GENETIC VARIATION AND EVOLUTIONARY HISTORY OF THE MACARONESIAN FERN *Asplenium hemionitis* L.

Cecília J. R. Durães

PhD Thesis

The University of Edinburgh

February 2005

ABSTRACT

This project investigates the population genetics of the fern *Asplenium hemionitis* L. with a view to elucidate its biogeographic history. *A. hemionitis* is distributed in the Macaronesian region (Azores, Madeira, Canary Islands and Cape Verde) and a few restricted coastal sites on the mainland (W Portugal and NW Africa). This diploid fern is present in the fossil record of southern Europe and it is considered a relict from the Tertiary. It has been suggested that during past glaciations, several plant species with Mediterranean distribution survived in the Macaronesian islands where the climate was warmer. These aspects make *A. hemionitis* an ideal species for the study of colonisation of the Macaronesian region and possible relationships with the mainland. Nuclear (isozymes) and organelle (chloroplast) markers have been used to investigate 105 populations covering most of the distribution range of *A. hemionitis*.

Population genetic analysis based on isozyme markers shows that whereas outcrossing appears to be the prevailing breeding system in populations of *A. hemionitis*, there is evidence for inbreeding. The amount of inbreeding verified is mostly due to intergametophytic selfing and may be associated with limited spore dispersal and fine structuring within populations. Populations sampled across the Macaronesian islands and the mainland maintain high levels of isozyme diversity, the highest levels being found on the Canary Islands and Madeira. This diversity is distributed mainly within rather than between populations. Estimations of genetic differentiation and distance show that Madeira is highly divergent from all the other regions but more from the mainland. The less divergent regions are the Canary Islands and Portugal, and the former and the Azores. The whole population set shows significant isolation by distance and, although less is verified within regions and islands, it is still significant.

Sequencing of the *trn*L and *rps*4 regions of the chloroplast genome (c.1900 bp) revealed 29 haplotypes and high diversity within all regions. Analysis of differentiation between populations and geographical regions shows that N_{ST} (takes into account genetic similarities between haplotypes) is higher than G_{ST} (based on variation in haplotype

frequencies alone). This result implies a certain degree of phylogeographic structure, showing that *A. hemionitis* has been evolving independently within the main archipelagos and the mainland. The integrated results of the isozyme and chloroplast DNA analysis are ultimately used to make inferences on the factors determining the observed levels and patterns of genetic diversity and the evolutionary history of *A. hemionitis*.

DECLARATION

This is my own work and any contributions by others are clearly cited.

Cecília J. R. Durães

Edinburgh, February 2005

ACKNOWLEDGMENTS

I would firstly like to thank my supervisor, Dr. Richard Ennos, whose guidance and support throughout my PhD have been invaluable, particularly so during my stressful writing up period. Whereas most people are content with one supervisor, I needed three! This work would never have been undertaken without Drs. Johannes Vogel (NHM) and Mary Gibby (RBGE), who not only conceived this project, but were also instrumental throughout its course.

Without the help of Fred Rumsey, Sally Henderson, Mark Carine and Patrick Acock (NHM), I wouldn't have been able to collect the thousands of plants used in this study and the experience would have been lonely and miserable. I will always have *frond* memories of those trips. High throughput isozyme analysis is a team effort! Harriet Hunt, Michael Grundmann, Steve Russell and Johannes Vogel (NHM) were my team members in the non-stop factory of grinding samples, cutting and scoring gels.

I would also like to thank Dr. Fátima Sales and Mr Ian Hedge for their friendship and encouragement.

Thank you Kiko, Marco, Rui, Carla and Fátima for visiting me in Edinburgh and keeping me sane in the absence of all things Portuguese. British Telecom would like to thank you too. Obrigado Mãe e Pai por todo o incentivo e por estarem sempre em casa aos Sábado à hora de almoço!

Thanks to Neil Gaiman and Morpheus for sharing their dreams.

Finally, I would like to thank Tony. You are really special.

This project was financed by Fundação para a Ciência e a Tecnologia, Portugal (grant reference SFRH/BD/3249/2000).

TABLE OF CONTENTS

APPENDICES

LIST OF ABBREVIATIONS

A	adenine
AFLP	amplified fragment length polymorphism
ATP	adenosine triphosphate
bp	base pair(s)
BP	before present
C	cysteine
CAPS	cleaved amplified polymorphic sequence
cpDNA	chloroplast DNA
dCAPS	derived cleaved amplified polymorphic sequence
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphate
E	East
EDTA	ethylenediaminetetraacetic acid
G	guanine
GDA	Genetic Data Analysis
H	histidine
IBD	isolation by distance
ISSR	Inter simple sequence repeat
ITS	internal transcribed spacer
M	molar
mM	millimolar
MSN	minimum spanning network
mtDNA	mitochondrial DNA
My	million years
N	North
N/S	North/South
nDNA	nuclear DNA
NHM	The Natural History Museum, London
NW	North West
PCR	polymerase chain reaction
PCR-RFLP	polymerase chain reaction-restriction fragment length polymorphism
pmol	picomole
PVP	Polyvinyl-pyrrolidone
RAPD	Randomly amplified polymorphic DNA
RBGE	Royal Botanic Garden Edinburgh
RFLP	restriction fragment length polymorphism
S	South
SNP	single nucleotide polymorphism
SSR	simple sequence repeat (microsatellite)
SW	South West
T	thymine
TE	Tris-EDTA buffer

TRIS	Tris[hydroxymethyl]aminomethane
UV	ultraviolet
W	West
y	year(s)

CHAPTER 1

INTRODUCTION

1.1 Colonisation and evolution on oceanic islands

Oceanic islands are considered hotspots for speciation and endemism, which arise through isolation or adaptive radiation, more often a combination of both (Humphries, 1979; Baldwin *et al.*, 1998; Emerson, 2002). The isolated Hawaiian archipelago constitutes a classical example, where about 90% of the angiosperms are endemic (Wagner *et al.*, 1990). The Macaronesian archipelagos, situated in the N Atlantic Ocean (Azores, Madeira, the Canary Islands and Cape Verde), are far less isolated, and thus contain fewer endemics [20% overall (Humphries, 1979)]; 40% in the Canary Islands (Santos-Guerra, 1999)]. Both Hawaii and Macaronesia, however, contain a smaller proportion of endemic pteridophytes, 69% (Palmer, 2002) and 18%, respectively (Hansen & Sunding, 1993).

The lower rate of speciation in insular pteridophytes as compared with angiosperms can be explained by higher rates of dispersal between established populations due to efficient dispersal of spores via wind (Ranker *et al.*, 1994). Interpopulational gene flow inhibits genetic divergence in the absence of strong diversifying selection, and so groups with more easily dispersed propagules (e.g. cryptogams) would be expected to exhibit slower rates of divergence and speciation than those with less vagile propagules (Ranker *et al.*, 1994). It has been shown by Muñoz *et al.* (2004) that wind is one of the main forces driving the current distribution of cryptogamic plants (pteridophytes, mosses, liverworts and lichens). The authors investigated c. 1850 cryptogam species distributed in islands of the southern hemisphere, all dispersing by spores, and found a stronger association of floristic similarities with maximum wind connectivity than with geographic proximity.

Traditional and molecular studies of several plant groups have revealed similar characteristics with respect to the colonisation and evolution on oceanic islands (Baldwin *et al.*, 1998; Hess *et al.*, 2000) which can be summarised as the following: a) the success of long-distance dispersal is inversely proportional to island remoteness, i.e., the closer the island to another land mass, the higher the probability of colonisation; b) introductions to distant archipelagos take place once; c) geographic isolation reduces gene flow between populations and over time, colonial populations become genetically

divergent from their parent population due to mutation, genetic drift and/or natural selection; d) return to continents appears to be rare.

These principles of island colonisation and evolution are based predominantly on studies of angiosperm groups; pteridophyte groups, however, have received far less attention. Identifying equivalent principles for pteridophytes requires detailed investigations at the interface of population and evolutionary biology. *Asplenium hemionitis* L. is a pteridophyte distributed in the Macaronesian islands, showing no apparent diversification. The various islands are arranged in a system of archipelagos that are at different distances from the mainland, where a few outlying populations are also present. *A. hemionitis* has been used here as a model species to determine whether the principles of island colonisation and evolution, established from studies on angiosperms, hold true for pteridophytes on the Macaronesian islands. In order to obtain information on gene flow, phylogeography and patterns of colonisation, evidence from the distribution of a variety of different molecular markers within and among populations of the species is employed.

1.2 The Macaronesian region: location, origin and ecological conditions

1.2.1 Definition and geography

Macaronesia is used to describe an area of natural delimitation that includes the archipelagos of the Azores, Madeira, the Salvage Islands, the Canary Islands and Cape Verde (Figure 1.1). This region has a biogeographical relationship with an area in South Morocco (NW Africa), which is considered to be a Macaronesian enclave on the mainland (Peltier, 1973; Sunding, 1979). Some authors, based on the phytogeographic affinities among the archipelagos and between them and the mainland, reject the term 'Macaronesia' as a biogeographical unit (Dias, 1996; Rivas-Martínez *et al.*, 1993; Schäfer, 2003) and suggest that it should be used only as a synonym for 'middle Atlantic islands'.

The five Macaronesian archipelagos are situated in the Atlantic Ocean, from about 39° N to about 15° N. The distances from the European and African continents vary from 115km (Canary Islands) to 1600km (Azores) and the islands are between 15km^2 (Salvage Islands) and 7000km^2 (Canary Islands). There is also a large variation in altitude, with the Salvage Islands having a maximum elevation of 183m whilst Pico de Teide (Canary Islands) is more than 3700m above sea level. Table 1.1 summarises the physical features of the Macaronesian islands.

Table 1.1 Physical characteristics of the Macaronesian islands (adapted from Sunding, 1979; Silva *et al.*, 2000; Schäfer, 2003).

Archipelagos	No. of islands	Area (km^2)	Maximum altitude (m)	Distance to mainland (km)	Distance to nearest archipelago (km)
Azores	9	2348	2351	1584	900
Madeira	3	728	1861	600	260
Salvage Islands	2	<15	183	360	170
Canary Islands	7	7273	3718	95	170
Cape Verde	10	4033	2829	500	1400

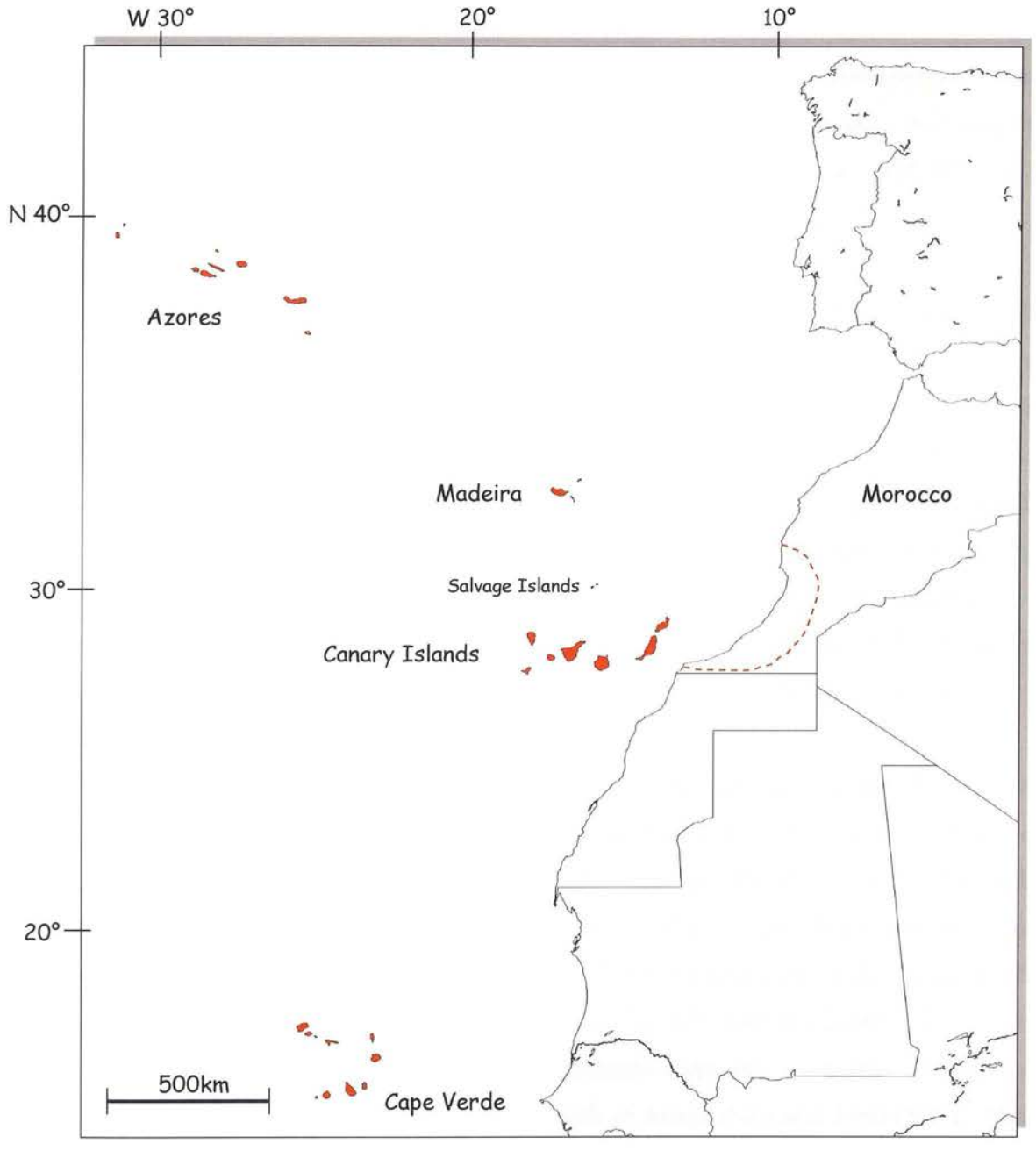


Figure 1.1 The phytogeographical region of Macaronesia. Constituting archipelagos are highlighted in red; dashed red line represents approximately the Macaronesian 'enclave' in south Morocco (adapted from Sunding, 1979).

