S & Ptacek MB. 2005. Wild-caught hybrids between sailfin and shortfin mollies (Poeciliidae, *Poecilia*): morphological and molecular verification. *Hidrobiologica* 15: 131-137.
- Konishi M & Takata K. 2004. Impact of asymmetrical hybridization followed by sterile F1 hybrids on species replacement in *Pseudorasbora*. *Conservation Genetics* 5: 463-474.
- Landegren U, Nilsson M & Kwok P-Y. 1998. Reading bits of genetic information: methods for single-nucleotide polymorphism analysis. *Genome Research* 8: 769-776.
- Letting DL, Fecteau DA, Haws TF, Reed SL, Hopkins RO, Coleman RD & Goddard KA. 1999. Unexpected ratio of allozyme expression in diploid and triploid individuals of the clonal hybrid fish *Phoxinus eos-neogaeus*. *Journal of Experimental Zoology* 284: 663-674.
- Levin DA, Francisco-Ortega J & Jansen RK. 1996. Hybridization and the extinction of rare plant species. *Conservation Biology* 10: 10-16.
- Liguori R, Argiriou A & De Simon L. 2003. A rapid method for detecting mutations of the human LDL receptor gene by complete cDNA sequencing. *Molecular and Cellular Probes* 17: 15-20.
- Lima NRW & Bizerril CRSF. 2002. Frequency-dependence of mating success in *Poeciliopsis monacha* (Pisces, Cyprinodontiformes) reproductive complex, Sonora, Mexico. *Brazilian Archives of Biology and Technology* 45: 161-169.
- Linch M. 2004. Intron evolution as a population-genetic process. *Proceedings of the National Academy of Sciences USA* 99: 6118-6123.
- Machordom A & Doadrio I. 2001. Evidence of a Cenozoic Betic-Kabilian connection based on freshwater fish phylogeography (*Luciobarbus*, Cyprinidae). *Molecular Phylogenetics and Evolution* 18: 252-263.
- Machordom A, Suárez J, Almodóvar A & Bautista JM. 2000. Mitochondrial haplotype variation and phylogeography of Iberian brown trout populations. *Molecular Ecology* 9: 1325-1338.
- Malvagia S, Morrone A, Caciotti A, Bardelli T, d'Azzo A, Ancora G, Zammarchi E & Donati MA. 2004. New mutations in the *PPBG* gene lead to loss of PPCA protein which affects the level of the b-galactosidase/neuraminidase complex and the EBP-receptor. *Molecular Genetics and Metabolism* 82: 48-55.
- Martins MJ, Collares-Pereira MJ, Cowx IG & Coelho MM. 1998. Diploids v. triploids of *Rutilus alburnoides*: spatial segregation and morphological differences. *Journal of Fish Biology* 52: 817-828.
- Mateos M & Vrijenhoek RC. 2005. Independent origins of allotriploidy in the fish genus *Poeciliopsis*. *Journal of Heredity* 96: 32-39.
- Miller RJ. 1962. Reproductive behavior of the stoneroller minnow, *Camptostoma anomalum pullum*. *Copeia* 2: 407-417.
- Mills SC & Reynolds JD. 2003. Operational sex ratio and alternative reproductive behaviours in the European bitterling, *Rhodeus sericeus*. *Behavioral Ecology and Sociobiology* 54: 98-104.
- Moya-Palomares ME. 2002. Evolución sedimentológica y geomorfológica de las Vegas Bajas del Guadiana entre Mérida y Badajoz (España). PhD Thesis. Universidad Complutense de Madrid, Madrid, 297p.

