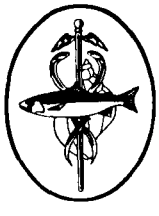
GENETIC ANALYSIS OF EUROPEAN SEABASS (*DICENTRARCHUS LABRAX* L.) FROM PORTUGUESE WATERS USING ALLOZYME AND MICROSATELLITE LOCI

A thesis presented for the degree of
Doctor of Philosophy to the University of Stirling

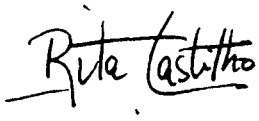
by
Rita Castilho
Institute of Aquaculture
University of Stirling



1998

Declaration

I declare that this thesis has been composed by myself. The work described here is of my own research and has neither been submitted nor accepted for any other degree.

Candidate: 

Supervisor:

Supervisor:

Date:

Acknowledgements

This work was only possible due to the full financial support given by the Unidade de Ciências e Tecnologias dos Recursos Aquáticos, University of Algarve, Portugal and the Junta Nacional Investigação Científica e Tecnológica, through two grants from Ciência (BD/2236/92) and Praxis XXI (BD/5820/95) programmes.

At the Institute of Aquaculture, University of Stirling, there are a lot of people to whom I am in debt. My main supervisor Dr. Brendan McAndrew, at the Institute of Aquaculture, was at the origin of this work. I would like to express my gratitude for having accepted me as his student, and for never loosing his temper when putting up with me every other day, complaining, suggesting, forcing some ideas. And contrary to my expectations he took very little time to correct some drafts, specially this thesis!

Dr. David Penman, also at the Institute was a big help in the early days of this work, since DNA methods were completely unknown to me. And he was always a source of continuous healthy discussions on just about everything, but specially the hottest topic: laboratory organisation and meetings.

A word of thanks to the examiners, Dr. John Taggart and Dr. Tom Cross, who revised the thesis, providing numerous suggestions that contributed to a real improvement of the contents.

I am in debt to the people, professional fishermen and others, who provided the biological material without which this work would have never even started, namely: M. Sobral (CRIPAveiro), J.A. Ferreira (Foz), M. Salgado and G. Salgado (Óbidos), F. Caetano and A.A. Silva (Milfontes) e B. Fanaia (Faro). I would also like to thank the people involved with me in sampling: M. Brito José, J. Reis, M. Reis, P. Lino and F. Soares.

I am in debt to Drs Garcia de Léon and F. Bonhomme, from the Laboratoire Génome et Populations, Université de Montpellier II, Montpellier, France, for making 3 pairs of primers available.

My visit to the Marine Gene Probe Laboratory, Dalhousie University, Nova Scotia, Canada, was a major turning point for me and also for introducing the molecular genetics at its state of the art stage at the Institute. I was fortunate enough to meet there

Doug Cook, the laboratory manager, technician and chief magician. He looked over my shoulder at all times to make sure this beginner was doing things right! I owe him a lot in many ways. Also to Mike O'Connell who's friendship and support was very encouraging.

I am grateful to Michel Raymond for providing software and also for being available at any time for friendly advise and discussions over the email.

I also would like to express my thanks to António Nogueira, who during his stays in Stirling was a good companion, who initiated me in the Internet, and who provided me a simple and efficient computer program for handling the references database.

Life in Stirling can be hard for southern Europeans, specially in winter.... Most of the hardships were overcome by a very friendly and supportive atmosphere that surrounded me. I cannot pass by without mentioning all that contributed to the weekend get togethers and to those lunches "a la una": Toya, Luis, Natalio, Bela, Meli, David, Cova, Jose, and of course Carlos. Our friendship lasts!

I would like to express my thanks to Maria Teresa Dinis, my mentor for a long time, she initiated me in research, and since them we have been good friends, building a special bond between us.

To my parents, who supported me since early times to pursue what I wanted, and who encouraged me to go to Stirling. They had always my best interest at heart. Iam sorry they never got to know the Scotish weather..... After all my complains, we toured Scotland during the hottest summer in 200 years!

To Filomena, who was always a constant presence. She kept my spirits high when needed, by email, telephone, fax and yes, even snail mail, talking about all the tricks laboratory work plays on you. It surely made a difference! I am perfectly aware that she endured pretty difficult times while I was writing the thesis at home.

My pets, 8 dogs and 7 cats provided good company in all those walks I needed to take to clear my mind at times.

ABSTRACT

Genetic differentiation among juvenile samples of the European seabass (*Dicentrarchus labrax*) from the coast of Portugal is reported by means of two types of genetic markers: allozymes and microsatellites. Repeat samples were taken from 5 different nursery grounds (Aveiro, Foz, Óbidos, Milfontes and Faro) along the coast of Portugal between November 1992 and February 1994. Starch-gel electrophoresis was used to assess the level and distribution of genetic variability of 38 loci. Six of these were found to be polymorphic at the 99% level and were used in population surveys: *AAT-3**, *ADA**, *GPI-1**, *GPI-2**, *G3PDH-2**, *SOD**. Statistical analysis revealed low but statistically significant multilocus F_{ST} (0.0108, $p<0.001$) values suggesting that population structuring exists along the Portuguese coast line. The results indicate that there is some restriction in gene flow between the more southerly population at Faro and all other sites to the North. Five microsatellite loci were screened in over 300 individuals. High levels of polymorphism (number of alleles observed per locus ranged from 20 to 41) and observed heterozygosities, ranging from 0.45 to 0.89 (mean over all loci = 0.71) were detected. Two loci displayed heterozygosity deficits (*Dla6* and *Labrax-9*) and were not used in population comparisons. Statistical analysis revealed low but statistically significant multilocus F_{ST} (0.0025, $p<0.001$) at the other three loci (*Dla11, Labrax-3* and *Labrax-8).* No clear geographic patterns emerged from these results. Overall allozymes performed well when compared to microsatellites, in detecting microgeographic genetic structure in this species. Microsatellites revealed high levels of polymorphism that should prove useful as markers in the management of wild and farmed seabass stocks in the future. The level of differentiation, low values of F_{ST}, detected among the sites is low but is typical of marine species which have a much greater chance of mixing.

ABBREVIATIONS USED IN THE PRESENT WORK

A	2'-deoxyadenosine
AP	Ammonium persulfate
ATP	Adenosine triphosphate
Bp	base pairs
C	2'-deoxycytidine
d H_2O	distilled water
dd H_2O	double distilled water
DNA	Deoxyribonucleic acid
EDTA	Ethylenediamine tetraacetic acid
EtBr	Ethydium bromide
EtOH	Ethanol
E.U.	European Union
F.A.O.	Food and Agriculture Organization
G	2'-deoxyguanosine
G.E.P.P.	Gabinete de Estudos e Planeamento das Pescas (Fisheries Study and Planning Group)
HCl	Hydrochloric acid
IAA	Isoamyl alcohol
ICES	International Council for the Exploration of the Sea
IPTG	β-D-isopropyl-thiogalactopyranoside
kbp	kilo base pairs
MGLP	Marine Gene Probe Laboratory
MtDNA	mitochodrial DNA
MTT	3,(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide
NBT	Nitroblue tetrazolium
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
rpm	rotations per minute
SDS	Sodium dodecyl sulphate
SSC	standard saline citrate

T	2'-deoxythimidine
T_a	annealing temperature
Taq	*Thermus aquaticus*
TEMED	N,N,N',N'-tetramethylethylenediamine
TEN	Tris-EDTA-NaCl
T_m	melting temperature
TRIS	Tris(hydroxymethyl) aminomethane
VNTR	Variable number of tandem repeats
WGAGFM	Working Group on the Application of Genetics in Fisheries and Mariculture
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

TABLE OF CONTENTS

PART 1 - GENERAL INTRODUCTION

1. INTRODUCTION

The analysis of genetic diversity and relatedness between or within different species, populations and individuals constitutes a central component for many disciplines in biological science. The study of the genetic variation in populations and its change, the following of allele frequencies in populations, real or theoretical, through time and space is known as population genetics. The main goals of population genetics are to describe how the frequency of an allele changes over time; to analyse the factors that lead to changes in gene (allele) frequencies and to determine how changes in gene (allele) frequencies affect genetic differentiation, evolution and speciation. The parental structure, the exchanges among demes, the identification of reproduction units and its interdependence, the spatial structure of the species and the capability of migration along short, medium or long distances are quite important questions in basic population biology science. Most of these aspects, however, are also quite important in applied aspects such as resource management, conservation biology, biodiversity and genetic improvement, just to cite a few. The question of how much a population is different from another has several good reasons: academic (one may want to know more about evolution of a certain species or a group of species); environmental (to assert if there a loss of genetic variation that endangers natural stocks); managerial (are stocks discernible for which different management practices should be put in place); aquacultural (are there genetically different farmed stocks or is it possible to derive domesticated strains from different natural stocks).

The knowledge of population structure is an important element in effective fisheries management. Accurate genetic stock identification provides managers with a picture of the biological relationships between populations, which will in turn allow for a more informed management policy for the resource. However, the absence of obvious geographic barriers as well as active and passive dispersal of organisms contributes generally to lower levels of genetic divergence in marine fishes compared to anadromous and freshwater. That is why marine populations discreteness has constituted an ambitious task in fisheries.

Effective fishery management relies upon identification of the populations of a species and this idea has been brought together with the concept of stock (May and Krueger, 1990). Spawning aggregations constitute fish stocks or populations because they are a group of conspecific fish that interact more frequently with one another than with individuals belonging to other conspecific populations (Bermingham, 1990). In genetic terms, a population would be "the largest aggregate of individuals whose genotypes can be considered as random draws from a single gene pool" (May and Krueger, 1990). This term is often used synonymously with stock by fishery managers to designate a particular fish population. Nevertheless, the possibility of genetic flow prevents, in most cases, this sort of isolation to be translated into the existence of stock-specific tags and most of the times only differences in allele frequencies are detectable (Bermingham, 1990).

Most marine biologists and fish managers would agree that a principal goal of fishery management is to obtain detailed information on fish stocks and to understand how these interact with each other and their environment (Ferris and Berg, 1987), and to evaluate the impact that particular fishery practices might have on the genetic structure of a stock (Larkin, 1981). Management of seabass requires detailed information on the number of spawning aggregations in Europe, the extent of larval dispersal from each and the potential threats to recruitment that may exist within each stock. It is clear that the combined effects of life history, dispersal capability and environmental limitations, will not only affect the geographic but also the genetic structure of the species and that will be reflected in the assays, be they isozymic, mitochondrial DNA or nuclear DNA based. The increasing commercial and recreational pressures on seabass requires that as much information on stock integrity from all sources should be available to enable sustainable management strategies to be developed.

Commercial ventures such as aquaculture enterprises cannot rely entirely on wild broodstock. The advantages of a well adapted and genetically well characterised broodstock is potentially invaluable for producers. This is one of the arguments for further genetic studies on seabass and there is an urgent need particularly for the

initiation of broodstock selection programmes and procedures for avoiding inbreeding (Carrillo *et al.*, 1995).

Throughout its range the seabass is an important commercial species, being exploited by commercial fishermen, aquaculture enterprises and sports anglers. The commercial importance of this species, contributed to the increase in information on seabass biology (see Pickett and Pawson, 1994) and population genetics (García de León, 1995; Allegrucci *et al.*, 1997; Naciri *et al.*, submitted) in recent years . The obvious economical importance of the seabass wild fishery and the potential for aquaculture, prompted several laboratories to study the seabass genetic structure both in the Mediterranean and the Atlantic. Different genetic markers, including allozymes, mtDNA and microsatellites have been used. Most studies report genetic differentiation of the Mediterranean *vs.* the Atlantic samples, and also, although to a lesser degree, some genetic differentiation within Mediterranean and Atlantic samples.

The coast of Portugal constitutes essentially the centre of the species range. It is important to obtain samples of seabass along the Portuguese coast, to determine the stock structure in that area and try to integrate the new data into an overall picture. In an ideal situation samples from adjacent localities in Spain should be obtained to allow direct comparison. However, that was not possible: wild seabass samples were difficult to obtain in many of the sampling locations in Portugal, for reasons explained later, and the same difficulties hindered the sampling in Spain. So comparisons between the present work and other previously published works will be carried out with due precaution.

In this study the aim was to make use of molecular data and integrate this with available information from the life cycle of seabass to try to contribute to the understanding of some of the actual problems of seabass population structure and biology. Among these, the most interesting are concerned with the following questions: Is the entire seabass population of the Portuguese coast a single panmictic population?

How many demes can be defined, how can they be defined? Is there any molecular data evidence to support or dismiss the 'homing' behaviour of seabass, e.g. do migratory marine seabass return to natal sites for spawning?

As it will be referred to in the population genetics section, the resolution power of each technique depends upon several aspects including the species characteristics, such as life cycle, egg and larval stages, migration patterns, spawning habits, etc., the geographic range of study, wide or confined, the sample sizes and a lot other aspects that are more related to the markers themselves. The existence of molecular markers is due to the divergence of the macromolecules that constitute the genome. This divergence is caused by mutations, which originate genetic variability. It is upon this variability that the three main evolutionary forces will act: genetic drift, selection and migration. Thus, at a given point in time, there can be several genetic variants within the same population that are the basis of a molecular polymorphism which can be analysed in terms of genetic structure.

Population genetic surveys of different geographic origins will generate much useful information of immediate and long term importance. Protein electrophoresis, by providing rapidly collected and purely genetic data sets, has opened new ways of examining fishery management problems. Researchers world-wide (Chapman and Brown, 1990; Carvalho and Pitcher, 1995; Ferguson, 1995; Park and Moran, 1995; Utter, 1995) recommend allozyme studies as a first step in understanding population structure. It is clear that protein electrophoresis is a good value for money technique that will continue to be very useful in fishery management. Nevertheless, it may not detect most of the genetic variation that may be of value in understanding stock dynamics, specially when considering highly mobile temperate species such as the bass. In fact, the interplay between dispersal capability and environmental limitation to gene flow, be these in the past or the present, has an important influence on the genetic structure reflected in genetic assays (Avise, 1994). In this regard, the use of microsatellite polymorphisms is considered as an appropriate molecular genetic tool to elucidate intraspecific structure (Park and Moran, 1995).

One important objective of the present work is the comparison of the information and performance generated by the two different molecular approaches, allozymes and microsatellites, which is an aspect not abundantly covered in the literature (Carvalho and Hauser, 1995).

Population genetics can approach the study of a given taxa by either characterizing the molecular basis of variation in particular systems or by applying the genetic data to phylogeny, phylogeography and organismal evolution. This work is a contribution to both aspects, as it has a microgeographic approach to the population genetics of the seabass, but can at the same time contribute to the understanding of the seabass biological history.

The work presented here is divided into several parts:

Part 1 - provides introductory material and background: a general introduction focusing on the present knowledge of seabass biology, exploitation and population genetics;

Part 2 - includes common materials and methods used in this study for both allozymes and microsatellites approaches, such as sampling procedures, characterisation of the sampling locations, and data analysis;

Part 3- presents allozyme methodology, results and discussion;

Part 4- is constituted by microsatellite DNA methodology, results and discussion and

Part 5 - provides a comparative analysis of allozyme and microsatellite performance, general discussion and conclusion, references and annexes which include individual genotypes for allozyme and microsatellite loci, as well as detailed recipes and protocols.

2. THE ECOLOGY OF THE SEABASS

The seabass (Figure 1) belongs to the order Perciformes, family Serranidae (sometimes referred to as Moronidae) and the genus *Dicentrarchus*. The seabass has been known by numerous synonymous designations, namely, *Perca labrax* (Linnaeus, 1766), *Labrax lupus* (Valenciennes, 1828) and *Morone labrax* (Boulenger,1890) (Pickett and Pawson, 1994). In 1966, Whitehead and Wheeler (1966) proposed the name *Dicentrarchus labrax* (Linnaeus, 1758). Most of the North American species which are close relatives of *D. labrax* live in fresh or brackish water (e.g. *Morone saxatilis*).

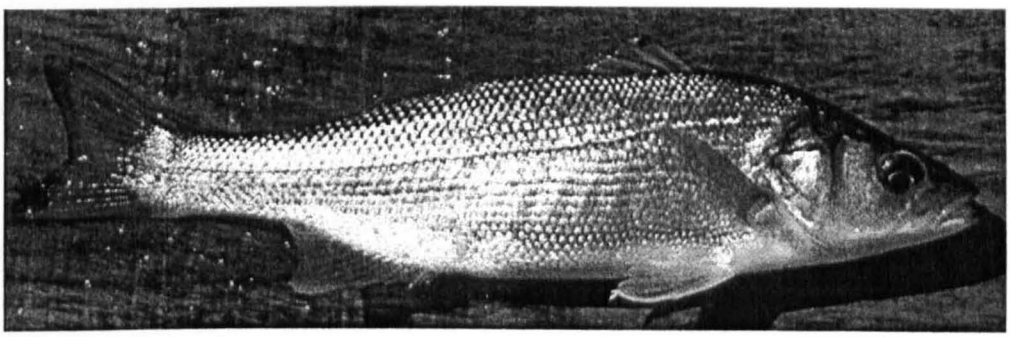


Figure 1. The seabass.
(© Ifremer/O. Barbaroux)

Seabass is a euryhaline and also eurythermic species. This means it tolerates a wide range of salinity, being possible to find it in low salinity water as well as in full strength sea water (Chervinski, 1979), and in water temperatures ranging from 5-28°C (Barnabé, 1980). This, according to some authors (Allegrucci *et al.*, 1994) opens up possibilities for freshwater cultivation.

The seabass is a highly mobile species, known to undergo migrations of over 500 km per year (Pawson and Pickett, 1990), with a European distribution ranging as far south as the Atlantic Ocean off Morocco (30°N) through the Irish Sea, North Sea and Baltic Sea, to southern Norway (60°N) and eastwards throughout the Mediterranean to the Black Sea (Whitehead *et al.*, 1986). The bass has been increasingly found in the

North of Scotland, due to outflows from power stations, that locally maintain water temperatures more favourable to the species throughout the year (*per. obs.*).

This species is common along rocky sea coasts and sandy beaches with waves and is also commonly found in estuaries and lagoons, often moving many km from the sea into low salinity and freshwater (Bertignac, 1988). *D. labrax* can be considered a seasonal-sedentary species because it can remain in the estuaries of rivers for quite long periods, although it must return to the sea before the spawning season (Demestre *et al.*, 1989).

2.1. LIFE CYCLE

Information on the life cycle of the seabass has been produced and compiled by a number of authors (see Barnabé, 1980; San Feliu, 1987; Bertignac, 1988).

The seabass is a multiple spawner, meaning that the ovaries contain more than one group of synchronous developing oocytes (Mayer *et al.*, 1990). The eggs are pelagic and small (1.02 - 1.39 mm) (Barnabé, 1980), and each female can produce from 290 x 10^3 to 955 x10^3 eggs/Kg of fish (Kennedy and Fitzmaurice, 1972; Bou Ain, 1977). The multiple spawning habit might explain the relative scarcity of bass eggs in plankton samples, when compared with large numbers found from other commercial fish species (Thompson and Harrop, 1987).

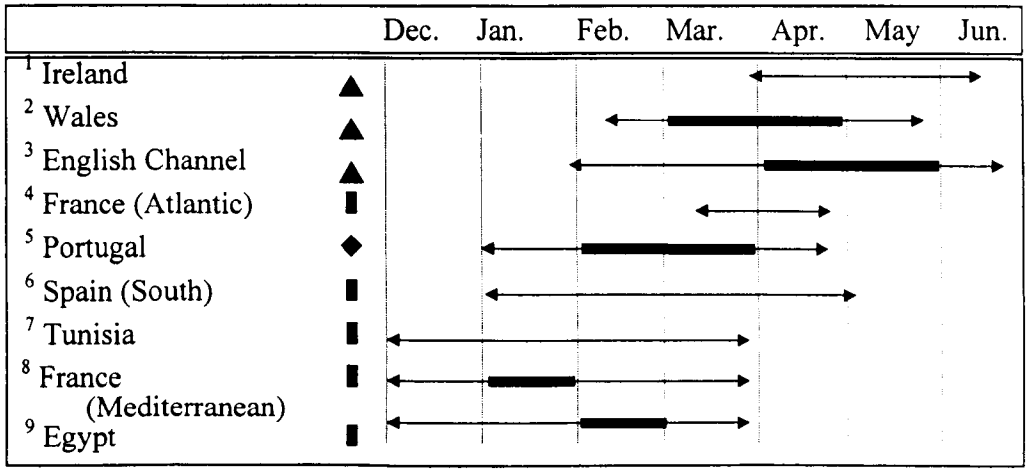
The spawning season seems to be latitude-related, depending on the water temperature (Kelley, 1988a; Jennings and Pawson, 1991), the maximum spawning activity occurs when the water temperature is around 8.5°C to 11°C, according to data for the English Channel (Thompson and Harrop, 1987). Over the species extended range spawning occurs from December to June (Table 1) (Kennedy and Fitzmaurice, 1968; Stequert, 1972; Bou Ain, 1977; Arias, 1980; Barnabé, 1987; Thompson and Harrop, 1987; Wassef and El Emary, 1989; Jorge, 1991; Jennings and Pawson, 1992).

Some authors (Kennedy and Fitzmaurice, 1968), think that it is likely that some spawning takes place whenever temporary assemblies of bass occur, in a variety of situations, including inshore waters (rivers and estuaries), rocky reefs (Barnabé, 1976;

Campillo *et al.*, 1989) at the mouth of rivers (Dieuzeide *et al.*, 1954), or coastal waters (Sabriye *et al.*, 1988) where tides create suitable conditions for spawning. It was observed in the English Channel that the highest density of bass eggs were predominantly offshore, thus indicating that the greatest spawning activity occurs away from the coast. Nevertheless, some bass eggs were also found inshore (Dando and Demir, 1985; Jennings and Pawson, 1992).

Table 1. Spawning season of the seabass.

[1]Kennedy and Fitzmaurice (1968); [2]Jennings and Pawson (1992); [3]Thompson and Harrop (1987); [4]Stequert (1972); [5]Jorge (1991); [6]Arias (1980); [7]Bou Ain (1977); [8]Barnabé (1987); [9]Wassef and El Emary (1989). ▬ Spawning peak. Data represented evaluated as:▲ number of eggs/m^2 (collection by drifting nets at the surface); ▮ gonad development; ◆ otholits' daily increments.



In experimental conditions seabass eggs take between 4 and 9 days to hatch (at 15°C and 9°C, respectively), and larvae are about 4.0-4.5 mm in length (Barnabé, 1976). Post-larvae up to 9.5 mm of standard length have been caught at sea. In estuaries the minimum size of larvae was 10 mm (Dando and Demir, 1985). Post-larvae enter the lagoons and estuaries in the UK in May-June (at 10-15 mm length) (Dando and Demir, 1985; Kelley, 1988b). In the North of Portugal (Aveiro) the arrival of post-larvae is observed between June-July, 3-4 months after spawning, (Rebelo, 1992) and by this time they are 30 mm in length (Gordo, 1989) although in Figueira da Foz, slightly to the south, they enter between March and July (Jorge, 1991). In the French Mediterranean area, from February to March there is mainly passive movements of larvae (7-8 mm), with the currents. However by May-June there is active movement of

fry (25-40 mm) into the lagoons (Campillo *et al.*, 1989). The mechanism by which larvae enter nursery areas is not known (Pawson and Pickett, 1987), although there seems to be a need to reach fresh water (Kelley, 1988b). The search for higher temperatures and lower salinity, could act as stimulus for the active migration of the young (Pickett and Pawson, 1994). Also, bass larvae have been captured at 15m below the water surface, where residual currents were faster than those at the sea surface suggesting the larvae are actively selecting beneficial currents that help their migration (Jennings and Pawson, 1992). It appears therefore, that there is a combination of passive movement with the currents and active searching for the nursery and that this might vary from site to site.

Larvae do not adopt a typical young bass lifestyle in the brackish water of estuaries until 2 or 3 months after hatching in UK waters (Pawson and Pickett, 1987). The abundance of seabass larvae has been associated in some instances with narrow ranges of physical and chemical parameters, 12°C temperature, very low salinity, high oxygen, almost neutral pH, and low transparency waters (Rebelo, 1992).

The age group 0+ is found predominantly in the estuarine environment during autumn and winter (Pawson and Pickett, 1987; Gordo, 1989; Rebelo, 1992). Around 50% of the total seabass sampling catch in the Aveiro lagoon, Portugal, between 1985 and 1987 belonged to age group 0+ (Gordo, 1989).

In the UK. the extent to which the young bass remain within an estuary as they grow, will vary with the type of estuary and the weather experienced in winter (Kelley, 1988b). This author presents a pattern of estuary-dependence with three main types, based largely on the estuary size:

- large (>4 Km^2 of water-area at low tide) at which bass stay largely within the estuary for their first four or five years;
- medium (1-4 Km^2), bass spend their first two years in the estuary, and during the warmer months until the age of 4+ to 5+;
- small (<1 Km^2), bass presence is limited to their first and second summer.

Of course this does not exclude the presence of young bass visiting from other areas and adult fish moving through to their wintering areas during autumn (Kelley, 1979).

In the Mediterranean, adaptation to lagoon/freshwater appears to be accompanied by high mortality and selection against a number of allozyme genotypes (Allegrucci *et al.*, 1997). It is unclear whether the selection is related to temperature, salinity or a combination of these factors.

Juveniles (<32 cm) are found to stay loyal to their native estuary as shown by tagging (Pickett and Pawson, 1994). Adolescent bass between 32 and 42 cm were recaptured in or near their original nursing areas, although some individuals have been recaptured in adult wintering areas (Pickett and Pawson, 1994). Beyond this size the adolescents enter a period of wide-ranging wandering after which they settle on an area for their future adult summer life, moving then into the adult pattern (Kelley, 1991). In fact, in UK waters, a large proportion of the young adult bass population undergo extensive migrations each year, some fish being recaptured in winter 300 to 500 km to the south and/or west of a summer tagging site (Pawson and Pickett, 1990).

The life cycle of seabass can be summarised in the following way (Figure 2):

- Spawning takes place in offshore and coastal waters;
- Post-larvae arrival in estuaries or lagoons a month after spawning;
- Larvae concentrate at the head of the 'salt-wedge';
- Larvae spreading back through the estuaries or lagoons (this appears to be associated with adaptation in Mediterranean populations);
- First summer: shallow creeks, marsh pools and tributary streams;
- Juveniles move into deeper areas in October;
- Adolescent movements occur: visits to other estuaries and future summer life area is adopted;
- Up to 4 or 5-years the individuals stay in or near the adopted estuary or lagoon;
- Maturity reached between 2-7 years (males) and 2-8 years (females) depending on the locations (see table 2);
- Brief visits of adults to estuaries or lagoons (Kelley, 1988b);
- Adult bass migrations, extensive in the UK, restricted in Mediterranean.

Much of this is based on British Isles studies almost at the extreme northern range of seabass distribution. Mediterranean populations appear to be somewhat different, with greater juvenile mobility and more sedentary adults (Chauvet *et al.*, 1992). Another important point is that there is no evidence in UK waters for freshwater adaptation, as seen in the Mediterranean (Allegrucci *et al.*, 1997), so the existence of "racial" difference is a possibility.

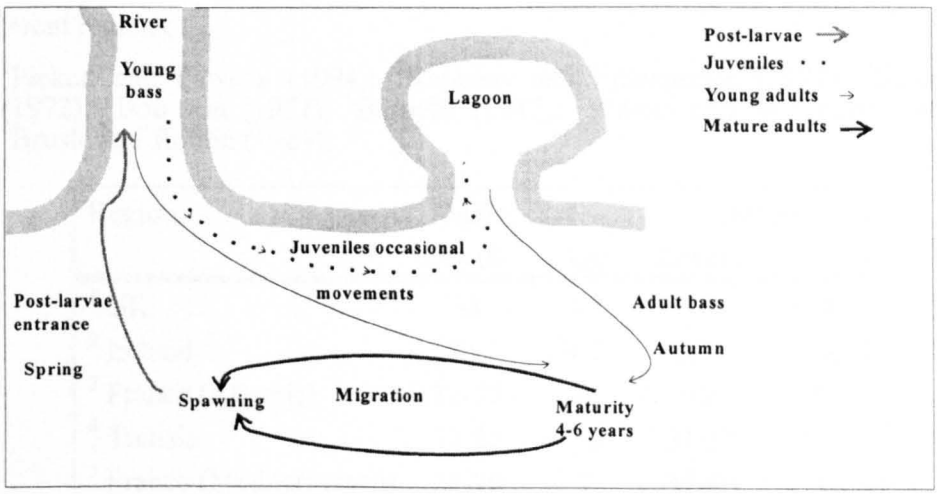


Figure 2. Life cycle of seabass.
The arrows represent the fish movements (based on Andrade, 1983; Pawson *et al.*, 1987; Kelley, 1988b; Rebelo, 1992).

The tendency to remain in a particular area throughout the year, appears to diminish as bass grow, and the accompanying seasonal movements are more extensive in the larger fish (Pawson *et al.*, 1987). Although there is no definite evidence, there are suspicions about adults maintaining localised fixed spawning ground (Dinis, *pers. comm.*). It could be the same area as where they spent their early life (Kelley, 1991) or a different one chosen after maturity is reached.

2.2. GROWTH

In the UK bass reach maturity at 4-6 years of age (35-40 cm) and may continue to spawn for up to 20 years thereafter (Pawson and Pickett, 1987). In more southerly locations, like France, Portugal or Egypt, some significant differences are observed in the length and age at first maturity (Table 2). There is a tendency for faster growth and

younger first maturity ages in the south of the range. Sea temperature seems to play an important role in differences between age of maturity (Kelley, 1988a). In temperature controlled fish farms, seabass mature sooner, males 19 months old and nearly 20 cm in length and females around 30 months old and between 27 and 31cm of length (Bruslé and Roblin, 1984).

Table 2. Length (cm) and age (years) of the first sexual maturity of wild seabass in different regions.

[1]Pickett and Pawson (1994); [2]Kennedy and Fitzmaurice (1972); [3]Stequert (1972); [4]Bou Ain (1977); [5]Barnabé (1987); [6]Wassef and El Emary (1989); [7]Brusle and Roblin (1984).

Region	Males		Females	
	Length	Age	Length	Age
[1] UK.	34	4-7	38	5-8
[2] Ireland	31.5	4-7	35.5	5-8
[3] France (Atlantic)	32-37	4	42	6
[4] Tunisia	23-25	2-3	31-33	4-5
[5] France (Mediterranean)	28-30	2	37-40	3
[6] Egypt	20	2	29	4
[7] Fish farm conditions (South)	19	2	28-32	2-3

Male and female bass appear to grow at the same rate until they reach maturity, which happens, on average, 1 to 1.5 years earlier in males (Bruslé and Roblin, 1984; Pawson and Pickett, 1987). Subsequently, females tend to be larger at a given age and for 4 or 5 years after maturing are usually heavier than males of the same length.

Variations in growth rate are observed in bass from different regions: in the Mediterranean the growth rate is higher than along the Biscay coast and is lowest in UK waters (Wassef and El Emary, 1989). The cause of the variations in growth is the different sea temperatures (Alliot *et al.*, 1983), but food availability (Wassef and El Emary, 1989) may also play an important role.

Parameters of von Bertalanffy's growth equation obtained for bass vary from region to region (Table 3) and reflect what was found for growth rates: it seems that growth is potentially higher in the southern regions of its distribution.

Table 3. von Bertalanffy's growth equations for bass from different localities.
[1] Barnabé (1987); [2] Wassef and El Emary (1989).

Region	Sex	L_∞	k	t_0
[1] Ireland	M	62.0	0.18	0.33
	F	73.1	0.15	0.14
[1] UK	M	87.5♦	0.09	-0.32
	F	70.2	0.14	-0.20
[1] France (Atlantic)	M	62.2	0.20	0.11
	F	60.5	0.23	0.18
[1] France(Mediterranean)	M	63.0	0.28	-0.18
	F	91.3	0.17	-0.26
[1] Tunisia	M	54.9	0.18	-0.98
	F	105.8	0.07	-1.49
[2] Egypt	M	78.1	0.08	-1.77
	F	87.8	0.06	-1.80

2.3. FEEDING HABITS

The seabass is a lively opportunistic predator taking crustaceans, fish and marine worms at any depth in the water column (Arias, 1980; Pawson and Pickett, 1987; Costa, 1988). The most common prey species for *D. labrax* are decapod crustaceans (benthic organisms), particularly *Carcinus maenas* (Kelley, 1987) and *Crangon crangon* (Barnabé, 1976; Kelley, 1987) and also isopod crustaceans and amphipods. Fish prey include benthic fish like sand-eels (Ammodytidae) and *Gobius minutus*, and pelagic species such as *Sardina pilchardus* and *Engraulis encrasicolus* are also an important component in seabass diet (Barnabé, 1976). Feeding preferences change as a function of fish size and although bigger fish still consume small prey such as shrimps, no amphipods are found in the stomach contents of fish above 40 cm, which may be due to the larger spaces between the branchial spines (Barnabé, 1976). In the Portuguese populations sampled the most common prey species were: the decapod, *Palaemon serratus*, among isopods, *Sphaeroma hokeri* (Cristo, *per. comm.*), *Sphaeromona*

♦ Note: In UK, males are bigger than females, that difference is due to the inclusion of a few very old individuals, aged between 13 and 21 years old (Pickett and Pawson, 1994, page 116).

rugicauda and *Cyallura* cf. *carinata* (Machado, *per. comm.*) and among the fish, *Engraulis encrasicolus* and several Mugilidae species.

Seabass are known to favour particular feeding locations, such as rocky reefs and sand banks, where prey species are abundant, and may accumulate in areas with high densities of food, specially around shoals of small fish. There seems to be little seasonal variation in the prey preferences of seabass, it is mainly the availability that will direct the diet, although feeding intensity is higher in spring and in summer and decreases during October-December (Arias, 1980; Kelley, 1987).

2.4. TAGGING STUDIES

Information on natural populations has come from various tagging studies mostly in UK and Irish waters often over many years (Kennedy and Fitzmaurice, 1972 - 895 fish tagged, 2.8% recaptured; Holden and Williams, 1974 - 954 fish tagged, 6.2% recaptured; Kelley, 1979 - 912 fish tagged, 9.4% recaptured; Pawson *et al.*, 1987 - 5933 fish tagged, 9.7% recaptured). These studies show that juvenile bass movements are usually within 80 Km of the release areas (Pawson *et al.*, 1987), and that this pattern starts to change quite dramatically as first maturity is attained. In fact, adult bass show extensive movements between summer feeding areas and winter pre-spawning areas and back in the spring time to the summer areas (Pickett and Pawson, 1994). The range of these displacements can be less than 100 Km, but generally average 400-500 Km, although it is possible that some bass may spend the winter time 800 Km away from their summer areas (Pickett and Pawson, 1994). In general terms it can be said that adult bass have a definite affinity for summer areas, and hence can have characteristic distribution ranges. But the real question here is: do adult bass show the same affinity for spawning areas that they show for summer feeding areas? Another tagging study was carried out on seabass during one single year (921 bass released, 2.2% recaptured) in the Golf of Lyon (Chauvet *et al.*, 1992). This work suggests that adult populations of seabass are sedentary or localised, but that juveniles are nomadic in this part of the Mediterranean. This difference could be related to the fact that environments in both studied areas, UK and Mediterranean, are very different.

In conclusion, the seabass has potential for wide population mixture at several life stages, i.e. via egg and larval dispersion as well as adult migration habits. Adult tagging studies did not show any integrity of stocks, only a preference for adult individuals to spend the summer in the same areas, year after year, but they also showed no evidence for mixing.

This type of study, capture and recapture, is very time consuming and demanding in terms of resources. A lot of conditions limit its application, namely: the means to tag (manpower and logistics) and the involvement of fishermen to report when and where a tagged fish is captured. It has to be said that most of the studies have also poor returns in terms of recaptures, when compared to the tagging effort. Also, that this kind of population biology study is regarded in this day and age as old fashioned. However, it would certainly be interesting to gain precise information on the location of spawning grounds and their constitution, and capture-recapture studies could indeed give some interesting answers. The studies up to now have failed to provide that information, because no adults were recaptured during winter time, due to the harsh weather conditions, that prevent fishermen going out to sea. Maybe a forthcoming solution is the combination of the principles of this technique without involvement of the recapture itself. That could be achieved by the use of advanced electronics technology, i.e., the insertion of a micro-transmitter into individual fish, that would allow the precise migration of a set of individuals to be measured remotely. They could be tagged before they leave the nurseries and thus it would be possible to trace their migration with a receiver, especially during the following winter when they gather for spawning. Also important would be the collection of complementary informative data on the biology of this species, as there are still some doubts on the preferred location of the spawning sites. As was said before, most authors found the greatest egg densities offshore, thus inferring that the spawning grounds must be in those areas. However, some eggs are also found inshore, and it may well be that these are a number of preferred spawning grounds.

It would be important to identify the number of locations of spawning sites and their stability from year to year. This type of study would have important implications for the management and conservation of seabass. Today, however, this type of information is unlikely to come from large boat based studies, but from the

identification of the levels of genetic differentiation in sampling regimes aimed at collections of individuals from selected sites over a number of years and indirectly inferring the possible genetic structures that might be present.

3. FISHERIES AND AQUACULTURE

3.1. FISHERIES

Southern European countries, particularly Greece, France and Italy, lead seabass landings, although, compared with other species, these are relatively small, near 24 x 10^6 Kg✣ in 1995 (Table 4).

Bass can be caught from the shore by angling with rod and line or beach seines and set nets, although the use of small boats is quite common, equipped with a variety of methods including gill nets, trolled lures and long-lines to capture seabass (Pawson and Pickett, 1987).

In certain areas, the seabass has a big impact on the local economy, through small-scale inshore fisheries, recreational anglers and fish farms. In some areas it is the most important fish species, such as in the Gulf of Lion (Campillo *et al.*, 1989).

The main market is to the hotel and restaurant trade, either directly from fishermen and merchants or via wholesale (Shaw and Curry, 1989b).

High fishery mortalities in juveniles, fish under 32 cm, were recorded during the eighties which resulted in the introduction of conservationist legislation (Pawson and Pickett, 1987; Pawson and Pickett, 1990; Kelley, 1991) such as:

- increase in the minimum landing size from 32 cm to 36 cm;
- restrictions on the use of gill and similar nets;
- prohibition on bass fishing in nursery areas for all or part of the year in the UK.

✣ In original statistics from (F.A.O., 1995), the unit used is metric ton, an American designation to distinguish it clearly from the traditional American ton. The metric ton is equivalent to the ton, which is equal to 1000 kilograms. (Rowlett, 1998, available in http://www.unc.edu/~rowlett/units).

Table 4. Wild landings (in 10^3 Kg) of seabass.

(F.A.O., 1995). Estimates in italics, shaded cells represent lack of information.

	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Channel Is.	15	14	*29*	48	*50*	46	36	*37*	83	*77*
Cyprus	1	0	3	10	15	15	29	33	20	99
Denmark			18	2	0	0	0	1	0	1
Egypt						720	720	1139	720	904
France	4206	4729	4081	3967	3344	3340	3614	4545	4704	5209
Greece	254	177	176	758	2204	2801	5664	7926	7363	10068
Ireland		4		1	1	0	0	0	0	0
Israel								100	300	700
Italy	550	750	930	1100	1050	1538	1826	2000	2850	3600
Malta						150	350	450	350	350
Morocco					24	96	119	193	332	533
Netherlands			11	2						
Portugal	462	325	167	60	93	79	61	140	196	307
Slovenia								6	13	34
Spain	419	440	480	441	572	520	506	*920*	*1000*	*1100*
UK.	129	128	171	201	191	262	157	249	549	722
Total	6036	6567	6066	6590	7544	9567	13082	17739	18480	23704

In Portugal around 97-99% of all seabass are caught by long-lines, and catches reach a peak between January and February (Figure 3). This constitutes also the peak of the spawning season in Portugal (see Table 1).

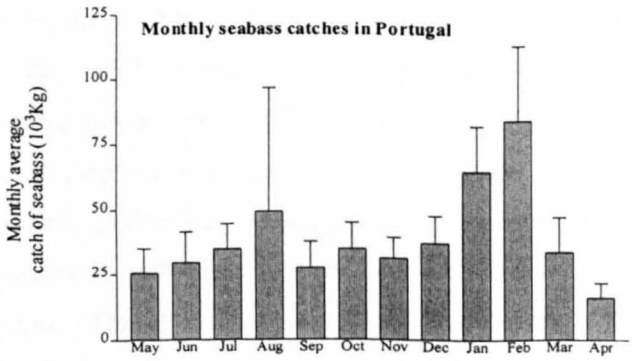


Figure 3. Seabass monthly catches averaged over 7 years (1988-1994) in Portugal.

Thin bars represent standard deviations of the mean. Sources: (G.E.P.P., 1988; G.E.P.P., 1989; G.E.P.P., 1990; G.E.P.P., 1991; G.E.P.P., 1992; G.E.P.P., 1993; G.E.P.P., 1994).

3.2. AQUACULTURE

Seabass culture is a comparatively new activity, undertaken in extensive systems in lagoons and salt ponds and more recently in sea cages, usually in the warmer parts of the species natural range, the Mediterranean and the south of the Iberian Peninsula.

Available statistics for productions levels are shown in Table 5. Wild catches and aquaculture production together amounted in 1994 to nearly $32x10^6$ Kg, with aquaculture production representing almost 42%. This increase has been made possible by cost effective improvements in broodstock management and husbandry practices. Optimisation of rearing techniques have included optimised stocking density, intensity and quality of light, photoperiod, tank colour and water quality have resulted in considerable reductions in larval abnormalities and increased survivals up to 50% in young fish (Chatain, 1991). Problems with larval spinal deformities, non functional swimbladder and malfunction of yolk sac reabsorption were hindering production and were mostly solved by a combination of factors: live feed, cleaning the water surface of oil to allow larvae to gulp air, and thus inflating a normal swimbladder (Chatain, 1991). Economic factors, such as production costs, marketing and ownership were also under firm restraint.

The increased survival of fry resulted in many more being available such that the price has since dropped by almost 50% in the last few years. There are still two main constraints to the cost effective production of table size fish (350 g): the generalised slow growth of this species and the high proportion of males in many fish farm populations. Males grow slower than females, a 2-year female reaches on average 0.5Kg, while a male of the same age is only about 0.3 Kg, and 70-90% of total production are males (Carrillo *et al.*, 1995). Overall, it can be said that fish farms are producing more commercial size seabass and as a consequence a slight decline in market profits has been observed (Figure 4). Nevertheless, the market price is still attractive, particularly when compared to seabream produced in the same geographical region. Therefore, it seems likely that the main farmed supply in the near future will continue to come from aquaculture facilities in the warmer European and north African waters where rearing is most cost effective.

Table 5. Seabass aquaculture production per country (10^3 Kg).

(F.A.O., 1996). Estimates in italics.

	1986	1987	1988	1989	1990	1991	1992	1993	1994
Algeria	*5*	*6*	7	*5*	2	4	*3*	*6*	*5*
Cyprus	1	0	3	10	15	15	29	33	20
Egypt						720	720	1139	720
France	90	140	145	250	300	414	550	1330	2138
Greece	90	*70*	*110*	300	1952	2530	5043	7345	6870
Israel								75	*145*
Italy	550	750	930	1100	1050	1538	1826	2000	2850
Malta						150	350	400	350
Morocco						56	121	120	107
Portugal	52	52	52	5	2	3	8	83	120
Slovenia									34
Spain	31	38	29	24	31	92	143	370	351
Tunisia	30	40	316	*300*	283	305	161	419	571
Total	849	1096	1592	1994	3635	5827	8954	13320	14281

There is already a certain amount of imported farmed Egyptian, Moroccan and Tunisian bass seen in the major European markets (Shaw and Curry, 1989a).

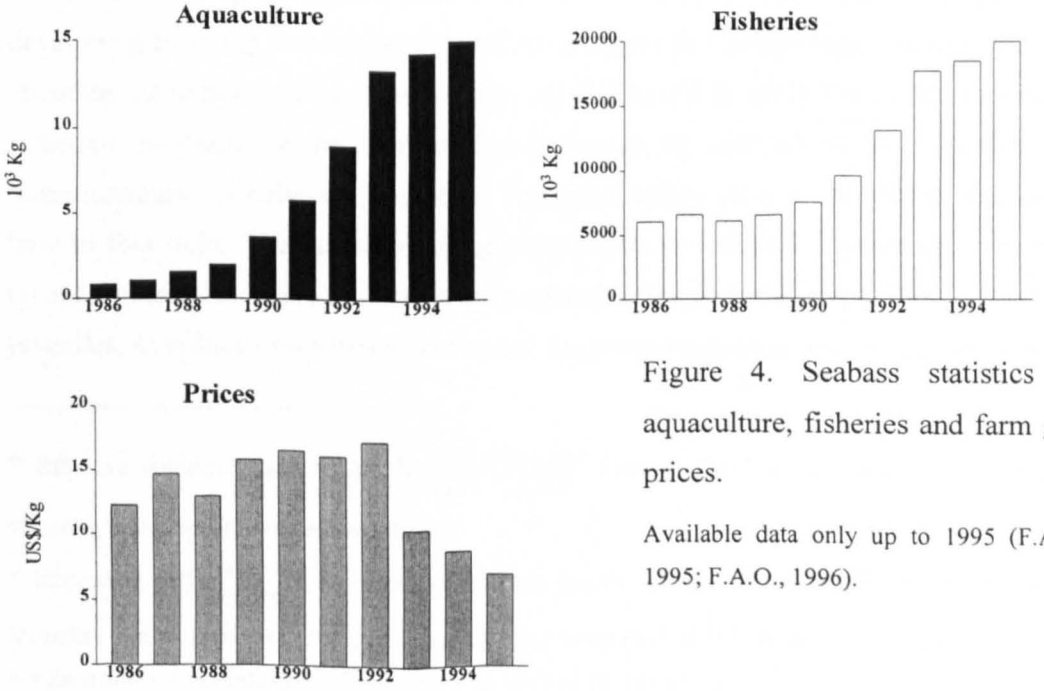


Figure 4. Seabass statistics on aquaculture, fisheries and farm gate prices.

Available data only up to 1995 (F.A.O., 1995; F.A.O., 1996).

To date the use of genetic techniques to improve or manage this species are almost totally absent from the industry. The usual husbandry practices tend to increase inbreeding as they tend to rely on very few individuals as progenitors, and hatcheries maintain broodstock populations with skewed sex-ratios, normal 1:2 or 1:3 female:male. Skewed sex-ratios dramatically reduce the *Ne*✣ of a population and increase the rate of inbreeding✦. The impact of these practices has already been linked to poor quality eggs and the low survival of larvae in hatcheries (Martínez *et al.*, 1990; Carrillo *et al.*, 1995). The lack of domesticated stocks, in the sense that there is no broodstock management or selection process in place by which farmers direct production towards faster growing, healthier, less aggressive fish, certainly also constrains production.

Today there are no commercial selectively improved seabass strains available and so fish farms have really no option, but to return to the wild to get, preferably, mature fish (that will be either induced to spawn or if allowed to adjust to captivity, will spawn spontaneously), or to exchange fish with other farms. In the very few cases where domesticated stocks exist, these as well as other hatcheries, relied entirely on wild caught fish or buying eggs and/or juveniles from other hatcheries. This often resulted in hatcheries being established from a very narrow genetic base. With the emphasis on developing breeding and husbandry techniques, genetic management has received little attention. Its importance has been recognised and now 5 farms in France are conducting selection programmes on seabass, co-ordinated by the SYSAAF (Syndicat des Sélectionneurs Avicoles et Aquacoles Français) which gives them advice and know-how in this field. The largest running programme was started 5 years ago, the most recent 2 years ago. Three objectives are targeted by these farms: to produce "improved" juveniles, to reduce on-growing cost and to improve the aquaculture product quality.

✣ Effective population size (Ne)= $\frac{4 \times Males \times Females}{Males + Females}$ (Tave, 1993); number of parents contributing offspring into the next breeding population.

✦ Inbreeding (F) = $\frac{1}{2Ne}$ is the mating of relatives that usually contributes to a decline in growth rate, fecundity and an increase in the percentage of deformed/abnormal fish; small Ne increases ΔF, so large numbers and 1:1 sex-ratio is the best way to reduce risk of inbreeding.

Coordinated efforts are now being made through the EU and MEDRAP organisations to identify the main constraints in bass culture and instigate targeted research into the genetic improvement of this species.

Having clearly established a market, with a constant demand, the seabass has future economic prospects and producers are looking forward to the contribution of genetic improvement. Breeders have clearly identified an acute need for the quantification of the variability in their stocks which are or will be submitted to selection and the identification of suitable breeding strains (identification of genetically different populations and their capacities).

These questions cannot be answered without further knowledge of the genetic variability present in wild populations, their breeding structure and a more detailed knowledge on the reproductive behaviour of the seabass.

4. POPULATION GENETICS

This chapter will firstly address the major applications of population genetics to fishery management and mention some examples; it will then refer to the genetic markers/techniques primarily used and typical examples, and finally the state-of-the-art in relation to the population genetics of the seabass.

4.1. POPULATION GENETICS APPLIED TO FISHERY MANAGEMENT

Most of the work done on population genetics addresses 4 fundamental problems: i) identification of morphologically cryptic species; ii) phylogenetic and phylogeographic analysis; iii) amount and spatial structure of genetic diversity and iv) temporal genetic change. Analysis of both protein and DNA polymorphisms in aquatic organisms have been employed mainly in studies of genetic variation within populations, genetic divergence between populations, mixed stock fisheries, hybridisation and introgression, kinship analysis, individuality and parentage and interspecific phylogeny.

Unquestionably one of the most prominent fields of marine genetic research has been the application of population genetics to fishery management. In the last decade or so, there have been books and scientific meetings dedicated to the subject. Two major reviews were published (Ryman and Utter, 1987; Carvalho and Pitcher, 1995) containing an assessment of the range of approaches available, their relative strength and weaknesses and the diversity of applications to which they are best suited. There have been a number of important meetings such as the "Biochemical Genetics and Taxonomy of Fishes" and the "Molecular Biology in Fish, Fisheries and Aquaculture", organised in 1991 and 1995, by the Fisheries Society of the British Isles. Organisations such as ICES (International Council for the Exploration of the Sea)/WGAGFA (Working Group on the Application of Genetics in Fisheries and Mariculture) have also dedicated intense efforts to analyse the contribution of populations genetics in fisheries and discussing guidelines that may be used by the people in the field (I.C.E.S., 1997; I.C.E.S., 1998).

Reviewing the most important advances resulting from the interaction between population genetics and fisheries research, one observes that the use of molecular genetic techniques has expanded dramatically over the past several years, largely due to the increased availability of techniques with different possibilities and the awareness of the value of genetic data. Research in population genetics and fisheries generally involves the characterisation of molecular genetic variants that can be used as genetic markers and it will be briefly mentioned in the next section, the many types of molecular markers that are being used in the field.

A survey of reports on genetic population structures of different marine fishes, shows a wide variation in the degree of intraspecific differentiation (Table 6). At one extreme, species such as the black perch (*Embiotica jacksoni*) display genetic differences among local populations at levels equivalent to those observed among congeneric species. At the other extreme are fishes such as the halibuts (*Hippoglossos spp.*) where no apparent differentiation has been detected, in spite of the extensive geographic sampling range. Between these extremes are a number of species, such as the herring (*Clupea harengus*) and cod (*Gadus morhua*), where at least some degree of genetic differentiation has been detected. The comparison, for instance, of the degree of population differentiation over the entire geographic range (north-eastern Atlantic and Mediterranean) of four flatfish species (*Platichthys flesus* – flounder; *Solea vulgaris* – Dover sole; *Scophthalmus rhombus* – brill and *Scophthalmus maximus* – turbot) with similar habitat and life history features, lead to entirely different conclusions. There was a high degree of differentiation in the flounder (Galleguillos and Ward, 1982; Borsa *et al.*, 1997), and sole (Kotoulas *et al.*, 1995a), no differentiation in the brill and an intermediate pattern in the turbot (Blanquer *et al.*, 1992). The comparison of the flounder, sole and turbot suggest that gene flow and consequently the population genetic structure is affected by the water temperature during the reproductive period, which is critical for the survival of the offspring during their pelagic stage (Kotoulas *et al.*, 1995a).

Table 6. Examples of differing degrees of genetic differentiation among fish species. (modified and expanded from Utter and Ryman, 1993).

Species scientific name	Species common name	Observation	Type of molecular marker	Reference
Solea vulgaris	Dover sole	Geographic unit of population structure lies within a radius of the order of 100km	Allozymes	(Kotoulas *et al.*, 1995a)
Platichthys flesus	Flounder	Four separate populations which are totally geographically isolated	Allozymes	(Borsa *et al.*, 1997)
Scophthalmus maximus	Turbot	Black Sea-Aegean Sea populations different from the rest of the species natural range	Allozymes	(Blanquer *et al.*, 1992; Bouza *et al.*, 1995)
Scophthalmus rhombus	Brill	No apparent differentiation between Atlantic and Mediterranean populations	Allozymes	(Blanquer *et al.*, 1992)
Pristimoides filamentosus	Snapper	No differentiation among populations separated by thousands of Km	Allozymes	(Shaklee and Samollow, 1984)
Anguilla anguilla	European Eel	No differentiation among populations	Microsatellites	(Daemen *et al.*, 1996)
Embiotoca jacksoni	Black perch	Genetic distance among local populations is of the same magnitude as that commonly observed among congeneric species	Allozymes	(Waples, 1987)
Clupea harengus	Atlantic herring	No apparent divergence of eastern and western Atlantic populations, genetically distinct populations in Norwegian fjords	Allozymes	(Grant, 1984) (Jørstad *et al.*, 1991)
Gadus morhua	Atlantic cod	Genetic divergence between Baltic and Atlantic populations	Allozymes Microsatellites	(Mork *et al.*, 1985) (Ruzzante *et al.*, 1996a)
Merluccius merluccius	Atlantic hake	Distinct Mediterranean and Atlantic populations	Allozymes	(Roldan and Pla, 1995)
Sciaenops ocellatus	Red drum	Genetic isolation of Atlantic and Gulf of Mexico populations	Allozymes	(Gold *et al.*, 1993)
Hippoglossos spp.	Halibut	No evidence of substructuring within either *H. hippoglossus* (Atlantic) or *H. stenolepis* (Pacific)	Allozymes	(Mork and Haug, 1983; Grant *et al.*, 1984)
Reinhardtius hippoglossoides	Greenland halibut	Barents Sea population different from Greenland waters	Allozymes	(Igland and Naevdal, 1995)
Salmo salar	Atlantic salmon	Levels of genetic variation more similar to that found in marine species, with a discontinuity between Europe and North America	Allozymes Minisatellites mtDNA	(Ståhl, 1983; Davidson *et al.*, 1989; Bermingham *et al.*, 1991; Taggart *et al.*, 1995)
Sardinella aurita	Round sardine	Mediterranean populations genetically closer to West Atlantic ones than to southern-east Atlantic	Allozymes	(Chikhi *et al.*, 1998)

However, despite the molecular genetic data generated in the last two decades and, as said, the increase use of DNA methodologies in stock structure analysis (Ovenden, 1990; Wright, 1993; Park and Moran, 1995), it seems that only information produced from anadromous salmonids has been unambiguously useful for fishery management (Carvalho and Hauser, 1995). The salmonids represent a rather special case (Carvalho and Hauser, 1995), with both financial and biological reasons to justify it: on one hand the salmonids have a high commercial value drawing research funding more easily, and on the other hand, their anadromous life history results in a large number of identifiable populations and their tetraploid ancestry increases the number of scorable loci. In fact, anadromous salmonids, mainly the Pacific salmon species have played a central role in the development and current application of procedures for population and stock analysis (Utter and Ryman, 1993).

When anadromous salmonids from different isolated and genetically distinct population mix in the ocean, they are frequently captured as a mixed stock, before completing their migration back to their natal streams for spawning. The fishery manager is usually concerned with identifying the origins and contribution of these stocks to avoid an excessive harvest of the less robust populations contributing to the fishery. In the case of Pacific salmon in particular this became possible using genetic markers because genetic research identified ample genetic variation at the allozymes level which could be used for stock differentiation (Utter, 1991) and for estimating relative stock contributions to mixed-stock fisheries (Shaklee *et al.*, 1990b) for instance. The success of molecular genetic techniques that enable the Pacific salmon runs to be managed, almost on a week by week basis is unique in the fisheries world. However, even within the salmonid group not all studies have proved so immediately useful. In the Atlantic salmon, *Salmo salar*, most studies have generally shown low levels of genetic variation with allozymes (Ståhl, 1983) and mitochondrial DNA (Davidson *et al.*, 1989) discriminating between North American and European Atlantic and Baltic populations on the basis of diagnostic mtDNA RFLP's (Bermingham *et al.*, 1991). This has forced workers to identify and use more variable markers as microsatellites (a class of variable number of tandem repeat genetic markers that will be presented in detail in

Part 4 of this thesis). This has uncovered high levels of genetic variation and observed heterozygosity at most scored loci (McConnell *et al.*, 1995b) of Atlantic salmon.

On the whole, most data generated has proven useful even if only to less immediately applied studies on fish populations (Carvalho and Hauser, 1995) although as referred there are already examples of direct application to fishery management.

4.2. MOLECULAR GENETIC MARKERS AND TECHNIQUES COMMONLY USED IN POPULATION GENETICS

One of the most important factors in population genetics in the context of fisheries management is the ability to select the most appropriate markers and effective technique(s) to answer a particular problem. In Table 7 are presented qualitative comparisons of molecular methodologies, although some of the evaluations are based on general cases and some are debatable. It seems evident from the literature that the relative value of different methods depends largely upon the species and the geographic scale to which they are applied, as well as the questions asked. No single marker/technique is suited to all applications, as the different genetic markers differ in their ability to detect genetic variability.

The development of techniques to study genetic polymorphism at the DNA level rather than at the protein level since the late 1970s have been rapid. This started with the discovery of restriction endonuclease digestion of DNA which was followed by Southern blotting, DNA sequencing and gene cloning and more recently PCR, random amplified polymorphic DNA (RAPD), mini and microsatellites. All these techniques provided opportunities to study different forms of DNA polymorphism with varying degrees of success.

Many commercially important species were first studied in terms of their population genetics by the use of protein electrophoresis. Extensive reviews on the application of isozymes to fisheries have been published (Shaklee *et al.*, 1990b; Utter, 1991; Utter and Ryman, 1993), although this technique is nowadays, regarded by some, as old fashioned. However, this type of polymorphism is well suited for many studies of intraspecific variation like determining mating systems, heterozygosity evaluation, genetic variation between geographically different population and also interspecific

variation like the detection of hybrid zones and in phylogenetics studies (Carvalho and Hauser, 1995; Park and Moran, 1995; Ward and Grewe, 1995). Therefore, electrophoretic analysis of protein has been applied to many actual fishery management problems in a number of species, providing basic information necessary to manage exploited fisheries properly (Allendorf *et al.*, 1987; Park and Moran, 1995; Utter, 1995; Ward and Grewe, 1995).

Table 7. Qualitative comparisons of molecular methods.

(modified from Park and Moran, 1995; I.C.E.S., 1997)

	Protein methods	MtDNA 1	MtDNA 2	MtDNA 3	VNTRs Single	VNTRs Multi	Microsatellites	RAPDs
Tissue	F/Fr	F	Fr	E	F/Fr/E		F/Fr/E/D	F/Fr/E
Tissue quality	High	H	L	L	High		Low	High
Specimens sacrifice	O	O	N	N	O		N	N
Percent of genome	< 5	$< 10^{-4}$			5-20		5-20	Unknown
Number of loci	10^2 - 10^3	1			10^4		10^4-10^5	Unknown
Cost (labour/cons.)	Low	M	H	H	High		High	Medium/High
Reproducibility	Good	Good			Good		Good	Poor
Automation	No	No	Y	Y	No[2]		Yes	No
Relative rate of evolution among categories	Mo	Rapid			Mo/rapid		Rapid	Rapid
Pedigree analysis	No	NS			Good		Good	Fair
Population genetics	Good	Good			Good	NS	Good	Fair
Hybrid zones	Good	Good			NS		NS	Fair
Phylogenetics	Fair[†]	Good			NS		NS	NS

1. RFLP analyses, 2. PCR/RFLP, 3. PCR /sequencing. H=High, M=Medium, L=Low, O=Often, Y=Yes, N=No, NS=Not suitable, F=Fresh, Fr=Frozen, E=Ethanol, D=Dried, Mo= Moderate. [†]Cichlids constitute probably the most successful application in fish (Sodsuk *et al.*, 1995), however in marine fish species there are very few applications to date. [2]Technical approaches to reduce time and hazards involved in screening minisatellite variation have been developed (McGregor *et al.*, 1996).

It was, for example, possible to obtain useful estimates of stock composition based on electrophoretic analysis of protein variation for several species: chum salmon (*Oncorhynchus keta*) (Beacham *et al.*, 1987); pink salmon (*Oncorhynchus gorbusha*) (Beacham *et al.*, 1985); chinook salmon (*Oncorhynchus tshawytscha*) (Shaklee *et al.*, 1990b) and sockeye salmon (*Oncorhynchus nerka*) (Grant and Utter, 1980).

Brown trout (*Salmo trutta*) population genetic structure, particularly from Lough Melvin (Ireland), has been extensively analysed by a number of techniques: allozymes (Ferguson and Taggart, 1991); mtDNA (McVeigh *et al.*, 1995) and minisatellites (Prodöhl *et al.*, 1992). Three distinct brown trout types exist in Lough Melvin: gillaroo, sonaghen and ferox, on the basis of their morphology (Cawdery and Ferguson, 1988) and different feeding preferences (Ferguson, 1986). Brown trout in particular is one of the most polymorphic of all fish species (Ferguson, 1989), allozyme electrophoresis usually reveals heterozygosities of less than 10%, and an average of two to three alleles (Ferguson *et al.*, 1995) per locus. mtDNA revealed 10 haplotypes (McVeigh *et al.*, 1995) and minisatellites showed 8 alleles (Prodöhl *et al.*, 1994). Overall allozymes, mtDNA and minisatellites were all able to detect the three brown trout types. MtDNA, however, outperformed the other two markers in identifying gillaroo or sonaghen individuals due to the relatively high frequencies of unique haplotypes (Ferguson *et al.*, 1995).

There are many other instances where allozymes were outperformed by other markers in detecting levels of genetic differentiation: i) orange roughy, (*Haplostethus atlanticus*), where restriction fragment length polymorphism of mtDNA did show significant differentiation (Smolenski *et al.*, 1993), when allozyme analysis did not (Elliot and Ward, 1992); ii) walleye pollock, (*Theragra chacogramma*), where mtDNA analysis revealed differentiation (Mulligan *et al.*, 1992), where allozymes did not (Grant and Utter, 1980). We will refer in more detail to other applications of this class of markers as well as the main limitations and advantages of the technique in the appropriate section of this thesis.

Restriction fragment length polymorphism (RFLP) of the mitochondrial DNA genome has been employed to determine genetic variability both between and within species in various organisms. In 1979, mtDNA was introduced to natural population genetic studies (Avise *et al.*, 1979; Brown *et al.*, 1979). The first marine fish surveyed using mtDNA RFLPs was the skipjack tuna, *Katsuwanus pelamis,* from the Atlantic and Pacific regions (Graves *et al.*, 1984) showing no population differentiation.

Several properties of mtDNA make this molecule suitable for population studies at that time (Berg and Ferris, 1984; Ferris and Berg, 1987; Bermingham, 1990; Billington and Hebert, 1991): relative ease of isolation of material; maternal inheritance and higher mutation rate compared with single copy nuclear DNA and a lack of recombination (Ferris and Berg, 1987). These properties make it particularly suitable for estimating levels of such factors as stock discrimination, migration and gene flow in natural populations (Bembo *et al.*, 1995; Magoulas *et al.*, 1996), hybridisation and introgression (Gonzalez-Villasenor *et al.*, 1986; Youngson *et al.*, 1992) and phylogenetic reconstruction (Billington *et al.*, 1990). However in some species complications in the analysis may arise due to the fact that the paternal mitochondria enter the fertilised eggs. A significant paternal contribution of mtDNA genotypes was reported in European anchovy, *E. encrasicolus* (Magoulas and Zouros, 1993) and mussel (Skibinski *et al.*, 1994).

Variation of life history inter- and intraspecifically is hypothesised to be reflected in genetic divergence of the populations. Generally, it is accepted that there are differences between salmon from different rivers leading to a stock concept: a group of salmon having a particular spawning ground at a particular time of the year, not interbreeding with other groups (Ihssen *et al.*, 1981). Some examples of the use of mtDNA in discriminating populations can be mentioned: i) the North American Atlantic salmon was discriminated from the fish of European origin by a minimum of seven restriction site differences (Bermingham *et al.*, 1991); ii) the chinook salmon (*Oncorhynchus tshawytscha*) populations from British Columbia were discriminated from the Alaskan sites in the Pacific Northwest (Wilson *et al.*, 1987); iii) genetically distinct population of striped bass, *Morone saxatilis*, were found in the Gulf of Mexico (Wirgin *et al.*, 1989); iv) mtDNA data were effective at discriminating between walleye

stocks (Ward *et al.*, 1989); v) genetically distinct populations of American shad, *Alosa sapidissima*, were found along the Atlantic coast of North America (Bentzen *et al.*, 1989); vi) marked mtDNA differences were detected between swordfish (*Xiphias gladius)* populations of the Atlantic and Mediterranean (Kotoulas *et al.*, 1995b).

The analysis of population structure using mtDNA in the red drum, *Sciaenops ocellatus,* an important recreational species in the Gulf of Mexico, and along the Atlantic coast of the Southeast of the United States, showed that the population as a whole was weakly subdivided, with semi-isolated subpopulations occurring along the south-east of Atlantic coast and in the Gulf of Mexico, less clear cut but useful information (Gold *et al.*, 1993).

Nevertheless, in spite of the advantages of mtDNA it is not a universal panacea, for instance it failed to detect genetic structure in: i) in deep-sea orange roughy, *Hoplostethus atlanticus,* collected from the east and west coasts of Tasmania (Ovenden *et al.*, 1989); ii) in albacore tuna, *Thunnus alalunga* from the north and south Atlantic (Graves and Dizon, 1987); iii) in yellow fin tuna, *Thunnus albacares* (Ward *et al.*, 1994a); iv) in Atlantic cod, *Gadus morhua* from the eastern and western Atlantic populations (Smith *et al.*, 1989). In the yellow fin tuna for instance, mtDNA showed no evidence of the structure detected using allozymes (Ward *et al.*, 1994a).

Protein or mtDNA markers are based on changes in DNA sequence generally as a result of point mutations involving base substitutions. However, more recently the focus has turned to another type of variation, that of differences in the number of repeated copies of a segment of DNA. Most genomic DNA of eukaryotes contains large amounts of highly repetitive sequences scattered throughout the genome. The properties of this class of markers will be referred to in greater detail in the microsatellite chapter of this thesis. The "minisatellite" repeats also known as VNTR (variable number of tandem repeat) loci, are dispersed tandem repeat units of about 9-64 bp in length. Polymorphism is caused by an increase or decrease in the length of particular arrays resulting in changes in the repeated number which is due to unequal crossovers during meiosis and DNA slippage during DNA replication (Jarman and Wells, 1989; Avise, 1994).

DNA fingerprinting techniques were introduced in 1985 (Jeffreys *et al.*, 1985) using Southern blot analyses of hypervariable DNA regions. The DNA fingerprint profiles result from restriction endonucleases cleaving total DNA extracted from an individual outside the repeat clusters. The DNA is separated in a gel and transferred to a membrane which is then probed by a given repeat sequence. The original probes i.e. 33.6 and 33.15 isolated from myoglobin intervening sequences, hybridise to 10-15 bp long conserved core sequences of the repeated units in the human genome revealing band profiles distinguishing all individuals except monozygotic twins (Jeffreys *et al.*, 1985). From each locus of an individual either one or two bands is revealed depending upon whether the locus is homozygous or heterozygous with respect to tandem repeat numbers. They are called multilocus DNA fingerprints if the probes hybridise to several loci scattered among the genomic DNA. The bands making up the fingerprint patterns are inherited in a Mendelian manner, each band in an individual DNA fingerprint profile is derived from either its biological father or mother. These multilocus DNA finger prints are however generally too variable for population genetic analysis - every individual is likely to be unique - but are very useful with pedigree or family groups.

The multilocus polymorphic makers were in some instances successful when applied to fishes, as in the case of the striped bass, *Morone saxatilis*, where a genetically distinct population was detected in the Apalachicola-Chattahoochee-Flint river system (Gulf of Mexico) (Wirgin *et al.*, 1991), providing justification for special fishery management measures. Nevertheless, its wider application in population genetic studies of fisheries, has remained problematic (Park and Moran, 1995) because the complexity of the patterns frequently produce ambiguous genetic estimates (Wright, 1993). This complexity of multi-locus fingerprints is mainly due to high number of bands, the intensity of hybridisation signal that might result from the co-migration of bands from the same locus or just similar sized DNA framents from different loci (Wright, 1993). Although standard population statistical measures cannot be applied to these data, heterozygosity, genetic distance, diversity and family information can be obtained from such patterns (Ferguson *et al.*, 1995).

To avoid this problem it is possible to clone a single minisatellite locus and hybridize this under high stringency conditions, to genomic DNA blots. Individual

DNA profile will have either one or two bands, depending on whether such an individual is homozygous or heterozygous. This improves the fidelity and interpretation of results but does require greater investment in probe development. Their abundance in the salmonid genome, for instance, is quite remarkable: minisatellites loci hybridising to the human-derived fingerprint probes (33.15 and 33.6, Jeffreys *et al.*, 1985) appear to be several times more abundant in salmonid fishes than in the human genome (Taggart and Ferguson, 1990b). In the beginning of 1990, this technique was first developed for salmonid and tilapia species (Taggart and Ferguson, 1990a; Bentzen *et al.*, 1991). The application of this technique has been particularly successful in Atlantic salmon: i) demonstrating significant genetic differentiation between Atlantic salmon from both sides of the North Atlantic (Taggart *et al.*, 1995); ii) genetic stock identification based on three minisatellite loci provided information on stock composition (Galvin *et al.*, 1995); iii) minisatellite data is more accurate than the allozyme data for a given sample size to trace the tributary origin of Atlantic salmon from the river Shannon in Ireland (Galvin *et al.*, 1995); iv) four minisatellite VNTR loci with several alleles (5-12), with heterozygosities ranging from 40-66% were used to differentiate between populations of western European Atlantic origin (Stone *et al.*, 1995).

The PCR (Polymerase Chain reaction) (Mullis and Faloona, 1987) is a tool for molecular biology that has revolutionised molecular genetics. The technique is so sensitive that allows the amplification of a particular DNA sequence many thousand-fold from a single DNA molecule. The product is sufficient for direct sequencing or restriction analysis. The most common PCR targets in fisheries are mtDNA and nuclear DNA sequences both microsatellite and randomly amplified sequences.

The analysis of randomly amplified polymorphic DNA (RAPD), where primers of random nucleotide sequence, of about 10 to 20 bp long, are used in PCR reactions to amplify anonymous regions of nuclear DNA, has shown low stringency and repeatability. The technique is very sensitive to slight changes in amplification conditions, and has had limited use in fisheries (Ferguson *et al.*, 1995; Park and Moran, 1995).