1.2.2 Origin and geology of the Macaronesian archipelagos

Many different opinions have been held on the origin of the Macaronesian archipelagos (especially the Canary Islands and Cape Verde), but the most recent consider the islands as the products of oceanic volcanism (Scheidegger, 2002; Abratis *et al.*, 2002; Anguita & Hernán, 2000). The substrates consist mainly of basaltic rocks and the soil is typically the product of basaltic lava and pyroclasts (Schäfer, 2003).

The Macaronesian archipelagos exhibit a wide range of geological ages (Table 1.2). The oldest islands are found in the archipelagos of Cape Verde (Maio may be as old as 16.0My) and the Canary Islands (Fuerteventura is at least 20.0My old). In the archipelago of Madeira, Porto Santo is the oldest island (14.0My) but Madeira is much younger (4.6My). The archipelago of the Azores includes the youngest islands, although Santa Maria could be around five to eight My old; however, the upper age limit may not be reliable (Féraud *et al.*, 1984). The most western of the Canary Islands (El Hierro and La Palma) were formed much later than the rest of the archipelago and, consequently, are younger islands, with ages of 0.8-1.0 and 2.0My, respectively.

For the Azores, Madeira and Canary Islands, there is geological evidence of a long and complex volcanic history. The island of Madeira suffered several cycles of volcanic activity, the last eruptions occurring around 6500y BP (Geldmacher *et al.*, 2000). In the Canary Islands, several periods of volcanic activity and quiescence have occurred in the past 20My on each island. The exception is La Gomera, the only island in the archipelago without major activity for the past 4My (Juan *et al.*, 2000). In Tenerife (Canary Islands), relatively recent (<2My) volcanic eruptions, generated massive lava flows that joined the previously separate massifs of Anaga (NE) and Teno (NW). These geologically old massifs (>4.5My) have not been covered by recent lava flows (Juan *et al.*, 2000). The Azores are in close proximity to the tectonically active Mid-Atlantic Ridge. The islands are hotspot volcanoes where seismic and volcanic eruptions frequently occur (Schäfer, 2003), especially in the central and eastern groups (situated on the Eurasian plate). The latest volcanic eruption has been reported from Faial (at the centre of the archipelago) in 1957/58. The western group (Flores and Corvo), situated

on the N American plate, is moving away from the Mid-Atlantic Ridge, and the last volcanic activity dates back to about 3000y (Schäfer, 2003).

Geological evidence indicates a possible ancient land connection between Madeira, the Canary Islands and the mainland. The eastern most Canary Islands (Fuerteventura and Lanzarote) are presently 95km from the African coast, but at periods during the last 20My, they have probably been as close as 65km (García-Talavera, 1999). There are also more than 20 large volcanic sea mounts located between the Canary Islands, Madeira, Salvage Islands and the continent, several of which are presently less than 100m below sea level (Carine *et al.*, 2004). Recent studies have shown that some of these sea mounts are at least 68My old (Geldmacher *et al.*, 2001) and that the N Atlantic sea level was c. 120m below present level during the last glacial maximum, c. 17000y BP (Fairbanks, 1989). It is plausible that these sea mounts may have acted as 'stepping stones' during the glacial periods when sea levels were lower (García-Talavera, 1997), thus facilitating the dispersal of plants and animals to more isolated islands.

Table 1.2 Age of the Macaronesian islands in millions of years (My); * no information; the islands are presented in a W-E direction. These data have been summarised in various papers (a-g).

Archipelagos	Islands	Islands age (My)
Azores	Flores	2.9[a]; 2.15[b]
	Corvo	2.15[b]
	Faial	0.73[a, b]
	Pico	0.3[a]; 0.4[b]
	São Jorge	0.55[a, b]
	Graciosa	2.5[a]; 0.62[b]
	Terceira	2.0[a]; 3.52[b]
	São Miguel	4.01[a, b]
	Santa Maria	8.12[a]; 5.1[b]
Madeira	Madeira	4.6[c]
	Desertas	3.6[c]
	Porto Santo	14.0[c]
Salvage Islands	Selvagens (3 islands)	12.0[c]
Canary Islands	El Hierro	0.8[d]; 1.0[e]
	La Palma	2.0[d, e]
	La Gomera	12.5[d]; 10.0[e]
	Tenerife	11.6[d, e]
	Gran Canaria	14.0[d, e]-16.0[e]
	Fuerteventura	21.0[d]; 20.0[e]
	Lanzarote	15.5[d, e]
Cape Verde	Santiago	(4.0-5.0) 8.5-9.8[f]
	Maio	8.0-16.0[g]
	Other islands	*

[a] Silva *et al.* (2000); [b] Schäfer (2003); [c] Geldmacher *et al.*, 2000; [d] Francisco-Ortega *et al.* (1996b); [e] Emerson *et al.* (2000a); [f] Furnes & Stillman (1987); [g] Silva *et al.* (1981)

1.2.3 Climatic and ecological conditions

The Canary Islands can be divided up into two major phytogeographical units (Humphries, 1979). The eastern islands (Lanzarote and Fuerteventura) are arid most of the time due to their low elevation (c.650m), and because they are in close proximity to Africa (95km) they are strongly influenced by Saharan climatic conditions. The western and central islands have much higher elevations (up to 3718m in Tenerife), are situated further from Africa (200-360km), have a more humid oceanic climate and are consequently far less influenced by the Saharan conditions.

In addition, the prevailing northerly winds (trade winds), are responsible for pronounced N/S differences in many of the islands and result in high ecological diversity within the archipelago (Humphries, 1979). Generally, while a persistent cloud layer and evergreen forests prevail on the northern slopes of the western islands, the south faces are drier and occupied predominantly by xerophytic, semi-desert and halophytic species (Bramwell & Bramwell, 2001). On both N and S faces, lowland coastal vegetation is characterised by xerophytic elements.

The island of Madeira, like the western Canary Islands, is also under the influence of prevailing northerly winds and thus shows a distinct N/S difference, very similar to that found in more western Canary Islands (Humphries, 1979). However, both temperature and humidity gradients are much lower than in the Canary Islands and so the vegetation zones are less differentiated. The islands of Porto Santo and Desertas are characterised by lowland xerophytic communities (Sjögren, 1973).

The Azores are much colder and wetter than any islands in the archipelagos of Madeira and Canary Islands (Humphries, 1979). The temperature is more or less constant throughout the year and the humidity is rather high (exceeds 95% for one fifth of the year). The climate is mainly regulated by a branch of the Gulf Stream and by the Azores anti-cyclone, a zone of high pressure frequently occurring in the Mid Atlantic (Schäfer, 2003). The influence of the trade winds is more noticeable on the eastern most islands (Santa Maria and São Miguel). In general, cloudy and foggy weather is

present for a great part of the year in all islands. Consequently, the Azores islands lack the N/S vegetation differentiation that characterises Madeira and the Canary Islands, and the only real variations are in the most part due to altitudinal gradients (Humphries, 1979).

The Cape Verde Islands also lack vegetation zonation, but the climatic and ecological conditions are widely different from those found on the Azores (Humphries, 1979). The archipelago is subtropical, uniformly arid and suffers from long periods of drought throughout the year. As a result, the vegetation is a mix of tropical, xerophytic and grassland species.

The Salvage Islands are rather small in size and in terms of altitude (180m). These features are responsible for their climate and ecological conditions, which are similar to those found on the low areas of the N-facing slopes of the western Canary Islands, i.e. arid.

In conclusion, the combination of prevailing humid northerly winds (trade winds) with the altitudinal gradients on islands has produced several distinct ecological zones with a high diversity of habitats (Bramwell, 1972; Humphries, 1979; Bramwell & Bramwell, 2001). These conditions are, ultimately, the major factors responsible for the rich, and to a great extent endemic, Macaronesian flora.

1.3 Flora of the Macaronesian region

In the recent years, the Macaronesian archipelagos (especially the Canary Islands) have been the objects of intensive studies. Based on geological and molecular data, scientists have been able to make inferences concerning the origin and evolution of the flora, focusing particularly on species diversification (e.g. adaptive radiation processes) and colonisation events. In the following section, the composition and distribution patterns of the Macaronesian flora will be considered. This in turn leads to the discussion of the origin and evolution of that flora, focusing on the source area of introductions and aspects of the geological and climatic history of Europe and northern Africa during the Tertiary period.

While considering the composition of the Macaronesian flora, it is important to consider that the Macaronesian region comprises five individual archipelagos, each with its own distinctive characteristics. Undoubtedly, they have common 'Macaronesian' features, and are hence included in the same phytogeographic region.

The flora of Macaronesia exhibits a number of typical oceanic island characteristics (Carine *et al.*, 2004), namely a high degree of endemism [20% overall (Humphries, 1979)]; 40% in the Canary Islands (Santos-Guerra, 1999)] and a predominant woody habit among endemics (c. 70% of Canary Islands endemics are woody; Aldridge, 1979). The endemic plants are certainly a significant group in Macaronesia, and will be considered in more detail in conjunction with the other major floristic elements (following Sunding, 1970; Bramwell, 1972; Bramwell & Richardson, 1973; Bramwell, 1976; Sunding, 1979). Of particular relevance is the *Laurisilva* forest, which involves a significant proportion of the natural habitat. This will be discussed in subsequent sections.

1.3.1 Floristic elements

The present-day flora of Macaronesia can be divided into a number of floristic or phytogeographic elements based on affinities with external (i.e. non-Macaronesian) *taxa*. In this context, the term 'floristic element' is used to designate species of the same present-day distribution, or species of the same floristic region, sometimes including in it, in a wider sense, aspects of history and origin (Sunding, 1979).

Within the region of Macaronesia (including all archipelagos) there are approximately 3200 flowering species (Sunding, 1979), 83 pteridophytes (Hansen & Sunding, 1993) and 600 bryophytes (Freitas & Brehm, 2001). The major floristic elements in the flora of Macaronesia are the Mediterranean, the Atlantic, the Saharo-Sindian, the tropical, the cosmopolitan and the endemic. The Mediterranean element comprises a large group of species, e.g. the classical example *Olea europaea* L. The Atlantic element includes mainly pteridophytes and it is more frequent in the Azores. The Saharo-Sindian element is especially prevalent in the lower and arid parts of the Canary Islands and Cape Verde. The tropical element is represented in all archipelagos, but is most prevalent in Cape Verde; it is especially well represented by Bryophytes (Sjögren, 2000). The endemic species can be further subdivided, according to their present-day relations to the flora of other parts of the world and includes both endemics of a wider part of the Macaronesian area and endemics of single archipelagos or single islands. The endemic element affinities are indicated in more detail in Table 1.3.

Table 1.3 Subdivision of the endemic element in Macaronesia and its various affinities (adapted from Sunding, 1979).

Groups of endemics	Characteristics and examples
Mediterranean affinity	A greater part of the endemic flora of Madeira and the Canary Islands but a relatively smaller part of the Azorean endemics; classical examples are woody *Echium* spp. and *Sonchus* spp.); for some of the *taxa* within this group, their closest relatives are today found in the east Mediterranean (e.g. *Sideritis* spp. and *Erysimum* spp.)
Atlantic affinity	Hygrophilous ferns and forest plants; relatively few species in the Canary Islands; several at more northern latitudes, especially in the Azores
Saharo-Sindian affinity	Predominantly in the lower, arid zones of the Canary Islands and in Cape Verde; examples include *Forsskahlea* spp. and *Kickxia* spp.
East African affinity	Shows a present-day disjunction from Macaronesia to East Africa; a well known example is the genus *Dracaena*, with disjunct species in Macaronesia, the Red Sea area and Socotra
South African affinity	Species whose closest relatives are only found in southern Africa, like the closely related genera *Phyllis* in Macaronesia and *Galopina* in South Africa
Afromontane affinity	Occurs in Macaronesia and some African mountains; exemplified by *Canarina canariensis* (L.) Vatke in the Canary Islands and the genus *Cryptotaenia* also in Nigeria and Tanzania regions
East Asian (-Australian) affinity	Some of the Macaronesian endemics which apparently have the least possibility of long distance dispersal, e.g. the genus *Appolonias* in Macaronesia and southern India; also *Picconia* in Macaronesia and its closest relative *Notelaea* in eastern Australia
American affinity	Found in all large Macaronesian island groups; examples are the *Bystropogon* species and their closest relatives in southern America

It is clear from the evidence gathered here that a large percentage of the Macaronesian flora is composed of Mediterranean species and, among the endemics, species with Mediterranean affinity. This indicates a close relationship between the Macaronesian and the Mediterranean floras (Sunding, 1979). The affinities with the Mediterranean area (southern Europe and northern Africa) are not restricted to the present-day plant distribution, but are extended to the past flora of both areas. Many of these past connections are manifested in an important plant community of Macaronesia, the *Laurisilva* forest, whose elements are related to the Tertiary fossil flora of southern Europe. Firstly, the characteristics of that community are considered and then its origins in the context of the Tertiary climate and flora changes in Europe and northern Africa.

1.3.2 The *Laurisilva* forest

The *Laurisilva* is a forest with a great biodiversity that is mostly endemic. It is stratified in different levels, mainly composed of trees and bushes with persistent foliage, dark green leaves (often flat, glabrous and coriaceous), perpendicularly disposed to daylight and having protected buds (Jardim & Fontinha, 2000). It is also very diverse in endemic animals, including insects that play a vital role in pollination. This evergreen forest grows slowly under a subtropical and Mediterranean climate of relatively high humidity (usually above 85%) and medium temperatures, which prevail on the North side of the islands (Bramwell & Bramwell, 2001).

The most characteristic *Laurisilva* trees belong to the Lauraceae family: *Laurus azorica* (Seub.) Franco, *Ocotea foetens* (Ait.) Baill, *Persea indica* (L.) K. Spreng. and *Appolonias barbujana* (Cav.) Bornm. Hence the designation *Laurisilva*, meaning 'forest of laurel' (originated from *Laurus silva*, the Latin words for, respectively, laurel and forest).

There are many other endemic trees and herbaceous species from different families associated with this forest (Jardim & Fontinha, 2000). Cryptogamic plants are very significant for both their diversity and covering area. Pteridophytes are abundant, especially in the deep valleys which are dark and humid. Bryophytes cover a large proportion of surfaces of the soil, rocky slopes, and fill in not only trunks but also the branches of trees and bushes. Lichens and fungi are also common; for example, the Madeiran endemic fungus *Laurobasidium laurii* (Geyler) Jülich is only found on *L. azoriza* bark (Jardim & Fontinha, 2000).