- Mulcahy DG, Spaulding AW, Mendelson JR & Brodie ED. 2006. Phylogeography of the flat-tailed horned lizard (*Phrynosoma mcallii*) and systematics of the *P. mcallii-platyrrhinus* mtDNA complex. *Molecular Ecology* 15: 1807-1826.
- Nelson J. 1994. Fishes of the World. John Wiley & Sons Inc., New York.
- Niu T, Qin ZS, Xu X & Liu JS. 2002. Bayesian haplotype inference for multiple linked single-nucleotide polymorphisms. *American Journal of Human Genetics* 70: 157-169.
- Pala I & Coelho MM. 2005. Contrasting views over a hybrid complex: between speciation and evolutionary "dead-end". *Gene* 347: 283-294.
- Pflieger W. 1965. Reproductive behavior of the minnows, *Notropis spilopterus* and *Notropis whipplii*. *Copeia* 1965(1): 1-9.
- Pitts CS, Jordan DR, Cowx IG & Jones NV. 1997. Controlled breeding studies to verify the identity of roach and common bream hybrids from a natural population. *Journal of Fish Biology* 51: 686-696.
- Placek MB. 2002. Patterns of inheritance of mating signals in interspecific hybrids between sailfin and shortfin mollies (Poeciliidae: *Poecilia Mollienesia*). *Genetica* 116: 329-342.
- Poncin P, Jeandarme J & Berrebi P. 1994. A behavioural study of hybridization between *Barbus barbus* and *Barbus meridionalis*. *Journal of Fish Biology* 45: 447-451.
- Poncin P, Philippart JC & Ruwet JC. 1996. Territorial and non-territorial spawning tactics in the bream, *Abramis brama* (Teleostei, Cyprinidae). *Journal of Fish Biology* 49: 622-626.
- Poncin P, Jeandarme J, Rinchar J & Kestemont P. 1997. Le comportement de reproduction du goujon, *Gobio gobio*, en aquarium: premiers résultats. *Bulletin Français de Pêche et de Pisciculture* 346: 547-555.
- Próspero MI & Collares-Pereira MJ. 2000. Nuclear DNA content variation in the diploid-polyploid *Leuciscus alburnoides* complex (Teleostei, Cyprinidae) assessed by flow cytometry. *Folia Zoologica* 49: 53-58.
- Rabito FG & Heins DC. 1985. Spawning behaviour and sexual dimorphism in the North American cyprinid fish *Notropis leedsii*, the bannerfin shiner. *Journal of Natural History* 19: 1155-1163.
- Reichard M, Jurajda P & Smith C. 2004a. Male-male interference competition decreases spawning rate in the European bitterling (*Rhodeus sericeus*). *Behavioural Ecology and Sociobiology* 56: 34-41.
- Reichard M, Smith C & Jordan WC. 2004b. Genetic evidence reveals density-dependent mediated success of alternative mating behaviours in the European bitterling (*Rhodeus sericeus*). *Molecular Ecology* 13: 1369-1578.
- Rhymer JM. 2006. Extinction by hybridization and introgression in anatine ducks. *Acta Zoologica Sinica* 52 (Supplement): 583-585.
- Ribeiro F, Cowx IG, Tiago P, Filipe AF, Moreira da Costa L & Collares-Pereira MJ. 2003. Growth and reproductive traits of diploid and triploid forms of the *Squalius alburnoides* cyprinid complex in a tributary of the Guadiana River, Portugal. *Archiv für Hydrobiologie* 156: 471-484.
- Ribeiro F, Beldade R, Dix M & Bochechas J. 2007. Carta Piscícola Nacional Direcção Geral dos Recursos Florestais - Fluviatilis, Lda. World Wide Web electronic publication, www.fluviatilis.com, version 01/2007.