Most of the current research effort has concentrated on microsatellites or simple sequence repeats (SSR), consisting of short (1 to 6 bp) tandem arrays (Tautz, 1989).

Similarly to allozymes, we will refer to this class of markers in detail in the appropriate section.

4.3. POPULATION GENETICS OF THE SEA BASS

There have been a number of studies that have attempted to study the population genetics of the European seabass. The first of these collected samples from natural populations of the seabass from the Atlantic (Concarneau, N[✣]=21), the French Mediterranean (Espiguette, N=20 and Mauguio, N=28), the Tunisian Mediterranean (Gabès, N=20), and from a French hatchery population (Sète, N=30) were screened for genetic differentiation (Benharrat *et al.*, 1984). That study included 34 allozyme loci, 15 of which were found to be monomorphic in all the samples, the other 19 were polymorphic in at least one of the samples. The overall conclusion was that the Mediterranean populations were relatively homogeneous, but there was genetic differentiation between Mediterranean and Atlantic populations. This differentiation was due to significant allele frequency differences at three loci (*EST-2**, *MEP-1** and *AAT-1**) and higher heterozygosity levels in the Atlantic. The sample from the hatchery was only polymorphic at 4 loci (*GPI-1**, *IDHP-2**, *PGM-2** and *SOD**) and this clear reduction in variation was considered to be due to low numbers of broodstock. This work estimated Nei's (1972) genetic distance between Atlantic and Mediterranean populations, to range between 0.0101 and 0.0152, and among Mediterranean populations the values range was 0.0011 and 0.0016.

The first Spanish seabass samples were assessed in 1990 (Martínez *et al.*, 1990) in a study where 11 polymorphic loci were found (*EST-2**, *AAT-3**, *AAT-4**, *LDH-1**, *LDH-2**, *MDH-1**, *MDH-2**, *PGM-2**, *To-1*[✧], *To-2* and *XDH-2**). This work included 2 hatchery samples: one from Tinamenor hatchery in Cantabria (North of Spain, Gulf of Biscay, Atlantic), with 20 individuals and another from Mar Menor (Mediterranean), with 100 individuals and a third sample of wild caught fish from Gulf of Cadiz (Atlantic), with 96 individuals. The wild sample revealed higher heterozygosity and

---

✣ N= sample size.

✧ Loci designations as in original papers, no full description given. May correspond to *SOD**.

polymorphism (H_{obs}[☆]=0.021 and P[✵]=0.1), while the population from the Mediterranean hatchery showed the lowest values (H_{obs}=0.008 and P=0.042). In another paper (Martínez *et al.*, 1991), presumably the same samples from Cantabria and Cadiz, where further analysed, although the number of fish in the sample from Cadiz varied (from 96 in the previous work to 100 in this). Also more details on the sample from the hatchery in Cantabria are given. This sample originated as eggs from a wild stock captured in Ria de Santander and was subsequently moved to Scotland. From the 61 loci scored in the hatchery sample, only 4 were polymorphic (*LDH-1**, *LDH-2**, *MEP-2** and *XDH-2**) (H_{obs}=0.015 and P=0.066). In the natural population, from the 49 loci studied, 5 presented more that one allele (*ADH**, *AAT-3**, *EST-3**, *MEP-1** and *SOD-2**) (H_{obs}=0.018 and P=0.102). The reduction in polymorphism observed in the hatchery sample, could have been due to bottlenecking and genetic drift occurring in the Tinamenor stock when it was captured from wild and moved to Scotland. In this work the authors claim the existence of two genetically divergent groups, sharing no common polymorphic loci (Martínez *et al.*, 1991). However, the exact provenance of the hatchery fish could be in doubt as many European seabass hatcheries mix broodstocks or eggs from different origins. The polymorphic loci observed in the Tinamenor sample are atypical of those observed in the other Atlantic populations so far analysed, but have been observed in Mediterranean populations. *LDH-1**, *LDH-2** and *MEP-2** were found to be polymorphic in the Mediterranean (Benharrat *et al.*, 1984) but not in the Atlantic (Benharrat *et al.*, 1984; Child, 1992). The genetic distance values between North and South of Spain samples were estimated as 0.012 with a standard deviation of 0.018.

The most northerly seabass sampling location to have been analysed to date is the UK. In UK populations, the *PGM** locus was found to be particularly important in the differentiating between seabass from the Irish Sea and the English Channel and Thames Estuary (Child, 1992). Other loci such as *SOD**, *EST** and *GPI-1** and *GPI-2** loci were also shown to be polymorphic but no significant differentiation could be detected. The position of the seabass from the Bristol Channel was not clear, results suggest that

☆ H_{obs} = observed heterozygosity.

✵ P = polymorphism.

bass from that region may be more similar to the English Channel rather than the Irish Sea bass. No differences could be detected among or between age groups at a site and between wintering adult bass and summer-caught bass (Child, 1992). No data on heterozygosity, F_{ST} values or genetic distances are given in the original paper.

Several interesting studies on seabass from Italian waters have been published since 1985: 19 enzyme systems were surveyed in an Italian hatchery (Pallestrina-Venezia) two of which were polymorphic, *PGM-2** and *SOD-1**, with two alleles each (Cervelli, 1985); another Italian hatchery population (Trapani) was assessed for enzymatic ontogenetic variation at six different enzyme loci, and it was found that *ADH** was active only in the liver of adults, *G3PDH** was active in only two tissues of adult individuals (muscle and liver), and that *GPI**, *LDH** and *MDH** are completely active 30 days after hatching meaning that cellular differentiation of tissues where these enzymes are predominant and most active is completed within this time (Basaglia *et al.*, 1989). However, no data on genetic variation is provided as it was beyond the scope of the paper. The acclimation of seabass to fresh water was studied (Allegrucci *et al.*, 1994) by the assay of 28 loci coding for 24 enzymes using cellulose acetate strips and seven of these loci were shown to be polymorphic (*ADK-3**, *CK-4**, *EST-2**, *FBP-2**, *G6PDH**, *MEP** and *NP**). Multivariate analysis of individual genotypes indicated that survival was not random with respect to genotype, suggesting a selective response to the changed environmental regime and the conditions of stress imposed by the acclimation which caused high mortality rates. These authors propose that several coadapted genetic pools exist representing "races" of seabass more or less well adapted to fresh water. The authors also mention that preliminary results on some wild populations of seabass from different geographic origins within the Tyrrhenian sea revealed the presence of geographic variation, but no details were presented in this paper.

The same group then used RAPDs on cultured populations and showed very high levels of polymorphism (Allegrucci *et al.*, 1995). The analysis of 126 polymorphic RAPD markers revealed strong polymorphisms (76.2%) and a genetic diversity between 0.22 and 0.25 in non acclimatised fish, levels that are maintained after acclimatisation. Changes in genetic frequencies before and after acclimatisation were

observed in 2 markers from the 1989 sample and 13 markers from the 1990 samples. Comparing the results from this work with the ones from the previous work (Allegrucci *et al.*, 1994) the authors state that the most important difference lies in the number of markers available to interpret the results. The number of allozyme loci used was quite small compared with the number of RAPD markers used.

In 1997, Allegrucci *et al.* (1997) reported a study of 28 allozyme loci in wild and artificially reared seabass samples, originating from 10 sites, either coastal lagoon or marine sites in the Mediterranean Sea and one wild population from the Portuguese coast (Aveiro). On average 30 individuals from each sample were scored, but no precise data on each sample size are given. In total 26 loci were polymorphic, and the results indicated significant population structure within Mediterranean seabass taken from marine sites, with a multilocus F_{ST} value of 0.34. However, little or no structure was detected when the seabass samples came from freshwater lagoons spread over the same geographic range. Six of the loci showed consistent allele frequency differences between marine and lagoon samples *(AH*, FBALD*, CK-4*,G6PDH*, MEP*, PNP*)* and four others contributed to most of the population differentiation (*ADA*, ADK-4* EST-2*,GDA**). The authors noted that apart from *AH** and *FBALD**, most of these loci display the differential mortality in trials associated with acclimatisation of young seabass to freshwater (Allegrucci *et al.*, 1994). Again multivariate analysis of individual allozymic profiles and of allele frequencies suggested that different arrays of genotypes prevail in lagoons compared to marine samples, particularly at those loci implicated in adaptation to freshwater. Interestingly, these authors also included a Portuguese sample from Aveiro and estimated an average genetic distance between it and the Mediterranean populations of D = 0.236. This large difference was caused by the presence of alternative alleles at six loci which led them to call into question the taxonomic status of the Portuguese population.

Direct sequencing (Patarnello *et al.*, 1993) of the cytochrome *b* mitochondrial gene was carried out in 40 individuals from 7 samples: a farm in the Northern Adriatic Sea (N=11), and wild samples from the Northern Adriatic (N =8), Southern Adriatic (N =6), Crete (N =4), Sicily (N =1), Sardinia (N =1) and Tyrrhenian Sea (N =11). The cytochrome *b* showed 6 transitional changes, which generated 5 different haplotypes.

Haplotype divergence ranged from 1 to 4 base changes, with an index of nucleotide heterogeneity ranging from between 0.47 and 0.51. Comparisons among wild populations showed a non homogeneous distribution of the two most frequent haplotypes, the absence of two haplotypes from wild populations that are present in farm populations and the exclusive presence of one haplotype in wild populations with a frequency of 35%. The results seem to indicate a structuring of Mediterranean bass populations into two groups, Eastern and Western, although as the authors admit the statistics are not significant (Fisher's exact test comparing wild population of East and West Mediterranean gave a value of $0.08>p>0.07$). The number of individuals per sample is extremely low, which can also affect the results.

More recently, evidence for population differentiation on a regional scale, using microsatellite markers was presented (García de León *et al.*, 1995; García de León *et al.*, 1997). These authors analysed 172 mature fish caught in four locations: three from Gulf of Lion (Rhône, N=50; Grau-du-Roi, N=50; Western Rhône, N= 35) and one from the Gulf of Valencia (N =37). The results suggest a slight genetic differentiation between the two gulfs, with an F_{ST} value of 0.007, $P<0.05$, indicating that the corresponding populations are likely to be dynamically independent. These studies also uncovered high levels of polymorphism and heterozygosity: the six dinucleotide microsatellite loci revealed between 13 and 30 alleles, with heterozygosities from 0.77 to 0.82, when all four populations were combined. This finding is in agreement with similar population studies with microsatellite markers in other species (Atlantic cod, *Gadus morhua*, Ruzzante *et al.*, 1996a) and led them to suggest that microsatellite variation will be a valuable tool in studies of genetic tagging, marker-assisted selection programs and identification of natural stocks (García de León *et al.*, 1995).

The genetic differentiation among seabass populations on a wider geographic scale, including samples from the North Sea, the Gulf of Bicay, the Lusitano-Moroccan Gulf, the Alboran Sea and the Western Mediterranean, with 6 microsatellites loci, show that the seabass populations cluster into two homogeneous groups (Naciri *et al.*, submitted). On one side, the Atlantic group, which includes the Alboran Sea east of the Gibraltar Strait, and on the other a Western Mediterranean group, with sample from Valencia, Sète, Marseille, Sicilly and Anaba in Algeria. The samples analysed included

a sample of 50 individuals from Aveiro provided to the Laboratoire Génome et Populations, at the University of Montpellier and constitutes a subset of the sample that are analysed in the present work. This particular sample clusters within the Atlantic sample and it is by no means genetically as different as the sample analysed by Allegrucci *et al.* (1997) with allozymes.

Four single-locus minisatellite probes were developed for seabass (Benedetti *et al.*, 1995), from 3 males and 3 females and with *MboI* restriction enzyme. The results from a sample of 18 adult wild fish show a high variability index (44-72%) and high heterozygosity (78-94%). The authors present their results as very promising especially in regarding to individual typing.

The data bank GenBank/EMBL contains two seabass sequences: RNAm of the growth hormone gene (615 bp long) (Doliana *et al.*, 1992) and RNAt-thr and pro (150 bp long) (Cecconi *et al.*, 1993).

Most, but not all, of the population genetic studies on the seabass have used allozyme electrophoresis and although there is little consistency in the loci and techniques used there appear to be significant differences between Atlantic and Mediterranean populations. These differences appear to be based on allele frequencies at *AAT-1** (Benharrat *et al.*, 1984), *IDHP** (Allegrucci *et al.*, 1997), *FBALD** (Allegrucci *et al.*, 1997) and *MEP** (Benharrat *et al.*, 1984; Allegrucci *et al.*, 1997). It is also important to note that in the few farmed populations analysed there seems to be a reduction in genetic variability when compared to wild samples and lower polymorphism (e.g. Benharrat *et al.*, 1984). However, it is quite clear that authors did not use the same procedures (e.g. gels made of starch vs. celulose acetate strip), the samples sizes were not the same and the data analyses were quite different, in that very simple statistics ($\chi 2$ test for observed and expected heterozygosity, and comparison of polymorphisms) were used (Benharrat *et al.*, 1984; Child, 1992), while some others (Allegrucci *et al.*, 1997; García de León *et al.*, 1997) estimate and test F_{ST} and present pairwise sample comparisons.

It is, therefore, of obvious interest to undertake a more in depth analysis of the genetic variability on the same samples, using different genetic markers and the estimation of gene flow, effective size and genetic substructuring.

The Portuguese seabass population lies at the centre of distribution of the species apparently with no obvious barrier to the species migration. Seabass spawning aggregations and nursery areas are spread all along the Atlantic coastline of Portugal. The Portuguese populations also lies between previous sampling sites which indicate that there might be some differentiation between the North and South of the Iberian Peninsula (Martínez *et al.*, 1991). It is therefore timely that the Portuguese populations should be analysed. Juvenile samples were collected from geographically spread out nursery sites along the coast and were subjected to a detailed analysis using allozyme and microsatellite markers to obtain estimates of population differentiation and substructuring gene flow and effective population size. The results from this allozyme and microsatellite study are assessed as individual data sets and a comparative analysis of the results from both markers is produce. Findings are discussed in the light of the latest population genetic and fisheries ideas.

PART 2 - GENERAL MATERIALS AND METHODS

5. SAMPLING

Adequate and representative sampling strategies in population genetics studies although of fundamental importance, have not received sufficient attention in the past. However even, the use of sophisticated statistical procedures and molecular genetic markers, available today, cannot make up for defective sampling (Hansen, 1998). The issue is regarded by population geneticists in general, as acute, so much so that, the W.G.A.G.F.M.* in this year's report (Hansen, 1998) suggests that a workshop on sampling sizes and the particular tests procedures to apply to population studies using highly variable molecular markers are so fundamental that this should be given a high priority.

Large numbers of individuals (Table 8) are needed when allozyme electrophoretic data are to be used to evaluate genetic differentiation between conspecific populations, species with little electrophoretic divergence or genetic changes within a population through time (Chakraborty and Leimar, 1987). This is because in such instances, reliable estimates of divergence require more accurate estimates of alleles frequencies at polymorphic loci. A clear example of this is that with samples sizes of 100 individuals, the differences in allele frequencies have to be 0.1 to 0.2 to be detected, with type I error* probability set at 5%, and type II error♦ probability set at 80% (Baverstock and Moritz, 1990). In practice, however, we find a great diversity of sample sizes in seabass work (Table 9), and of course in the work with many other species (Table 10). A rule of thumb suggests that there should be less than 5% probability of not detecting an allele present at a frequency of 5% (Allendorf and Phelps, 1981). This would correspond in practical terms to sampling 60 genes, or a sample size of 30 for diploid loci. However, this was based on allozymes in which the number of alleles per locus was usually between 1 and 4, with the majority of polymorphic loci with 2 to 3 alleles.

---

* Working Group on the Application of Genetics in Fisheries and Mariculture

* Rejection of the null hypothesis when it is in fact true.

♦ No rejection of null hypothesis when it is false.

This situation has completely changed with the advent of highly polymorphic markers, where the number of alleles per locus is typically very high, often above 20. With a large number of alleles, as in microsatellite loci, some sample sizes are bound not to be sufficient, to allow statistical tests with enough power to detect Hardy-Weinberg deviations.

Table 8. Theoretical sample sizes considered for genetic population studies based on allozymes.

[*] When allelic frequencies are anticipated to only vary between 5 to 15% among populations.

Sample size	Number of loci	Reference
30	—	(Aebersold *et al.*, 1987)
≥25	—	(Leary and Booke, 1990)
50-100	≥6 polymorphic	(Shaklee *et al.*, 1990b)
40-50	—	(May and Krueger, 1990)
≥80[*]	—	(May and Krueger, 1990)

Some authors have suggested that pilot studies should be carried out, before the commencement of the main work (Baverstock and Moritz, 1990; May and Krueger, 1990; Shaklee *et al.*, 1990b; Hansen, 1998) involving a small number of samples from geographically distant areas, and small sample sizes ($N \approx 50$). The preliminary results from this survey will give important indications as to the planning of the sampling area and sample sizes of the main work.

However, even when a sampling design is carefully laid-out, it is more often than not that practical limitations hinder sampling, e.g. weather conditions that prevent small commercial boats or bigger vessels fishing or even the unwillingness of fishermen to cooperate. Also, it has to be said, that most people try to do the best of what they can

[*] When allelic frequencies are anticipated to only vary between 5 to 15% among populations.

when sampling, and thus, frequently, the sample collection is more a function of opportunity than of design. In these conditions, although compensation for such constraints is difficult, improvements can be achieved by taking into account questions concerning the biology of the species, namely the reproductive biology, the social structure, the dispersal ability during several life-history stages including migration behaviour and the existence of larval, feeding or spawning grounds for instance (Carvalho, 1998; Hansen, 1998). These aspects are of utmost importance and there is an example, whereby modifications of the sampling strategy, combined with molecular and morphometric techniques, allowed the detection of allozymic distinct populations of European anchovy (*Engraulis encrasicolus*) (Bembo *et al.*, 1996; discussed in Carvalho, 1998).

Another important aspect of sampling is the recording of biological information on the individual fish sampled, as this information can be used to show if differentiation is for instance associated with specific cohorts/spawning populations or sex. Minimal information should include age, sex, length, weight, maturity, precise location and date of sampling (Hansen, 1998).

Table 9. Examples of type of genetic marker, sample sizes, number of loci, polymorphism, geographic distance ranges of seabass genetic studies.

Type of marker	Sample size	No. of loci	*P* (%)	Range (Km)	Reference
Allozymes	20-100; X=57 ± 60	49; 61	6.1-6.6	1400	(Martínez *et al.*, 1991)
Allozymes	20-48; X=34 ± 20	34	8.8-17.6	2800	(Benharrat *et al.*, 1984)
Allozymes	13-105; X=30 ± 21	21	NA	1000	(Child, 1992)
Allozymes	X=30	28	25.0-35.7	5000	(Allegrucci *et al.*, 1997)
Microsatellites	35-50; X=43 ± 8	6	100	600	(García de León *et al.*, 1997)
mtDNA	1-11; X=6 ± 4	-	-	1600	(Patarnello *et al.*, 1993)

Sample sizes: minimum-maximum number of individuals in a sample, X = average sample sizes ± standard deviation; *P* (%): Percentage of polymorphic loci, 0.95 criterion meaning that the most common allele has a frequency ≤ 0.95; NA – data not available; Range: geographic distance between the most distant samples .

Table 10. Some examples of sample sizes, number of loci studied and geographic range of genetic studies on different fish species.

Species	Sample size	N. loci	Range (Km)	Reference
Gadus morhua	96-100 (X=98 ± 1.9)	19	769-8307	(Mork *et al.*, 1985)
Oncorhynchus mykiss	71-100 (X=91 ± 13)	12	150	(Snowdon and Adam 1992)
Sciaenops ocellatus	30-95 (X=56 ± 18)	44	30-960	(King and Pate, 1992)
Solea vulgaris	8-120 (X=46 ± 35)	23	200-7900	(Kotoulas *et al.*, 1995a)
Sardinella aurita	30- 313 (X=127 ± 108)	25	2200-5000	(Chikhi *et al.*, 1998)

Sample sizes: minimum-maximum number of individuals in a sample, X = average sample sizes ± standard deviation; Range: minimum and maximum geographic distance ranges between samples.

5.1. COLLECTION

Collecting fish for research purposes is an activity limited by certain constraints, among the most obvious are: local abundance of fish and means to collect them, in this case being dependent on the existence of artisanal fisheries and willingness of local fisherman to cooperate. As it turned out, the main estuaries and lagoons along the coastline, namely: Aveiro, Figueira da Foz, Óbidos, Vila Nova de Milfontes and Faro (Figure 5) were the basis of the population survey. Seabass is a highly priced fish, so sampling was targeted on juveniles and special permits from the Direcção Geral das Pescas were provided to fishermen, to legally catch undersize fish. Even so, a lot of them declined to catch fish for this study, because the permits would reveal they possessed the appropriate, but illegal nets and therefore the Maritime Police would be aware of that fact. This obviously hampered the scope of the planned sampling programme.

Fish were caught between February 1993 and March 1994 (Table 11), and each locality is represented by sample sizes from 65 to 213.

Table 11. Source and number of seabass analysed genetically.
(see Figure 1 for location of sampling sites).

Locality	Sampling date	Sample size
Aveiro	February 1993	119
Aveiro	March 1994	46
Foz	November 1993	121
Óbidos	February 1993	104
Milfontes	February 1993	73
Milfontes	February 1994	140
Faro	April 1992	65

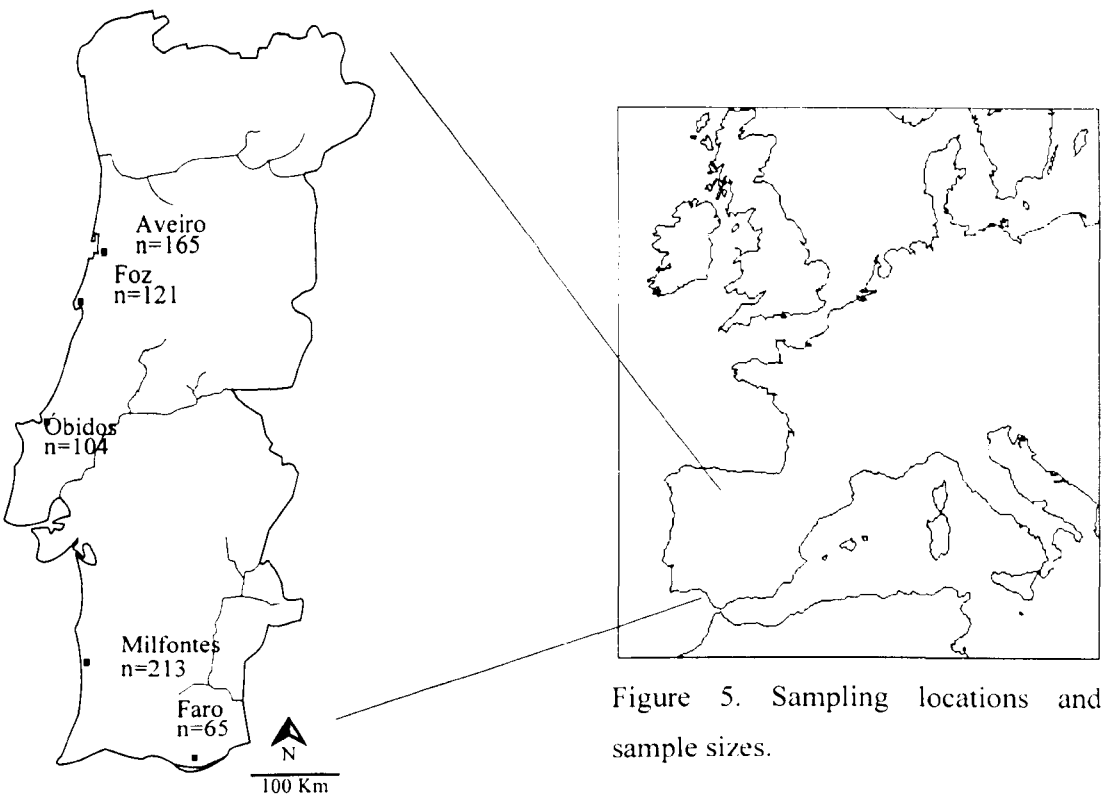


Figure 5. Sampling locations and sample sizes.

Given that the assessment of the stability of genetic frequencies in population genetics studies is particularly important, the sampling program included whenever possible repeat samples from the same site in different years.

Specimens were collected from natural populations by local fishermen using mainly two kinds of net fishing techniques: gill nets laid from a boat or by placing nets across small subsidiary channels of the main estuary, and collecting the fish trapped during low-tide. Fish were held on ice, where available, but were frozen to -18°C as quickly as possible on arrival to shore.

Fish were subsequently transported to the University of Algarve, during March-April each year from each site, where they were kept frozen at -20°C until dissection, which took place shortly after.

5.2. BRIEF CHARACTERISATION OF SAMPLING LOCATIONS

5.2.1. Aveiro Lagoon System (Aveiro)

The Aveiro lagoon (Figure 6) is of recent geological formation, from the tenth century (Rebelo, 1992), when the coastline defined a long and narrow bay with the common estuary of three rivers (Barrosa, 1980). Data on present topography and physical characteristics of the Aveiro lagoon system are described as follows (Barrosa, 1980), (Rebelo, 1992) and (Moreira *et al.*, 1993): The lagoon with a maximum length of 45 Km and width of about 11 Km, has an artificial narrow entrance, two adjacent long sea walls and a central area of many small islands and channels. The total area of the Aveiro lagoon varies from 43 Km^2 at low tide and 47 Km^2 at high tide, with a water volume of $70x10^6$ m^3 of the lagoon area is very shallow, with a depth of about 1 m, the exception being the navigational channels and near the mouth. The entrance of the lagoon is a man-made permanent connection between the

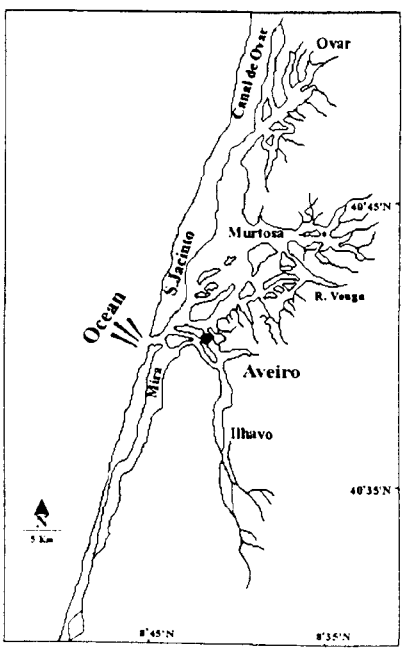


Figure 6. Aveiro estuary.

- Remelha (sampling location).

lagoon and the ocean, allowing the beneficial effects of tides (25 to $90x10^6$ m^3 of seawater, and a freshwater contribution flow of 3 to 60 m^3 s^{-1} dependent on seasonal precipitation and runoff patterns) to be felt within the system, namely drainage and water renovation, with tidal amplitudes between 1 and 3 m. This flow shows the limited freshwater input in relation to the ocean influence. These effects are, however, only significant near the mouth and the central parts of the main channels. There appears to be three published works specifically on the seabass of the Aveiro lagoon (Andrade, 1983; Andrade, 1986; Gordo, 1989). Other works (Arruda *et al.*, 1988; Rebelo, 1992) were on more general aspects of fish populations from that site. The contribution of the marine migratory species in Aveiro lagoon to the total number of species is less than in the Atlantic and Mediterranean ecosystems (Rebelo, 1992); the recent geological formation of the lagoon may have contributed to this observation. Seabass was among the most abundant species by weight at Murteira (Arruda *et al.*, 1988). Most of the individuals caught (76%) were less than 1 year old, 20% were 1 year old and only 4% were 2 years old, aged by otoliths and scales, respectively (Andrade, 1983; Gordo, 1989). This indicates the importance of this system as a nursery area.

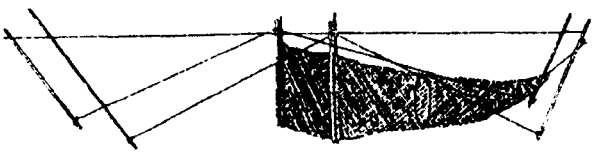


Figure 7. "Botirão". (Arruda *et al.*, 1988)

Samples in Aveiro were taken at Remelha, at the Ilhavo Canal (one of the deepest and more polluted zones of this lagoon, mainly due to urban effluents). The fishery took place, over a period of a week, in February 1993 and February 1994, with a trap called "botirão" (Figure 7), a fishing gear characteristic of the fisherman in Aveiro. It is a conical fish trap of about 15m long, with a mouth of 2 m high and 5 m wide. Within the net there is a "valve" which allows fish to enter the conical end, the cod-end, but prevents them from exiting.

5.2.2. Mondego river estuary (Figueira da Foz)

The Mondego river (Figure 8) has a total length of 234 Km with an extended estuary area of 6644 Km^2. The yearly average water temperature is 13°C, ranging from 8°C to 23.5°C. The estuary is in a region of humid temperate climate, with predominant winds from the North. Major alterations in the system, like the building of dams, the deepening of the river floor, irrigation channels have dramatically changed the prevalent characteristics of the estuary about ten years ago (Jorge, 1991). Among the main changes are the disappearance of cockle (*Cerastoderma edule* L.) banks, the decrease in salinity and the diminishing area of salt marsh.

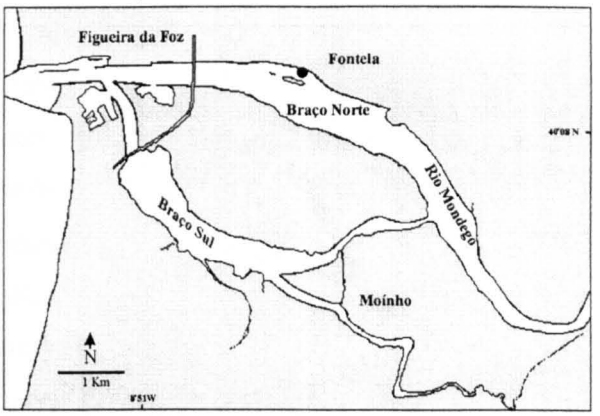


Figure 8. Mondego estuary.
• Sampling location.

Although fishery statistics for within estuary captures do not exist, a seabass hook line fishery is known to be of seasonal importance, during the summer, and all year round illegal fishing with fixed gill net rigged in small tributary channels is also known. According to Jorge (1991) larvae are observed as early as March but the peak is in July when numbers reach their maximum. This coincides with the disappearance of older juveniles year classes, moreover all seabass captured within the estuary, were immature. During December there is a generalised decrease in seabass captures, suggesting their movement to deeper areas.

For the present work, samples were obtained during November 1992, for a number of consecutive days, by a local fisherman, operating alone from a small boat using gill nets. All captured fish were immature.

5.2.3. Óbidos Lagoon (Óbidos)

The Óbidos Lagoon (Figure 9) is located on the Portuguese west coast, 100 Km North of Lisbon. It is a shallow semi-enclosed body of water, oriented northwest-southeast, with its longitudinal axis perpendicular to the coast line. This lagoon system was until the 18th century represented on charts as a navigable estuary, and there are references that the baymouth was formed during the medieval period (Henriques, 1988).

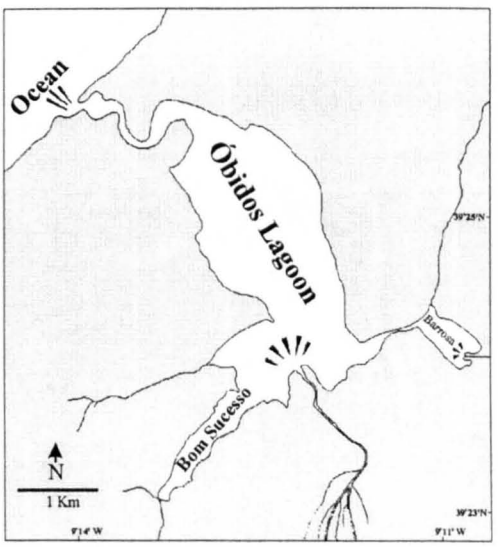


Figure 9. Óbidos estuary.

The lagoon occupies a surface of 6 Km^2, with a depth from 0 and 4 m. In this lagoon the major hydrodynamic agents are the waves and tidal currents, establishing a shallow and narrow entrance subjected to sand obstructions. Physical and chemical data (Quintino and Rodrigues, 1989) shows an increase in water temperatures in summer from 18°C near the entrance to 25°C in the inner basin, and more uniform values of around 13°C in winter.

The temperature difference in summer is mainly due to the shallow depths of the inner basin. Salinities vary between around 25‰ in winter and 36 ‰ in summer. The fauna composition of the lagoon is related to Atlantic and Mediterranean coastal or lagoon systems.

A total of 104 fish were caught during February 1993, in various locations mainly in the centre of the lagoon (the reason why the sampling location is not indicated in the figure), by gill net from a small boat.

5.2.4. River Mira estuary (Vila Nova de Milfontes)

The Mira estuary (Figure 10) is a small system on the south-western Portuguese coast, with a basin of 1576 km^2. Tidal range is from 1.1 and 3.4 m. The effective runoff is calculated to be $2x10^8$ m^3 because of the existence of an upstream reservoir. Limited and strongly seasonal freshwater flow and circulation is dominated by bi-directional tidal currents (Andrade, 1986) allowing it to be viewed as a tidal lagoon with a limited freshwater flow.

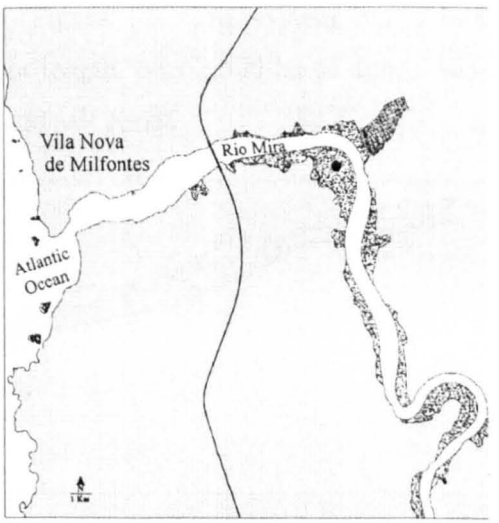


Figure 10. Milfontes estuary.
• Sampling location.

The fish, 213 in total, all juveniles, were caught with a single trap over a day in February 1993 and the same in 1994. The trap, the "tapa-esteiros" (Figure 11), is a fishing gear traditionally used in places where small subsidiary channels of the main estuary remain with water in low tide. It is a fish trap that takes advantage of the retreating water movements as low tide approaches and prevents the fish from leaving the channels. Within the net there is a "valve" which allows fish to enter the conical end, the cod-end, but prevents them from exiting.

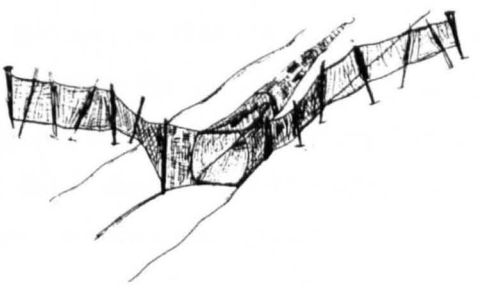


Figure 11. "Tapa-esteiros"

5.2.5. Ria Formosa Lagoon System (Faro)

The Ria Formosa Lagoon is located in an interface region between the Mediterranean, North Africa and the Atlantic (Figure 12). The system has a total surface, during high tide, of 11.800 ha, 55 Km in length, plus 2.500 ha of dunes, sand and salt marshes and 2.000 ha of fish reservoirs and salt ponds.

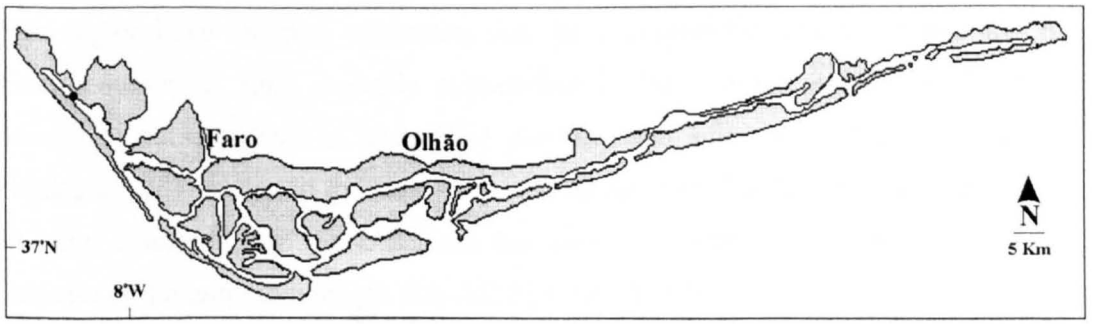


Figure 12. Ria Formosa lagoon system.

• Sampling location

The absence of big fresh water inflows and the strong renewal of sea water allows a great stability of salinity values. Its variations, either particular or seasonal, are mainly due to winter rains and evaporation. Water temperature varies between 10°C and 28°C, which in fact constitutes particularly favourable climate conditions. According to local fishermen, autumn and winter are seasons of great activity for seabass. This is possibly not only because there are good feeding conditions but also because many fish get together for spawning in various places along the Portuguese coast. Seabass prefer shallow waters, sandy or rocky bottoms. From this sampling site only 65 juvenile fish were caught, all by "tapa-esteiros". In Faro, former salt ponds were converted to fish ponds for extensive aquaculture. Local people take advantage of the fact that lagoon systems are good nursery areas, to have fish growing in captivity. Recruitment in these facilities relies on wild fish allowed entering the ponds when the tide level and the season is appropriate by the lifting of the pond's monk trap. Fish are usually captured by emptying the ponds, during low tide, particularly around Easter time where the prices, specially for exports are higher. The species that are the most abundant in these ponds are: seabream (*Sparus aurata*), seabass (*D. labrax*), mullets (*Mugil* spp.) although

several other species can also be found in much lower numbers, such as the red mullet (*Mullus surmuletus*). Because the sample taken in 1992 in Faro was small, a second sample was needed, so on two different occasions in 1993 and 1994, I was present during the emptying of ponds. At those times not a single seabass was caught.

The fact that recruitment to these ponds relies on wild fish, and that very few seabass are being caught in this way, similarly fishermen catches from the wild in the Faro region have dropped, suggesting that the recruitment of seabass in this area has been jeopardised. One probable explanation is the overexploitation in the easily accessible nursery areas of undersized juveniles and adults by fishermen, that were finding a ready market in the area. This unsustainable exploitation has also taken place in other countries, e.g. the U.K., and has lead to an regulatory measures such as a increase in the minimum length size (MLS) from 32 to 36 cm in 1991.

5.3. TISSUE STORAGE AND DATA COLLECTION

Dissection of individuals took place at the University of Algarve, between 1 to 6 months after the fish were caught. From each individual a few grams of tissue from muscle, from above the lateral line, liver, heart and eye were taken. Records of total length, total weight and sex were kept for each individual. Otoliths as well as scales were collected and kept dry in paper envelopes. Protein structure is quite labile, producing poor electrophoretic results if excessive denaturation of proteins is allowed. For that reason, tissue samples were kept in Eppendorf tubes and held in liquid nitrogen waiting for transport to Scotland, where they were stored in -70°C freezer.

5.4. AGE DETERMINATION

Most species of fish which live in waters with fluctuating temperature regimes have distinct seasonal patterns of growth (Pawson and Pickett, 1987). This results in the appearance of ring like patterns, in certain anatomical structures, which can be used to assess age, if a particular pattern can be identified with a season. Scales and otoliths are particularly well suited to this purpose. Warm water conditions either natural or artificial such as cooling water outflows from power stations, will produce enhanced

growth for that year. In bass, although scales are easier to age read than otoliths (Kelley, 1988a), it is still a laborious task compared to length determination. In fact, length can also serve as a measure of age: in a population with several age classes, the distribution of a morphometric character such as length can be considered also as a compound of age distributions and these are analogues. The identification of each distribution is dependent on the knowledge of each distribution's parameter and their relative proportions. In this work, length is used as a measure of age, and the resolution of the modal components of the length distributions was determined by ANAMOD (Nogueira, 1992). This software is based on an improvement of the probability paper method (Harding, 1949). The modal classes identified were then checked for age using otoliths and scales from individuals near the tails of the length distributions. In this way, the time consumed in determining each individual's age by reading otoliths or scales is greatly reduced, because only a few selected individuals in the tails or area of overlap between modes need to be assessed.

6. DATA ANALYSIS

How are sets of genotypes and allele frequency data used in studies of population or stock identification? The analysis includes genetic characterisation of the samples, assessment of Hardy-Weinberg deviations, determination of heterogeneity of allele and genotype frequencies between samples, and ideally, an estimation of genetic structuring, gene flow, similarities and distances indexes. Only in some instance, when the geographical range is appropriate and results justify, clinal variation of allelic frequencies is presented. This section contains a description of the procedures used in this work together with references, whenever appropriate, to other relevant procedures that were and still are used.

6.1. GENETIC CHARACTERISATION OF SAMPLES

6.1.1. Individual genotypes for each locus in each population

Genotypic frequencies indicate solely the organisation of alleles in genotypes. Considering a population of N individuals, with 2 alleles at one locus, A and a, the possible genotypes are AA, Aa and aa, and $N=N_{AA}+N_{Aa}+N_{aa}$. Genotype frequencies will be as follows:

$$p_{AA} = N_{(AA)}/N\,;\; p_{aa} = N_{(aa)}/N\,;\; p_{Aa} = N_{(Aa)}/N$$

6.1.2. Determination of allelic frequencies

Allelic frequencies will give us the proportion of each allele taking into account all the alleles of the same gene. Frequencies for alleles A and a are determined by:

$$p_A = \left(N_{(AA)} + 1\,2N_{(Aa)}\right)/N \quad \Rightarrow \quad p_A = p_{AA} + 1/2\,p_{Aa}$$

$$p_a = \left(N_{(aa)} + 1\,2N_{(Aa)}\right)/N \quad \Rightarrow \quad p_a = p_{aa} + 1/2\,p_{Aa}$$

6.1.3. Polymorphic loci

The genetic polymorphism can be defined as the occurrence, in the same population, of two or more alleles at the same locus, both with an appreciable frequency (Cavalli-Sforza and Bodmer, 1971). Today we use two levels of allele frequency to consider a locus polymorphic: when the most frequent allele has a frequency ≤99% ($P_{0.99}$) or ≤95% ($P_{0.95}$). Polymorphism corresponds then to the proportion of polymorphic loci: $P = x/l$, where x is the number of polymorphic loci and l the total number of loci studied, and is commonly expressed as percentage.

6.1.4. Heterozygosity

The frequency of heterozygotes is important, since each heterozygote carries different alleles and represents the existence of variation. Heterozygosity (H_{obs}) is defined as the proportion of observed heterozygotes (h) at a given locus: $H_{obs} = h/N$, where N is the total individuals scored for that locus, and the average heterozygosity, $\overline{H}_{obs}$ is the averaged sum of the observed heterozygosities over all loci. The value of the theoretical expected heterozygosity under Hardy-Weinberg proportions, can be estimated as: $H_{exp} = 1 - \sum_{i=1}^{k} p_i^2$, with p_i, being the frequency of allele i and k the number of alleles at a locus (Nei, 1975). Similarly, the average expected heterozygosity ($\overline{H}_{exp}$) is the averaged sum of the heterozygosities over all loci, thus: $\overline{H}exp = \left(1 - \sum_{i=1}^{k} p_i^2\right)/l$. As the allelic frequencies are strongly dependent on sample size (Shriver *et al.*, 1995) it is highly desirable that an unbiased expected heterozygosity can be estimated. In that way the number of individuals sampled are taken into account giving an unbiased measure of genetic variability. The gene diversity or unbiased expected heterozygosity (H_{unb}) is estimated by: $H_{unb} = \frac{N}{N-1}\left(1 - \sum_{i=1}^{k} p_i^2\right)$, where N is the number of individuals sampled (Nei, 1978).

All these are available from software packages such as Biosys (Swofford and Selander, 1989), Genetix (Belkhir *et al.*, 1996 - 1998), Popgene (Yeh *et al.*, 1997) and others.

6.1.5. Testing Hardy-Weinberg equilibrium

If we consider a large random mating population with no selection, mutation or migration, no matter the values of the initial genotype frequencies of one such population, gene transmission by itself does not modify those frequencies. This rule is the fundamental theorem of Hardy-Weinberg's Law (Hardy, 1908; Weinberg, 1908). This theorem considers that for diploid populations, the genotypic frequencies of a generation will reach an equilibrium in one generation time and can be represented by a binomial (for two alleles) or multinomial (for more than two alleles) expression of the allelic frequencies:

binomial - $(p_i + p_j)^2 = p_i^2 + p_j^2 + 2 \times p_i p_j$

multinomial – $(p_i + p_j + p_k + ...)^2 = {p_i}^2 + {p_j}^2 + {p_k}^2 + 2 \times p_i p_j + 2 \times p_i p_k + 2 \times p_j p_k + \cdots$

The Hardy-Weinberg equilibrium is a fundamental model in population genetics, allowing comparisons of observed and expected allele frequencies and thus to find out whether the preconditions necessary for the equilibrium are violated. So the question here is how well genotype frequencies at every locus studied in each population approach the expected frequencies, as for instance, heterozygote deficiency can be generated by regular systems of inbreeding or by population structuring.

Among the most popular Hardy-Weinberg equilibrium tests is the Workman and Niswander test (Workman and Niswander, 1970), which is done by $\chi 2$ goodness-of-fit to determine if F_{IS} (that measures Hardy-Weinberg departure within subpopulations) is significantly different from zero. This parameter measures the reduction in heterozygosity due to non-random mating within the subpopulation, and thus helps to detect departures from Hardy-Weinberg by measuring the amount of heterozygote deficiency or excess observed in the sample. Values of F_{IS}, in most natural populations, are typically close to zero, which indicates random mating within subpopulations. A

positive F_{IS} value indicates a deficit of heterozygotes and therefore some within-population structuring. To test this, one calculates $\chi 2$ according to $\chi^2 = N(F_{IS})^2(k-1)$, where N is the total number of individuals sampled, k is the number of alleles at the locus, with degrees of freedom $(d.f.) = \frac{[k(k-1)]}{2}$. This computation, however, has a serious pitfall, since it is seriously affected by small observed and expected numbers (Lessios, 1992), such as the ones seen in genetic data sets, where even if the sample sizes are relatively big, the observed and expected numbers of some alleles or genotypes are small. This fact induced researchers to pool genotype classes, to overcome that problem, or to introduce numerical corrections (Cannings and Edwards, 1969).

With the intention to overcome this disadvantage in the $\chi 2$ test, several authors explored the possibilities of the Fisher's exact test (Fisher, 1935). Exact tests are generally used for small sample sizes, when there is the greatest chance of having small expected numbers in the $\chi 2$ test formulation (Weir, 1996). In fact, this test was originally proposed to deal with a table of 2 x 2, and it was afterwards expanded to tables with more than two rows and/or columns (Zar, 1984). Nevertheless, these could not deal with common population genetics data sets, that typically have several populations and many loci. This is because the number of cases grows exponentially with the number of alleles and populations and computation becomes quite slow. In cases were there are rare alleles at a locus (such as the case of microsatellite loci), expected numbers can be small even in moderately large samples, and exact tests are most desirable in such circumstances.

Testing Hardy-Weinberg equilibrium should focus on a specific biological and/or genetic hypothesis (Lessios, 1992). These specific hypotheses require particular statistical procedures to perform high power tests. Bootstrap tests allow the testing of a broad range of hypothesis, but it is of limited use for smaller samples (less than 20 individuals) with low genetic variation (Van Dongen and Backeljau, 1997). These authors recognise that bootstrapping cannot replace exact tests when there is a combination of negative and positive F_{IS} values, which may result in undetected Hardy-Weinberg deviations by the bootstrap method.

The other one-tailed tests were compared and two tests were found to perform well under different hypotheses (Rousset and Raymond, 1995):

to test a specific hypothesis - a score test related to Robertson and Hill's estimator of the inbreeding coefficient (*U*-test) is close to optimal for the detection of heterozygote deficiency or excess. A PC-based software GENEPOP (version 3.1b) (Raymond and Rousset, 1995b) performs the score *U*-test on one or several samples for the alternative hypothesis $U = \sum_{i=1}^{k} \frac{n_{ii}}{p_i} - N$, where k are the number of alleles at a locus, n_{ii} are the genotype (A_i A_i) numbers, p_i observed allele frequencies of allele i, and N sample size.

to test the null hypothesis - Fisher's exact test or probability test (Raymond and Rousset, 1995a; Rousset and Raymond, 1995). This test as stated before is not adversely affected by small expected values and may be the only one valid when sample sizes are small and or alleles are rare (Lessios, 1992). The probability test calculates the exact probability of the observed sample to be drawn from the population by chance if the null hypothesis held true. Then all the possible tables with the same margin totals are tabulated and only the ones with the same or smaller probabilities are considered. If the sum of these probabilities is less than or equal to the significance level, α, then the null hypothesis, i.e., independence of allelic distribution across populations, is rejected and there is evidence of population differentiation (for an example see Zar, 1984). This test is also available in GENEPOP (Raymond and Rousset, 1995b) (GENEPOP, option 1, sub-option 3) implemented as an analogous test to Fisher's exact test on a two-by-two contingency table, but extended to a contingency table of R x C size (Guo and Thompson, 1992). The test is performed using a modified version of the Markov-chain random walk algorithm described Guo and Thomson (1992), that is more efficient from a computational point of view (Raymond and Rousset, 1995b; Schneider *et al.*, 1997). This was the procedure used in the present work to detect significant departure from Hardy-Weinberg equilibrium.

Schneider *et al.* (1997), describe in detail the procedure followed in the present work, where a contingency table is first built, with observed and expected genotype

frequencies as entries. The contingency table for a general case is represented in Table 12, and the probability of the table is presented in equation:

$$\Pr = \frac{N!\prod_{i=1}^{k} n_{i*}!}{(2N)!\prod_{i=1}^{k}\prod_{j=1}^{i} N_{ij}!}\,2^{H}$$, where N denotes the number of individuals; k, the number of alleles; n_{i*}, the number of i alleles; N_{ij}, the counts of the genotypes A_iA_j and H is the number of heterozygote individuals.

Table 12. Contingency table for exact test for a diallelic situation (3x2).

Genotypes		Observed	Expected	Genotype counts
		C_1	C_2	
R_1	*ii*	N_{ii}	$\left[\left(N_{ii}+\frac{1}{2}N_{ij}\right)/N\right]^2 \times N$	$R_1C_1+R_1C_2$
R_2	*ij*	N_{ij}	$2\times\left[\left(N_{jj}+\frac{1}{2}N_{ij}\right)/N\right]^2 \times \left[\left(N_{ii}+\frac{1}{2}N_{ij}\right)/N\right]^2 \times N$	$R_2C_1+R_2C_2$
R_3	*jj*	N_{jj}	$\left[\left(N_{jj}+\frac{1}{2}N_{ij}\right)/N\right]^2 \times N$	$R_3C_1+R_3C_3$
Totals		N	N	2N

Instead of enumerating all possible contingency tables, a Markov chain is used to efficiently explore the space of all possible tables. This Markov chain consists in a random walk in the space of all contingency tables produced. The probability of obtaining a particular table corresponds to its actual probability under the null hypothesis of Hardy-Weinberg equilibrium. A table is modified by selecting the table's two distinct rows *i1*, *i2* and two distinct columns *j1*, *j2* at random; the new table is obtained by decreasing the counts of the cells (*i1*, *j1*) (*i2*, *j2*) and increasing the counts of the cells (*i1*, *j2*) (*i2*, *j1*) by one unit, leaving the marginal genotype counts

unchanged. This procedure is repeated a number of times to obtain all possible contingency tables having identical marginal counts. In order to start from a random initial position in the Markov chain, the chain is explored for a pre-defined number of steps (the dememorization phase) before the probabilities of the switched tables are compared to that of the initial table. The number of dememorization steps should be big enough to allow the Markov chain to "forget" its initial state, and make it independent from its starting point. The number of dememorization steps performed for all the tests was 10000. The *P*-value of the test is then taken as the proportion of the visited tables having a probability smaller or equal to the observed contingency table.

A standard error on *P* is estimated by subdividing the total amount of required steps into B batches (Guo and Thompson, 1992). A P-value is calculated separately for each batch. Let us denote it by P_i (i=1,...,B). The estimated standard error (s.d.) is then calculated as: $s.d.(P) = \sqrt{\frac{\sum_{i=1}^{B}(P - P_i)^2}{B(B-1)}}$. In the present work a 100 batches and 1000 iterations per batch were used.

6.2. POPULATION DIFFERENTIATION

The two main points of concern when studying population genetic differentiation: are *how to estimate the differentiation* and *how to test those estimates* in order to obtain statistical significance.

The parameters most widely used for population differentiation are Wright's *F* statistics (Wright, 1951), F_{IS} , F_{ST} and F_{IT}, that describe the arrangement of genetic variation in a subdivided infinite size population. These three coefficients allocate the genetic variability to the total population level (T), subdivisions (S) and individuals (I), and are calculated as follows:

$F_{IS} = \frac{H_s - H_I}{H_s}$ - measures the reduction in heterozygosity of an individual due to non-random mating within its subpopulation, i.e., the proportion of total genetic variability due to differences within populations;

$F_{ST} = \frac{H_T - H_s}{H_T}$ - measures the reduction in heterozygosity of a subpopulation due to random genetic drift, i.e., the proportion of total genetic variability due to differences between populations;

$F_{IT} = \frac{H_T - H_I}{H_T}$ - measures the reduction in heterozygosity of an individual relative to the total population, i.e., the proportion of total genetic variability due to differences across populations;

with H_S being the expected heterozygosity under Hardy-Weinberg equilibrium; H_I, being the observed heterozygosity in subpopulations and H_T being the observed heterozygosity in the whole population.

The hierarchical *F*-statistics are types of inbreeding coefficients, but they differ in the reference populations. F_{IS} is concerned with inbreeding in individuals (I) relative to the subpopulation to which they belong to (S), F_{ST} with inbreeding in subpopulations (S), relative to the total population (T) and F_{IT} with inbreeding in individuals (I), relative to the total population (T). The three indexes are interrelated so that: $F_{ST} = \frac{F_{IT} - F_{IS}}{1 - F_{IS}}$. The proportion of the total genetic variation is explained by differences within populations and between populations. When dealing with real populations, however, it is the Weir and Cockerham (1984) approach that should be used, because these authors propose non biased estimates of the original Wright's statistics: f for F_{IS}, θ for F_{ST}, and F for F_{IT} (Table 13). These estimates are not affected by numbers of alleles per locus, population sample size and number of populations observed (Weir and Cockerham, 1984). θ is calculated according to the following equation:

$\theta = \frac{1}{k}\sum_{u=1}^{k}\frac{a_u}{\sum_{u=1}^{k}(a_u + b_u + c_u)}$, where k is the number of alleles at the locus, and a_u, b_u and c_u are the among samples, among individuals within samples and within individual estimates of components of variance, respectively, of a nested analysis of variance on allele frequencies (Weir and Cockerham, 1984).

Table 13. Relationship between and Weir and Cockerham (1984) estimators.

Wright (1951)	Weir & Cockerham (1984	Correlation between	Interval
F_{IT}	F = 1- C/(A+B+C)	alleles of individuals of all populations	[-1,+1]
F_{IS}	f = 1- C/(B+C)	alleles of individuals within one population	[-1,+1]
F_{ST}	θ = A/(A+B+C)	alleles of individuals of one population relative to all other populations	[0,+1]

The meaning of these parameters for allele *i*, is: A*i*, variance component of allele frequency between populations; B*i*, variance component of allele frequency between individuals within each population; C*i*, variance component of allele frequency between gametes within each individual. These calculations are detailed in Weir and Cockerham (1984).

Nevertheless, and as in most papers today, it will be Wright's original designations that we will use in this work. F_{ST} values help us to understand the degree of population differentiation among samples.

Several procedures are available to test for population differentiation, among them various exact tests (see Rousset and Raymond, 1997) which are preferable to traditionally asymptotic tests, for reasons already presented in the previous section. Exact tests are based on the hypothesis of random sampling of genes, which for diploid organisms may or may not be valid. In the present work we opted for the use of an allelic goodness-of-fit test, for two main reasons: the sampling of genotypes needs not to be independent, and because tests acknowledging the identity of alleles are more powerful than test ignoring it, as allelic statistics take into account whether 2 genotypes share an allele or not, and genotype statistics only consider whether genotypes are identical or not (Goudet *et al.*, 1996).

The hypothesis to test was that of a random distribution of k different alleles among r populations (Raymond and Rousset, 1995a; Goudet *et al.*, 1996; Rousset and Raymond, 1997). The test performed is analogous to Fisher's exact test on a 2 x 2 contingency table extended to a r x k contingency table (Table 14). All potential states

of the contingency table are explored with a Markov chain similar to that described for the case of Hardy-Weinberg test. During this random walk between the states of the Markov chain, we estimate the probability of observing a table less or equally likely than the observed sample configuration under the null hypothesis of panmixia. If the sum of these probabilities is less than or equal to α (the significance level), then the null hypothesis, i.e., independence of allelic distribution across populations, is rejected and there is evidence of population differentiation. As previously described, an estimation of the error on the P-value was done by partitioning the total number of steps into a given number of batches.

For each pairwise comparison between samples, a combination of the probabilities over all loci was estimated using Fisher's method (Fisher, 1970).

Table 14. Contingency allelic table for a 3 sample, 3 allele situation.

Population/ alleles	***1***	***j***	***k***	**Totals**
1	n_{11}	n_{1j}	n_{1k}	$R_1 = n_{11}+n_{1j}+ n_{1k}$
i	n_{i1}	n_{ij}	n_{ik}	$R_i = n_{i1}+n_{ij}+ n_{ik}$
r	n_{r1}	n_{rj}	n_{rk}	$R_r = n_{r1}+n_{rj}+ n_{rk}$
Totals	C_1	C_j	C_k	$N = \sum_{i=1}^{r}\sum_{j=1}^{k} R_i C_j$

Each row represents a sample and each column an allele (adapted from Goudet *et al.*, 1996).

$$\Pr = \frac{\prod_{i=1}^{r} n_{i.}! \prod_{j=1}^{k} n_{.j}!}{N! \prod_{i=1}^{r}\prod_{j=1}^{k} n_{ij}!} = \frac{n_{1.}!...n_{r.}!n_{.1}!...n_{.k}!}{N!n_{ik}!...n_{1k}!...n_{rj}!n_{rk}!},$$ (Goudet *et al.*, 1996; Rousset and Raymond, 1997)) with the following notation: n_{ij} is the number of individuals with allele j in sample i; n_i is the sample size of population i; n_j is the total number of individuals with allele j, N is the total number of individuals sampled, r the number of samples, k the number of alleles.

6.3. CORRECTION OF SIGNIFICANCE LEVELS

Tables of statistical tests are commonly analysed in population genetics studies. The marking of component tests as statistically significant, based on their single-test significance values may be inappropriate, since chance alone will produce some significant tests (see Rice, 1989 for simulations). In fact, one should define a family of tests, for instance, to test Hardy-Weinberg proportions among 8 loci from 16 samples. In order to maintain the table significance for each single locus test (0.05), each individual test needs to be more stringent, being adjusted accordingly to the number of single tests the table is composed of (16). The sequential Bonferroni technique was subsequently introduced (Holm, 1979) to increase the power of the individual tests, that consists of: selecting a significance level (α); ranking the P-values from the smallest (P_1) to the largest (P_t); comparing P_1 with α /t, t being the number of tests. If P_1 is smaller or equal to α/t, then the test indicates significance at the "table-wide" α level, if not the test indicates non-significance. If and only if $P_1 \leq \alpha/t$, the next test can be made, comparing P_2 with $\alpha/(t\text{-}1)$. So, for the example given above, the initial α=0.05/16=0.003. If one of the 16 tests yields a $P \leq 0.003$, and all the others have a $P>0.003$, then only that test is statistically significant.

6.4. ISOLATION BY DISTANCE

Mantel's test (Mantel, 1967) was used to determine if there was a correlation between two semi-matrices, one of genetic distance and another one of correspondent geographical distances (GENEPOP option 6, sub-option 5) (Raymond and Rousset, 1995b and references therein). The principle of the Mantel test is to permute lines or columns of the two semi-matrices to obtain the test statistics distribution under the null hypothesis i.e., independence between the two variables.

The geographic distances between localities were estimated approximately by measuring the coast line with a curvemeter on a 1:1 000 000 map between river mouths. The estimates of genetic distances used for this procedure were the F_{ST} pairwise values.

6.5. GENETIC DISTANCE

When genetic data are available from several populations, it is natural that comparisons are made in order to see how similar or different these populations are. Several measures of genetic distance are available (Cavalli-Sforza and Edwards, 1967; Nei, 1972; Nei, 1978; Reynolds *et al.*, 1983). Although the genetic distance proposed by Nei (1972) is one of the most commonly used, for intraspecific divergence the estimation according to Reynolds' (1983) (D=-ln (1-θ)) is more appropriate for short-term evolution, when the divergence between populations with a common ancestral population may be regarded as being solely due to drift. This method generates a genetic distance coefficient that reflects allele frequency differences. Thus, in the present work, the genetic relationships between all pairs of populations were estimated using both Nei's distance, to allow comparisons with other published works, and Reynolds *et al.* (1983) distance.

Again, several tree construction methods are available, based on neighbor-joining or on cluster analysis (Sneath and Sokal, 1973) like the WPGMA (weighted pair group method using arithmetic averages) and the UPGMA (unweighted pair group method using arithmetic averages). The relationships among samples inferred from Reynolds *et al.* (1983) were summarised with a neighbour-joining tree (Saitou and Nei, 1987). Neighbor-joining (Saitou and Nei, 1987) is conceptually related to cluster analysis, but does not assume that the evolutionary rate is the same in all lineages. In the neighbor-joining method, a modified distance matrix is constructed in which the separation between each pair of nodes is adjusted on the basis of their average divergence from all other nodes, which has the effect of normalizing the divergence of each sequence for its average clock rate (Swofford *et al.*, 1996). Then, the algorithm combines the least distant pair of sequences as defined by this modified matrix (and thus not as defined by the original matrix) into a new one. This new OTU (Operational Taxonomic Unit) is added to the tree while the replaced OTUs and their respective branches are removed from the tree. This process converts the newly added OTU into a terminal node on a tree of reduced size. At each stage in the process, two terminal OTUs are replaced by one new. The process is complete when only two OTUs remain, separated by a single

branch. A worked example can be found in Swofford et al. (1996). Besides the fact that the neighbor-joining method is very effective in recovering the correct tree topology (Saitou and Nei, 1987), its main advantage is its speed.

For tree reliability, allele frequencies were bootstrap 1000 times using SEQBOOT program of the PHYLIP computer package (Felsenstein, 1993). Bootstrap is a general re-sampling method to estimate the distribution of statistics based on independent, identically distributed observations. The resultant bootstrapped frequencies were used to calculate Nei's (Nei, 1972) and Reynolds' (Reynolds *et al.*, 1983) distances. A neighbor-joining tree was then constructed by TREECON (version 1.3b) (Van de Peer and De Wachter, 1994) to depict relationships among populations.

6.6. ESTIMATION OF GENE FLOW

Gene flow can be defined as the transfer of genetic material between populations. This can be achieved through passive movements of zygotes or the displacement of individuals at any stage of their life cycle. Gene flow has been quantified through the estimation of the migration rate (m) which in turn is defined as the proportion of migrants in a population.

The island model (Wright, 1943) was proposed to describe a specific situation: in each generation any gene samples from a population has a probability m of having migrated from any one of an infinite number of other populations. When these various "islands" are of finite size N_e (the number of breeding individuals) an equilibrium is established between loss of variation due to drift within islands, and gain of variation by migration from other islands. The equilibrium is $F_{ST} = \frac{1}{1+4N_e m}$. A remarkable feature of this equation is that the F_{ST} value is independent of the mutation rate and the number of alleles (Hartl and Clark, 1989). The drawback of using F_{ST} or related quantities to estimate migration rates is, of course, that it is assumed that population differentiation is due to gene flow. In the absence of direct observation on gene flow, this may be the only alternative but it does mean that alternative evolutionary scenarios leading to the same pattern of gene frequencies cannot be eliminated (Weir, 1996).

Another model is the "stepping-stone" model (Kimura and Weiss, 1964) in which populations are organised in discrete colonies or demes with migration between them, the fundamental difference from Wright's isolation by distance model is that is assumed that only migration into adjacent demes is allowed.

Most of the estimates of gene flow are based on N_em, the absolute number of individuals exchanged between populations. Under neutrality and equilibrium, divergence between populations is due to the number of migrants. N_em can be estimated from F_{ST} (WRIGHT, 1969) and derived formulae and from private alleles (Slatkin, 1985).

It is important to note that these estimates are not precise, but nevertheless they give us a hint of the order of magnitude of exchanges between populations.

The effective number of migrants per generation (N_em) was estimated as: ($N_em = \frac{1}{4}\left(\frac{1}{\overline{F_{ST}}} - 1\right)$ (Wright, 1951) and also by Takahata's modification (Takahata, 1983): $N_em = \frac{1}{[4\alpha]} \times \left[\frac{1}{4}\left(\frac{1}{\overline{F_{ST}}} - 1\right)\right]$, where $\alpha = [n/(n-1)]^2$ and n is the number of populations sampled and $\overline{F_{ST}}$ is the F_{ST} mean value.

6.7. ESTIMATION OF EFFECTIVE SIZES VS. MUTATION MODELS

The concept of effective size (N_e) refers to the size of an ideal population (Hartl and Clark, 1989) that has identical genetic characteristics, such as homozygosity, allelic frequencies and allelic variances of the observed population. N_e is usually much smaller than the total population, and that has to do with different sex ratio, demographic and fertility fluctuations.

An important aspect of the estimation of effective size, is that this calculation is not independent of the mutation model that best fits the mutation pattern of the locus. This poses more problems in the works concerning microsatellite loci, as the mutation process is not clearly elucidated. It was suggested (Shriver *et al.*, 1995) that both

infinite-allele model (IAM, Kimura and Crow, 1964) and the stepwise model (SMM, Kimura and Ohta, 1978) can be used to describe microsatellite variability. According to the IAM, every mutation is to a new allelic type different from those previously present in the species. The SMM assumes that, mutation occurs through the gain or loss of a single repeat unit, meaning that some mutations will actually generate alleles already present in the population. Both models usually give similar results when heterozygosity is below 0.5, but above that their results may differ, in which case both models should be used and compared. Actually, in fish heterozygosity levels in microsatellites can be quite high, well above 0.5 and mutation rates and mechanisms are poorly known. These facts favour the simultaneous use of both models (Table 15) more so when it is suggested that the mode of mutation at microsatellite loci will likely be case/population and locus specific (O'Connell and Wright, 1997).

According to some authors (Weir, 1996), in the IAM an equilibrium can be reached between the loss of variation by drift and the introduction of variation by mutation in a finite population of size N. In such a situation, heterozygosity can be expressed as $H_{exp} = 4N_e\mu/(4N_e\mu)+1$ (Crow, 1986), and in turn N_e is: $Ne = \left[H_{exp}\ \left(1 - H_{exp}\right)\right]/4\mu$, H_{exp} being the expected heterozygosity, μ the mutation rate and N_e the effective size. Ewens formula (Ewens, 1972) (formula 11 and associated table 1): $E(k) = \sum_{i=1}^{2n-1} \frac{\theta}{(\theta + i)}$, gives the expected number of alleles (k) as a function of the value $\theta = 4N\mu$, with n the number of individuals, N the effective size and μ the mutation rate. The calculations are somewhat complex because it is not possible to have θ as a function of n and k. A small program initially written by Laurent Excoffier and modified by Lounès Chikhi was used and gives an accurate value of θ, from which it is possible to estimate the effective size as proposed.

Table 15. Estimation of effective sizes according to different mutation models.

Mutation model	Based on	Reference	Formula
Infinite allele	Sample size, allele number	Ewens (1972)	$E(k) = \sum_{i=1}^{2n-1} \frac{\theta}{(\theta + i)}$
Infinite allele	Mean expected heterozygosity	Crow (1986)	$Ne = [H_{exp} \ (1 - H_{exp})] / 4\mu$
Stepwise	Mean expected heterozygosity	Ohta and Kimura (1973)	$Ne = [(1\ 1 - H_{exp})^2] / 8\mu$

In the alternative mutation model SMM, the alleles in the same mobility class would not necessarily be identical by descent only identical in state (Ohta and Kimura, 1973). Ohta and Kimura (1973) followed Bulmer's observation (Bulmer, 1971) that at many loci there is one common allele with intermediate mobility and several less common alleles roughly symmetrically distributed in mobility about the common allele. Although the stepwise mutation model is apparently not applicable to electrophoretically distinguishable alleles (adjacent alleles in general do not differ by a single charge state), it might well be applicable to alleles at microsatellite loci. There is some evidence from pedigree data for a stepwise pattern of mutation (Weber and Wong, 1993), and also with observations from regularities in the distributions of the frequencies of alleles, with one common allele with intermediate mobility and several less common alleles roughly symmetrically distributed in mobility about the common allele. We assume that the alleles at one locus are selectively equivalent and that there is random mating in the population. Ohta and Kimura (1973) proposed that in equilibrium, the population expected heterozygosity can be estimated by: $H_{\exp} = 1 - \left(1/(1 + 8N_e\mu)^{0.5}\right.$, hence Ne becomes: $Ne = [(1\ 1 - H_{exp})^2] / 8\mu$. As in the infinite allele model it is considered that there is an equilibrium between drift and mutation.

Experimental microsatellite mutation rates observed range between 10^{-5} to 10^{-3} (Dallas, 1992; Ellegren, 1995) (see also Bruford and Wayne, 1993; reviewed in Weber and Wong, 1993) for various species, although there is no data for fishes. We will use 10^{-4}, used previously in a European seabass paper (García de León *et al.*, 1997), that will allow results to be compared.

6.8. EWENS-WATTERSON NEUTRALITY TEST

Based on the infinite allele model (IAM) and on Ewens (Ewens, 1972) sampling theory, the Ewens-Watterson neutrality test, performs a test of selective neutrality in a population at equilibrium. The testing procedures, computed by POPGENE (Yeh *et al.*, 1997), requires the generation of random neutral samples, which are generated using the algorithm by Manly (1985). The significance of the probability of the observed sample is found by comparing it to random neutral samples with the same number of alleles of identical sizes. The probability of the sample's selective neutrality is obtained as the proportion of random samples which are less probable than the observed sample (Slatkin, 1994; Slatkin, 1996).

6.9. ESTIMATION OF NULL ALLELES' FREQUENCIES

Heterozygote deficiencies relative to Hardy-Weinberg expectations are sometimes observed at microsatellite loci. Such deficiencies can be due to a population subdivision effect (Wahlund, 1928), inbreeding due to consanguineous mating or to the presence of null alleles. In microsatellites, null alleles can arise when mutations within the flanking regions of the microsatellite, prevent primers from binding (Callen *et al.*, 1993), impede amplification and result in underestimation of heterozygosity. In this case, genotypes that appear to be homozygous are in fact heterozygous. A simple method for estimating the frequency (r) of a null allele was developed (Brookfield, 1996) : $r = \frac{H_{\exp} - H_{obs}}{1 + H_{\exp}}$.