The island of Madeira harbours an extremely well preserved *Laurisilva* (Figure 1.2) which also exists in a much smaller extension in the central and western Canary Islands (Figure 1.2). In the Azores, Cape Verde, Iberia and Morocco few elements of this typical forest are present (Freitas & Brehm, 2001). Although the forest remnants are well preserved, it is known that they occupied a more extensive area in the recent past (Aubreville, 1976; Bouza *et al.*, 2002; Jardim & Fontinha, 2000; Schäfer, 2003). Much

of the Macaronesian laurel forest has been destroyed in the last 500 years, since Man colonised the various archipelagos. The extant forest is fragmented by agri-silviculture and overgrazing. On Madeira, it extends across 15000ha of land, roughly 16% of the island's surface. On the Canary Islands, the *Laurisilva* occupies only about 6000ha (2%), which is still more than the remaining forest found on the Azores (1%), on the islands of São Miguel and Terceira. In 1999, the remaining *Laurisilva* in Madeira was declared as a World Heritage Patrimony under the protection of UNESCO (Jardim & Fontinha, 2000). The Laurisilva of the Canary Islands is now also protected under the same classification (Bramwell & Bramwell, 2001).

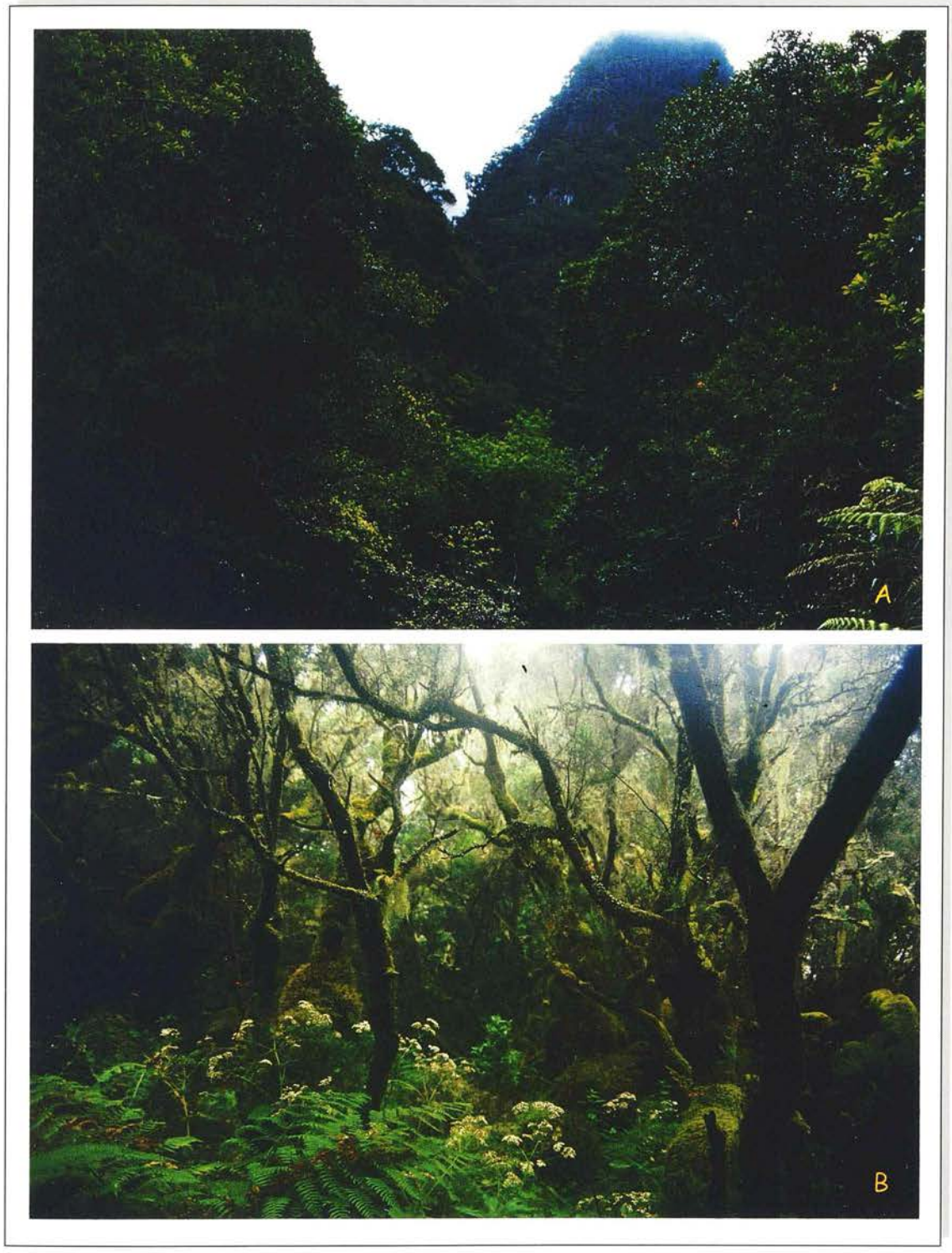


Figure 1.2 The *Laurisilva* forest. A. Madeira (from Jardim & Fontinha, 2000); B. El Hierro (Canary Islands).

1.4 Origin and evolution of the Macaronesian flora

An understanding of the climatic and biogeographic changes in Europe and Africa during the Tertiary and Quaternary periods was established as early as the 19th century, based on a multidisciplinary approach of palaeobotany and palaeoclimatology. However, the key points, particularly those of biogeographic relevance, have been reviewed and summarised more recently for the Eurasian and African regions by several authors, e.g. Axelrod (1975), Quézel (1978), Cronk (1992), Mai (1995), Hewitt (1999) and Tiffney & Manchester (2001). The following summary is largely based on these publications.

1.4.1 Climatic and phytogeographic changes in Europe and northern Africa during the Quaternary periods

The Atlantic Ocean began to open with the rotation of Africa and S America away from N America in the early Jurassic (180My BP). About 80My later, around the mid-Cretaceous, Eurasia and North America were still rather close to each other. Forests of tropical and subtropical requirements covered low-middle latitudes of N America and Eurasia during the late Cretaceous (c. 70-65My BP). They largely represent ancient oak-laurel-palm forests that lived under monsoonal climate with precipitation concentrated in the warm season and with warm summers and frostless winters.

The Tertiary started with moderate climates in the Palaeocene (65My BP). At this time, both Eurasia and the Sahara were still dominated by tropical humid forest. The warming progressed until the early Eocene (c. 50My BP), which was the warmest time in the Tertiary. At this stage, forests of thermophilic, frost-intolerant species were most abundant up to latitudes of 50-60°. From this time on, the climate cooled in a series of steps, the most pronounced of which approximately coincides with the Eocene-Oligocene border (c. 33.7My BP). The Eocene-Oligocene cooling corresponded with a major shift in vegetation, with thermophilic and evergreen *taxa* in Europe becoming less

common and communities of more seasonal climates becoming widespread. At this time, the Sahara was essentially occupied by a temperate rain forest and the Mediterranean region by a subtropical woodland savannah.

A gradual warming occurred into the early Miocene (c. 20My BP), leading to a resurgence of subtropical *taxa* and the expansion of evergreen and thermophilic *taxa* in Europe and N Africa. This warming interval was short and much cooler than the warm period in the Eocene, and consequently these floras lacked several paratropical elements that had been present at that stage. Starting in the middle Miocene (c.15My BP), cooling returned and progressed stepwise to the Pleistocene ice ages (c. 2My-10000y BP), marking the beginning of the Quaternary period. This led to sequential modernisation of the flora of Europe, which involved the spread of deciduous trees and herbs, and the loss of thermophilic, evergreen elements. A phase of aridity in N Africa that started at the Miocene was accentuated during the Pliocene with the establishment of desert climate in the Sahara and a lowering of the temperature in the Mediterranean region. The result was the emergence, both in S Europe and N Africa, of the typical sclerophyllous Mediterranean vegetation.

During the Pleistocene, the cyclic drastic cooling periods are directly associated with the glaciations in Europe and the pluvial phases in Africa. Although the impact of these pluvial phases permitted the exchange of species between the Mediterranean and tropical regions of Africa, those elements did not reach Europe. During these relatively cold and wet periods, typically Mediterranean species were pushed to the coastlines of S Europe and N Africa. In these areas, the Iberian Peninsula and the Macaronesian islands served as the main refuges. A more detailed account of plant colonisation of the Macaronesian islands will be given next, considering their time of emergence and position in the Atlantic Ocean.

1.4.2 Plant colonisation of the Macaronesian region

The Macaronesian islands began emerging in the Atlantic Ocean at the start of the Miocene (c. 20My BP). The first islands to appear were the eastern Canary Islands and Cape Verde, followed by the Madeira archipelago and then the Azores and western Canary Islands. During this period (c. 20-2My BP), the flora of southern Europe and northern Africa maintained a subtropical environment and forests of evergreen and thermophilic *taxa* were the predominant vegetation. Being the closest continental areas to the newly emerged Macaronesian islands, S Europe and N Africa provided the majority of species that initially colonised those islands (Bramwell & Richardson, 1973; Sunding, 1979). This type of vegetation was similar to the present-day *Laurisilva*.

The Tertiary flora of S Europe and N Africa has been studied in great detail. In a number of localities in Spain, France, Italy, Austria, Hungary and Georgia, plant fossils have been found resembling extant *taxa* from Macaronesia (Table 1.4), especially from Madeira and the Canary Islands (Ferguson, 1974; Sunding, 1979; Rodrigo Pérez, 1992; Bramwell & Bramwell, 2001). Fossils of typical Macaronesian plants like *Laurus*, *Persea*, *Ocotea*, *Appolonias*, *Dracaena* and several ferns date from the Miocene and Pliocene and may be up to 20My old (Bramwell & Bramwell, 2001). There is also fossil evidence that a forest similar to the *Laurisilva* existed on the W coast of Morocco (Aubreville, 1976).

At the end of the Tertiary (c. 2My BP), the *Laurisilva* forest was established in the Macaronesian islands. During the Pleistocene cooling periods, the *Laurisilva*-like forest regions disappeared from Europe and N Africa, leaving only remains in the archipelagos of the Azores, Madeira and the Canary Islands (Sunding, 1979; Bramwell & Bramwell, 2001). These islands were highly suitable for humid subtropical vegetation to survive because Quaternary climatic fluctuations were buffered by the oceanic climate. However, some species now with a mainly Macaronesian distribution, can still be found in a few places in the Iberian Peninsula [W Portugal and S Spain (Bramwell & Richardson, 1973)], as well as on the Atlantic coast of Morocco (Quézel, 1978).

It is reasonable to assume that a greater part of the Macaronesian endemic element, irrespective of its present-day affinities to different regions, must be considered a relict flora of once wider distribution in S Europe and N Africa (Bramwell, 1976; Sunding 1979; Cronk, 1992). More recently, several DNA-based studies have revealed that some of these lineages considered relictual, especially woody *taxa*, are of more recent origin [e.g. *Argyranthemum* (Francisco-Ortega *et al.* 1996a, 1996b, 1997), *Bencomia* alliance (Helfgott *et al.*, 2000), *Echium* (Böhle *et al.*, 1996), *Pericalis* (Panero *et al.*, 1999)]. The picture that is emerging is that of a mixed flora of Tertiary relicts and *taxa* that have evolved and radiated following the colonisation of the islands. It has also been shown, using DNA sequencing, that some mainland representatives of Macaronesian *taxa* [e.g. *Convolvulus fernandesii* Pinto da Silva & Teles (Carine *et al.*, 2004)], are not relictual but in fact back-colonisations from Macaronesia.

Table 1.4 Fossil *taxa* found in European Tertiary deposits whose modern relatives have a mainly Macaronesian distribution (from Sunding, 1979; Rodrigo Pérez, 1992).

Fossil *taxon*	**Corresponding Macaronesian *taxon***
Adiantum reniforme	*Adiantum reniforme*
Appolonias aquensis	*Appolonias barbujana*; *A. ceballosi*
Asplenium hemionitis	***Asplenium hemionitis***
Clethra berendtii	*Clethra arborea*
Dracaena brogniartii	*Dracaena draco*
Dracaena narbonensis	*Dracaena draco*
Ilex canariensis	*Ilex canariensis*
Laurus abchasica	*Laurus azorica*
Maytenus canariensis	*Maytenus canariensis*
Myrsine spp.	*Myrsine* spp.
Ocotea heerii	*Ocotea foetens*
Persea indica	*Persea indica*
Phoenix sp.	*Phoenix canariensis*
Picconia excelsa	*Picconia excelsa*
Pinus consimilis	*Pinus canariensis*
Pinus resurgens	*Pinus canariensis*
Smilax targionii	*Smilax canariensis*
Viburnum pseudotinus	*Viburnum tinus* subsp. *rigidum*
Woodwardia radicans	*Woodwardia radicans*

Most of the Tertiary species that colonised Macaronesia are now extinct in Europe and Africa. Also, several of those species have undergone extensive diversification, many with different representatives in different islands (Bramwell & Bramwell, 2001). However, a few examples exist of species that still maintain remnant populations in Europe and/or Africa and have somehow remained unchanged. One such example is the fern *Asplenium hemionitis*, and for those reasons it has been chosen for a detailed investigation of population genetics and phylogeography.

1.5 Distribution and biology of *Asplenium hemionitis* L.

1.5.1 Geographical distribution

A. hemionitis (Figure 1.3) is one of the few true Macaronesian elements in the European fern flora. It is distributed on all islands of the Azores, Madeira and the Canary Islands (Hansen & Sunding, 1993); although reaching Cape Verde, its occurrence is rare and can only be found in four of the islands (Lobin *et al.*, 1998). In the mainland it has isolated populations in W Portugal (area of Sintra and Mafra, Nogueira & Ormonde, 1986) and NW Africa (Morocco and Algeria, Salvo & Ouyahya, 1999). The geographical distribution is illustrated in Figure 1.4.