- Ribeiro O, Lautensach H, Daveau S. 1987. Geografia de Portugal - I. A posição geográfica e o território. Edições João Sá da Costa, Lisboa, 334p.
- Robertson AV, Ramsden C, Niedzwiecki J, Fu J & Bogart JP. 2006. An unexpected recent ancestor of unisexual *Ambystoma*. *Molecular Ecology* 15: 3339-3351.
- Rodríguez-Vidal J, Cáceres L & Ramirez A. 1993. Modelo evolutivo de la red fluvial cuaternaria en el suroeste de la Península Ibérica. Actas da 3ª Reunião do Quaternário Ibérico, Coimbra, pp 93-96.
- Rosenfield JA, Nolasco S, Lindauer S, Sandoval C & Kodric-Brown A. 2004. The role of hybrid vigor in the replacement of Pecos pupfish by its hybrids with sheepshead minnow. *Conservation Biology* 18: 1589-1598.
- Rowe KC, Heske EJ, Brown PW & Paige KN. 2004. Surviving the ice: northern refugia and postglacial colonization. *Proceedings of the National Academy of Sciences* 101: 10355-10359.
- Saitoh K, Kim I-S & Lee E-H. 2004. Mitochondrial gene introgression between spined loaches via hybridogenesis. *Zoological Science* 21: 795-798.
- Salzburger W, Brandstätter A, Gilles A, Parson W, Hempel M, Sturmbauer C & Meyer A. 2003. Phylogeography of the vairone (*Leuciscus souffia*, Risso 1826) in central Europe. *Molecular Ecology* 12: 2371-2386.
- Schartl M, Wilde B, Schlupp I & Parzefall J. 1995. Evolutionary origin of a parthenoform, the Amazon Molly *Poecilia formosa*, on the basis of a molecular genealogy. *Evolution* 49: 827-835.
- Schlötterer C. 2004. The evolution of molecular markers – just a matter of fashion? *Nature Reviews in Genetics* 5: 63-70.
- Shumway C. 1999. A neglected science: applying behavior to aquatic conservation. *Environmental Biology of Fishes* 55: 183-201.
- Smith C, Reynolds JD & Sutherland WJ. 2000. Population consequences of reproductive decisions. *Proceedings of the Royal Society of London Series B* 267: 1327-1334.
- Smith C, Rippon K, Douglas A & Jurajda P. 2001. A proximate cue for oviposition site choice in the bitterling (*Rhodeus sericeus*). *Freshwater Biology* 46: 903-911.
- Smith C, Reichard M, Jurajda P & Przybylski M. 2004. The reproductive ecology of the European bitterling (*Rhodeus sericeus*). *Journal of Zoology* 262: 107-124.
- Stephens M, Smith NJ & Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* 68: 978-989.
- Svärdson G. 1952. Spawning behaviour of *Leuciscus rutilus* (Linné). *Reports of the Institute of Freshwater Research Drottningholm* 33: 199-203.
- Sykotis GP, Neumann S, Georgopoulos NA, Sgourou A, Papachatzopoulou A, Markou KB, Kyriazopoulou V, Paschke R, Vagenakis AG & Papavassiliou AG. 2003. Functional significance of the thyrotropin receptor germline polymorphism D727E. *Biochemical and Biophysical Research Communications* 301: 1051-1056.