PART 3 - ALLOZYME VARIATION

7. INTRODUCTION

The analysis of morphological variation during the first half of this century gave way to the study of genetic variation at the molecular level in the 1960s. This was a major step in the evaluation of variability in natural populations, which were thought to be quite homogenous. This has opened a new approach to understand population biology and the evolution of organisms. Since this novel technique is simple for population studies, it permitted a wealth of Mendelian data to be investigated in a short period of time.

Protein electrophoresis is the migration of proteins under the influence of an electrical field. The rate at which a particle moves in the medium is directly related to its net charge and inversely related to its size. Protein molecules usually have a net charge and are thus amenable to separation by electrophoresis. Differences in protein migration are used to get insights into the amount of genetic variation that exists within populations and the amount of differentiation between populations or between species.

It has been generally accepted that most protein polymorphisms (allozymes) are selectively neutral (Nei, 1978; Kimura, 1983; Utter, 1991), or nearly neutral (Ohta, 1992). However, there are known exceptions, namely in the *Mytilus* complex (Hilbish and Koehn, 1985), in *Fundulus heteroclitus* (Powers, 1990) and more recently in *Neocyttus rhomboidalis* (Elliot *et al.*, 1998). Also in the European seabass evidence of selection was produced (Allegrucci *et al.*, 1994; Allegrucci *et al.*, 1997) considering five loci, *CK-4**; *EST-2**, *G6PDH**, *MEP** and *NP** to be "under selection". Nevertheless, the majority of polymorphisms remain untested, so it is legitimate, as a starting working assumption, to consider the neutrality of alleles (Allendorf and Phelps, 1981).

Under the neutrality hypothesis, allele and genotype frequencies respond primarily to mutation, migration and genetic drift. As a result, neutral polymorphism provides basic knowledge to answer the population genetic questions involved with isolation, migration and evolution of populations.

Isozymes are functionally similar, but separable forms of enzymes, encoded by one or more loci (Markert, 1975). Within a locus, the variants found, i.e. the alleles, are named allozymes. Starch gel electrophoresis was first used to analyse human blood groups (Smithies, 1955), and its application in fishes followed shortly by Sick (1961) and others. To this day, the application of this technique has been quite extensive and isozyme data is presently available for many species, including marine fish (Rao *et al.*, 1989; Agnèse, 1991; Kotoulas *et al.*, 1995a; Mamuris *et al.*, 1998), crustaceans (Benzie *et al.*, 1992; Pannacciulli *et al.*, 1997) and molluscs (Jarne *et al.*, 1991; Carvalho *et al.*, 1992; Allcock *et al.*, 1997), just to name a few. This amount of data represents a considerable advantage to people planning new isozyme studies as many systems have been optimized and therefore it is relatively easy to maximise results by directing efforts towards already proven systems.

Protein electrophoresis has proven useful in accessing genetic variation at different levels in marine teleosts. For resolving taxonomic issues, allozyme analysis has proven to be a very satisfactory method, useful in distinguishing between cryptic species, species that are indistinguishable by morphometry, e.g. *Albula* spp., bonefish, *Saurida* spp. and *Synodus* spp., lizardfish (Shaklee *et al.*, 1982) and *Sprattus* spp., Clupeidae (Smith and Robertson, 1981) and to confirm existing taxonomies, e.g. Atherinids (Creech, 1991). At the intraspecific or population level, there is a wealth of data published: blue fin tuna (*Thunnus thynnus*) (Smith and Clemens, 1973); Pacific herring (*Clupea pallasi*) (Grant and Utter, 1984), Atlantic herring (*Clupea harengus*) (Jørstad *et al.*, 1991); round sardinella (*Sardinella aurita*) (Chikhi *et al.*, 1998); striped bass (*Morone saxatilis*) (Fabrizio, 1987); common sole (*Solea vulgaris*) (Kotoulas *et al.*, 1995a); Atlantic flounder (*Platichthys flesus*) and Pacific flounder (*P. stellatus*) (Borsa *et al.*, 1997); red drum (*Sciaenops ocellatus*) (Gold *et al.*, 1993); Pacific halibut (*Hippoglossus stenolepis*) and Atlantic halibut (*H. hippoglossus*) (Grant *et al.*, 1984); Pacific cod (*Gadus macrocephalus*) (Grant *et al.*, 1987); Atlantic cod (*Gadus morhua*) (Mork *et al.*, 1985) and many more.

Although one might predict that the amount of genetic variation observed using DNA rather than protein markers is greater, *per se* because the portion of the genome analysed is more representative, in many cases an increased resolution in the population

subdivision is not always seen. In fact, mitochondrial DNA (mtDNA) markers have produced encouraging results in the majority of cases where allozymes fail to produce evidence for genetic differentiation. However, there are examples of the reverse situation, i.e., significant population differentiation shown by allozymes and not by mtDNA markers: in American oysters (*Crassostrea virginica*) populations, from Chesapeake (Buroker, 1983); Pacific populations of yellow fin tuna, *Thunnus albacares* (Ward *et al.*, 1994a); Australian and New Zealand jackass morwong, *Nemadactylus macropterus* (Elliot and Ward, 1992); Mexico and Atlantic Ocean red drum, *Scianops ocellatus* (Gold *et al.*, 1993). Also with Salmonids isozymes have produced results at both ends of the scale: the technique has been successfully applied for stock discrimination among Pacific salmon (*Oncorhynchus* spp.) (Beacham *et al.*, 1985; Beacham *et al.*, 1987; Utter *et al.*, 1987a; Wood *et al.*, 1989; Shaklee *et al.*, 1990b) on one hand and has revealed low levels of genetic variation in Atlantic salmon (Utter, 1981; Ståhl, 1987) on the other hand.

In general terms, the advantages of allozymes studies as compared with DNA markers are: i) low cost for chemical and labour; ii) many individuals can be scored for several allozyme loci within a short timespan; and iii) codominance - both alleles in a diploid organism are usually clearly identifiable and heterozygotes can be discriminated from homozygotes, which is a pre-requisite for estimation of allele frequencies in population genetic studies. However, there are also a number of limitations to allozyme studies. A new allele will only be detected as a polymorphism if a nucleotide substitution has resulted in a amino acid substitution, which in turn affects the electrophoretic mobility of the molecule. Because of the redundancy of the genetic code (64 codons coding 20 amino-acids) and the fact that not every amino acid replacement leads to a charge difference, only 30% of all nucleotide substitutions are detected by electrophoresis. Therefore allozyme analysis underestimates the genetic variability. Theoretically, studies of genetic polymorphism at the DNA level should yield greater levels of variability than that of protein polymorphism (Nei, 1987; Hallerman and Beckmann, 1988). The use of allozyme electrophoresis restricts the part of the genome that can be analysed as it only detects genetic changes that affect genes that actively

express proteins detectable with histochemical stains. These genes are only a small portion of the whole genome of an organism and so it is legitimate to expect that the amount of genetic variation that can be detected by other DNA methods will increase the resolution. The type of tissue used (e.g. muscle, liver, eye, heart) often requires the sacrifice of the individuals in order to collect the tissue samples. Also the storage of samples must be at a temperature of -20°C, for up to months, or at -70°C for longer storage, as the structure of proteins is labile and excessive denaturation must be prevented after specimen collection.

7.1. MATERIALS AND METHODS

Protein variation was assayed by horizontal starch gel (Pasteur *et al.*, 1987; Morizot and Schmidt, 1990) (Figure 13). Because protein electrophoresis is a widely used technique in genetics it has been widely described in detail in many laboratory manuals (e.g. Pasteur *et al.*, 1987; Leary and Booke, 1990). Here is a short description, referring only to the main conditions that were particular to this work.

Thirty enzyme systems were assessed by starch gel electrophoresis. Nomenclature for protein-coding loci followed international recommendations (Shaklee *et al.*, 1990a) and is referred to in Table 16. The nomenclature used can be summarised in the following way:

- Enzyme abbreviations are represented by non-italicised uppercase.
- Gene symbols are distinguished from enzyme by two topographic conventions: they are italicised and include an asterisk (e.g. *SOD**).
- Alleles are represented by their relative electrophoretic mobility, preceded by an asterisk and italised (i.e. *SOD*100*, *SOD*50*, means the *SOD* has two alleles: the most common allelic product at the locus, allele 100, and a slower allele 50).
- When two or more forms of the same enzyme are known, these isozymes are identified by an Arab number that follows the enzyme abbreviation and is separated from it by an hyphen (e.g. *GPI-1**), according to the gel position of their products, 1 being the most cathodal.

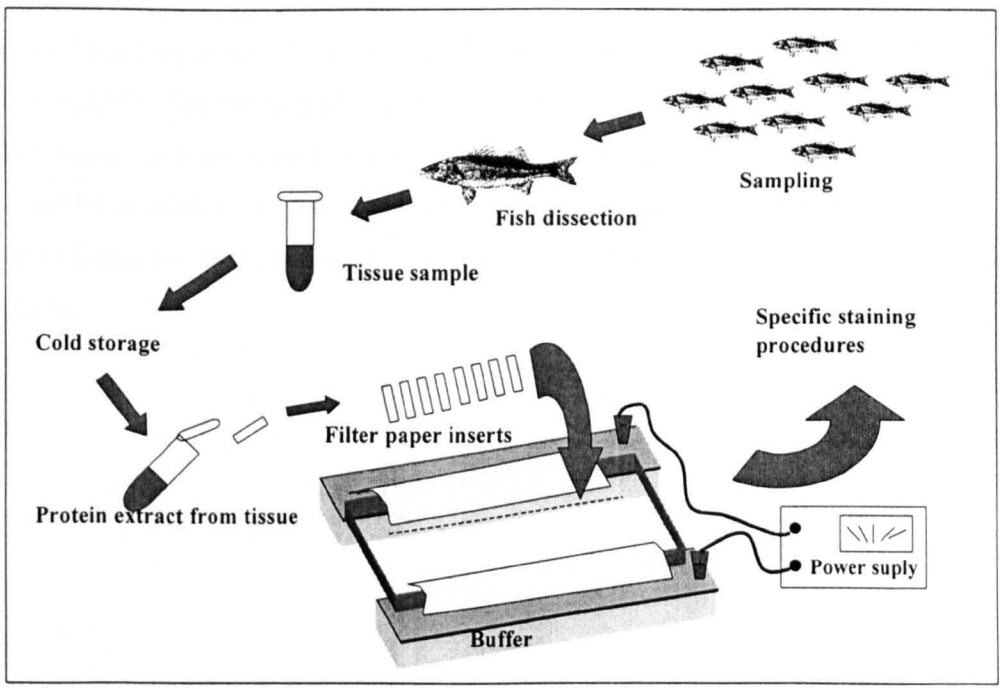


Figure 13. Standard steps for obtaining genotypic data from electrophoresis. (Adapted from Utter *et al.*, 1987b).

Gels were prepared as a 10% suspensions of hydrolysed starch (SIGMA S-4501). Gel and electrode buffers used were: Continuous Tris-Citrate, pH 8.0 (Pasteur *et al.*, 1987); Tris EDTA Borate pH 8.5 (McAndrew and Majumdar, 1983) and Citrate Aminopropyl Morpholine, pH 6.0 (Clayton and Tretiak, 1972) (see Annexe 1 for formulations).

Histochemical staining methods were first adapted from a earlier work, because that work was done in the same species and at the same laboratory as the present work (Martínez, 1991). Working recipes were cut down in order to reduce the chemical's concentrations used. When poor resolution or low activity was observed, recipes from alternative sources were used (Harris and Hopkinson, 1976; Pasteur *et al.*, 1987; Murphy *et al.*, 1990). All the recipes that gave the best results and that were used for the routine screening are detailed in Annexe 1.

All laboratory procedures were performed at the Institute of Aquaculture, University of Stirling. A set of 50 individuals, from the first sample collected (Faro – 1992), was used for the preliminary analysis, to determine the optimal combination of gel buffer, tissue and staining recipes. This was done, in spite of the existence of previous works available on this same species. It is known that variations in water quality and chemicals used can cause changes in the staining, thus altering results and its resolution.

Table 16. Name, abbreviations and number of the 30 enzyme systems studied.

Number as recommended by the International Union of Biochemistry, Nomenclature Committee (1984). (Committee), 1984)

NAME	ABBREVIATION	NUMBER (E.C.)
Acid phosphatase	ACP	3.1.3.2
Aconitase hydratase	AH	4.2.1.3
Adenosine deaminase	ADA	3.5.4.4
Adenylate kinase	AK	2.7.4.3
Alcohol dehydrogenase	ADH	1.1.1.1
Aspartate aminotransferase	AAT	2.6.1.1
Catalase	CAT	1.11.1.6
Creatine kinase	CK	2.7.3.2
Diaphorase	DIA	1.6.2.2
Esterase	EST	3.1.1.-
Fructose biphosphate aldolase	FBALD	4.1.2.13
Fumarate hydratase	FH	4.2.1.2
Glucose-6-phosphate dehydrogenase	G6PDH	1.1.1.49
Glucose-6-phosphate isomerase	GPI	5.3.1.9
Glutamate dehydrogenase	GLUDH	1.4.1.-
Glycerol-3-phosphate dehydrogenase	G3PDH	1.1.1.8
Hexokinase	HK	2.7.1.1
Isocitrate dehydrogenase	IDHP	1.1.1.42
L-Iditol dehydrogenase	IDDH	1.1.1.14
L-Lactate dehydrogenase	LDH	1.1.1.27
Malate dehydrogenase	MDH	1.1.1.37
Malic enzyme	MEP	1.1.1.40
Mannose-6-phosphate isomerase	MPI	5.3.1.8
6-Phosphogluconate dehydrogenase	PGDH	1.1.1.44
Peptidase	PEPA	3.4.-.-
Phosphoglucomutase	PGM	5.4.2.2
Purine-nucleoside phosphorylase	PNP	2.4.2.1
Pyruvate kinase	PK	2.7.1.40
Superoxide dismutase	SOD	1.15.1.1
Xanthine dehydrogenase	XDH	1.2.1.37

7.2. RESULTS

Although allozyme based works have been produced on this species, at the same laboratory as the present work, as said before, it was realised from the first few electrophoresis runs, that probably due to sample conditions, an optimisation of the process was needed, as satisfactory results were not obtained with some of the existing recipes.

Samples caught in 1992 and 1993 were screened for all loci, whereas samples from 1994 were screened only for the polymorphic loci.

Genotypes of polymorphic loci of all individuals included in the analysis are detailed in Annexe 2.

7.2.1. Monomorphic systems

Fourteen enzyme systems were found to be monomorphic (Table 17 and Figure 14). It was not possible to ascertain the protein structure or even to confirm the number of loci due to the lack of polymorphism. However, comparison of the patterns observed with the literature, lead us to propose, for further calculations (such as heterozygosities) a total number of 29 monomorphic loci.

Table 17. Number of loci found, optimal buffer and tissue for *D. labrax* monomorphic isozyme systems.

[¶]Conservative approach (see Figure 14)

ENZYME	LOCI	BUFFER	TISSUE
ADH	3	TEB	Liver (*ADH-1**, *ADH-3**)
			Muscle (*ADH-2**)
CK	2	TEB	Muscle
DIA	2	CAM	Liver
EST	2[¶]	CTC	Heart
FH	2	CTC	Muscle
GLUDH	1	CTC	Liver
IDDH	2	CTC	Liver
LDH	3	CTC	Eye
MDH	2	CTC	Muscle
MPI	2	CTC	Liver
PEPA	2	TEB	Liver
PGM	3	CAM	Liver
PNP	1	CTC	Liver
XDH	2	CAM	Liver
Total Enzyme Systems			14
Total Monomorphic Loci			29

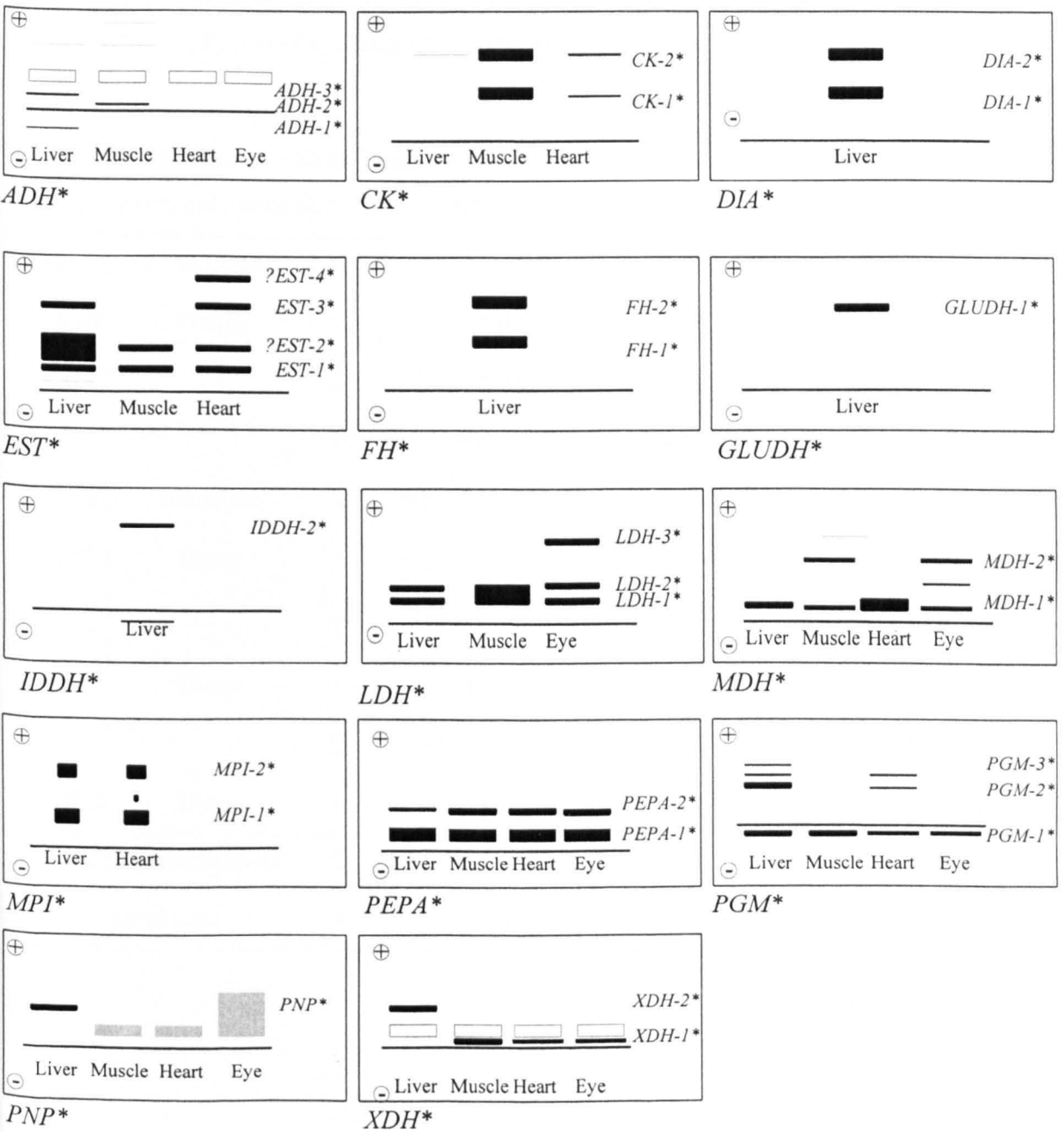


Figure 14. Diagrammatic representation of zymograms of monomorphic enzyme systems found in *Dicentrarchus labrax*.

7.2.2. Polymorphic systems

Five enzyme systems, with a total of 9 loci, 6 of which were polymorphic, were found (Table 18). These are described in detailed in the following sections.

Table 18. Number of loci found, corresponding alleles, optimal buffer and tissue for *D. labrax* polymorphic isozyme systems.

ENZYME	STRUCTURE	LOCI	ALLELES	BUFFER	TISSUE
AAT	Dimer	1	*100	CAM/TEB	Liver
		2	*100	CAM/TEB	Liver
		3	*90, *100	CAM/TEB	Liver
ADA	Monomer	1	*45, *75, *100, *150	CAM/TEB	Liver/Eye
G3PDH	Dimer	1	*100	CTC	Muscle (*G3PDH-1**)
		2	*80, *100, *125	CTC/TEB	Liver (*G3PDH-2**)
GPI	Dimer	1	*100, *150	CTC	Muscle
		2	*80, *100, *110	CTC	Muscle
SOD	Dimer	1	*50, *100	TEB	Liver (on *ADH**)
Total Enzyme Systems		5			
Total Loci		9			

AAT

Aspartate aminotransferase, E.C. 2.6.1.1., Dimer, Liver/CAM (3 loci).

In the seabass the best results for AAT* were obtained with liver tissue with CAM buffer. The product of *AAT-1**, was cathodal, while *AAT-2** and *AAT-3**, were anodal, with an heterodimer band observed. Both *AAT-1** and *AAT-2** were monomorphic and *AAT-3** was polymorphic with 2 alleles (**90, *100*) (Figure 15).

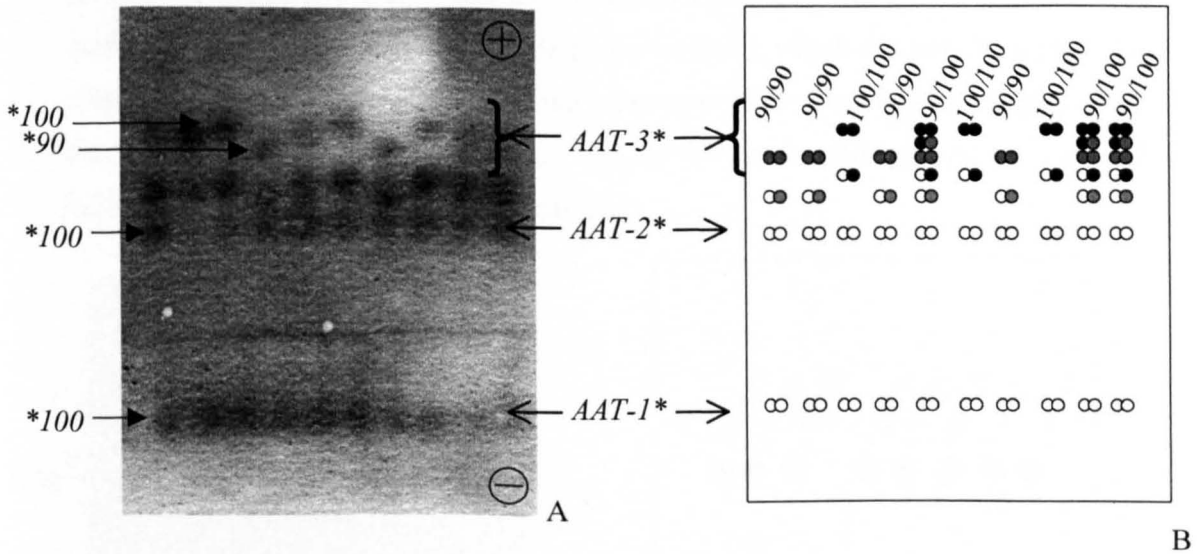


Figure 15. Aspartate aminotransferase polymorphism.

(A) Zymogram exhibiting diallelic variation at the *Aspartate aminotransferase-3** locus (*AAT-3**). *AAT-1** and *AAT-2** are monomorphic. (B) Diagrammatic interpretation showing banding patterns for the dimeric enzyme with the products of 3 loci encoding the enzyme appearing in the same tissue. *AAT-2** and *AAT-3** form interlocus heterodimers. As it is a dimeric protein, each polypeptide is formed by the combination of two sub-units.

Allele 100 is represented by •, allele 90 by ● , and *AAT-2** by ○. An individual with genotype *100/100* will show a band •• and an individual with genotype *90/90*, will show a band ●● . Also, in each case an intermediate band, the heterodimer •○ or ●○will appear, respectively. A heterozygote individual will show 3 bands, resulting from the combination of the 3 sub-units (••, ●● and ○○) and 2 heterodimers (•○ and ●○).

ADA

Adenosine deaminase, E.C. 3.5.4.4, Monomer, Liver/CAM/TEB (1 loci).

All tissues were run simultaneously, liver, muscle and eye on CTC buffer. Three different recipes were tested (Pasteur *et al.*, 1987; Morizot and Schmidt, 1990; Martínez, 1991). One band was detected in the three tissues with this last stain, while the other stains were difficult to read. Four alleles were found (**45, *75, *100* - Figure 16 - and **150* - not shown). In the first few runs with individuals from Faro, to ascertain the running conditions and the recipes, there was a problem with one or both enzymes used (Xanthine oxidase and Nucleoside phosphorylase), which did not work properly, ruining the results. As a consequence, the subsequent liver deterioration, resulting from the process of consecutive thawing, led to the successful scoring of only 22 individuals, out of a total of 65 in that sample.

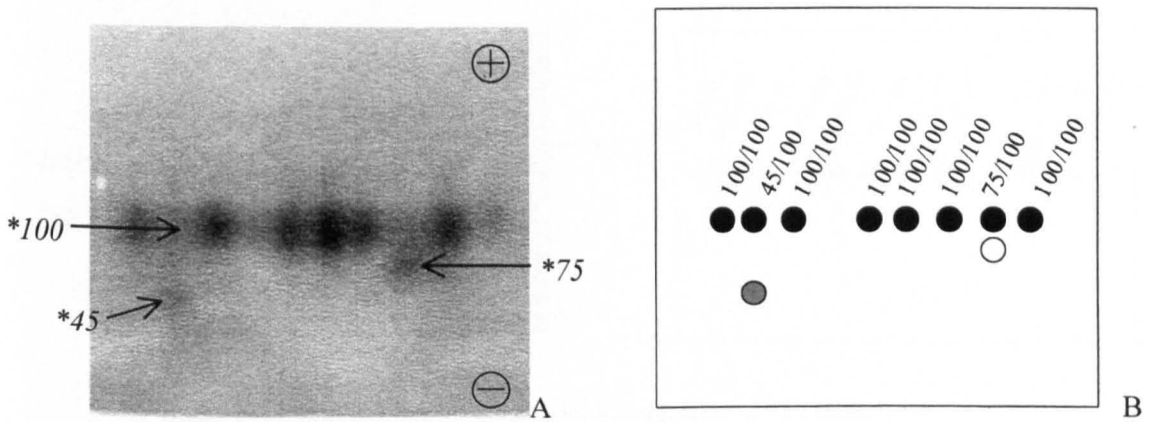


Figure 16. Adenosine deaminase in bass liver.

(A) Zymogram exhibiting variation at the Adenosine deaminase. (B) Diagrammatic interpretation showing banding patterns for the monomeric enzyme, i.e., a protein made up of one polypeptide chain. Homozygotes have one band and heterozygotes two bands. Allele **100* is represented by ●, allele **75* by ○, and allele **45* by ◉.

G3PDH

Glycerol-3-phosphate dehydrogenase, E.C. 1.1.1.8, Dimer, Muscle (*G3PDH-1**)-Liver (*G3PDH-2**)/CTC/TBE.

Two anodal loci were scored for this enzyme each represented by a single band with different electrophoretic mobility and tissue predominance, *G3PDH-1** in muscle and G3PDH-2* in liver. Best results were obtained with a recipe (slightly modified from Morizot and Schmidt, 1990 see annexe 1). Although *G3PDH-1** appeared to be monomorphic, *G3PDH-2** was polymorphic with 3 alleles (**80, *100, *125*) (Figure 17). There were 18 individuals that were not possible to score, which was probably due to liver degradation, after the process of thawing and freezing.

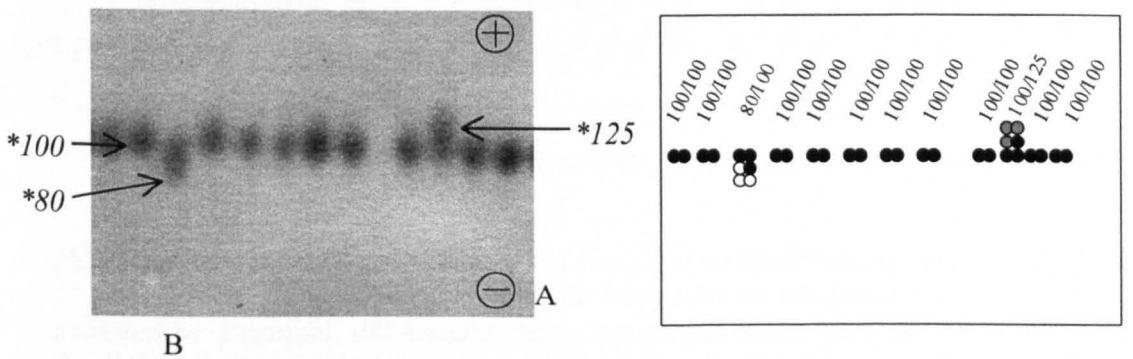


Figure 17. Glycerol-3-phosphate dehydrogenase-*2** in bass liver .

(A) Zymogram exhibiting variation at the *Glycerol-3-phosphate dehydrogenase-2**. (B) Diagrammatic interpretation showing banding patterns for the dimeric enzyme. Allele **100* is represented by ●, allele **80* by○ , and allele **125* by● . An individual with genotype *100/100* will show a band ●● , an individual with genotype *80/80*, will show a band ○○ and individual with genotype *125/125*, will show a band ●● .

Heterozygotes will show 3 bands in total, 2 resulting from the combination of the homomers (●● ,○○ and ●●) and an intermediate heterodimer (○● , ●○ or ●●).

GPI

Glucose Phosphate Isomerase, E.C. 5.3.1.9, Dimer, Muscle/CTC (2 loci).

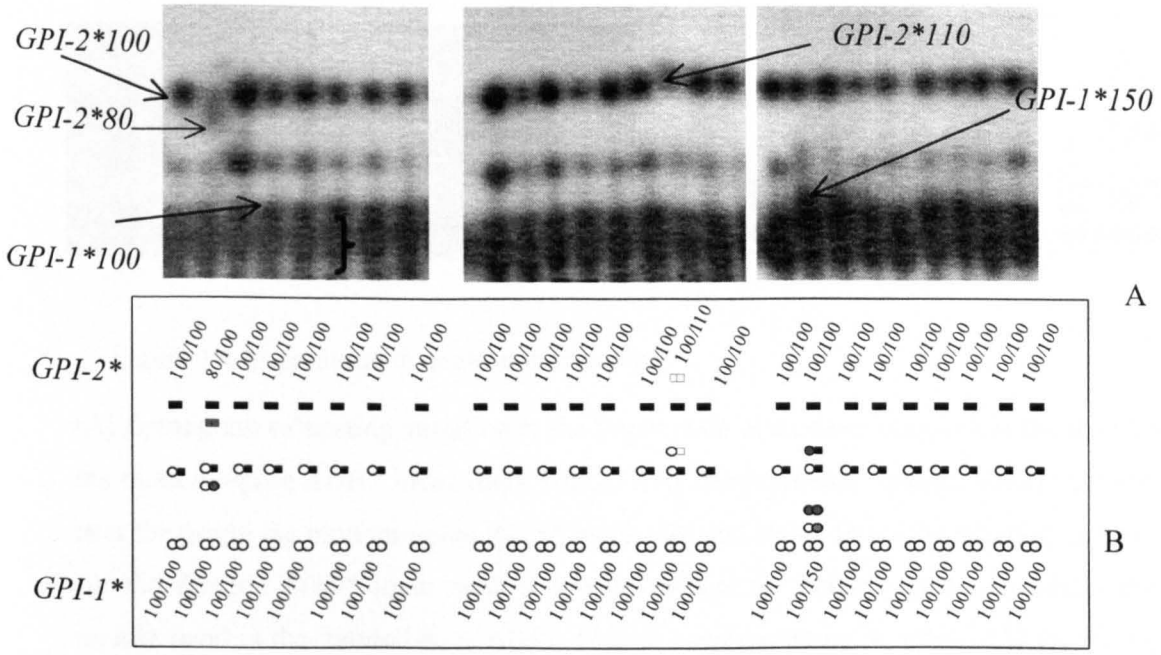


Figure 18. Glucose Phosphate Isomerase (*GPI-1* and GPI-2**) in muscle tissue.

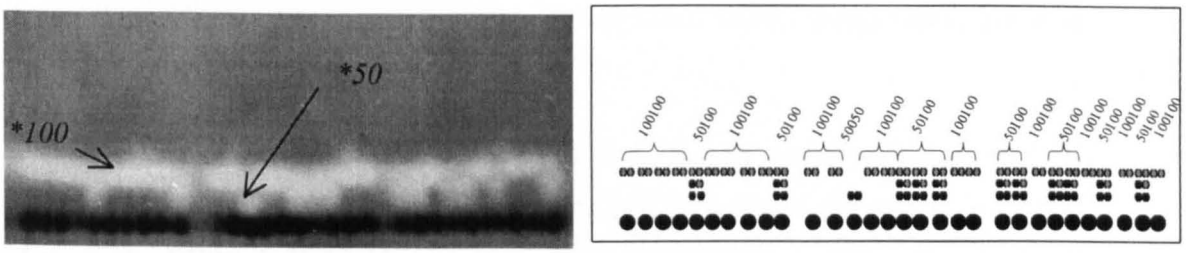
(A) Zymogram exhibiting variation at the two Glucose Phosphate Isomerase loci: *GPI-1** and *GPI-2**. Also observable was the formation of sub-bands on *GPI-1** locus indicated by } symbol. (B) Diagrammatic interpretation showing banding patterns for the dimeric enzyme with the products of 2 loci encoding the enzyme in the same tissue. *GPI-1** and *GPI-2** form interlocus heterodimers. As it is a dimeric protein, each polypeptide is formed by the combination of two sub-units.

Allele *GPI-1*100* is represented by ○ , and allele *GPI-1*150* is represented by ●. Allele *GPI-2*80* is represented by ▪ , allele *GPI-2*100* is represented by ▪ and allele *GPI-2*110* is represented by □ . Thus, an individual with genotype *GPI-1*100/100*, *GPI-2*100/100* will show three bands (○○ , ○▪ and ▪▪).

Two loci were scored: *GPI-1*** in muscle, with 2 alleles (**100, *150*) and *GPI-2***, with best resolution in liver, but also scorable in muscle, with 3 alleles (**80, *100, *110*) (Figure 18). The dimeric nature of this protein was shown with the observable products of the loci and their intermediate heterodimers. There was, near the origin, the formation of a sub-banding pattern.

SOD

Superoxide dismutase, E.C. 1.15.1.1., Dimer, Liver/CTC (1 locus) on ADH.



A B

Figure 19. Superoxide dismutase in bass liver.

(A) Zymogram exhibiting variation at the Superoxide dismutase locus (clear bands) and the monomorphic *ADH** locus (dark bands). (B) Diagrammatic interpretation showing near the origin the monomorphic *ADH** bands (●), and above these the banding patterns for the dimeric *SOD** locus with a 3 band pattern for heterozygotes, in which the middle band is the heterodimer. Allele **100* is represented by ●, allele **50* by ●. An individual with genotype *100/100* will show a band ●●, an individual with genotype *50/50*, will show a band ●●. Heterozygotes will show 3 bands in total, 2 resulting from the combination of the homomers (●● and ●●) and an intermediate heterodimer (●●).

A very clear set of bands was obtained, when staining the gels for ADH. The light bands indicating the presence of SOD appear very clearly. One locus with 2 alleles (**50*, **100*) was detected, as shown in Figure 19 .

7.2.3. Enzyme systems not considered for analysis

A total of 11 enzyme systems were not used in the analysis: ACP, AH, AK, CAT, FBALD, G6PDH, HK, IDHP, MEP, PGDH and PK. ACP, AH, CAT, FBALD, IDHP and MEP showed some evidence of polymorphism, but none were used in the genetic assessment. These systems would not stain consistently and if they did often gave poorly resolved smeared bands, making it difficult to interpret the gels. A description of the observations is offered instead.

It was not possible to get reproducible results with the ACP system, even though a combination of tissues, buffers and stains were tried.

Two bands were apparent after staining for AK system, although the bands in liver and muscle tissue appeared as smears. The top band in liver tissue was common to G6PDH, HK and PGDH systems.

With the AH system, a combination of liver, muscle and heart tissue on CTC buffer with four different stains and buffers recipes were tested (Pasteur *et al.*, 1987; Morizot and Schmidt, 1990; Murphy *et al.*, 1990; Martínez *et al.*, 1991). One invariant band was detected in liver using the four stains; none to 2 bands could be detected in muscle respectively and three bands in heart for all stains. The best resolution was obtained with the "Pasteur" stain.

CAT staining with sodium thiosulphate, potassium sulphate and acetic acid (Murphy *et al.*, 1990) was the most effective. There was an interesting pattern in heart although not reproducible and too much activity in liver. This staining turns the gel blue and the catalase activity is detected by the appearance of white bands, but they disappeared rapidly. Diluting the sample does not improve the resolution or reducing some of the recipe components was not successful in slowing the reaction, as scoring or photographs have to be made very fast before the whole gel turns dark blue.

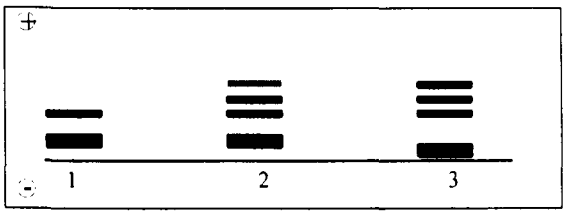
Very poor resolution was obtained for FBALD, even with different combinations of buffers and tissues. Three tissues were tried: liver, muscle and heart and two stains were tested (Murphy *et al.*, 1990; Martínez *et al.*, 1991). The latter gave better results, but even so, there appears to be too much activity to have a clear gel reading and reproducibility was low. A dilution of samples did not improve the final results.

*G6PDH** was stained with four different recipes (Pasteur *et al.*, 1987; Morizot and Schmidt, 1990; Murphy *et al.*, 1990; Martínez *et al.*, 1991) and it appears that two loci were present in liver and muscle but these were ambiguous because of poor resolution, with a great deal of band smearing.

In the HK system it was possible to detect activity in liver tissue with a single strong band appearing, however the same band (in position, colour, intensity and form)

was also found with staining which had fructose or mannose as substrate or even with no substrate at all. The band also appeared when staining for AK, G6PDH and PGDH systems.

Three different patterns were observed in seabass heart for *IDHP** as shown in Figure 20. It was not possible to get reproducible results even with different staining recipes.



*IDHP**

Figure 20. Pattern found in liver tissue with IDHP* staining.

In the MEP system a pattern of three bands was observed in muscle, although the expected tetrameric nature of this protein was not observed (Morizot and Schmidt, 1990). The bands were well separated, although having an undulating appearance.

The pattern observed for PGDH was similar to the one described by Martínez *et al.* (1991), a slow faint band and a fast stronger band, although as previously stated, the stronger band was common to AK, HK, and G6PDH systems.

*PK** isozyme shows a smear of bands in liver tissue as well as in eye, it is only in heart that one or possibly two loci are present and in muscle two or three loci.

7.2.4. Age/length structure

To assess the temporal stability of allele frequencies, a modal component analysis was undertaken to quickly identify the cohorts present in the samples. This procedure was performed for two reasons: avoid ageing every single individual and to allow the inclusion of individuals for which neither scales or otoliths samples were available. Only after this allocation to a year class were the temporal stability tests performed.

To assure the correct allocation of individuals, the modal classes identified were aged *a posteriori* with otoliths and scales from a group of individuals chosen from the tails or overlap in the distributions. Scale reading for age determination is easier than otolith reading in the seabass, but both procedures are labour intensive. The same scale ring pattern described by Andrade (1983) for seabass was observed in this study.

Analysis of the age structure showed that several spawning cohorts were present in some of the populations sampled (Table 19 and Figure 21). The 1993 and 1994 sample from Aveiro could be placed into three cohorts, 1990, 1991 and 1992. The 1993 and 1994 samples from Milfontes could be placed into 1991, 1992 and 1993 cohorts. G-tests were performed to evaluate how well the expected distributions fitted the observed ones. Results showed no significant differences ($P>0.05$) (Table 20) in all tests. The smaller sample sizes or the presence of a single cohort at the other sites meant that this analysis could not be undertaken in all samples.

Table 19. Average length and standard deviation of cohorts in samples.

	Component/ Cohort	Average length (cm)	Standard Deviation	Min.	Max.	n
Foz	1/1991	11.62	0.925	10.70	12.55	107
November 1993	2/1990	15.36	0.802	14.55	16.16	2
Aveiro	1/1992	15.5	1.45	14.03	16.86	103
February 1993	2/1991	21.8	2.10	19.67	23.88	6
	3/1990	24.7	0.28	24.45	25.00	6
Aveiro	1/1992	22.1	0.62	21.13	23.5	15
March 1994	2/1991	23.9	0.80	23.11	24.72	24
Obidos	1/1991	22.42	1.079	21.34	23.50	98
February 1993	2/1990	27.25	0.580	26.67	27.83	3
Milfontes	1/1991	23.57	1.245	22.32	24.81	50
February 1993						
Milfontes	1/1993	13.49	1.036	12.45	14.53	84
February 1994	2/1992	19.61	1.963	17.65	21.57	25
Faro	2/1991	20.73	0.816	19.91	21.55	53
April 1992						

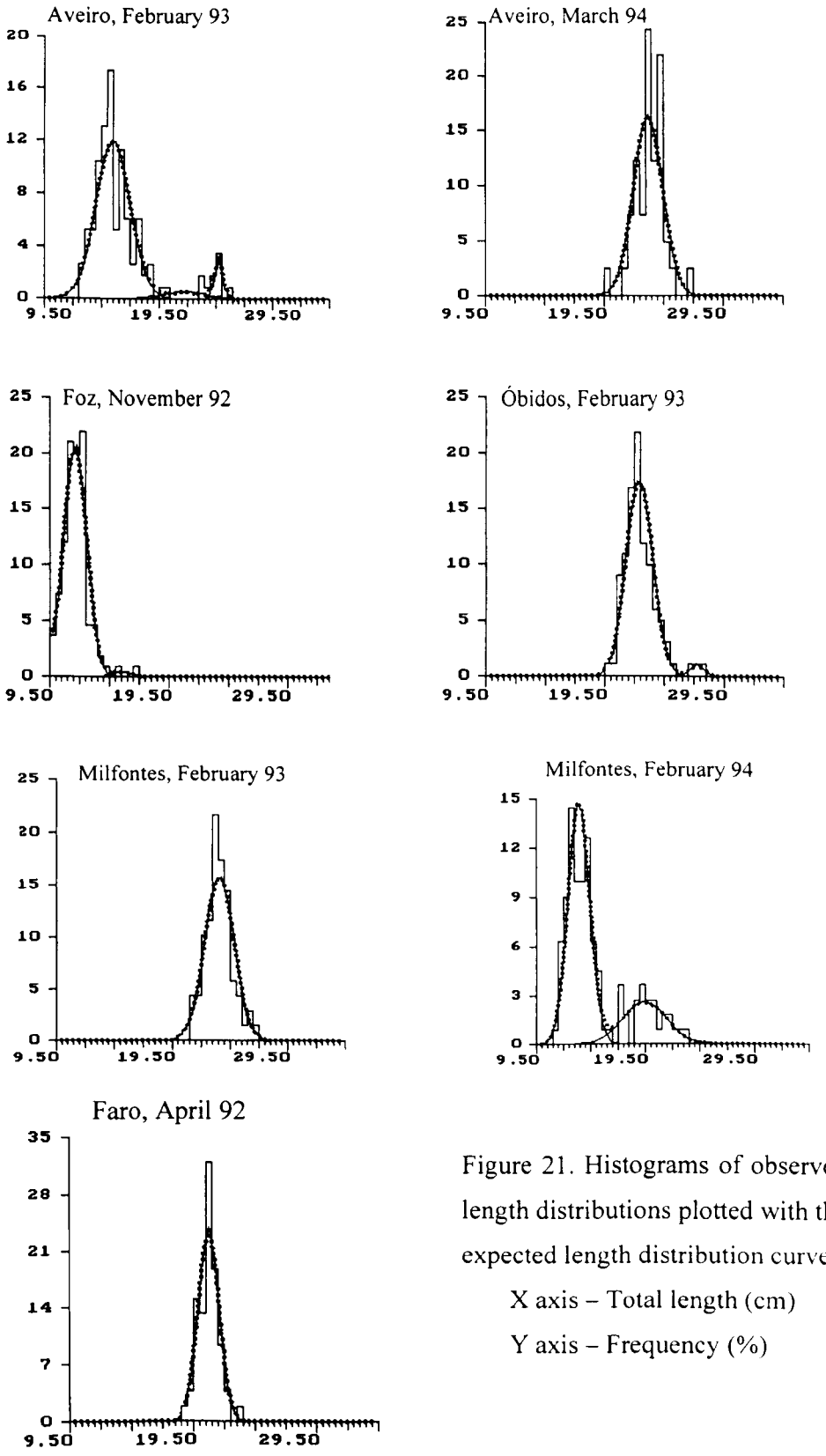


Figure 21. Histograms of observed length distributions plotted with the expected length distribution curves.

X axis – Total length (cm)

Y axis – Frequency (%)

Table 20. G-test probability results for adjustment of expected length distribution to observed distributions displayed in Figure 21.

Location	Foz	Aveiro	Aveiro	Óbidos	Milfontes	Milfontes	Faro
Collection date	11/93	02/93	03/94	02/93	02/93	02/94	04/93
Probability	0.371	0.067	0.06	0.453	0.738	0.824	0.723

7.2.5. Hardy-Weinberg equilibrium

The analysis was divided into three parts: an analysis of the total catch from each sampling year, an analysis of cohorts and an analysis of each sample.

Within sampling years

Tests of Hardy-Weinberg equilibrium using the probability test did not indicate a departure from the equilibrium for fish caught in 1992 (all caught in Faro), 1993 and 1994 (Table 21 – Genepop option 1, sub-option 3). All individual samples for each locus are in Hardy-Weinberg equilibrium.

Table 21. Probability values of Hardy-Weinberg probability test for sampling years. Asterisk indicate the level of significance: $^{*}P<0.05$ after Bonferroni adjustment for simultaneous comparisons; — means only one allele was present or two alleles were detected but one was represented by only one copy or no data was available.

Locus	1992	1993	1994
*AAT-3**	0.3119	0.3319	0.3417
*ADA**	—	1.0000	1.0000
*G3PDH-2**	—	1.0000	0.0558
*GPI-1**	1.0000	1.0000	1.0000
*GPI-2**	—	1.0000	1.0000
*SOD**	0.3974	0.5439	0.3016
All	0.6529	0.9917	0.5883

Within cohorts

Hardy-Weinberg probability tests were performed for two of the 3 cohorts identified in Aveiro (1991 and 1992), as cohort 1990 was represented by only six individuals. The results yield no significant departures from Hardy-Weinberg expectations (Table 22 - Genepop option 1, sub-option 3).

Table 22. Hardy-Weinberg probability test results and F_{IS} values for Aveiro cohorts.

P, probability values; — means only one allele was present or two alleles were detected but one was represented by only one copy.

	Aveiro (n=30) Cohort 1991		Aveiro (n=118) Cohort 1992	
Loci	*P*	F_{IS}	*P*	F_{IS}
*ADA**	—	—	1	-0.017
*AAT-3**	1	-0.150	0.1711	+0.132
*G3PDH-2**	—	—	1	-0.009
*GPI-1**	1	-0.029	1	-0.020
*GPI-2**	—	—	1	+0.077
*SOD**	1	+0.077	0.7166	+0.034
All	1 (df=6)		0.9794 (df=12)	

Similar results were obtained for the 3 cohorts (1991, 1992, 1993) identified in Milfontes samples (Table 23 - Genepop option 1, sub-option 3).

Table 23. Hardy-Weinberg probability test results and F_{IS} values for Milfontes cohorts.

P, probability values; — means only one allele was present or two alleles were detected but one was represented by only one copy.

	Milfontes (n=50) Cohort 1991		Milfontes (n=25) Cohort 1992		Milfontes (n=84) Cohort 1993	
Loci	*P*	F_{IS}	*P*	F_{IS}	*P*	F_{IS}
*ADA**	1	-0.005	1	-0.006	—	—
*AAT-3**	1	+0.023	0.2861	+0.391	0.1826	-0.161
*G3PDH-2**	—	—	1	-0.006	—	—
*GPI-1**	1	-0.043	1	-0.012	—	—
*GPI-2**	—	—	—	—	—	—
*SOD**	0.4693	-0.157	0.1147	-0.203	0.3840	+0.189
All	0.9925 (df=8)		0.6551 (df=10)		0.3525 (df=4)	

Therefore, for both samples where different cohorts were identified (Aveiro and Milfontes), no significant deviations from Hardy-Weinberg expectations were detected at any polymorphic locus.

In Table 24, F_{IS} values are presented for each sample, each polymorphic locus and also multilocus values. In this table, only *G3PDH-2** in the sample from Milfontes is significant, all other tests are non-significant, with probability values represented in Table 25 (Genepop option 1, sub-option 3).

When samples are pooled together for a global Hardy-Weinberg test, results indicate that no deviations to the Hardy-Weinberg proportions are found, $P = 0.7843$.

Table 24. F_{IS} values estimated for each sample, each locus and multilocus situation.

— means only one allele was present or two alleles were detected but one was represented by only one copy or no data was available.

Locus/Samples	Aveiro	Foz	Óbidos	Milfontes	Faro	Total
*AAT-3**	0.0766	-0.1109	0.1696	-0.0508	0.1409	0.0106
*ADA**	-0.0143	-0.0196	-0.0118	-0.0054	—	-0.0134
*G3PDH-2**	-0.0086	0.0000	—	0.2784*	—	0.1361
*GPI-1**	-0.0282	-0.0531	-0.0251	-0.0294	-0.0164	-0.0344
*GPI-2**	-0.0159	-0.0085	-0.0251	0.0024	—	-0.0155
*SOD**	0.0315	0.0279	0.1116	-0.1273	0.1097	-0.0009

Table 25. Hardy-Weinberg probability test values estimated for each sample, each locus and multilocus situation.

— means only one allele was present or two alleles were detected but one was represented by only one copy or no data was available.

Locus/Samples	Aveiro	Foz	Óbidos	Milfontes	Faro	Total
*AAT-3**	0.3438	0.2674	0.1050	0.5690	0.3119	0.2387
*ADA**	1.0000	1.0000	1.0000	1.0000	—	1.0000
*G3PDH-2**	1.0000	—	—	0.0488	—	0.1963
*GPI-1**	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
*GPI-2**	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
*SOD**	0.7498	0.7954	0.2625	0.0820	0.3974	0.3930
All	0.9973	0.9791	0.7082	0.4322	0.6529	0.1356

7.2.6. Polymorphism and heterozygosity

If Hardy-Weinberg deviations between cohorts had been observed, that effect had to be taken into account and partitioned polymorphism and heterozygosity values would have been shown. However, no cohort or sampling year Hardy-Weinberg deviations were detected and data analysis with pooled cohorts within each sampling site can be performed without being affected by cohort or sampling year effect. This is the reason why the polymorphism and heterozygosity parameters are presented at this stage and not earlier.

Allele frequencies of polymorphic loci are presented in Table 26. The average percentage of polymorphic loci with the $P_{0.99}$ criteria was 12.1%, ranging from 7.9 to 15.8%, whereas at the $P_{0.95}$ level it was 5.8%, ranging from 5.3 to 7.9%.

The percentage of heterozygous loci per individual, taking into account all loci scored, averaged 2%, the minimum value being at Faro (1.2%) and the maximum at Foz (2.7%) (Table 27).

Table 26. *D. labrax* allele designation (relative mobility), allele frequencies and heterozygosity at the six polymorphic loci.

(*N* =number of individuals analysed for each locus; $H_{exp.}$ =Expected heterozygosity; $H_{unb.}$ =Expected unbiased heterozygosity (Nei 1978); $H_{obs.}$ = Observed heterozygosity).

Locus/alleles		Aveiro	Foz	Óbidos	Milfontes	Faro
*AAT-3**						
(N)		159	121	101	208	65
	*90	0.2107	0.2851	0.2426	0.2404	0.0769
	*100	0.7893	0.7149	0.7574	0.7596	0.9231
$H_{exp.}$		0.3326	0.4077	0.3675	0.3652	0.1420
$H_{unb.}$		0.3337	0.4093	0.3693	0.3661	0.1431
$H_{obs.}$		0.3082	0.4545	0.3069	0.3846	0.1231
*ADA**						
(N)		165	120	104	211	22
	*45	0.0121	0.0083	0.0096	0.0071	0.0000
	*75	0.0091	0.0208	0.0096	0.0024	0.0000
	*100	0.9758	0.9708	0.9760	0.9905	1.0000
	*150	0.0030	0.0000	0.0048	0.0000	0.0000
$H_{exp.}$		0.0477	0.0570	0.0473	0.0188	0.0000
$H_{unb.}$		0.0478	0.0572	0.0475	0.0189	0.0000
$H_{obs.}$		0.0485	0.0583	0.0481	0.0190	0.0000
*G3PDH-2**						
(N)		165	121	104	213	51
	*80	0.0091	0.0000	0.0000	0.0117	0.0098
	*100	0.9848	0.9959	1.0000	0.9836	0.9902
	*125	0.0061	0.0041	0.0000	0.0047	0.0000
$H_{exp.}$		0.0300	0.0082	0.0000	0.0324	0.0194
$H_{unb.}$		0.0300	0.0083	0.0000	0.0325	0.0196
$H_{obs.}$		0.0303	0.0083	0.0000	0.0235	0.0196
*GPI-1**						
(N)		165	120	103	211	63
	*100	0.9697	0.9458	0.9709	0.9692	0.9762
	*150	0.0303	0.0542	0.0291	0.0308	0.0238
$H_{exp.}$		0.0588	0.1025	0.0566	0.0597	0.0465
$H_{unb.}$		0.0589	0.1029	0.0568	0.0599	0.0469
$H_{obs.}$		0.0606	0.1083	0.0583	0.0616	0.0476
*GP1-2**						
(N)		165	120	103	213	64
	*80	0.0182	0.0125	0.0291	0.0047	0.0078
	*100	0.9788	0.9875	0.9709	0.9953	0.9922
	*110	0.0030	0.0000	0.0000	0.0000	0.0000
$H_{exp.}$		0.0416	0.0247	0.0566	0.0093	0.0155
$H_{unb.}$		0.0418	0.0248	0.0568	0.0094	0.0156
$H_{obs.}$		0.0424	0.0250	0.0583	0.0094	0.0156
*SOD**						
(N)		161	99	104	211	65
	*50	0.1460	0.2576	0.2212	0.1991	0.1769
	*100	0.8540	0.7424	0.7788	0.8009	0.8231
$H_{exp.}$		0.2493	0.3825	0.3445	0.3189	0.2912
$H_{unb.}$		0.2501	0.3844	0.3462	0.3196	0.2935
$H_{obs.}$		0.2422	0.3737	0.3077	0.3602	0.2615

Table 27. *D. labrax* genetic variability at 39 loci in all populations.

(H_{exp} = Expected heterozygosity; H_{unb} = Expected unbiased heterozygosity (Nei 1978); H_{obs} = Observed heterozygosity; $P(0.95)$ means that a locus is considered polymorphic if the frequency of the most common allele does not exceed 0.95; $P(0.99)$ means that a locus is considered polymorphic if the frequency of the most common allele does not exceed 0.99; *s.d.* = standard deviation).

	Mean sample size/locus	H_{exp}	H_{unb}	H_{obs}	$P(0.95)$	$P(0.99)$	Mean no. of alleles per locus
AVEIRO	164.7	0.0200	0.0201	0.0193	0.053	0.158	1.3
s.d	*(0.2)*	*(0.0670)*	*(0.0672)*	*(0.0633)*			*(0.1)*
FOZ	119.5	0.0259	0.0260	0.0271	0.079	0.132	1.2
s.d	*(0.7)*	*(0.0903)*	*(0.0907)*	*(0.0951)*			*(0.1)*
ÓBIDOS	103.8	0.0230	0.0231	0.0205	0.053	0.132	1.2
s.d	*(0.1)*	(0.0809)	(0.0813)	(0.0701)			*(0.1)*
MILFONTES	212.7	0.0212	0.0212	0.0226	0.053	0.105	1.2
s.d	*(0.2)*	(0.0776)	(0.0778)	(0.0843)			*(0.1)*
FARO	44.6	0.0135	0.0137	0.0123	0.053	0.079	1.1
s.d	*(3.3)*	(0.0522)	(0.0526)	(0.0467)			*(0.1)*

7.2.7. Population differentiation

Allelic counts observed in each sample are presented in Table 28, and the results of sample differentiation tests performed on those counts are presented in Table 29 (Genepop option 3, sub-option 1). The results indicate that allele frequencies at *AAT-3** and *SOD** were distributed heterogeneously (*P-value*<0.05 of H_0: the allelic distribution is independent across samples) among the five samples. Although the multisample, multilocus probability value is significant (*P-value*< 0.00001), the single locus effect of *AAT-3** is the major contribution.

Table 28. Allelic counts in all samples in *D. labrax*.

Locus	Samples	Alleles				
*ADA**		**100*	**45*	**75*	**150*	Total
	Aveiro	322	4	3	1	330
	Foz	233	2	5	0	240
	Óbidos	203	2	2	1	208
	Milfontes	418	3	1	0	422
	Faro	44	0	0	0	44
	Total	1220	11	11	2	1244
*AAT-3**		**100*	**90*	Total		
	Aveiro	251	67	318		
	Foz	173	69	242		
	Óbidos	153	49	202		
	Milfontes	316	100	416		
	Faro	120	10	130		
	Total	1013	295	1308		
*G3PDH-2**		**100*	**80*	**125*	Total	
	Aveiro	325	4	1	330	
	Foz	241	1	0	242	
	Óbidos	208	0	0	208	
	Milfontes	419	5	2	426	
	Faro	101	1	0	102	
	Total	1294	11	3	1308	
*GPI-1**		**100*	**150*	Total		
	Aveiro	320	10	240		
	Foz	227	13	330		
	Óbidos	200	6	206		
	Milfontes	409	13	422		
	Faro	213	3	126		
	Total		1279	45	1324	
*GPI-2**		**100*	**80*	**110*	Total	
	Aveiro	323	6	1	330	
	Foz	237	3	0	240	
	Óbidos	200	6	0	206	
	Milfontes	424	2	0	426	
	Faro	127	1	0	128	
	Total	1311	18	1	1330	
*SOD**		**100*	**150*	Total		
	Aveiro	275	47	322		
	Foz	147	51	198		
	Óbidos	162	46	208		
	Milfontes	338	84	422		
	Faro	107	23	130		
	Total	1029	251	1280		

Table 29. Estimated multisample Wright *F*-statistics and G-based test probability of sample allelic differentiation in *D. labrax*.

	F	*f*	θ	Allelic differentiation	
Locus	(F_{IT})	(F_{IS})	(F_{ST})	Probability	S.E.
*AAT-3**	0.0288	0.0106	0.0184***	0.00001	0.00001
*ADA**	-0.0131	-0.0135	0.0003	0.51871	0.00382
*G3PDH-2**	0.1365	0.1361	-0.0001	0.57024	0.00258
*GPI-1**	-0.0344	-0.0344	-0.0001	0.51246	0.00261
*GPI-2**	-0.0122	-0.0155	0.0032	0.14102	0.00265
*SOD**	0.0063	-0.0009	0.0072*	0.02570	0.00154
All Loci	0.0145	0.0038	0.0108**	0.0002	

Multisample Wright *F*-statistics as estimated by GENEPOP (option 6, sub-option 1). Allelic distribution across samples tested with a G-based test (GENEPOP, option 3 sub-option 1), yielding the probability of rejecting the null hypothesis (H_0: the allelic distribution is identical across samples), presented here with associated standard errors (S.E.). All samples are included in the analysis. F_{ST} single locus probability estimates are indicated: $^{*}P<0.05$; $^{***}P<0.001$. F_{ST} probability estimates for multilocus considered significant following sequential Bonferroni adjustment for 6 simultaneous tests are indicated: ** initial α (0.01/6)=0.0017.

Pairwise analysis will bring out the most striking differences of the five sample sites for single locus differentiation. The Bonferroni correction was applied for 10 tests per locus, which results in an associated α of (0.05/10) 0.005. The results (Table 30) show that Faro is significantly different (*P-value*<0.001) from all other populations at the *AAT-3** locus.

Table 30. Probability values of single locus pairwise comparisons.

Locus	Population pair	Probabilities
*AAT-3**	Foz-Faro	0.00000*
*AAT-3**	Milfontes-Faro	0.00000*
*AAT-3**	Óbidos-Faro	0.00002*
*AAT-3**	Aveiro-Faro	0.00061*
*SOD**	Foz-Aveiro	0.00194*
*GPI-2**	Óbidos-Milfontes	0.01668
*SOD**	Aveiro-Óbidos	0.03659
*ADA**	Foz-Milfontes	0.04846

Only initially significant results are shown (*P*<0.05). Probability values as estimated by GENEPOP (option 3, sub-option 2), were judged significant following sequential Bonferroni adjustment (initial α=0.005) for ten simultaneous comparisons (0.05/10) are indicated with an asterisk (*).

Values of F_{ST} estimated for pairs of populations peaked at 0.05 for Foz and Faro (Table 31 below diagonal), although from the 10 possible comparisons, only two were statistically significant (Foz-Faro and Milfontes-Faro, $P<0.001$), after Bonferroni correction. Negative F_{ST} estimates may arise between genetically very similar samples. In such cases, F_{ST} values and N_em estimates are meaningless and the infinity symbol was used for the latter in Table 31.

Table 31. Matrix of pairwise F_{ST} and N_em values.

	Aveiro	Foz	Óbidos	Milfontes	Faro
Aveiro	-	14	58	100	12
Foz	0,01797 ^{NS}	-	∞	59	5
Óbidos	0,00434 ^{NS}	-0,00037 ^{NS}	-	∞	8
Milfontes	0,00249 ^{NS}	0,00424 ^{NS}	-0,00187 ^{NS}	-	8
Faro	0,02114 ^{NS}	0,05259 ***	0,03220 ^{NS}	0,03085 ***	-

F_{ST} values as estimated by GENEPOP (option 6, sub-option 2) are given below the diagonal. The number of migrants, N_em, exchanged under the island model hypothesis, $N_em = (1- F_{ST})/4F_{ST}$ (Wright, 1951), is given above the diagonal. Computation of N_em from negative F_{ST} estimates are represented by ∞. Tests considered significant after following adjustment for 10 simultaneous tests ($\alpha = 0.05/10 = 0.005$) are indicated below diagonal: NS, Not significant; ***P<0.001.

Estimates of the effective number of migrants per generation (N_em) was 22.9 by Wright's formula and 14.65 by Takahata's formula, calculated with the mean of F_{ST}.

Because there is an effect of the sample from Faro on the genetic structure results presented in Table 29, 30 and 31, that can bias the results, a multisample test was performed on all samples excluding Faro (Table 32). Indeed it seems to be so as no significant results were found.

Table 32. Estimated "all samples but Faro" Wright *F*-statistics and G-based test probability of sample allelic differentiation in *D. labrax*.

	F	*f*	θ	Allelic differentiation	
Locus	(F_{IT})	(F_{IS})	(F_{ST})	Probability	S.E.
*ADA**	-0.0135	-0.0141	0.0006	0.3372	0.00909
*AAT-3**	0.0061	0.0047	0.0013	0.2163	0.01297
*G3PDH-2**	0.1479	0.1470	0.0010	0.4559	0.00712
*GPI-1**	-0.0354	-0.0358	0.0004	0.3902	0.00700
*GPI-2**	-0.0124	-0.0166	0.0041	0.0864	0.00524
*SOD**	0.0032	-0.0125	0.0092	0.0166	0.00461
All Loci	0.0012	-0.0030	0.0042	0.0375	

"All samples but Faro" Wright *F*-statistics as estimated by GENEPOP (option 6, sub-option 1). Allelic distribution across samples tested with a G-based test (Genepop, option 3, sub-option 1), yielding the probability of rejecting the null hypothesis (H_0: the allelic distribution is identical across samples), presented here with associated standard errors (S.E.). Faro sample not included in the analysis. No significant F_{ST} probability multilocus estimates were found following the sequential Bonferroni adjustment for 6 simultaneous tests, initial α (0.01/6)=0.0017.

When pairwise comparisons were computed (Table 33), distant populations were no more significantly differentiated than closer ones. The Mantel test (GENEPOP, option 6, sub-option 5) on a F_{ST} vs geographical distance matrix between all population pairs gives a *P*-value of 0.131, with 10000 permutations, meaning that the genetic distance is not positively correlated with geographic distance.

Table 33. Matrix of pairwise F_{ST} and geographic distance.

	Aveiro	Foz	Óbidos	Milfontes	Faro
Aveiro	0	60	155	425	635
Foz	0,01797	0	95	365	575
Óbidos	0,00434	-0,00037	0	270	480
Milfontes	0,00249	0,00424	-0,00187	0	210
Faro	0,02114	0,05259	0,03220	0,03085	0

F_{ST} values are given below the diagonal. Geographical distance in Km.

For comparison purposes, Nei's (1972) genetic distance was estimated as well as Reynolds' genetic distance as this distance is more appropriate in cases of intraspecific divergence (see data analysis section) (Table 34).

Table 34. Nei's (1972) (below diagonal) and Reynolds' (1983) (above diagonal) genetic distances estimated from polymorphic loci. Data not bootstrapped.

	AVEIRO	FOZ	ÓBIDOS	MILFONTES	FARO
AVEIRO	-	0.01814	0.00432	0.00250	0.02138
FOZ	0.0034	-	-0.00037	0.00428	0.05407
ÓBIDOS	0.0013	0.0008	-	-0.00187	0.03273
MILFONTES	0.0008	0.0011	0.0003	-	0.03134
FARO	0.0034	0.0089	0.0052	0.0048	-

One potential problem when interpreting bootstrap values associated to genetic distance trees, is the artificial robustness of nodes that can be generated when bootstrapping across a limited number of loci (O'Connell *et al.*, 1998b), such as in the present work. The small variance estimate is generated by the possibility that the number of permutations may be much less than the number of re-sampling events. In the present work re-sampling of six loci can in fact lead to erroneous interpretations, and to see if there was any variation in the bootstrap values, fewer replications were made a second time (100 instead of 500). In both cases the values associated varied only slightly remaining, however, above 90%.

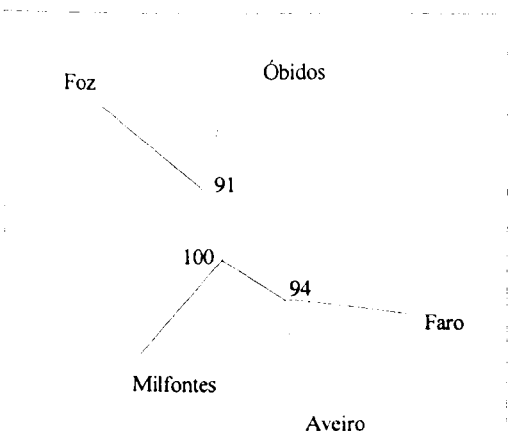


Figure 22. Consensus neighbor-joining tree based on Reynolds genetic distance illustrating the genetic relationships among populations.

Values at nodes represent the percentage of times that the particular node occurred in the 100 trees generated by bootstrapping the original allele frequencies.

The Ewens-Waterson neutrality test was applied only to *AAT-3** and *SOD** as these are the loci that can possibly detect evidence of selection. The other 4 loci are rarely polymorphic and hence of almost no informative value in this type of analysis. The test revealed no single locus evidence for departure from neutrality, as can be observed in Table 35.

Table 35. Ewens-Waterson neutrality test.

Locus	*n*	*k*	Obs. F	Min F	Max F	Mean*	SE*	L95*	U95*
*AAT-3**	1308	2	0.6507	0.5000	0.9985	0.8694	0.0265	0.5042	0.9985
*SOD**	1280	2	0.6847	0.5000	0.9984	0.8706	0.0266	0.5050	0.9984

n, total number of alleles; *k*, different number of alleles; Obs. F, Observed frequency; SE, standard error; L95, lower 95% confidence interval; U95, upper 95% confidence interval; * statistics calculated using 10 000 simulated samples.

7.3. Discussion

The sampling strategy followed in the present work, aimed at increasing the probability of detect genetic differences among geographically close samples, by trying to get "a sufficiently large sample". The level of genetic differentiation usually encountered in marine fish is generally smaller than the differentiation among anadromous or freshwater fish (Ferguson *et al.*, 1995). This can usually be attributed to higher gene flow among subpopulations, probably due to the relative absence of physical barriers to dispersal in the sea (Carvalho, 1998). So the geographical sampling range and the sample sizes should be associated, because the closer the marine fish populations are, the more likely it is that the genetic differentiation among them is lower, and thus the more difficult it will be to detect the differences. That is, lower samples sizes will probably mean that the intraspecific variation could go undetected. In the present work, as we proposed to sample a relatively narrow geographical range (see

Table 9), where the maximum distance between samples (Aveiro-Faro) was 635 Km, we looked for relatively high sample sizes.

In fact, when comparing the sample sizes shown in Table 9, with the ones from the present work, we can see that on average sample sizes are much higher than any other previous work. Four of the five samples have more than 100 individuals, and it was only Faro that was below that number, for reasons already explained in the materials and methods section. However, it should be restated, that the reliance on commercial artisanal fishermen to obtain samples, has many drawbacks, and that the end-result of a well thought sampling strategy, is more often than not, deficient.

Another important factor when using allozymes is tissue degradation, either upon fishing, and before freezing of the samples becomes possible, or during laboratory manipulations. This degradation may affect the electrophoretic mobility and resolution possibly leading to the misreading of gels. This is specially true in fragile tissue such as the liver, which autolyses quickly. It is not uncommon for an electrophoresis run to be hindered by something like power failure, accidental electricity cut-off, gel overheating, deteriorated buffer or chemicals. In this case the run has to be repeated, and another thawing and freezing of the samples has to done. Due to inexperience, no tissue aliquots were taken, which would have allowed fresh samples to be held in reserve for unexpected problems. In view of the present work's experience, tissue aliquots are greatly recommended.

7.3.1. Polymorphic loci

All population genetic studies on seabass have found a number of polymorphic loci, however the extent of the polymorphism varies, as seen in the Population Genetics chapter. For instance, in a very recent work, Allegrucci *et al.*(1997), found 26 polymorphic loci in 13 different locations including the Mediterranean and the Atlantic. The Atlantic being represented by one sample from Aveiro. The higher polymorphism found might relate to the fact that cellulose acetate and not starch electrophoresis resulted in increased resolution. Moreover, the higher number of polymorphic loci cannot be attributed to larger samples, as the average sample size was 30 individuals, which is less than the smallest sample in this study. Allegrucci *et al.*(1997) used a

alphabetic nomenclature for the alleles, and therefore it is not possible to conclude if the alleles detected in some of the loci previously reported as monomorphic, is due to the increase resolution. In fact, if the alleles reported have very similar migration distances, it would be difficult to tell them apart with starch gel electrophoresis. All polymorphic loci in the present study, except *SOD**, which was not assessed in the Italian study, were also found to be polymorphic by Allegrucci *et al.* (1997). This is to say that comparisons between works have to take into account, not only the results but also the methods to achieve them. From all the works previously published on seabass allozymes, it is the Martínez *et al.* (1991) that reports the most "extreme" results, claiming reproductive isolation between the 2 populations assessed. It will be further referred this work in the appropriate context but it should be said that the work was based on two samples, one of which of only 20 individuals and originated from a hatchery, reported as being part of a F1 progeny from the wild. Not only is the sample size small, but the performance of the sample must be questioned as mixing with other populations might have occurred in a captive regime. This sample is so very much different in many respects from other wild Atlantic samples, that it is legitimate to question if the sample is representative of the wild Atlantic populations. In the following discussion the above mentioned points should be kept in mind.