1.5.2. Reproductive biology

Being a pteridophyte, *A. hemionitis* disperses and reproduces sexually from haploid spores. These are produced in high quantities (in the order of five million per year; see page 113, 5.1.2) by the adult plant (sporophyte) on the abaxial side of the fronds, which all have the potential to become fertile. The spores are of one type (homosporous), very small and light, and so easily transported by wind currents. The outer wall of the spores is ornamented (Figure 1.3); this potentially assists in wind transport and holding to a suitable substrate, thus helping germination (Jermy, 1984; Ormonde, 1996). Once germination starts, a haploid gametophyte develops (Figure 1.3) and produces the male and female gametangia. Following fertilisation, an independent diploid sporophyte starts developing and will become the adult plant. The gametophyte and the sporophyte are then two independent generations of a homosporous pteridophyte life cycle. Since a single gametophyte can produce both male and female gametes, there is potential for intra-gametophytic selfing, an extreme form of inbreeding that produces a completely homozygous individual. The other breeding systems present in pteridophytes (Klekowski, 1979) are outcrossing (cross-fertilisation between different individual sporophytes) and inter-gametophytic selfing (cross-fertilisation between gametophytes

derived from the same individual sporophyte). Aspects of the pteridophyte life cycle and breeding systems are further discussed in Chapter 4.

5.1.3 Morphological variation

In Europe and Macaronesia, the fern *A. hemionitis* can be readily distinguished from the other members of the genus by the particular shape of its fronds (Figure 1.3). The fronds are entire and 5-lobed, much like ivy leaves (*Hedera* sp.); this is very different from the dissected shapes characteristic of many European *Asplenium*. *A. hemionitis* can show some morphological polymorphism, and thus the species has received numerous (10) intraspecific taxonomic treatments (Bolle, 1864; Schneider, 1892; Bormüller, 1904; Kunkel, 1965, 1966, 1978; Benl, 1969). All varieties described are, to some extent, slight modifications of the classic 5-lobed frond shape (Figure 1.3). They have all been described based on single individuals found in the Canary Islands. However, different frond shapes can be found sometimes in the same plant (pers. obs.) and at different development stages. Also, in more favourable ecological conditions (abundant shade and humidity), some individuals can grow to be much larger than the average, showing extra-long lobes and stipes. It has not been possible to identify divergent lines of *A. hemionitis* which can be associated with different islands or regions at either the species or intraspecific levels.

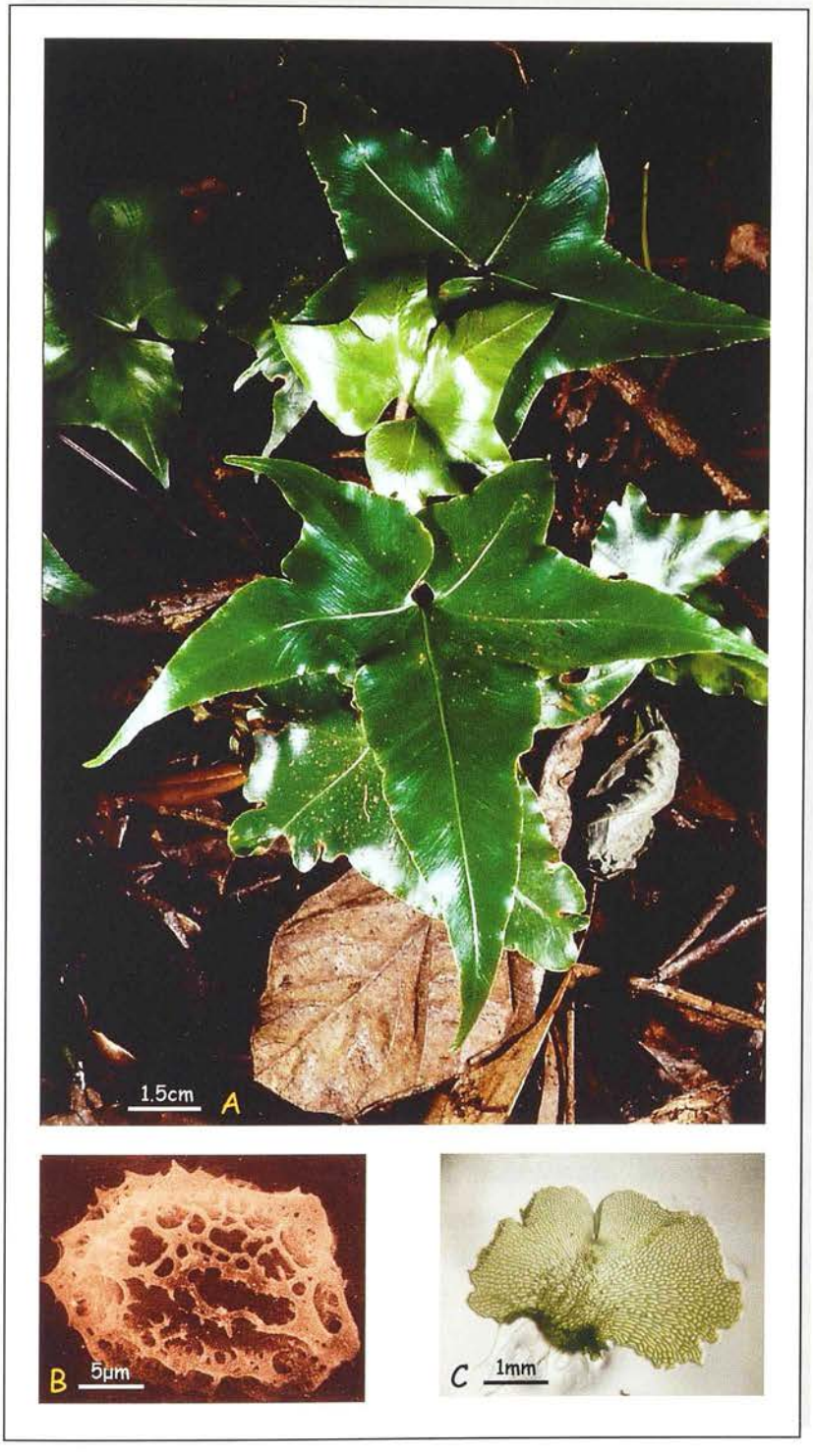


Figure 1.3 Life stages of the fern *Asplenium hemionitis*. A. Sporophyte (photograph by J. Vogel); B. Spore (from Ormonde, 1996); C. Gametophyte.

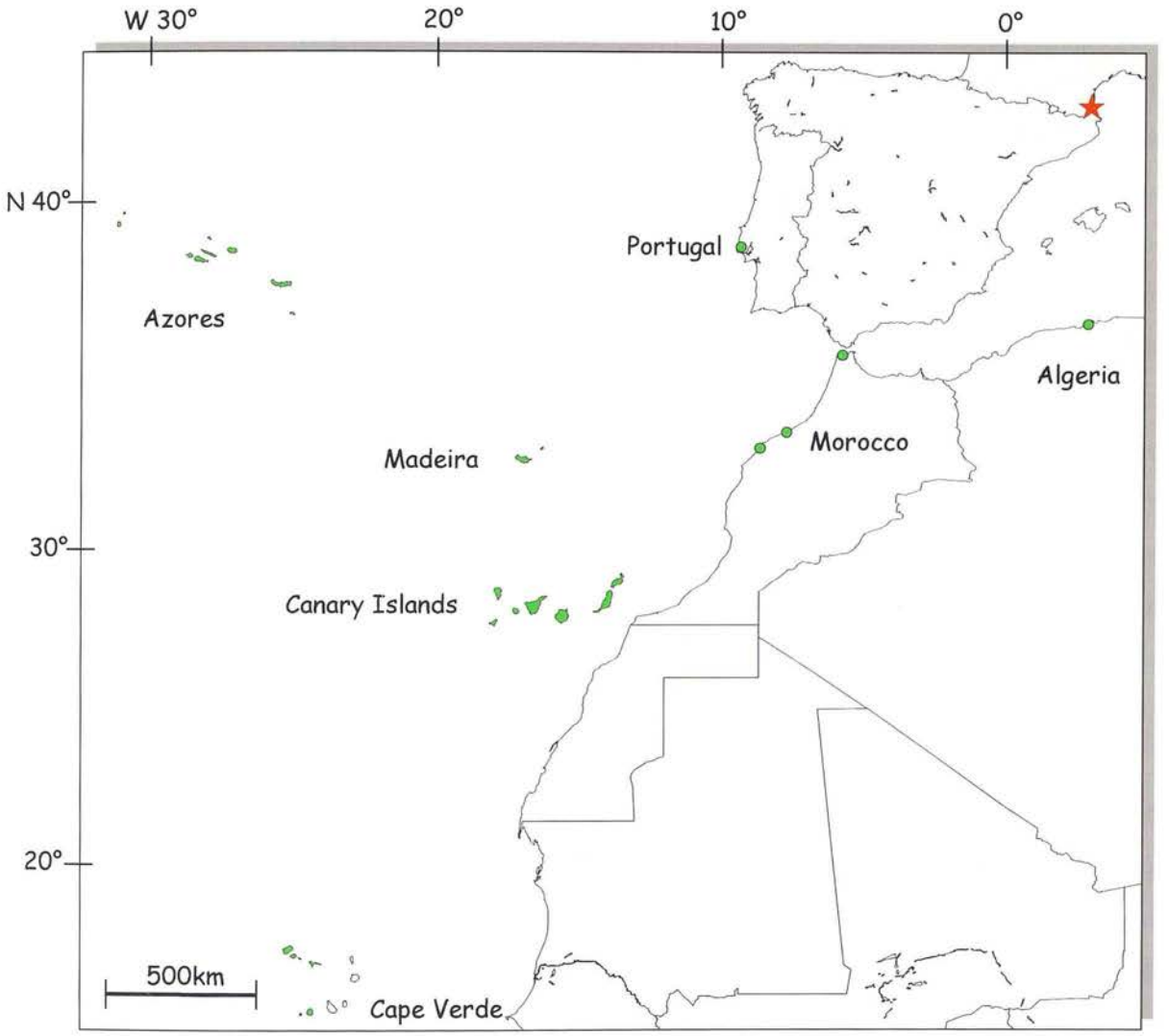


Figure 1.4 Geographical distribution of *A. hemionitis*. Islands where it is present are shaded in green; presence on the mainland is represented by green dots; red star represents the location in S France where a potential fossil has been recorded.

1.5.4 Ecology

A. hemionitis is a perennial, calcifuge fern that is found growing in a variety of habitats, from forest grounds and cave walls to road verges and man-made walls (pers. obs.). However, in the different distribution areas, it is likely to occupy specific habitats. In the Canary Islands it is mostly found on natural habitats, as an element of the *Laurisilva*, where it grows on forest grounds and slopes, walls of deep gullies and caves. It can also be occasionally found on banks of abandoned fields and road verges. In W Portugal (Sintra), it is virtually confined to man-made walls and old steps constructed in steep areas, with the occasional single individual establishing on forest ground or the base of rocks. In the Azores and Madeira, where both natural and artificial habitats are available, it is more frequent in semi-natural communities. In the Azores it is frequently found on road verges and damp banks and in Madeira it is found on 'levada' (water collection canals) walls, damp banks and slopes in abandoned vineyards and fields. In both archipelagos, it can also be found occasionally on man-made walls when there is enough soil available to become established.

There is less information available on the ecology of *A. hemionitis* in Cape Verde and Morocco and it is based on scarce literature and herbarium material. In Cape Verde, where humid and shady forests equivalent to *Laurisilva* are lacking, *A. hemionitis* is reported to grow only in rock crevices and walls of small caves (Lobin *et al.*, 1998). In Morocco, this fern grows on rocky walls of humid coastal forests (Salvo & Ouyahya, 1999). More precise information on the ecology of this fern in Morocco is based on the collections made by Fred Rumsey (NHM) on the NW coast (Tanger). Two populations were found growing on a shady and humid semi-natural woodland near springs and abandoned cultivation terraces. It seems clear that *A. hemionitis* favours the insular humid and warm oceanic conditions that prevail in Macaronesia. The isolated continental sites where *A. hemionitis* grows mimic this niche extremely well and are thus regarded as 'mainland islands'.

In Sintra, Azores and Madeira, field surveys have failed to find *A. hemionitis* growing in natural habitats equivalent to those of the Canary Islands. However, earlier collections

of *A. hemionitis* in these three regions (Durães, 2000) suggest that it is present there in natural sites. It is clear that Man's influence and habitat destruction has significantly restricted this species distribution and ecological range.

1.5.5 Phylogenetic and taxonomic relationships

A. hemionitis is a diploid fern (2n=72, Lovis *in* Reichstein, 1981; Queirós & Ormonde, 1987) for which hybridisation events, although frequent within European *Asplenium* (Reichstein, 1981), have never been recorded. This suggests a certain degree of genetic distance from other European and Macaronesian *Asplenium*. In fact, several recently published phylogenies of this genus, using both nuclear (n) and chloroplast (cp) DNA data (Pintér *et al.*, 2002; Van den Heede *et al.*, 2003; Schneider *et al.*, 2004), have not been able to establish a fully resolved position of *A. hemionitis* within the genus.

1.5.6 Fossil record

Fossils of *A. hemionitis* are referred by Sunding (1979) and Rodrigo Pérez (1992) as present in European Tertiary deposits (Table 1.4), however, precise locations are not given. Following an exhaustive survey of literature on the European fossil record, one single reference to *A. hemionitis* has been found (Saporta, 1865). The author reports the presence of two fossil fronds in Miocene (23.8-5.3My) deposits from S France (Figure 1.4). The fossil nomenclature and description are somewhat confounding and doubtful, but it is most likely to *A. hemionitis* that the author is referring. Moreover, the presence of fossils of other Macaronesian endemic species, e.g. *Laurus* sp., *Persea* sp. and *Dracaena* sp. in the same area, further supports this conclusion.

1.5.7 Conservation status

A. hemionitis has a wide distribution range but its populations are localised and small. For this reason, its conservation status has received considerable attention. In the Canary Islands and Madeira, although present in all islands of both archipelagos, it is considered rare (Ormonde, 1990; 1991). In Cape Verde, the number of islands and sites where is has been recorded is decreasing, and so it has been considered critically endangered (Lobin *et al.*, 1998). In Morocco it is also rare (Fennane & Tattou, 1998; Salvo & Ouyahya, 1999) and in Portugal it is considered vulnerable and endangered (Dray, 1985; Directive Fauna-Flora-Habitat 92/43/CEE, Annexe 4; Berne Convention, 1979).