- Taborsky M. 1999. Conflict or cooperation: what determines optimal solutions to competition in fish reproduction? *In* Almada VC, Oliveira RF & Gonçalves EJ (Eds.) Behaviour and conservation of littoral fishes. ISPA, Lisbon, pp 301-349.
- Takatani T, Takaoka N, Tatsumi M, Kawamoto H, Okuno Y, Morita K, Masutani T, Murakawa K & Okamoto Y. 2001. A novel missense mutation in human lactate dehydrogenase b-subunit gene. *Molecular Genetics and Metabolism* 73: 344-348.
- Takeuchi H, Kondo M, Daimon M, Susa S, Ueoka K, Uemura O & Togari H. 2001. Neonatal-onset hereditary coproporphyria with male pseudohermaphroditism. *Blood* 98: 3671-3873.
- Toone JR, Applegarth DA, Levy HL, Coulter-Mackie MB & Lee G. 2003. Molecular genetic and potential biochemical characteristics of patients with T-protein deficiency as a cause of glycine encephalopathy (NKH). *Molecular Genetics and Metabolism* 79: 272-280.
- Tsigenopoulos CS, Kotlik P & Berrebi P. 2002. Biogeography and pattern of gene flow among *Barbus* species (Teleostei: Cyprinidae) inhabiting the Italian peninsular and neighbouring Adriatic drainages as revealed by allozyme and mitochondrial sequence data. *Biological Journal of the Linnean Society* 75: 83-99.
- Turner G. 1986. Teleost mating systems and strategies. *In* Pitcher TJ (Ed.) The behaviour of Teleost fishes. Croom Helm Ltd, London, pp 253-274.
- Ünver B & Erk'akan F. 2005. A natural hybrid of *Leuciscus cephalus* (L.) and *Chalcalburnus chalcoides* (Güldenstädt) (Osteichthyes-Cyprinidae) from Lake Tödürge (Sivas, Turkey). *Journal of Fish Biology* 66: 899-910.
- Vinyoles D, Robalo JJ, Sostoa A, Almodovar A, Elvira B, Nicola GG, Fernández-Delgado C, Sousa-Santos C, Doadrio I, Sarda-Palomera F & Almada VC. 2007. Spread of the alien bleak *Alburnus alburnus* (Linnaeus, 1758) (Actinopterygii, Cyprinidae) in the Iberian Peninsula: the role of reservoirs. *Graellsia* 63.
- Vives SP. 1990. Nesting ecology and behavior of Hornyhead Chub *Nocomis biguttatus*, a keystone species in Allequash Creek, Wisconsin. *American Midland Naturalist* 124: 46-56.
- Vrijenhoek R. 1989. Genetic and ecological constraints on the origins and establishment of unisexual vertebrates. *In* Dawley R & Bogart J (Eds.) Evolution and Ecology of Unisexual Vertebrates. New York State Museum Bulletin 466, New York, pp 24-31.
- Wedekind C. 1996. Lek-like spawning behaviour and different female mate preferences in roach (*Rutilus rutilus*). *Behaviour* 166: 681-695.
- Wilson EO. 2000. Sociobiology. The new synthesis. The Belknap Press of Harvard University Press, Cambridge, 697p.
- Wirtz P. 1999. Mother species-father species: unidirectional hybridization in animals with female choice. *Animal Behaviour* 58: 1-12.
- Yakovlev VN, Slynko YV, Grechanov IG & Krysanov EY. 2000. Distant hybridization in fish. *Journal of Ichthyology* 40: 298-311.

Zardoya R & Doadrio I. 1998. Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. *Proceedings of the Royal Society B* 265: 1365-1372.

Zhang P, Sheng H, Morabia A & Gilliam TC. 2003. Optimal step length EM algorithm (OSLEM) for the estimation of haplotype frequency and its application in lipoprotein lipase genotyping. *Bioinformatics* 4:3.

Chapter 11

Concluding Remarks

The results of the research conducted revealed that the reproductive behaviour must have had a crucial role not only at the origin of *S. alburnoides* but also along its evolutionary history, being determinant to population structuring by influencing the genetic dynamics within the complex and also between *S. alburnoides* and the other *Squalius* species with which it has been interacting.

The specific objectives that were put forward were accomplished and the obtained results may be summarized as follows:

i) The description of a more practical method for the determination of the genome constitution of the *S. alburnoides* forms, based on the analysis of the double peaks generated by heterozygous indels during the sequencing of the beta-actin nuclear gene, allowed several approaches to the main goal of this dissertation: the investigation of the reproductive behaviour of *S. alburnoides* and of its implications on the evolutionary history of the complex. Indeed, the identification of individual genome constitutions was indispensable for the prosecution of the morphological, ethological and phylogeographical researches conducted;