In this section the polymorphic loci will be discussed from 2 angles, one the comparison essentially with Martínez *et al.* (1991) work in terms of band patterns and another comparing results with all works on seabass allozymes in terms of monomorphism/polymorphism and number of alleles found. However, the section on population differentiation will deal in more detail with differences and similarities of alleles frequencies between the present work and the other studies.

AAT

The AAT system showed interesting similarities and differences when compared with the Spanish samples (Martínez *et al.*, 1991). The results showed a general pattern similar to the ones referred to by Martínez *et al.* (1991) for a sample from the north of Spain (Tinamenor), in which AAT has one cathodal locus of mitochondrial origin and two anodal loci of cytosolic origin. The polymorphism detected in the Portuguese

samples was more closely related, however, to the South of Spain (Cadiz) samples: 2 alleles detected in *AAT-3** (**93* and **100* - Cadiz; **90* and **100* - Portugal) than to the northern sample. It seems fair to assume that a difference in nomenclature between the present work and Benharrat *et al.* (1984)'s work, led to a different designation of the same locus: *AAT-1** locus, with 3 alleles (**80*, **100* and **120*) found in Benharrat *et al.* (1984), is equivalent to the *AAT-3** locus, with 2 alleles (**90* and **100*) of the present work. In Allegrucci *et al.* (1997) work this system, has 2 loci, *AAT-1** and *AAT-2** monomorphic in the sample from Aveiro, being polymorphic in some of the Mediterranean samples, with 2 alleles each. It is possible to assume that these monomorphic loci correspond to the *AAT-1** and *AAT-2** of the present work, also monomorphic, although it may be otherwise.

ADA

The ADA system showed a strong band in all tissues, found to be similar to the one described by Martínez *et al.* (1991), although this system is polymorphic in the Portuguese samples, no polymorphism was found in the Spanish samples. ADA was also found to be polymorphic by Allegrucci *et al.* (1997) in the Aveiro sample, with two of the possible four alleles that were detected among all populations in that work. Two of the samples in the present study, Aveiro included, did show 4 alleles.

G3PDH

The G3PDH system showed a pattern of a strong band in all tissues, similar to the one described by Martínez *et al.* (1991). This system was not assessed by Benharrat *et al.* (1984) nor by Allegrucci *et al.* (1997). It was however scored by Child (1992) who found it to be monomorphic.

GPI

Interlocus heterodimers are usually formed by the GPI protein (Morizot and Schmidt, 1990), and this was clearly observed in this study. Also observable was the formation of sub-bands on *GPI-1** locus, due to the expression of non-genetic secondary isozymes, that may vary according to tissue location, age and electrophoresis buffer used (Murphy *et al.*, 1990). *GPI-1** often yields two anodal sub-bands beyond each homomer or

heteromer, confusing the scoring. The presence of interlocus heterodimers usually provide a good basis for scoring.

Differences are observed in this system when comparing samples from Portugal and Spain: a) there was polymorphism in both loci in muscle whereas in the Spanish samples no polymorphisms were observed and b) only *GPI-1** was visible in muscle tissue in the Spanish samples (Martínez *et al.*, 1991), but both loci were detected in muscle tissue in this sample. These differences could be explained by distinct types of muscle as white muscle and red muscle can express different loci. Benharrat *et al.* (1984) found in the Atlantic, *GPI-1** to be monomorphic and *GPI-2** to be polymorphic, with 2 alleles (**100* and **150*). The GPI-2* locus was indeed only polymorphic in the Atlantic, and not in the Mediterranean. Child (1992) and Allegrucci *et al.* (1997) found both loci polymorphic in the Atlantic, and Allegrucci *et al.* (1997) found only two samples from the Mediterranean polymorphic at *GPI-2**.

SOD

*SOD** can often be scored on gels stained for other enzymes producing formazan precipitates as *FBALD**, *ADH**, *G3PDH**, *AH**, etc., and so, there is no need for a specific stain. Zones of activity were visualized as white or clear bands on a blue background and the contrast can be increased by exposure to light. There is a supernatant/cytosolic, dimeric form and the mitochondrial (tetramer) form of this enzyme (Harris and Hopkinson, 1976). Martínez (1991) found one locus (dimer) in the northern population, and two in the southern population (dimer and tetramer), while only one locus was detected in this study. Benharrat *et al.* (1984) found 2 alleles (dimer) in the Atlantic sample, and 3 in the Mediterranean samples (*25, **100* and **175*). The Cadiz population from Martínez *et al.* (1991) work exhibits 3 alleles (**63*, **100* and **130*) and the sample from northern Spain was monomorphic. Child (1992) found 2 alleles in the only locus detected, in the populations around the British Isles. Similarly to Benharrat *et al.* (1984) and the present work, 2 alleles were scored, with mobilities of **50* and **100*, and it could be assumed that the species has 3 alleles, a slow, a medium and a fast migrating allele of the dimeric form.

7.3.2. Monomorphic loci

ADH

*ADH**, as with all NAD-dependent dehydrogenases, needs to have its substrate specificity confirmed to prevent confusion with other non-specific NAD enzymes. These zones appear in the absence of stain substrate (Morizot and Schmidt, 1990), and inhibition by Pyrazol can help in checking this, which was done. It is usually resolved as a single isozyme in liver tissues, often migrating cathodally, but other products can be found (Morizot and Schmidt, 1990). The pattern observed was not similar to the one described by Martínez *et al.* (1990, 1991). Though no explanation can be offered for this occurrence, it would be desirable to score on the same gel this system for a direct comparison. Benharrat *et al.* (1984) and Child (1992) did not score this system, and Allegrucci *et al.* (1997) also found it to be monomorphic for the sample from Aveiro. Martínez *et al.* (1991) however, found this system to be polymorphic, with 3 alleles in the sample from Cadiz.

CK

Staining for *CK** showed bands, which presumably correspond to 2 loci active in muscle tissue. Although CTC, CAM and TEB buffers resolved this system, the best resolution was attained with TEB buffer. Martínez (1991) observed 2 loci active in muscle tissue with CTC and CAM buffer. It is usually described as a dimeric system, although muscle CK heterozygotes in most fishes produce only the two homodimeric allozymes (Morizot and Schmidt, 1990). Muscle creatine kinase is generally present in sufficient quantity to resolve by general protein stain (Murphy *et al.*, 1990). From previously published works, it was only Allegrucci *et al.* (1997) who scored the CK system, finding CK-4* to be polymorphic with 2 alleles, in liver tissue, in the sample from Aveiro. All the Mediterranean samples were also polymorphic, with 2-3 alleles.

DIA

In poeciliid livers this system has been occasionally observed as two zones of activity (Morizot and Schmidt, 1990). The pattern found is similar to the one in Martínez *et al.* (1991). No other previously published works scored this system.

EST

A complex pattern of bands was observed in esterases of liver and muscle. Three recipes were tested: Martínez (1991), Murphy *et al.* (1990) and Morizot and Schmidt (1990). With all three stains there is a consistent pattern of two bands (a slow and a fast band) in liver. The main problem lies in the bands between these two. The fast band in muscle seems to correspond to this middle band. In heart 4 bands were observed, therefore there appears to be 4 loci (but a conservative approach was adapted in Table 17), although Martínez *et al.* (1991) and Benharrat *et al.* (1984) report the presence of 5 loci. Unfortunately no reproducible results for the intermediate bands were obtained. This enzyme is known to be quite labile and it is a possibility that the sample condition was not the best for this enzyme. Most authors found 2 loci, with Benharrat *et al.* (1984) scoring it as monomorphic in Atlantic and polymorphic in the Mediterranean. Child (1992) scored both loci as monomorphic in the Atlantic, Martínez *et al.* (1991) scoring both loci as polymorphic with 2 alleles in the sample from the north of Spain. Allegrucci *et al.* (1997) did not analysed this system. Because of these equivocal results, this system deserves further attention.

FH

Two FH bands were detected, with a pattern similar to the one found by Martínez *et al.* (1991). Allegrucci *et al.* (1997) scored this system as very weakly polymorphic in one Mediterranean sample. No other previously published works scored this system.

GLUDH

One *GLUDH** locus in liver was detected, with a pattern similar to the one described by Martínez *et al.* (1991) and in general agreement to what was expected: a single isozyme is expressed most strongly in liver tissue extracts often still weak. The structure seems to be quite variable, described as a tetramer (Murphy *et al.*, 1990) and as an hexameric (Aebersold *et al.*, 1987). No other previously published works scored this system.

IDDH

A single gene product is expressed predominantly in liver. Staining for other NAD-dependent or "nothing" dehydrogenases may occur (inhibitors of LDH and ADH are included in the stain recipe) (Morizot and Schmidt, 1990). Two *IDDH** bands were detected in liver, corresponding to one anodal and one cathodal locus, showing no variation. Martínez *et al.* (1991) describes a mix of the bands: one cathodal locus in samples from the north of Spain, 2 cathodal and 1 anodal loci in Cadiz populations. No other previously published works score this system.

LDH

Three LDH gene products have been well characterized in fishes, two with broad tissue specificity and a third restricted to liver or eye, depending upon taxon. Heterotetramer formation among gene products ranges from absent to unrestricted (Morizot and Schmidt, 1990). However, only two bands were observed in liver and eye and one large band in muscle corresponding to *LDH-1** and *LDH-2**, and a third band in eye, *LDH-3** with the Martínez (1991) stain, no heterotetramer bands were observed. These patterns were similar to Martínez *et al.* (1991). Benharrat *et al.* (1984) also found this system to be monomorphic in the sample from Atlantic, but polymorphic in some samples from the Mediterranean. Child (1992) found this system to be monomorphic and Allegrucci *et al.* (1997) did not scored this enzyme.

MDH

*MDH** presents one locus coding the mitochondrial protein form and two loci coding the supernatant/cytosolic form in humans (Harris and Hopkinson, 1976). All three can usually be resolved using muscle tissue but that is not the case in seabass, where only 2 loci were scored. It can be proposed that only the two loci coding the supernatant/cytosolic form were found, because of the intensity of heterodimeric band found in eye. This pattern was generally similar to the one found by Martínez (1991) for the seabass from the north of Spain. Benharrat *et al.* (1984) found 2 loci, both monomorphic in the Atlantic sample as did Child (1992). Allegrucci *et al.* (1997) found

one locus, monomorphic in the sample from Aveiro and polymorphic in some of the Mediterranean samples.

MPI

*MPI** presents one product in virtually all tissues (Morizot and Schmidt, 1990). Products of LDH may also appear as faint bands following staining and LDH activity can be suppressed by adding 50 mg of pyruvic acid (Murphy *et al.*, 1990). The same pattern as Martínez (1991) was observed with the only difference residing in the relative position of the muscle band, at the level of the two lower bands of liver and heart in Martínez (1991). From earlier published works, only Allegrucci *et al.* (1997) scored this system as polymorphic, in 2 out of 12 seabass populations including Aveiro.

PEPA

To visualise *PEPA** two different substrates: L-Leucyglycyl-Glycine and L-Leucyl-Glycine were tried. As can be seen in Figure 14, band a) is present in both substrates, band b) only in L-Leucyl-Glycine and band c) although very faint, is present only in L-Leucyglycyl-Glycine. These three bands could correspond to three different loci, as in Martínez (1991), although he used different substrates and found a slightly different pattern. Only Child (1992) scored this system, and found it to be monomorphic.

PGM

Only one *PGM** locus product (often with more anodal "satellite" bands) has been characterized in poeciliids and is detectable in virtually all tissues. Phenotypes in many other fishes suggest that additional locus products remain to be resolved (Morizot and Schmidt, 1990). Three different stain recipes were tested: Martínez (1991), Murphy *et al.* (1990) and Morizot and Schmidt (1990). A very strong cathodal band was observed in liver and muscle using all three stains. A slow band appears in both tissues with the "Martínez" stain, but only in liver with other two stains. There is a further set of bands in liver with the "Martínez" and "Morizot" stains, but no clear interpretation was achieved. Martínez (1991) does not record the presence of the cathodal band. Martínez *et al.* (1991) and Allegrucci *et al.* (1997) found this system to be monomorphic.

Benharrat *et al.* (1984) found it to be polymorphic for both loci in the Atlantic, with 2 and 3 alleles respectively. Moreover, the sample from the Atlantic was the only one to be polymorphic at the *PGM-1** locus. Child (1992) also found it to be polymorphic with 2 alleles. Indeed, as said in the Population Genetics chapter, this system alone provided evidence for stock separation between bass from the Irish Sea and bass from the Bristol Channel, English Channel and Thames estuary (Child, 1992). Allegrucci *et al.* (1997) scored only one locus, monomorphic in the Atlantic population and polymorphic in all Mediterranean populations, with up to 3 alleles. Because of the number of alleles found in these works, and only on the basis of that, it seems that *PGM-1** found in Benharrat *et al.* (1984) corresponds to *PGM** in Child (1992) (2 alleles) and that *PGM-2** found in Benharrat *et al.* (1984) corresponds to *PGM** in Allegrucci *et al.* (1997) (3 alleles).

XDH

There seem to be two loci, although with a different band pattern from the one Martínez (1991) described. In common there is weak activity, and the relative position of muscle (*XDH-1**) and heart (*XDH-2**). Martínez *et al.* (1991) found *XDH-2** to be polymorphic in the sample from northern Spain, with 2 alleles and monomorphic in Cadiz sample, similar to what was found in this work. No other works on seabass allozymes scored this system.

7.3.3. Loci not considered for the analysis

Unsatisfactory results were attained for ten systems, probably due to enzyme denaturation or oxidation or failure to find the perfect combination of tissue, buffer and recipe that would allow good scoring conditions.

ACP

Martínez *et al.* (1991) describe a 3 loci pattern: *ACP-1** (liver and heart), *ACP-2** (muscle and stomach), *ACP-3** (heart, stomach, spleen).There are two types of ACP's: "red cell" which show pink zones of activity, and "tissue" ACP isozymes (Murphy *et al.*, 1990). According to Morizot and Schmidt (1990) this enzyme system deserves detailed study in fishes. These authors report a cathodally migrating dimeric enzyme in

liver (probably a uridine 5'-monophosphate phosphohydrolase), as well as weak anodal zones of activity in other tissues. Benharrat *et al.* (1984) found this enzyme to be polymorphic in the Atlantic and Mediterranean samples with 2 alleles.

AH

Two loci occur for AH in vertebrates. Two forms: (1) mitochondrial (muscle), (2) cytosolic (liver), with examples (in poeciliids) of "satellite" bands generated more cathodally than the native isozyme (Morizot and Schmidt, 1990). Being a monomeric enzyme the middle band observed in heart is interpreted to be a different locus. The heart pattern was similar to Martínez *et al.* (1991), although only the top liver band and the lower muscle band were observed. Allegrucci *et al.* (1997) found 2 alleles in the Atlantic population and in most of the Mediterranean populations, only the F2 populations from broods collected in local sites having 3 alleles.

CAT

CAT staining with sodium thiosulphate, potassium sulphate and acetic acid (Murphy *et al.*, 1990) was the most effective. The bands appeared very quickly, in a matter of seconds and had to be scored immediately. Martínez (1991) found a 5 band pattern, which he attributed to 2 loci. No other previously published works scored this system.

FBALD

Difficulties in resolving clearly and reproducibly *FBALD** have been experienced before (Martínez, 1991) which lead to him not considering this system. Allegrucci *et al.* (1997) was able to score this locus as polymorphic, with 2 alleles in the Atlantic population and 2-3 alleles in the Mediterranean populations with a total of 4 different alleles found. Interesting is the fact that the most common allele in the sample from Aveiro, is not the most common in any of the other bass populations assessed by Allegrucci *et al.* (1997), and although it does not constitute an absolute diagnostic marker, in terms of presence vs. absence, provides nonetheless appropriate indications, in terms of allelic frequencies.

G6PDH

*G6PDH** showed a pattern similar to the one described by Martínez *et al.* (1991). The structure of this enzyme has been referred to as a mixture of dimers and tetramers in fishes (Morizot and Schmidt, 1990). Benharrat *et al.* (1984) found this system to be polymorphic, with 2 alleles, the most common one with a frequency greater than 0.97 among Atlantic and Mediterranean populations, whereas Martínez *et al.* (1991) and Child (1992) found this locus to be monomorphic in the sample from Tinamenor and the British Isles respectively.

HK

Martínez (1991) detected 2 bands, corresponding to 2 loci in muscle and heart, one in eye (*HK-1**) and liver (*HK-2**). Activity was only detected in liver tissue, and as in Martínez (1991) only one single band was evident. HK is still a poorly resolved systems in fishes (Morizot and Schmidt, 1990) and maybe that is why no other previously published works scored this system.

IDHP

Martínez (1991) and Cervelli (1985) report the existence of two *IDHP** loci, with no variation. In this study three different patterns were observed in seabass individuals, although it was impossible to allocate alleles to specific bands. The explanation for the patterns observed is possibly the one put forward by Utter *et al.* (1987) for rainbow trout: isoloci (different loci of the same multilocus enzyme system, which happens to have identical electrophoretical mobility). With identical mobilities, the isoloci codominant expression can be masked even when it is known which of the two is the polymorphic locus making it most difficult to assign alleles to specific loci. Benharrat *et al.* (1984) and Child (1992) scored 2 locus, finding them to be monomorphic in the Atlantic samples, although the Mediterranean samples assessed by Benharrat *et al.* (1984) were polymorphic. Allegrucci *et al.* (1997) scored one locus, with 3 alleles. It is interesting to see that this is potentially a very important locus in terms of marker, as in the Atlantic population, alleles A and B have frequencies of 0.417 and 0.583, and in the Mediterranean populations allele A is quite rare or completely absent.

MEP

Mitochondrial and supernatant/cytosolic forms of MEP system are know in humans (Harris and Hopkinson, 1976). There is often sufficient breakdown of NADP to NAD in liquid stocks in prolonged storage that NAD-dependent MDH activity will be resolved in addition to MEP. If there is any doubt as to the identity of the MEP, a control slice from the same gel should be stained specifically for MDH, to ascertain which zones of activity are MDH (Murphy *et al.*, 1990). MEP electrophoretic patterns in fishes are often complex and difficult to assign to specific loci (Morizot and Schmidt, 1990). Martínez *et al.* (1991) found a two band pattern for every tissue except for heart where only one stronger, more cathodal band than the other two was found. He concluded that there were 3 loci. The author found both loci to be polymorphic one in each of the analysed samples: *MEP-1**, with 3 alleles in the sample from Cadiz and *MEP-2**, with 2 alleles in the Tinamenor sample. Benharrat *et al.* (1984) also found both loci polymorphic, with 2 alleles each in the Atlantic sample, although *MEP-2** was found to be monomorphic in the Mediterranean. Interestingly the most common allele in the Atlantic sample (*105**), at a frequency of 0.67, is different from the most common allele in the Mediterranean samples (*100**), at a frequency between 0.80 and 0.95. Also Allegrucci *et al.* (1997) found something interesting along the same lines, while scoring only one locus: the most common allele in the Aveiro sample is absent from all Mediterranean populations except one, in which its frequency is 0.02.

PGDH

Although Morizot and Schmidt (1990) say that a single gene product of *PGDH** is present in virtually every tissue only Child (1992) scored 2 loci for this system, both monomorphic. The overall pattern found in the present work was similar to the one described by Martínez (1991) with no variation observed. Benharrat *et al.* (1984) observed allelic variation in both the Atlantic and Mediterranean samples (on the whole with 3 alleles), although only 1 or 2 present at each sample. Allegrucci *et al.* (1997) scored only one locus, with 2 alleles in the sample from Aveiro, from a total of 4 alleles present in the Mediterranean samples.

PK

It was not possible to determine with certainty the number of *PK** loci. It will be necessary to try with fresh tissue and different buffer/recipe combinations in order to bring out the number of loci present in this system. Only Allegrucci *et al.* (1997) scored this system, being monomorphic in all samples scored.

7.3.4. Age/length structure

Year-classes and length results were comparable to previously published studies of seabass in Aveiro (Table 36) (Andrade, 1983). The method for resolving the modal components of the length distributions determined by ANAMOD (Nogueira, 1992), seems to be applicable to the seabass, only being necessary to read otoliths and scales in a few selected individuals near the edge of the distribution instead of the whole sample, and in doing so saving time and labour.

Table 36. Range of lengths in the first 3 year classes of *D. labrax*.

(Andrade, 1983)

Year Class	*N*	Min.-Max. length	Mean	*s.d.*
0+	139	<19.1		
1+	36	17.1-25-5	20.88	2.49
2+	7	28.0-33.0	31.14	1.58

N, sample size; *s.d.*, standard deviation.

7.3.5. Population genetics

Hardy-Weinberg equilibrium is a requirement that must be fulfilled before most other tests and analysis applied are valid. That is why, tests to assess Hardy-Weinberg conformity between the fish caught in different sampling years and between cohorts of fish from the same location were performed before any pooling of samples.

The probability test used, calculates the probability of a given data set under the null hypothesis of random union of gametes. This test causes the rejection of the null hypothesis, when the data set is among the least probable outcomes.

Carvalho and Hauser (1995) propose that the generally good fit to Hardy-Weinberg expectations provide indirect evidence for lack of selective forces affecting allozyme markers. Lessios (1992), however, thinks that it is also possible that factors that are causing deviations from expectations pull genotype frequencies in opposite directions, so that the end result is non-significant differences between observed and expected genotypes.

It can be observed that some of the probability values in Tables 22 to 25 are equal to 1 for Hardy-Weinberg tests, even though the single locus population F_{IS} values are not zero, thus indicating that the observed heterozygosities are not equal to the expected heterozygosities. The probability values should then be less than one. The explanation for the probability values found lies in the very low level of polymorphism at the loci in question.

Temporal stability

It can be said that the non-significant results in tests for Hardy-Weinberg expectations between sampling years (Table 21, page 94), combined with no significant differences between the different cohorts (Tables 22 and 23, page 95), suggests little temporal fluctuation, providing some, although weak, evidence that selection is not operating at these loci, within this species, at the level of resolution one can have with this type of marker. This was supported by the neutrality test performed, that showed *AAT-3** and *SOD** to be selectively neutral. This information is important to take into account, although many studies do not do so. Child (1992), however, performed $\chi 2$ tests within and between 3 age groups (juveniles, <32 cm; adolescents, 32-42 cm; and adults, >42 cm), and found no significant differences.

Polymorphism and heterozygosity

Most allozyme studies that have sampled seabass from both Atlantic and Mediterranean origin, have found heterozygosity values higher in the Atlantic and lower in the Mediterranean (Table 37). Interestingly this same trend is also observed when the sample from Faro, is analysed apart from the other samples (values in Table 37). This sample is the most likely to suffer past or present Mediterranean influence due to geographical proximity. Of course sample sizes can influence these results, and it is important to keep in mind that the sample from Faro was considerably smaller (average number of individuals over loci = 44.6) than the other samples in the present work. On average higher heterozygosity values may be explained by a larger population size and a more efficient mixing of individuals in the Atlantic or in other words higher gene flow.

Heterozygote deficits are, however, observed in many samples, independently from their origin (see also Table 37), although many of those deviations are not statistically significant. Generally, Hardy-Weinberg deviations based on heterozygote deficits, can be explained by one of a number of reasons. One is that the deficits are due to the Wahlund effect, meaning that the samples may contain individuals from one or a number of sub-units with different allele frequencies. The seabass has been shown to be a somewhat fragmented species in genetic terms, possibly more than might be expected, it being a marine fish with pelagic eggs and a migratory behaviour. Another possible explanation is the lower fitness of heterozygotes, Allegrucci *et al* (1997) detected evidence of selection over loci by showing differences between "lagoon" and "marine" genotypes. It was not possible to observe this in the present work, since all samples were contained juveniles caught inside estuaries or lagoons. The neutrality test performed indicates that the data set does not reveal any effect of selection, and thus the loci can be considered as neutral. Allegrucci *et al.* (1997) found six of the loci showing consistent allele frequency differences between marine and lagoon samples *(AH*, FBALD*, CK-4*, G6PDH*, MEP*, PNP*)* and four others contributing to most of the population differentiation (*ADA*, ADK-4* EST-2*,GDA**). From all these only ADA* was also found to be polymorphic in the present work, although it was not contributing significantly to the differentiation observed.

Table 37. Mean observed ($\overline{H}_{obs}$) and expected heterozygosity ($\overline{H}_{exp}$) and polymorphism (P) values from seabass allozymes surveys.

Origin	$\overline{H}_{obs}$	$\overline{H}_{exp}$	$P_{0.95}$ (%)
Atlantic			
Benharrat (1984)	0.052	0.055	17.6
Allegrucci (1997)	0.112	0.112	32.1
Martínez (1991)✦	0.017	0.020	7.1
Child (1992) †	0.024	0.025	10.35
Mediterranean			
Benharrat (1984)	0.012-0.027	0.021-0.028	8.8-11.8
Allegrucci (1997)	0.077-0.130	0.089-0.145	25.0-35.7
Martínez (1991)	0.002	0.002	0.0
Cervelli (1985)	0.023	—	5.9
Present study			
All samples	0.012-0.027	0.014-0.026	5.3-7.9
Only Faro	0.012	0.014	5.3
All but Faro	0.019-0.027	0.020-0.026	5.95

✦Estimated from raw data, pooling Atlantic samples. † Data not presented in original work.

Population differentiation

The results show significant sub-structuring of juvenile seabass populations along the Portuguese coast, based on the fact that the sample from Faro, specially at the *AAT-3** locus is very different from those samples further north. It is therefore important to note that three of the six allozyme loci studied in the present work are rarely polymorphic and most of the differences observed depend, as said, on a single sample from Faro (and on a single locus *AAT-3**) which was compromised by poor storage. It should thus be acknowledged that the sampling regime and the sample quality jeopardise in part more substantiated results from this work.

The sub-structuring detected is shown by single locus pairwise tests in which Faro was significantly different from all other sites at *AAT-3** and by the significant F_{ST}

values for the pairwise comparisons between Faro and Foz and Faro and Milfontes. Reanalysis of the data without the Faro sample show that there was still significant allelic variation between the remaining four sites at the *SOD** locus. The allele frequency of the most common allele at *AAT-3** in Faro (0.92) was more similar to the average observed by Benharrat *et al.* (1984) in their Mediterranean samples. The frequency of the common allele at the other Portuguese sites in this survey ranged between 0.715-0.789 which is very similar (0.76) to that observed by Martínez *et al.* (1991) in Cadiz, a site closer to the Mediterranean than Faro.

The Faro sample came from a single collection in 1992; subsequent attempts to collect more fish in this area met with no success. The evidence from other Portuguese sites is that there was temporal stability in allele frequencies between different cohorts at a site and/or annual samples taken from the same site which gave no *a priori* reason to question the Faro result. Although large changes in allele frequency have been observed by Allegrucci *et al.* (1997) in some enzyme systems implicated in the adaptation of fry to fresh water in lagoons, *AAT-3** is not one of the locus implicated. The lower frequencies of the common allele in the populations on either side of Faro suggest that a more extensive sampling of this vast lagoon system is essential to confirm whether the observed frequency is temporally stable.

In highly mobile species such as the seabass (Pawson and Pickett, 1990), in an area with no obvious physical barriers to movement, gene flow between adjacent localities could be high due to both active movement of adult and juvenile fish and the passive movement of fertilized eggs and larvae. The level of differentiation (F_{ST}) is not high and does not exceed that expected for marine species with high dispersal capabilities which usually show a range of F_{ST} values from 0 to 0.028 (Waples, 1987). Previous work on allozymes has shown multilocus θ values for the seabass as high as 0.34 (based on 18 polymorphic loci) in widely distributed Mediterranean populations (Allegrucci *et al.*, 1997), a high value given the dispersal potential of the species. Our results, however, show that only around 1.1% of the overall genetic variation was due to differences between populations. Pairwise θ values between all sampling sites ranged from 0-0.053 (see Table 33), much lower but from a much smaller geographical range than those observed for Mediterranean populations (0.068-0.565) by Allegrucci *et*

al.(1997). This suggests that the populations of seabass are probably not panmictic, but rather a dynamic mosaic of subpopulations.

The estimated average number of migrants (15) exchanged per generation was very small and likely to be a tiny proportion of the population size of this species, although no estimate is available. Any exchange is probably one-dimensional, with fish moving between sub-populations/spawning aggregations spaced along the coastline, exchange being more common between neighbouring rather than distant sites. However, the correlation between geographical distance and genetic distance was not positive, and maybe due to the micro-geographic nature of the present work. In fact, the distances are quite small and the seabass has a migratory potential of several hundred km, which can mean that although the more common exchanges are between neighbouring sites, the more distant sites could still be well within the dispersal capabilities of the species. This estimated level of migration, although theoretically ensuring all alleles will be shared by all sub-populations over time may be totally inadequate to maintain identical alleles frequencies in dynamic natural populations. Wright (1969) points out that identical allele frequencies cannot be maintained simply by the exchange of one or very few migrants per generation and that a small F_{ST} value can appear among subpopulations with significant amounts of interchange.

One reason for studying these Portuguese sites is that they lay between the nominally reproductively isolated populations studied by Martínez *et al.* (1991). The Tinamenor sample of Martínez *et al.* (1991) was thought to originate from local wild caught but farmed fished, this though is now suspect in the light of comparisons with Portuguese samples to the south and those of Benharrat *et al.* (1984) to the north at Concarneau. The polymorphic *LDH** loci observed in the Tinamenor sample are not seen in any other Atlantic samples but are common in Mediterranean samples (Benharrat *et al.*, 1984). Without this Tinamenor sample, the other Atlantic sites become much more similar and share common polymorphic loci (*AAT-3**,*GPI-1**, *GPI-2** and *SOD**). The question on the origin of the sample analysed by Martínez *et al.* (1991) remains unsolved. Allegrucci *et al.* (1997), using cellulose acetate gel electrophoresis, identified 13 polymorphic loci (but did not include *SOD** and *AAT-3** was monomorphic) in a sample of Portuguese bass from Aveiro. In that sample, both

*GPI-1** and *GPI-2** were polymorphic, contrasting with the majority of Mediterranean samples were these loci are generally monomorphic. In fact, out of 11 different Mediterranean sampling sites, only in one sample from the Gulf of Lion was *GPI-1** polymorphic and one sample from Sardinia was *GPI-2** polymorphic. Also Benharrat *et al.* (1984) only found *GPI-2** to be polymorphic in the Atlantic sample. Child (1992) too found *GPI-1** and *GPI-2** to be polymorphic around the British Isles. One can, therefore say that other loci such as *AAT-3** and *SOD** are commonly polymorphic in Atlantic samples, although not always (*AAT-3** was monomorphic in Allegrucci *et al.* (1997) work, as said before).

In the present study, 3 of the most important systems for differentiating Atlantic from Mediterranean populations, found by other authors, namely FBALD, IDHP and MEP systems, were not used because of the inconsistency of results and poor resolution. It would be quite important to improve the running conditions and/or change from starch to cellulose acetate to improve resolution.

In the seabass literature, two genetic distance indexes have been used based only on polymorphic loci, Nei's (Nei, 1972) and Reynolds' (Reynolds, 1983), and although the Reynolds distance is more appropriate for populations weakly differentiated, both indexes have been computed for comparison with previous works. It can be seen (Table 38) that most estimations have low values, independent of the index used, as might be when dealing with samples of a marine fish species. However, the values obtained by Allegrucci *et al.* (1997) are strikingly different from this trend. In fact, although both indexes are concurrent the maximum values reached are high when compared with all other published works. The genetic distance between Aveiro and the Mediterranean samples was D=0.236 (Nei's distance), of the same order of magnitude observed between different species of fish. This is in contrast to another study that compared Atlantic and Mediterranean bass directly (Benharrat *et al.*, 1984) which used starch electrophoresis and obtained estimates that ranged from D= 0.0101 to 0.0152. These differences may be due to the higher resolution power of the electrophoresis technique used, giving higher polymorphism and heterozygosity values (see Table 37). It is interesting to see that between Atlantic and Mediterranean samples, the genetic distance values are around 10 fold the values from the Mediterranean samples alone. In this

study the geographic proximity of the sampling sites and the nature of the fish being dealt with would allow us to predict that the genetic distance values would be at the lower end of the range of values estimated up to now. This was indeed confirmed and moreover the genetic relationships among sites could not be explained through the geographical distance between sampling locations. The distance coefficients, which assume that genetic drift is the predominant structuring agent, clusters Foz, Óbidos and Milfontes closer together, and sets Faro apart from all other samples.

The present study suggest a stable population structure in juvenile fish, which may or may not be a true reflection of what is happening in the adult spawning stock. Seabass spawning depends on temperature with more southerly populations tending to spawn first, and as the season progresses the spawning peak moves towards the north (Kelley, 1988b; Jennings and Pawson, 1991). Asynchrony in spawning time may well be one factor that would help to maintain differentiation between sub-populations by reducing the degree of possible overlap (Richardson, 1983; Milton and Shaklee, 1987). Seabass fry enter the lagoon systems and estuaries along the Portuguese coast 2-3 months after spawning has occurred offshore. Although the fry appear to enter the nursery area quicker compared to other areas, the time elapsed would still allow a large amount of mixing and dispersion to occur to the eggs and larvae. The results from Allegrucci *et al.* (1997) suggest it may not be realistic to expect all fry to enter lagoons. We may have, as in the Mediterranean, an adult spawning stock comprising freshwater and marine reared juveniles with very different allele frequencies. The apparent stability observed in lagoon populations may be related to adaptation to this environment, though the hydrographic data indicating fluctuations in salinity make this unlikely.

A more detailed study is required to examine the relationship between the juveniles and nearby spawning aggregations and to assess the level of migration of juveniles and their recruitment to the adult spawning population. Such studies on adult seabass using allozymes would be expensive because of the numbers required and the fact that the fishermen are unwilling to allow tissue samples to be removed from these highly priced fish before they go to market.

Table 38. Comparison of Nei's (Nei, 1972) and Reynolds' (Reynolds, 1983) distances of seabass samples, based on polymorphic loci only.
Range of minimum and maximum values, given when available.

Samples		Distance	Reference
Tinamenor-Cadiz	*Nei*	0.012	(Martínez *et al.*, 1991)
Mediterranean	*Nei*	0.0011-0.0016	(Benharrat *et al.*, 1984)
Atlantic-Mediterranean	*Nei*	0.0101-0.0152	(Benharrat *et al.*, 1984)
Aveiro - Mediterranean	*Nei*	0.236	(Allegrucci *et al.*, 1997)
Atlantic	*Nei*	0.0003-0.0089	Present work
Atlantic	*Reynolds*	-0.0019-0.0541	Present work
Mediterranean	*Reynolds*	0.040 – 0.565	(Allegrucci *et al.*, 1997)

The use of highly polymorphic microsatellite loci, as it will be seen in the next chapter, means that non-disfiguring biopsies (scales, blood) can be removed from commercial catches without the need to buy fish, dramatically reducing the cost of any study on adult populations.

PART 4 - MICROSATELLITE VARIATION

8. MICROSATELLITE DNA

"Questions about reproductive success, levels of population variation, substructuring and hybridization are currently being addressed through the study of microsatellite variation in natural populations. The potential of microsatellite loci as precise tools to measure genetic variation and gene flow in natural populations seems considerable" (Bruford and Wayne, 1993).

8.1. INTRODUCTION

Three different classes of repetitive and highly polymorphic DNA have been traditionally recognised: satellite; minisatellite and microsatellite DNA. Minisatellites and microsatellites also commonly referred to as Variable Number of Tandem Repeats sequences, or in a more concise way, VNTRs (Nakamura *et al.*, 1987).

Historically, the term satellite DNA was applied to a satellite band in nucleic acid material recovered at specific densities in Caesium chloride gradients spun to equilibrium (Kit, 1961). It turned out that the density difference reflected a difference in composition. These bands had a higher than usual content of G+C bases. In fact, it was shown that the bands were composed of highly repetitive DNA sequences and that these were, for the most part, localised to heterochromatic or centromeric regions of the chromosomes (Purdue and Gall, 1970). Nowadays, the term is applied to any repetitive DNA sequence with a monomer unit of a few to several thousand base pairs (bp) that is tandemly organized in a head to tail fashion, regardless of whether it is separable in a buoyant density gradient analysis (Singer, 1982).

Minisatellite DNA (Jeffreys *et al.*, 1985), are units of 9-65 bp, frequently G rich, with specific alleles varying in size. Some minisatellites display the greatest variability known for DNA sequences, and therefore extremely high rates of mutation can be assumed with heterozygosities approaching 100% (Wright, 1993). There is evidence suggesting that minisatellites may have evolved from microsatellites (Wright, 1994). Multilocus DNA fingerprinting techniques allow the visualization of many loci at the

same time, resulting in a highly specific (at the individual level), but very complex bar code type of pattern. This technique is extremely powerful for detecting individual variation, but it is not optimal for population level studies or for establishing pedigrees. Single-locus fingerprinting, on the other hand, is simpler and more powerful because only one locus is scored at a time. Each individual shows a simpler 2 band pattern, one band inherited from each parent. The genomic organization of minisatellites (see Wright, 1993 and references therein) that have been studied, show that in humans they appear to be clustered in terminal regions, while in mouse they are interstitially distributed. The application of this technique includes differentiating individuals, populations and pedigree within a closed population.

Microsatellites (Litt and Luty, 1989; Tautz, 1989; Weber and May, 1989), or simple sequence repeats (SSR), are short arrays with repeats one to six bases, between 20 and 300 bp long. More commonly they consist of mono-, di-, tri or tetra-nucleotide repetitions, such as $(GT)_n$ or $(GTA)_n$. Because of the small size they are readily amplified from minute amounts of tissue by Polymerase Chain Reaction (PCR). Allele size of microsatellites can be determined with 1 bp accuracy and are characterized by their number of repeats at the locus analysed (Weber, 1990):

Perfect – alternate tandem repeats without interruption and without adjacent repeats of another nature, e.g. GTGTGTGTGTGTGTGTGTGTGTGT;

Imperfect – two or more runs of uninterrupted repeats separated by no more than three consecutive non-repeat bases, e.g. GTGTAAGTGTGTAAGTAAGT;

Compound – runs of repeats separated by no more than three consecutive non-repeat bases from a run of ≥ 5 uninterrupted dinucleotide or longer repeat length, e.g. GTGTGTGTGTCACACACACA.

Any kind of combination between these families is possible. There is some evidence showing that compound and interrupted loci tend to be less polymorphic (Queller *et al.*, 1993; Estoup *et al.*, 1995). Reviews on DNA fingerprinting of fishes (Wright, 1993; O'Reilly and Wright, 1995) and microsatellite DNA in fishes (O'Connell and Wright, 1997) were produced largely in response to the interest shown in these markers by researchers working in the field of fish population genetics. Early on many of these researchers had the perception that these genetic markers could help to shed

new light on some of the problems being addressed with no success using existing techniques.

The Polymerase Chain Reaction or PCR (Mullis and Faloona, 1987), makes a huge number of copies of DNA fragments from a single molecule of DNA. Because VNTRs are flanked by unique sequences, if these are known, it is possible to design and synthesise complementary primers to these sequences and the VNTRs can then be amplified. Moreover, this technique allows the use of frozen samples and alcohol preserved samples. Today total nucleic acid extractions are undeniably fast and reproducible, and DNA amplification through PCR is now almost a standard procedure. The use of microsatellite markers has become the technique of choice particularly when only small amounts of DNA are available, as it allows the screening of archived collections such as scales, alcohol or even formalin preserved specimens as well as the non destructive sampling of live fish.

The importance of microsatellites as genetic markers lie in the properties they exhibit: amplification by PCR, abundance and distribution, high polymorphism and heterozygosity and simple mode of inheritance are some of the most important ones. Microsatellites are generally more suitable to amplification by PCR than minisatellites, because of their smaller size, although exceptions may exist (McGregor *et al.*, 1996) and new technical developments have also recently enlarged the size of PCR products that can routinely be obtained. The abundance of microsatellites is quite high with an even distribution of repeats in the eukaryotic genomes separated by an average distance of 50-100 Kb (Slettan *et al.*, 1993). In fish, for instance, microsatellites are estimated to occur every 7 kb in Atlantic cod (Brooker *et al.*, 1994) and 11-56 kb in Atlantic salmon (McConnell *et al.*, 1995b). Extensive allelic variation and high levels of heterozygosity have been observed in microsatellites (Table 39), perhaps because microsatellite DNA is non-coding, it may not be subjected to selection pressures, and thus allows a high mutation rate. Allelic variation is thought to be generated by intra-allelic polymerase slippage during DNA replication (Levinson and Gutman, 1987; Litt and Luty, 1989; Tautz, 1989; Bruford and Wayne, 1993). But although the nature and mechanisms of mutations in microsatellites are crucial for the development of statistical procedures for data analysis on the genetic structure of populations, to date, the pattern of mutation is

still poorly understood (Jarne and Lagoda, 1996). Some studies suggest that the most likely process in microsatellite mutation is the stepwise mutation model, in which an allele can change by either loss or gain of a single repeat (Valdes *et al.*, 1993). Others consider that this model incompletely describes the mutational process, suggesting a two-phase model, incorporating infrequent mutational changes of two or more repeat units (DiRienzo *et al.*, 1994) along with single-step mutation.

In terms of mutation rates, estimates can be as high as 10^{-4} to 10^{-3} per locus/gamete/generation (Lehmann *et al.*, 1996). A preliminary estimate of mutation rates in the seabass (García de León, 1995) yield a value of 2.4 to 9.6 $x10^{-3}$ based on 786 larval genotypes at 2 loci (*Labrax-3* and *Labrax-13*), the only known estimate in a teleost. It is generally accepted that microsatellite mutational rates can lie between 10^{-6} and 10^{-4} (Bruford and Wayne, 1993) based on estimates from mice and pigs (Dallas, 1992; Ellegren, 1995). Another advantage of microsatellites is that they follow a Mendelian inheritance pattern. This has been established by breeding studies in several species, like cod (Brooker *et al.*, 1994), threespine stickleback (*Gasterosteus aculeatus*) (Rico *et al.*, 1993) and seabass (García de León *et al.*, 1995).

The fact that these markers have a codominant expression, allows the visualization of all the genotypes, with consequent calculations of observed genotypic and allelic frequencies.

When compared to other DNA markers routine assaying is technically simple and quick, although the initial work of detecting the repeats and designing primers can be laborious and time consuming. Another advantage is that these markers tend to benefit from sequence automation and multiplexing (amplification of multiple loci in the same reaction and running the products in the same gel) (Olsen *et al.*, 1996; Wenburg *et al.*, 1996).

Table 39. Average sample sizes, average number of alleles per locus, number of loci, expected heterozygosity, number of variable loci and Hardy-Weinberg deviation in microsatellite marine fish studies.

(Adapted and expanded from O'Connell and Wright, 1997).

Species name	Species common name	Sample size	Alleles per locus	Number of loci	$H_{exp.}$	H-W deviation	F_{ST}	Reference
Dicentrarchus labrax	Seabass	43	17	6	0.84	Het.def.	0.007	(García de León *et al.*, 1997)
Clupea harengus	Atlantic herring	50	32-49	4	0.910	none	0.035	(Turan, 1997)
Clupea pallasi	Pacific herring	50	33	5	0.889	Het.def.	0.036	(O'Connell *et al.*, 1998b)
Gadus morhua	Atlantic cod	59	41	6	0.898	Het.def.	0.015	(Bentzen *et al.*, 1996)
Gadus morhua	Atlantic cod	60	20	5	0.864	None reported	Not reported	(Ruzzante *et al.*, 1996b)
Scophthalmus maximus	Turbot	49	7	3	0.678	Het.def.✣	0.009*	(Coughlan *et al.*, 1998)
Merlangius merlangus	Whiting	73	20	3	Not reported	Het.def.	0.006	(Rico *et al.*, 1997)
Anguilla anguilla	European eel	50	16	5	0.764	Het.def.	0.006	(Daemen *et al.*, 1996)

✣ Statistically not significant.

* Between wild samples.

Microsatellite primers developed for one species frequently amplify polymorphic loci in related species, avoiding all the primer development process: primers of Atlantic cod (*Gadus morhua*) amplify microsatellites of haddock (*Melanogrammus aeglefinus*) (Brooker *et al.*, 1994), coalfish (*Pollacius virens*) (Brooker *et al.*, 1994) and whiting (*Merlangius merlangus*) (Rico *et al.*, 1997); rainbow trout (*Oncorhynchus mykiss*) primers amplify in chinook salmon (*Oncorhynchus tshawytscha*), sockeye salmon (*Oncorhynchus nerka*) and coho salmon (*Oncorhynchus kisutch*), Atlantic salmon (*Salmo salar*) and Artic charr (*Salvelinus alpinus*) (Bentzen *et al.*, 1991; Morris *et al.*, 1995), for instance.

As with all classes of genetic markers, there are advantages and disadvantages. After referring to the advantages, some of the problems should now be addressed to, namely: sample sizes, potential for scoring error and null alleles.

When microsatellite loci with extremely high numbers of alleles are used, the information provided by them may be of very limited value for population structure analysis. Indeed, given the very high numbers of alleles at microsatellite loci in the majority of fishes investigated to date, some increase in the sample size will be required when one considers the number of alleles relative to sample size of the majority of studies to date (O'Connell and Wright, 1997). However, it may be unrealistic to expect sample sizes to increase in a way that most of the alleles are represented at least 5 times. Published works on microsatellites show that on one hand, most of the alleles occurred at frequencies less than 5%, and on the other hand, the majority of individual populations are described with less than 50% of the possible alleles being observed (O'Connell and Wright, 1997). Some authors take the option of binning the alleles, that is grouping alleles within a certain size class (Bentzen *et al.*, 1991), but of course that this method leads to a loss of information and is very subjective as to the classes' establishment.

Dinucleotide microsatellites are prone to stutter, and this can block the scoring of adjacent alleles. In population genetics this can *per se* lead to an excess of homozygotes that is completely artificial. This may be overcome either by i) using tri or tetranucleotide repeats, which are easier to score, as they are much less prone to stutter

and also because most adjacent alleles will be at least a 4 bp apart or by ii) using dinucleotide loci with a small product size of less than 120 bp which also tend to have less stuttering and to be more separated in the gel. An automated approach can identify stutter bands on the basis of signal strength.

If a point mutation occurs in the flanking region where the primer binds or large insert/deletion events happen within that same region, one may expect that the primer will not be able to bind to the DNA template, and thus no amplification will take place. The frequencies of null alleles can be calculated based on the frequencies of homozygous null genotypes (individuals for which no PCR product or only extremely weak amplification is generated). These null alleles have had a major impact on the variation patterns observed so far (O'Connell and Wright, 1997). It also seems that there is an increased likelihood that large products will generate more PCR failures, because of poor template quality (O'Connell and Wright, 1997). The occurrence of null alleles can be detected through a Hardy-Weinberg test, that will reveal a higher number of homozygotes relative to those expected.

The first research applications for microsatellites was in the production of genetic maps that allowed the association of defective human genes with phenotypes. A couple of examples attest the importance of these markers in human genetics: diseases such as diabetes (Copeman *et al.*, 1995), Wilson disease (Bowcock *et al.*, 1994), supravalvular aortic stenosis (SVAS) (Olson *et al.*, 1993) and human lipoprotein lipase (LPL) (Wood *et al.*, 1993) were found to be closely linked to specific microsatellite length sequences. Some diagnostic procedures, such as the ones for myotonic dystrophy (Taylor *et al.*, 1989) and for prenatal prediction of spinal muscular atrophy (SMA) (Morrison *et al.*, 1993) are based on microsatellite markers. Additional detection of disease genes for Huntington's disease, facioscapulohumural muscular dystrophy, piebaldism, Hurler/Scheie syndrome, one form of retinitis pigmentosa, polycystic kidney disease (Riess *et al.*, 1994) will be much improved by the presence of 60 closely related microsatellites on chromosome 4, where all those genes were already shown to be localised. It appears likely that repeat sequence mutation is a common source of human

disease - particularly for those disorders which follow dominant inheritance (fragile X-syndrome, DM - myotonic dystrophy) (Richards and Sutherland, 1994).

There is the possibility that microsatellite markers are either directly involved or closely linked to defective genes. In fact, all markers are potentially subject to "hitchhiking" selection (Ferguson, 1995). The view that microsatellites are selectively neutral may not be an absolute after all. A recent view is that there is increasing evidence that most "neutral" genetic markers, including allozymes, mtDNA and microsatellites, are subject to varying amounts of selective constraints (Stepien and Kocher, 1997). This, however, does not diminish its application in population genetics (Ferguson, 1995), as selection coefficients of 1 – 5 % will not be detectable in the framework of most studies. In fact, Ferguson (1995) is of the opinion that "in terms of finding markers for distinct populations, polymorphisms subject to selection may be of more value". This author explains quite simply that a selectively neutral allele needs 40 000 generations to become fixed in a population of effective size of 10 000 individuals, whereas an allele with a selective advantage of only 5%, only needs 396 generations.

8.2. APPLICATION OF MICROSATELLITES TO AQUACULTURE AND FISHERIES RESEARCH

With the development and implementation of new techniques for analysing DNA in a direct way, the time when enzyme electrophoresis was the only or the state of the art technique for inspecting genetic variation is ending. This does not mean that allozymes are obsolete, the technique has still advantages for preliminary screening of genetic variability, interspecies differentiation and phylogeographic studies. They have proved to be advantageous in detecting inter-population differences in allozyme frequencies, but have not been useful in detecting fixed or high-frequency diagnostic alleles (Ferguson, 1995). However, some molecular tools also present major limitations, for instance the tedious methods used to isolate the mtDNA molecule (Chapman and Brown, 1990), and lack of variability shown by the D-loop and the Cytochrome *b* mitochondrial gene in salmonids (Ferguson *et al.*, 1995), or the limited heterozygosity of restriction fragment length polymorphism for a particular locus (presence or absence). Other approaches involving direct assessment of DNA variation are known to

offer a further degree of accuracy to the field of stock assessment, fisheries and aquaculture and have already been mentioned in the population genetics chapter. The establishment of precise goals for the work to be carried out determines the appropriate technique to be used. The availability of more sophisticated technologies does not necessarily imply that these will suit all genetic research work.

The characteristics of microsatellite make these markers particularly valuable for stock discrimination and populations genetics and for pedigree analysis in aquaculture management programs. This type of marker has been used as a sensitive indicator of population structure in insects (Hugues and Queller, 1993), mammals (Taylor *et al.*, 1994) and fish (for a review see O'Connell and Wright, 1997). In fact, microsatellites are likely to revolutionise selective breeding, as it allows communal rearing of different genotypes without any physical tags, reducing greatly the impact of environmental variability which can affect the estimation of heritability and the effectiveness of family selection. Five hypervariable loci such as minisatellites and microsatellites with 10 alleles per locus may generate 500 million possible composite genotypes (Ferguson, 1995). This should cover all possible DNA patterns generated from investigated individuals for parentage analysis. When progeny from 100 rainbow trout (*Oncorhynchus mykiss*) crosses were communally reared for one year, and the smallest and largest fish screened with 4 microsatellites, over 91% of these fish could be traced back to their parents (Herbinger *et al.*, 1996). This example shows the potential of selection programs based on microsatellite data. In fact, in aquaculture, the use of microsatellites for genome mapping and marker assisted selection will greatly increase the efficiency of future breeding programs, for commercially important aquaculture species (Lie *et al.*, 1994).

In general terms, perhaps the most immediate application of microsatellites is in the making of genetic maps. Goodfellow (1993) acknowledges that microsatellite markers are nearly ideal for map construction, not only because there are large numbers of them, showing high levels of polymorphism, but also because they appear to be more randomly distributed than minisatellites, at least in humans. They also appear to be ideal markers for genetic improvement experiments, and other breeding trials involving half-sibs or multifamily rearing because they contain many different alleles so should

maximize the chances of recognizing segregation in any particular individual or family (Goodfellow, 1993). Microsatellites are currently used for mapping applications in fish, such as tilapia, where 30 linkage groups have been reported (Kocher *et al.*, 1998), zebrafish (*Brachydanio rerio*) (Postlewait *et al.*, 1994), Atlantic salmon (*Salmo salar*) (Slettan *et al.*, 1997) and rainbow trout (*Oncorhynchus mykiss*) (Postlethwait *et al.*, 1994). The establishment of detailed linkage maps will provide the grounds for finding the genetic basis of quantitative traits, like growth and disease resistance (Robic *et al.*, 1994). Most of these phenotypes are polygenic, that is, they are under the control of multiple loci. These loci are commonly referred to as quantitative trait loci (QTL). The detection of associations between phenotypes that do not segregate with Mendelian ratios, and RFLP patterns or microsatellite alleles, will indicate that one or more of the polygenes that controls the trait is located in the chromosomal region near the marker. The first recorded QTL in fish was in rainbow trout (Jackson, 1995), for upper temperature tolerance.

To date a number of microsatellites for aquacultured species are available, as is the case of rainbow trout (*Oncorhynchus mykiss*) (Herbinger *et al.*, 1996), seabass (*Dicentrarchus labrax*) (García de León, 1995; Castilho and McAndrew, 1998)] , turbot (*Scophthalmus maximus*) (Coughlan *et al.*, 1998), Atlantic salmon (*Salmo salar*) (McConnell *et al.*, 1995a) and seabream (*Sparus aurata*) (Batargias and Zouros, 1993).

The higher number of microsatellite alleles compared to allozymes and the lower number of alleles compared to minisatellites may make them useful for population studies in some circumstances, for instance, inbred and/or bottlenecked populations and population subdivision where genetic differentiation is limited. However, that situation also poses some problems in terms of statistical analysis, as often the microsatellites hypervariability may hinder the results as large number of alleles: i) create statistical problems in the form of large sampling errors, that can only be attenuated by an increase in the sampling effort and ii) imply higher mutation rates at these loci which may result in convergent mutations, and thus identity of allelic state will not mean identity by descent (DiRienzo *et al.*, 1994).

One of the first species to be studied with microsatellites was rainbow trout (*Oncorhynchus mykiss*) (Nielsen *et al.*, 1994), in a survey comparing the performance

of mtDNA and microsatellites. This work showed similar patterns of differentiation with both markers. The Atlantic salmon (*Salmo salar*) has been extensively studied by a number of markers: allozymes, mtDNA, minisatellites and microsatellites. MtDNA (Bermingham *et al.*, 1991), minisatellite (Taggart *et al.*, 1995) and microsatellite (McConnell *et al.*, 1995a) data revealed significant differences between European and north American Atlantic populations. However, microsatellites allowed further discrimination between Bay of Fundy area and other Atlantic salmon populations within eastern Canada (McConnell *et al.*, 1997). In a more geographically restricted scale, between two adjacent populations from Lac St. Jean, in Canada (Tessier *et al.*, 1995), mtDNA was more diagnostic than any single microsatellite locus, although both mtDNA and microsatellite data showed significant differences.

In population genetic studies, the reports on microsatellites of Atlantic cod (*Gadus morhua*) constitute an example of the usefulness of these genetic markers. There is extensive allozyme and mtDNA data on genetic variation and its application to cod stock discrimination (Mork *et al.*, 1985; Carr and Marshall, 1991). To date this shows no evidence of population subdivision except at the macrogeographic scale, across ocean basins or across deep oceanic channels (Jamieson, 1975; Cross and Payne, 1978). The use of microsatellites, changed this view completely, Brooker *et al.* (1994) identified and characterised a number of microsatellite loci which were used in subsequent studies. It was found a group of microsatellite loci exhibited highly significant differences in allele frequency and heterozygosities among samples taken from different locations in the north west Atlantic (Bentzen *et al.*, 1996; Ruzzante *et al.*, 1996b). Moreover, another study suggests that the genetic composition of the cod stocks in the Western Atlantic remains stable over time (Ruzzante *et al.*, 1996a).

Microsatellite variation has also been useful in showing fine structure not thought to be present in other species. Grant and Utter (1987) had shown macrogeographic differences between Gulf of Alaska and Bering Sea stocks of Pacific herring (*Clupea pallasi*) (Grant and Utter, 1984). However, a recent microsatellite study (O'Connell *et al.*, 1998a) showed clear structure within stocks in Prince William Sound. In turbot (*Scophthalmus maximus*), the use of 3 microsatellite loci, helped to confirm low levels of population genetic differentiation between Irish and Norwegian wild populations

(Coughlan *et al.*, 1998), observed in allozyme studies (Blanquer *et al.*, 1992; Bouza *et al.*, 1995). The analysis of the stock composition of the north Atlantic population of whiting (*Merlangius merlangus*) with 3 microsatellite loci suggests relatively low levels of differentiation among samples (5 samples with 50-87 individuals each) (Rico *et al.*, 1997). However, given that all samples showed a significant excess of homozygotes, with estimated frequencies of null alleles of 0.102, 0.116 and 0.143, at the three loci the authors call for careful interpretation of the results, especially because an unambiguous explanation for the observed excess of homozygotes was not possible. Rico *et al.* (1997) highlight the need for a better understanding of the microsatellite's evolution in natural populations in order for this type of markers to be more useful as a discriminatory tool in fisheries management. In a study of the European eel (*Anguilla anguilla*) 4 out of 5 loci showed no differentiation between samples from Ireland and Italy (Daemen *et al.*, 1996), confirming previous findings using allozymes that eels likely comprise a single panmictic unit (Yahyaoui *et al.*, 1983).

It is possible that the lack of divergence at several microsatellite loci (Bentzen *et al.*, 1996) could have derived through convergent evolution counterbalancing the effects of genetic drift, indicating that these markers may well be less informative for broad-scale geographic surveys (Carvalho, 1998).

In this study the objective is to determine if there is any intraspecific population differentiation, at a regional scale (samples separated by less than 1000 Km). Because of the seabass life-cycle characteristics (with pelagic eggs and larvae) and a migratory behaviour, the genetic differentiation between proximate populations *a priori* might be expected not to show an obvious discrete population structure. However the results obtained from allozymes indicate the possible existence of a degree of population structure. Even in the situation where genetic structure is not discrete, that does not necessarily mean that the population is freely interbreeding, as deficient sampling or limitations of the molecular tool employed may well be responsible (Carvalho, 1998). In the present study, it was decided to further analyse the samples with microsatellites, a tool that has shown in marine fish species, unprecedented high levels of

heterozygosity and polymorphism as shown in Table 39, increasing in principle, the possibility of resolving any genetic structuring that might exist.

8.3. MATERIALS AND METHODS

The initial screening work to find microsatellite loci for the seabass was carried out at the Marine Gene Probe Laboratory, Dalhousie University, Nova Scotia, Canada, under the technical supervision of Doug Cook. For most species microsatellite marker systems can be developed within 4-6 months. Actually, with "beginners luck", all the development was completed in about 14 days. The design of primers and all the routine screening was carried out afterwards in the molecular genetics laboratory at the Institute of Aquaculture, University of Stirling, Scotland. This facility, was being established upon my return from Canada. It was still several months before it was completely functional, with its new thermo-cycler, vertical electrophoresis apparatus and radiation clearance. All detailed protocols and recipes used are in Annex 3. Here is an overview of materials and methods (Figure 23).

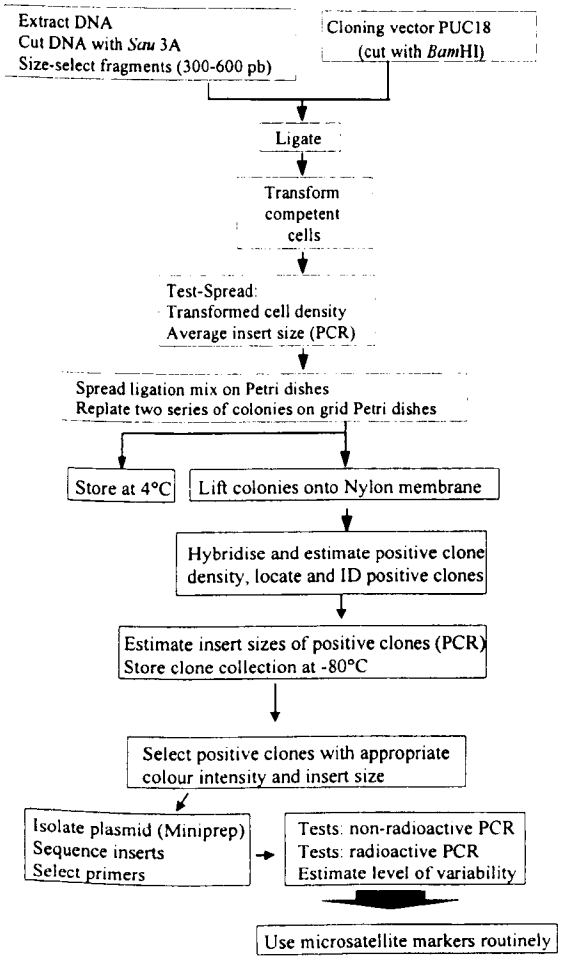


Figure 23. Main steps in the isolation of microsatellite markers. (Adapted from Estoup and Turgeon, 1996).

8.3.1. Source of genomic DNA

The key step in extracting nucleic acids is the removal of proteins. This is achieved first by sample digestion with proteolytic enzymes such as proteinase K, followed by extracting aqueous solutions of nucleic acids with phenol and/or chloroform. The protocol for extraction of total nucleic acids was adapted from Chapman and Powers (1984). A set of well preserved seabass tissue samples, were selected as sources of good quality DNA, a total of 8 liver samples and 5 muscle samples were used.

50 mg of tissue was extracted in a 1.5 mL Eppendorf tube, in 465 µL of lysis buffer TEN (0.1 M Tris, 0.1 EDTA, 0.25 NACl) to which 20 µL of 20% SDS (20%stock), 20 µL of RNAse (10 mg/mL stock) and 20 µL of Proteinase K (stock 10 mg/mL) were added. The SDS disrupts the cell membranes by removing lipid molecules. The High EDTA content removes magnesium ions that are essential for preserving the overall structure of the cell envelope as well as inhibiting cellular enzymes that could degrade DNA. The RNAse breaks down the RNA present in the sample. The contents were homogenized with a micropestle and incubated for 1 hour at 55°C.

Equal volume of phenol:chloroform:isoamyl alcohol (1:1:1/24) (organic solvents that dissolve proteins) was added to each tube (≈500 µl) and each tube was gently shaken. The tubes were then centrifuged at 14.000 rpm for 10 min. The supernatant was transferred to new 1.5 mL Eppendorf tubes. Two volumes of chloroform:isoamyl alcohol (24:1) were added to each tube and gently shaken until the two phases were mixed. The tubes were then centrifuged at 14.000 rpm for 5 min. The supernatant was transferred to a new tube and 1/5 volumes of 5 M NaCl was added. Two volumes of ethanol 99% at 4°C were added and the tube contents inverted. Contents were spun down by centrifuging at 14.000 rpm for 10 min and the ethanol was discarded. The pellet was washed twice with 70% ethanol to remove the phenol. Between each wash, a centrifugation at 14.000 rpm for 5 min was carried out. The pellet was dried in a

vacuum desiccator for about 10 min. TE-buffer (50 µl) was then added to resuspend the pellet, and resuspension occurred overnight at 37°C.

The resuspended DNA was quantified by spectrophotometry. A solution of double stranded DNA of 50 µg/mL has an absorbance of 1 at 260 nm. It follows that it is possible, knowing the dilution factor used and the absorbance of a given sample, to estimate the concentration of DNA. Most of the samples had a DNA concentration between 0.02-3.93 µg/µL, except one sample, from liver, which had 7.99 µg/µL. This sample was chosen as the source for genomic DNA for the construction of the genomic library. This protocol was further used for routine DNA extractions prior to PCR amplification.

8.3.2. Cloning methodology

The protocols used in this work were all based on the protocols used at the time (1994) at the Marine Gene Probe Laboratory. Those protocols were adapted over the years when necessary from Sambrook *et al.* (1989) by students and technicians of that Institution. This work greatly benefited from their experience. Detailed protocols as well as recipes for cloning can be found, as said, in Annex 3.

Constructing a genomic library

A genomic library was created by digesting 50µg DNA (6.5 µL in the present case) of one individual with high concentrations of *Sau*3A I endonuclease that recognizes a 4 bp sequence (↓GATC). Fractioning was done by electrophoresis of the DNA in a 1% low melting point agarose gel using a preparative comb (one small well and one big well). All of the digest was loaded on the big well and a standard 100bp ladder was loaded in the small well. The electrophoresis was run at 40 V, for 6 to 7 hours, with 1 x TAE buffer.

Fragments in the size range of 300-600 bp were excised from the gel and DNA was recovered by phenol-chloroform extractions. The pieces of the gel were cut into smaller chunks and placed in several sterile Corex (glass) tubes. These were frozen at -80°C for 2 hours. Three volumes of saturated phenol (3 times the volume of the gel fragment, 1g = 1mL) were added. Tubes were vortexed until the agarose was

solubilized. After that, the tubes were spun for 20 minutes at 14000 rpm. The aqueous phase was removed and placed in a new Corex tube. An equal volume of chloroform was added. The tubes were vortexed and spun for 3-5 minutes at 14000 rpm. The aqueous phase was again removed and placed in a new Corex tube. Sodium acetate was added to a final concentration of 0.3M as well as 3 volumes of cold absolute ethanol were added. The tubes were stored at -80°C for 30 minutes. Finally they were spun for 30 minutes at 14 000 rpm. The supernatant was removed and discarded, the pellet washed with 70% ethanol and air dried. The DNA pellet was resuspended in 500μL TE and transferred to 2mL screwcap tubes and the sodium acetate/alcohol step repeated. The dried pellet was then resuspended in 21μL TE. The assessment of the recovered DNA, was done, by eye, with 1μL of the DNA, on a 1% agarose gel.

Ligation was performed overnight at room temperature by T4 ligase between pUC18 cut with *Bam*H I / BAP (BRL Gibco) and the DNA fragments, using 3 different ratios vector:insert (1:0.5, 1:1 and 1:2). pUC18 was diluted prior to use, to a working concentration of 20 ng/μl. The ligation reactions were stopped by heat (65°C for 20 minutes). A fifth of the total volume of each ligation reaction (5 μl) (plus 1 μl loading buffer) was run on 1% agarose gel and stained in EtBr to ascertain the success of the ligation (by seeing a smear above the pUC18 band). The ligation reactions were stored at -20°C until transformation.

Competent *Escherichia coli* cells, strain MAX EFFICIENCY DH5α (BRL Gibco) were used for the transformation. The reason for the use of these cells was a practical one, as they were in stock, at the MGPL. In conjunction with pUC18 plasmid they allow the selection of recombinant cells by color which is a major advantage. pUC18 is a prokaryotic cloning vector containing multiple cloning sites located within the *lacZ* α-fragment of the cells, resulting in the disruption of β-galactosidase activity by cloned inserts. This enzyme is responsible for converting lactose to glucose plus galactose in the normal bacterium. Enzyme activity can be assayed with a chromogenic substrate, such as 5-bromo-4-chloro-indolyl-β-D-galactoside (X-Gal), which is colorless, but is converted to an intense blue product, bromochloroindole as a result of β-galactosidase activity. As pUC18 also carries an ampicillin resistance gene, non-transformed cells are not able to grow, the non recombinant colonies are blue and the

colonies containing the vector are white. Competent cells were thawed on ice, and 40μL were aliquoted per tube reaction. As a control, to determine transformation efficiency, 2 μL of control DNA (pUC19) was added to a 15 mL Falcon tube. The transformation was undertaken on different concentrations (1 and 2 μL) of the three different insert:vector ratio ligation reactions in 15 mL Falcon tubes. The tubes were incubated on ice for 30 minutes, and heat-shocked for 45 seconds in a 42°C water bath without shaking. The tubes were placed on ice for 2 minutes, after which 0.9mL at room temperature S.O.C. medium was added. The tubes were then incubated at 37°C with shaking at 225 rpm for 1 hour. 50 μL of 2% X-gal solution was added to ready prepared LB-ampicillin Petri plates, while the transformations were been shaken. Two different volumes of transformed cells (100μL and 200μL) were plated (18 + control) and grown overnight at 37°C.

Replica filters or lifts were made from each Petri plate by placing an 82 mm diameter supported nylon filter (Hybond-N-Amersham) on the surface of the Petri plate containing 18 h growth transformed cells. After lifting, each membrane was placed for 7 min on Whatman 3MM paper saturated with denaturing solution (1.5M NaCl with 0.5 M NaOH) at room temperature. Membranes were then transferred to Whatman 3MM paper saturated with neutralizing solution (1.5M NaCl, 0.5 Tris-HCl pH 7.2 and 0.001M EDTA) for 3 min and this step was repeated. Finally, the filters were placed in 2xSSC for 5 minutes, left to air dry and the DNA was immobilized by baking the filters at 80°C for 2 hours

Each membrane was screened by colony hybridization using a radiolabelled oligonucleotide probe: $(GT)_{15}$. 50 mL volume of pre-hybridization solution (50 mL 20x SSPE, 8 mL 50x Denhardt's, 5 mL 20% SDS, 20 mg RNA in 500 mL of distilled water) was performed for 2 hours at 62°C. After this step, the solution was discarded and 15 mL fresh pre-hybridization solution was added together with 10 μl of end-labeled primer (1μg of $(GT)_{15}$, 1 μl of buffer, 2 μl ddH_2O, 30 μCi $[\gamma^{32}P]ATP$ (4500 Ci/mmol; 10 mCi/ml) and 0.5μl T4-PNK at 37°C for 30 min). The hybridization bottles used could hold 8 to 12 filters. The primer labeling was prepared using Polynucleotide kinase from bacteriophage T4 (T4-PNK) to catalyse a reaction type known as forward

reaction, which is particularly convenient for use with chemically-synthesized oligonucleotides. T4-PNK catalyses the transfer of the terminal γ-phosphate group from a ribonucleoside triphosphate donor ([γ-^{32}P]ATP) to the 5'-terminus of the DNA molecule. Hybridization was done overnight at 62°C.

Following hybridization, the solution was discarded and the filters were washed once in 100 mL of washing solution (50 mL 20xSSC and 5 mL 20% SDS in 500 mL of distilled water at room temperature) for 15 minutes and twice in 100 mL of a stock of 5 mL 20xSSC, 5 mL 20% SDS in 500 mL distilled water, for 15 minutes, first at room temperature and finally at 62°C. Filters were then blotted dry and following a 6 hour exposure to X-ray film (Kodak XAR-5) with intensifying screens at -80°C, the signal was obtained.

After aligning the X-ray film to identify the colony or colonies from which the hybridization signal was obtained, a number of colony picks were made from individual colonies. Only recombinant colonies, the white colonies, have a signal, and the selected ones were preferably well isolated from surrounding colonies to ascertain that they correspond to only one colony. On the basis of this, clones were chosen and a pick was made into separate 15 mL tubes containing 5 mL of LB medium with ampicillin (0.5mL of 200mg/mL stock solution). After 48 hours the plasmid DNA from each of the clones was extracted with phenol-chloroform and mini-preparations following the classic alkaline lysis (Sambrook *et al.*, 1989). 5 μL of each sample was run on a 1.0% agarose gel to assess DNA quality and relative quantity.