1.6. Using molecular markers to investigate contemporary and historical processes in plant populations

Molecular markers (proteins or sequences of DNA which show quantifiable variation) are employed to assign unambiguous genotypes to individual organisms (Lowe *et al.*, 2004). Variation detected by molecular markers can be usefully termed 'genetic marker variation' to differentiate it from the quantitative variation detected by analysis of phenotypic measurements, which may be more informative about the adaptive potential of populations (Newton *et al.*, 1999). Because molecular markers represent differences in the genetic information within or between populations of organisms, they can be used to describe their genetic structure. The genetic structure of plant populations is the distribution of genotypes in space and time and results from both present processes and past history (Hewitt, 1999). In these terms, molecular marker variation describes the behaviour and evolution of genes in populations and allows us to identify the genetic consequences of different events.

The distribution of genetic diversity within a species is the result of evolutionary forces like mutation, genetic drift and migration (Lowe *et al.*, 2004). Mutations are changes in the DNA sequence (point mutations and insertion/deletion), therefore creating allelic variation and increasing genetic diversity. The amount and patterns of genetic diversity are ultimately determined by drift, gene flow (migration) and natural selection. Drift is the random fluctuations in allele frequency from generation to generation and causes populations to become different from one another just by chance. The smaller the population, the more powerful drift is. This is counteracted by migration, which maintains connectivity between individuals and populations. Over time, effective gene flow will cause newly created allelic variation to spread to other populations (Slatkin, 1987). Conversely, if gene flow between populations is somehow restricted, populations will differentiate and become genetically distant. Most biodiversity research uses neutral molecular markers (not under any selective advantage) and so selection is assumed to be negligible.

There is general agreement that transitory reductions in effective population size lead to a reduction of genetic diversity (Leberg, 1992; Widmer & Lexer, 2001). For example, dispersal into a new habitat can lead to the foundation of new populations. Such populations typically consist of a single or few individuals and, therefore, represent a small subsample of the genetic diversity of the source population. This process is called a founder event. Reductions in population size can also occur due to habitat reduction, either induced by climatic and geological changes or by Man. In the case of size recovery, the genetic diversity of those populations is, as in founder events, a subsample of the original diversity. This process is called a genetic bottleneck.

The breeding system of a species also influences the partitioning of genetic diversity within and between populations. The general trend is for outcrossing angiosperms with wind or animal dispersed propagules to maintain most of their genetic diversity within rather than among populations (Hamrick & Godt, 1990; Soltis & Soltis, 1990a, c; Hamrick *et al.*, 1992). In contrast, for inbreeding angiosperms with similar dispersal capacities, the genetic diversity is mostly partitioned among populations.

Therefore, by examining the levels and distribution of genetic markers in plant populations, it is possible to make inferences about contemporary processes like gene flow or the breeding system, and the history of the populations throughout the geographic range. For the past decades, isozymes have been the markers of choice for population genetics in natural populations of plants (Weeden & Wendel, 1990; Soltis & Soltis, 1990a, c; Hamrick & Godt, 1990; Hamrick *et al.*, 1992). In the recent years, the development of DNA-based tools, most notably the introduction of the PCR (polymerase chain reaction) technique, has allowed the direct analysis of DNA variation. These markers are particularly valuable to estimate genetic diversity parameters which are used to deduce gene flow patterns and events like bottlenecks and founder effects. However, it is not possible to determine the phylogenetic relationships between alleles with most of these techniques (e.g. isozymes, RAPD, AFLP). Using such tools as DNA sequencing, it is possible to examine the geographic distribution of genealogical lineages, especially those at the intraspecific level, an approach termed phylogeography (Avise, 1998). In the context of this study, phylogeography involves

the analysis of the spatial distribution of alleles, taking into account their phylogenetic relationships. This approach has been used to provide insights into how recent natural events (e.g. glaciations) have shaped patterns of intraspecific variation within a wide range of organisms (Newton *et al.*, 1999). A classical example is the European *Quercus*, an important tree genus for which large data sets are available (Ferris *et al.*, 1998; Dumolin-Lapègue *et al.*, 1997a; Petit *et al.*, 2002). Phylogeographic analysis based on cpDNA data indicated that *Quercus* species have been isolated in southern refugia in Europe during the Pleistocene ice-ages, from where northern Europe was colonised in warming periods. This interglacial mixing of lineages did not migrate south at the onset of the next glaciation, resulting in long-term isolation of southern refugia and divergence of lineages (Ferris *et al.*, 1999). *A. hemionitis* is also a species which is thought to have survived in refugia, namely the Macaronesian islands and restricted sites on the mainland. A phylogeographic approach has, therefore, the potential to reveal historical patterns of isolation and colonisation events.

Different methods of assessing genetic variation, such as various types of molecular markers and plant genomes, have different characteristics and may provide different kinds of information. For this reason, the most commonly used genetic markers are reviewed separately in Chapter 2.

1.7 Questions addressed in this study

This section presents a series of questions which will be addressed in subsequent chapters. The first list (I) corresponds to questions concerning the breeding system of *A. hemionitis*. These will be discussed in more detail in Chapter 4, including characterisation of specific objectives. Lists II and III correspond with questions concerning the genetic diversity, population structure and phylogeography of *A. hemionitis*. These relate to Chapter 5 (isozyme analysis) and Chapter 6 (cpDNA analysis).

I. Investigation of the breeding system and fine scale genetic structure (Chapter 4)

Inbreeding causes populations to become dominated by homozygous individuals and fixed for determined alleles. Consequently, the breeding system significantly determines how genetic diversity is partitioned within and among populations. An investigation of the breeding system of *A. hemionitis* is, therefore, essential for the interpretation of genetic diversity and population structure data. Estimations of the inbreeding coefficient employing isozyme markers and breeding experiments are used to answer the following questions:

i) What is the prevailing breeding system in *A. hemionitis*?
ii) How does the breeding system vary among populations and regions?
iii) Is there a correlation between the breeding system and different ecological conditions?
iv) If inbreeding is occurring, is it mostly attributed to inter- or intra-gametophytic selfing?
v) Is inbreeding occurring as a consequence of small scale (intrapopulational) genetic structuring of genotypes?
vi) Is there restriction of gene flow within populations?
vii) How is the breeding system affecting *A. hemionitis* genetic structure?

II. Investigation of genetic diversity, population differentiation and isolation by distance (Chapter 5)

The distribution range and population size of *A. hemionitis* have likely been reduced due to habitat destruction and/or past climatic changes. A decrease of population size is normally associated with a decrease of genetic diversity. The present distribution of *A. hemionitis* over a series of different archipelagos implies a certain degree of population isolation and restriction of gene flow (movement of genes among populations). This is the major parameter preventing populations from differentiating over time and can have profound impacts on the structure of genetic diversity. Isozyme markers are here employed to examine the levels of genetic diversity, genetic differentiation and genetic distance at different geographic scales in order to answer the following questions:

i) How much diversity is found within populations, islands and regions (archipelagos and mainland)?
ii) Is the amount of genetic diversity associated with recent events, e.g. genetic bottlenecks or founder effects?
iii) How does the global genetic diversity found in *A. hemionitis* compare with that found in other pteridophytes?
iv) What are the levels of genetic differentiation across the range of *A. hemionitis*?
v) Is genetic structuring organised differently at different distribution hierarchies, i.e. within and among populations, islands and archipelagos?
vi) Is restriction of gene flow occurring principally among populations within islands, among islands within archipelagos, or among archipelagos?
vii) Is restriction of gene flow explained by an isolation by distance model?
viii) What are the genetic relationships between different islands and regions?

III. Investigation of cpDNA diversity and phylogeography (Chapter 6)

A phylogeographic approach is particularly appealing in the case of *A. hemionitis*, since its geographic range comprises several populations in archipelagos that have been

isolated for a long period of time. Using markers like DNA sequences, it is possible to detect ancient polymorphism but also new genetic variants and their phylogenetic relationships. Drawing a parallel between history and variation in molecular markers will allow us to understand the sequence of events that shaped the present populations. Sequencing of two cpDNA regions was carried out to answer the following questions:

i) How much cpDNA diversity is found within populations, islands and regions (archipelagos and mainland)?
ii) How are the various chloroplast haplotypes distributed and how much cpDNA diversity is shared by the different regions?
iii) What are the levels of cpDNA differentiation at different geographic scales, e.g. populations, islands and regions?
iv) Is cpDNA differentiation structured phylogenetically due to independent evolution of distant populations?
v) Where are the most ancestral haplotypes located?
vi) Which is the most plausible region for the origin of *A. hemionitis*? Has it originated on the mainland or has it originated in the Macaronesian islands from where the mainland was then colonised?
vii) Are the populations in Sintra (Portugal) introduced or can they be considered relictual?
viii) How have the mainland populations been affected by the Pleistocene glaciations?

CHAPTER 2

SELECTION OF EFFECTIVE MOLECULAR MARKERS TO INVESTIGATE GENETIC STRUCTURE AND DEMOGRAPHIC PROCESSES IN *Asplenium hemionitis*

2.1 Introduction

Genetic markers can be classified into five broad groups (Lowe *et al.*, 2004): morphological (e.g. flower colour, leaf shape), cytological (e.g. chromosome number and structure), biochemical (secondary products like flavonoids and terpens), protein (e.g. isozymes) and DNA (sequence polymorphism and repeat). Given that the present study is focused on population genetics and phylogeography of a single, morphologically well defined *taxon*, only protein and DNA-based markers were considered as suitable.

Molecular genetic markers are defined as proteins or sequences of DNA used to reveal genetic variation. These markers can be screened to uncover key attributes of their state or composition and thus provide information about allelic variation at a given *locus*. Over the years, advances in molecular technology have led to the introduction of a multitude of molecular markers which are variable at many different levels. These markers have been applied to many biological questions, ranging from population genetics to gene mapping, phylogenetic reconstruction, paternity testing and forensic applications (Schlötterer, 2004).

The ideal genetic marker to investigate the genetic structure and history of plant populations should be discretely polymorphic, highly heritable (i.e. show no environmental or development influences) and show simple codominant inheritance (i.e. in a diploid, both alleles at a *locus* should be visible in the heterozygous condition). Furthermore, the marker should detect silent nucleotide changes (i.e. synonymous mutations), changes in coding and non-coding portions of the genome, and evolutionary homologous changes (i.e. similar due to descent from a common ancestor) (Lowe *et al.*, 2004). The marker should ultimately be easy to score using a simple, inexpensive and reproducible technique. None of the marker systems currently used in plant population studies has all these ideal characteristics (Lowe *et al.*, 2004). Each genetic marker system is characterised by a unique combination of advantages and disadvantages. These markers differ in the amount of variability they express, modes of inheritance and

the ease and costs of their development and application. The choice of a marker system is dictated significantly by its suitability to approach a particular question. The general view is that the rise and fall in popularity of different molecular markers reflects the continuous improvement in the way in which genetic variation is assayed, that is, the latest marker systems are the most informative ones. Nevertheless, in reviewing the history of molecular markers and their pros and cons, Schlötterer (2004) argues that there are only a few conceptually different classes of marker and that recently developed high-throughput methods might not be unconditionally superior to more traditional approaches. The markers described here fall into three classes: protein-based, DNA sequence polymorphism and DNA repetition variation. Their characteristics are discussed individually in sections 2.3 and 2.4.

2.2 Plant genome types

The plant genome consists of three components, the nuclear DNA (nDNA), the chloroplast DNA (cpDNA) and the mitochondrial DNA (mtDNA). Nuclear and organelle genomes (cpDNA and mtDNA) are localised, respectively, in the cell nucleus and cytoplasm and contain different amounts of DNA that have different evolutionary characteristics (Table 2.1). Attributes like ploidy level, mode of inheritance and structural mutation rates will determine their potential as genetic markers. The population genetic behaviour of the different genomes affects the level of genetic variation maintained.

The nuclear genome, although variable in size, is the largest of the genomes available for genetic analysis. The nDNA is organised as a diploid assemblage of varying numbers of linear chromosomes that are biparentally inherited. The cpDNA is organised as a single circular molecule and each gene is normally present only once. Plant mtDNA is much larger and more complex than cpDNA. In many plants it appears to exist in circular, linear and more complex molecular arrangements (Backert *et al.*, 1997). Each cell can contain 10^2 to 10^4 copies of these organelle DNA molecules (Page

& Holmes, 1998). Organelle genomes are predominantly maternally inherited, with the exception of gymnosperms and a few angiosperms in which the cpDNA is paternally inherited. Cases of biparental inheritance of the cpDNA have also been shown (Harris and Ingram, 1991). The genus *Asplenium* shows plastid inheritance through the mother (Vogel *et al.*, 1998b).

DNA is a mixture of coding and non-coding regions that evolve at different rates. Non-coding regions diverge more than coding regions, in both nDNA (Nei, 1990) and organelle DNA (Taberlet *et al.*, 1991; Demesure *et al.*, 1995). The faster a region of DNA evolves the more likely is the occurrence of polymorphisms. In plants, mutation rates in the chloroplast genome are three to five times higher than in the mitochondrial genome, but one half to one third of that found in the nuclear genome (Wolfe *et al.*, 1987). Gene diversity increases with both population size and mutation rate and, for an equivalent population size and mutation rate, gene diversity is always lower for haploid than for diploid markers. Therefore, because of their haploid nature, lower levels of gene diversity are expected in organelle genomes than in diploid nuclear genomes for genes with the same rate of neutral mutation (Ennos *et al.*, 1999). Uniparental inheritance has consequences for the partitioning of genetic diversity of organelle markers, making these often more sensitive for detecting geographic structure than nuclear markers.

Table 2.1 Relevant characteristics of the three plant genomes for population genetic and phylogeographic analysis (adapted from Lowe *et al.*, 2004 and Ennos *et al.*, 1999)

Attribute	Nuclear genome	Chloroplast genome	Mitochondrial genome
Structure	Linear chromosomes	Circular	Circular, linear and more complex shapes
Size (kb)	$5.00X10^4 - 3.07X10^8$	71 – 214	200 – 2400
Ploidy level	At least diploid (higher plants)	Haploid	Haploid
Inheritance	Biparental	Uniparental maternal (most angiosperms) paternal (gymnosperms)	Uniparental (mostly maternal)
Substitution rate[1]	$3.5X10^{-9}$	$0.86 - 1.20X10^{-9}$	$0.36 - 0.50 X10^{-9}$
Substitution rate compared with plant mtDNA	11.4	2.4	1.0
Foreign sequences	Common	Rare	Common
Structural mutations	Common	Rare	Common
Recombination	Present	Intramolecular	Inter and intramolecular

1 Substitutions per synonymous site per year; these are mean values for the genome, as substitutions rates vary from region to region within a genome.