ii) The analyzed *S. alburnoides* forms (AA, PA, PAA and PPA) showed significant morphological differences. It was also demonstrated that hybrids expressed both parental genomes (P and A), being clearly distinct from the nonhybrid AA form (representing the paternal ancestor of the complex) and from *S. pyrenaicus*, the maternal ancestor. Moreover, a morphological continuum was detected, with the hybrids occupying an intermediate position between AA and PP forms and with the PAA and PPA forms more closely related to AA and PP, respectively. This observation suggested the existence of a gene-dosage effect on the phenotype that was evident, for instance, in the increasing frequency of biseriated teeth (a diagnostic trait of the maternal ancestor) with the increasing proportion of P-genomes in the hybrids. The differentiation among the *S. alburnoides* forms in terms of morphology and colouration might be on the basis of the existence of some degree of spatial segregation. Ultimately, these phenotypic and ecological differences may influence the likelihood of the distinct mating patterns and, thus, affect the evolutionary dynamics of the complex;

iii) The ethological study of the reproduction of *S. alburnoides* maintained in semi-natural conditions showed the existence of non-assortative matings: females preferred to mate with non-hybrid males to the detriment of the also reproductively available hybrid males. This female mate choice was probably based on the vigorousness of the displays performed by these small males and has apparently crucial implications in the genetic dynamics of the complex. First, by achieving more fertilizations than hybrid males, non-hybrids ensure their persistence since the only pathway by which non-hybrids may be reconstituted is through crosses between PAA females and males of their own kind (AA) and, thus, once extinguished from a population they will never be reconstituted again. Secondly, crosses between AA males and PAA females are the ones that most contribute to increase the genetic variability of the complex since, besides the uncommon balanced tetraploids, these are likely the only type of males and females that perform meiosis;

iv) In an interspecific context, non-hybrid males were extremely active in courting and fertilizing the eggs of *S. torgalensis* females and assumed a sneaking behaviour in the presence of *S. pyrenaicus* courting pairs. This observation also seems to have important consequences to the genetic dynamics of the complex. If, by adopting a sneaking mating strategy, non-hybrid males were successful in fertilizing other *Squalius* eggs, it was expected that introgressions of other *Squalius* genes in the complex would

occur. Indeed, the role of non-hybrid males as promoters of introgression of new genes was corroborated by the fact that, when comparing the levels of introgression between populations that were sympatric with *Squalius* species distinct from *S. pyrenaicus* (so that the mtDNA of the sympatric species was distinct from the mtDNA carried by the complex), populations having non-hybrid males (Quarteira and Sado) showed much higher values than the ones observed in populations from where non-hybrid males were absent (Douro and Mondego). This suggested that, although hybrid *S. alburnoides* males may cross with other *Squalius*, non-hybrid males might be more successful in doing so, contributing, in the case of the Quarteira population, to a massive introgression of *S. aradensis* genes, almost replacing the typical *S. pyrenaicus*-like mtDNA of the complex. Thus, the extent of the introgressions in the complex seems to be positively correlated with the presence of non-hybrid males, which may be related to the adoption of a sneaking strategy by these males in the presence of other *Squalius* courting pairs. Contrastingly to previous studies that suggested an almost genetic uniformity at the mitochondrial level, it was demonstrated that the introgression of genes from non-*S. pyrenaicus* species was significant in all the studied populations;

v) The sequencing of the beta-actin gene allowed the discovery of a very rarely occurring non-hybrid AA female that highlighted the theoretical possibility of re-emergence of an extinct species (the paternal ancestor of *S. alburnoides*) from its descendent hybrids, although carrying the mtDNA of other species (*S. pyrenaicus*);

vi) The determination of the inherited nuclear genomes by the developed method based on the analysis of the double peaks generated by heterozygous indels allowed the confirmation of the production of haploid and diploid sperm by AA and PA males, respectively, and the first demonstration of the production of clonal sperm by a triploid PAA male. In contrast to many vertebrate hybrid lineages, the sperm of triploid hybrid males of *S. alburnoides* is viable and fully functional. This result, together with the available data concerning the gametogenic processes of the other forms of the complex, highlighted the important role of *S. alburnoides* females as regulators of the ploidy level. Indeed, the extremely common triploid females usually produce haploid eggs, which contribute to avoid the uncontrolled growth of the ploidy number beyond viable levels. Moreover, in contrast to hybrid males, by performing meiosis, triploid females are acting