Manual dideoxy-chain termination sequencing was performed with the T7 sequencing Kit from Pharmacia, according to the manufacturers instructions, using ^{35}S label. Denaturing acrylamide (6% w/v acrylamide; 0.3%w/v bis-acrylamide; 7M Urea; 1X TBE) gels were run on 1 X TBE on a S2 GIBCO BRL, vertical electrophoresis apparatus (gel size 30 x 40 cm), for 3.5 hours at 60 mA. The gel was fixed in a glacial acetic acid: methanol solution: water (1:1:10) for an hour and the gel dried in a Biorad gel dryer for 2 hours before being exposed overnight to Kodak X-omat R film.

Primer designing

Several criteria for primer design are usually accepted:

a) 3'end of primers should not be complementary,
b) if possible 3'end of primers C or G (C-G clamp),
c) approximately 20 bases long,
d) they should have similar melting temperature (2°C maximum difference),
e) when possible avoid internal repeats,
f) product size ideally between 100-150bp including primers,
g) approximately 50% GC content.

The annealing temperature is typically taken to be 5°C below melting temperature (T_m), which in turn was calculated using the following formula: T_m =[(A+T) x 2 + (G+C) x 4].

Two computer programs were used to select PCR primers, PRIMER 0.5 (Lincoln *et al.*, 1991) and OSP (Hillier and Green, 1991) both allowed the introduction of the above mentioned criteria.

From this point forward all work was conducted in the Institute of Aquaculture, University of Stirling.

Polymerase chain reaction (PCR)

The initial evaluation of each primer set was conducted using 30 DNA templates from randomly chosen individuals and the plasmid containing the insert from which the primer sequences were derived.

Reactions were carried out in a total volume of 10 μL using microtitre plates (Costar). Primers were used at a final concentration of 0.6μM. Only 10% of the reverse primer was end-labelled. The labelling reaction was performed with 2 units of T4 Polynucleotide kinase (New England Biolabs) and 5 pmoles of 3.000 Ci/mmol [γ^{32}P]-ATP for 30 minutes at 37°C.

The reaction mixes contained the DNA template (1 μL corresponding approximately to 5 ng), 1.5 mM $MgCl^{2+}$ (1 μL), 0.6μM of forward and reverse primers (0.2 μL), 200 μM of each nucleotide (0.2 μL), 0.5 unit of *Taq* Red hot DNA polymerase (AB Technologies) (0.05 μL), 1x PCR buffer containing 50 mM KCl, 10 mM Tris-HCl pH 8.3 (supplied by AB Technologies) (1 μL). Details of the cycles used are in the results section.

Stop dye (Pharmacia Sequencing kit) 5μl was added to the samples on completion of PCR cycles. The PCR products were then denatured at 94°C for 10-20 minutes and 3 μl of the sample was loaded on a 6% denaturing acrylamide gel. The accurate scoring allele sizes of each of the 60 individuals loaded per gel, were determined by reference to three sets of M13mp18 DNA sequence size ladder (Yanish-Perron *et al.*, 1985), i.e., ACGT (M13) – 10 fish – A (M13) – 10 fish – C (M13) – 10 fish – ACGT (M13)– 10 fish – G (M13) – 10 fish – T (M13) – 10 fish – A (M13) ACGT (M13).

6% denaturing acrylamide gels (gel size 30 x 40 cm) were run at 60 mA for 3 hours. The glass plates were then carefully separated and a sheet of Whatman 3MM paper placed on top of the gel. Because the gel is moist, the paper will adhere to it, and it is possible to lift the gel intact from the glass plate, by pulling the paper off. The gel was then wrapped in cling film, and placed in a film cassette with 2 intensifying screens and exposed to X-OMAT AR film from Kodak for 24 hours at -80°C.

8.4. RESULTS

8.4.1. Microsatellites and primer characterisation

The test ligations provided an adequate number of recombinant colonies, roughly 2750. Sixty clones (2% of all clones) were chosen based on their signal size and colony isolation. The percentage of recombinant colonies was high (90%), only about 10% of the colonies were blue. The library represented over 1.2 x 10^6 bp of the seabass genomic DNA, if the average insert size was 450 bp.

Forward short-run sequencing was performed on 57 clones out of the 60 (Table 40), of which 11 were sequenced well, 29 had bands that were visible but the sequence was not clear enough for primers to be designed. The clones sequencing quality (good readable sequences) varied a lot. The DNA sequence quality from the 60 selected clones was thus classified as good, poor and illegible (Table 40), where poor means that the bands were either too faint or fuzzy, or combined multiple stops (bands at the same position in two or three lanes), and illegible means that bands were present but the resolution was such that only very uncertain band identification could be made. Poor sequencing results using this technique could be caused by a number of factors, e.g., insufficient or dirty DNA template, insufficient enzyme activity, contamination of the sequencing reaction with protein or salt, non denatured sample before gel loading and poor contact between film and gel. Agarose electrophoresis of the template DNA, stained with Ethidium Bromide showed that some of the cloned sequences were at a low concentration which obviously caused some of the problems. Time constraints in this project meant that the optimisation of sequencing reactions was not the priority job only the good sequences were used in developing primers for the project.

In total 32 microsatellites could be identified and classified (Table 41). The average number of repeats was 21.2 ± 10.4. Among these 30 were $(CA)_n$•$(GT)_n$ sequences, one was a $(GA)_n$•$(CT)_n$, and another a $(TGA)_n$ repeat. Assuming that the microsatellites are evenly distributed in the cloned fraction of the genome, and that that fraction is representative of the whole genome, the average distance between neighbouring microsatellites can be estimated by dividing the total length of the screened DNA by the total number of $(GT)_n$ microsatellites isolated. It is estimated that microsatellite GT repeats occur in seabass on average every 41 kb.

The choice of the microsatellites to use in this or any other study, is in practical terms somewhat limited to the criteria used for the primer design. Some flanking regions are too close to the insertion point of the plasmid and thus there are not enough nucleotide bases to design a primer. In other flanking regions it is difficult to get, simultaneously, the right base length range, annealing temperature and CG clamps, for a good primer design.

Two pairs of primers were designed from clones 6 and 11 which were named as *Dla6* and *Dla11*, after the species name and the clone number in which were they were found (Figure 24). Although quite a considerable number of microsatellite sequences were detected, 32 in total, some of the clones would have to be grown again and DNA extracted and/or sequencing reactions would have to be repeated in order to obtain sequences that would offer no doubts as to the base sequence present for accurate primer design. Even so, there is no guarantee that the PCR conditions will meet the requirements for a successful amplification. Because of the costs and time constraints and the need to compare the allozyme data with another set of markers it was decided to use primers that were developed for this species. We were given access to the primers developed, and kindly given, by F.J. García de León and F. Bonhomme from the Laboratoire Génome et Populations, Université de Montpellier II, Montpellier, France (García de León *et al.*, 1995): *Labrax-3*, *Labrax-6*, *Labrax-8*, *Labrax-9*, *Labrax-13*, *Labrax-17* and *Labrax-29*. However, it was *Labrax-3*, *Labrax-8* and *Labrax-9*, that amplified successful. The *Labrax-9* primer set was not used in a population genetics survey on seabass by the same authors (García de León *et al.*, 1997) because of its low variability (F. Bonhomme, pers. comm.). However, this is not the case in the present study.

Table 40. Seabass DNA clones forward short-run sequenced: observations.

Clone	Motif	Number	Type (Weber, 1990)	Repeat	Observations
1					B
2					B
3					B
4	TG	15	Perfect		BD
5	TG	19	Perfect		BD
6	TG	26	Perfect		AD
7					F
8	TG	14,5	Imperfect	(TG)5C(TG)9	AD
9	TG	10	Perfect	(CA)3 CTCAT (AC)10	ADH
10					BE
11	TG	16 and 21	Perfect; perfect		AD
12	AG	10	Imperfect	(AG)2A(AG)7	AD
13	TG	18	Imperfect	(AC)3 AG (AC)7 AT (AC)6	AD
14	TG	10.5 +14	Perfect; Imperfect	(AC)14; (TG)5 G (GT)5	AD
15	TG	26	Perfect		AD
16	TG	14	Perfect		BD
17					BE
18					F
19	TG	32	?		BD
20					C
21	TG	30	Perfect		BD
22	TG	+30	Imperfect		BD
23	TG	40	Perfect		BD
24	TG	14	Imperfect	(GT)8 AT (GT)3 G (GT)3	AD
25	TG	?	Imperfect		BD
26					BE
27	TG	34	Perfect		BD
28	TG	16	Perfect		BD
29					AE
30					BE
31	TG	41,5	Imperfect	(GT)7 GAA (TG)28 AA (TG)4	BD
32	TG	44	Imperfect		BD
33	TG	24	Imperfect	(TG)4 TA (TG)4 TA (TG)4 CA (TG)9	BD
34	TG	?	Imperfect		BD
35					BE
36					BE
37					BE
38	TG	43	Perfect		BD
39					C
40					F
41	TGA + TG	?	?		BD
42					F
43	TG	?	?		BD
44					BE
45					F
46					F
47					F
48	TG	16	Perfect		BD
49					C
50					G
51					G
52					F
53					BE
54					BE
55	TG	17	Imperfect		BD
56					F
57					C
58					F
59					G
60	TG	11	Perfect		BD

A – good sequence, B – poor sequence, C – illegible sequence, D – microsatellite present, E – no microsatellite present, F – no sequence, G – not sequenced; H – microsatellite too close to the plasmid insertion.

The use of these primers also allowed direct comparison with other studies on in the seabass (García de León *et al.*, 1997; Naciri *et al.*, submitted). Characteristics of all the primers routinely used in the present work are detailed in Table 42.

Table 41. Summary of results from 60 clones.

Total number of clones sequenced:	57
Number of clones not assayed	3
Number of clones with poor sequence results	29
Number of clones with no microsatellites:	11
Number of clones with microsatellites:	30
Number of microsatellites found:	32
Average number of repeats:	22.7
Standard deviation of the number of TG repeats†:	±10,83
Number of TG microsatellites:	30
Number of AG microsatellites:	1
Other repeats: TGA	1
Number of perfect repeats:	16
Number of imperfect repeats:	12

† Conservative estimation from the clones where sequencing quality allowed.

The cycling program (Hybaid thermocycler) was optimized to produce the best possible results, consisting of 3 minutes denaturing at 94°C, followed by 30 cycles of:

a) denaturing at 94°C, for 2 minutes,

b) annealing for 45 seconds at 58°C for all primers except for *Labrax-8* which annealed at 52°C and

c) extension at 72°C for 45 seconds.

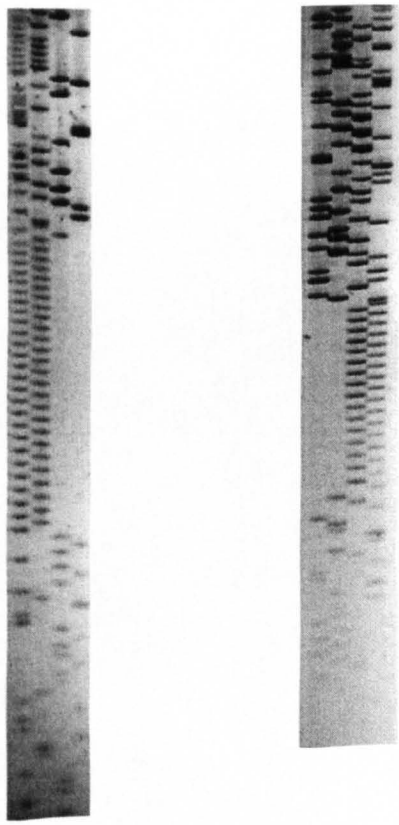
Because of the amplification product size, it was only possible to multiplex *Dla6* and *Labrax-9* on a gel. Full multiplexing (simultaneous amplification and run) between these two primer pairs was not achieved successfully, even though the annealing temperatures were the same.

The sequences of *Dla 6* and *Dla11* have been submitted to the GenBank/EMBL database and have accession numbers Y13158 and Y13159, respectively (Castilho and McAndrew, 1998).

Table 42. Details of the microsatellite loci used.

Primer forward (F') and reverse (R') sequences of the 5 loci, along with their respective annealing temperatures (Ta), expected product size in base pairs.

Locus	Repeat motif	Primer sequences 5'-3'	Ta (°C)	Size
Dla11	$(GT)_{16}$	F' – CACCTCTAATGCTTCCATGC	58	118
		R' – CGAATGCGCTACAAATCTGC		
Dla6	$(AC)_{26}$	F' – AATACGGTGGTGAATCAGTG	58	99
		R' – GCTGTTGTCTTGCTGCATAG		
Labrax-3	$(GT)_{25}GC(GT)_7(AT)_3(GT)_3$	F' – AAACAGTCTTTCAAGTGGTC	58	161
		R' – ATGGACAACTGCTGTCATAG		
Labrax-8	$(AC)_{19}$	F' – TGAGGAAGGTTTGAGAGAC	52	194
		R' – TTCTGCTCCTTAGATGAAC		
Labrax-9	$(AC)_{52}$	F' – TACAGCACCTCTTGAGAAGGG	58	190
		R' – GGCGTACTGCAGGAAAACAG		



Dla 6 *Dla 11*

Figure 24. Partial sequence of clones 6 and 11 containing microsatellites *Dla6* and *Dla11*, respectively.

8.4.2. Genetic variability of microsatellite loci

Only unambiguous bands were scored, and in the case of stutter bands, which overlap between alleles, scoring was based on comparing the intensity and number of stutter bands for each individual at each locus. To ensure consistency in PCR reactions, some samples were run twice. It was not possible to have the microsatellites scored independently by more than one person, however, all individuals were scored twice at different times.

The number of genotypes scored for each sample at each locus are presented in Table 43. If compared with the actual sample size, it can be observed that a number of individuals failed to amplify, either at one locus, or all loci. This observation will be taken into consideration in the estimation of the null alleles. However, it is important to say that *Dla11* failed to amplify in part of the sample from Milfontes and all the individuals in the sample from Faro. It is possible there was a problem with template quality, specifically in those samples, but *Dla11* subsequently failed to amplify in samples that had previously been amplified successfully. A logical explanation would be that for some reason the primers' sequence degenerated, loosing its functionality, however, this pair of primers was kept in the same storage conditions as the other primers and no obvious reason for any change could be found. Although not mentioned in the literature, the fact that some primers stop to function, is not new or unique, but no explanation is found (D. Cook, pers. comm.). The possible solutions for this problem would have involved using new primers, from a fresh batch, as well as sequencing of the templates that do not amplify. This, however, was not possible to do, due to time and financial constraints.

Table 43. Number of genotypes scored in each sample and at each locus.

Locus/Sample	Foz	Aveiro	Óbidos	Milfontes	Faro	Average
Dla11	80	87	56	32	-	64
Dla6	91	59	78	111	17	71
Labrax-3	85	75	83	128	40	82
Labrax-8	74	76	82	115	21	74
Labrax-9	44	75	73	134	16	68
Average	75	74	74	104	24	72

All five microsatellite loci surveyed were highly polymorphic in all seabass populations (Figure 25). All genotypes scored are presented in Annex 4. The number of unique alleles observed per locus ranged from 20 for *Dla11* to 41 for *Labrax-3* (mean ± s.d. = 31.2 ± 8.4; Table 44). Observed heterozygosities were high, ranging from 0.45 for *Dla6* to 0.89 for *Labrax-3* (mean = 0.71).

The allele sizes of all loci varied by increments of two base pairs, except for *Dla6*, which showed occasionally alleles that differed by a single base pair.

The allele frequency distributions (presented in Annex 5) were uni or multimodal, always exhibiting many rare alleles (Figure 26). There were two major patterns that describe the allele frequency distribution of the microsatellite loci studied. *Dla11* showed a single mode, around allele 117. The other loci, show a number of modes separated in some cases by very similar interval repeat number, namely *Labrax-3*, 18-20 repeats; *Labrax-8*, 14-18 repeats; *Labrax-9*, 20 repeats.

Table 44. Repeat structure, size range, observed allele numbers and observed and expected heterozygosity per microsatellite locus.

Locus	Repeat structure	Size range	Allele numbers	Observed/ Expected Heterozygosity
Dla11	$(GT)_{16}$	99-141	20	0.84/0.85
Dla6	$(AC)_{26}$	55-115	28	0.45/0.73
Labrax-3	$(GT)_{25}GC(GT)_7(AT)_3(GT)_3$	116-202	41	0.89/0.92
Labrax-8	$(AC)_{19}$	190-248	29	0.87/0.94
Labrax-9	$(AC)_{52}$	144-254	38	0.52/0.71
Multilocus				0.71/0.83

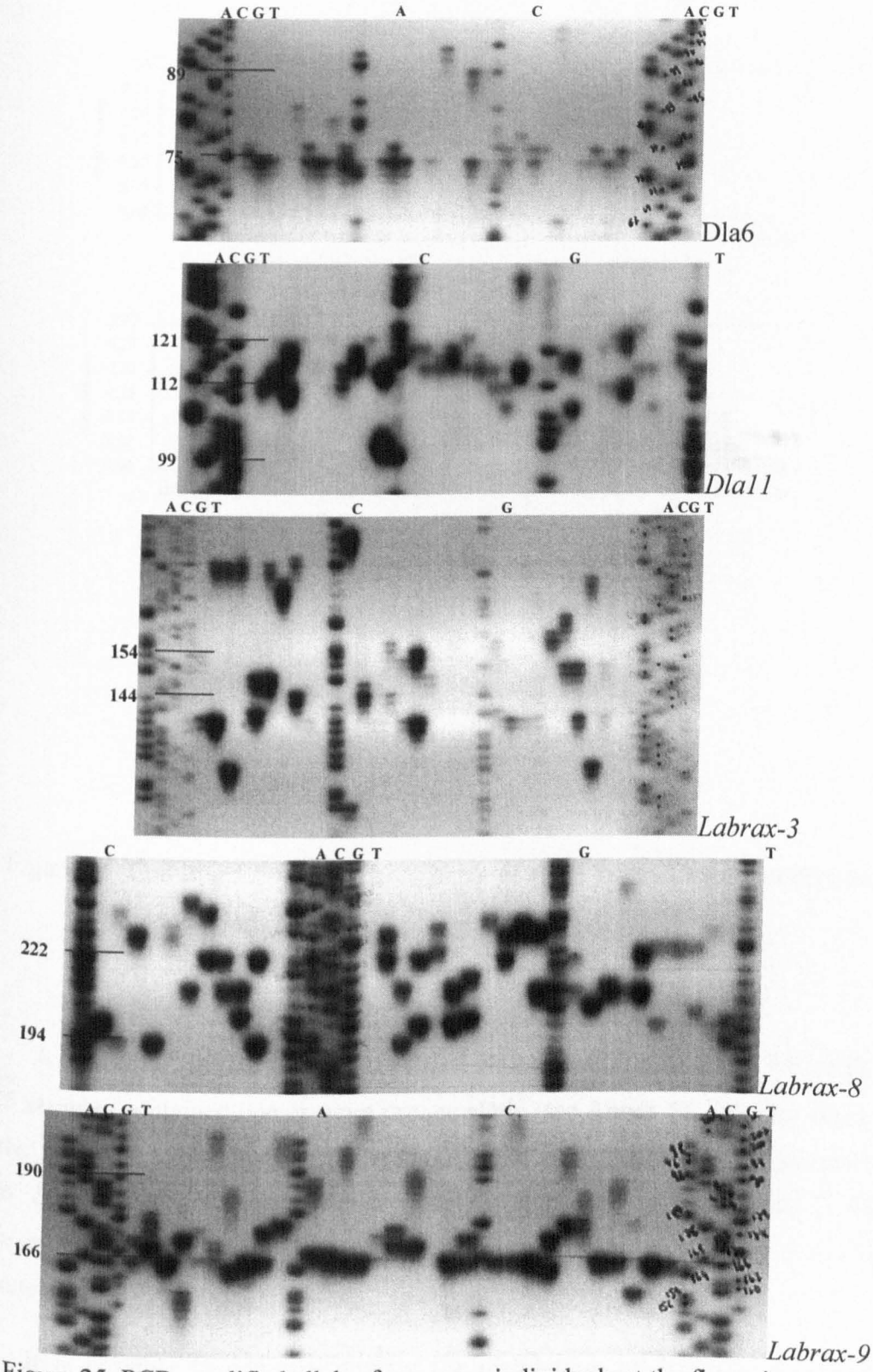


Figure 25. PCR amplified alleles from some individuals at the five microsatellite loci analysed.

The size of the PCR products were estimated by co-migration with those of a M13 clone of known sequence. The sequence ladder is indicated by the letters A, C, G, and/or T. Approximate sizes in base pairs are shown also. Some hand written marks used in scoring can also be seen.

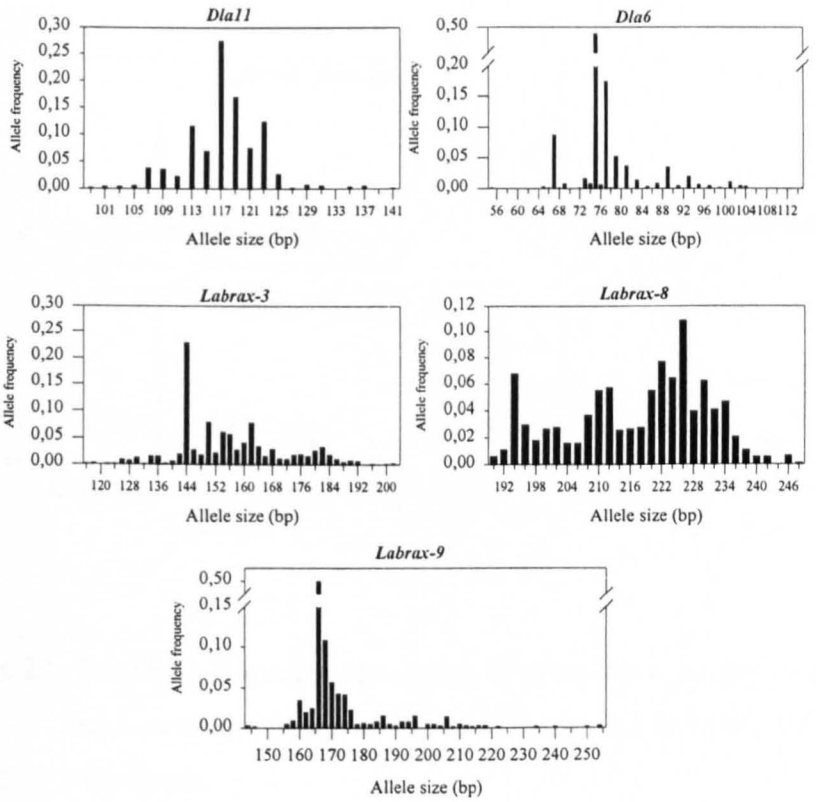


Figure 26. Combined frequency histograms of allele sizes in base numbers when all 5 populations are pooled.

A total of 37 private alleles were found with a mean frequency of 0.0108±0.0075, all 5 samples contained one or more private allele (see Annex 5). Faro has one private allele, whereas Aveiro and Milfontes have 8 and Foz and Óbidos 10 private alleles each. *Labrax-8* was the locus with the smallest number of private alleles, 1, whereas *Labrax-3* had 6, *Dla6* had 9, *Dla11* had 10 private alleles, and *Labrax-9* showed 11 private alleles.

It is possible to compare directly the distribution of *Labrax-3* and *Labrax-8* obtained here and the one shown in García de León *et al.* (1997) (Figure 27) with French Mediterranean samples.

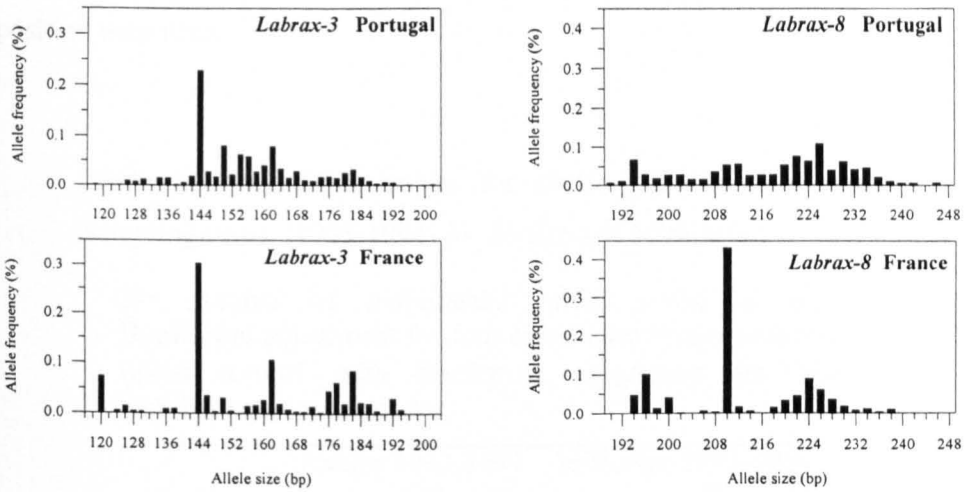


Figure 27. Combined frequency histograms of allele sizes (in base numbers) of the loci *Labrax-3* and *Labrax-8*, in the Portuguese and French Mediterranean populations.

8.4.3. Temporal variability

Aveiro and Milfontes were the two locations sampled in two consecutive years, as it was only possible to obtain one sample from the other 3 sites. An allelic differentiation probability test was performed within each locations, between sampling years, to assess temporal stability of allelic frequencies. No significant differences in allele frequencies were found after Bonferroni adjustment for multiple tests (Table 45).

Evaluation of whether there was variation in allele frequencies among year classes, when cohort numbers allowed (Aveiro and Milfontes), that might confound the interpretation of genetic variation among geographic samples was performed. The modal age components already identified in a previous chapter, showed that several spawning cohorts were present in some of the populations sampled. The cohorts at the same site were tested for temporal stability of allele frequencies. The results (Table 46) show no differences between cohorts at Aveiro and a significant multilocus result

obtained when comparing the cohorts from Milfontes. This is due primarily to the single locus effect from *Labrax-8*. Collectively the results suggest a fairly stable genetic composition over time.

Table 45. Probability values for allelic differentiation across sampling years (1993-1994) for Aveiro and Milfontes.

N= number of individuals; initial α=0.0125 with Bonferroni adjustment for four simultaneous comparisons; initial α=0.01 with Bonferroni adjustment for five simultaneous comparisons.

	Aveiro 1993-1994		Milfontes 1993-1994	
N	119	46	73	140
Locus	*P*		*P*	
Dla11	0.12484		—	
Dla6	0.79953		0.05490	
Labrax-3	0.61102		0.05608	
Labrax-8	0.52633		0.04346	
Labrax-9	0.22604		0.53184	

Table 46. Allelic differentiation probability values between cohorts.

N= number of individuals (Genepop option 3, sub-option 1).

Locus/Population	Aveiro		Milfontes		
Cohorts	91	92	91	92	93
N	10	80	43	19	55
Dla11	0.29950		-		
Dla6	0.92917		0.03565		
Labrax-3	0.56569		0.23097		
Labrax-8	0.76515		0.00874		
Labrax-9	0.57037		0.65709		
Multilocus	0.8912		0.0103		

This single locus result will be kept in mind when analysing the population structure and loci neutrality. For the rest of the analysis, as it has been done by other authors (Small *et al.*, 1998), fish originating from a single geographic location, will be treated as a single sample.

8.4.4. Geographic variation in alleles: size and frequency

The mean number of alleles was very high, ranging from 14.8 in Faro to 23.4 in Milfontes. This is most probably correlated with sample size: the sample from Faro is the smallest one and the one from Milfontes the largest as can be seen from Table 47, there is a trend for larger sample to have a larger number of alleles.

Table 47. Sample sizes, mean alleles number ± standard deviation and observed ($\overline{H}_{obs}$) and expected ($\overline{H}_{exp}$) heterozygosity values (Nei, 1978).

Population	Mean sample size	Mean number of alleles per locus	Heterozygosity	
			$\overline{H}_{obs}$	$\overline{H}_{exp}$
Aveiro	74	19.8 ± 6.0	0.69 ± 0.18	0.84 ± 0.09
Foz	75	20.0 ± 8.8	0.71 ± 0.25	0.83 ± 0.11
Óbidos	74	22.0 ± 7.8	0.74 ± 0.21	0.84 ± 0.09
Milfontes	104	23.4 ± 9.9	0.74 ± 0.21	0.82 ± 0.11
Faro	24	14.8± 10.0	0.67 ± 0.30	0.81 ± 0.16
All populations	341	31.2 ± 8.4	0.71 ± 0.21	0.83 ± 0.10

A more detailed description of variation with allele numbers and observed and expected heterozygosity per locus in each population is presented in Tables 48 and 49.

Table 48. Allele numbers per microsatellite locus in each population.

	Number of alleles					
Locus	Aveiro	Foz	Óbidos	Milfontes	Faro	All
Dla11	14	15	11	8	-	20
Dla6	18	15	17	19	8	28
Labrax-3	28	33	30	32	28	41
Labrax-8	24	25	26	28	17	29
Labrax-9	15	12	26	30	6	38

Table 49. Single locus microsatellite observed and expected heterozygosity in each population.

Locus	Observed/expected heterozygosity				
	Aveiro	Foz	Óbidos	Milfontes	Faro
Dla11	0.80/0.85	0.90/0.86	0.88/0.84	0.78/0.81	-
Dla6	0.54/0.75	0.36/0.68	0.40/0.74	0.42/0.71	0.53/0.75
Labrax-3	0.87/0.91	0.84/0.91	0.92/0.91	0.93/0.92	0.93/0.92
Labrax-8	0.80/0.94	0.93/0.94	0.83/0.93	0.90/0.94	0.91/0.92
Labrax-9	0.45/0.74	0.52/0.75	0.69/0.75	0.64/0.71	0.31/0.59
Multilocus	0.69/0.84	0.71/0.83	0.74/0.83	0.74/0.82	0.67/0.80

The size frequency distributions of microsatellite alleles varied only slightly among the samples (Figure 28a and 28b).

8.4.5. Departures from Hardy-Weinberg expectations

As can be observed in Table 50 when samples are pooled together (multi-sample n= 341) and when all loci are combined (multilocus), HW proportions are not met (Genepop option 1, sub-option 3). *Dla 11* is the only locus not to present significant Hardy-Weinberg deviation at any sampling location, loci *Dla6* and *Labrax-9* are the main loci responsible for this effect.

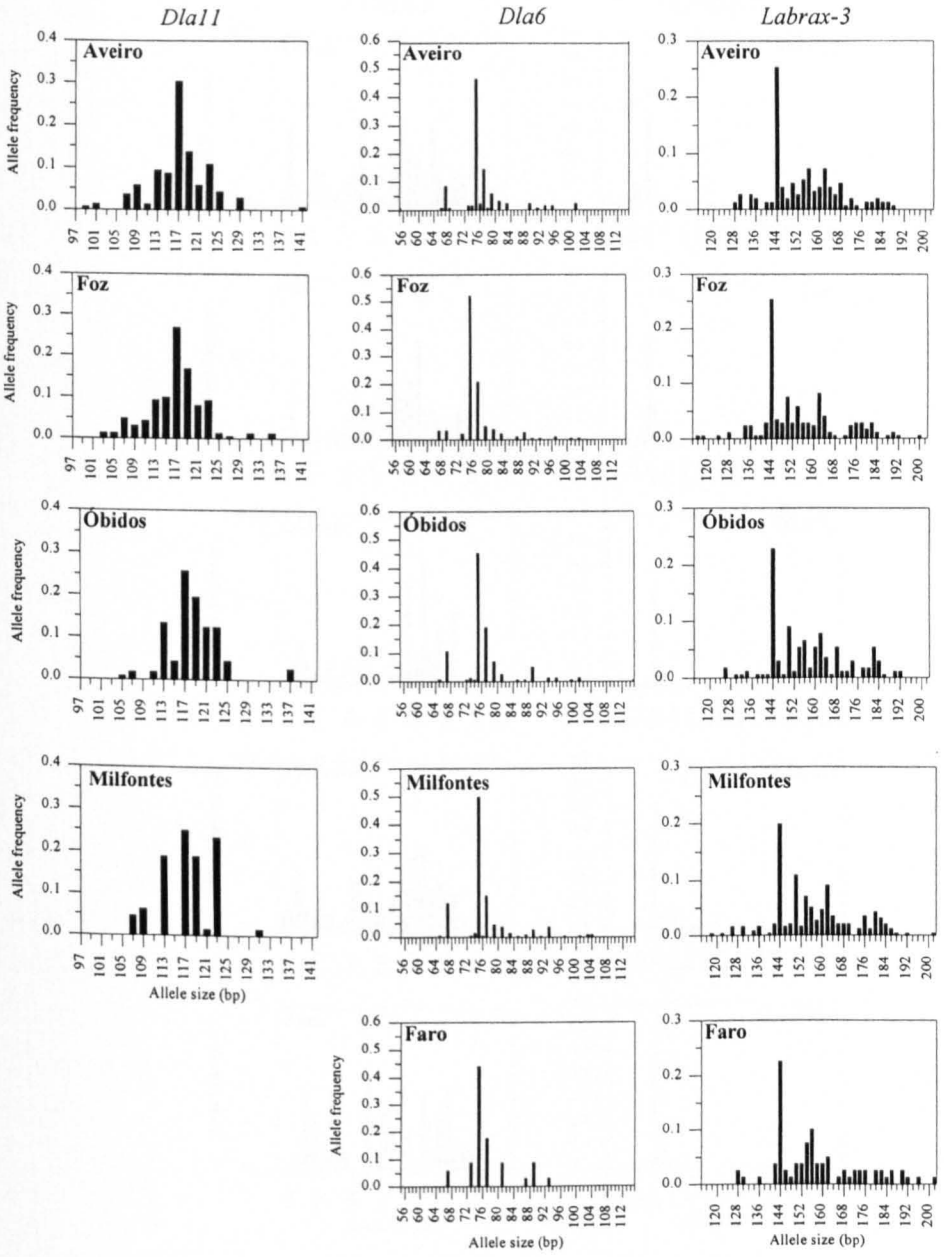


Figure 28a. Histograms of allele frequencies in each sampling location, loci *Dla11*, *Dla6* and *Labrax-3*.

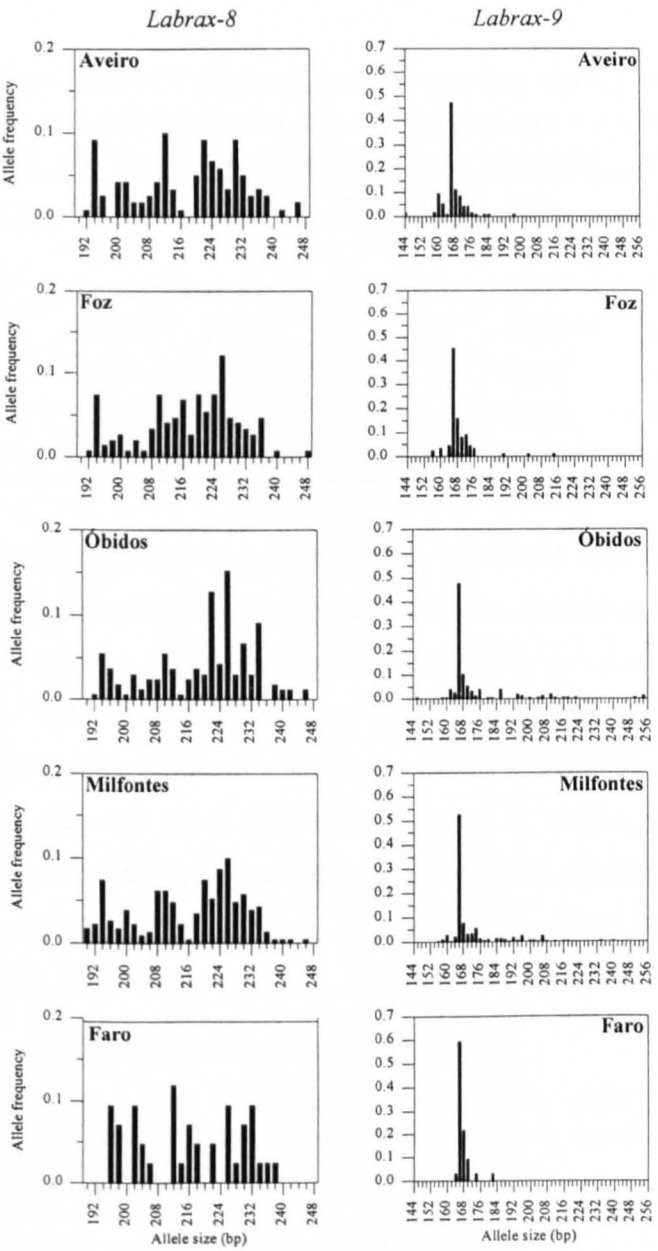


Figure 28b. Histograms of allele frequencies in each sampling location, loci *Labrax-8* and *Labrax-9*.

Table 50. F_{IS} values and statistical significance in each population and in the whole sample.

F_{IS} values estimated by GENEPOP (option 6.1) and probability test (H_0: random union of gametes) performed by GENEPOP (option 1.3).Asterisks indicate the level of significance: *P<0.05 and **P<0.01 after Bonferroni adjustment for simultaneous comparisons.

	Aveiro	Foz	Óbidos	Milfontes	Faro	Multi-sample
Dla11	0.068	-0.042	-0.033	0.046	-	0.004
Dla6	0.281**	0.468**	0.464**	0.405**	0.318**	0.407**
Labrax-3	0.050	0.085**	0.001	-0.006	0.003**	0.025**
Labrax-8	0.157	0.015	0.113*	0.047	0.045	0.074**
Labrax-9	0.400**	0.310**	0.090	0.096	0.493	0.197**
Multilocus	0.170**	0.141**	0.125**	0.109**	0.183**	0.128**

The percentage of hypothetical null genotypes, calculated from the number of individuals that failed to amplify is presented in Table 51. Individuals that failed to amplify at all loci were not taken into consideration, as it is more probable that amplification failure is due to template quality than to other reasons. Thus, in this estimate, it is assumed that the failure of amplification is due to a mutation in the primer region of the microsatellite and not to other factors. As can be observed, the values of hypothetical null alleles are extremely high. Brookfield's (1996) method for estimating null alleles frequencies (Table 52) based on observed and expected heterozygosities also produces high null allele frequencies.

Table 51. Frequency of hypothetical null genotypes.

Calculated from the number of individuals that failed to amplify. † Because of the fact that *Dla11* did not amplify part of the Milfontes sample and all of the Faro sample, no estimates are provided.

Locus/Pop.	Aveiro	Foz	Óbidos	Milfontes	Faro	Average
Dla11	0.214	0.279	0.430	†	†	0.308
Dla6	0.393	0.180	0.170	0.286	0.638	0.334
Labrax-3	0.223	0.198	0.120	0.118	0.064	0.145
Labrax-8	0.232	0.270	0.170	0.230	0.511	0.283
Labrax-9	0.330	0.604	0.260	0.155	0.638	0.397

Table 52. Estimated frequencies of null alleles.

Calculation based on observed and expected heterozygosities (Brookfield, 1996). † Because of the fact that *Dla11* did not amplify part of the Milfontes sample and all of the Faro sample, no estimates are provided.

Locus/Pop.	Aveiro	Foz	Óbidos	Milfontes	Faro	Average
Dla11	0.032	0.000	-0.016	†	†	0.000
Dla6	0.120	0.190	0.195	0.170	0.136	0.162
Labrax-3	0.021	0.037	0.000	-0.005	0.000	0.010
Labrax-8	0.077	0.010	0.052	0.026	0.026	0.038
Labrax-9	0.167	0.131	0.040	0.041	0.186	0.113

8.4.6. Estimation of effective sizes

Calculating the N_e values (Table 53) based on the sample size and the number of alleles, with Ewens (1972) formula the effective population sizes ranged between 15.000 and 17.800. If, on the other hand, we calculate the N_e value based on the expected mean heterozygosity, with Crow (1986) (N_e =(H/1-H)/4μ), the range is between 10.000 and 13.000. Under the stepwise model the values are 3 times larger, roughly between 30.000 and 48.000.

Table 53. Estimation of effective sizes according to different mutation models.

Estimates including all loci based on the lowest (Min. het.) and highest (Max. het) expected multilocus heterozygosity values.

		All loci	
Mutation model	Reference	Min. het.	Max. het.
Infinite allele	(Ewens, 1972)	15.000	17.800
Infinite allele	(Crow, 1986)	10.000	132.100
Stepwise	(Ohta and Kimura, 1973)	30.000	47.600

8.4.7. Population subdivision and gene flow

Multisample Wright *F*-statistics as estimated by GENEPOP (option 6, sub-option 1) and allelic distribution across samples tested with a G-based test results (GENEPOP, option 3.1) are presented in Table 54.

A multi-locus test in the allelic distribution across populations results in a significant global rejection of the null hypothesis, indicating thus heterogeneous distribution of alleles. All loci contribute to this significant result, except *Labrax-3*.

Table 54. Estimated multisample Wright *F*-statistics and G-based test probability of sample allelic differentiation in *D. labrax*.

	F	f	θ	Allelic differentiation	
Locus	(F_{IT})	(F_{IS})	(F_{ST})	Probability	S.E.
Dla11	0.0084	0.0040	0.0044	0.00080	0.00062
Dla6	0.4062	0.4074	-0.0021	0.00022	0.00014
Labrax-3	0.0245	0.0250	-0.0006	0.09190	0.01611
Labrax-8	0.0783	0.0741	0.0045*	0.00000	0.00000
Labrax-9	0.1985	0.1967	0.0021	0.01642	0.00732
All Loci	0.1295	0.1281	0.0016	<0.0001	
♦	0.0379	0.0355	0.0025*	<0.0001	

Multisample Wright *F*-statistics as estimated by GENEPOP (option 6, sub-option 1). Allelic distribution across samples tested with a G-based test (GENEPOP, option 3 sub-option 1), yielding the probability of rejecting the null hypothesis (H_o: the allelic distribution is identical across samples), presented here with associated standard errors (S.E.). All samples are included in the analysis. ♦only *Dla11*, *Labrax-3* and *Labrax-8*. F_{ST} single locus probability estimates are indicated: *P<0.05. F_{ST} probability estimates for multilocus considered significant following sequential Bonferroni adjustment for 6 simultaneous tests are indicated: * initial α (0.05/6)=0.0083.

The multilocus Wright's F-statistics estimated considering *Dla11*, *Labrax-3* and *Labrax-8*, which were the loci that contributed the least to the heterozygote deficit, are: F_{IT} = 0.0379, F_{IS} = 0.0355 and the F_{ST} is 0.0025.

Single locus pairwise population comparisons for genic differentiation produced 18 marginally (prior to correction for simultaneous tests) significant results, which after

Bonferroni adjustment, only 8 remain significantly different (Table 55). These results show that the differences are due to 3 loci, *Dla11*, *Labrax-3* and *Labrax-8* and that most adjacent localities, except Aveiro and Foz, have at least one locus with significant allele frequency differences.

Table 55. Single locus allelic pairwise population comparisons.

Locus	Population pair	Probability
Dla11	Óbidos - Milfontes	0.00282*
Dla11	Foz -Milfontes	0.00743*
Labrax-8	Foz-Faro	0.00007*
Labrax-8	Milfontes - Faro	0.00029*
Labrax-8	Foz - Óbidos	0.00165*
Labrax-8	Aveiro - Faro	0.00289*
Labrax-3	Foz- Milfontes	0.00221*
Labrax-3	Foz - Óbidos	0.00411*

Only significant results after Bonferroni correction are shown; initial α=0.05/6=0.0083 for Dla11 and initial α=0.05/10=0.005 for other loci. Probability values as estimated by GENEPOP (option 3, sub-option 2).

Values of F_{ST} estimated for pairs of populations peaked at 0.005 for Foz and Milfontes (Table 56 below diagonal), although from the 10 possible comparisons, none were statistically significant. Negative F_{ST} estimates may arise between genetically very similar samples. In such cases, F_{ST} values and N_em estimates are meaningless and the infinity symbol was used for the latter in Table 56. The estimated average number of migrants per generation estimated from the multilocus F_{ST} (0.0016) is 156, if estimated by Wright's formula or 100 if estimated by Takahata's. That number is reduced to 100 and 64, respectively, if we count only the loci with lower heterozygote deficiencies (*Dla 11*, *Labrax-3* and *Labrax-8*).

Table 56. Matrix of pairwise F_{ST} and N_em values.

	Aveiro	Foz	Óbidos	Milfontes	Faro
Aveiro	-	∞	417	83	∞
Foz	-0.0009	-	167	51	208
Óbidos	0.0006	0.0015	-	93	500
Milfontes	0.0030	0.0049	0.0027	-	58
Faro	-0.0016	0.0012	0.0005	0.0043	-

F_{ST} values as estimated by GENEPOP (option 6, sub-option 2) are given below the diagonal. The number of migrants, N_em, exchanged under the island model hypothesis, $N_em = (1- F_{ST})/4F_{ST}$ (Wright, 1951), is given above the diagonal. Computation of N_em from negative F_{ST} estimates are represented by ∞. No tests were considered significant following adjustment for 10 simultaneous tests ($\alpha = 0.05/10 = 0.005$) are indicated below diagonal.

The Mantel test (GENEPOP, option 6.5) on a F_{ST} *vs* geographical distance matrix between all population pairs (Table 57) gives a *P*-value of 0.558, with 10000 permutations, meaning that the genetic distance is not positively correlated with geographic distance, and therefore distant populations were no more significantly differentiated than closer ones.

Table 57. Matrix of pairwise F_{ST} and geographic distance.

	Aveiro	Foz	Óbidos	Milfontes	Faro
Aveiro	0	60	155	425	635
Foz	-0.0009	0	95	365	575
Óbidos	0.0006	0.0015	0	270	480
Milfontes	0.0030	0.0049	0.0027	0	210
Faro	-0.0016	0.0012	0.0005	0.0043	0

F_{ST} values are given below the diagonal. Geographical distance in Km.

For comparison purposes, Nei's (1972) genetic distance was estimated as well as Reynolds' (1983) genetic distance as it is thought to be more appropriate in cases of intraspecific divergence (Table 58).

Table 58. Nei's (1972) (below diagonal) and Reynolds' (1983) (above diagonal) genetic distances estimated from original and not bootstrapped data.

	AVEIRO	FOZ	ÓBIDOS	MILFONTES	FARO
AVEIRO	-	-0.00085	0.00061	0.00305	-0.00162
FOZ	0.040	-	0.00152	0.00489	0.00121
ÓBIDOS	0.047	0.046	-	0.00272	0.00054
MILFONTES	0.053	0.058	0.047	-	0.00432
FARO	0.180	0.175	0.186	0.207	-

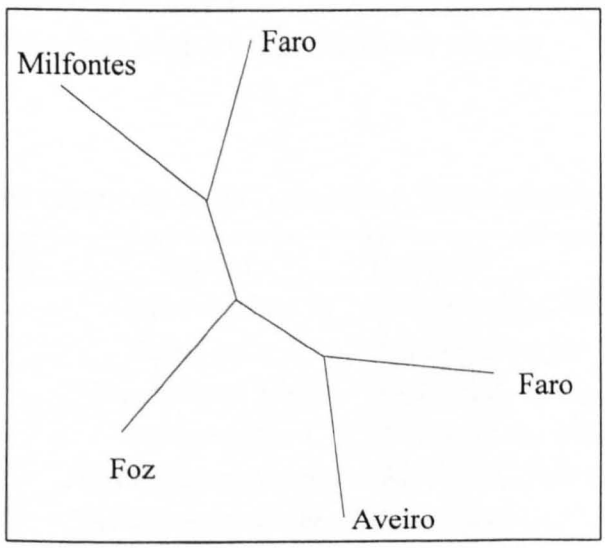


Figure 29. Consensus neighbor-joining unrooted tree based on Reynold´s genetic distance based on variation at 4 microsatellite loci in 5 samples.

Dla11 locus was not included as there are no values for the Faro sample and Phylip does not handle missing values. No percent bootstrap values at nodes are presented, as they were less to 50%.

8.5. DISCUSSION

Under the assumptions that microsatellites are evenly distributed in the cloned fraction of the genome, and that that fraction is representative of the whole genome, the average distance between neighbouring $(GT)_n$ microsatellites in the seabass is in the order of 41 kb. This result being dependent on several assumptions must be treated with caution. However, it gives an indication that microsatellites are highly abundant in the *D. labrax* genome, similar to what has been observed in other teleost species (Table 59).

Full multiplexing was not successfully achieved with *Dla6* and *Labrax-9*, even though the annealing temperature of the primers was the same, probably because of some form of primer-primer interaction. It was not feasible to multiplex any other combination of microsatellite primers as the size range of the alleles would have overlapped. This should not be a problem with the use of florescently labelled primer systems in which different loci can have different colour markers.

Table 59. Estimated occurrence of $(GT)_n$ repeats in teleost species.

Distribution (kb)	Species	Reference
7	*Gadus morhua*	(Brooker *et al.*, 1994)
11-56	*Salmo salar*	(McConnell *et al.*, 1995b)
12	*Brachydanio rerio*	(Goff *et al.*, 1992)
23	*Salmo trutta*	(Estoup *et al.*, 1993)
26	*Merlangius merlangus*	(Rico *et al.*, 1997)
35	*Pseudotropheus zebra*	(Van Oppen *et al.*, 1997)
41	*Dicentrarchus labrax*	Present work
90	*Salmo salar*	(Slettan *et al.*, 1993)

Microsatellite loci *Dla6* and *Labrax-9* gave the smallest diversity values, despite having the largest microsatellites in terms of number of repeats. This is not in agreement with the suggestion by Valdes *et al.* (1993) that microsatellites with larger

numbers of repeats are likely to show the higher diversity values. The populations surveyed in this study revealed very high levels of diversity with heterozygosity values ranging from 0.45 to 0.89. These levels of diversity are similar to those reported for microsatellite loci in other teleosts (Table 39). It should be stressed that the high number of alleles observed, 20 to 41, corresponding in fact to an average of 88% of the potential number of alleles (defined as: allelic range divided by 2). In fact, in *Labrax-8*, all possible allelic stages are filled. This very high number of alleles can generate sampling errors. Some authors consider the binning process as a statistically robust alternative, however it has its drawbacks, such as the subjectivity of the class constitution process. In this work it was chosen not to bin alleles so that direct comparisons could be performed with other works, particularly Garcia de León *et al.* (1997).

Sampling considerations are of great consequence in population genetic studies to provide trustworthy resolution of population differentiation. Particularly in studies involving microsatellites, the number of individuals analysed greatly influences the number of alleles that can be found at each locus, and as we have seen that number can be quite high. Computer simulations, as proposed by Parker (unpublished data in Kornfield and Parker, 1997) can be performed to evaluate the acceptable number of individuals to determine allele frequencies with statistical confidence. This author proposes a process of iterative sampling of alleles with replacement from randomly generated frequency distributions. With that method, sample identity (Nei's unbiased distance) is calculated after each additional allele is sampled. It is then possible to determine the number of individuals beyond which additional sample effort will not bring any more information. At this point, it is not possible to assess if the sample sizes were adequate, and thus the resolution of the sample relationships may be incomplete.

In that regard, it is interesting to observe, that the Faro sample had the lowest number of alleles, on average 14.8 in contrast to 23.4 in Milfontes. The most likely reason is the small size of the Faro sample, as can be observed there is a trend for larger samples to have larger numbers of alleles (Table 43 and 48). Despite the small sample size, it is quite remarkable that the 40 genotypes scored for *Labrax-3* in Faro, had 28 different alleles present, out of a total of 41 alleles scored in all samples for that same locus.

Recently it was proposed (Kornfield and Parker, 1997) that the presence of certain classes of alleles called saltines, may contain important genetic structure information. These saltines are classes of alleles at the extremes of the alleles size distribution and not present in certain samples. They may have be originated through a saltatory mutational process and are suggestive of a founder mutation (Maciel, 1995). Again, sample sizes are of extreme importance to detect very low frequency alleles such as saltines. *Melanochromis paralelus*, a Malawi rocky cichlid fish, known as mbuna, shows a saltine between 150-180 bp at the UME002 locus whereas in the other 5 mbuna species the smallest allele is over 250 bp (Kornfield and Parker, 1997). Inspection of the allele frequency distribution in Figures 28a and 28b and Annex 5, shows no evident saltines, as although there are many private alleles at the ends of the distribution, there is no particular agglomerate of private alleles in any of the samples. Genetic distribution may contribute to shape the distribution of these rare classes, as regeneration of drift eliminated saltines would be extremely improbable.

A highly significant heterozygote deficit was observed for *Dla6* in all samples. All other loci, except *Dla11*, showed a heterozygote deficit in at least one sample. Some authors (Pemberton *et al.*, 1995) consider that larger alleles are more likely to fail to amplify (null alleles). However, *Dla6* is the locus with the shortest amplification products (55-115 bp), the failure to amplify is probably due to other reasons such as point mutation in the primers annealing sequence.

There are numerous works on microsatellites that refer to heterozygote deficiencies, in fish (Table 60) and other organisms. Many workers started to realise that this deficiency is present in many samples, and a number of possible explanations other than null alleles are postulated, among these:

a) biased transmission of these markers from one generation to another, i.e., non Mendelian transmission;
b) Wahlund effect, mix of subpopulations or cohorts;
c) reproductive behaviour conductive to non random mating (natal homing, for instance);
d) inbreeding;

e) selection against certain heterozygotes;

f) mis-scoring of alleles in heterozygous individuals due to stutter bands and

g) homoplasy, given the high mutation rate of microsatellite loci, an individual can be homozygous for a pair of alleles of the same size which may not be the product of a single mutation event.

Some of these possible explanations are discussed next. The non Mendelian transmission of alleles cannot be ruled out *a priori*, but the expected inheritance pattern for 7 seabass microsatellite loci, including *Labrax-3*, *Labrax-8* and *Labrax-9* was shown by García de León *et al.* (1995). The inheritance of other 2 microsatellite loci, *Dla11* and *Dla6*, could not be tested as no facilities for breeding bass were available at Stirling. However, there is no particular reason to assume that *Dla11* and/or *Dla6* likely to be inherited in a non Mendelian fashion. Although null alleles are considered the most likely explanation for these heterozygote deficits it could also be due to population admixture. García de León (1997) deduced from a comparison between the levels of differentiation between samples (F_{ST}=0.007) in the French Mediterranean and the mean F_{IT} value (0.06) that the Wahlund effect could be rejected. This was because it was unlikely that sub-units inside the samples, would create F_{ST} values around 10 times greater than those found between the samples. Using the same reasoning and with an even greater difference between F_{ST}=0.002 and F_{IS}=0.128, makes it very unlikely that subpopulations exist at each site in the Portuguese study.

However, considering the life-cycle and age of the fish sampled, it would not really correspond to population admixture but it could be explained by the occasional visits of young bass to neighbouring estuaries or lagoons. If this was the case then positive F_{IS} values would be expected at a majority of the loci at a particular site, and this is in fact observed.

Table 60. Examples of studies in which a heterozygote deficit was observed.

Organism	Hypothesis	Reference
Gadus morhua	Mix of cohorts	(Ruzzante *et al.*, 1996a)
Gadus morhua	Population subdivision	(Ruzzante *et al.*, 1996b)
Anguilla anguilla	Mix of subpopulation within samples	(Daemen *et al.*, 1996)
Oncorhynchus mykiss	Co-migration of alleles of the same size but with single base substitutions	(Nielsen *et al.*, 1994)
Merlangius merlangus	Null alleles, mis-scoring, homoplasy, mix of subpopulation within samples	(Rico *et al.*, 1997)

Little is know about the mating behaviour of seabass, except that it forms shoals during the spawning season (Pickett and Pawson, 1994) and both sexes seem to shed their gametes in the water column (Barnabé, 1976). Therefore, it is difficult to envisage traits on which assortative mating or inbreeding could be based on (García de León *et al.*, 1997), except possibly homing behaviour. Do bass show a preference for certain spawning sites? It is not known, traditional spawning sites are known but not if it is the same bass spawning in the same place year after year, or if the sites remain more or less constant and adult seabass are hopping between different spawning locations. In fact, if bass shows some degree of fidelity, this could constitute a possible explanation.

Another possibility is the non-neutrality of the microsatellite loci. There is no direct evidence either way for the neutrality or non-neutrality of the microsatellite loci used, but 3 points against selection can be made at this time: a) as dinucleotide repeats they will not be a part of a protein coding region; b) it is statistically improbable that all the 5 loci are involved indirectly in selection through close linkage to selectively important genes; c) the high polymorphism levels of the loci is contradictory to fitness effects at particular alleles. The application of the Ewens-Watterson neutrality test to the present microsatellite data set is arguable as there are great disparities between observed and expected heterozygosities with hypothetical high levels of null alleles.

This would make *Dla6*, *Labrax-8* and *Labrax-9* non-neutral loci, which is probably not true, as more straight forward explanations, can be given for the heterozygote deficiencies observed.

So a number of different reasons can explain the observed heterozygote deficit in the present study, among the most likely: null alleles, migration of juveniles and/or fidelity of adult bass to spawning sites. With the present data set it is not possible to conclude unambiguously which of these possibilities, or which combination of these, fully explains the heterozygote deficit observed.

Significant allele frequency differences between sites were observed at all loci except *Labrax-3*. The separate estimate of multilocus F_{ST}, considering *Dla11*, *Labrax-3* and *Labrax-8*, which were the loci that contributed least to the heterozygote deficit, (0.0025), is higher than the estimate that takes into account all loci (0.0016) and both are statistically significant. These results suggest that the seabass along the Portuguese coast does not constitute a single panmictic population.

From a general comparison of these results with the ones found in the Mediterranean French populations some interesting differences can be seen to (Table 61). In the two loci common to this study and that of García de León *et al.* (1997) (*Labrax-3* and *Labrax-8*), higher number of alleles were present in the Portuguese samples compared to the French samples (see earlier Figure 30): in *Labrax-3*, 41 vs. 30 alleles and in *Labrax-8*, 29 vs. 23 alleles, for sample sizes of 411 in *Labrax-3* and 368 in *Labrax-8* in Portuguese samples and 172 in French samples. It is interesting to compare the distribution of alleles in these 2 loci. The *Labrax-3* distribution is very similar in both studies, sharing the same most frequent alleles: 144, 150, 156, 162 and 182. The exception is the frequency of allele 120 (≈7%) in the French samples and its total absence in the Portuguese samples. In *Labrax-8* the allelic frequency distribution is quite different between Portuguese Atlantic and French Mediterranean samples. While in Portugal no single allele has a frequency higher than 0.11 (allele 226), in France the most common allele has a frequency of nearly 0.45 (allele 210). García de León *et al.* (1997) did not estimate any null alleles frequencies, but they also detected heterozygote deficits in *Labrax-3* and *Labrax-8*.

The comparison that involve F_{ST} and derived values such as N_e and m, is only possible to conduct involving all the loci scored by García de León *et al.* (1997), as it is not possible to obtain the multilocus statistics involving only *Labrax-3* and *Labrax-8* from the French work. For that reason the Table 61 was constructed including the available data from García de León *et al.* (1997), which includes a total of 6 loci, and the present work's data divided in two sets: all loci and loci that did not show single-locus Hardy-Weinberg deviations.

The effective population size estimates over all loci give very similar values, all within the same order of magnitude. The values estimated from the mean expected heterozygosity over all loci are very similar to the ones found by García de León *et al.* (1997), mainly because the expected heterozygosities are similar.

The estimated number of migrants per generation, is 156, if estimated by Wright's formula or 47 if estimated by Takahata's, which is by itself quite high, but it is more important to compare it against the effective population sizes (Table 53). In this case the proportion of migrants is minute compared to the whole effective population, the highest estimate would be between 1.6% (156/10000) and 2.4% (100/4200), but an estimate based on an average $N_e m = (156+100)/2$ and N_e $(28800+175000)/2$ values would be 0.13%. This value is very close to the one found by García de León *et al.* (1997) (0.11%).

These levels of gene flow do not seem to completely prevent population differentiation. In fact, although multilocus F_{ST}'s here estimated are considerably lower than the ones found by García de León *et al.* (1997), they are nevertheless statistically significant. When only *Dla11*, *Labrax-3* and *Labrax-8* loci are considered in the present study, estimates of effective population size are within the same order of magnitude in both studies. The estimation method of the effective population size is based on heterozygosity values and these were very similar in both data sets.

Overall it can be said that the microsatellite markers show high heterozygosity deficit values within samples and results should thus be interpreted with caution. Allelic differentiation across samples is significantly different for 4 out of the 5 loci studied. Fst values are small but significant when estimated from 3 loci that contributed the least to the heterozygosity deficits observed. The results from Mantel test and from the genetic distance based tree did not conform with the model of genetic differentiation by distance at the microgeographic level.

Table 61. Comparative results from the present work and García de León *et al.* (1997) for important parameters in population genetics.

♦ Taking into account only *Dla 11*, *Labrax-3* and *Labrax-8*, which contributed the least for the heterozygote deficit; ✶ an estimate based on an average Wright's N_em and average N_e values; ✢ Values here presented were the lowest *Ne* values estimated from minimum expected heterozygosities using the 3 models (Ewens, 1972; Ohta and Kimura, 1973; Crow, 1986); ▲ Values here presented were the highest *Ne* values estimated from maximum expected heterozygosities using the 3 models (Ewens, 1972; Ohta and Kimura, 1973; Crow, 1986); ♣ Computed from 4 loci (excluding *Dla11*). *$P<0.05$.

Parameters	Present work		García de León *et al.* (1997)	
	All loci	♦		
Sample size	341	288		172
Minimum observed heterozygosity	0.67	0.82		0.77
Maximum observed heterozygosity	0.84	0.92		0.82
Average observed heterozygosity	0.71	0.88		0.79
Minimum expected heterozygosity	0.81	0.90		0.82
Maximum expected heterozygosity	0.84	0.94		0.85
Average expected heterozygosity	0.81	0.92		0.84
Average N_em (Wright)	156	100		37
Average N_em (Takahata)	100	64		11
Multilocus F_{IS}	0.128	0.036		0.046
Multilocus F_{ST}	0.0016	0.0025*		0.007*
Average minimum N_e ✢	10 000	4 200		11 000
Average maximum N_e ▲	47 600	346 000		54 000
Average N_e	28 800	175 000		32 500
Average m ✶	0.13%	0.06%		0.11%
Reynolds distance	*Atla* 0-0.00489♣		Atla-Med	0.0269-0.1091
			Med	0.0000-0.0996

PART 5

9. COMPARATIVE ANALYSIS OF ALLOZYME AND MICROSATELLITE MARKERS

Recent developments of microsatellite DNA markers have raised the question of their relative utility over other markers such as allozyme, mtDNA and minisatellites.

Among the pre-requisites for suitable genetic markers are traits such as codominance, single-locus coding, Mendelian inheritance, unambiguous reading and manageable finances and time. Of course, no single marker will be ideal in all applications and scoring, suitable level of variation, stable expression, not demanding in terms of time and cost, and at some point a choice has to be made. Comparative performance studies are, thus very important, as they will allow some reference towards similar species life cycle, expected levels of variability and so on, and thus can give indications as to the best marker to use in a given situation. Comparative allozyme and microsatellite performance studies, done on the same data set are not abundant in the literature (see Barker *et al.*, 1997; Estoup, 1998).

In the present study, it is relevant to compare the performance of allozymes and microsatellites at the microgeographic differentiation level. Differences in resolution power are expected *a priori*, mainly because of the different mutation rates, higher in microsatellites, resulting in very high levels of variation in microsatellite loci as compared to allozymes. Microsatellites usually exhibit much higher number of alleles than allozymes, with higher levels of heterozygosity. This fact, suggests that, as said in the previous chapter, microsatellites are potentially more sensitive than allozymes to changes in population breeding structure and rate of dispersion.

The goal of this chapter is to compare the level of within and between population variation assessed by the two different marker systems.

To assess the resolution power of both markers used in this study of population structure, one can see differences in the estimates obtained for variability, differentiation and evolutionary relationships.

Within population observed mean microsatellite heterozygosity ranged from 0.40 to 0.93 (average 0.711 ± 0.212) and that estimated from the polymorphic protein-coding loci from 0 to 0.45 (average 0.020 ± 0.005) (Table 62).

Table 62. Genetic variability parameters estimated from estimated from allozyme and microsatellite loci in seabass samples.

H_{obs} = observed heterozygosity, H_{exp}= expected heterozygosity; ♦ Values computed from *Dla11, Labrax-3* and *Labrax-8* loci.

Parameters/type of marker	Allozymes	Microsatellites	
		All loci	♦
Average H_{obs}	0.020 ± 0.005	0.71 ± 0.21	0.83 ± 0.13
Average H_{exp}	0.021 ± 0.005	0.83 ± 0.10	0.88 ± 0.08
Average allele number/locus	1.2 ± 0.1	31 ± 8.4	23 ± 8.2
Average sample size	129 ± 63.5	69 ± 34.8	74 ± 29.3
Reynolds distance	0 - 0.054	0 - 0.005	0 - 0.008

More important is perhaps to look at the expected heterozygosity values at population level as in Table 63, allowing a correlation with sample sizes. It can be observed that the smallest sample (Faro) is also the sample with lowest expected heterozygosity both with allozyme and microsatellite markers. In microsatellites there is definite trend for the more variable samples to also correspond to the larger sample sizes. In allozymes that trend is not observed, which indicate that it is more important for microsatellites to have larger sample sizes in order to identify as many alleles as possible.

Table 63. Expected heterozygosity multilocus values estimated from estimated from allozyme and microsatellite loci in seabass samples.

Samples	Allozymes	Sample size	Microsatellites	
			All loci	Sample size
Aveiro	0.020 ± 0.067	165	0.69 ± 0.18	74
Foz	0.026 ± 0.090	120	0.71± 0.25	75
Óbidos	0.023 ± 0.081	104	0.74 ± 0.21	74
Milfontes	0.021± 0.078	213	0.74 ± 0.21	104
Faro	0.014 ± 0.052	45	0.67 ± 0.30	24

The allozyme loci *ADA**, *G3PDH-2**, *GPI-1** and *GPI-2** are low heterozygosity loci, with an average value of 0.037 ± 0.027, and *AAT-3** and *SOD**, are high heterozygosity loci with an average value of 0.312 ± 0.091. Among the microsatellite loci, *Dla6* and *Labrax-9* showed relatively low heterozygosity values with an average value of 0.485 ± 0.119, and *Dla11*, *Labrax-3* and *Labrax-8* showed higher heterozygosities with average value of 0.872 ± 0.053. The microsatellite loci are 10 times as variable as allozyme loci, when considering only the lower heterozygosity values and over 2.5 times as variable as allozyme loci, when considering only the higher heterozygosities.

Allozyme loci polymorphic at the 99% level provide little information for population genetic studies, while loci polymorphic at the 95% level are informative. All microsatellite loci fall into this category, but sample sizes need to be higher in order to include most of the alleles present in the populations. Most microsatellites should be informative as long as sample sizes are appropriate.

Microsatellite loci showed genotype frequencies that deviate significantly from Hardy-Weinberg proportions, due mainly to an excess of homozygotes. As seen in the

microsatellite chapter, this may have several explanations, but no conclusions can be reached at this point. The samples should be rerun a number of times to be sure that they are really null alleles and not PCR artifacts. An alternative solution would be to sequence the samples that are not amplifying.

The heterogeneity of multilocus allelic distributions across the range of populations examined was more apparent with microsatellite than with allozyme loci (Table 64). In some pairwise comparisons significant allelic differentiation was detected using allozyme data but not microsatellite data and vice-versa. In fact, from the 10 possible comparisons it can be observed that:

- 3 give the same results, whether allozymes or microsatellites are concerned;
- 5 show non significant results with allozymes and significant differences with microsatellite loci - which could be explained by a higher resolution power of microsatellites in detecting genetic differentiation;
- 2 give contradictory results, significant with allozymes and not significant with microsatellite loci; these involve the samples from Faro-Foz and Faro-Óbidos; maybe the smaller sample size of Faro has played an important role diminishing the probabilities of sampling most of the alleles present in the population, as can be seen in Table 48 (page 160).

In fact the sample from Faro was involved in all the significant allozymes pairwise comparison results. Indeed this sample when analysed with allozyme markers, appears to be atypical, in the sense that it is the only sample that shows consistent single-locus pairwise allelic differentiation. This could, however, be at least partly attributable to smaller sample size of the Faro sample when compared to the other populations.

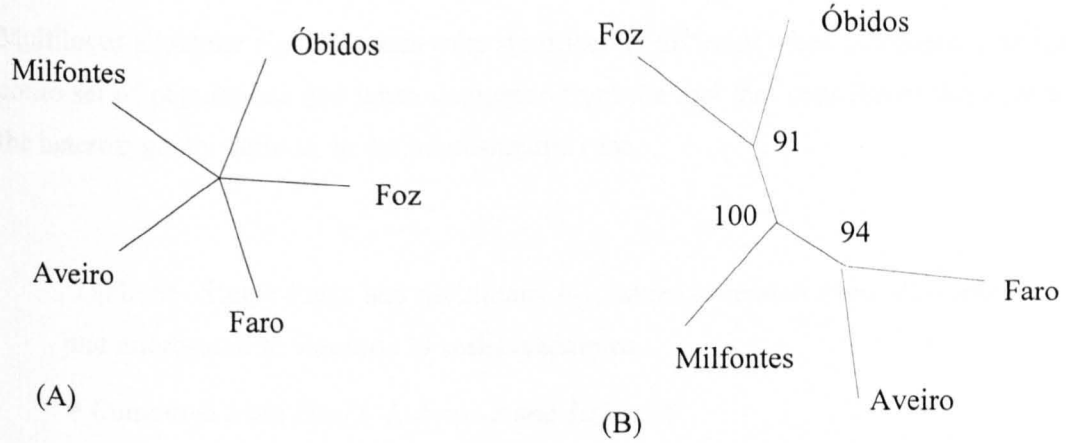
Table 64. Probability of homogeneity of allele frequencies estimated from pairwise tests.

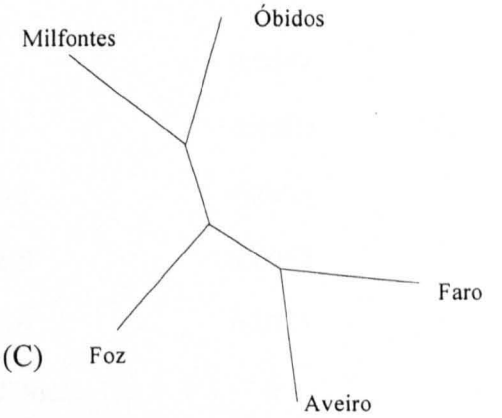
Microsatellite multilocus probability values given above diagonal, allozyme multilocus probability values given below diagonal. Probability estimates considered significant following sequential Bonferroni adjustment for 10 simultaneous tests are indicated: NS - not significant; * P<0.05; ** P<0.01.