2.3 Protein-based markers

The first true molecular markers to be established were based on separation of enzymatic proteins, by their charge and molecular weight, when they migrate through an electrical field. This technique, designated protein electrophoresis, is coupled with visualisation of proteins via specific histochemical stains (Hunter & Markert, 1957). These proteins are classified as isozymes, and, since their first use investigating population polymorphism of *Drosophila* and humans in the 1960s, have become widely used genetic markers in both plant and animal diversity studies.

Isozymes are all the different functional forms of an enzyme, present in the same individual that are encoded by different *loci*. When these forms are encoded by different alleles of a single *locus* they are designated allozymes. Isozymes and allozymes are not degraded forms of a protein, but, respectively, the product of different *loci* or alleles. These encode slightly different enzymes, adapted to variable metabolic conditions and localised in two or more different subcellular compartments (Soltis *et al.*, 1983). Isozyme variants are attributed to nucleotide substitutions that cause changes in amino acid net charge and conformation, and, therefore, alter their electrophoretic mobility. The electrophoretic verification of such amino acid alterations provides a mean of monitoring changes in the nucleotide sequence of the respective coding gene. Following electrophoresis, enzymatic activity is detected by soaking the gel in an enzyme-specific histochemical stain (contains a substrate, necessary cofactors and a dye) which produces discrete bands at positions of enzyme activity. The electrophoretic phenotype, designated zymogram, is interpreted as a genotype.

Isozymes are codominant markers, and data are usually interpreted as *loci* and alleles based on the knowledge of enzyme structure and *taxon* ploidy level. Processes like gene duplication, polyploidy, interlocus interactions and null alleles can complicate the interpretation of the banding patterns (Weeden & Wendel, 1990). Banding patterns also become more complex by the presence of ghost bands that may result from technical aspects, e.g. freezing-defrosting cycles and gel overheating (Kephart, 1990).

Isozyme electrophoresis is a cost effective method that is technically easy and safe. Although about 100 enzyme systems have been developed (Murphy *et al.*, 1996), most studies use ca. 20-30 enzyme systems that encode 15-50 *loci*. A large set of individuals (ca. 40-50) can be analysed for 5-8 enzyme systems in each assay. A multitude of methods and studies is available to help interpretation, analysis and comparison of data. However, isozymes can express low levels of polymorphism and gene variation can be underestimated due to codon redundancy and synonymous nucleotide substitutions (Nei, 1987). Also, the selective neutrality of isozymes is still an unresolved question (Ouborg *et al.*, 1999), since there is evidence for selection at various isozyme *loci* (Kreitman & Akashi, 1995; Baker, 2000).

Allozymes function well as nuclear species-specific markers in the quantification of heterozygosity, genetic diversity, genetic differentiation and gene flow. They are still the only easily used marker system to investigate genetic diversity in polyploids. Since isozyme data cannot be ordered (i.e. the genealogical pattern of relationships among alleles cannot be inferred), DNA markers with a phylogenetic signal are normally preferred for phylogenetic studies.

2.4 DNA-based markers

Most of the existing data on plant population genetic structure comprise studies of nuclear markers, mainly isozyme variants with low mutation rates (Ennos *et al.*, 1999). In recent years there has been a great development not only of techniques that assess DNA variation directly using novel DNA markers, but also by exploitation of the ability to detect variation in the two organelle genomes of plants, the chloroplast and mitochondria genomes. Another important advantage that DNA-based markers have over isozymes is that they allow the number of mutations between different alleles to be quantified (Schlötterer, 2004). A major breakthrough for the use of DNA-based markers was the introduction of the polymerase chain reaction (PCR). This technique allows any genomic region to be amplified and analysed in many individuals without

the requirement for cloning or isolation of large amounts of ultrapure genomic DNA. The following sections review the most commonly used methods in studies of plant population genetics and phylogeography. These methods are described based on Lowe *et al.* (2004), Schlötterer (2004), Baker (2000) and Wolfe & Liston (1998).

2.4.1 DNA sequence polymorphism

2.4.1.1 *Restriction fragment length polymorphism (RFLP)*

RFLPs were the first type of DNA marker to be studied. Variation in primary DNA structure can be detected using restriction endonucleases. These enzymes recognise and cut specific nucleotide motifs in a DNA sequence and are, therefore, capable of reducing complex DNA to a population of fragments with discrete sizes. The number of bases in the restriction site (four base-cutting enzymes produce more fragments than six base-cutting enzymes) and the genome base composition determine the number of restriction sites identified in a genome. The larger the number of restriction enzymes used, the greater the probability of finding polymorphisms. Length polymorphisms can then be detected, following Southern blotting of digested DNA, using radioactively labelled DNA probes that hybridise to a single target sequence in the genome.

RFLP analysis measures DNA variation originated by length mutations (indels) or single nucleotide changes, which may cause the loss or gain of a restriction site. RFLP data can be scored considering probes to represent *loci* (alleles are defined by a specific probe-enzyme combination), in terms of presence or absence of restriction fragments, as fragment direct analysis or occurrence analysis, and the analysis of the position of restriction sites (site occurrence or site mutation).

The use of site data means that RFLP data can be ordered and interpreted phylogenetically. Furthermore, RFLP markers are codominant, making it possible to determine variation at a single *locus*. The level of polymorphism achieved with RFLPs can be high, but, ultimately, the combination of the species under study, the probes and

restriction enzymes used in the analysis will determine the amount of variation detected. The applications of RFLPs include estimations of genetic diversity, population structure and gene flow as well as hybridisation, introgression and polyploidy events. Additionally, RFLP ordered data can be valuable markers for phylogeographic and phylogenetic investigations. Disadvantages in the use of RFLPs include limited suitable hybridisation probes to detect polymorphism and the requirement of relatively large amounts of sample DNA (grams). Population genetics studies using RFLPs can be expensive and time consuming since large sample sizes are required.

A recent modification of the RFLP technique uses PCR to amplify specific DNA regions. In theory, PCR-RFLPs technique (also designated CAPS – cleaved amplified polymorphic sequences) is very simple. Firstly, primers are chosen which are known to amplify a fragment of DNA from the species under investigation. Conserved primers, such as the various mitochondrial and chloroplast primers described by Taberlet *et al.* (1991), Demesure *et al.* (1995), Dumolin-Lapégue *et al.* (1997b) and Hamilton (1999) are especially useful for PCR-RFLPs. Following amplification, the DNA fragment is digested with restriction enzymes. The products of digestion are visualised directly by gel electrophoresis. This approach increases the number of sequences that can be studied using RFLPs but reduces the information that may be detected compared with standard RFLPs, since mutations outside the relatively small, amplified region will not be found. Although potentially less informative, this approach has its advantages since only small amounts of DNA are required (nanograms) and methylation problems associated with some restriction enzymes are eliminated (PCR products are unmethylated). Also, universal primers for specific *loci* are becoming widely available.

In the absence of suitable restriction sites in the amplified DNA fragment, or when the relevant site is only cleaved by an expensive or unreliable endonuclease, polymorphisms can be detected using a particularly useful approach designated dCAPS (derived CAPS). Here a mutagenic primer is used, which when combined with the sequence of DNA amplified, either creates or removes a restriction site. Generally, the enzyme chosen has a recognition site consisting of six bp. Preferably, the sixth bp corresponds to the polymorphism and the mutation within the primer is as far away as

possible from this position (i.e. the first or second bp within the recognition site). This approach requires previous knowledge of the DNA sequence to detect polymorphisms.

2.4.1.2 *Amplified fragment length polymorphism (AFLP)*

AFLP analysis is a PCR-based technique introduced by Vos *et al.* (1995). It involves the selective amplification of an arbitrary subset of restriction fragments generated by double digestion of DNA with a frequently cutting and a rarely cutting restriction endonuclease. Fragment ends are modified by the addition of double-stranded adaptors, which provide the primer sites for subsequent PCR amplification. In the first phase of PCR amplification (preselective amplification), primers are used which are complementary to the adaptors but have an additional base pair. In the second phase (selective amplification), the preselective PCR product is used as the template for amplification with selective primers that are identical to the preselective primers except for the addition of one to three additional selective bases and either a radioactive or fluorescence label. The multiple rounds of amplification reduce the complexity of the PCR product population which can then be separated on polyacrylamide gels or an automated sequencer.

The AFLP method generates a number of bands that are determined by the number of bases in the variable part of the selective primer, genome composition and complexity. Single nucleotide changes that affect restriction sites causing the AFLP primers to mispair, thus preventing amplification; AFLP will also detect indels and base rearrangements that affect the presence or size of restriction fragments. AFLP markers are usually analysed as dominant markers and alleles detected as a band presence or absence. However, codominant AFLP markers may be detected because of small insertions or deletions in the restriction fragments (Vos *et al.*, 1995). Although allelic relations may not be immediately obvious, Wolfe & Liston (1998) state that AFLP markers derived from nuclear DNA are codominant and thus inherited in a Mendelian fashion.

AFLPs markers can be highly polymorphic, are well dispersed through the genome and do not require *a priori* sequence knowledge. Also, this technique is highly reproducible and reliable since the primers anneal perfectly to their target sequences. Product profiles can be automatically scored when separated on an automated sequencer. However, the technical requirements make it an expensive and laborious method. Other problems involve determination of product homologies and *locus* designation.

Most applications of AFLPs have been for genome mapping and breeding studies. Lately, AFLPs have been extensively used in ecological genetics to estimate gene diversity, population structure and clonality. Since these markers cannot be ordered, they are not suitable for use in phylogenetic and phylogeographic studies.

2.4.1.3 *Randomly amplified polymorphic DNA (RAPD)*

RAPD is another technique based on standard PCR methodology, with the distinction that usually only one primer is used. These primers are arbitrary decamer oligonucleotide sequences that are used to randomly amplify regions of the genome. Primer binding sites are scattered throughout the nuclear and organelle genomes, in all classes of DNA from single to multiple-copy DNA and in coding and non-coding regions. A discrete PCR product is produced when, at an appropriate annealing temperature, the single primer binds to sites on opposite strands of the genomic DNA that are within an amplifiable distance, generally less than 3000 bp. The PCR products are visualised directly by gel electrophoresis.

The presence or absence of specific PCR products is assumed to represent mutations in the primer-binding sites of the genomic DNA. These processes include nucleotide substitutions that create or abolish primer sites, deletion or inversion of either priming sites or segments between priming sites, and formation of secondary structures between priming sites. Bands on a gel are scored as present or absent, meaning that RAPDs are dominant markers. Sometimes bands can vary in intensity within and between lanes which can complicate analysis.

RAPD markers generate a large number of bands with a potential high level of polymorphism. It is a cheap, simple technique that does not require sequence information. Nevertheless, these markers can be unreliable and have been extensively criticised on technical and theoretical basis (Jones *et al.*, 1997; Harris, 1999). The reproducibility and repeatability of RAPD assays is poor mainly due to unspecific amplification of the short primers and the low annealing temperatures used. Other limitations include issues associated with product homology, genome sampling and non-independence of *loci*.

Despite the limitations, RAPDs have been frequently used to estimate gene diversity, population structure and clonality. Although RAPDs have been employed for evolutionary biology and phylogeographic studies, their use is controversial since the output data cannot be ordered. Also, when RAPDs analysis is conducted on more distantly related individuals, homology between bands can be low.

2.4.1.4 *Single nucleotide polymorphism (SNP)*

SNPs are individual point mutations in the genome. There are vast numbers of these, some of which also give rise to RFLPs, but many of which do not because the sequence in which they lie is not recognised by any restriction enzyme. Within a population, however, most nucleotide positions are invariant (Schlötterer, 2004). Therefore, *a priori* information about the presence of allelic variation at a given genomic position is required. Various strategies have been pursued to identify SNPs. One approach is a screen of expressed sequence tags to detect polymorphic sites. Another approach to identify SNPs throughout the genome is the generation of whole shotgun genome sequences, using a pool of individual donors for the genomic DNA to be sequenced. More efficient is the use of reduced representation shotgun sequencing, in which only a reduced fraction of the genome (e.g. separated by size fractioning after a restriction digest) can be sequenced for multiple individuals (Schlötterer, 2004).

SNPs are usually scored as biallelic. The information content of a single SNP is limited, since there is a high possibility that all members a population are homozygous for an individual SNP (Brown, 1999), or if one of the alleles occurs at a very low frequency (Schlötterer, 2004). Also, often more than two alleles can be observed for a single SNP, which violates the assumption that they are biallelic (Schlötterer, 2004).

SNPs are easy to type and new analytical approaches are being developed at present. The main drawbacks are the high costs involved in SNP isolation, low information content and substantial mutational rate heterogeneity among sites. The main applications of SNPs are fine genome mapping and inference of past demographic events, such as population expansion and gene flow.

2.4.1.5 *DNA sequencing*

Despite not being a marker in the narrow sense, DNA sequence analysis is the most direct method to detect DNA polymorphism. This approach targets specific DNA regions which are amplified by PCR and then subjected to cycle sequencing. During this process, dideoxynucleotides (nucleotide analogs that prevent further extension of the DNA strand) are incorporated into the extending DNA strand. Once incorporated, DNA synthesis is terminated. The result is a population of DNA sequences of different lengths, which can then be separated on polyacrylamide gels. In manual sequencing, DNA strands are radioactively labelled and in automated sequencing fluorescent dyes with different colours are used.

Sequencing produces fragments of DNA in which the exact order of the nucleotides is determined. When multiple samples are compared it is possible to detect addition of nucleotides, substitutions, deletions and duplications. Data are scored directly as the separate nucleotide bases, which can be read from autoradiographs in manual sequencing or as peaks on an automated DNA sequencer.

Sequence determination generates very high quality information with the highest level of resolution possible. Although this method was initially time consuming and expensive, recent advances in sequencing technology allow sequence analysis of many DNA fragments for many individuals. The description of universal primers (e.g. Taberlet *et al.*, 1991; Demesure *et al.*, 1995; Dumolin-Lapegue *et al.*, 1997b; Hamilton, 1999) has allowed sequencing and comparison of *taxa* without knowledge of their DNA sequence. Nonetheless, DNA sequencing is still significantly more expensive than the other techniques and some DNA sequences may prove very difficult to sequence due to the occurrence of secondary structures in the DNA strand. To overcome cost problems, polymorphism information can be used for RFLP or SNP analysis, and, potentially, large number of samples can be screened.