as generators of genetic variability in the complex, a role that is shared with nonhybrid males and balanced tetraploids;

vii) The sequencing of the beta-actin gene also confirmed that the paternal ancestor of the complex have been a species close to *Anaecypris hispanica* (as previously proposed), likely extinct, belonging to the European *Alburnus* lineage;

viii) The phylogeographical study conducted showed that the origin of the *S. alburnoides* must have been posterior to the differentiation of the other *Squalius* species, which had a star-like radiation from a central and very widely distributed *S. pyrenaicus* clade. This radiation of peripheral derived clades (*S. aradensis*, *S. torgalensis*, *S. carolitertii*, *S. valentinus* and *S. malacitanus*) seems to have occurred from the Miocene to the Pliocene and was corroborated by data on the paleohydrography of Iberia;

ix) The hypothesis of multiple and independent origins of *S. alburnoides* was not confirmed and it was proposed that the complex had a single origin in the Upper Pleistocene (around 700.000 years ago). Under the most plausible scenario, considering the estimated divergence times of the populations and the present day distribution of other species, *S. alburnoides* was originated when the Tagus-Guadiana drainage (carrying the *S. pyrenaicus* maternal ancestor) captured an endorheic lake where the *Anaecypris*-like ancestor would have been isolated. Thus, the origin of the complex was likely a result of the paleohydrographical rearrangements of the Iberian Peninsula. The absence of *Anaecypris*-like mtDNA in the complex may be explained if interspecific crosses between *S. pyrenaicus* females and *Anaecypris*-like males were systematically more frequent than the reverse ones. This hypothesis was corroborated by the sneaking mating strategies that reconstituted non-hybrid males adopt in the presence of *S. pyrenaicus* courting pairs. This behaviour may, thus, be seen as an ancestral trait that was likely present in the paternal ancestor species and that re-emerged when AA males were reconstituted from the hybrids. Later, the paternal ancestor of the complex seems to have suffered considerable declines (motivated by ethological, climatic and/or ecological factors) that most probably resulted in its extinction, irreversibly eliminating the chance of incorporation of *Anaecypris*-like mtDNA in the complex;

x) The dispersal of *S. alburnoides* likely occurred after all the forms of the complex have been originated and the paternal ancestor has become extinct. The complex seems to have dispersed northwards to Mondego and Douro through independent routes from the Tagus; south-westwards to Sado also via Tagus; and southwards to Guadalquivir and Quarteira via the upper and lower sections of Guadiana, respectively;

xi) With the dispersal to river basins outside the distribution range of *S. pyrenaicus*, the complex became sympatric with other *Squalius* species and, as a consequence of interspecific crosses, their nuclear and mitochondrial genes were introgressed in the complex. In contrast to the almost unidirectional introgression of mitochondrial genes that occurred when the complex was emerging, present day introgressions of nuclear genes are bidirectional, with important consequences to the definition of the taxonomic boundaries of the sympatric *Squalius* species. Indeed, PPA females, as they produce P-eggs, may be responsible, for instance, for the introgression of *S. pyrenaicus* genes in *S. carolitertii* populations, contributing for the genetic and phenotypic homogenization of both species.

Despite the occurrence of atypical gametogenic modes and of populations with extremely biased sex ratios, the reproduction of *S. alburnoides* seems to always involve the union of eggs and sperm and is not strictly dependent upon other *Squalius* species. Thus, it allows questioning whether this complex should continue to be designated as "asexual". The presence of fertile fish of both sexes and recombination in male and female forms argues against such labelling.