	Aveiro	Foz	Óbidos	Milfontes	Faro
Aveiro	—	NS	*	*	NS
Foz	NS	—	**	*	NS
Óbidos	NS	NS	—	**	NS
Milfontes	NS	NS	NS	—	**
Faro	NS	**	*	**	—

The measure of genetic distance applied to this data set yielded variable results, in which the topology of both allozyme and microsatellite loci is not consistent with the geographical distribution of the samples (Figure 30). In fact the most geographically distant samples appear clustered together (Aveiro and Faro) in both trees. It can be hypothesised that, as Allegrucci *et al.* (1997) suggest for the Mediterranean, seabass populations may be grouped according to their ecological origin: marine and coastal lagoons. The Aveiro and Faro samples have a lagoon origin, which could contribute to the similarity between these samples. However, Óbidos also of a lagoon origin does not appear to cluster together with either Aveiro or Faro. The number of samples and the number of individuals in each sample in the present work was not sufficient to give a conclusive answer , but in future work, the ecological origin of the sample must be taken into consideration.

Figure 30. Hypothetical and actual relationships among the five samples of *D. labrax*.





(A) The null hypothesis: no definable relationships exist among the samples. (B) The relationships among the samples based on neighbor-joining analysis of Reynolds genetic distance inferred from 6 polymorphic allozyme loci. (C) The relationships among the samples based on neighbor-joining analysis of Reynolds genetic distance inferred from 4 polymorphic microsatellite loci. Numbers indicate the percentage of bootstrap replicates (out of 100) in which associations were supported.

Reynolds distances were almost ten-times larger for allozymes than for microsatellites, increasing the neighbor-joining tree robustness.

The higher mutation rates of microsatellites do not seem have a large effect on values of a parameter such as F_{ST}. The level of genetic structure observed in the present

study is almost 7 times larger for allozymes (F_{ST} = 0.0108) then for microsatellites (F_{ST} = 0.0016) (Table 65). The F_{ST} value is expected to be lower when high mutation rates are involved (Rousset, 1996) and indeed that is in agreement with the findings. Multilocus allozyme F_{ST} estimates were significantly different when computed over the entire set of populations and when computed from the loci that contributed the least to the heterozygosity deficits, in the microsatellite case.

Table 65. Single-locus and multilocus F_{ST} values estimated from allozyme and microsatellite variation in seabass samples.

♦ Computed from *Dla11, Labrax.-3* and *Labrax-8.*

Allozymes		Microsatellites	
Locus	F_{ST}	Locus	F_{ST}
*AAT-3**	0.0184*	*Dla11*	0.0044
ADA	0.0003	*Dla6*	-0.0021
*G3PDH-2**	-0.0001	*Labrax-3*	-0.0006
*GPI-1**	-0.0001	*Labrax-8*	0.0045
*GPI-2**	0.0032	*Labrax-9*	0.0021
*SOD**	0.0072*	All loci	♦
Multilocus	0.0108**	0.0016	0.0025*

From Table 66, it can be observed that for many of the general analysis performed in this study, both markers respond in the same way: detection of allelic differentiation among samples; no differences between age-classes; no correlation between genetic and geographical distance and poor clustering of geographically closely related samples using NJ dendrograms.

Clearly of great importance is the number of markers and individuals that must be sampled to provide reliable resolution at the intraspecific level.

Table 66. Comparison of microsatellite and allozyme markers for different analysis performed in this study.

(Adapted from Estoup, 1998).

Parameters	Allozymes	Microsatellites
F_{ST}	0.0108, S	0.0016, NS
N_em	23	156
Occurrence of private alleles (alleles unique to one sample)	No/Rare	Yes
Detection of allelic differentiation among samples	Yes, S	Yes, S
Detection of differences between age-classes	NS	NS
Correlation between genetic and geographical distance	No	No
Clustering of geographically closely related samples using NJ dendrograms	Poor	Poor

NS – non-significant; S – significant.

10. FINAL DISCUSSION AND CONCLUSIONS

Population genetics deals with many aspects of the grouping of individuals, by genetic means. A species population structure, the way individuals of the same species interact genetically can go from a single breeding population (panmixia) to multiple disjunctive populations, with all the possible intermediate stages. Panmixia means that all of the individuals in question constitute a single reproductive unit. This circumstance is expected when there is only one spawning area or when individuals do not return to their locations of birth for spawning (May and Kruger, 1991). The existence of geographically restricted spawning aggregations in some species suggests the existence of genetically discrete populations, but the potential genetic flow arising from the passive movement of pelagic eggs and larvae and the active dispersal of juveniles and adults can counteract this (Bentzen *et al.*, 1989). Any structure will be basically defined by the action of several mechanisms that may contribute to the species being subdivided. Among these genetic drift and natural selection are the most significant ones, with reproductive barriers of several kinds (geographical, hydrological, behavioural etc) also playing an important role.

On average, marine fish subpopulations are generally less genetically differentiated (F_{ST} = 0.062) than freshwater fish subpopulations (F_{ST} = 0.222) with anadromous species showing intermediate levels (F_{ST} = 0.108) (Gyllensten, 1985; Ward *et al.*, 1994b). Consequently, it might be said that the level of gene flow among marine subpopulations is greater than that of freshwater subpopulations. This may reflect the relative absence of physical barriers limiting genetic exchange among populations in the marine environment, establishing the general idea that marine species, especially those with migratory habits and with an egg and larval stages submitted to great current movements and therefore more probable stock mixes, would have a very low level of intraspecific differentiation.

A conclusion that a particular population structure exists for a certain species ought to take into consideration not only the analysis of genetic data, but should also

take into account knowledge about the life history of species, especially migratory and reproductive behaviour (May and Kruger, 1990).

It was during the 1990s that most of the effort in describing the genetic structure of the seabass has advanced, which was due to the interesting results from pioneer work done on this species and the implementation of new genetic markers, such as microsatellites. Some research groups have developed work on this species using mainly two types of genetic markers, allozyme and microsatellites. Studies based on allozymes and microsatellites do give different levels of genetic differentiation when analysed in detail. However, these studies seem to indicate that seabass populations have some degree of genetic structuring.

In this work, Portuguese estuarine juvenile seabass were sampled in order to assess the genetic structure of the species along the Portuguese coast. The main question involved in this work is: Is the entire seabass population of the Portuguese coast a single panmictic population? The main conclusion that can be drawn from the present results is that the null hypothesis of no genetic differentiation among populations of seabass from the Portuguese coast can be rejected. The seabass populations do not belong to a large homogeneous population, but rather exhibit subtle genetic differences from one another. In fact, these results are consistent with the existence of population differentiation in seabass, even at the fine geographic level here assessed, indicating however that the extent of genetic differentiation between samples is of low magnitude. Both allelic frequencies distributions (see Table 26 and Figures 28a and 28b) and F_{ST} measures indicate that the proportion of total genetic variation occurring among samples is quite small: between 0.11% to 0.25% (Table 65). The low level of genetic differentiation in seabass may stem, at least in part, from gene flow, but may also reflect selection, genetic drift and particularly in microsatellites loci convergent mutations. Gene flow is not able to maintain genetic homogeneity among subpopulations. In other words, there may be some restriction to the gene flow that is likely to cause genetic differentiation, but selection and genetic drift can also act independently in each of the nursery grounds especially if there is age-specific mortality, strong selection or population crashes. It is predicted that the observed structure can also be due to some degree of adult seabass homing to particular spawning

sites which enable their young to drift and actively swim into suitable nursery areas. Larval seabass enter nursery areas in the lagoons and estuaries very early in their lives, and it might be assumed that they arrive from nearby spawning aggregations. However, the degree of dispersion depends very much on the local current systems in the area. It would be appropriate and interesting at this point to put more emphasis on the study of the seabass life-cycle, in order to definitely establish its migration behaviour.

Estimates of gene flow using allozyme data, suggest moderate levels of gene flow among populations, with an average of 14.75 effective migrants per generation exchanged between localities, although this value is substantially increased to 100 with microsatellite data. This estimate of N_em is derived using Wright's (1943) island model of migration where gene flow is assumed to proceed between an infinite number of subpopulations, and is conservative if the true population structure conforms to an isolation by distance model, which is not the case in the present work. In fact, the results from the present study do not show a pattern of population genetic isolation by distance, whereby neighbouring localities are usually genetically similar to each other. Aveiro and Faro, for instance, do not show significant differences in the allelic frequency distribution among them. The fact that there is no correlation between genetic distance and geographic distance, lead us to propose an island model whereby each population has the possibility to exchange individuals with every other population.

The level of population resolution likely using allozymes may prove inadequate because of the low levels of variation observed, specially if the separation matrix is starch instead of cellulose acetate. In fact, Allegrucci *et al.* (1997) found 25 to 37.5% polymorphic loci using cellulose acetate, which is considerably higher than previous studies that used starch electrophoresis, namely Cervelli (1985) (5.9%) and Benharrat *et al.* (1984) (8.8-11.8%), in the Mediterranean. Thus, in future works, it is advisable to continue to use allozymes, as they have provided evidence suggesting the existence of two types of seabass populations, within the Mediterranean (sea and lagoon). Also highly polymorphic microsatellite loci have now been developed for the seabass (García de León *et al.*, 1995; Castilho & McAndrew, 1998) meaning that non-disfiguring biopsies (scales, blood) can be removed from commercial catches without the need to buy fish, dramatically reducing the cost of any study on adult populations.

In microsatellites loci heterozygosities are higher than ones observed with allozymes, revealing the importance of microsatellites variability and its potential usefulness in population genetic studies and in aquaculture selective breeding studies. Microsatellites have proven appropriate for research of population genetic structure at the microgeographic scale, however it is desirable to enlarge the sample sizes in order to increase the resolution power, decreasing the allele frequency sampling error that arises from the high number of alleles present at each locus. Because of the levels of gene flow detected, it is probably better to have fewer samples with many individuals, rather than more samples with less individuals. This will allow a better estimation of local allele frequencies, with increasing the statistical power of tests to detect smaller allele frequency differences. The microsatellite polymorphism observed should allow the detection of a greater level of genetic differentiation between populations separated by very small geographic distances (if such occur), if the sample sizes are increased.

High heterozygote deficits found in this as well in many other works in microsatellite loci may be an indication of the possible presence of null alleles. The extremely high estimates of null genotypes deduced from non-amplified samples if correct would make the allele frequency data presented quite unreliable. However, there are equally plausible technical reasons for non-amplification. No attempts at re-scoring all individuals that failed to amplify were made and it is unlikely that all 5 loci show substantial null allele frequencies, and García de León *et al.* (1997) does not mention evidence of null alleles at the loci examined in common. With this in view, technical problems hindering the amplification remain as the more likely explanation. With the increasing automation of microsatellite loci detection by automated sequencing it will also be possible to obtain the DNA sequences involved and thus to assess the real null allele frequencies involved.

The level of differentiation detected in the present study is indicative of a typical marine species with low values of F_{ST}. Some differentiation in population structure is detected at the allozyme level and microsatellite level but the two methods detect different structuring. Three allozyme loci are rarely polymorphic and most of the differences depend on a single sample from Faro (and on a single locus *AAT-3**) which was compromised by poor storage. It should thus be acknowledged that the sampling

regime and the sample quality jeopardise in part more substantiated results from this work.

In conclusion, the exact population discreteness cannot be ascertained with the present results, neither can the precise reasons for the differentiation observed identified. Despite the limitations of the data set and assuming the technical problems are random and induce no obvious bias, the data set suggests low level of differentiation between juvenile seabass populations.

These results may have therefore, various implications for the population biology of seabass, the management of seabass stocks, and also on the utility of allozymes and microsatellite as a tool in population genetics of *a priori* low level of genetic differentiation. Even if, as far as fishery managers are concerned, the gene flow estimated may seem too much to be able to genetically distinguish these populations, the data are valuable in fishery management as well as in the biology context. It will be important to compare the described substructure of the less mobile juveniles with the highly mobile adult populations nearby that can enhance our ability to discriminate populations. Particular attention should be paid to spatio-temporal sampling strategy in future works. The emphasis on sampling adult individuals during the spawning period, using different types of markers is desirable but cost and fishermen reticence to allow fish to be damaged will mean that microsatellites will be the only practical method.

11. REFERENCES

Aebersold, P. B., Winans, G. A., Teel, D. J., Milner, G. B. and Utter, F. M. (1987). Manual for starch gel electrophoresis: a method for the detection of genetic variation. NOAA Technical Report NMFS.

Agnèse, J. F. (1991). Taxonomic status and genetic differentiation among West African populations of the *Chrysichthys auratus* complex (Pisces, Siluriformes), based on protein electrophoresis. *Aquaculture and Fisheries Management*, **22**(2)**:** 229-237.

Allcock, A. L., Brierley, A. S., Thorpe, J. P. and Rodhouse, P. G. (1997). Restricted gene flow and evolutionary divergence between geographically separated populations of the Antarctic octopus *Pareledone turqueti*. *Marine Biology*, **129:** 97-102.

Allegrucci, G., Caccone, A., Cataudella, S., Powell, J. R. and Sbordoni, V. (1995). Acclimation of the European sea bass to freshwater: monitoring genetic changes by RAPD polymerase chain reaction to detect DNA polymorphisms. *Marine Biology*, **121:** 591-599.

Allegrucci, G., Fortunato, C., Cautadela, S. and Sbordoni, V. (1994). Acclimation to freshwater of the Sea Bass: evidence of selective mortality of allozyme genotypes. *In*: *Genetics and Evolution of Aquatic Organisms* (Beaumont, A. R., ed.). Chapman & Hall, London.

Allegrucci, G., Fortunato, C. and Sbordoni, V. (1997). Genetic structure and allozyme variation of sea bass (*Dicentrarchus labrax* and *D. punctatus*) in the Mediterranean Sea. *Marine Biology*, **128:** 347-358.

Allendorf, F. W. and Phelps, S. R. (1981). Use of allelic frequencies to describe population structure. *Canadian Journal of Fisheries and Aquatic Sciences*, **38:** 1507-1514.

Allendorf, F. W., Ryman, N. and Utter, F. M. (1987). Genetics and fishery management. Past, present and future. *In*: *Population Genetics & Fishery Management* (Ryman, N., and Utter, F. M., eds.). Washington Sea Grant Program, University of Washington, Seattle.

Alliot, E., Pastoureaud, A. and Thebault, H. (1983). Influence de la température et de la salinité sur la croissamce et la composition corporelle d'alevins de *Dicentrarchus labrax*. *Aquaculture*, **31:** 181-194.

Andrade, J. P. (1983). Contribuicao para o conhecimento da biologia (determinacao da idade e estudo do crescimento) do robalo *Dicentrarchus labrax* (L.) (Pisces, Moronidae) da Ria de Aveiro. *Publ. Cient. Univ. Algarve.*, **1:** 1-21.

Andrade, J. P. (1986). Morphological study of the sea-bass, *Dicentrarchus labrax*, of the Aveiro lagoon (Portuguese coast). *Publ. Cient. Univ. Algarve.*, **2:** 1-23.

Arias, A. (1980). Growth, food and reproductive habits of seabream (*Sparus aurata* L.) and sea bass (*Dicentrarchus labrax* L.) in the 'esteros' (fish ponds) of Cadiz. *Invest. Pesq.(Barc.)*, **44**(1): 59-83.

Arruda, L. M., Andrade, J. P. and Cunha, M. M. (1988). Abundance, diversity and community structure of the fish population in the Ria de Aveiro (Portugal). *Oceanologica Acta*, **11**(3): 235-240.

Avise, J. C. (1994). *Molecular markers, natural history and evolution*. Chapman and Hall, London.

Avise, J. C., Giblin-Davidson, C., Laerm, J., Patton, J. C. and Lansman, R. A. (1979). Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis*. *Proceedings of the National Academy of Sciences USA*, **76:** 6694-6698.

Barker, J. S. F., Moore, S. S., Hetzel, D. J. S., Evans, D. and Byrne, K. (1997). Genetic diversity of Asian water buffulo (*Bubalus bubalis*): microsatellite variation and a comparison with protein coding-loci. *Animal Genetics*, **28**(2): 103-115.

Barnabé, G. (1976). Contribution à la connaissance de la biologie du loup *Dicentrarchus labrax* (L.) (Poisson, Serranidae). These Doctorat Thesis. USTL, Montpellier.

Barnabé, G. (1980). *Exposé synoptique des donnés biologiques sur le loup ou bar Dicentrarchus labrax (Linné,1758)*. *In*: *Synopsis, Pêches* (F.A.O., ed.). F.A.O., Rome.

Barnabé, G. (1987). L'elevage du loup et de la dourade. *In*: *Aquaculture*. Lavoisier, Paris.

Barrosa, J. O. (1980). A Ria de Aveiro e a sua Influencia na Economia da Regiao. *Recursos Hídricos*, **1**(1-2): 55-70.

Basaglia, F., Salvatorelli, G., Santulli, A. and D'amelio, V. (1989). Ontogenetic variations of some enzymes in *Dicentrarchus labrax* (Serraniade). *Bolletino Zoologico*, **56:** 1-6.

Batargias, C. and Zouros, E. (1993). Development and the potential applications of highly polymorphic microsatellite and cDNA nuclear markers in gilthead seabream, *Sparus aurata. In*: *Fourth Congress of the European Society for Evolutionary Biology, Evolution 93*, Montpellier.22-28 August. .

Baverstock, P. R. and Moritz, C. (1990). Sampling design. *In*: *Molecular Systematics* (Hillis, D. M., and Moritz, C., eds.). Sinauer Associates, Massachusetts.

Beacham, T. D., Gould, A. P., Withler, R. E., Murray, C. B. and Barner, L. B. (1987). Biochemical genetic survey and stock identification of chum salmon (*Oncorhynchus keta*) in British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, **44:** 1702-1713.

Beacham, T. D., Withler, R. E. and Gould, A. P. (1985). Biochemical stock identification of chum salmon (*Oncorhynchus gorbuscha*) in southern British Columbia and Puget Sound. *Canadian Journal of Fisheries and Aquatic Sciences*, **42:** 1474-1483.

Belkhir, K., Borsa, P., Goudet, J., Chikhi, L. and Bonhomme, F. (1996 - 1998). GENETIX, logiciel sous WindowsTM pour la génétique des populations. Laboratoire Génome et Populations, CNRS UPR 9060, Université de Montpellier II, Montpellier (France).

Bembo, D. G., Carvalho, G. R., Cingolani, N., Arneri, E., Gianetti, G. and Pitcher, T. (1996). Allozymic and morphometric evidence for two stocks of the European anchovy *Engraulis encrasicolus* in Adriatic waters. *Marine Biology*, **126:** 529-538.

Bembo, D. G., Carvalho, G. R., Snow, M., Cingolani, N. and Pitcher, T. (1995). Stock discrimination among European anchovies *Engraulis encrasicolus*, by means of PCR-amplified mitochondrial DNA analysis. *Fishery Bulletin*, **94:** 31-40.

Benedetti, P., Pinna, A. and Colombo, L. (1995). Isolamento di sonde per l'analisi unilocus di DNA minisatellite nel branzino, *Dicentrarchus labrax* L. *In*: *PIM per*

le zone lagunari dell'Adriatico settentrionale. (Berletti, M., Rossi, R., and Spreafino, E., eds.). 27-34.

Benharrat, K., Pasteur, N., Siau, Y. and Bouian, A. (1984). Polymorphisme biochimique de loups (*Dicentrarchus labrax*) originaires de quatre populations naturelles et d'un élevage. *In*: *Recherches Biologiques en Aquaculture*. 17-27.

Bentzen, P., Brown, G. C. and Leggett, W. C. (1989). Mitochondrial DNA polymorphism, population structure and life history variation in American Shad (*Alosa sapidissima*). *Canadian Journal of Fisheries and Aquatic Sciences*, **46:** 1446-1454.

Bentzen, P., Harris, A. S. and Wright, J. M. (1991). Cloning of hypervariable minisatellite and simple sequence microsatellite repeats for DNA fingerprinting of important aquaculture species of Salmonids. *In*: *DNA fingerprinting: approaches and applications* (Burke, T. B., Dolf, G., Jeffreys, A. J., and Wolff, R., eds.). Basel, Switzerland, Birkhauser Veriag.

Bentzen, P., Taggart, C. T., Ruzzante, D. E. and Cook, D. (1996). Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the northwest Atlantic. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**(12)**:** 2706-2721.

Benzie, J. A. H., Frusher, S. and Ballment, E. (1992). Geographical variation in allozyme frequencies of populations of *Penaeus monodon* (Crustacea: Decapoda) in Australia. *Australian Journal of Marine and Freshwater Research*, **43**(4)**:** 715-725.

Berg, W. J. and Ferris, S. D. (1984). Restriction emdonuclease analysis of salmonid mitochondrial DNA. *Canadian Journal of Fisheries and Aquatic Sciences*, **41:** 1041-1047.

Bermingham, E. (1990). Mitochondrial DNA and the analysis of fish population structure. *In*: *Electrophoretic and Isoelectric Focusing Techniques in Fisheries Management* (Whitmore, D. H., ed.). CRC Press, Boca Ratton.

Bermingham, E., Forbes, S. H., Friedlan, K. and Pla, C. (1991). Discrimination between Atlantic salmon (*Salmo salar*) of North American and European origin using restriction analysis of mitochondrial DNA. *Canadian Journal of Fisheries and Aquatic Sciences*, **48:** 884-893.

Bertignac, M. (1988). Le bar, effort de peche en hausse...gare au stock. *In*: *Equinose*.

Billington, N. and Hebert, P. D. N. (1991). Mitochondrial DNA diversity in fishes and its implications for introductions. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**(Suppl): 80-94.

Billington, N., Hebert, P. D. N. and Ward, R. D. (1990). Allozyme and mitochondrial DNA variation among three species of *Stizostedium* (Percidae): phylogenetic and zoogeographical implications. *Canadian Journal of Fisheries and Aquatic Sciences*, **47:** 1093-1102.

Blanquer, A., Alayse, J. P., Berrada-Rkhami, O. and Berrebi, P. (1992). Allozyme variation in turbot (*Psetta maxima*) and (*Scophthalmus rhombus*) (Osteichthyes, Pleuronectiformes) throughout their range in Europe. *Journal of Fish Biology*, **41:** 725-736.

Borsa, P., Blanquer, A. and Berrebi, P. (1997). Genetic structure of the flounders *Platichthys flesus* and *P. stellatus* at different geographic scales. *Marine Biology*, **129:** 233-246.

Bou Ain, A. (1977). Contribution a l'etude morphologique, anatomique et biologique de *Dicentrarchus labrax* (Linne, 1758), *Dicentrarchus punctatus labrax* (Linne, 1758) et *Dicentrarchus punctatus* (Bloch, 1792) des cotes tunisiennes. Thèse de Doct. Specialite Thesis. Fac. Sc. Tunis, Tunis.

Bouza, C., Sanchez, L. and Martinez, P. (1995). A comparative gene diversity analysis between natural populations and productive stocks of turbot (*Scophthalmus maximus*). *In*: *Molecular Biology in Fish, Fisheries and Aquaculture* (Matthews, R. A., and Thorpe, J. E., eds.). Fisheries Society for the British Isles, Plymouth, U.K. .

Bowcock, A. M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J. R. and Cavalli-Sforza, L. L. (1994). High resolution human evolutionary trees with polymorphic microsatellites. *Nature*, **368:** 455-457.

Brooker, A. L., Cook, D., Bentzen, P., Wright, J. M. and Doyle, R. D. (1994). Organization of microsatellites differs between mammals and cold-water teleost fishes. *Canadian Journal of Fisheries and Aquatic Sciences*, **51:** 1959-1966.

Brookfield, J. F. Y. (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, **5:** 453-455.

Brown, W. M., George, M. and Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences USA*, **76:** 1967-1971.

Bruford, M. W. and Wayne, R. K. (1993). Microsatellites and their application to population genetic studies. *Current Opinions in Genetic Development.*, **3**(6)**:** 939-943.

Bruslé, J. and Roblin, C. (1984). Sexualite du loup *Dicentrarchus labrax* en condition d'elevage controlé. *In*: *L'aquaculture du bar et des Sparides* (Barnabé, G., and Billard, R., eds.). INRA, Paris.

Buroker, N. E. (1983). Genetic differentiation and population structure of the American oyster *Crassostrea virginica* (Gmelin) in Chesapeake bay. *Journal of Shellfish Research*, **3:** 153-167.

Callen, D. F., Thompson, A. D., Shen, Y., Phillips, H. A., Richards, R. I., Mulley, J. C. and Sutherland, G. R. (1993). Incidence and origin of "null" alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics*, **52:** 922-927.

Campillo, A., Aldebert, Y., Bigot, J.-L. and Liorzou, B. (1989). Données sur la distribution des principales éspèces commerciales du Golf du Lion. IFREMER. Paris.

Cannings, C. and Edwards, A. W. F. (1969). Expected genotypic frequencies in a small sample: deviation from Hardy-Weinberg equilibrium. *American Journal of Human Genetics*, **21:** 245-247.

Carr, S. M. and Marshall, H. D. (1991). Detection of intraspecific DNA sequence variation in the mitochondrial cytochrome b gene of Atlantic cod (*Gadus morhua*) by the polymerase chain reaction. *Canadian Journal of Fisheries and Aquatic Sciences*, **48:** 48-52.

Carrillo, M., Zanuy, S., Prat, F., Cerda, J., Ramos, J. and Mananos, E. B., N. (1995). Sea Bass (*Dicentrarchus labrax*). *In*: *Broodstock Management and Egg and Larval Quality* (Bromage, N. R., and Roberts, R. J., eds.). Blackwell Science, Oxford.

Carvalho, G. and Hauser, L. (1995). Molecular genetics and the stock concept in fisheries. *In*: *Molecular Genetics in Fisheries* (Carvalho, G., and Pitcher, T. J., eds.). Chapman and Hall, London.

Carvalho, G. R. (1998). Molecular genetic analysis of stock structure in marine pelagic fishes. International Council for the Exploration of the Sea. Cork.

Carvalho, G. R. and Pitcher, T. J. (1995). *Molecular Genetics in Fisheries* (Carvalho, G. R., and Pitcher, T. J., eds.). Chapman & Hall, London.

Carvalho, G. R., Thompson, A. and Stoner, A. L. (1992). Genetic diversity and population differentiation of the shortfin squid *Illex argentinus* in the south-west Atlantic. *Journal of Experimental Marine Biology and Ecology*, **158:** 105-121.

Castilho, R. and McAndrew, B. J. (1998). Two polymorphic microsatellite markers in the European seabass, *Dicentrarchus labrax* (L.). *Animal Genetics*, **29:** 151-152.

Cavalli-Sforza, L. L. and Bodmer, W. F. (1971). *The genetics of human populations*. W. H. Freeman, San Francisco.

Cavalli-Sforza, L. L. and Edwards, A. W. F. (1967). Phylogenetic analysis: models and estimation procedures. *Evolution*, **21:** 550-570.

Cawdery, S. A. H. and Ferguson, A. (1988). Origins and differentiation of three sympatric species of trout (*Salmo trutta* L.) in Lough Melvin. *Polskie Archiwum Hydrobiologii*, **35:** 267-277.

Cecconi, F., Venanzetti, F. and Mariattini, P. (1993). Sequence of the mitochondrial tDNA thr and tRNA pro genes from European seabass, *Dicentrarchus labrax*. *Nucleic Acids Research*, **21**(9): 2253.

Cervelli, M. (1985). Primi risultati dell'analisi genetica di alcuni sistemi enzimatici di *Dicentrarchus labrax* L. *Oebalia*, **XI N.S.:** 45-48.

Chakraborty, R. and Leimar, O. (1987). Genetic variation within a subdivided population. *In*: *Population Genetics & Fishery Management* (Ryman, N., and Utter, F. M., eds.). Washington Sea Grant Program, University of Washington, Seattle.

Chapman, R. W. and Brown, B. C. (1990). Mitochondrial DNA isolation methods. *In*: *Electrophoretic and Isoelectric Focusing Techniques in Fisheries Management* (Whitmore, D. H., ed.). CRC Press.

Chatain, B. (1991). Current status of the French intensive rearing techniques for the sea bass (*Dicentrarchus labrax*) and seabream (*Sparus auratus*). *In*: *13th Conference of ESCPB on Research for Aquaculture: Fundamental and Applied Aspects* (ESCPB, ed.), Antibes. 148p.

Chauvet, C., Lasserre, G., Bach, P. and Besseau, L. (1992). Resultats d'une experience de marquage de trois especes de poissons cotiers du golfe du lion: *Dicentrarchus labrax*, *Sparus aurata* et *Liza aurata*. *Cybium*, **16**(1): 3-11.

Chervinski, J. (1979). Preliminary experiments on the adaptability of juvenile European sea bass (*Dicentrarchus labrax* (L.)) and gilthead sea bream (*Sparus aurata* (L.)) to brackish water. *Bamidgeh*, **31:** 14-14.

Chikhi, L., Bonhomme, F. and Agnese, J.-F. (1998). Low genetic variability in a widely distributed and abundant clupeid species, *Sardinella aurita* . New empirical results and interpretations. *Journal of Fish Biology*, **52:** 861-878.

Child, A. R. (1992). Biochemical polymorphism in bass, *Dicentrarchus labrax,* in the waters around the British Isles. *Journal of the Marine Biology Association of the U.K.*, **72:** 357-364.

Clayton, J. W. and Tretiak, D. N. (1972). Amine-citrate buffers for pH control in starch gel electrophoresis. *Journal of the Fisheries Research Board of Canada*, **29:** 1169-1172.

Committee), I. I. U. o. B. N. (1984). *Enzyme nomenclature*. Academic Press, Orlando, Florida.

Copeman, J. B., Cucca, F., Hearne, C. M., Cornall, R. J., Reed, P. W., Roenningen, K. S., Undlien, D. E., Nistico, L., Buzzetti, R. and Tosi, R. (1995). Linkage desiquilibrium mapping of a type I diabetes susceptibility gene (IDDM7) to chromosome 2q31-q33. *Nature Genetics*, **9**(1): 80-85.

Costa, M. J. (1988). The Tagus and Mira Estuaries (Portugal) and their role as spawning and nursery areas. *Journal of Fish Biology*, **33**(SA): 249-250.

Coughlan, J. P., Imsland, A. K., Galvin, P. T., Fitzgerald, R. D. and Cross, T. F. (1998). Microsatellite DNA variation in wild populations and farmed strains of turbot from Ireland and Norway: a preliminary study. *Journal of Fish Biology*, **52:** 916-922.

Creech, S. (1991). An electrophoretic investigation of population of *Atherina boyer* Risso, 1810 and *A. presbyter* Cuvier 1829 (Teleostei: Atherinidae): genmetic evidence in support of the two species. *Journal of Fish Biology*, **39:** 807-816.

Cross, T. F. and Payne, R. H. (1978). Geographic variation in Atlantic cod, *Gadus morhua*, off eastern North America: a biochemical systematica approach. *Journal of the Fisheries Research Board of Canada*, **35:** 117-123.

Crow, J. F. (1986). *Basic concepts in population, quantitative and evolutionary genetics*. W.H. Freeman & Company, N.Y.

Daemen, E., Volckart, F., Hellemans, B. and Ollevier, F. (1996). The genetic differentiation of the European eel (*Anguilla anguilla* L.) on the european continental shelf. *Royal Belgium Academy of Sciences*, **65**(1)**:** 39-42.

Dallas, J. F. (1992). Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. *Mammalian Genome*, **5:** 32-38.

Dando, P. P. and Demir, N. (1985). On the spawning and nursery grounds of bass (*Dicentrarchus labrax*) in the Plymouth area. *Journal of the Marine Biology Association of the U.K.*, **65:** 159-168.

Davidson, W. S., Birt, T. P. and Green, J. M. (1989). A review of genetic variation in Atlantic salmon, *Salmo salar* L., and its importance for stock identification, enhancement programmes and aquaculture. *Journal of Fish Biology*, **34:** 547-560.

Demestre, M., Recasens, L. and Sanchez, P. (1989). Evolution de la pecherie pendant les vingt dernieres annees dans les lagunes du delta de l'Ebre (Catalogne, Espagne). *Bulletin de la Societé de Zoologie Française*, **114**(4)**:** 21-29.

Dieuzeide, R., Novella, M. and Roland, J. (1954). Catalogue des poissons des côtes algériennes. 2: Ostéoptérygiens. *Bulletin de la Station d' Aquiculture et Pêches de Castiglione*, **5:** 258.

DiRienzo, A., Peterson, A. C., Garza, J. C., Valdes, A. M., Slatkin, M. and Freimer, N. (1994). Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences USA*, **91:** 3166-3170.

Doliana, R., Bontolossi, R. and Colombo, M. (1992). Cloning and sequencing of European sea bass (*Dicentrarchus labrax*) growth hormone cDNA using polymerase chain reaction and degenerate oligonucleotides. *DNA seq.*, **3:** 185-189.

Ellegren, H. (1995). Mutation rates at porcine microsatellite loci. *Mammalian Genome*, **6:** 376-377.

Elliot, N. G., Lowry, P. S., Grewe, P. M., Innes, B. H., Yearsley, G. K. and Ward, R. D. (1998). Genetic evidence for depth and spacially separated stocks of deep-water spikey oreo in Australasian waters. *Journal of Fish Biology*, **52:** 796-816.

Elliot, N. G. and Ward, R. D. (1992). Enzyme variation in orange roughy, *Hoplostethus atlanticus* teleostei: Trachichthyidae), from southern Australian and New Zealand waters. *Australian Journal of Marine and Freshwater Research*, **43:** 1561-1571.

Estoup, A. (1998). Comparative analysis of microsatellite and allozyme markers: a case study investigating microgeographic differentiation in brown trout (*Salmo trutta*). *Molecular Ecology*, **7:** 339-353.

Estoup, A., Presa, P., Krieg, F., Vaiman, D. and Guyomard, R. (1993). (CT)n and (GT)n microsatellites: a new class of genetic markers for *Salmo trutta* L. (brown trout). *Heredity*, **71:** 488-496.

Estoup, A., Tailliez, C., Cornuet, J. M. and Solignac, M. (1995). Size homoplasy and mutational processes of interruped microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Molecular Biology and Evolution*, **12**(6)**:** 1074-1084.

Estoup, A. and Turgeon, J. (1996). Microsatellite markers: isolation with non-radioactive probes and amplification. INRA, Jouy-en-Josas.

Ewens, W. J. (1972). The sampling theory of selectively neutral alleles. *Theoretical Population Biology*, **3:** 87-112.

F.A.O. (1995). *Fisheries statistics, catches and landings. In*: *F.A.O. Fisheries Series, n°48*. F.A.O., Rome.

F.A.O. (1996). *Fisheries Circular*. F.A.O., Rome.

Fabrizio, M. C. (1987). Growth-invariant discrimination and classification of striped bass stocks by morphometric and electrophoretic methods. *Transactions of the American Fisheries Society*, **116:** 728-736.

Felsenstein, J. (1993). PHYLIP. *In*: *Phylogeny Inference Package*. University of Washington, Seattle.

Ferguson, A. (1986). Lough Melvin - a unique fish community. *Occasional Papers in Irish Science and Technology*, **1:** 1-17.

Ferguson, A. (1989). Genetic differences among bron trout, *Salmo trutta*, stocks and their importance for the conservation and management of the species. *Freshwater Biology*, **21:** 35-46.

Ferguson, A. (1995). Molecular Genetics in Fisheries: current and future perpectives. *In*: *Molecular Genetics in Fisheries* (Carvalho, G. R., and Pitcher, T. J., eds.). Chapman and Hall, London.

Ferguson, A. and Taggart, J. B. (1991). Genetic differentiation among the sympatric brown trout (*Salmo trutta*) populations of Lough Melvin, Ireland. *Biological Journal of the Linnean Society*, **43:** 221-237.

Ferguson, A., Taggart, J. B., Prodöhl, P. A., McMeel, O., Thompson, C., Stone, C., McGinnity, P. and Hynes, R. A. (1995). The application of molecular markers to the study and conservation of fish populations, with special reference to *Salmo*. *Journal of Fish Biology*, **47**(Suppl. A)**:** 103-126.

Ferris, S. D. and Berg, W. J. (1987). The utility of mitochondrial DNA in fish genetics and fishery management. *In*: *Population Genetics & Fishery Management* (Ryman, N., and Utter, F. M., eds.). Washington Sea Grant Program, University of Washington, Seattle.

Fisher, R. A. (1935). The logic of inductive inference. *J. Roy. Stat. Soc.*, **98:** 39-54.

Fisher, R. A. (1970). *Statistical methods for Research Workers*. *In*: *Statistical methods for Research Workers*. Oliver & Boyd, Edinburgh.

G.E.P.P. (1988). Recursos de Pesca. Ministério da Agricultura e Pescas. Lisboa.

G.E.P.P. (1989). Recursos de Pesca. Ministério da Agricultura e Pescas. Lisboa.

G.E.P.P. (1990). Recursos de Pesca. Ministério da Agricultura e Pescas. Lisboa.

G.E.P.P. (1991). Recursos de Pesca. Ministério da Agricultura e Pescas. Lisboa.

G.E.P.P. (1992). Recursos de Pesca. Ministério da Agricultura e Pescas. Lisboa.

G.E.P.P. (1993). Recursos de Pesca. Ministério da Agricultura e Pescas. Lisboa.

G.E.P.P. (1994). Recursos de Pesca. Ministério da Agricultura e Pescas. Lisboa.

Galleguillos, R. A. and Ward, R. D. (1982). Genetic and morphological divergence between populations of the flatfish *Platichthys flesus* (L.). *Biological Journal of the Linnean Society*, **17:** 395-408.

Galvin, P., McKinnell, S., J.B., T., Ferguson, A., O'Farrell, M. and Cross, T. F. (1995). Genetic stock identification of Atlantic salmon using single locus minisatellite DNA profiles. *Journal of Fish Biology*, **47**(Suppl. A): 186-199.

García de León, F. J. (1995). Marqueurs hypervariables chez le loup (*Dicentrarchus labrax L.*): applicationsaux programmes d'amelioration génétique et a l'étude de populations naturelles. *In*: *Centre d'Oceanologie de Marseille*. Thèse de Doctorat Thesis. Université d'Aix-Marseille II, Marseille.

García de León, F. J., Chikhi, L. and Bonhomme, F. (1997). Microsatellite polymorphism and population subdivision in natural populations of European sea bass *Dicentrarchus labrax* (Linnaeus, 1758). *Molecular Ecology*, **6:** 51-62.

García de León, F. J., Dallas, J. F., Chatain, B., Canone, M., Versini, J. J. and Bonhomme, F. (1995). Development and use of microsatellite markers in seabass, *Dicentrarchus labrax* (Linnaeus, 1758) (Perciformes: Serranidae). *Molecular Marine Biology and Biotechnology*, **4**(1): 62-68.

Goff, D. J., Galvin, K., Katz, H., Westerfield, M., Lander, E. S. and Tabin, C. J. (1992). Identification of polymorphic simple sequence repeats in the genome of zebrafish. *Genomics*, **14:** 200-202.

Gold, J. R., Richardson, L. R., Furman, C. and King, T. L. (1993). Mitochondrial DNA differentiation and population structure in red drum (*Sciaenops ocellatus*) from the Gulf of Mexico and Atlantic Ocean. *Marine Biology*, **116:** 175-185.

Gonzalez-Villasenor, L. I., Burkhoff, A. M., Corces, V. and Powers, D. A. (1986). Characterization of cloned mitochondrial DNA from the teleost *Fundulus heteroclitus* and its usefulness as an interspecies hybridization probe. *Canadian Journal of Fisheries and Aquatic Sciences*, **43:** 1866-1872.

Goodfellow, P. N. (1993). Microsatellites and the new genetic maps. *Current Biology*, **3**(3): 149-151.

Gordo, L. S. (1989). Age, growth and sexuality of sea bass, *Dicentrarchus labrax* (Linnaeus, 1758) (Perciformes, Moronidae) from Aveiro Lagoon, Portugal. *Scientia Marina*, **53**(1): 121-126.

Goudet, J., Raymond, M., De Meeüs, T. and Rousset, F. (1996). Testing differentiation in diploid populations. *Genetics*, **144:** 933-940.

Grant, W. S. (1984). Biochemical population genetics of Atlantic herring *Clupea harengus*. *Copeia*, : 357-364.

Grant, W. S., Chang, I. Z., Kobayashi, T. and Stahl, G. (1987). Lack of genetic stock discretion in Pacific cod (*Gadus macrocephalus*). *Canadian Journal of Fisheries and Aquatic Sciences*, **44**(3): 490-498.

Grant, W. S., Teel, D. J. and Kobayashi, T. (1984). Biochemical population genetics of Pacific halibut (*Hippoglossus stenolepis*) and comparison with Atlantic halibut (*H. hippoglossus*). *Canadian Journal of Fisheries and Aquatic Sciences*, **41:** 1083-1088.

Grant, W. S. and Utter, F. M. (1980). Biochemical genetic variation in walleye pollock, *Theragra chalcogramma*, and population structure in the southeastern Bering Sea and Gulf of Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, **37:** 1093-1100.

Grant, W. S. and Utter, F. M. (1984). Biochemical population genetics of Pacific Herring (*Clupea pallasi*). *Canadian Journal of Fisheries and Aquatic Sciences*, **41**(6): 856-864.

Graves, J. E. and Dizon, A. E. (1987). Mitochondrial DNA sequence similarity of Atlantic and Pacific albacore tuna (*Thunnus alalunga*). *Canadian Journal of Fisheries and Aquatic Sciences*, **46:** 870-873.

Graves, J. E., Ferris, S. D. and Dizon, A. E. (1984). Close genetic similarity of Atlantic and Pacific skipjack tuna (*Katsuwonus pelamis*) demonstrated with restriction endonuclease analysis of mitochondrial DNA. *Marine Biology*, **79**(3): 315-319.

Guo, S. W. and Thompson, E. A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, **48:** 361-372.

Gyllensten, U. (1985). The genetic structure of fish: Differences in the intraspecific distribution of biochemical genetic variation between marine, anadromous and freshwater species. *Journal of Fish Biology*, **26:** 691-699.

Hallerman, E. M. and Beckmann, J. S. (1988). DNA-level polymorphism as a tool in fisheries science. *Canadian Journal of Fisheries and Aquatic Sciences*, **45:** 1075-1087.

Hansen, M. M. (1998). Practical sampling strategies in studies of genetic population structure of marine and anadromous species. International Council for the Exploration of the Sea. Cork.

Harding, J. P. (1949). The use of probability paper for the graphical analysis of polymodal frequency distributions. *Journal of the Marine Biological Association of the U.K.*, **28:** 141-153.

Hardy, G. H. (1908). Mendelian proportions in a mixed population. *Science*, **28:** 49-50.

Harris, H. and Hopkinson, D. A. (1976). *Handbook of Enzyme Electrophoresis in Human Genetics*. North-Holland Publishing Company, Amsterdam.

Hartl, D. L. and Clark, A. G. (1989). *Principles of Population Genetics*. Sinauer Associates, Massachusetts.

Henriques, A. G. (1988). Simulação em modelo matemático do regime sedimentologico do estuario do rio Mondego (porto da Figueira da Foz) (Tema 2). *Recursos Hídricos*, **9**(3): 13-34.

Herbinger, C. M., Doyle, R. W., Pitman, E. R., Paquet, D., Mesa, K. A., Morris, D. B., Wright, J. M. and Cook, D. (1996). DNA fingerprint based analysis of paternal and maternal effects on offspring growth and survival in communally reared rainbow trout. *Aquaculture*, **137:** 245-256.

Hilbish, T. J. and Koehn, R. K. (1985). The physiological basis of natural selection at the *LAP* locus. *Evolution*, **39:** 1302-1317.

Holden, M. and Williams, T. (1974). The biology, movements and populations dynamics of bass, *Dicentrarchus labrax* , in English waters. *Journal of the Marine Biology Association of the U.K.*, **54:** 91-107.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6:** 65-70.

Hugues, C. R. and Queller, D. C. (1993). Detection of highly polymorphic microsatellite loci in a species with allozyme polymorphism. *Molecular Ecology*, **2:** 131-137.

I.C.E.S. (1997). Report of the Working Group on the Application of Genetics in Fisheries and Mariculture. International Council for the Exploration of the Sea. Mariculture Committe. Gdansk.

I.C.E.S. (1998). Report of the Working Group on the Application of Genetics in Fisheries and Mariculture. International Council for the Exploration of the Sea. Mariculture Committe. Cork.

Igland, O. and Naevdal, G. (1995). Analysis of genetic differentiation between Greenland halibut, *Reinhartius hippoglossoides* Walbum 1792, from the Barents Sea, East and West Greenland. *In*: *Molecular Biology in Fish, Fisheries and Aquaculture* (Matthews, R. A., and Thorpe, J. E., eds.). Fisheries Society of the British Isles, Plymouth, U.K. .

Ihssen, P. E., Booke, H. E., Casselman, J. M., Meglade, J. M., Payne, N. R. and Utter, F. M. (1981). Stock identification: Materials and methods. *Canadian Journal of Fisheries and Aquatic Sciences*, **38:** 1838-1855.

Jackson, T. R. (1995). Linkage analysis of molecular markers and a search for quantitative trait loci for upper temperature tolerance in rainbow trout (*Oncorhynchus mykiss*). PhD. Thesis. University of Guelph, Guelph.

Jamieson, A. (1975). Enzyme types of Atlantic cod stocks on the North American banks. *In*: *Isozymes* (Markert, C. L., ed.). Academic Press, New York.

Jarman, A. P. and Wells, R. A. (1989). Hypervariable minisatellites: recombinators or innocent bystanders? *Trends Genetics*, **5:** 367-371.

Jarne, P., Berrebi, P. and Guelorget, O. (1991). Variabilité génétique et morphometrique de cinq populations de la palourde *Ruditapes decussatus* (Molusque, Bivalve). *Oceanologica Acta*, **11**(4)**:** 401-407.

Jarne, P. and Lagoda, P. J. L. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, **11**(10 Oct.)**:** 424-429.

Jeffreys, A. J., Wilson, V. and Thein, S. L. (1985). Hypervariable minisatellite regions in human DNA. *Nature*, **314:** 67-73.

Jennings, S. and Pawson, M. G. (1991). The development of bass, *Dicentrarchus labrax*, eggs in relation to temperature. *Journal of the Marine Biology Association of the U.K.*, **71**(1)**:** 107-116.

Jennings, S. and Pawson, M. G. (1992). The origin and recruitment of seabass, *Dicentrarchus labrax* , larvae to nursery areas. *Journal of the Marine Biology Association of the U.K.*, **72:** 199-212.

Jorge, I. (1991). Contribuição para o conhecimento da ictiofauna do estuário do Mondego. I.N.I.P. Lisboa.

Jørstad, K. E., King, D. P. F. and Naevdal, G. (1991). Population structure of Atlantic herring, *Clupea harengus* L. *Journal of Fish Biology*, **39**(Suppl.A): 43-52.

Kelley, D. F. (1979). Bass populations and movements on the west coast of the U.K. *Journal of the Marine Biology Association of the U.K.*, **59:** 896-936.

Kelley, D. F. (1987). Food of Bass in U.K. waters. *In*: *Journal of the Marine Biology Association of the U.K.*

Kelley, D. F. (1988a). Age determination in bass and assessment of growth and year-class strength. *Journal of the Marine Biology Association of the U.K.*, **68:** 179-214.

Kelley, D. F. (1988b). The importance of estuaries for sea-bass, *Dicentrarchus labrax* (L.). *Journal of Fish Biology*, **33**(Suppl): 25-33.

Kelley, D. F. (1991). The Sea Bass. *Nat. Dev.*, : 8-11.

Kennedy, M. and Fitzmaurice, P. (1968). Ocurrence of eggs of bass *Dicentrarchus labrax* on the southern coast of Ireland. *Journal of the Marine Biology Association of the U.K.*, **48:** 585-592.

Kennedy, M. and Fitzmaurice, P. (1972). The biology of bass, *Dicentrarchus labrax* , in Irish waters. *Journal of the Marine Biology Association of the U.K.*, **52:** 557-597.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.

Kimura, M. and Crow, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics*, **49:** 725-738.

Kimura, M. and Ohta, T. (1978). Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences USA*, **75:** 2868-2872.

Kimura, M. and Weiss, G. H. (1964). The stepping-stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, **49:** 561-576.

King, T. L. and Pate, H. O. (1992). Population structure of spotted seatrout inhabiting the Texas Gulf coast: an allozymic perspective. *Transactions of the American Fisheries Society*, **121:** 746-756.

Kit, S. (1961). Equilibrium sedimentation in density gradients of DNA preparations from animal tissues. *Journal of Molecular Biology*, **3:** 711-716.

Kocher, T. D., Lee, W. J., Sobolewska, H., Penman, D. and McAndrew, B. (1998). A genetic linkage map of a cichlid fish, the Tilapia (*Oreochromis niloticus*). *Genetics*, **148:** 1225-1232.

Kornfield, I. and Parker, A. (1997). Molecular systematics of a rapidly evolving species flock: the mbuna of Lake Malawi and the search for phylogenetic signal. *In*: *Molecular Systematics of Fishes* (Kocher, T. D., and Stepian, C. A., eds.). Academic Press, London.

Kotoulas, G., Bonhomme, F. and Borsa, P. (1995a). Genetic structure of the common sole *Solea vulgaris* at different geographic scales. *Marine Biology*, **122:** 361-375.

Kotoulas, G., Magoulas, A., Tsimenides, N. and Zouros, E. (1995b). Marked mitochondrial DNA differences between Mediterranean and Atlantic populations of the swordfish, *Xiphias gladius*. *Molecular Ecology*, **4:** 473-481.

Larkin, P. A. (1981). A perspective on population genetics and salmon management. *Canadian Journal of Fisheries and Aquatic Sciences*, **38:** 1469-1475.

Leary, R. F. and Booke, H. E. (1990). Starch gel electrophoresis and species distinction. *In*: *Methods for fish biology* (Schreck, C. B., and Moyle, P. B., eds.). American Fisheries Society, Bethesda, Maryland.

Lehmann, T., Hawley, W. A. and Colloins, F. H. (1996). An evaluation of evolutionary constraints on microsatellite loci using null alleles. *Genetics*, **144:** 1155-1163.

Lessios, H. A. (1992). Testing electrophoresis data for agreement with Hardy-Weinberg expectations. *Marine Biology*, **112:** 517-523.

Levinson, G. and Gutman, G. A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution*, **4:** 203-221.

Lie, O., Slettan, A., Grimholt, U., Lundin, M., Syed, M. and Olsaker, I. (1994). Fish gene maps and their implications for aquaculture. *Animal Biotechnology*, **5**(2)**:** 209-218.

Litt, M. and Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics*, **44**(3)**:** 397-401.

Maciel, P. (1995). Correlation between CAg repeat length and clinical features in the MAchado-Joseph disease. *American Journal of Human Genetics*, **57:** 54-61.

Magoulas, A., Tsimenides, N. and Zouros, E. (1996). Mitochondrial DNA phylogeny and the reconstruction of the population history of a species: the case of the european anchovy (*Engraulis encrasicolus*). *Molecular Biology and Evolution*, **13**(1): 178-190.

Magoulas, A. and Zouros, E. (1993). Restriction/site heteroplasmy in anchovy (*Engraulis encrasicolus*) indicates incidenta biparental inheritance of mitochondrial DNA. *Marine Biology Evolution*, **10:** 319-325.

Mamuris, Z., Apostolidis, A. P. and Triantaphyllidis, C. (1998). Genetic protein variation in red mullet (*Mullus barbatus*) and striped red mullet (*M. surmuletus*) populations from the Mediterranean Sea. *Marine Biology*, **130:** 353-360.

Manly, B. F. J. (1985). *The statistics of natural selection. In*: *The statistics of natural selection*. Chapman and Hall, London.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Reseach*, **27:** 209-220.

Markert, C. L. (1975). Biology of isozymes. *In*: *Isozymes Molecular Structure* (Markert, C. L., ed.). Academic Press, N.Y.

Martínez, G. (1991). Analisis genetico de poblaciones de lubina (*Dicentrarchus labrax* L.) mediante electroforesis de isoenzimas. *In*: *Tesis doctoral*. Y Thesis. Universidad de Malaga, Malaga.

Martínez, G., McEwen, I., McAndrew, B. J. and Alvarez, M. C. (1991). Electrophoretic analysis of protein variation in two Spanish populations of the european seabass, *Dicentrarchus labrax* L. (Pisces, Moronidae). *Aquaculture and Fisheries Management*, **22:** 443-455.

Martínez, G., Mores, A., Thode, G. and Alvarez, M. C. (1990). Differencias genéticas entre poblaciones de lubina (*Dicentrarchus labrax*) mediante electroforésis de isoenzimas. *In*: *Actas III Congresso National de Acuicultura*. 313-318.

May, B. and Krueger, C. C. (1990). Use of allozyme data for population analysis. *In*: *Electrophoretic and Isoelectric Focusing Techniques in Fisheries Management* (Whitmore, D. H., ed.). CRC Press.

Mayer, I., Shackley, S. E. and Witthames, P. R. (1990). Aspects of the reproductive biology of the bass, *Dicentrarchus labrax* L. II. Fecundity and pattern of oocyte development. *Journal of Fish Biology*, **36:** 141-148.

McAndrew, B. and Majumdar, K. C. (1983). Tilapia stock identification using electrophoretic markers. *Aquaculture*, **30:** 249-261.

McConnell, S. K., Hamilton, L., Morris, D., Cook, D., Paquet, D., Bentzen, P. and Wright, J. M. (1995a). Isolation of salmonid microsatellite loci in Atlantic salmon and their application to the populations genetics of Canadian East coast stocks. *Aquaculture*, **137**(137): 19-30.

McConnell, S. K., O'Reilly, P., Hamilton, L., Wright, J. M. and Bentzen, P. (1995b). Polymorphic microsatellite loci from Atlantic salmon (*Salmo salar*) : Genetic differentiation of North American and European populations. *Canadian Journal of Fisheries and Aquatic Sciences*, **52:** 1863-1872.

McConnell, S. K. J., Ruzzante, D. E., O'Reilly, P. T., Hamilton, L. and Wright, J. M. (1997). Microsatellite loci reveal highly significant genetic differentiation among Atlantic salmon (*Salmo salar* L.) stock from the east coast of Canada. *Molecular Ecology*, **6:** 1075-1089.

McGregor, D., Galvin, P. and Cross, T. (1996). PCR amplification of a polymorphic minisatellite VNTR locus in whiting (*Merlangius merlangus* L.). *Animal Genetics*, **27:** 49-51.

McVeigh, H. P., Hynes, R. A. and Ferguson, A. (1995). Mitochondrial DNA differentiation of sympatric populations of brown trout, *Salmo trutta* L. from Lough Melvin, Ireland. *Canadian Journal of Fisheries and Aquatic Sciences*, **52** (8):1617-1622.

Milton, D. A. and Shaklee, J. B. (1987). Biochemical genetics and population structure of Blue Grenadier, *Macruronus novaezelandiae* (Hector) (Pisces: Merluccidae), from Australian waters. *Australian Journal of Marine and Freshwater Research*, **38:** 727-742.

Moreira, M. H., Queiroga, H., Machado, M. M. and Cunha, M. R. (1993). Environmental gradients in a southern Europe estuarine system: Ria de Aveiro, Portugal. Implications for soft bottom macrofauna colonization. *Netherlands Journal of Aquatic Ecology*, **27**(2-4): 465-482.

Morizot, D. C. and Schmidt, M. E. (1990). Starch gel electrophoresis and histochemical visualization of proteins. *In*: *Electrophoretic and Isoelectric Focusing Techniques in Fisheries Management* (Whitmore, D. H., ed.). CRC Press.

Mork, J. and Haug, T. (1983). Genetic variation in halibut (*Hippoglossus hippoglossus* L.) from Norwegian waters. *Hereditas*, **98:** 167-174.

Mork, J., Ryman, S., Utter, F. and Sundnes, G. (1985). Genetic variation in Atlantic cod (*Gadus morhua*) throughout its range. *Canadian Journal of Fisheries and Aquatic Sciences*, **42:** 1580-1587.

Morris, D. B., Richard, K. R. and Wright, J. M. (1995). Microsatellites from rainbow trout (*Oncorhynchus mykiss*) and their use for genetic studies of salmonids. *Canadian Journal of Fisheries and Aquatic Sciences*, : .

Morrison, K. E., Daniels, R. J., Suthers, G. K., Flynn, G. A., Francis, M. J., Grewal, P. K., Dennis, C., Buckle, V. and Davies, K. E. (1993). Two novel microsatellite markers for prenatal prediction of spinal muscular atrophy (SMA). *Human Genetics*, **92**(2)**:** 133-138.

Mulligan, T. J., Chapman, R. W. and Brown, B. L. (1992). Mitochondrial DNA analysis of walleye pollock, *Theragra chalcogramma* , from the eastern Bering Sea and Shelikof Strait, Gulf of Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, **49:** 319-326.

Mullis, K. B. and Faloona, F. (1987). Specific synthesis of DNA in vitro via polymerase catalyzed chain reaction. *Methods Enzymol*, **155:** 335-350.

Murphy, R. W., Sites, J. W., Buth, D. G. and Haufler, C. H. (1990). Proteins I: Isozyme electrophoresis. *In*: *Molecular Systematics* (Hillis, D. M., and Moritz, C., eds.). Sinauer Associates, Massachussets.

Naciri, M., Lemaire, C., Borsa, P. and Bonhomme, F. (submitted). A DNA-microsatellite analysis of the genetic transition around Gibraltar Strait in the European sea bass *Dicentrarchus labrax*.

Nakamura, Y., Leppert, M., O' Connell, P., Wolff, R., Holm, T., Culver, M. and Martin, C. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science*, **235**(4796)**:** 1616-1621.

Nei, M. (1972). Genetic distance between populations. *American Naturalist*, **106:** 283-292.

Nei, M. (1975). *Molecular Population Genetics and Evolution.* North-Holland, Amsterdam.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89:** 583-590.

Nei, M. (1987). Genetic distance and molecular phylogeny. *In*: *Population Genetics & Fishery Management* (Ryman, N., and Utter, F. W., eds.). Washington Sea Grant Program, University of Washington, Seattle.

Nielsen, J. L., Gan, C. A., Wright, J. M., Morris, D. and Thomas, W. K. (1994). Biogeographic distributions of mitochondrial and nuclear markers for southern steelhead. *Molecular Marine Biology and Biotechnology*, **3**(5)**:** 281-293.

Nogueira, A. J. A. (1992). ANAMOD - Extracção dos componentes modais de distribuições de frequências de variáveis biométricas. Provas de Aptidão Pedagógica e Capacidade Científica Thesis. Universidade do Algarve, Faro.

O'Connell, M., Dillon, M. C. and Wright, J. M. (1998a). Development of primers for polymorphic microsatellite loci in the Pacific herring (*Clupea harengus pallasi*). *Molecular Ecology*, **7:** 358-360.

O'Connell, M., Dillon, M. C., Wright, J. M., Bentzen, P., Merkouris, S. and Seeb, J. (1998b). Genetic structuring among Alaskan Pacific herring populations identified using microsatellite variation. *Journal of Fish Biology*, **53:** 150-163.

O'Connell, M. and Wright, J. M. (1997). Microsatellite DNA in fishes. *Reviews in Fish Biology and Fisheries*, **7:** 331-363.

Ohta, T. (1992). The nearly neutral theory of molecular evolution. *Annual Reviews in Ecology and Systematics*, **23:** 263-286.

Ohta, T. and Kimura, M. (1973). The model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a genetic population. *Genetic Research*, **22:** 201-204.

Olsen, J. B., Wenburg, J. K. and Bentzen, P. (1996). Semiautomated multilocus genotyping of Pacific salmon (*Oncorhynchus* spp.) using microsatellites. *Molecular Marine Biology and Biotechnology*, **5**(4)**:** 259-272.

Olson, T. M., Michels, V. V., Lindor, N. M., Pastores, G. M., Weber, J. L., Schaid, D. J., Drsicoll, D. J., Feldt, R. H. and Thibodeau, S. N. (1993). Autosomal dominant

supravalvular aortic stenosis: localization to chromosome 7. *Human Molecular Genetics*, **2**(7): 869-873.

O'Reilly, P. and Wright, J. M. (1995). The evolving technology of DNA fingerprinting and its application to fisheries and aquaculture. *Journal of Fish Biology*, **47**(Suppl A): 29-55.

Ovenden, J. R. (1990). Mitochondrial DNA and marine stock assessment: a review. *Australian Journal of Marine and Freshwater Research*, **41:** 835-853.

Ovenden, J. R., Smolenski, A. J. and White, R. W. G. (1989). Mitochondrial DNA restriction site variation in Tasmanian populations of orange roughy (*Hoplostethus atlanticus*), a deep-water marine teleost. *Australian Journal of Marine and Freshwater Research*, **40:** 1-9.

Pannacciulli, F. G., Bishop, J. D. D. and Hawkins, S. J. (1997). Genetic structure of populations of two species of *Chthamalus* (Crustacea: Cirripedia) in the north-east Atlantic and Mediterranean. *Marine Biology*, **128:** 73-82.

Park, L. K. and Moran, P. (1995). Developments in molecular genetic techniques in fisheries. *In*: *Molecular Genetics in Fisheries* (Carvalho, G. R., and Pitcher, T. J., eds.). Chapman and Hall, London.

Pasteur, N., Pasteur, G., Bonnhomme, F., Catalan, J. and Britton-Davidian, J. (1987). *Manuel technique de génétique par electrophorèse des proteínes*. Technique et documentation. Lavoisier, Paris.

Patarnello, T., Bargelloni, L., Caldara, F. and Colombo, L. (1993). Mitochondrial DNA sequence variation in the European seabass, *Dicentrarchus labrax* L.: evidence of differential haplotype distribution in natural and farm population. *Molecular Marine Biology and Biotechnology*, **2**(6): 333-337.

Pawson, M. G., Kelley, D. F. and Pickett, G. D. (1987). The distribution and migrations of bass, *Dicentrarchus labrax* L. in waters around England and Wales as shown by tagging. *Journal of the Marine Biology Association of the U.K.*, **67:** 183-217.

Pawson, M. G. and Pickett, G. D. (1987). The bass (*Dicentrarchus labrax*) and management of its fishery in England and Wales. Ministry of Agriculture Fisheries and Food. Lowestoft.

Pawson, M. G. and Pickett, G. D. (1990). Management of bass (Dicentrarchus labrax L.) fisheries in the U.K. *In*: *Fisheries in the year 2000. Proceedings of the 21st*

anniversary conference of the Institute of Fisheries Management (O'Grady, K., Butterworth, A. J. B., Spillett, P. B., and Domaniewsi, J. C. J., eds.), Royal Holloway and Bedford New College. 189-198.

Pemberton, J. M., Slate, J., Bancroft, D. R. and Barret, J. A. (1995). Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology*, **4:** 249-252.

Pickett, G. D. and Pawson, M. G. (1994). *Sea Bass. In*: *Sea Bass*. Chapman & Hall, London.

Postlewait, J. H., Johnson, S. L., Midson, C. N., Talbot, W. S., Gates, M., Ballinger, E. W., Africa, D., Andrews, R., Carl, T., Eisen, J. S., Horne, S., Kimmel, C. B., Hutchinson, M., Johnson, M. and Rodriguez, A. (1994). A genetic linkage map for the zebrafish. *Science*, **264:** 699-703.

Powers, D. A. (1990). The adaptive significance of allelic isozyme variation in natural populations. *In*: *Electrophoretic and Isoelectric Focusing Techniques in Fisheries Management* (Whitmore, D. H., ed.). CRC Press.

Prodöhl, P. A., Taggart, J. B. and Ferguson, A. (1992). Genetic variability within and among sympatric brown trout (*Salmo trutta*) populations: multi-locus DNA fingerprinting analysis. *Hereditas*, **117:** 45-50.

Prodöhl, P. A., Taggart, J. B. and Ferguson, A. (1994). Single locus inheritance and joint segregation analysis of minisatellite (VNTR) DNA loci in brown trout (*Salmo trutta*). *Heredity*, **73:** 556-566.

Purdue, M. L. and Gall, J. (1970). Chromosomal localization of mouse satellite DNA. *Science*, **168:** 1356-1358.

Queller, D. C., Strassmann, J. E. and Hughes, C. R. (1993). Microsatellites and kinship. *Trends in Ecology and Evolution*, **8**(8)**:** 285-288.

Quintino, V. and Rodrigues, A. M. (1989). Environment gradients and distribution of macrozoobenthos in three Portuguese coastal systems: Obidos, Albufeira and Alvor. *In*: *Reproduction, Genetics And Distributions of Marine Organisms.* (Ryland, J. S., and Tyler, P. A., eds.). Int. Symp. Ser. 441-450.

Rao, M. R. K., Padhi, B. K. and Khuda-bukhsh, A. R. (1989). Lactate dehydrogenase isozymes in fifty-two species of teleostean fishes: taxonomic significance of Ldh-C gene expression. *Biochemical Systematics Ecology*, **17**(1)**:** 69-76.

Raymond, M. and Rousset, F. (1995a). An exact test for population differentiation. *Evolution*, **41:** 823-833.

Raymond, M. and Rousset, F. (1995b). Genepop (ver 1.2): a population genetics software for exact tests and ecuminicism. *Journal of Heredity*, **86:** 248-249.

Rebelo, J. E. (1992). The ichthyofauna and abiotic hydrological environment of the Ria de Aveiro, Portugal. *Estuaries*, **15**(3)**:** 403-413.

Reynolds, J., Weir, B. S. and Cockerham, C. C. (1983). Estimation of co-ancestry coefficient: basis for a short-term genetic distance. *Genetics*, **105:** 767-779.

Rice, W. R. (1989). Analysing tables of statistical tests. *Evolution*, **43**(1)**:** 223-225.

Richards, R. I. and Sutherland, G. R. (1994). Simple repeat DNA is not replicated simply. *Nature Genetics*, **6**(2)**:** 114-116.

Richardson, B. J. (1983). Distribution of protein variation in Skipjack Tuna (*Katsuwonus pelamis*) from the central and south-western Pacific. *Australian Journal of Marine and Freshwater Research*, **34:** 231-251.

Rico, C., Inbrahim, K. M., Rico, I. and Hewitt, G. M. (1997). Stock composition in North Atlantic population of whiting using microsatellite markers. *Journal of Fish Biology*, **51:** 462-475.

Rico, C., Zadworny, D., Kuhnlein, U. and Fitzgerald, G. J. (1993). Characterization of hypervariable microsatellite loci in the threespine stickleback *Gasterosteus aculeatus*. *Molecular Ecology*, **2**(4)**:** 271-272.

Riess, O., Winkelmann, B. and Epplen, J. T. (1994). Toward the complete genomic map and molecular pathology of human chromosome 4. *Human Genetics*, **94**(1)**:** 1-18.

Robic, A., Dubois, C., Milan, D. and Gellin, J. (1994). A rapid method to isolate microsatellite markers from cosmid clones. *Mammalian Genome*, **3:** 177-179.

Roldan, M. I. and Pla, C. (1995). More data about genetic variability in Atlantic and Mediterranean populations of European hake, *Merluccius merluccius*: evidence for a mixing zone. *In*: *Molecular Biology in Fish, Fisheries and Aquaculture* (Matthews, R. A., and Thorpe, J. E., eds.). Fisheries Society of the British Isles, Plymouth, U.K. .

Rousset, F. (1996). Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics*, **142:** 1357-1362.

Rousset, F. and Raymond, M. (1995). Testing heterozygote excess and deficiency. *Genetics*, **140:** 1413-1419.

Rousset, F. and Raymond, M. (1997). Statistical analysis of population genetic data: new tools, old concepts. *Trends in Ecology and Evolution*, **12**(8): 313-317.

Ruzzante, D. E., Taggart, C. T. and Cook, D. (1996a). Spatial and temporal variation in the genetic composition of a larval cod (*Gadus morhua*) aggregation: cohort and genetic stability. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**(12): 2695-2705.

Ruzzante, D. E., Taggart, C. T., Cook, D. and Goddard, S. V. (1996b). Genetic differentiation between inshore and offshore Atlantic cod (*Gadus morhua*) off Newfoundland: a test, and evidence of temporal stability. *Canadian Journal of Fisheries and Aquatic Sciences*, **53:** 634-645.

Ryman, N. and Utter, F. (1987). *Population Genetics & Fishery Management* (Ryman, N., and Utter, F., eds.). Washington Sea Grant Program, Washington.

Sabriye, A. S., Reay, P. J. and Coombs, S. H. (1988). Sea-bass larvae in coastal and estuarine plankton. *Journal of Fish Biology*, **33**(Supp.A): 231-233.

Saitou, N. and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenies. *Molecular Biology and Evolution*, **4:** 406-425.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular cloning, a laboratory manual*. Cold Spring Harbour Laboratory Press, New York.

San Feliu, J. M. (1987). La lubina, *Dicentrarchus labrax* (L). *In*: *La acuicultura marina en la Comunidad Valenciana*. Exmo. Aymto. Castellón de la Plana.

Schneider, S., Kueffer, J.-M., Roessli, D. and Excofier, L. (1997). Arlequin: A software for population genetic data analysis. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva, Geneva.

Shaklee, J. B., Allendorf, F. W., Morizot, D. C. and Whitt, G. S. (1990a). Gene nomenclature for protein-coding loci in fish. *Transactions of the American Fisheries Society*, **119:** 2-15.

Shaklee, J. B., Phelps, S. R. and Salini, J. (1990b). Analysis of fish structure and mixed-stock fisheries by the electrophoretic characterization of allelic isozymes. *In*: *Electrophoretic and Isoelectric Focusing Techniques in Fisheries Management* (Whitmore, D. H., ed.).

Shaklee, J. B. and Samollow, P. B. (1984). Genetic variation and population structure in a deepwater snapper, *Pristimoides filamentosus*, in the Hawaiian archipelago. *Fisheries Bulletin*, **82:** 703-713.

Shaklee, J. B., Tamaru, C. S. and Waples, R. S. (1982). Speciation and evolution of marine fishes studied by the electrophoretic analysis of proteins. *Pacific Science*, **36**(2)**:** 141-157.

Shaw, S. A. and Curry, A. (1989a). Markets for seabream and seabass. *In*: *Markets in Europe for selected aquaculture species: salmon, trout, seabream, seabass* (Globefish, F., ed.), Roma.

Shaw, S. A. and Curry, A. (1989b). *Markets in Europe for selected aquaculture species: Salmon, trout, seabream, seabass. In*: *Globefish Res. Programme.* Ices Mar. Sci. Symp.

Shriver, M. D., Jin, L., Boerwinkle, E., Deka, R., Ferrell, R. E. and Chakraborty, R. (1995). A novel measure of genetic distance for highly polymorphic tandem repeat loci. *Molecular Biology and Evolution*, **12**(5)**:** 914-920.

Sick, K. (1961). Haemoglobin polymorphism in fishes. *Nature*, **192:** 894-896.

Singer, M. F. (1982). Highly repeated sequences in mammalian genomes. *International Reviews of Cytology*, **76:** 67-112.

Skibinski, D. O. F., Gallagher, C. and Beynon, C. M. (1994). Mitochodrial DNA inheritance. *Nature*, **368:** 817-818.

Slatkin, M. (1985). Rare alleles as indicators of gene flow. *Evolution*, **39**(1)**:** 53-65.

Slatkin, M. (1994). An exact test for neutrality based on the Ewens sampling distribution. *Genetic Research*, **64:** 71-74.

Slatkin, M. (1996). A correction to the exact test based on the Ewens sampling distribution. *Genetic Research*, **68:** 259-260.