DNA sequence analysis is used in estimation of gene diversity, population structure, hybridisation, introgression and gene flow. In terms of phylogeography and phylogenetics, DNA sequence analysis provides the maximum amount of useful information since the characters are ordered.

2.4.2 DNA repeat variation

Repetitive DNA has been widely used as a source of DNA markers. A large part of the non-coding DNA in eukaryotes consists of sequence motifs tandemly repeated many times. One type of DNA in this class is termed satellite DNA and consists of motifs of two bp to 40 Kb that are repeated hundreds or thousands of times (Page & Holmes, 1998). Two similar types of repetitive DNA sequence are minisatellites and microsatellites which, in comparison with satellite DNA, are only moderately repetitive. These repetitions in DNA sequence can be produced by mutation, unequal crossing-over and DNA slippage (a process that can occur when DNA strands mispair during replication and recombination) (Page & Holmes, 1998). The most remarkable feature of minisatellites and microsatellites is their extraordinary variability.

2.4.2.1 *Minisatellites*

Minisatellite DNA consists of small to moderate size units (up to 200 bp long) that are tandemly repeated along the chromosome, reaching sizes of 50 Kb. The first step of minisatellite analysis involves digestion of genomic DNA with restriction enzymes. These enzymes must be able to cleave the flanking DNA into relatively small fragments but leave the minisatellite DNA intact. Samples are run on agarose gels and, following Southern blotting, hybridised with a minisatellite core probe. This procedure produces a barcode-like hybridisation pattern that reflects changes in the number of repeat units.

Minisatellites are extremely variable and a large number of alleles (i.e. repeat copy number) can be found within a population. The alleles at each *locus* are inherited in a simple Mendelian fashion and act as codominant characters (Carter, 2000).

Minisatellites have been successfully used in forensics and paternity testing, but mainly in animal studies. Despite their high polymorphism, minisatellites have not been widely used in population genetics and genome mapping. The complex banding patterns prevent the assignment of alleles to a given *locus* and standard population genetics analysis can not be applied (Schlötterer, 2004). The non-random distribution in the genome limits the use of minisatellites for genome mapping. Minisatellites still remain as a reliable technique for parentage analysis but technical challenges make PCR-based methods preferable for studies in population biology.

2.4.2.2 *Microsatellites or simple sequence repeats (SSRs)*

Similar to minisatellites, microsatellites are also tandemly repeated sequences (10-50 copies), but their repeat motifs are shorter (one to four nucleotide repeats). They occur very frequently, are highly polymorphic and are assumed to be randomly distributed throughout the nuclear and organelle genome (Schlötterer, 2004). The typical small repeat regions are easily amplified by PCR with primers designed to conserved regions

flanking the variable microsatellites. PCR products are separated on agarose, polyacrylamide or automated DNA sequencing gels.

The polymorphism detected in microsatellites results from length variation brought about by changes in the number of repeat units. Usually, gel bands (alleles) are regularly spaced but, exceptionally, large jumps in allele size can occur which may be due to mutations in the flanking regions (Jarne & Lagode, 1996). Microsatellites are codominant markers that can be analysed as *locus*/allele systems as each one represents a single *locus*. Alleles can differ by as little as one bp, so it is necessary to have adequate standards to check allele designations. The occurrence of PCR artefacts ('stutter' bands that result from slipped-strand misrepair) complicate the automated scoring of microsatellite alleles (Schlötterer, 2004).

The relative abundance, codominant inheritance and high polymorphism (detectable in both nuclear and organelle genomes) of microsatellites have made them the most popular genetic markers for intrapopulation studies, mapping and paternity testing. As a drawback, microsatellites are expensive to identify in each new species and primer pairs tend to be species specific. Also, mutation mechanisms of microsatellites can be complex, creating difficulties for population genetic analysis.

Microsatellite markers have been widely applied for estimation of gene diversity and population structure. The high number of alleles per *locus* makes them ideal markers for analysis of gene flow. Yet, their use in phylogenetic analysis is limited due to problems of homoplasy and recurrent generation of alleles of identical length.

A modification of microsatellites is the ISSR (inter simple sequence repeat) approach. This method is based on length variation of amplification products of genomic DNA found between microsatellites (SSRs). Assuming high frequency and random distribution of microsatellites across the genome, it is likely that two SSRs, with a common motif and oriented on opposing DNA strands, are found within an amplification distance (Wolfe & Liston, 1998). Single primers with sequences complementary to SSRs sequences will amplify a high number of polymorphic bands.

ISSR markers are highly polymorphic, selectively neutral and are well dispersed throughout the genome. They are inherited as dominant or codominant genetic markers and are interpreted as biallelic with alleles designated as band present or absent (Wolfe & Liston, 1998). These markers have been used in plant cultivar identification, genomic mapping projects, gene flow analysis and clonal diversity.

2.5 Overview and selection of markers

Compared with seed plants, relatively few studies of levels and distribution of genetic diversity within and between populations have been published in pteridophytes (Pryor *et al.*, 2001). Most of the existing data comprise studies of isozyme markers. These have been successfully used to investigate genetic variation, population structure, gene flow, breeding systems, polyploidy and issues related with conservation. Only a few studies have used DNA-based markers like cpDNA PCR-RFLPs (Ji *et al.*, 1994), RAPDs (Korpelainen, 1996; Schneller *et al.*, 1998), AFLPs (Keiper & McConchie, 2000) and microsatellites (Pryor *et al.*, 2001) to investigate genetic diversity in pteridophytes. More recently, a comprehensive study has used cpDNA sequence data to explore patterns of diversity and phylogeography of the fern *Asplenium ceterach* (Trewick *et al.*, 2002).

A preliminary study using isozymes showed high levels of genetic diversity in *A. hemionitis*. In this study, 13 enzyme systems were analysed and the 22 *loci* identified were all polymorphic, proving that isozyme electrophoresis had great potential to investigate the population genetics of *A. hemionitis*. At spatial scales of within and between populations, the distribution of highly variable markers, such as microsatellites, would give insight into the amount of gene flow and population structure (Ouborg *et al.*, 1999). However, it is not suitable to use these same markers to study populations at large geographical scales, like in the case of *A. hemionitis*. The results may reveal non-overlapping sets of alleles in different populations, indicating that these populations are completely differentiated (Ouborg *et al.*, 1999) and not much can be said about the

effect of distance on the amount of differentiation. In fact, differentiation at such spatial scales is determined almost completely by genetic drift and mutational divergence. At these spatial scales, it is also important to determine the relatedness and phylogeny of populations rather than merely the level of gene flow (which is probably very low if it exists at all). Less variable markers, like slowly evolving isozymes and cpDNA, are then more suited.

In order to determine phylogeographic patterns and phylogenetic relationships within *A. hemionitis*, a molecular marker with a phylogenetic signal, that is, ordered rather than unordered, is required. Markers like RAPDs, AFLPs and ISSRs are generally dominant and it is very difficult to determine specific variation at a single *locus* and, therefore, the phylogenetic relationships between alleles (Newton *et al.*, 1999). It is also not possible to determine the phylogenetic relationships between alleles with isozymes and microsatellites (Newton *et al.*, 1999). Intraspecific phylogenies can only be constructed from intraspecifically variable markers like DNA sequences or RFLP data (Lowe *et al.*, 2004).

Organelle markers, because of their unusual inheritance patterns and modes of evolution, have the potential to provide novel insights into the migration history, dispersal and evolution of plant *taxa* (Ennos *et al.*, 1999). Traditionally, variation in organelle genomes has been detected as RFLPs, usually now used in a modified form as PCR-RFLP analysis. Chloroplast DNA PCR-RFLPs trials on *A. hemionitis* revealed no variation, so cpDNA sequencing remained as the best potential method to determine intraspecific variation and reveal phylogeographic patterns. Ideally, data from both nuclear and organelle markers should be combined to give clearer insights on the population genetics and demographic history of the species under study (Ennos *et al.*, 1999). The ITS (internal transcribed spacer) region of the nuclear ribosomal DNA, frequently used in phylogenetic studies, proved very difficult to amplify using PCR. Additionally, trials to amplify the nuclear gene glyceraldehyde 3-phosphate dehydrogenase (Wall, 2002) have also failed. The alternative was to sequence at least two regions of the cpDNA, which, when combined, could show a better phylogenetic resolution.

CHAPTER 3

MATERIALS AND METHODS

3.1 Experimental investigation of the breeding system

A set of preliminary breeding experiments was conducted to complement and facilitate inferences on the breeding system made from isozyme data. The breeding experiments were designed to test the capacity of *A. hemionitis* for self-fertilisation and outcrossing.

3.1.1 Collection of spores

Fronds bearing ripe sporangia were collected in the field into labelled greaseproof paper packets and left drying at room temperature. Pressed fronds from collections made prior to this study (populations M1 and M2, Table 3.2) served as an additional source of spores. This material was provided by the Natural History Museum, London, and was also placed in greaseproof paper envelopes.

3.1.2 Germination of spores and growth of gametophytes

3.1.2.1 *Preparation of gel medium*

A gel medium (pH 7.0) was prepared with Phytagel and each of the stock nutrient solutions following Dyer (1979, Appendix 1). The solution was heated in a microwave oven for the minimum time for the Phytagel to completely dissolve, with frequent swirling. The mixture was cooled to approximately 60°C and Nystatin fungicide (10^7 units l^{-1}) was added. The solution was then poured onto 5cm diameter Petri dishes and left to set.

3.1.2.2 *Spore sowing*

The spores were tapped to one end of the envelope and most of the sporangial debris was discarded. The individuals randomly selected as source of spores are indicated in Chapter 4. Spores were transferred to small labelled tubes and the ends were covered with four layers of lens cleaning tissue. The tubes were held over a prepared Phytagel plate and tapped sharply to release the spores through the layers of tissue. Gel plates were checked for a uniform spore density (c. 2 gametophytes per 1.5cm^2) under a stereo microscope (40x) and sealed with Parafilm. The plates were transferred to a growth chamber at 23°C and on a 16 hour/8 hour light/dark cycle.

3.1.3. Growth of sporophytes

Four different experimental treatments were set up to study the ability of *A. hemionitis* for: i) intra-gametophytic selfing; ii) inter-gametophytic selfing; iii) outcrossing within populations and iv) outcrossing between populations. The respective arrangement of gametophytes in cells is given in table 3.1.

A few weeks after sowing, protonemal gametophytes were isolated as required (Table 3.1) by transfer to cells of 4 x 6 cell plates. Each cell (total volume 5ml) was filled with 2.5ml of fresh Phytagel medium. Each plate was covered with its lid, sealed with Parafilm and placed in the growth chamber (same conditions as before). After an initial two week period, the plates were watered regularly with tap water to promote fertilisation.

Table 3.1 Experiments designed to investigate the breeding system of *A. hemionitis*.

Experiment	**Procedure: combinations of gametophytes in each cell**
i) Intra-gametophytic selfing	single gametophytes
ii) Inter-gametophytic selfing	pairs of gametophytes from the same individual
iii) Outcrossing within populations	pairs of gametophytes from different individuals of the same population
iv) Outcrossing between populations	pairs of gametophytes from different populations

3.2 Isozyme electrophoresis

3.2.1 Plant material and sampling strategy

Plant material for isozyme electrophoresis involved frond samples of *A. hemionitis* obtained from collections at the sites of distribution. This technique requires fresh samples for enzyme extraction.

A. hemionitis collections were made during various field trips to the main areas of the species distribution: Portugal, Morocco, Azores, Madeira and the Canary Islands (Figures 3.1 to 3.10). Since only a small amount of material per individual plant (ca. $1cm^2$) is needed for enzyme extractions, only one frond per plant was collected. *A. hemionitis* populations are typically small in size, with 4-30 individuals (sporadically ≥30); therefore all individuals per population were sampled. Fronds from each population were placed in separate sealed and labelled plastics bags to maintain humidity. Each site/population was given a number and additional letters were used to denote subpopulations. In total, 2073 individual plants representing 105 populations (average size of 20 individuals) were analysed for isozyme variation (Table 3.2). *A. hemionitis* fronds keep fresh for several days. In the case of long field trips, the collections were posted to the laboratory. The samples were kept at 4°C until enzymes were extracted (maximum of one week).

Table 3.2 Sites sampled for *A. hemionitis* and samples used for isozyme electrophoresis. **n** is the number of individuals sampled at a given site; * information unavailable.

Main area	Site/Island	Code	n	Habitat	Alt (m)	Longitude (N)	Latitude (W)
Portugal	Sintra	S1	6	Man-made wall	220	09.23.07	38.47.49
	Sintra	S2	11	Man-made wall	220	09.23.09	38.47.49
	Sintra	S3	22	Man-made wall	210	09.23.12	38.47.49
	Sintra	S4	19	Man-made wall	210	09.23.11	38.47.50
	Sintra	S5	17	Man-made wall	200	09.23.13	38.47.50
	Sintra	S6	5	Man-made wall	200	09.23.15	38.47.51
	Sintra	S7	7	Man-made wall	200	09.23.16	38.47.51
	Sintra	S8	9	Man-made wall	370	09.23.30	38.47.14
	Sintra	S9	29	Man-made wall	420	09.23.18	38.47.31
	Sintra	S10	17	Man-made wall	420	09.23.18	38.47.33
	Sintra	S11	15	Man-made wall	415	09.23.18	38.47.34
	Sintra	S12	17	Man-made wall	410	09.23.19	38.47.36
	Sintra	S13	65	base of rocks and steps	470	09.23.18	38.47.30
	Sintra	S14	116	Man-made wall	420	09.23.18	38.47.37
Azores	Flores	F1	10	roadside	*	*	*
	Flores	F2	23	roadside	*	*	*
	Flores	F3	16	roadside	100	31.12.32	39.30.44
	Flores	F4	11	roadside	25	31.07.38	39.27.23
	Flores	F5	14	roadside	130	31.15.00	39.26.23
	Flores	F6	9	roadside	*	*	*
	Corvo	C1	13	road verge	260	31.05.11	39.41.16
	Corvo	C2	10	road verge	360	31.05.25	39.26.23
	Pico	P1	30	road verge	*	*	*
	Pico	P2	15	road verge	230	28.15.34	38.29.45
	Pico	P3	4	road verge	230	28.15.30	38.29.45
	Faial	FA1	34	road verge	147	28.37.19	38.33.13
	Faial	FA2	14	road verge	223	28.38.25	38.36.31
	Faial	FA3	11	road verge	223	28.38.29	38.36.31
	Faial	FA4	28	road verge	113	28.40.30	38.27.50
	Faial	FA5	10	road verge	113	28.40.25	38.27.50
	Faial	FA6	30	road verge	114	28.41.49	38.38.13
	Faial	FA7	26	road verge	276	28.45.27	38.36.06
	Faial	FA8	14	road verge	276	28.45.30	38.36.06
	Faial	FA9	27	road verge	193	28.37.23	38.33.11
	Faial	FA10	17	Man-made wall	193	28.37.20	38.33.11
	Faial	FA11	30	road verge	262	28.43.45	38.37.34
	Faial	FA12	24	road verge	262	28.43.47	38.37.34
	Faial	FA13	5	road verge	295	28.43.22	38.37.30
	São Miguel	SM1	26	roadside	175	25.17.46	37.50.57
	São Miguel	SM2	10	rocky bank	100	25.16.43	37.51.09
	São Miguel	SM3	9	roadside	100	25.16.44	37.51.09
	São Miguel	SM4	30	road verge	253	25.21.19	37.49.51
	São Miguel	SM5	24	road verge	230	25.20.09	37.50.41
	São Miguel	SM6	15	road verge	260	25.15.46	37.50.53
	São Miguel	SM7	21	road verge	257	25.13.36	37.51.02
	São Miguel	SM8	8	road verge	260	25.13.00	37.51.59
	São Miguel	SM9	29	road verge	172	25.08.26	37.49.36
	São Miguel	SM10	21	Man-made wall	295	25.08.39	37.48.20
	São Miguel	SM11	11	road verge	266	25.08.43	37.49.30
	São Miguel	SM12	10	road verge	266	25.08.45	37.49.30
	São Miguel	SM13	26	road verge	265	25.09.18	37.47.32

Table 3.2 Continued.