Although the researches conducted have contributed to highlight significantly some more important aspects of the reproduction of the *S. alburnoides* complex and of the effect of the reproductive behavioural patterns on the morphology and evolutionary dynamics of the complex, they have also raised specific questions and hypothesis that should be addressed in future studies, namely:

- The biological species concept is somewhat challenged by the interesting role of *S. alburnoides* as a vehicle of introgression of genes belonging to a species (*S. pyrenaicus*) that is geographically isolated from the introgressed one (*S. carolitertii*). Future studies with other suitable nuclear DNA markers should focus on the extent of these introgressions in

order to evaluate the genetic integrity of the currently taxonomically recognized *Squalius* species;

- Also regarding the relationship between *S. alburnoides* and other species, it would be interesting to search, in the Guadiana river basin, for hybrids with *A. hispanica* and with the recently introduced *Alburnus alburnus*;

- The introgression of non-*S. pyrenaicus* genes in the complex may be expressed by differences in the phenotype of the *S. alburnoides* belonging to distinct populations. As this hypothesis was formulated based on the analysis of a small number of individuals, further morphological studies should be conducted mainly on PPA specimens, which are expected to better express the contribution of the P-genome to the phenotype. Additionally, the hypothetical existence of a geographical variation of the expression of the A-genome on the phenotype should also be confirmed with broader samples of *S. alburnoides* non-hybrids;

- It would be interesting to test if the morphological differentiation of the *S. alburnoides* forms is correlated with spatial segregation, namely, if the small non-hybrids are indeed more well adapted to semi-lentic conditions and higher water temperatures;

- The nuclear gene used in this study proved to be useful to detect ancient connections between populations that were undetectable with the faster evolving *cytb* gene. However the lack of a validated molecular clock as the one used for the mtDNA analysis precluded the estimation of divergence times, a situation that should be solved in the future;

- Phylogeographical studies on other Iberian Cyprinidae genera such as *Barbus* and ex-*Chondrostoma* should be conducted and their results integrated to test the validity of the colonization routes now hypothesized for *Squalius*;

- Exhaustive mtDNA sequencing of the captured *S. alburnoides* non-hybrids should be conducted to discard the hypothesis of the eventual occurrence of survivors of the paternal species of the complex;

- The sampling effort in the distribution area of *S. alburnoides* should be intensified aiming to detect the real relative frequencies of the various forms of the complex in all the populations. With the described method based on beta-actin gene sequences, this study

would not require a massive sacrifice of individuals (as would happen if using allozymes) nor would be extremely expensive and time consuming (as would happen if using microsatellites);

- Despite what was observed in the presence of all the possible ploidy forms of *S. alburnoides* in a situation of high density of fishes, it would be interesting to test whether or not diploid and triploid females adopt similar behaviours and identical mate choices. Additionally, future experiments should also be conducted to characterize the behavioural pattern of females towards diploid, triploid and tetraploid hybrid males, in the absence of nonhybrid males;

- Using DNA-fingerprinting techniques, the fertilization rates of sneaking non-hybrid males in interspecific contexts should be evaluated;

- The gametogenic processes of some *S. alburnoides* forms that remain unknown should be clarified, as this information is crucial to the understanding of the genetic dynamics of the complex. Moreover, the oogenesis of the PAA females requires further attention since, as they are the most common type of females in all the populations, it is important to know the relative proportions of the three types of eggs that may be produced. To do this, surveys should be conducted in natural populations or at least simulating the natural conditions with the best fidelity, since it is known that meiosis may be sensitive to environmental factors and, thus, extrapolations from the results obtained in laboratorial conditions may not reflect the reality;

- Further investigations, using other molecular markers, should be conducted to corroborate the estimated date for the origin of the complex (more recent than proposed by previous authors) and the proposed single origin hypothesis.

Finally, although the results presented in this dissertation contributed with innovative data to the current knowledge on the reproduction and evolutionary history of *S. alburnoides* and confirmed some aspects that have been put forward by previous investigations, as this Iberian complex is an interesting model of evolution-in-action, the state of the art on it is a dynamic process and, therefore, further researches might always be conducted in several different domains of investigation.