Slettan, A., Olsaker, I. and Lie, O. (1993). Isolation and characterization of variable (GT) sub(n) repetitive sequences from Atlantic salmon, *Salmo salar* L. *Animal Genetics*, **24**(3)**:** 195-197.

Slettan, A., Olsaker, I. and Lie, O. (1997). Segregation studies and linkage analysis of Atlantic salmon microsatellites using haploid genetics. *Heredity*, **78**(6)**:** 620-627.

Small, M. P., Beacham, T. D., Withler, R. E. and Nelson, J. (1998). Discriminating coho salmon (*Oncorhynchus kisutch*) populations within the Fraser River, British Columbia, using microsatellite DNA markers. *Molecular Ecology*, **7:** 141-155.

Smith, A. C. and Clemens, H. B. (1973). A population study by proteins from the nucleus of bluefin tuna eye lens. *Transactions of the American Fisheries Society*, **102:** 578.

Smith, P. J., Birley, A. J. and Jamieson, A. (1989). Electrophoretic studies and the stock concept in marine fishes. *Journal of Fish Biology*, **35:** 345-346.

Smith, P. J. and Robertson, D. A. (1981). Genetic evidence for two species of sprat (*Sprattus*) in New Zealand waters. *Marine Biology*, **62:** 227-233.

Smithies, O. (1955). Zone electrophoresis in starch gels: group variations in the serum proteins of normal individuals. *Biochemistry Journal*, **61:** 629-641.

Smolenski, A. J., Ovenden, J. R. and White, R. W. G. (1993). Evidence of stock separation in southern hemisphere orange roughy (*Hoplostethus atlanticus*, Trachidae). *Marine Biology*, **116:** 219-230.

Sneath, P. H. A. and Sokal, R. R. (1973). *Numerical taxonomy*. W.H.Freeman, San Francisco.

Snowdon, R. J. and Adam, K. D. (1992). Allozyme variation in Rainbow-trout from Central North Island, New-Zealand. *Transactions of the American Fisheries Society*, **121**(6): 810-814.

Sodsuk, P. K., McAndrew, B. J. and Turner, G. F. (1995). Evolutionary relationships of the Lake Malawi *Oreochromis* species: evidence from allozymes. *Journal of Fish Biology*, **47:** 321-333.

Ståhl, G. (1983). Differences in the amount and distribution of genetic variation between natural populations and hatchery stocks of Atlantic salmon. *Aquaculture*, **33:** 23-32.

Ståhl, G. (1987). Genetic population structure of Atlantic salmon. *In*: *Population Genetics & Fishery Management* (Ryman, N., and Utter, F. M., eds.). Washington Sea Grant Program, Universitandum NMFS-SEFC-199, Seattle.

Stepien, C. A. and Kocher, T. D. (1997). Molecules and morphology in studies of fish evolution. *In*: *Molecular Systematics of Fishes* (Kocher, T. D., and Stepien, C. A., eds.). Academic Press, London.

Stequert, B. (1972). Contribution à l'etude du bar *Dicentrarchus labrax* (Linne) des reservoirs a poissons de la region d'Archachon. Y Thesis. These 3eme cycle, Fac. Sc. Bordeaux, France.

Stone, C. E., Taggart, J. B. and Ferguson, A. (1995). Single locus minisatellite DNA variation in European populations of Atlantic salmon (*Salmo salar* L.). *In*: *Molecular Biology in Fish, Fisheries and Aquaculture* (Matthews, R. A., and Thorpe, J. E., eds.). Fisheries Society of the British Isles, Plymouth, U.K. .

Swofford, D. L., Olsen, G. J., Waddell, P. J. and Hillis, D. M. (1996). Phylogenetic inference. *In*: *Molecular Systematics* (Hillis, D. M., Moritz, C., and Mable, B. K., eds.). Sinauer Associates, Sunderland.

Swofford, D. L. and Selander, R. B. (1989). Biosys-1 (ver. 1.7): a computer program for the analysis of allelic variation in population genetics and biochemical systematics. *Journal of Heredity*, **72:** 281-283.

Taggart, J. B. and Ferguson, A. (1990a). Hypervariable minisatellite DNA single locus probes for the Atlantic salmon *Salmo salar* L. *Journal of Fish Biology*, **37:** 991-993.

Taggart, J. B. and Ferguson, A. (1990b). Minisatellite DNA fingerprints of salmonid fishes. *Animal Genetics*, **21:** 337-389.

Taggart, J. B., Verspoor, E., Galvin, P., Moran, P. and Ferguson, A. (1995). A minisatellite DNA marker for discriminating between European and North American Atlantic salmon (*Salmo salar* L.). *Canadian Journal of Fisheries and Aquatic Sciences*, **52:** 2305-2311.

Takahata, N. (1983). Gene identity and genetic differentiation of populations in the finite island models. *Genetics*, **104:** 497-512.

Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, **17:** 6463-6471.

Tave, D. (1993). *Genetics for Fish Hatchery Managers*. Van Nostrand Reinhold, New York.

Taylor, A. C., Sherwin, W. B. and Wayne, R. K. (1994). Genetic variation of microsatellite loci in a bottlenecked species: the northern hairy-nosed wombat *Lasiorhinus krefftì*. *Molecular Ecology*, **3:** 277-290.

Taylor, G. R., Noble, J. S., Hall, J. L., Stewart, A. D. and Mueller, R. F. (1989). Hypervariable microsatellite for genetic diagnosis. *Lancet*, **2**(8660): 454.

Tessier, N., Bernatchez, L., Presa, P. and Angers, B. (1995). Gene diversity analysis of mitochondrial DNA, microsatellites and allozymes in landlocked Atlantic salmon. *Journal of Fish Biology*, **47**(Suppl. A): 156-163.

Thompson, B. R. and Harrop, R. (1987). The distribution and abundance of bass (*Dicentrarchus labrax*) eggs and larvae in the English Channel and southern North Sea. *Journal of the Marine Biology Association of the U.K.*, **67:** 263-274.

Turan, C. (1997). Population structure of Atlantic herring, *Clupea harengus L.*, in the North-East Atlantic using phenotypic and molecular approaches. PhD Thesis. University of Hull, U.K.

Utter, F. and Ryman, N. (1993). Genetic markers and mixed stock fisheries. *Fisheries*, **18**(8): 11-21.

Utter, F., Teel, D., G., M. and McIsaac, D. (1987a). Genetic estimates of stock composition of 1983 chinook salmon, *Oncorhynchus tshawytscha*, harvest off the Washington coast and the Columbia river. *Fishery Bulletin*, **85:** 13-23.

Utter, F. M. (1981). Biological criteria for definition of species and distinct intraspecific populations of anadroumous salmonids under the U.S. Endangered Species Act of 1973. *Canadian Journal of Fisheries and Aquatic Sciences*, **38:** 1626-1635.

Utter, F. M. (1991). Biochemical genetics and fisheries management: an historical perspective. *Journal of Fish Biology*, **39**(Suppl): 1-20.

Utter, F. M. (1995). Perspectives of molecular genetics and fisheries into the 21st century. *In*: *Molecular Genetics in Fisheries* (Carvalho, G. R., and Pitcher, T. J., eds.). Chapman and Hall, London.

Utter, F. M., Aebersold, P. and Winans, G. (1987b). Interpreting genetic variation detected by electrophoresis. *In*: *Population Genetics & Fishery Management* (Ryman, N., and Utter, F. M., eds.). Washington Sea Grant Program, Universitandum NMFS-SEFC-199, Seattle.

Valdes, A. M., Slatkin, M. and Freimer, N. B. (1993). Allele frequencies at microsatellite loci: The stepwise mutation model revisited. *Genetics*, **133**(3): 737-749.

Van de Peer, Y. and De Wachter, R. (1994). TREECON for Windows: a software package for the construction and drawing of evolutionary tree for the Microsoft Windows environment. *Computer Applications in Bioscience*, **10:** 569-570.

Van Dongen, S. and Backeljau, T. (1997). Bootstrap tests for specific hypothesis at single locus inbreeding coefficients. *Genetica*, **99:** 47-58.

Van Oppen, M. J. H., Rico, C., Deutsch, J. C., Turner, G. F. and Hewitt, G. M. (1997). Isolation and characterization of microsatellite loci in the cichlid fish *Pseudotropheus zebra*. *Molecular Ecology*, **6:** 387-388.

Wahlund, S. (1928). The combination of population and the appearance of correlation examined from the stand point of the study of heredity. *Hereditas*, **11:** 65-106 (in German).

Waples, R. S. (1987). A multispecies approach to the analysis of gene flow in marine shore fishes. *Evolution*, **41**(2)**:** 385-400.

Ward, R. D., Billington, N. and Hebert, P. D. N. (1989). Comparison of allozyme and mitochondrial DNA variation in populations of Walleye, *Stizostedium vitreum*. *Canadian Journal of Fisheries and Aquatic Sciences*, **46:** 2074-2084.

Ward, R. D., Elliot, N. G., Grewe, P. and Smolenski, A. (1994a). Comparison of allozyme and mitochondrial DNA variation in yellowfin tuna (*Thunnus albacares*) from the Pacific Ocean. *Marine Biology*, **118:** 531-536.

Ward, R. D. and Grewe, P. M. (1995). Appraisal of molecular genetic techniques in fisheries. *In*: *Molecular Genetics in Fisheries* (Carvalho, G. R., and Pitcher, T. J., eds.). Chapman and Hall, London.

Ward, R. D., Woodwark, M. and Skibinski, D. O. F. (1994b). A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. *Journal of Fish Biology*, **44**(2)**:** 213-232.

Wassef, E. and El Emary, H. (1989). Contribution to the biology of bass *Dicentrarchus labrax* L. in the Egyptian Mediterranean waters off Alexandria. *Cybium (3e serie)*, **13**(4)**:** 327-345.

Weber, J. L. (1990). Informativeness of human (dC-dA)n.(dG-dT)n polymorphisms. *Genomics*, **7:** 524-530.

Weber, J. L. and May, P. E. (1989). Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, **44:** 388-396.

Weber, J. L. and Wong, C. (1993). Mutation of human tandem repeats. *Human Molecular Genetics*, **2:** 1123-1128.

Weinberg, W. (1908). Uber Vererbungsgesetze beim Menschen. *Z. Abst. V. Vererb.*, **1:** 277-330.

Weir, B. S. (1996). Intraspecific differentiation. *In*: *Molecular Systematics* (Hillis, D. M., and Moritz, C., eds.). Sinauer Associates, Massachusetts.

Weir, B. S. and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, **38:** 1358-1370.

Wenburg, J. K., Olsen, J. B. and Bentzen, P. (1996). Multiplexed systems of microsatellites for genetic analysis in coastal cutthroat trout (*Oncorhynchus clarki clarki*) and steelhead (*Oncorhynchus mykiss*). *Molecular Marine Biology and Biotechnology*, **5**(4)**:** 273-283.

Whitehead, P. J. and Wheeler, A. C. (1966). The generic names used for the sea bass of Europe and North America (Pisces, Serranidae). *Annali di Museo Civile da Storia Naturale Genova*, **76:** 23-40.

Whitehead, P. J. P., Bauchot, M. L., Hureau, J. C., Nielsen, J. and Tortonese, E. (1986). *Fishes of the North-eastern Atlantic and the Mediterranean* (UNESCO, ed.).

Wilson, G. M., Thomas, W. K. and Beckenbach, A. T. (1987). Mitochondrial DNA analysis of Pacific northwest populations of *Oncorhynchus tshawytscha*. *Canadian Journal of Fisheries and Aquatic Sciences*, **44:** 1301-1305.

Wirgin, I. I., Grunwald, C., Garte, S. J. and Mesing, C. (1991). Use of DNA fingerprinting in the identification and management of a striped bass population in the southeastern United States. *Transactions of the American Fisheries Society*, **120**(3)**:** 273-282.

Wirgin, I. I., Proenca, R. and Grossfield, J. (1989). Mitochondrial DNA diversity among populations of striped bass in the southeastern United States. *Canadian Journal of Zoology*, **67:** 891-907.

Wood, C. C., Rutherford, D. T. and McKinnell, S. (1989). Identification of sockeye salmon (*Oncorhynchus nerka*) stocks in mixed stock fisheries in British Columbia

and Southeast Alaska using biological markers. *Canadian Journal of Fisheries and Aquatic Sciences*, **46:** 2108-2120.

Wood, S., Schertzer, M., Hayden, M. and Ma, Y. (1993). Support for founder effect for two lipoprotein lipase (LPL) gene mutations in French Canadians by analysis of GT microsatellites flanking the LPL gene. *Human Genetics*, **91**(4)**:** 312-316.

Workman, P. L. and Niswander, J. D. (1970). Studies on southwestern Indian tribes. II. Local genetic differentiation in the *Papago*. *American Journal of Human Genetics*, **22:** 24-49.

Wright, J. M. (1993). DNA fingerprinting of fishes. *In*: *Biochemistry and Molecular Biology of Fishes* (Hochachka, and Mommsen, eds.). Elsevier Science Publishers, Amsterdam.

Wright, J. M. (1994). Mutation at VNTRs: are minisatellites the evolutionary progeny of microsatellites? *Genome*, **37:** 345-347.

Wright, S. (1943). Isolation by distance. *Genetics*, **28:** 114-138.

Wright, S. (1951). The genetical structure of populations. *Annual Eugenics*, **15:** 323-354.

Wright, S. (1969). The theory of gene frequencies. *In*: *Evolution and the Genetics of Populations*. University of Chicago Press, Chicago, Illinois.

Yahyaoui, A., Bruslé, J. and Pasteur, N. (1983). Étude du polimorphisme biochimique de deux population naturelles (MAROC Atlantique et Roussillon) de civelles et anguillettes d' *Anguilla anguilla* L. *In*: *Actes de Colloques: Bases biochimiques de l'aquaculture* (IFREMER, ed.), Montpellier. 373-390.

Yanish-Perron, C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp 18 and pUC19 vectors. *Gene*, **33:** 103-119.

Yeh, F. C., Yang, R.-C., Boyle, T., Ye, Z. H. and Mao, J. X. (1997). POPGENE. Molecular biology and Technology Centre, University of Alberta, Alberta, Canada.

Youngson, A. F., Knox, D. and Johnstone, R. (1992). Wild adult hybrids of *Salmo salar* L. and *Salmo trutta* L. *Journal of Fish Biology*, **40:** 817-20.

Zar, J. H. (1984). *Biostatistical analysis*. Prentice Hall, Englewoods Cliff, New Jersey.

ANNEXES

ANNEX 1

BUFFERS AND AND STAIN RECIPES FOR ALLOZYME ELECTROPHORESIS

1. Electrode and Gel Buffers

CTC	**Continuous Tris-citrate** (Pasteur *et al.*, 1987)	
Electrode:	0.25 *M* Tris	30.29 g/l; 75.73 g/2l; 151.45g/5l
	0.075 *M* Citric acid	15.76 g/l; 39.40 g/2l; 78.80g/5l
	pH 8.0	
Gel:	Dilute electrode buffer 1:25 (12.8ml:320ml) (9.2ml:230ml)	
	pH 8.0	
Both pH buffers adjusted with NaOH		

TEB	**Tris EDTA Borate** (McAndrew and Majumdar, 1983)	
Electrode:	0.50 *M* Tris	60.57 g/l; 151.43 g/2.5l; 302.85 g/5l
	0.016 *M* EDTA	5.99 g/l; 14.98 g/2.5l; 29.95 g/5l
	0.24 *M* Boric acid	15.00 g/l; 37.50 g/2.5l; 75.00 g/5l
	pH 8.5	
Gel:	Dilute electrode buffer 1:10 **(32 ml:320 ml)** (23ml:230ml)	
	pH 8.5	
Both pH buffers adjusted with Tris or Boric Acid		

CAM	**Citrate Aminopropyl Morpholine** (Clayton and Tretiak, 1972)	
Electrode:	0.04 *M* Citric acid monohydrate	8.406g/l; 21.02g/2.5l; 42.03g/5l
	pH 6.1	
Gel:	Dilute electrode buffer 1:20 **(16 ml:320 ml)** (11.5ml:230ml)	
	pH 6.0	
Both pH buffers adjusted with n-(3-aminopropyl) Morpholine		

DTC	**Discontinuous Tris-citrate** (Ridgway *et al.*, 1970)		
Electrode:	**A:** 0.06 *M* LiOH	2.52 g/l;	5.04g/2l
	0.3 *M* Boric acid	18.55 g/l;	37.10g/2l
	pH 8.1		
	B: 0.03 *M* Tris	3.63 g/l;	7.26g/2l
	0.005 *M* Citric acid	1.05 g/l;	2.10g/2l
	pH 8.5		

Gel is made with 9:1 gel buffer:electrode buffer, i.e., dilute **32 ml** of electrode buffer in **288 ml** of gel buffer. (23ml:207 ml)

2. STAINING SOLUTIONS

NAD (10 mg/ml)

ml	100	80	50	20
g	1	0.8	0.5	0.2

NADP (5 mg/ml)
MTT (5 mg/ml)
NBT (5 mg/ml)

ml	100	80	50	20
g	0.5	0.4	0.25	0.1

PMS (1 mg/ml)

ml	100	80	50	20
g	0.1	0.08	0.05	0.02

Mg Cl_2 $6H_2O$ 0.1 *M*

ml	100	80	50	20
g	2	1.6	1	0.4

DL - Isocitric Acid (50 mg/ml)

ml	100	80	50	25	12.5	10
g	5	4	2.5	1.25	0.63	0.5

DL - Malic Acid (60 mg/ml)

H_2O	ml	100	50
Malic acid	g	26.8	13.4
NaOH	g	16	8.0

! Dangerous reaction !
Pour half the volume of the **water on the acid**. The container must be surrounded by ice. Add slowly the NaOH. Do not let the solution heat. After dissolution, add the remaining water and **adjust pH 7.0**.

$Na_2S_2O_3 5H_2O$ (Sodium thiosulfate)
0.06*M* = 1.5 g/100 ml

Citric acid
0.005*M* = 1.05 g/l, pH 8.5

3. BUFFERS

TRIS-HCl

1 litre	0.025 *M*	0.03*M*	0.05*M*	0.1*M*	0.2*M*	0.2*M*
Tris (g/l)	0.3	3.63	6.05	12.1	24.2	24.2
HCl (to pH)	**8.5**	-	**8.6**	**8.0**	**8.0**	**9.0**

	0.4*M*	0.5*M*	0.5*M*	4*M*?
Tris (g/l)	48.4	60.55	60.55	484
HCl (to pH)	**8.0**	**7.6**	**8.0**	**8.0**

PO_4

	0.05 *M*	0.05 *M*	0.04*M*	0.1*M*	0.1*M*	0.1*M*
NaH_2PO_4 (g/l)	4.05	0.45	3.12	12.5	9.98	5.22
Na_2HPO_4 (g/l)	5.82	16.56	2.84	2.8	5.11	20.25
HCl (to pH)	**6.5**	**7.8**	**7.0**	**6.0**	**6.5**	**7.0**

	0.2*M*	0.2 *M*
NaH_2PO_4 (g/l)	20.8	0
Na_2HPO_4 (g/l)	9.47	28.2
HCl (to pH)	**6.5**	**9.1**

Phosphate Na/HCl, 0.2 *M*, pH 7.5 (adjust pH with HCl 1N)
Na_2HPO_4, anhidre = 28.4 g/l

Acetate/Na OH, 0.15 *M*, pH 5.0

Acetic acid =	9.3 ml or	2.6 ml
Na OH (tablets)=	5.1 g	1.42 g
H_2O = up to	900 ml	250 ml

Citrate-NaOH (Sodium citrate, S-4641 or BDH 30128)
0.1*M* = 29.41 g/l, pH 4.5

4. FIXING SOLUTIONS

Agar fixing

Ethanol	1.0L
H_2O	1.0L
Glacial acetic acid	0.2L

Gel fixing
Glycerol 10%

4. RECIPES

AAT	Aspartate aminotransferase	**E.C.2.6.1.1.**
Dimer		**Liver/CAM (3** loci)
Tris		300 mg
L-Aspartic acid		65 mg
α-Ketoglutaric acid		20 mg
Pyridoxal-5-phosphate (**P5P**)		10 mg
Polyvinyl Polypyrrolidone (**PVP**)		10 mg
Fast blue **BB** salt		25 mg
Distilled water (just before staining)		25 ml
Add 1.5 % agar		10 ml
Incubate at warm temperature		

ACP	**Acid Phosphatase**	E.C.3.1.3.2.
Monomer ("red cell") or Dimer ("tissue")		**Heart/TEB**
Stirling stain		
Incubate gel slice for 30 min in buffer Acetate/NaOH, 0.15M, pH 5.0, 5°C, before revelation.		
Drain the gel. (Pasteur *et al.*, 1987)		
Add:		
Buffer Citrate-NaOH 4.5 pH		25 ml
Fast Blue **BB** salt		50 mg
Stock solution (2.5 ml acetone, 2.5 ml water, 50 mg α-naphthyl phosphate)		5 ml
Add 1.5 % agar		

"Red cell" and "tissue" ACP isozymes.
Red cell show pink zones of activity (Murphy *et al.*, 1990).
"This enzyme system deserves detailed study in fishes". In liver there is a cathodally migrating dimeric enzyme. Weak, anodal zones of activity are detected in several tissues. Liver ACP is probably a uridine 5'-monophosphate phosphohydrolase (Morizot and Schmidt, 1990).

ADA Monomer	**Adenosine deaminase**	E.C 3.5.4.4. **Liver/CAM (1** loci)
Adenosine		40 mg
Distilled water		5 ml
Buffer PO_4 Na/HCl 0.2M pH 7.5		5 ml
Xanthine oxidase (**XOD**)		0.4U=22 µl
Nucleoside phosphorilase (**NP**)		0.2U= 6 µl
MTT		2 ml
PMS		10 ml
Add 1.5 % agar		10 ml

ADH Dimer	**Alcohol dehydrogenase**	E.C.1.1.1.1 **TEB/Liver** (loci **1 and 3**) **TEB/Muscle**(locus**2**)
Buffer Tris/HCl 0.05M pH 8.6		10 ml
Absolute ethanol		300 µl
Pyruvic acid (inhibits LDH)		50 µl
NAD		2 ml
MTT		1 ml
PMS		1 ml
NBT		1 ml
Add 1.5 % agar		10 ml
Incubate in dark for 15 hours minimum		

Usually resolved as a single isozyme in liver tissues, often migrating cathodally. Other locus products may be found in muscle and other tissues (Morizot and Schmidt, 1990). ADH are inhibited by Pyrazol (5 mg). As all NAD-dependent dehydrogenases, substrates specificity should be confirmed to prevent confusion with zones of formazan which appear in the absence of stain substrate (Morizot and Schmidt, 1990).

AH Monomer	**Aconitate hydratase**	E.C. 4.2.1.3. **Heart/CTC (3 loci)**
Cis aconitic acid		1 ml
Buffer Tris 0.5M pH 8.0		10 ml
IDH		2 U = 50 µl
NADP		0.5 ml
Mg Cl_2 1M		2.5 ml
MTT		1 ml
PMS		2 ml
Add 1.5 % agar		10 ml

Mitochondrial and supernatant/cytosolic forms are known (Harris and Hopkinson, 1976). Two loci for AH in vertebrates. Two forms: (1) mitochondrial (muscle), (2) cytosol (liver). There are examples (poeciliids) of "satellite" bands generated more cathodally than the native isozyme (Morizot and Schmidt, 1990).

AK	**Adenylate kinase**	E.C.2.7.4.3.
Monomer		**Liver/CTC (1 loci)**
Glucose		80 mg
ADP		30 mg
Buffer Tris HCl 0.2M pH 8.0		10 ml
$Mg\ Cl_2$ 1M		1 ml
NADP		2 ml
HK		3.5 U = 1 µl
G6PDH		10 U = 5 µl
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

CAT	**Catalase**	E.C. 1.11.1.6.
Tetramer (?)		**Heart/CTC**
Allow the gel slice to warm to **ambient temperature.**		
$Na_2\ S_2\ O_3\ 5H_2O$ (Sodium thiosulphate)		7.5 ml
Hydrogen peroxide 3%		15 ml
H_2O		40 ml
Buffer PO_4 ,0.1 *M*, pH7.0		5 ml
Incubate at ambient temperature for 1 minute (or less if bubbling is observed), drain and **add together**:		
KI (Potassium iodide) 0.09*M*		25 ml
Glacial acetic acid		0.5 ml
H_2O		25 ml
Flush the gel slice quickly with water to remove KI as soon as the white zones of CAT activity are visible against the blue background. **Photograph immediately**. **Do not store**.		
Add 1.5 % agar		10 ml

Sodium thiosulphate: 248.17; 1.49g/100 ml; 9g/500ml
KI: 4.5g/300ml + 6 ml Acetic Acid

CK	**Creatine kinase**	E.C. 2.7.3.2.
Dimer*		**Muscle/TEB**
Phosphocreatine 1*M*		20mg
Glucose		40 mg
ADP		20 mg
Buffer Tris HCl 0.2M pH 8.0		10 ml
HK		15 U = 3 µl
G6PDH		3.5U=17.5µl
NADP		1 ml
Mg Cl_2		250 µl
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

*although muscle CK heterozygotes in most fishes produce only the two homodimeric allozymes (Morizot and Schmidt, 1990).
Muscle creatine kinase present in sufficient quantity to resolve by general protein stain.

DIA	**Diaphorase**	E.C. 1.6.2.2.
?		**Liver/CAM (2 loci)**
2,6-Dichlorophenol-indophenol (using filter paper) : Distilled water*		4 mg : 2.5 ml
Buffer Tris HCl 0.2 *M* pH 8.0		25 ml
NADH		10 mg
MTT		1 ml
Add 1.5 % agar		10 ml
Incubate at warm temperature for 30 min. then move to dark room temperature conditions until next day		

*Solution previously prepared and kept in +6°C.
Poorly understood isozyme systems in fishes. In poeciliid livers has been occasionally observed two zones of activity (Morizot and Schmidt, 1990).

EST	**Esterases**	E.C. 3.1.1.-
?		**Heart/CTC (4 loci)**
Soak gel slice in buffer PO_4 0.05M pH 6.5 at 4°C for 15 min.		
Pour off buffer.		
Buffer PO_4 0.05 pH 6.5		5 ml
Fast blue **RR** salt		15 mg
Add **acetone : α-Naphthyl acetate***		1 ml
Add 1.5 % agar		10 ml

- α-Naphthyl acetate : acetone : H_2O = 1g : 50 ml :50 ml

FBALD Tetramer	**Fructose biphosphate aldolase**	E.C. 4.1.2.13. **TEB/Liver**
(Murphy *et al.*, 1990)		
Buffer Tris HCl 0.2M pH 8.0		50 ml
Fructose -1,6- Diphosphate		80 mg
NAD		2 ml
Add before using:		
G-3-αPDH		200 U
MTT		1 ml
PMS		5 ml
Add 1.5 % agar		50 ml

FH Tetramer	**Fumarate hydratase**	E.C. 4.2.1.2. **Muscle/CTC (2 loci)**
Fumaric acid		60 mg
Buffer Tris HCl 0.5*M* pH 8.0		10 ml
Sodium pyruvate		1 ml
NAD		2 ml
MDH		60 U = 7 µl
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

MDH may be visualised, control slice stained for MDH may be necessary to identify zones of FH activity (Morizot and Schmidt, 1990).

G3PDH Dimer	**Glycerol-3-phosphate dehydrogenase**	E.C. 1.1.1.8 **Muscle(locus 1)/CTC Liver (locus 2)/TEB**
DL-α-Glycerophosphate (Na_2 salt) 1.0*M**		140 mg
Buffer Tris HCl 0.2*M* pH 8.0		10 ml
NAD		2 ml
PMS		1 ml
NBT		2 ml
Add 1.5 % agar		10 ml

G6PDH *	**Glucose-6-Phosphate dehydrogenase**	E.C. 1.1.1.49 **Liver/CTC (2 loci)**
Glucose-6-Phosphate		20 mg
Buffer Tris HCl 0.2*M* pH 8.0		10 ml
NADP		1 ml
$MgCl_2$ 1*M*		1 ml
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

*Perhaps a mixture of dimers and tetramers in fishes (Morizot and Schmidt, 1990).

GLUDH **	**Glutamate dehydrogenase**	E.C. 1.4.1.-.* **Liver/CTC (1 locus)**
L-Glutamic acid Na		100 mg
Buffer Tris HCl 0.2*M* pH 8.0		10 ml
NADP		1 ml
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

* E.C. 1.4.1.2. (Murphy *et al.*, 1990).
E.C. 1.4.1.3. (Morizot and Schmidt, 1990).

** Tetramer (Murphy *et al.*, 1990).
Hexameric (Aebersold *et al.*, 1987).

A single isozyme is usually observed, expressed most strongly in liver tissue extracts, expression is often week, and care should be taken to avoid confusion with other NAD-dependent or "nothing" dehydrogenases (Morizot and Schmidt, 1990).

GPI Dimer	**Glucose Isomerase**	**Phosphate**	E.C. 5.3.1.9 **Muscle/CTC (2 loci)**
Frutose-6-Phosphate			20 mg
Buffer Tris HCl 0.2*M* pH 8.0			10 ml
Glucose-6-Phosphate Dehydrogenase			1.4 U =10 µl
NADP			1 ml
$MgCl_2$ 1*M*			100 µl
MTT			1 ml
PMS			1 ml
Add 1.5 % agar			10 ml

Interlocus heterodimers are usually formed (Morizot and Schmidt, 1990).
Staining usually reaches maximum intensity in <5 minutes even at room temperature and should be closely monitored (Morizot and Schmidt, 1990).

HK	**Hexokinase**	E.C. 2.7.1.1
Dimer		**Liver/CTC (1 locus)**
Glucose		100 mg
ATP		20 mg
Buffer Tris HCl 0.2*M* pH 8.0		10 ml
NADP		2 ml
$MgCl_2$ 1*M*		1 ml
Glucose-6-Phosphate Dehydrogenase		5U = µl
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

Fructose or Mannose instead of Glucose gave the same results.

IDDH	**Iditol Dehydrogenase**	E.C. 1.1.1.14.
?		**CTC/Liver (2 loci)**
Sorbitol		350 mg
Buffer Tris HCl 0.1*M* pH 8.0		10 ml
Pyruvate (inhibits LDH)+Pyrazol		1 ml
NAD		2 ml
MTT		0.5 ml
PMS		0.5 ml
NBT		1 ml
Add 1.5 % agar		25 ml

A single gene product is expressed predominantly in liver. Staining for other NAD-dependent or "nothing" dehydrogenases may occur (inhibitors of LDH and ADH should be included in the stain recipe) (Morizot and Schmidt, 1990).

IDHP	**Isocitrate dehydrogenase**	E.C. 1.1.1.42.
Dimeric		**Heart/CTC (2 loci)**
Buffer Tris HCl 0.2*M* pH 8.0		10 ml
Mg Cl_2		1 ml
NADP		1 ml
DL-Isocitric acid Na		1 ml
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

LDH	Lactate dehydrogenase	E.C. 1.1.1.27. **Eye/CTC (3 loci)**
Buffer Tris HCl 0.2*M* pH 8.0		5 ml
DL-Lactate Na		1 ml
NAD		1 ml
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

Three LDH gene products have been well characterised in fishes, two with broad tissue specificity and a third restricted to liver or eye, depending upon taxon. Heterotetramer formation among locus products ranges from absent to unrestricted (Morizot and Schmidt, 1990).

MDH	Malate dehydrogenase	E.C. 1.1.1.37 **Muscle/CTC (2 loci)**
Buffer Tris HCl 0.2*M* pH 8.0		5 ml
DL-Malic acid		1 ml
NAD		1 ml
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

One locus coding mitochondria form and two loci coding supernatant/cytosolic form (Harris and Hopkinson, 1976; Morizot and Schmidt, 1990). All three can usually be resolved using muscle tissue.

MEP	Malic enzyme	E.C. 1.1.1.40. **Muscle/CTC (2 loci)**
Buffer Tris HCl 0.2*M* pH 8.0		10 ml
0.1 $MgCl_2$ $6H_2O$		1 ml
DL-Malic acid		2 ml
NADP		1 ml
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

Mitochondrial and supernatant/cytosolic forms are known (Harris and Hopkinson, 1976). There is often sufficient breakdown of NADP to NAD in liquid stocks in prolonged storage that NAD-dependent MDH activity will be resolved in addition to MEP. If there is any doubt as to the identity of the MEP, a control slice from the same gel should be stained specifically for MDH, to ascertain which zones of activity are MDH (Murphy *et al.*, 1990).

ME electrophoretic patterns in fishes are often complex and difficult to assign to specific loci (Morizot and Schmidt, 1990).

Zones of MDH activity often appear on ME-stained gels, necessitating control gel slices stained for MDH (Morizot and Schmidt, 1990).

MPI	Mannose Phosphate Isomerase	E.C. 5.3.1.8. Liver/CTC (2 loci)
Mannose-6-Phosphate		20 mg
Buffer Tris HCl 0.2*M* pH 8.0		10 ml
Glucose-6-phosphate dehydrogenase		10 U= 4µl
Glucose-6-Phosphate isomerase		25 U= 9µl
NADP		1 ml
Mg Cl_2 1*M*		50 µl
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

One locus product is present in virtually all tissues (Morizot and Schmidt, 1990). Products of LDH may appear as faint bands following staining. LDH activity can be suppressed by adding 50 mg of pyruvic acid (Murphy *et al.*, 1990).

PEPA (A) or (B)	Dipeptidase	E.C. 3.4.-.-. Liver/TEB (2 loci)
Peroxidase		10 mg
L-aminoacid oxidase		5 mg
3-amino-9-ethylcarbazole DMSO	**dissolved** in 3ml	12.5 mg
Buffer Tris HCl 0.2*M* pH 8.0		10 ml
Mg Cl_2 1*M*		1 ml
LEU-LEU (A) or LEU-GLY (B)		20 mg
Add 1.5 % agar		10 ml

DMSO=Dimetylsulphoxide

PGDH Dimer	6-Phosphogluconate dehydrogenase	E.C. 1.1.1.44. Liver/CTC (2 loci)
6-Phosphogluconate-Na		15 mg
Buffer Tris HCl 0.2*M* pH 8.0		5 ml
NADP		1 ml
Mg Cl_2 1*M*		1 ml
MTT		1 ml
PMS		1 ml
NBT		1 ml
Add 1.5 % agar		10 ml

A single gene product is present in virtually every tissue (Morizot and Schmidt, 1990).

PGM	**Phosphoglucomutase**	E.C. 5.4.2.2.
?		**Liver/CAM (3 loci)**
Glucose-1-Phosphate		50 mg
Buffer Tris HCl 0.2*M* pH 8.0		10 ml
NADP		1 ml
Glucose-6-PDH		1.4U=1 µl
Mg Cl_2 1*M*		350 µl
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

Only one locus product (often with more anodal "satellite" bands) has been characterised in poeciliids and is detectable in virtually all tissues. Phenotypes in many other fishes suggest that additional locus products remain to be resolved (Morizot and Schmidt, 1990).

PK	**Pyruvate Kinase**	E.C. 2.7.1.40
Tetramer		**CTC**
ADP		15 mg
D-Glucose		45 mg
Phospho(*enol*)pyruvate		10 mg
Buffer Tris HCl 0.2*M* pH 8.0		10 ml
NAD		1 ml
HK		20U= 4 µl
G6PDH		30U=12µl
Mg Cl_2 1*M*		1 ml
NBT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

PNP	**Purine nucleoside phosphorylase**	E.C. 2.4.2.1.
Trimer		**CTC (1 loci)**
Inosine		58 mg
Buffer PO4 0.1*M* pH 7.0-7.5		10 ml
XO		15µl
MTT		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml

SOD	**Superoxide dismutase**	E.C. 1.15.1.1.
*		**Liver/CTC (1 loci)**

*Mitochondrial (tetramer) and supernatant/cytosolic (dimer) forms are known (Harris and Hopkinson, 1976; Morizot and Schmidt, 1990).

As SOD can often be scored on gels stained for other enzymes producing formazan precipitates as **ADH**, G3PDH, AH, etc., there is no need for a specific stain. Zones of activity are visualised as white or clear bands on a blue background; contrast can be increased by exposure to light.

XDH	**Xanthine dehydrogenase**	E.C. 1.2.1.37.
?		**Liver/CAM (2 loci)**
Hypoxanthine suspended in acetone*		35 mg
Buffer Tris-HCl 0.2*M* pH 8.0		10 ml
NAD		1.5 ml
MTT (and/or NBT)		1 ml
PMS		1 ml
Add 1.5 % agar		10 ml
Completely revealed on next day.		

*Hypoxanthine is quite insoluble, Brewer (1970 in Murphy *et al.*, 1990) recommended heating the substrate in the buffer and wait for cooling. Alternatively, Richardson *et al.* (1986 in Murphy *et al.*, 1990) recommended suspending the hypoxanthine in acetone (40 mg/ml acetone). For some species as much as 1 g of hypoxanthine may be required for optimal resolution (Murphy *et al.*, 1990).

This staining also resolves XO.

ANNEX 2

GENOTYPES OF ALLOZYME POLYMORPHIC LOCI

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
FOZ	November 1993	107	100100	090100	100100	100100	100100	100100
FOZ	November 1993	108	100100	090100	100100	100100	100100	050100
FOZ	November 1993	109	100100	090100	100100	100150	100100	100100
FOZ	November 1993	110	100100	100100	100100	100100	100100	050100
FOZ	November 1993	111	100100	100100	100100	100100	100100	050100
FOZ	November 1993	112	075100	100100	100100	100100	100100	100100
FOZ	November 1993	113	100100	090100	100100	100100	100100	050100
FOZ	November 1993	114	100100	090100	100100	100100	100100	050100
FOZ	November 1993	115	100100	100100	100100	100100	100100	100100
FOZ	November 1993	116	100100	100100	100100	100100	100100	100100
FOZ	November 1993	117	100100	090100	100100	100100	100100	050100
FOZ	November 1993	118	100100	090100	100100	100150	100100	050100
FOZ	November 1993	119	100100	100100	100100	100100	100100	100100
FOZ	November 1993	120	100100	090100	100100	100100	100100	050100
FOZ	November 1993	121	100100	090100	100100	100100	100100	100100
FOZ	November 1993	122	100100	100100	100100	100100	100100	050100
FOZ	November 1993	123	000000	090100	100100	100100	100100	100100
FOZ	November 1993	124	100100	100100	100100	100150	100100	050100
FOZ	November 1993	125	075100	090100	100100	100100	100100	100100
FOZ	November 1993	126	100100	090090	100100	100100	100100	100100
FOZ	November 1993	127	100100	100100	100100	100100	100100	050100
FOZ	November 1993	128	100100	090100	100100	100100	100100	100100
FOZ	November 1993	129	100100	090100	100100	100100	100100	100100
FOZ	November 1993	130	100100	100100	100100	100150	100100	100100
FOZ	November 1993	131	100100	100100	100100	100100	100100	050100
FOZ	November 1993	132	100100	100100	100100	100100	100100	100100
FOZ	November 1993	133	100100	090090	100100	100150	100100	100100
FOZ	November 1993	134	100100	100100	100100	100100	100100	100100
FOZ	November 1993	135	075100	100100	100100	100100	080100	100100
FOZ	November 1993	136	100100	090100	100100	100100	100100	100100
FOZ	November 1993	137	100100	090100	100100	100100	100100	100100
FOZ	November 1993	138	100100	100100	100100	100100	100100	050100
FOZ	November 1993	139	100100	090090	100100	100100	100100	050100
FOZ	November 1993	140	100100	090100	100100	100100	100100	050100
FOZ	November 1993	141	100100	100100	100100	100100	100100	100100
FOZ	November 1993	142	075100	090090	100100	100100	100100	050100
FOZ	November 1993	143	100100	100100	100100	100100	100100	100100
FOZ	November 1993	144	100100	090100	100100	100100	100100	100100
FOZ	November 1993	145	100100	100100	100100	100100	100100	050100
FOZ	November 1993	146	100100	100100	100100	100100	100100	050100
FOZ	November 1993	147	100100	090100	100100	100100	100100	050100
FOZ	November 1993	148	100100	100100	100100	100100	100100	050100
FOZ	November 1993	149	100100	090100	100100	100150	100100	050100
FOZ	November 1993	150	100100	100100	100100	100100	100100	050100
FOZ	November 1993	151	100100	100100	100100	100100	100100	100100
FOZ	November 1993	152	100100	100100	100100	100100	100100	050100
FOZ	November 1993	153	100100	090100	100100	100100	100100	100100
FOZ	November 1993	154	100100	090100	100100	100100	100100	050050
FOZ	November 1993	155	100100	100100	100100	100100	100100	050100
FOZ	November 1993	156	100100	090100	100100	100100	100100	100100
FOZ	November 1993	157	100100	090100	100100	100100	100100	050050
FOZ	November 1993	158	100100	090100	100100	100100	100100	050100

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
FOZ	November 1993	159	100100	100100	100100	100100	100100	050100
FOZ	November 1993	160	100100	100100	100100	100100	100100	100100
FOZ	November 1993	161	100100	090100	100100	100100	100100	100100
FOZ	November 1993	162	100100	100100	100100	100100	100100	050050
FOZ	November 1993	163	100100	100100	100100	100100	100100	050100
FOZ	November 1993	164	100100	090100	100100	100100	100100	050100
FOZ	November 1993	165	100100	090100	100100	100100	100100	100100
FOZ	November 1993	166	100100	090100	100100	100100	100100	100100
FOZ	November 1993	167	100100	100100	100100	100100	100100	100100
FOZ	November 1993	168	100100	090100	100100	100100	100100	100100
FOZ	November 1993	169	100100	090100	100100	100100	100100	050050
FOZ	November 1993	170	100100	100100	100100	100100	100100	100100
FOZ	November 1993	171	100100	100100	100100	100100	100100	100100
FOZ	November 1993	172	100100	100100	100100	100100	100100	100100
FOZ	November 1993	173	100100	090100	100100	100100	100100	050100
FOZ	November 1993	174	100100	100100	100100	100100	100100	050100
FOZ	November 1993	175	100100	090100	100100	100100	100100	100100
FOZ	November 1993	176	100100	090100	100100	100100	100100	100100
FOZ	November 1993	177	100100	090100	100100	100100	100100	100100
FOZ	November 1993	178	100100	100100	100100	100100	100100	100100
FOZ	November 1993	179	100100	100100	100100	100100	100100	100100
FOZ	November 1993	180	100100	100100	100100	100100	100100	050100
FOZ	November 1993	181	100100	100100	100100	100100	100100	100100
FOZ	November 1993	182	100100	090100	100100	100100	100100	050050
FOZ	November 1993	183	100100	100100	100100	100100	100100	100100
FOZ	November 1993	184	100100	090100	100100	100100	100100	100100
FOZ	November 1993	185	100100	100100	100100	100100	100100	100100
FOZ	November 1993	186	100100	090100	100100	100100	100100	100100
FOZ	November 1993	187	100100	090100	100100	100150	100100	050100
FOZ	November 1993	188	100100	090100	100100	100100	100100	100100
FOZ	November 1993	189	100100	100100	100100	100150	100100	050050
FOZ	November 1993	190	100100	090100	100100	100100	100100	100100
FOZ	November 1993	191	100100	090100	100100	100100	100100	050050
FOZ	November 1993	192	100100	090100	100100	100100	100100	100100
FOZ	November 1993	193	100100	100100	100100	100100	100100	050100
FOZ	November 1993	194	045100	090090	100100	100100	100100	100100
FOZ	November 1993	195	100100	100100	100100	100100	100100	100100
FOZ	November 1993	196	100100	100100	100100	100100	100100	100100
FOZ	November 1993	197	100100	100100	100125	100100	100100	100100
FOZ	November 1993	198	100100	090100	100100	100150	100100	100100
FOZ	November 1993	199	100100	100100	100100	100100	100100	100100
FOZ	November 1993	200	100100	090100	100100	100150	100100	050100
FOZ	November 1993	201	100100	090100	100100	100100	100100	050100
FOZ	November 1993	202	100100	100100	100100	100100	100100	050100
FOZ	November 1993	203	100100	100100	100100	100150	100100	100100
FOZ	November 1993	204	100100	090100	100100	100100	100100	050100
FOZ	November 1993	205	100100	090100	100100	100150	100100	100100
FOZ	November 1993	206	100100	100100	100100	000000	000000	000000
FOZ	November 1993	207	100100	090100	100100	100100	100100	000000
FOZ	November 1993	208	100100	090100	100100	100100	100100	000000
FOZ	November 1993	209	100100	090100	100100	100100	100100	000000
FOZ	November 1993	210	100100	090100	100100	100100	100100	000000

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
FOZ	November 1993	211	045100	090100	100100	100100	100100	000000
FOZ	November 1993	212	100100	100100	100100	100100	100100	000000
FOZ	November 1993	213	100100	090090	100100	100100	100100	000000
FOZ	November 1993	214	100100	090090	100100	100100	100100	000000
FOZ	November 1993	215	100100	090100	100100	100100	100100	000000
FOZ	November 1993	216	100100	100100	100100	100100	100100	000000
FOZ	November 1993	217	100100	100100	100100	100100	100100	000000
FOZ	November 1993	218	100100	100100	100100	100100	080100	000000
FOZ	November 1993	219	100100	100100	100100	100150	100100	000000
FOZ	November 1993	220	100100	100100	100100	100100	080100	000000
FOZ	November 1993	221	100100	100100	100100	100100	100100	000000
FOZ	November 1993	222	100100	100100	100100	100100	100100	000000
FOZ	November 1993	223	100100	100100	100100	100100	100100	000000
FOZ	November 1993	224	100100	090100	100100	100100	100100	000000
FOZ	November 1993	225	075100	100100	100100	100100	100100	000000
FOZ	November 1993	226	100100	090100	100100	100100	100100	000000
FOZ	November 1993	227	100100	100100	100100	100100	100100	000000
AVEIRO	February 1993	228	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	229	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	230	100100	000000	100100	100100	100100	000000
AVEIRO	February 1993	231	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	232	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	233	100100	090100	100100	100100	100100	050100
AVEIRO	February 1993	234	100100	100100	100100	100100	080100	050100
AVEIRO	February 1993	235	045100	100100	100100	100100	100100	100100
AVEIRO	February 1993	236	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	237	100100	100100	100100	100150	100100	000000
AVEIRO	February 1993	238	100100	100100	100100	100150	100100	050100
AVEIRO	February 1993	239	100100	100100	100100	100100	100100	000000
AVEIRO	February 1993	240	100100	100100	100100	100100	100100	050050
AVEIRO	February 1993	241	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	242	075100	100100	100100	100100	100110	100100
AVEIRO	February 1993	243	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	244	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	245	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	246	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	247	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	248	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	249	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	250	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	251	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	252	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	253	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	254	100100	090100	100100	100100	100100	050100
AVEIRO	February 1993	255	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	256	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	257	100100	090100	100100	100100	100100	050100
AVEIRO	February 1993	258	100100	100100	100100	100100	080100	100100
AVEIRO	February 1993	259	075100	090100	100100	100150	100100	100100
AVEIRO	February 1993	260	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	261	100100	100100	100100	100100	100100	100100

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
AVEIRO	February 1993	262	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	263	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	264	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	265	100100	090100	100100	100150	100100	100100
AVEIRO	February 1993	266	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	267	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	268	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	269	100100	090100	100100	100100	100100	050100
AVEIRO	February 1993	270	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	271	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	272	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	273	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	274	100100	100100	100100	100150	100100	100100
AVEIRO	February 1993	275	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	276	100100	100100	100100	100100	080100	100100
AVEIRO	February 1993	277	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	278	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	279	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	280	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	281	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	282	100100	100100	100100	100100	080100	100100
AVEIRO	February 1993	283	100100	090090	100100	100100	100100	100100
AVEIRO	February 1993	284	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	285	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	286	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	287	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	288	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	289	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	290	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	291	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	292	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	293	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	294	100100	090090	100100	100150	100100	100100
AVEIRO	February 1993	295	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	296	100100	090090	100125	100100	100100	100100
AVEIRO	February 1993	297	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	298	100100	090100	100100	100100	100100	050100
AVEIRO	February 1993	299	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	300	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	301	100100	100100	100100	100100	080100	100100
AVEIRO	February 1993	302	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	303	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	304	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	305	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	306	045100	090100	100100	100100	100100	050100
AVEIRO	February 1993	307	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	308	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	309	100150	100100	100100	100100	100100	100100
AVEIRO	February 1993	310	100100	090090	100100	100100	100100	100100
AVEIRO	February 1993	311	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	312	100100	100100	100100	100150	100100	100100
AVEIRO	February 1993	313	100100	100100	100100	100100	100100	100100

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
AVEIRO	February 1993	314	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	315	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	316	075100	090100	100100	100150	100100	050100
AVEIRO	February 1993	317	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	318	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	319	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	320	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	321	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	322	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	323	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	324	100100	100100	080100	100100	100100	100100
AVEIRO	February 1993	325	045100	100100	100100	100100	100100	100100
AVEIRO	February 1993	326	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	327	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	328	100100	100100	100100	100100	100100	050050
AVEIRO	February 1993	329	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	330	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	331	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	332	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	333	045100	090090	100100	100100	100100	100100
AVEIRO	February 1993	334	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	335	100100	090090	100100	100100	100100	100100
AVEIRO	February 1993	336	100100	100100	100125	100100	100100	100100
AVEIRO	February 1993	337	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	338	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	339	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	340	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	341	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	342	100100	100100	100100	100100	100100	100100
AVEIRO	February 1993	343	100100	090100	100100	100150	100100	050100
AVEIRO	February 1993	344	100100	090100	100100	100100	100100	100100
AVEIRO	February 1993	345	100100	100100	100100	100100	100100	050100
AVEIRO	February 1993	346	100100	090100	100100	100150	100100	100100
AVEIRO	March 1994	844	100100	100100	100100	100100	100100	050100
AVEIRO	March 1994	845	100100	100100	100100	100100	100100	050050
AVEIRO	March 1994	846	100100	100100	100100	100100	100100	100100
AVEIRO	March 1994	847	100100	100100	100100	100100	100100	100100
AVEIRO	March 1994	848	100100	100100	100100	100100	100100	100100
AVEIRO	March 1994	849	100100	100100	100100	100100	100100	100100
AVEIRO	March 1994	850	100100	090100	100100	100100	100100	100100
AVEIRO	March 1994	851	100100	100100	100100	100100	100100	050100
AVEIRO	March 1994	852	100100	100100	100100	100100	100100	100100
AVEIRO	March 1994	853	100100	090100	100100	100100	100100	000000
AVEIRO	March 1994	854	100100	100100	080100	100100	100100	100100
AVEIRO	March 1994	855	100100	100100	100100	100100	100100	050100
AVEIRO	March 1994	856	100100	100100	100100	100100	100100	100100
AVEIRO	March 1994	857	100100	100100	100100	100100	100100	100100
AVEIRO	March 1994	858	100100	100100	100100	100100	100100	100100
AVEIRO	March 1994	859	100100	100100	100100	100100	100100	100100
AVEIRO	March 1994	860	100100	100100	100100	100100	100100	100100
AVEIRO	March 1994	861	100100	090100	100100	100100	100100	050100
AVEIRO	March 1994	862	100100	090100	100100	100100	100100	100100

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
AVEIRO	March 1994	863	100100	100100	100100	100100	100100	050100
AVEIRO	March 1994	864	100100	090100	100100	100100	100100	050100
AVEIRO	March 1994	865	100100	090100	100100	100100	100100	100100
AVEIRO	March 1994	866	100100	000000	100100	100100	100100	050100
AVEIRO	March 1994	867	100100	000000	100100	100100	100100	050100
AVEIRO	March 1994	868	100100	000000	100100	100100	100100	100100
AVEIRO	March 1994	869	100100	000000	100100	100100	100100	100100
AVEIRO	March 1994	870	100100	090100	100100	100100	100100	100100
AVEIRO	March 1994	871	100100	100100	100100	100100	080100	100100
AVEIRO	March 1994	872	100100	100100	100100	100100	100100	050100
AVEIRO	March 1994	873	100100	000000	100100	100100	100100	100100
AVEIRO	March 1994	874	100100	090100	100100	100100	100100	100100
AVEIRO	March 1994	875	100100	090100	100100	100100	100100	100100
AVEIRO	March 1994	876	100100	090100	100100	100100	100100	100100
AVEIRO	March 1994	877	100100	090100	100100	100100	100100	050100
AVEIRO	March 1994	878	100100	090090	100100	100100	100100	100100
AVEIRO	March 1994	879	100100	090100	080100	100100	100100	100100
AVEIRO	March 1994	880	100100	090100	100100	100100	100100	050100
AVEIRO	March 1994	881	100100	090090	100100	100100	100100	100100
AVEIRO	March 1994	882	100100	090100	100100	100100	100100	050100
AVEIRO	March 1994	883	100100	090090	100100	100100	100100	100100
AVEIRO	March 1994	884	100100	090100	100100	100100	100100	050100
AVEIRO	March 1994	885	100100	090100	100100	100100	100100	100100
AVEIRO	March 1994	886	100100	100100	100100	100100	100100	050050
AVEIRO	March 1994	887	100100	090100	100100	100100	100100	100100
AVEIRO	March 1994	888	100100	090100	100100	100100	100100	100100
AVEIRO	March 1994	889	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	347	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	348	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	349	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	350	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	351	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	352	100100	000000	100100	000000	000000	100100
OBIDOS	February 1993	353	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	354	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	355	100100	000000	100100	100100	100100	100100
OBIDOS	February 1993	356	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	357	100100	100100	100100	100150	100100	100100
OBIDOS	February 1993	358	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	359	100100	100100	100100	100100	100100	050050
OBIDOS	February 1993	360	045100	100100	100100	100100	100100	050100
OBIDOS	February 1993	361	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	362	100150	090100	100100	100100	080100	050100
OBIDOS	February 1993	363	100100	090090	100100	100100	100100	050050
OBIDOS	February 1993	364	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	365	075100	100100	100100	100100	100100	100100
OBIDOS	February 1993	366	100100	100100	100100	100100	080100	100100
OBIDOS	February 1993	367	100100	090100	100100	100100	100100	050100
OBIDOS	February 1993	368	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	369	100100	100100	100100	100100	100100	050050
OBIDOS	February 1993	370	100100	090100	100100	100100	100100	100100

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
OBIDOS	February 1993	371	100100	090100	100100	100100	100100	050050
OBIDOS	February 1993	372	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	373	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	374	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	375	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	376	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	377	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	378	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	379	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	380	045100	100100	100100	100100	100100	100100
OBIDOS	February 1993	381	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	382	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	383	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	384	100100	090100	100100	100100	100100	050050
OBIDOS	February 1993	385	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	386	100100	090090	100100	100100	100100	100100
OBIDOS	February 1993	387	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	388	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	389	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	390	100100	090090	100100	100100	100100	050100
OBIDOS	February 1993	391	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	392	100100	100100	100100	100100	080100	050100
OBIDOS	February 1993	393	100100	090100	100100	100100	100100	050100
OBIDOS	February 1993	394	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	395	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	396	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	397	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	398	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	399	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	400	100100	090100	100100	100150	100100	050100
OBIDOS	February 1993	401	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	402	100100	090100	100100	100100	080100	100100
OBIDOS	February 1993	403	100100	090090	100100	100100	100100	100100
OBIDOS	February 1993	404	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	405	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	406	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	407	100100	000000	100100	100100	100100	050100
OBIDOS	February 1993	408	100100	100100	100100	100100	100100	050050
OBIDOS	February 1993	409	075100	100100	100100	100100	100100	100100
OBIDOS	February 1993	410	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	411	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	412	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	413	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	414	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	415	100100	090100	100100	100150	100100	050100
OBIDOS	February 1993	416	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	417	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	418	100100	090100	100100	100100	100100	050100
OBIDOS	February 1993	419	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	420	100100	090090	100100	100100	100100	100100
OBIDOS	February 1993	421	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	422	100100	100100	100100	100100	100100	100100

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
OBIDOS	February 1993	423	100100	090090	100100	100100	100100	100100
OBIDOS	February 1993	424	100100	090100	100100	100100	100100	050100
OBIDOS	February 1993	425	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	426	100100	090090	100100	100100	100100	100100
OBIDOS	February 1993	427	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	428	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	429	100100	090100	100100	100100	100100	050100
OBIDOS	February 1993	430	100100	090090	100100	100100	100100	100100
OBIDOS	February 1993	431	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	432	100100	100100	100100	100150	100100	050050
OBIDOS	February 1993	433	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	434	100100	090100	100100	100100	100100	100100
OBIDOS	February 1993	435	100100	090100	100100	100150	100100	050100
OBIDOS	February 1993	436	100100	090100	100100	100100	100100	050100
OBIDOS	February 1993	437	100100	090100	100100	100100	100100	050100
OBIDOS	February 1993	438	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	439	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	440	100100	100100	100100	100150	100100	050100
OBIDOS	February 1993	441	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	442	100100	100100	100100	100100	080100	100100
OBIDOS	February 1993	443	100100	090090	100100	100100	100100	050100
OBIDOS	February 1993	444	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	445	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	446	100100	100100	100100	100100	080100	050100
OBIDOS	February 1993	447	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	448	100100	100100	100100	100100	100100	100100
OBIDOS	February 1993	449	100100	100100	100100	100100	100100	050100
OBIDOS	February 1993	450	100100	090100	100100	100100	100100	050100
MILFONTES	February 1993	451	100100	090100	100100	100100	100100	100100
MILFONTES	February 1993	452	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	453	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	454	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	455	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	456	100100	090100	100100	100100	100100	100100
MILFONTES	February 1993	457	100100	100100	100100	100150	100100	100100
MILFONTES	February 1993	458	100100	090100	100100	100100	100100	050100
MILFONTES	February 1993	459	100100	100100	100100	100150	100100	050100
MILFONTES	February 1993	460	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	461	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	462	100100	100100	100100	100150	100100	050100
MILFONTES	February 1993	463	100100	090090	100100	100150	100100	100100
MILFONTES	February 1993	464	100100	090100	100100	100100	100100	100100
MILFONTES	February 1993	465	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	466	100100	090090	100100	100100	100100	100100
MILFONTES	February 1993	467	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	468	100100	090100	100100	100100	100100	050100
MILFONTES	February 1993	469	075100	090100	100100	100100	100100	050100
MILFONTES	February 1993	470	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	471	100100	090100	100100	100100	100100	050050
MILFONTES	February 1993	472	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	473	100100	100100	100100	100100	100100	050100

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
MILFONTES	February 1993	474	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	475	100100	090100	100100	100100	100100	100100
MILFONTES	February 1993	476	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	477	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	478	100100	090100	100100	100100	100100	050100
MILFONTES	February 1993	479	045100	100100	100100	100100	100100	100100
MILFONTES	February 1993	480	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	481	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	482	100100	090100	100100	100100	100100	100100
MILFONTES	February 1993	483	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	484	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	485	100100	090100	100100	100100	100100	100100
MILFONTES	February 1993	486	100100	090100	100100	100100	100100	050100
MILFONTES	February 1993	487	100100	090100	100100	100100	100100	100100
MILFONTES	February 1993	488	100100	100100	100125	100100	100100	100100
MILFONTES	February 1993	489	100100	090100	100100	100100	100100	100100
MILFONTES	February 1993	490	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	491	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	492	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	493	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	494	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	495	100100	090100	080100	100100	100100	100100
MILFONTES	February 1993	496	100100	090100	100100	100100	100100	050100
MILFONTES	February 1993	497	100100	090100	100100	100100	100100	050100
MILFONTES	February 1993	498	100100	100100	100100	100150	100100	050100
MILFONTES	February 1993	499	100100	090100	100100	100150	100100	050050
MILFONTES	February 1993	500	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	501	100100	100100	100100	100100	100100	050050
MILFONTES	February 1993	502	100100	090100	100100	100100	100100	050100
MILFONTES	February 1993	503	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	504	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	505	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	506	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	507	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	508	100100	090100	100100	100100	100100	050100
MILFONTES	February 1993	509	100100	090100	100100	100100	100100	100100
MILFONTES	February 1993	510	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	511	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	512	100100	090100	100100	100100	100100	100100
MILFONTES	February 1993	513	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	514	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	515	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	516	100100	090090	100100	100100	100100	100100
MILFONTES	February 1993	517	100100	100100	100100	100100	100100	050100
MILFONTES	February 1993	518	100100	090100	100100	100100	100100	050100
MILFONTES	February 1993	519	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	520	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	521	100100	100100	100100	100150	100100	100100
MILFONTES	February 1993	522	100100	100100	100100	100100	100100	100100
MILFONTES	February 1993	523	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	600	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	601	100100	100100	100100	100100	100100	100100

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
MILFONTES	February 1994	602	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	603	100100	090100	080100	100100	100100	050100
MILFONTES	February 1994	604	100100	100100	100100	100150	100100	100100
MILFONTES	February 1994	605	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	606	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	607	100100	100100	100100	100150	100100	100100
MILFONTES	February 1994	608	100100	000000	100100	100150	100100	100100
MILFONTES	February 1994	609	100100	100100	100100	100150	100100	050100
MILFONTES	February 1994	610	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	611	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	612	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	613	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	614	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	615	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	616	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	617	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	618	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	619	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	620	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	621	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	622	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	623	100100	090100	100100	100100	080100	100100
MILFONTES	February 1994	624	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	625	100100	100100	100100	100100	080100	100100
MILFONTES	February 1994	626	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	627	100100	090100	100100	100150	100100	050100
MILFONTES	February 1994	628	045100	100100	100100	100100	100100	100100
MILFONTES	February 1994	629	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	630	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	631	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	632	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	633	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	634	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	635	100100	100100	080100	100100	100100	100100
MILFONTES	February 1994	636	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	637	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	638	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	639	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	640	000000	000000	100100	100100	100100	000000
MILFONTES	February 1994	641	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	642	100100	000000	100100	100100	100100	100100
MILFONTES	February 1994	643	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	644	000000	000000	100100	100100	100100	000000
MILFONTES	February 1994	645	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	646	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	647	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	648	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	649	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	650	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	651	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	652	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	653	100100	090090	100125	100100	100100	100100

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
MILFONTES	February 1994	654	100100	090090	100100	100100	100100	100100
MILFONTES	February 1994	655	100100	090090	100100	100100	100100	100100
MILFONTES	February 1994	656	100100	100100	100100	000000	100100	050100
MILFONTES	February 1994	657	100100	100100	100100	000000	100100	100100
MILFONTES	February 1994	658	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	659	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	660	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	661	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	662	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	663	100100	090090	100100	100100	100100	100100
MILFONTES	February 1994	664	100100	100100	100100	100150	100100	100100
MILFONTES	February 1994	665	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	666	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	667	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	668	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	669	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	670	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	671	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	672	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	673	100100	090100	080080	100100	100100	050100
MILFONTES	February 1994	674	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	675	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	676	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	677	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	678	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	679	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	680	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	681	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	682	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	683	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	684	100100	090090	100100	100100	100100	100100
MILFONTES	February 1994	685	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	686	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	687	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	688	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	689	100100	090090	100100	100100	100100	100100
MILFONTES	February 1994	690	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	691	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	692	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	693	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	694	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	695	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	696	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	697	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	698	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	699	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	700	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	701	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	702	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	703	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	704	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	705	100100	100100	100100	100100	100100	050100

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
MILFONTES	February 1994	706	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	707	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	708	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	709	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	710	100100	090090	100100	100100	100100	100100
MILFONTES	February 1994	711	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	712	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	713	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	714	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	715	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	716	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	717	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	718	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	719	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	720	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	721	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	722	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	723	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	724	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	725	100100	100100	100100	100100	100100	050050
MILFONTES	February 1994	726	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	727	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	728	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	729	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	730	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	731	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	732	100100	000000	100100	100100	100100	100100
MILFONTES	February 1994	733	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	734	100100	100100	100100	100100	100100	100100
MILFONTES	February 1994	735	100100	100100	100100	100100	100100	050100
MILFONTES	February 1994	736	100100	090100	100100	100100	100100	100100
MILFONTES	February 1994	737	045100	090100	100100	100100	100100	050100
MILFONTES	February 1994	738	100100	090100	100100	100100	100100	050100
MILFONTES	February 1994	739	100100	100100	100100	100100	100100	100100
FARO	February 1992	11	000000	100100	100100	100100	100100	050100
FARO	February 1992	12	000000	100100	100100	100100	100100	100100
FARO	February 1992	13	000000	100100	100100	100100	100100	100100
FARO	February 1992	14	000000	100100	100100	100100	100100	050100
FARO	February 1992	15	000000	100100	100100	100100	100100	100100
FARO	February 1992	16	000000	100100	100100	100100	100100	100100
FARO	February 1992	17	000000	100100	100100	100100	100100	100100
FARO	February 1992	18	000000	100100	100100	100100	080100	050100
FARO	February 1992	19	000000	100100	100100	100100	100100	100100
FARO	February 1992	20	000000	100100	100100	100100	100100	050100
FARO	February 1992	21	000000	100100	100100	100100	100100	050100
FARO	February 1992	22	000000	100100	100100	100100	100100	100100
FARO	February 1992	23	000000	100100	100100	100100	100100	100100
FARO	February 1992	24	000000	100100	100100	100100	100100	100100
FARO	February 1992	25	000000	100100	100100	100100	100100	100100
FARO	February 1992	26	000000	100100	000000	100100	100100	100100
FARO	February 1992	27	000000	100100	000000	100100	100100	050050

Sample	Date	Ref	*ADA**	*AAT-3**	*G3PDH-2**	*GPI-1**	*GPI-2**	*SOD**
FARO	February 1992	28	000000	100100	000000	100100	100100	100100
FARO	February 1992	29	000000	100100	000000	100100	100100	050100
FARO	February 1992	30	000000	100100	100100	000000	000000	100100
FARO	February 1992	31	000000	100100	100100	100100	100100	100100
FARO	February 1992	32	000000	100100	100100	100100	100100	100100
FARO	February 1992	33	000000	100100	100100	100100	100100	050100
FARO	February 1992	34	000000	100100	100100	000000	100100	100100
FARO	February 1992	35	000000	100100	100100	100100	100100	100100
FARO	February 1992	36	000000	100100	100100	100100	100100	100100
FARO	February 1992	37	000000	100100	100100	100100	100100	100100
FARO	February 1992	38	000000	100100	100100	100100	100100	050100
FARO	February 1992	39	000000	100100	100100	100150	100100	100100
FARO	February 1992	40	000000	100100	100100	100100	100100	050100
FARO	February 1992	41	000000	100100	000000	100100	100100	100100
FARO	February 1992	42	000000	100100	000000	100150	100100	100100
FARO	February 1992	43	000000	100100	000000	100100	100100	100100
FARO	February 1992	44	000000	100100	000000	100100	100100	100100
FARO	February 1992	45	000000	100100	000000	100100	100100	050050
FARO	February 1992	46	000000	100100	000000	100100	100100	050100
FARO	February 1992	47	000000	100100	000000	100100	100100	100100
FARO	February 1992	48	000000	100100	000000	100100	100100	100100
FARO	February 1992	49	000000	100100	000000	100100	100100	100100
FARO	February 1992	50	000000	100100	000000	100100	100100	100100
FARO	February 1992	51	000000	100100	100100	100150	100100	100100
FARO	February 1992	52	000000	100100	100100	100100	100100	050100
FARO	February 1992	53	000000	100100	100100	100100	100100	100100
FARO	February 1992	54	100100	100100	100100	100100	100100	050100
FARO	February 1992	55	100100	100100	100100	100100	100100	050100
FARO	February 1992	56	100100	090100	100100	100100	100100	050100
FARO	February 1992	57	100100	100100	100100	100100	100100	100100
FARO	February 1992	58	100100	100100	100100	100100	100100	100100
FARO	February 1992	59	100100	090100	100100	100100	100100	050050
FARO	February 1992	60	100100	100100	100100	100100	100100	050100
FARO	February 1992	61	100100	100100	100100	100100	100100	100100
FARO	February 1992	62	100100	100100	080100	100100	100100	100100
FARO	February 1992	63	100100	100100	100100	100100	100100	100100
FARO	February 1992	64	100100	100100	100100	100100	100100	100100
FARO	February 1992	65	100100	090100	100100	100100	100100	100100
FARO	February 1992	66	100100	090100	100100	100100	100100	100100
FARO	February 1992	67	100100	100100	100100	100100	100100	100100
FARO	February 1992	68	100100	090100	100100	100100	100100	100100
FARO	February 1992	69	100100	090100	100100	100100	100100	050100
FARO	February 1992	70	100100	100100	100100	100100	100100	100100
FARO	February 1992	71	100100	090100	100100	100100	100100	100100
FARO	February 1992	72	100100	090100	100100	100100	100100	050100
FARO	February 1992	73	100100	100100	100100	100100	100100	100100
FARO	February 1992	74	100100	090090	100100	100100	100100	100100
FARO	February 1992	75	100100	100100	100100	100100	100100	100100

ANNEX 3

PROTOCOLS AND RECIPES FOR

MICROSATELLITE GENOMIC LIBRARY CONSTRUCTION, CLONE SEQUENCING AND AMPLIFICATION AND DETECTION OF POLYMORPHISMS

MICROSATELLITE LIBRARY PREPARATION

1.1. SOURCE OF GENOMIC DNA

1. To obtain DNA from "difficult" samples, that is samples that there is a suspicion that the DNA is degraded, the tissue may be previously homogeneised with a mortar and pestle with liquid nitrogen.
2. Place the sample in 465 µL of the lysis buffer (100/100/250 TEN) in a 1.5 mL Eppendorf tube.
3. Add 20µL of 20%SDS and 20 µL of RNAse (from a 10 mg/mL stock).
4. Homogenize the sample with a micropestle.
5. Add 20 µL of Proteinase K (from a 10 mg/mL stock) and homogenise again.
6. Leave 1 hour at 55°C, or overnight at 37°C.
7. Add 500 µL of Phenol:Chloroform:IAA (1:1:1/24).
8. Mix well and spin 10 min for 14 000 rpm.
9. Transfer aqueous phase (upper) to another 1.5 mL Eppendorf tube, assuring that no phenol is also transferred.
10. Add 500 µL of Choloroform:IAA (24:1), mix well.
11. Spin for 5 minutes at 14 000 rpm.
12. Transfer the supernatant to a fresh Eppendorf tube.
13. Add 1/5 volumes of 5M NaCl.
14. Add 2 volumes of 99% cold EtOH and invert a few times.
15. Allow precipitation to occur within 30 min at –20°C.
16. Spin for 10 minutes at 14 000 rpm.
17. Discard EtOH
18. Wash the pellet twice with 70% cold EtOH inverting the tube contents a few times.
19. Spin for 5 minutes at 14 000 rpm.
20. Discard EtOH.
21. Dry in a speed-vacuum.
22. Suspend pellet in 50 µL of TE, leave overnight at 37°C for complete ressuspension.

1.2. RESTRICTION

Material	Chemicals	Solutions
Gel rig with preparative comb	Genomic DNA (50μg)	TAE
Corex tubes (autoclaved)(~ 8 tubes per size selection)	Restriction Enzyme	Ethidium Bromide (0.5 μg/mL)
Razor blade	OPA	Chloroform:isoamyl alcohol (24:1)
	Low Melting Point Agarose (LMP)	Ethanol (absolute and 70%)
	100 bp ladder	3M NaAcetate
	Phenol	TE
	Chloroform	
	Isoamyl alcohol	
	NaAcetate	
	Ethanol	
	Ethidium bromide	
	Tris	
	EDTA	

Procedure:

1. Digest 50μg DNA overnight. The final volume of the digest should be about 250-300μL.
2. Add 60 μL 6X Loading dye.
3. Prepare a 1% LMP agarose gel using the preparative comb (one small well and one really big well).
4. Run all of the digest:

   Electrophorese:
   Buffer: 1 x TAE (20 mL 50xTAE + 980 mL ddH_2O)
   Gel: 1%
   Sample: 250-300 μl digested sample + 60 μl 6xloading dye
   Ladder: 1 μl 100 bp ladder + 10 μl 1xTAE + 3 μl 6xloading dye
   Run at 40 V, for 6 to 7 hours

5. After it has run cut the ladder lane and some of the digest lane and stain in Ethidium Bromide and visualize with UV.
6. Take a picture of the gel piece with a ruler along side for reference when cutting the size selection out.

7. Align the ruler with the rest of the gel and cut out the appropriate section of gel corresponding with the size selection 300-600bp.
8. Cut the pieces of the gel in smaller chunks and place in sterile Corex tubes.
9. Place the tube at -80°C for 2 hours.
10. Add 3 volumes of saturated phenol (3 times the volume of the gel fragment, 1g = 1mL).
11. Vortex the tube until the agarose is solubilized.
12. Balance the tubes and spin for 20 minutes at 14 000 rpm.
13. Remove the aqueous phase and place in a new Corex tube. Repeat the phenol extraction if the interphase is wide or bits of the interphase are in the your aqueous layer.
14. Add an equal volume of chloroform.
15. Vortex and spin 3-5 minutes at 14 000 rpm.
16. Remove the aqueous phase and place in a new Corex tube.
17. Add Na Acetate to a final concentration of 0.3M.
18. Add 3 volumes of cold absolute ethanol.
19. Place at-80°C for 30 minutes.
20. Spin for 30 minutes at 14 000 rpm.
21. Remove supernatant. Wash the pellet with 70% EtOH.
22. Air dry.
23. Resuspend the pellet in 500µL TE and transfer to a 2mL screwcap tube.
24. Add 3M Sodium Acetate to a final concentration of 0.3M.
25. Add 3 volumes of absolute EtOH.
26. Place at -80°C for 30 minutes.
27. Spin for 15 minutes at 14 000 rpm.
28. Remove supernatant and wash pellet with 70% EtOH.
29. Dry pellet.
30. Resuspend in 21µL TE.
31. Run 1µL of 1% agarose to quantify the recovered DNA.

1.3. LIGATION

Material	Chemicals	Solutions
Eppendorf 0.5 mL	pUC18 BamHI/BAP	TAE
Water bath 65°C	T4DNA Ligase	
Gel rig	OPA	
2 teeth comb	ATP	
	DNA	
	Agarose	
	Loading buffer	
	100 bp ladder	
	Ethidium bromide	

Procedure:

1. Dilute pUC18 to 20 ng/µl;
2. Make 10mM stock ATP (**0.5µl ATP+4.5 µl H_2O**)
3. Setup 3 ligation reactions using 3 ratios of vector:insert for the size range of 300-600 bp:

Ratio	1:0.5	1:1	1:2
DNA (µl)	0.50	1.00	2.00
dilute vector (µl)	1.00	1.00	1.00
OPA (µl)	2.50	2.50	2.50
10 mM ATP (µl)	1.25	1.25	1.25
T4 DNA Ligase (µl)	1.50	1.50	1.50
water (µl)	18.25	17.75	16.75
Total volume (µl)	25.00	25.00	25.00

4. Incubate at RT, O/N;
5. Stop reaction by heating at 65°C for 20 minutes;
6. Run 5 µl of ligation reaction (+1 µl loading buffer) on 1% agarose;
7. Stain in EtBr and visualize - look for smear above pUC18 band;
8. Store ligation reactions at -20°C until transformation.

1.4. TRANSFORMATION

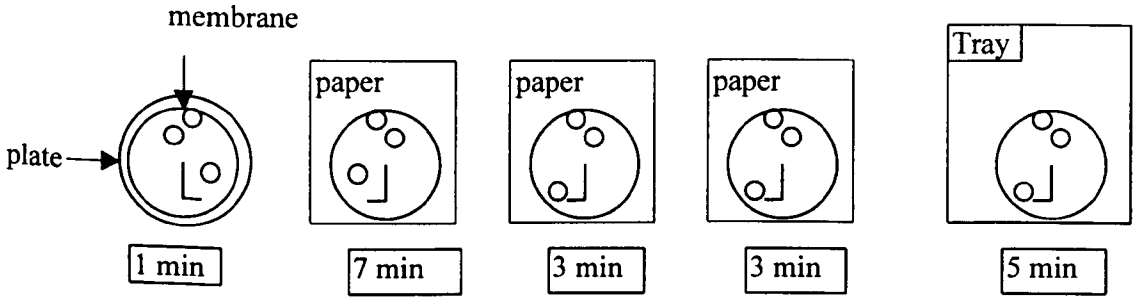
Material	Chemicals	Solutions
Polypropylene tubes (Falcon 2059)	Tryptone	LB liquid media
Plates	Yeast	S.O.C Media
Bunsen	NaCl	2% X-gal solution
Plates rotor	Agar	ampicillin [100mg/mL]
Nalgene disposable filterware 115ml	Ampicillin	1M NaCl
42°C waterbath	Competent cells	1M KCl
37°C shaker	KCl	2M Mg^{++} (filter ster.)
LB Plates	Glucose	1M Glucose (auto, 4°C)
Autoclaved botle (for S.O.C.)	Dimethylformamide	Glycerol stock
	x-gal	2% x-gal
	Glycerol	

1. Chill the polypropylene tubes and heat water bath to 42°C.
2. Thaw competent cells on ice (each tube has 200 µl of cells).
3. Aliquot 40µL competent cells per tube.
4. Refreeze any unused cells at -70°C.
5. As a control (and to determine transformation efficiency) add 2 µL control DNA (puc19). Move pipette through cells while dispensing. Gently tap tube to mix.
6. For DNA from ligation reactions add 1 µL to cells. Also do a second transformation using 2µL of ligation mix. Move pipette through cells while dispensing. Gently tap tubes to mix.
7. Incubate cells on ice for 30 minutes.
8. Heat-shock the cells for 45 seconds in a 42°C water bath; DO NOT SHAKE.
9. Place on ice for 2 minutes.
10. Add 0.9mL of room temperature S.O.C. Medium.
11. Shake at 225 rpm (37°C) for 1 hour.
12. While the transformations are shaking add 50 µL 2% X-gal to each of the plates required
13. Plate 200µL and 100µL of the undiluted transformation mix and 100µL of the diluted transformation mix on plates.
14. Incubate at 37°C overnight.

1.5. BLOTTING

Material	Chemicals	Solutions
Amersham Hybond N nylon membrane, 0.45	NaCl	denaturing solution
µ, 82 mm, 50/pk, cat:RPN 82N		
numbered circular nylon filters	NaOH	neutralizing solution
3 strips of Whatman filter paper	Tris-HCl	2X SSC
3 strips of Saran wrap	EDTA	
1 tray	Na_3citrate	
80°C oven		

Protocol:



1. Place the three pieces of Whatman paper on the three pieces of Saran wrap. Soak the first with denaturing solution, and the second two with neutralizing solution. Fill the tray with 2XSSC.
2. Carefully place membrane on agar surface. Mark the membrane and agar using a sterile needle.
3. Remove the membrane after 1 minute and place, colony side up, on the Whatman paper with denaturing solution. Leave for 7 minutes.
4. Place the membrane, colony side up, on the first Whatman paper soaked in neutralizing solution. Leave for 3 minutes then repeat with the second neutralizing solution soaked filter paper.
5. Place filter in 2X SSC for 5 minutes.
6. Transfer, colony side up, to a sheet of Whatman paper and allow to air dry.
7. Bake the filters at 80°C for 2 hours.

1.6. HYBRIDIZATION

Material	Chemicals	Solutions
colony blots	microsatellite probe	20XSSPE
Hybaid oven (at 62°C)	[γ-^{32}P]ATP	50X Denhardt's
Hybridization bottles	OPA	20XSDS
Corex tubes	T4-PNK	20XSSC
Twisers	20 mg/mL RNA	solution **A** (2xSSC/0.2%SDS)
	probe (GT_{15})	solution **B**
		(5xSSPE/2x Denhardt's /0.5%SDS/
		100µg/mL RNA)
		solution **C** (0.2xSSC/0.2%SDS)

Procedure:

1. Label primer (this recipe gives enough for **1** bottle so multiply this recipe as required - one can get 8-10 filters in a bottle):

1µg (GT)15 ≈ 3µL (stock 0.346 µg/µl)
1µL OPA buffer
2µL ddH_2O
30 µCi [γ-^{32}P]ATP ≈ 3µL
0.5µL T4-PNK

2. Incubate at 37°C for 30 min- 18 hours. Heat at 65°C for 15 minutes to kill the kinase.
3. Scrub filters with your finger tips in 50 mL of solution **A.**
4. Prehybridize filters in 50 mL of solution **B** (prehybridization) for 2 hours at 62°C in the hybridization bottles.
5. Discard prehybridization solution.
6. Add 15mL of pre-hybridization solution **B** to each bottle. Add 10µL of labeled probe to each bottle (so the pre-hybridization becomes hybridization solution) .
7. Hybridize at 62°C overnight.
8. Discard hybridization solution. Do two 15 min washes with solution **A** at room temperature.
9. Check the filters with the counter for the intensity of the signal.
10. If the signal is strong enough do a 15 minute wash at 62°C with solution **C**.
11. Check the signal again, if the signal is still very strong and appears to be evenly distributed you can perform another 15 minute wash at 62°C with solution **C**.

12. Place the filters on Saran wrap and cover with another piece of Saran wrap.

13. Place filters in a cassette with film and an intensifying screen. Leave overnight at -80°C.

14. Expose the autorad.

15. Align the filters with the autorad and mark the location of the needle holes on the autorad to be able to align the highlighted colonies with the original plate.

16. Align the original plate with the appropriate position on the autorad. Align the marks on the autorad with the needle holes on the plate.

17. Where a colony matches with a highlight on the autorad pick the colony with a sterile toothpick streak on tubes with LB.

18. Grow the colonies up overnight.

19. Make glycerol stocks and do plasmid preps and proceed to sequencing.

Washes:

1. Solution **A** - 15 min, RT;
2. Solution **C** - 15 min, RT;
3. Solution **C** - 15 min, 62°C.

Hot primer:

1. In a labelled 0.5 mL Eppendorf tube put 1 µL OPA buffer, 2 µL ddH_2O and 0.5 µL PNK.
2. Take 1µL of primer stock ((GT_{15}) = 0.346 µg/µl) and dilute up to 100µl;
3. Add 3 µL of diluted primer to the Eppendorf tube.
4. Go to the radiation station and to the primer tube, add 3 µL of [γ-^{32}P]ATP.

   The primer is labelled and ready to use.

1.7. ALKALINE LYSIS FOR MINI-PREP OF PLASMID DNA

Material	Chemicals	Solutions
Polypropylene tubes (Falcon 2059)	Tryptone	LB liquid media
Nalgene disposable filterware 115ml	Yeast	ampicillin
37°C shaker (225 rpm)	NaCl	1M Glucose
Pipette and tips	Ampicillin	1 M Tris.Cl pH 8.0
Centrifuge	Glucose	0.5M EDTA pH 8.0
Eppendorf tubes	Tris	1M NaOH
	EDTA	20% SDS
	NaOH	5 M Potassium acetate
	SDS	5 M NaCl
	Potassium acetate	Isopropanol
	Glacial acetic acid	70% ethanol
	NaCl	100% ethanol
	Isopropanol	
	Ethanol	
	Glycerol	

1. Transfer a single bacterial colony into 5 mL of LB + Amp medium in a loosely capped 15 mL tube (culture tubes).

   Incubate the culture overnight at 37°C with 225 rpm.
2. Prepare 3 sets of tubes (glycerol stocks and pelleting).
3. Take 500µL of each culture and mix with 500µL of 80% glycerol. **Vortex strongly** and freeze at -80°C.
4. Fill all the tubes with 1.5 mL of culture.

   Centrifuge at maximum speed for 30 sec.

   Discard the upper layer.

   Pour another 1.5 mL over the pellet and repeat centrifugation.

   Discard the upper layer.

   Pour another 1.5 mL over the pellet and repeat centrifugation.

   Alternatively use the universal tubes, centrifuge at maximum speed for 5 min.
5. Remove medium with a transfer pipette (bulb fin pipette), leaving the pellet as dry as possible.
6. Ressuspend in 200 µL of ice-cold solution I (see annex 1-15) by vigourous vortexing.

7. Add 400 µL of freshly prepared solution II (see annex 1-15):

	25 mL	2.5 mL
1M NaOH	5 mlL	500 µL
20% SDS	1.3 mL	130 µL
water	19.7 mL	1.9 mL

8. Close the tube and mix by inverting the tube 5 times (fast).
9. Add 300 µL of ice-cold solution III (see annex 1-15).
10. Close tube, vortex gently for 10 sec.
11. Store on ice for 3-5 min.
12. Centrifuge at maximum speed for 5 minutes.
13. Transfer supernatant to a fresh tube with a transfer pipette.

    Keep tubes on ice.

    At this point you can freeze (-20°C) the tubes and do the extraction the next day.

14. Add 1/50 vol. of 5M NaCl (1mL - 20 µL)
15. Add 2 vol. of ice-cold isopropanol and mix by vortexing.

    Allow the mixture to stand for 2 minutes.
16. Centrifuge at 12.000rpm for 5 minutes.
17. Remove the supernatant by aspiration. Stand the tube in an inverted position on a paper towel to allow the fluid to drain. Care must be taken because the pellet is not strongly hold in the botton of the tube.
18. Rinse the pellet of double-stranded DNA with 1 mL of 70% ethanol at 4°C. Remove the supernatant as described in 17, and allow the pellet of nucleic acid to dry in the air for 10 min.
19. Redissolve the nucleic acids in 50 µL of TE (pH 8.0). Vortex briefly. Store at -20°C.

1.8. SEQUENCING

Material	Chemicals	Solutions
Water baths (65°C and 37°C)	NaCl	5M NaCl
Timer	Potassium acetate	3M Na Acetate
Ice	Ethanol	70% ethanol
Pipette and tips	Isotope	100% ethanol
Centrifuge	T7 Sequencing kit	
Eppendorf tubes (0.5 mL)		
Microwell plates		
250 mL cylinder		
1 l cylinder		
2 l beaker		

Procedure:

As manufacturer's instructions.

Gel preparation:

100 mL 6% w/v acrylamide, 0.3% w/v bis-acrylamide, 7M Urea; 1 X TBE
330 µL 10% AP (see annex 3-21)
40 µL TEMED

Running Conditions

Electrode buffer: 10XTBE 100 mL
H_2O up to 1000 mL

1. Pre-heat the electrode buffer, 5 min in microwave or pre-run gel for 1 hour. Fixed mA = 50-60 mA; 1200-1300 V.
2. Load buffer.
3. Put comb in place.
4. Denature samples (only first set) - 2 min, >80°C.
5. Load 1.5 µL of each sample. Load with T to serve as a marker of the gel orientation), then ACGT.
6. Turn power on, denature second set of samples.
7. The gel will run until the blue dye reaches the end of the plate (3 hours).

Fixing gels

1. Take the plates out of the gel rig and put them in a container (box dimension: 38x42x15cm), with fixing solution. Take the top plate out, carefully.
2. Fix the gel for 1 hour in 1L.
3. Take the gel and plate and place them on the bench. Place a sheet of Whatman No.3 paper over the gel. Cut to size. Lift carefully, so that the gel clings to the paper without breaking.
4. Wrap in cling film.

5. Dry gel for 2-2.5 hours, at 80°C, with paper facing downwards.

Exposing

Expose film to gel overnight (film type: Kodak X-OMAT AR - grey film - cat. 165.1512, 35x43cm; blue film can also be used, expose longer 1 to 2 days).

1.9. RECIPES

Phenol equilibration

Phenol must be equilibrated to a pH of > 7.8 because DNA will partition into the organic phase at acid pH.

Melt phenol at 68°C.

Add hydroxyquinone to a final concentration of 0.1% (This compound is an antioxidant, a partial inhibitor of RNase and a weak chelator of metal ions).

To the melted phenol add an equal volume of buffer 0.1M Tris (pH8) - stir on magnetic stirrer for 15 minutes.

Remove upper aqueous phase.

Add an equal volume of 0.1M Tris (pH 8) to the phenol.

Stir on magnetic stirrer for 15 minutes.

Remove upper aqueous phase with a glass pipette attached to a vacuum line equipped with traps.

Repeat extraction until the pH of the phenolic phae is >7.8 (pH paper).

After final aqueous phase has been removed add 0.1 volume of 0.1 M Tris (pH 8) containing 0.2% ß-mercaptoethanol (200 µL).

(storage up to 1 month)

Phenol:Chloroform:IAA

5 mL of Phenol ; 5 mL of 24:1 Chloroform:IAA; 10 mL of TE.

Use lower phase.

Chloroform:IAA (24:1)

96 mL Chloroform; 4 mL of IAA.

10 X TBE

	10 X	5 X
Tris Base	121 g	60.5 g
EDTA	7.4 g	3.7 g
Boric Acid	53.4 g	26.7 g

to a final volume of 1 L pH 8.3

50X TAE

242g Tris; 57.1 mL glacial acetic acid; 100 mL 0.5 M EDTA (pH 8).

Adjust volume to 1L.

TE buffer (pH8)

Composition: 10mM Tris Cl (pH8); 1mM EDTA (pH8)

Preparation;

5mL 1M Tris (pH8)

1mL 0.5M EDTA (pH8)

Adjust volume to 500 mL

2M NaOH

8g NaOH dissolved in 100 mL ddH_2O; 16g NaOH dissolved in 200 mL ddH_2O.

5M NaCl

Dissolve 292.2 g of Na Cl in 800 mL H_2O.

Adjust the volume to 1L with H_2O.

0.5M EDTA (pH 8)

Add 186.1 g disodium EDTA · $2H_2O$ to 800 mL of H_2O.

Stir vigorously on a magnetic stirrer.

Adjust the pH to 8 with NaOH (approx 20g of NaOH pellets).

The disodium salt of EDTA will not go into solution until the pH of the solution is adjusted to approximately 8 by the addition of NaOH.

3M Sodium Acetate (NaAc) pH 5.2

Dissolve 408.1 g of sodium acetate·$3H_2O$ in 800 mL of H_2O.

Adjust the pH to 5.2 with glacial acetic acid.

Adjust the volume to 1L with H_2O.

5M Potassium acetate

49.1 g of potassium acetate.

Adjust the volume to 100 mL.

Adjust the pH with 2M acetic acid to 7.5.

1M $MgCl_2$

Dissolve 203.3g of $MgCl_3 \cdot 6H_2O$ in 800 mL of H_2O.

Adjust the volume to 1L with H_2O.

101.65 g in 500 mL H_2O.

50.8 g in 250 mL H_2O.

For LB (Luria-Berani Medium) plates (40 required)

To 960 mL of deionized H_2O add

Bacto-tryptone	10 g
Bacto Yeast Extract	5 g
NaCl	10 g

Shake until the solute have dissolved.

Adjust the pH to 7 with 5N NaOH (approx 0.2 mL).

Sterilize by autoclaving for 20 minutes at 15 lb/sq in a liquid cycle.

Add 100 mg (1mL of stock) of ampicillin.

LB liquid media

Same as plates but leave out the agar - should only need 500mL

SOC Medium

Bacto-tryptone	10g
Bacto-yeat extract	2.5 g
NaCl	0.25 g

Autoclave for 20 minutes, cool to 60°C.

Add 20 mL of a sterile 1M solution of glucose.

Alkaline lysis buffers for mini-prep of plasmid DNA

Solution I (500 mL):

50 mM glucose =	25 mL 1M glucose
25 mM Tris Cl (pH8)=	12.5 mL 1M Tris Cl
10 mM EDTA (pH 8)=	10 mL 0.5 M EDTA

Adjust volume to 500 mL

Solution II (500 mL):

0.2N NaOH (freshly diluted from a 10N stock)= 10 mL

1% SDS= 25 mL 20% SDS

Adjust the volume to 500 mL

Solution III (100 mL)

5M potassium acetate	60 mL
glacial acetic acid	11.5 mL
H_2O	28.5 mL

The resulting solution is 3M with respect to potassium and 5M with respect to acetate

1M Glucose

18g glucose disolved in 90 mL deionized H_2O.

Adjust the volume to 100 mL.

Sterilize by filtration through a 0.22 μm filter.

6X loading buffer (pg 6.12 Maniatis)

0.25% bromophenol blue

0.25% xylene cyanol TE

15% Ficoll (Type 400 Pharmacia) in H_2O

Denaturation buffer

Composition: 1.5M NaCl; 0.5M NaOH

Preparation:

87.66 g NaCl

20 g NaOH

Adjust volume to 1L with ddH_2O

Hybridization buffer

Composition: 1mM EDTA (pH8); 0.263M Na_2HPO_4; 1% BSA (Bovine Serum Albumin); 7% SDS.

Preparation:

2 mL 0.5M EDTA (pH 8).

526.3 mL 0.5M Na_2HPO_4.

10 g BSA.

70 g SDS (or 350 mL 20% SDS).

Adjust volume to 1L.

0.5M Na_2HPO_4 (pH7.2)

70.98 g Na_2HPO_4.

3 mL 85% H_3PO_4 (85%).

Adjust volume to 1L.

Neutralizing solution

Composition: 1.5m NaCl; 0.5M Tris (pH7.2); 0.001M Na_2EDTA.

Preparation:

87.66 g NaCl.

500 mL 1M Tris (pH 7.2).

2mL 0.5M EDTA (pH8).

Adjust volume to 1L.

Extraction Buffer

Composition: 0.1M Tris (pH8); 0.1M EDTA (pH 8); 0.25M NaCl; 1% SDS

Preparation:

100 mL 1M Tris

200 mL 0.5M EDTA

50 mL 5M NaCl

50 mL 20% SDS

Adjust volume to 1L

Lysis Buffer (Phenol Extraction)

100 mM NaCl

500 mM Tris pH8

20 mM EDTA (0.5M pH8)

25 mL 20% SDS

+ 420 mL ddH_2O

1M Tris

Dissolve 121.1g of Tris base in 800 mL of H_2O.

Adjust the pH to the desired value by adding concentrated HCl:

pH	HCl
7.4	70 mL
7.6	60 mL
8.0	42 mL

Adjust the volume of the solution to 1L with H_2O.

S.O.C. Media

To 97mL distilled water add: 2g tryptone equivalent
0.5g yeast extract
1 mL 1M NaCl
0.25 mL 1M KCl

Stir to dissolve.

Autoclave and cool to RT.

Add 1mL 2M Mg^{++} stock and 2mL 1M glucose.

Rinse 0.2μm Nalgene disposable filter with water.

Filter media through 0.2μm filter into autoclaved bottle.

2% X-gal (total volume 1 mL)

Dissolve 20 mg X-gal in 1 mL Dimethylformamide.

Glycerol (total volume 100 mL)

Dissolve 80 mL glycerol in 20 mL water.

Autoclave.

To keep the transformation mix

Add 900 μL of LB+Amp.

Leave in the shaker 37°C, overnight.

Take 500μLculture add 500μL of 80% glycerol.

Freeze at -80°C.

Ampicillin (total volume 10 mL, aliquoted in 10 Eppendorf tubes)(200mg/mL)

2g ampicillin, 10 mL water - for LB broth use 0.5 mL (100 mg)/L

Blotting Solutions

Denaturing solution	**Neutralizing solution**	**20XSSC**
1.5M NaCl	1.5M NaCl	3M NaCl
0.5M NaOH	0.5M Tris-HCl pH7.2	0.3M Na_3citrate
	0.001M EDTA	

Hybridization solutions

Solution **A** (2X SSC / 0.2% SDS)

50 mL 20X SSC
5 mL 20% SDS
water up to **500ml**

Solution **B** (5X SSPE / 2X Den / 0.5%SDS / 100μg/mL RNA)

50 mL 20X SSPE
5 mL 20% SDS
8 mL 50X Denhardt
water up to **200 mL**

Solution **C** (0.2xSSC/0.2%SDS)

5 mL 20X SSC
5 mL 20% SDS
water up to **500 mL**

20X SSC

Dissolve 175.3g of NaCl and 88.2 g of sodium citrate in 800 mL of H_2O

Adjust the pH to 7.0 with a few drops of the 10N solution of NaOH

Adjust the volume to 1L with H_2O

350.6 g NaCl, 176.4 g Sodium citrate to 2L

20% SDS

Dissolve 200 g of electrophoresis grade SDS in 900 mL H_2O

Heat to 68°C to assist dissolution

Adjust pH to 7.2 by adding a few drops of concentrated HCl

Adjust volume to 1L with H_2O

50X Denhardt

5g Ficoll

5g polyvinylpyrrolidone

5g Bovine Serum Albumin

Dissolve in 300mL H_2O.

Once dissolved, bring up to 500 mL total with H_2O.

Filter and dispense into 25 mL aliquots.

Store at –20°C.

20X SSPE

174g NaCl.

88.2g sodium citrate.

Dissolve in 800mL H_2O.

Adjust pH to 7.4 with NaOH (6.5mL of a 10N solution).

Adjust volume to 1 liter.

Dispense into aliquots.

Sterilize by autoclaving.

Acrylamide gels (30 x 40 cm)

A commercial ready-to-use 6% w/v acrylamide/Bis 0.3% acrylamide, 7 M Urea, 1x TBE solution was used (Scotlab SL-9238).

Before pouring add to 120 mL of ready-to-use acrylamide solution:

10% AP	330 µl
TEMED	48 µl

Pour gel.

Gel will be ready in 40 min to 1 hour.

AP can be kept in stock for a week at RT or fridge.

10% Ammonium persulfate (AP)

Ammonium persulfate 1g

H_2O to 10 mL

The solution may be stored in aliquots of 1 mL for several weeks.

Fixing gels solution

10% Acetic acid
10% Methanol
80% water

Staining agarose gels with ethidium bromide

The gel is immersed in electrophoresis buffer or H_2O containing EtBr (0.5µg/mL, of a stock)

200 mL ,10 µL dye

400 mL, 20 µL dye

1L, 50 µL dye

For 30 to 45 minutes at room temperature.

ANNEX 4

GENOTYPES OF MICROSATELLITE LOCI

Locality	Date	Ref.	Dla11	L3	L8	L9	Dla6
Foz	November 1993	107	000000	000000	200222	000000	067075
Foz	November 1993	108	000000	000000	208226	000000	077079
Foz	November 1993	109	000000	000000	212232	000000	075075
Foz	November 1993	110	117117	133133	226226	168190	077077
Foz	November 1993	111	117121	141171	228224	174174	075075
Foz	November 1993	112	115117	153153	214214	000000	075075
Foz	November 1993	113	115119	147165	214226	000000	075081
Foz	November 1993	114	103119	143161	218236	166170	075077
Foz	November 1993	115	000000	141141	214240	000000	075077
Foz	November 1993	116	105121	151155	216216	166166	075077
Foz	November 1993	117	111115	141141	198216	156168	075075
Foz	November 1993	118	111115	157177	198216	172172	089089
Foz	November 1993	119	115117	113141	226248	166170	077077
Foz	November 1993	120	117119	149149	226236	156166	081081
Foz	November 1993	121	115117	149149	210236	166160	077097
Foz	November 1993	122	115121	131175	228236	166172	075075
Foz	November 1993	123	121121	155187	220226	000000	075087
Foz	November 1993	124	000000	000000	000000	166166	000000
Foz	November 1993	125	115117	171173	228232	160166	075075
Foz	November 1993	126	111111	175131	228194	166166	000000
Foz	November 1993	127	113115	141155	194234	166166	081097
Foz	November 1993	128	111123	141147	224230	166170	075075
Foz	November 1993	129	111121	141141	200200	168174	075079
Foz	November 1993	131	115117	159141	000228	170170	075075
Foz	November 1993	132	119123	139141	234236	166170	075075
Foz	November 1993	133	117119	141141	194220	166166	073073
Foz	November 1993	134	117115	000000	000000	000000	000000
Foz	November 1993	135	117123	139151	212228	168168	075075
Foz	November 1993	136	113000	000000	192224	000000	075083
Foz	November 1993	137	119123	141159	216226	166166	067081
Foz	November 1993	138	113119	141181	212216	166166	000000
Foz	November 1993	139	119123	147171	210224	166166	075075
Foz	November 1993	140	119107	131141	194230	000000	075089
Foz	November 1993	141	119123	143143	226226	000000	075075
Foz	November 1993	142	117119	159163	234236	166166	075075
Foz	November 1993	143	113113	141141	208216	168168	075075
Foz	November 1993	144	107119	145145	218224	000000	073073
Foz	November 1993	145	107123	159159	200230	000000	067067
Foz	November 1993	146	117119	137141	194220	166172	077077
Foz	November 1993	147	119123	141151	220230	202214	075075
Foz	November 1993	148	107131	141141	220224	176176	077077
Foz	November 1993	149	117119	145151	194220	168174	069075
Foz	November 1993	150	113117	141151	210228	166166	075075
Foz	November 1993	152	113117	131151	000000	000000	069069
Foz	November 1993	154	117113	141151	208224	166172	075075
Foz	November 1993	155	109123	000000	000000	000000	075075
Foz	November 1993	156	000000	000000	000000	000000	075075
Foz	November 1993	159	000000	000000	000000	000000	077077
Foz	November 1993	160	000000	000000	000000	000000	075075
Foz	November 1993	161	117119	135141	000000	000000	000000
Foz	November 1993	162	000000	147159	210224	166172	075075
Foz	November 1993	163	000000	000000	000000	000000	069075
Foz	November 1993	164	000000	000000	000000	000000	077077
Foz	November 1993	165	000000	000000	000000	000000	079079
Foz	November 1993	166	000000	125139	000000	000000	077077

Locality	Date	Ref.	Dla11	L3	L8	L9	Dla6
Foz	November 1993	167	000000	145159	000000	000000	079093
Foz	November 1993	168	107109	153185	208220	166168	069069
Foz	November 1993	169	117119	157179	224226	160172	075079
Foz	November 1993	170	117119	000127	220230	000000	075075
Foz	November 1993	171	000000	000141	000000	000000	075075
Foz	November 1993	172	000000	000000	000000	000000	077077
Foz	November 1993	173	117123	000000	000000	000000	067077
Foz	November 1993	174	117115	141173	222232	168172	077101
Foz	November 1993	175	000000	155153	194210	168168	075081
Foz	November 1993	176	117117	141139	212214	000000	075075
Foz	November 1993	177	107131	125189	220196	000000	081103
Foz	November 1993	178	117115	147159	196210	164164	075075
Foz	November 1993	179	000000	173197	220212	000000	075075
Foz	November 1993	181	000000	141159	226236	000000	075075
Foz	November 1993	183	111113	147161	204222	000000	077083
Foz	November 1993	184	000000	141169	210224	000000	077077
Foz	November 1993	185	000000	153163	000000	000000	075075
Foz	November 1993	186	107115	159177	000000	000000	075079
Foz	November 1993	188	117123	151155	204224	000000	077077
Foz	November 1993	189	115123	000000	000000	000000	000000
Foz	November 1993	190	121123	000000	000000	000000	077077
Foz	November 1993	191	000000	000000	214226	000000	075075
Foz	November 1993	192	119119	141147	000194	000000	079077
Foz	November 1993	193	000000	000000	202216	000000	000000
Foz	November 1993	194	000000	133141	000000	000000	000000
Foz	November 1993	195	117127	000000	000000	000000	075075
Foz	November 1993	196	000000	141141	222232	000000	077089
Foz	November 1993	197	000000	000177	000000	000000	075091
Foz	November 1993	198	107119	139159	000000	000000	000000
Foz	November 1993	200	000000	147179	226234	000000	000000
Foz	November 1993	201	113115	157179	000232	000000	000000
Foz	November 1993	202	117117	161177	214226	000000	075075
Foz	November 1993	203	113121	000000	000000	000000	000000
Foz	November 1993	204	000000	173175	212228	000000	083083
Foz	November 1993	205	113117	161143	000226	000000	075079
Foz	November 1993	207	119121	133141	210226	000000	075077
Foz	November 1993	208	117123	143147	000000	000000	075075
Foz	November 1993	209	103117	141159	000226	000000	000000
Foz	November 1993	210	121135	141179	218226	000000	000000
Foz	November 1993	211	121117	141159	222232	000000	075075
Foz	November 1993	212	119117	121181	000232	000000	000000
Foz	November 1993	213	121119	141147	194222	000000	000000
Foz	November 1993	214	117125	141159	194210	000000	077077
Foz	November 1993	215	117119	000125	000230	000000	075087
Foz	November 1993	216	117123	147157	204210	000000	075075
Foz	November 1993	217	113117	149179	000000	000000	075077
Foz	November 1993	218	113109	143145	210226	000000	000000
Foz	November 1993	219	117135	000000	000000	000000	000000
Foz	November 1993	220	105119	147161	000000	166176	000000
Foz	November 1993	221	109119	141147	194222	164166	067077
Foz	November 1993	222	000000	115175	000000	164166	075089
Foz	November 1993	223	113121	151141	206230	166166	000000
Foz	November 1993	224	000000	141175	216218	166166	075075
Foz	November 1993	225	113125	171187	194216	166170	075077
Foz	November 1993	226	117117	173161	198222	168168	075075

Locality	Date	Ref.	Dla11	L3	L8	L9	Dla6
Foz	November 1993	227	109117	161151	208220	000000	075075
Aveiro	February 1993	228	000000	141141	196226	166168	067075
Aveiro	February 1993	229	119123	141127	214224	000000	077077
Aveiro	February 1993	230	113117	141141	208222	166166	077077
Aveiro	February 1993	231	107123	143157	194212	000000	067101
Aveiro	February 1993	232	000000	161155	000000	166166	067067
Aveiro	February 1993	233	117119	151165	000226	166174	075075
Aveiro	February 1993	234	000000	163169	000220	162166	067079
Aveiro	February 1993	235	121113	157141	000226	172172	075075
Aveiro	February 1993	236	000000	159131	200214	166166	077079
Aveiro	February 1993	237	000000	159159	220234	166166	000091
Aveiro	February 1993	238	115117	159177	194210	166160	077075
Aveiro	February 1993	239	000000	155163	000224	000000	000000
Aveiro	February 1993	240	000000	161161	000224	166160	000000
Aveiro	February 1993	241	115113	153157	000210	166166	075075
Aveiro	February 1993	242	125117	000000	000000	000000	000000
Aveiro	February 1993	243	117115	133171	000000	166166	075093
Aveiro	February 1993	244	119109	125131	000000	170170	075075
Aveiro	February 1993	245	117121	000000	000000	000000	000000
Aveiro	February 1993	246	109121	141161	220246	160170	067075
Aveiro	February 1993	247	117129	000000	000000	000000	000000
Aveiro	February 1993	248	117115	141155	000226	166166	067075
Aveiro	February 1993	249	115109	000163	000000	000000	000000
Aveiro	February 1993	250	117117	153149	212232	160160	075077
Aveiro	February 1993	252	125125	151125	194234	168168	075075
Aveiro	February 1993	253	117119	000177	204210	166160	077089
Aveiro	February 1993	254	117117	155159	210226	160170	073073
Aveiro	February 1993	255	000000	141153	220226	000000	000000
Aveiro	February 1993	256	117119	141157	000000	000000	000000
Aveiro	February 1993	257	117119	000163	000000	000000	000000
Aveiro	February 1993	259	141129	153165	228232	000000	000000
Aveiro	February 1993	260	113117	149145	000200	000000	075075
Aveiro	February 1993	261	113113	000000	000000	000000	000000
Aveiro	February 1993	262	113119	145141	000000	000000	000000
Aveiro	February 1993	265	113117	141141	000000	000000	000000
Aveiro	February 1993	266	101101	000000	000000	000000	083101
Aveiro	February 1993	267	099119	143165	220228	000000	000000
Aveiro	February 1993	268	117123	000000	000000	000000	075075
Aveiro	February 1993	269	121123	000000	000000	000000	000000
Aveiro	February 1993	270	119119	151147	194238	000000	077077
Aveiro	February 1993	273	000000	000000	000000	000000	000073
Aveiro	February 1993	274	117117	000000	000000	000000	000000
Aveiro	February 1993	276	117111	153153	000000	000000	000075
Aveiro	February 1993	277	117109	000000	000000	000000	000000
Aveiro	February 1993	281	119119	000000	000000	000000	000000
Aveiro	February 1993	282	123123	000000	000000	000000	000000
Aveiro	February 1993	283	000000	000000	000000	000000	000067
Aveiro	February 1993	284	117000	000000	000000	000000	000000
Aveiro	February 1993	285	107113	000000	000000	168168	000075
Aveiro	February 1993	288	109119	000000	000000	000000	000000
Aveiro	February 1993	300	107117	133153	202202	162168	081083
Aveiro	February 1993	301	117117	141147	194238	168168	075075
Aveiro	February 1993	302	113121	179175	230230	160170	000000
Aveiro	February 1993	303	000000	151181	194194	168168	075077

Locality	Date	Ref.	Dla11	L3	L8	L9	Dla6
Aveiro	February 1993	304	000000	141141	228232	160170	000000
Aveiro	February 1993	305	000000	161169	212242	000000	075075
Aveiro	February 1993	306	115121	141159	222238	162170	075075
Aveiro	February 1993	307	115117	161159	212222	160176	067067
Aveiro	February 1993	308	119129	149181	202212	160164	000000
Aveiro	February 1993	309	117115	151169	194222	168168	068076
Aveiro	February 1993	310	000000	165179	230222	162162	076078
Aveiro	February 1993	311	000000	165157	194212	000000	000079
Aveiro	February 1993	312	000000	141141	200222	166184	074076
Aveiro	February 1993	313	000000	141155	224232	166196	075077
Aveiro	February 1993	314	115123	141175	200212	144144	075095
Aveiro	February 1993	315	000000	141153	202216	166178	079079
Aveiro	February 1993	316	113115	143147	236236	172172	000000
Aveiro	February 1993	317	115123	151185	222230	000000	075075
Aveiro	February 1993	318	111123	127141	236230	166166	066074
Aveiro	February 1993	319	000000	141147	234212	166176	000000
Aveiro	February 1993	320	000000	165179	212222	166172	000000
Aveiro	February 1993	321	000000	141165	208208	000000	000000
Aveiro	February 1993	322	109129	141159	212212	166166	079079
Aveiro	February 1993	323	117117	127141	246210	158166	000075
Aveiro	February 1993	324	119123	149167	224212	170170	000000
Aveiro	February 1993	325	000000	000159	200224	166174	000075
Aveiro	February 1993	326	117125	141183	224224	166166	000075
Aveiro	February 1993	327	117123	000000	204224	166166	075089
Aveiro	February 1993	328	117123	000000	230200	000000	000000
Aveiro	February 1993	329	117107	141141	202196	166166	000000
Aveiro	February 1993	330	117107	131153	000000	162166	081083
Aveiro	February 1993	331	000000	141147	194236	166166	000000
Aveiro	February 1993	332	115123	000177	230230	158168	075075
Aveiro	February 1993	333	117119	141159	194220	166174	075101
Aveiro	February 1993	334	121125	139143	196224	166166	075075
Aveiro	February 1993	335	117119	137153	230230	166166	075089
Aveiro	February 1993	336	117119	131159	206206	166166	000000
Aveiro	February 1993	337	113119	143157	226226	000000	000000
Aveiro	February 1993	338	109109	137139	232232	000000	000000
Aveiro	February 1993	339	113123	151143	214222	174174	000000
Aveiro	February 1993	341	117117	147159	210222	166182	055093
Aveiro	February 1993	342	117117	141147	222226	166166	075079
Aveiro	February 1993	343	125123	127153	220230	166170	000000
Aveiro	February 1993	344	000000	000141	214228	166166	077081
Aveiro	February 1993	346	117121	141163	192226	166166	075081
Aveiro	March 1994	857	117117	177183	208222	000000	075075
Aveiro	March 1994	862	115119	000000	194212	166166	075075
Aveiro	March 1994	865	111115	000000	000000	162166	075077
Aveiro	March 1994	866	111115	000141	200214	172172	075077
Aveiro	March 1994	868	115117	141151	000224	160166	067077
Aveiro	March 1994	870	121121	000157	204210	170170	075075
Aveiro	March 1994	871	115117	000000	212222	160170	000000
Aveiro	March 1994	872	111111	133163	194222	166166	075075
Aveiro	March 1994	876	113115	000000	224232	160166	077075
Aveiro	March 1994	877	111123	141145	200212	160170	000000
Aveiro	March 1994	878	115117	000000	202216	162168	000000
Aveiro	March 1994	880	119123	000000	236236	160170	000000
Aveiro	March 1994	881	117119	000141	222230	162170	075075
Aveiro	March 1994	883	119123	000175	212212	160176	000000

Locality	Date	Ref.	Dla11	L3	L8	L9	Dla6
Aveiro	March 1994	885	113119	141141	212234	160164	091095
Aveiro	March 1994	886	117119	000155	212222	166184	000000
Aveiro	March 1994	887	113113	000000	208208	166174	075075
Aveiro	March 1994	889	117119	000141	200222	166176	000000
Óbidos	February 1993	347	121113	000153	208226	166168	089101
Óbidos	February 1993	348	000000	139165	192218	166176	075075
Óbidos	February 1993	349	000000	153151	194218	166166	075075
Óbidos	February 1993	350	119121	153165	222226	166250	000000
Óbidos	February 1993	351	123111	159141	202232	164166	089089
Óbidos	February 1993	352	119123	141161	210222	166176	075075
Óbidos	February 1993	353	117119	153151	210222	166200	000000
Óbidos	February 1993	354	000000	151129	208228	162206	077077
Óbidos	February 1993	356	117123	161187	206218	196218	000000
Óbidos	February 1993	357	117121	153141	208216	166210	077077
Óbidos	February 1993	358	117123	157159	206222	166212	067067
Óbidos	February 1993	359	121121	135147	226226	168180	075075
Óbidos	February 1993	360	000000	141155	218226	186186	075075
Óbidos	February 1993	361	113121	141159	224238	166166	075077
Óbidos	February 1993	362	117121	000000	000000	000000	000000
Óbidos	February 1993	363	117123	127141	222222	166254	075077
Óbidos	February 1993	364	117113	147177	202210	170170	075075
Óbidos	February 1993	365	119115	147169	194222	166172	067067
Óbidos	February 1993	366	117115	141171	210246	166166	075089
Óbidos	February 1993	367	117125	171183	000000	000000	000067
Óbidos	February 1993	368	113125	147159	212222	164170	075075
Óbidos	February 1993	369	117115	147159	216224	168206	075081
Óbidos	February 1993	370	115137	000000	000000	000000	000000
Óbidos	February 1993	371	113117	000000	000226	146168	075075
Óbidos	February 1993	372	115119	141165	000000	000000	000000
Óbidos	February 1993	373	113121	147159	000000	166254	067065
Óbidos	February 1993	374	000000	141153	000000	000000	000000
Óbidos	February 1993	375	000000	167181	000000	000000	073075
Óbidos	February 1993	376	000000	159179	000000	000000	000000
Óbidos	February 1993	377	000000	159159	000000	000000	000000
Óbidos	February 1993	378	105123	000000	210220	166196	075075
Óbidos	February 1993	379	117123	141161	000000	000000	000000
Óbidos	February 1993	380	117121	141157	222230	166210	075075
Óbidos	February 1993	381	119121	141165	000000	174194	067067
Óbidos	February 1993	382	000000	159181	226226	174186	077079
Óbidos	February 1993	383	000000	141147	220220	166194	000000
Óbidos	February 1993	384	000000	141141	222234	166204	075075
Óbidos	February 1993	385	117137	141179	000000	166166	077077
Óbidos	February 1993	386	000000	161165	210230	162172	079079
Óbidos	February 1993	387	117125	141167	210238	176172	075075
Óbidos	February 1993	388	117125	177181	208228	166166	077077
Óbidos	February 1993	389	000000	147153	220194	166210	077077
Óbidos	February 1993	390	107000	123165	222234	172186	000000
Óbidos	February 1993	391	119121	141175	224194	166166	081081
Óbidos	February 1993	392	000000	123141	194194	166222	000000
Óbidos	February 1993	393	000000	147157	196234	166216	075079
Óbidos	February 1993	394	117123	179179	220242	168186	075075
Óbidos	February 1993	395	000000	147151	218234	164194	095095
Óbidos	February 1993	396	121123	141157	216222	166160	077077
Óbidos	February 1993	397	113119	143155	204224	166170	067115

Locality	Date	Ref.	Dla11	L3	L8	L9	Dla6
Óbidos	February 1993	398	000000	151151	234234	166168	079089
Óbidos	February 1993	301	000000	141165	194224	182186	075075
Óbidos	February 1993	400	000000	141157	228230	166172	067067
Óbidos	February 1993	401	113113	141179	196222	166166	075075
Óbidos	February 1993	402	117119	131179	226234	166168	077077
Óbidos	February 1993	403	119123	177181	226234	166166	077077
Óbidos	February 1993	404	107137	143151	226226	166166	075075
Óbidos	February 1993	405	117119	153179	234246	166166	075075
Óbidos	February 1993	406	117123	171175	226226	176176	077085
Óbidos	February 1993	407	117119	147181	206226	000000	075075
Óbidos	February 1993	410	113119	123189	226226	166168	075077
Óbidos	February 1993	411	000000	145149	222222	166170	075075
Óbidos	February 1993	413	000000	147159	222242	000000	093093
Óbidos	February 1993	414	113119	141155	210226	000000	067079
Óbidos	February 1993	415	000000	141157	212240	166168	077089
Óbidos	February 1993	416	000000	000151	226226	000000	000000
Óbidos	February 1993	417	000000	000000	218234	000000	077077
Óbidos	February 1993	418	117117	000000	000000	000000	000000
Óbidos	February 1993	419	000000	000000	202240	162168	077077
Óbidos	February 1993	420	113113	143159	222234	166166	077081
Óbidos	February 1993	421	117117	141141	226230	000000	075075
Óbidos	February 1993	422	000000	141179	222230	166166	075075
Óbidos	February 1993	423	119121	141171	000000	000000	000067
Óbidos	February 1993	424	119123	157161	212228	166166	075089
Óbidos	February 1993	425	000000	153165	232238	166166	079079
Óbidos	February 1993	426	119123	153141	222230	168166	075075
Óbidos	February 1993	427	000000	131175	224234	166170	075067
Óbidos	February 1993	428	000000	141153	000000	000000	000000
Óbidos	February 1993	429	000000	000000	000000	000000	067075
Óbidos	February 1993	430	000000	000147	194232	166168	075077
Óbidos	February 1993	431	000000	000000	202216	000000	067079
Óbidos	February 1993	432	000000	143179	226226	166166	079077
Óbidos	February 1993	433	000000	141147	226226	000000	075075
Óbidos	February 1993	434	117117	143147	196230	166166	075079
Óbidos	February 1993	435	000000	000000	204234	000000	000000
Óbidos	February 1993	436	000000	157189	212232	158166	075077
Óbidos	February 1993	437	000000	141141	196210	166166	075073
Óbidos	February 1993	438	000000	000000	000000	000000	000000
Óbidos	February 1993	439	117119	000147	194222	000000	075075
Óbidos	February 1993	440	123125	153161	214230	166166	075075
Óbidos	February 1993	441	000000	151159	212234	162166	075075
Óbidos	February 1993	442	119119	141141	226232	166176	089087
Óbidos	February 1993	443	000000	000000	196228	000000	000000
Óbidos	February 1993	444	000000	137169	196230	000000	067075
Óbidos	February 1993	445	113117	141163	206234	162168	067075
Óbidos	February 1993	446	000000	000155	212230	000000	075075
Óbidos	February 1993	447	000000	165171	202198	162166	067075
Óbidos	February 1993	448	113119	151187	200224	168168	099101
Óbidos	February 1993	449	111121	141149	198198	170170	072074
Óbidos	February 1993	450	119107	147157	222230	164166	000000
Milfontes	February 1993	451	109113	141151	210234	166170	075075
Milfontes	February 1993	452	000000	147147	226232	000000	075075
Milfontes	February 1993	453	000000	159183	196228	000168	075075
Milfontes	February 1993	454	000000	159183	194222	166166	075075

Locality	Date	Ref.	Dla11	L3	L8	L9	Dla6
Milfontes	February 1993	455	117123	185133	222234	166166	067075
Milfontes	February 1993	456	117117	151151	226234	166166	075075
Milfontes	February 1993	457	109123	127141	224232	166176	077077
Milfontes	February 1993	458	109123	141165	202228	166166	075081
Milfontes	February 1993	459	117123	123181	220226	166166	076077
Milfontes	February 1993	460	000000	000000	210220	000000	077077
Milfontes	February 1993	461	109123	141181	220226	166170	075075
Milfontes	February 1993	462	117117	141173	228236	192234	075081
Milfontes	February 1993	463	117121	163179	194210	166166	067093
Milfontes	February 1993	464	123123	000000	214226	162180	000000
Milfontes	February 1993	465	123131	141179	222230	166174	075075
Milfontes	February 1993	466	119123	141179	220230	166218	067075
Milfontes	February 1993	467	107119	163181	222232	160160	081081
Milfontes	February 1993	468	113119	151123	000000	172174	067067
Milfontes	February 1993	469	123123	151163	210226	206212	067075
Milfontes	February 1993	470	113119	145153	210226	166166	067075
Milfontes	February 1993	471	000000	177179	208230	166168	074093
Milfontes	February 1993	472	000000	141143	208222	166206	077077
Milfontes	February 1993	473	000000	141147	194222	164168	067077
Milfontes	February 1993	474	113113	141159	196232	166206	075075
Milfontes	February 1993	475	117119	000000	220246	166206	075075
Milfontes	February 1993	476	000000	151141	198224	166196	075101
Milfontes	February 1993	477	117117	161157	230230	166156	075093
Milfontes	February 1993	478	000000	000000	000000	000000	000000
Milfontes	February 1993	479	000000	157159	198232	166168	074074
Milfontes	February 1993	480	113119	157161	220228	174196	067089
Milfontes	February 1993	481	000000	000000	000000	166166	000000
Milfontes	February 1993	482	117119	131163	208222	166174	081081
Milfontes	February 1993	483	117123	147189	220228	166192	075075
Milfontes	February 1993	484	107113	147147	200224	196204	067075
Milfontes	February 1993	485	113123	159177	200230	166172	077077
Milfontes	February 1993	486	000000	147179	224228	000000	065067
Milfontes	February 1993	487	113119	139151	222226	166174	067079
Milfontes	February 1993	488	000000	149157	208228	166240	075073
Milfontes	February 1993	489	000000	000000	000000	000000	000000
Milfontes	February 1993	490	119123	147157	210234	166196	104104
Milfontes	February 1993	491	117117	151153	200234	166192	075079
Milfontes	February 1993	492	113123	147155	234234	166164	067067
Milfontes	February 1993	493	000000	145165	210226	168172	067083
Milfontes	February 1993	494	000000	141147	210224	170174	075089
Milfontes	February 1993	495	000000	141147	212224	166166	075075
Milfontes	February 1993	496	113117	141147	210224	166168	077077
Milfontes	February 1993	497	000000	143143	212228	166166	103103
Milfontes	February 1993	498	107119	181179	212224	166166	093075
Milfontes	February 1993	401	000000	141153	224224	166166	000000
Milfontes	February 1993	500	113119	159161	000000	166166	075075
Milfontes	February 1993	501	000000	000000	000000	000000	000000
Milfontes	February 1993	502	000000	147139	208218	166168	075075
Milfontes	February 1993	503	117119	159165	206230	174188	075077
Milfontes	February 1993	504	000000	139147	000000	166166	000000
Milfontes	February 1993	505	000000	141183	224224	166172	067067
Milfontes	February 1993	506	000000	157161	206226	206200	075075
Milfontes	February 1993	507	000000	000141	192234	166168	000000
Milfontes	February 1993	508	000000	000141	212224	166170	067075
Milfontes	February 1993	509	000000	147157	192234	168186	079079

Locality	Date	Ref.	Dla11	L3	L8	L9	Dla6
Milfontes	February 1993	510	000000	000000	000000	000000	000000
Milfontes	February 1993	511	000000	141151	000000	166166	075075
Milfontes	February 1993	513	000000	000141	000000	166166	067075
Milfontes	February 1993	515	000000	115173	214240	174186	075075
Milfontes	February 1993	516	000000	147159	000000	160160	000000
Milfontes	February 1993	517	000000	133159	224236	166196	000000
Milfontes	February 1993	518	000000	000179	198232	166202	075075
Milfontes	February 1993	520	000000	141175	224238	166178	075075
Milfontes	February 1993	521	000000	141147	202226	166172	067075
Milfontes	February 1994	600	000000	141147	194230	166206	075077
Milfontes	February 1994	601	000000	167173	194220	166216	000000
Milfontes	February 1994	602	000000	147181	194220	172176	077077
Milfontes	February 1994	603	000000	159159	194202	192192	075075
Milfontes	February 1994	604	000000	145179	200230	166166	075075
Milfontes	February 1994	605	000000	141171	212236	164180	075077
Milfontes	February 1994	606	000000	127145	194218	168168	000000
Milfontes	February 1994	607	000000	141153	228226	166166	075075
Milfontes	February 1994	608	000000	153175	196210	166180	087087
Milfontes	February 1994	609	000000	141151	208224	166166	000000
Milfontes	February 1994	610	000000	155173	194214	166172	000000
Milfontes	February 1994	611	000000	127177	200200	166168	067067
Milfontes	February 1994	612	000000	141165	196196	194194	067077
Milfontes	February 1994	613	000000	141167	000000	166184	075075
Milfontes	February 1994	614	000000	155199	200218	184158	075075
Milfontes	February 1994	615	000000	141153	220224	160166	093093
Milfontes	February 1994	616	000000	141159	194198	168166	000000
Milfontes	February 1994	617	000000	141149	204210	166166	079081
Milfontes	February 1994	618	000000	123159	208218	166168	075075
Milfontes	February 1994	619	000000	133141	208208	166166	075075
Milfontes	February 1994	620	000000	000000	200220	166172	000000
Milfontes	February 1994	621	000000	119141	218226	170178	077077
Milfontes	February 1994	622	000000	131177	192218	166166	075075
Milfontes	February 1994	623	000000	151173	202222	158172	075075
Milfontes	February 1994	624	000000	147171	196220	166174	000000
Milfontes	February 1994	625	000000	141153	212220	170202	083083
Milfontes	February 1994	626	000000	141177	214220	164184	075077
Milfontes	February 1994	627	000000	147151	190208	166166	075075
Milfontes	February 1994	628	000000	147151	190190	166174	075081
Milfontes	February 1994	629	000000	155167	000238	176200	075077
Milfontes	February 1994	630	000000	153161	214222	168186	000000
Milfontes	February 1994	631	000000	141147	000216	168168	075075
Milfontes	February 1994	632	000000	159159	000224	166196	075077
Milfontes	February 1994	633	000000	149159	000206	166208	075075
Milfontes	February 1994	636	000000	151143	206218	170188	093093
Milfontes	February 1994	638	000000	157151	192224	166166	075089
Milfontes	February 1994	639	000000	157157	230190	166172	000000
Milfontes	February 1994	640	000000	000141	000000	166166	075077
Milfontes	February 1994	641	000000	139139	192210	166206	079079
Milfontes	February 1994	642	000000	141159	000226	166166	075077
Milfontes	February 1994	643	000000	141141	226218	164170	077077
Milfontes	February 1994	644	000000	147161	208230	174196	067097
Milfontes	February 1994	645	000000	000151	000000	174184	000000
Milfontes	February 1994	646	000000	141153	208226	166166	075075
Milfontes	February 1994	647	000000	141151	208220	166186	075077
Milfontes	February 1994	648	000000	153161	212226	158174	067075

Locality	Date	Ref.	Dla11	L3	L8	L9	Dla6
Milfontes	February 1994	649	000000	153155	220228	166166	075077
Milfontes	February 1994	650	000000	141147	194194	000000	089089
Milfontes	February 1994	651	000000	000000	226226	000000	075091
Milfontes	February 1994	652	000000	161159	000000	000000	000000
Milfontes	February 1994	653	000000	127141	230230	000000	079079
Milfontes	February 1994	654	000000	000141	000000	166166	000000
Milfontes	February 1994	655	000000	000141	194242	166176	075075
Milfontes	February 1994	656	000000	141177	212224	166174	000000
Milfontes	February 1994	657	000000	141147	212228	166166	067075
Milfontes	February 1994	659	000000	159163	194226	166168	067075
Milfontes	February 1994	660	000000	141177	194232	166166	075089
Milfontes	February 1994	661	000000	000177	000000	166166	077077
Milfontes	February 1994	662	000000	159177	194222	166166	075077
Milfontes	February 1994	663	000000	155173	212226	166166	075075
Milfontes	February 1994	665	000000	151167	194232	166168	079085
Milfontes	February 1994	666	000000	141157	216222	166164	077075
Milfontes	February 1994	668	000000	141175	202226	166188	075075
Milfontes	February 1994	669	000000	167173	000000	000000	000000
Milfontes	February 1994	670	000000	157161	208220	000000	000000
Milfontes	February 1994	672	000000	000000	000000	000000	075075
Milfontes	February 1994	673	000000	141171	000000	166166	000000
Milfontes	February 1994	674	000000	159173	000000	000000	075075
Milfontes	February 1994	676	000000	000000	000000	166166	000000
Milfontes	February 1994	677	000000	153177	226224	166160	075075
Milfontes	February 1994	679	000000	141151	000224	000000	000000
Milfontes	February 1994	680	000000	145177	000000	000000	000000
Milfontes	February 1994	682	000000	133165	204226	000000	000000
Milfontes	February 1994	683	000000	000000	000000	166190	000000
Milfontes	February 1994	684	000000	000151	000000	166166	000000
Milfontes	February 1994	685	000000	141159	000000	166168	000000
Milfontes	February 1994	686	000000	123173	000000	166166	000000
Milfontes	February 1994	687	000000	000155	000000	166166	000000
Milfontes	February 1994	689	000000	141177	000000	000000	000000
Milfontes	February 1994	691	000000	137147	000000	000000	000000
Milfontes	February 1994	692	000000	141155	000000	000000	000000
Milfontes	February 1994	698	000000	000153	000000	000000	000000
Milfontes	February 1994	709	000000	149179	000000	000000	000000
Milfontes	February 1994	712	000000	000000	000000	166166	000000
Milfontes	February 1994	713	000000	000145	210234	170174	075075
Milfontes	February 1994	715	000000	147159	000000	000000	000000
Milfontes	February 1994	716	000000	000000	000000	160168	000000
Milfontes	February 1994	718	000000	000000	000000	000170	000000
Milfontes	February 1994	719	000000	000000	000000	166166	000000
Milfontes	February 1994	724	000000	141153	000000	000000	000000
Milfontes	February 1994	725	000000	000147	000000	160166	000000
Milfontes	February 1994	735	000000	000000	200212	166168	000000
Milfontes	February 1994	736	000000	000000	210232	166166	000000
Faro	February 1992	11	000000	141151	000000	000000	000000
Faro	February 1992	12	000000	000155	000000	000000	000000
Faro	February 1992	13	000000	141153	000000	000000	000000
Faro	February 1992	14	000000	141181	000000	000000	000000
Faro	February 1992	15	000000	141177	000000	000000	000000
Faro	February 1992	16	000000	141183	000000	000000	000000
Faro	February 1992	18	000000	169189	000000	000000	000000

Locality	Date	Ref.	Dla11	L3	L8	L9	Dla6
Faro	February 1992	21	000000	125159	196206	000000	000000
Faro	February 1992	22	000000	141141	000000	000000	000000
Faro	February 1992	23	000000	000000	000228	000000	000000
Faro	February 1992	24	000000	000153	000000	000000	000000
Faro	February 1992	25	000000	141171	000000	000000	000000
Faro	February 1992	26	000000	143155	202228	000000	000000
Faro	February 1992	27	000000	151151	000000	000000	000000
Faro	February 1992	28	000000	000000	212216	000000	000000
Faro	February 1992	29	000000	139193	000000	000000	000000
Faro	February 1992	30	000000	000000	202226	000000	000000
Faro	February 1992	31	000000	151187	000000	000000	000000
Faro	February 1992	32	000000	147159	000000	000000	000000
Faro	February 1992	34	000000	155199	000000	000000	000000
Faro	February 1992	38	000000	139173	000000	000000	000000
Faro	February 1992	41	000000	153127	000000	000000	000000
Faro	February 1992	43	000000	000141	000000	000000	000000
Faro	February 1992	47	000000	141155	000000	000000	000000
Faro	February 1992	52	000000	141171	000000	000000	000000
Faro	February 1992	54	000000	141153	000000	000000	000000
Faro	February 1992	55	000000	141133	000000	000000	000000
Faro	February 1992	56	000000	141153	000000	000000	000000
Faro	February 1992	57	000000	141153	000000	000000	000000
Faro	February 1992	58	000000	000167	202214	166174	067073
Faro	February 1992	59	000000	143139	196226	166164	075093
Faro	February 1992	60	000000	149165	232236	168168	073075
Faro	February 1992	61	000000	141153	198232	166166	077077
Faro	February 1992	62	000000	125151	212218	166166	000000
Faro	February 1992	63	000000	141153	232232	170170	077081
Faro	February 1992	64	000000	167187	202226	166166	077077
Faro	February 1992	65	000000	145153	196196	166182	073075
Faro	February 1992	66	000000	151169	212230	166166	075075
Faro	February 1992	67	000000	147183	230238	000000	075075
Faro	February 1992	68	000000	149165	222216	166168	075075
Faro	February 1992	69	000000	141147	222234	000000	075075
Faro	February 1992	70	000000	159163	000212	168168	075081
Faro	February 1992	71	000000	157157	204218	166170	067081
Faro	February 1992	72	000000	141179	198204	166166	075077
Faro	February 1992	73	000000	157179	216226	166166	075075
Faro	February 1992	74	000000	173177	212230	168168	089089
Faro	February 1992	75	000000	149159	198212	166166	087089

ANNEX 5

MICROSATELLITE ALLELE FREQUENCIES

Locus: *Dla11*

Pop/Alleles	99	101	103	105	107	109	111	113	115	117	119	121	123	125	127	129	131	135	137	141
Foz	0.000	0.000	0.012	0.012	0.050	0.031	0.043	0.099	0.099	0.267	0.168	0.081	0.093	0.012	0.006	0.000	0.012	0.012	0.000	0.000
Aveiro	0.007	0.014	0.000	0.000	0.036	0.058	0.014	0.094	0.086	0.309	0.137	0.058	0.108	0.043	0.000	0.029	0.000	0.000	0.000	0.007
Óbidos	0.000	0.000	0.000	0.009	0.027	0.000	0.018	0.133	0.044	0.257	0.195	0.124	0.124	0.044	0.000	0.000	0.000	0.000	0.027	0.000
Milfontes	0.000	0.000	0.000	0.000	0.047	0.063	0.000	0.188	0.000	0.250	0.188	0.016	0.234	0.000	0.000	0.000	0.016	0.000	0.000	0.000
Faro	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Locus: *Dla6*

Pop/Alleles	55	65	66	6	68	69	72	73	74	75	76	77	78	79	81	83	85	87	89	91	93
Foz	0.000	0.000	0.000	0.033	0.000	0.033	0.000	0.022	0.000	0.522	0.000	0.209	0.000	0.049	0.038	0.022	0.000	0.011	0.027	0.005	0.005
Aveiro	0.008	0.000	0.008	0.087	0.008	0.000	0.000	0.024	0.016	0.472	0.024	0.134	0.008	0.063	0.031	0.024	0.000	0.000	0.024	0.016	0.016
Óbidos	0.000	0.006	0.000	0.120	0.000	0.000	0.006	0.013	0.006	0.449	0.000	0.190	0.000	0.070	0.025	0.000	0.006	0.006	0.051	0.000	0.013
Milfontes	0.000	0.005	0.000	0.122	0.000	0.000	0.000	0.005	0.014	0.500	0.005	0.149	0.000	0.045	0.036	0.014	0.005	0.009	0.027	0.005	0.036
Faro	0.000	0.000	0.000	0.059	0.000	0.000	0.000	0.088	0.000	0.441	0.000	0.176	0.000	0.000	0.088	0.000	0.000	0.029	0.088	0.000	0.029

Locus: *Labrax-3*

Pop/Alleles	113	115	119	121	123	125	127	129	131	133	135	137	139	141	143	145	147	149	151	153	155
Foz	0.006	0.006	0.000	0.006	0.000	0.017	0.006	0.000	0.023	0.023	0.006	0.006	0.029	0.253	0.034	0.029	0.075	0.029	0.057	0.029	0.029
Aveiro	0.000	0.000	0.000	0.000	0.000	0.012	0.025	0.000	0.025	0.019	0.000	0.012	0.012	0.259	0.037	0.019	0.043	0.025	0.049	0.068	0.037
Óbidos	0.000	0.000	0.000	0.000	0.018	0.000	0.006	0.006	0.012	0.000	0.006	0.006	0.006	0.222	0.029	0.006	0.099	0.012	0.058	0.070	0.023
Milfontes	0.000	0.004	0.004	0.000	0.015	0.000	0.015	0.000	0.007	0.015	0.000	0.004	0.019	0.211	0.015	0.022	0.107	0.015	0.074	0.052	0.030
Faro	0.000	0.000	0.000	0.000	0.000	0.024	0.012	0.000	0.000	0.012	0.000	0.000	0.036	0.226	0.024	0.012	0.036	0.036	0.071	0.107	0.048

Locus: *Labrax-8*

Pop/Alleles	190	192	194	196	198	200	202	204	206	208	210	212	214	216	218	220	222	224	226	228	230
Foz	0.000	0.006	0.077	0.013	0.019	0.026	0.006	0.019	0.006	0.032	0.071	0.039	0.045	0.065	0.026	0.071	0.052	0.071	0.129	0.052	0.045
Aveiro	0.000	0.008	0.086	0.023	0.000	0.047	0.039	0.016	0.016	0.023	0.047	0.094	0.031	0.008	0.000	0.055	0.086	0.078	0.078	0.031	0.086
Óbidos	0.000	0.006	0.055	0.036	0.018	0.006	0.030	0.012	0.024	0.024	0.055	0.036	0.006	0.024	0.036	0.030	0.127	0.042	0.158	0.030	0.067
Milfontes	0.017	0.021	0.072	0.025	0.017	0.038	0.021	0.008	0.017	0.059	0.059	0.047	0.021	0.008	0.034	0.072	0.051	0.093	0.102	0.047	0.055
Faro	0.000	0.000	0.000	0.091	0.068	0.000	0.091	0.045	0.023	0.000	0.000	0.136	0.023	0.068	0.045	0.000	0.045	0.000	0.091	0.045	0.068

Locus: *Labrax-9*

Pop/Alleles	144	146	156	158	160	162	164	166	168	170	172	174	176	178	180	182	184	186	188	190	192
Foz	0.000	0.000	0.023	0.000	0.034	0.000	0.045	0.455	0.159	0.080	0.091	0.045	0.034	0.000	0.000	0.000	0.000	0.000	0.000	0.011	0.000
Aveiro	0.017	0.000	0.000	0.017	0.095	0.052	0.009	0.474	0.112	0.086	0.043	0.043	0.017	0.009	0.000	0.009	0.009	0.000	0.000	0.000	0.000
Óbidos	0.000	0.007	0.000	0.007	0.007	0.041	0.027	0.479	0.103	0.055	0.034	0.014	0.041	0.000	0.007	0.007	0.000	0.041	0.000	0.000	0.000
Milfontes	0.000	0.000	0.004	0.011	0.030	0.004	0.022	0.522	0.081	0.037	0.037	0.056	0.015	0.007	0.011	0.000	0.015	0.015	0.011	0.004	0.019
Faro	0.000	0.000	0.000	0.000	0.000	0.000	0.031	0.594	0.219	0.094	0.000	0.031	0.000	0.000	0.000	0.031	0.000	0.000	0.000	0.000	0.000

Locus: ***Dla11***

Pop/Alleles	99
Foz	0.000
Aveiro	0.007
Óbidos	0.000
Milfontes	0.000
Faro	-

Locus: ***Dla6***

Pop/Alleles	55	95	97	99	101	103	104	115
Foz	0.000	0.000	0.011	0.000	0.005	0.005	0.000	0.000
Aveiro	0.008	0.016	0.000	0.000	0.024	0.000	0.000	0.000
Óbidos	0.000	0.013	0.000	0.006	0.013	0.000	0.000	0.006
Milfontes	0.000	0.000	0.005	0.000	0.005	0.009	0.009	0.000
Faro	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Locus: ***Labrax-3***

Pop/Alleles	113	157	159	161	163	165	167	169	171	173	175	177	179	181	183	185	187	189	193	197	199
Foz	0.006	0.023	0.080	0.040	0.011	0.006	0.000	0.006	0.023	0.029	0.029	0.023	0.029	0.011	0.000	0.006	0.011	0.006	0.000	0.006	0.000
Aveiro	0.000	0.043	0.074	0.037	0.037	0.043	0.006	0.019	0.006	0.000	0.019	0.025	0.019	0.012	0.012	0.006	0.000	0.000	0.000	0.000	0.000
Óbidos	0.000	0.053	0.076	0.035	0.006	0.053	0.012	0.012	0.029	0.000	0.018	0.018	0.053	0.029	0.006	0.000	0.012	0.012	0.000	0.000	0.000
Milfontes	0.000	0.044	0.085	0.033	0.019	0.019	0.019	0.000	0.011	0.033	0.011	0.044	0.033	0.019	0.011	0.004	0.000	0.004	0.000	0.000	0.004
Faro	0.000	0.036	0.048	0.000	0.012	0.024	0.024	0.024	0.024	0.024	0.000	0.024	0.024	0.012	0.024	0.000	0.024	0.012	0.012	0.000	0.012

Locus: ***Labrax-8***

Pop/Alleles	190	232	234	236	238	240	242	246	248
Foz	0.000	0.045	0.026	0.045	0.000	0.006	0.000	0.000	0.006
Aveiro	0.000	0.047	0.023	0.031	0.023	0.000	0.008	0.016	0.000
Óbidos	0.000	0.030	0.091	0.000	0.018	0.012	0.012	0.012	0.000
Milfontes	0.017	0.038	0.042	0.013	0.008	0.004	0.004	0.004	0.000
Faro	0.000	0.091	0.023	0.023	0.023	0.000	0.000	0.000	0.000

Locus: ***Labrax-9***

Pop/Alleles	144	194	196	200	202	204	206	208	210	212	214	216	218	222	234	240	250	254
Foz	0.000	0.000	0.000	0.000	0.011	0.000	0.000	0.000	0.000	0.000	0.011	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Aveiro	0.017	0.000	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Óbidos	0.000	0.021	0.014	0.007	0.000	0.007	0.014	0.000	0.021	0.007	0.000	0.007	0.007	0.007	0.000	0.000	0.007	0.014
Milfontes	0.000	0.007	0.026	0.007	0.007	0.004	0.026	0.004	0.000	0.004	0.000	0.004	0.004	0.000	0.004	0.004	0.000	0.000
Faro	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000