Main area	Site/Island	Code	n	Habitat	Alt (m)	Longitude (N)	Latitude (W)
Azores	São Miguel	SM14	7	road verge	41	25.14.58	37.44.59
	São Miguel	SM15	20	road verge	319	25.41.40	37.50.34
Madeira	Madeira	M1	5	path side	475	17.10.40	32.51.30
	Madeira	M2	8	*	200	17.07.45	32.49.52
	Madeira	M3	67	Man-made wall	950	16.55.25	32.46.40
	Madeira	M4	10	garden	850	16.53.02	32.44.20
	Madeira	M5	18	base of Man-made wall	100	17.06.32	32.49.18
	Madeira	M6	12	base of rocks in vineyard	150	17.08.23	32.50.44
	Madeira	M7	13	rock face	400	17.08.02	32.49.49
	Madeira	M8	14	Man-made walls and base	400	17.08.02	32.49.46
	Madeira	M9	18	rocky slope	400	17.08.02	32.49.43
	Madeira	M10	17	base of Man-made wall	400	17.08.02	32.49.36
	Madeira	M11	27	base of Man-made wall	400	17.08.02	32.49.38
	Madeira	M12	5	rocky slope	400	17.08.02	32.49.30
	Madeira	M13	13	Man-made wall	450	17.08.02	32.49.25
	Madeira	M14	4	rocky slope	400	17.06.30	32.49.19
C. Islands	Tenerife	T1	31	forest slope	818	16.10.57	28.33.10
	Tenerife	T2	6	forest slope	875	16.10.34	28.33.36
	Tenerife	T3	12	disturbed forest slope	1012	16.16.45	28.31.59
	Tenerife	T4	25	forest slope	1005	16.16.43	28.32.02
	Tenerife	T5	9	forest slope	990	16.16.37	28.32.10
	Tenerife	T6	10	road verge	805	16.17.02	28.31.41
	Tenerife	T7	17	forest slope	610	16.12.05	28.32.57
	Tenerife	T8	49	gentle forest slope	787	16.11.02	28.33.06
	Tenerife	T9	11	road verge	782	16.17.06	28.31.27
	Tenerife	T10	35	forest steep slope	800	16.10.54	28.33.20
	Tenerife	T11	25	sides of deep gully	744	16.50.27	28.19.20
	Gran Canaria	GC1	30	forest bank	707	15.35.31	28.05.25
	Gran Canaria	GC2	38	Man-made wall	*	15.35.27	28.04.42
	Gran Canaria	GC3	30	disturbed forest slope	*	15.35.27	28.04.41
	Gran Canaria	GC4	6	forest bank	821	15.35.27	28.04.46
	Gran Canaria	GC5	14	road side slope	1009	15.34.48	28.02.46
	El Hierro	EH1	10	*	*	*	*
	El Hierro	EH2	5	walls of small cave	1000	17.56.27	27.47.42
	El Hierro	EH3	35	walls of small cave	1003	17.56.31	27.47.50
	El Hierro	EH4	7	walls of small cave	1003	17.56.31	27.47.51
	El Hierro	EH5	8	forest path slope	605	17.59.22	27.45.04
	El Hierro	EH6	5	forest slope	620	17.59.18	27.45.10
	El Hierro	EH7	27	forest slope	675	17.59.16	27.45.10
	El Hierro	EH8	7	forest slope	675	17.59.17	27.45.08
	El Hierro	EH9	23	forest ground	753	18.04.27	27.44.05
	La Palma	LP1	26	base of rocks and walls	460	17.47.55	27.44.04
	La Palma	LP2	12	humid bank	430	17.47.56	28.47.19
	La Palma	LP3	36	forest and path slope	713	17.48.00	28.47.19
	La Palma	LP4	9	steep forest slope	715	17.48.07	28.47.11
	La Palma	LP5	30	road verge and walls	1550	17.55.07	28.46.58
	La Palma	LP6	10	small gully	800	17.49.34	28.48.44
	La Palma	LP7	13	forest path slopes	450	17.46.33	28.45.55
	La Gomera	LG1	29	*	*	*	*
	La Gomera	LG2	19	forest ground	780	17.14.05	28.09.11
	La Gomera	LG3	30	Man-made walls and	831	17.17.44	28.09.46
	La Gomera	LG4	11	rocks	736	17.12.28	28.07.34
Morocco	Tanger	TA1	39	walls of steep gully	*	05.48.50	35.46.51
	Tanger	TA2	11	disturbed forest bank	*	05.48.49	35.46.51
				disturbed forest bank			

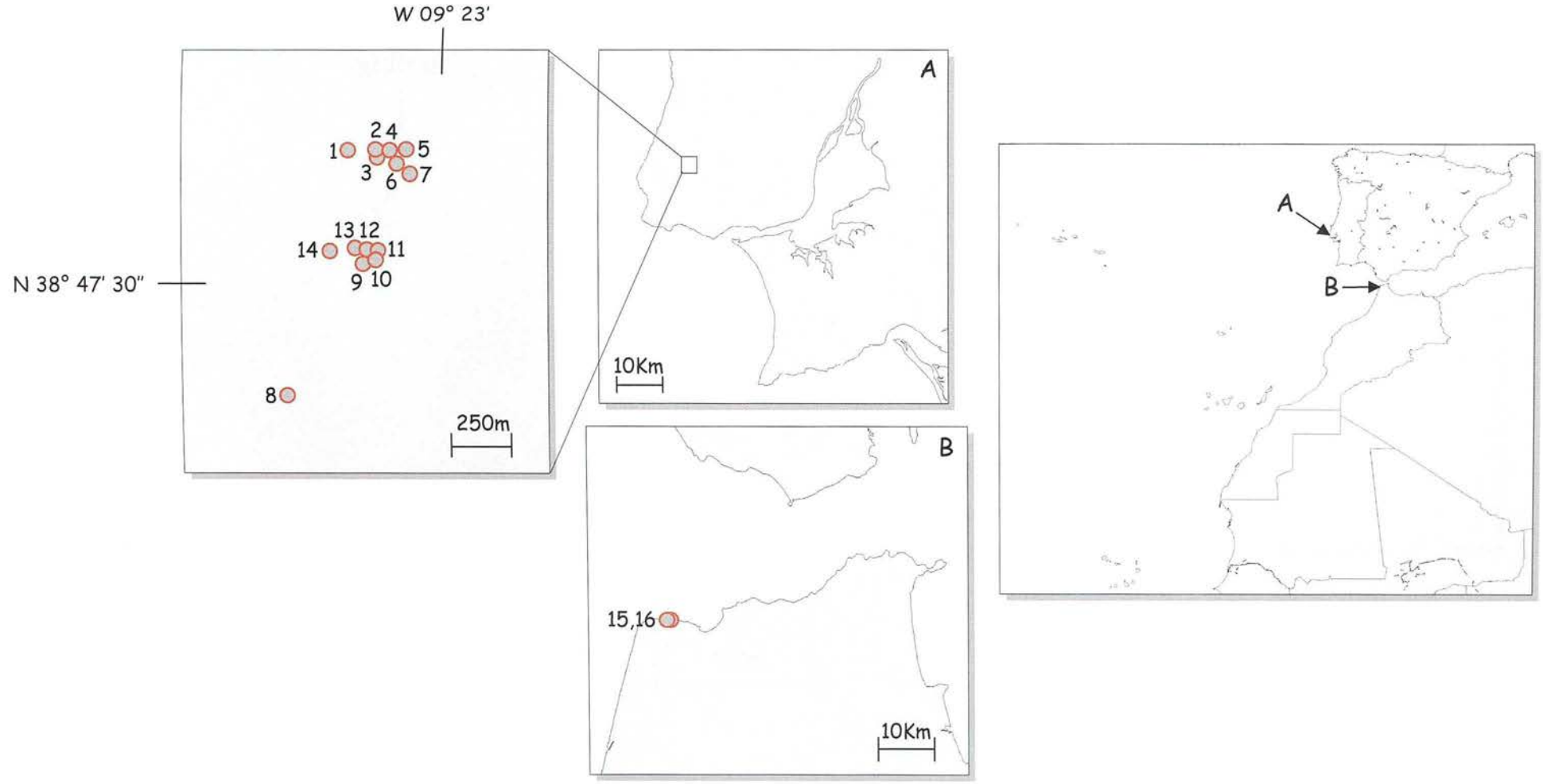


Figure 3.1 Distribution of populations of *A. hemionitis* sampled in A) Sintra (Portugal) and B) Tanger (Morocco, N 35°46'51'', W 05°48'47''). Each dot represents a population; 1-14 = S1-S14; 15-16 = TA1-TA2 (references according to Table 3.2).

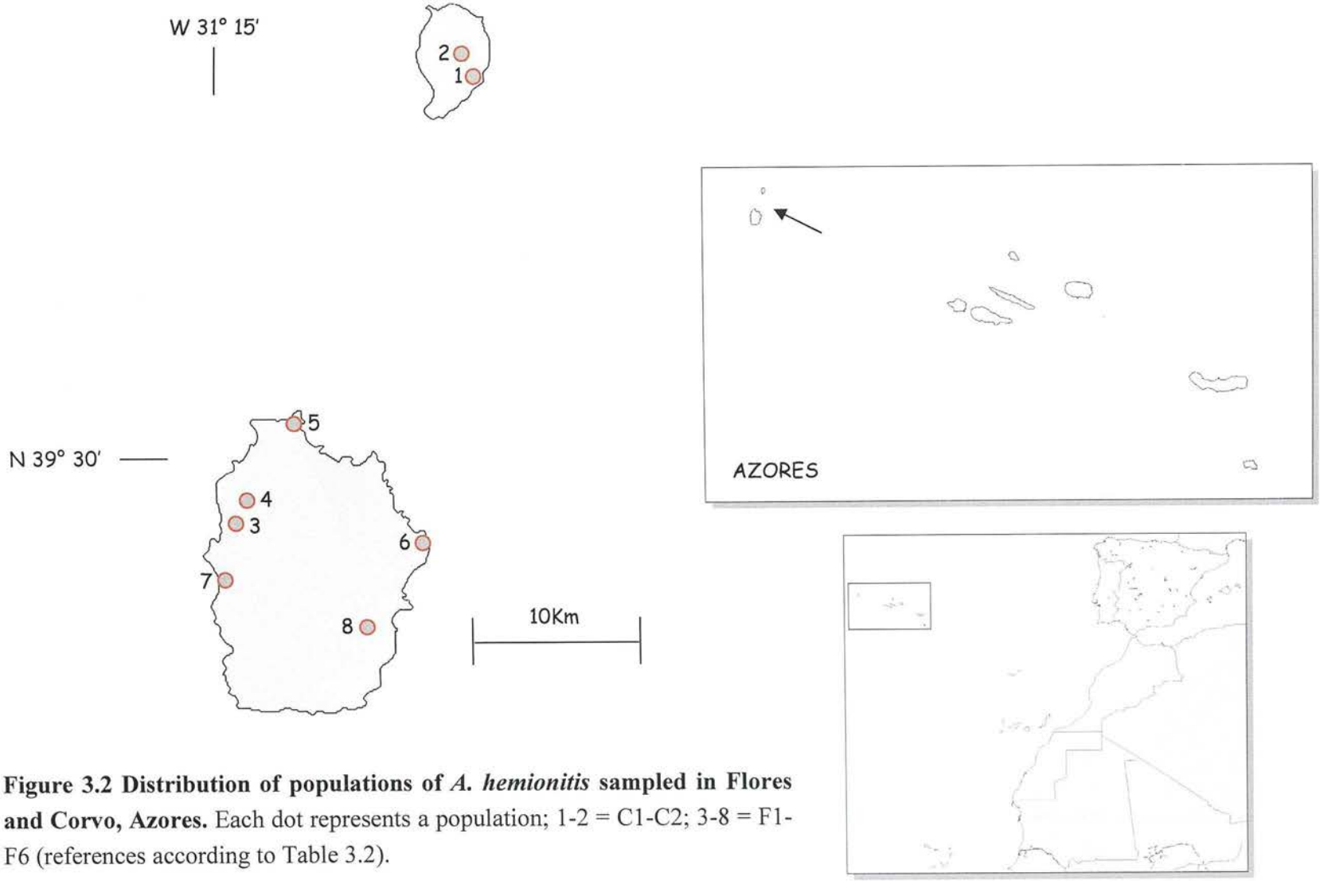


Figure 3.2 Distribution of populations of *A. hemionitis* sampled in Flores and Corvo, Azores. Each dot represents a population; 1-2 = C1-C2; 3-8 = F1-F6 (references according to Table 3.2).

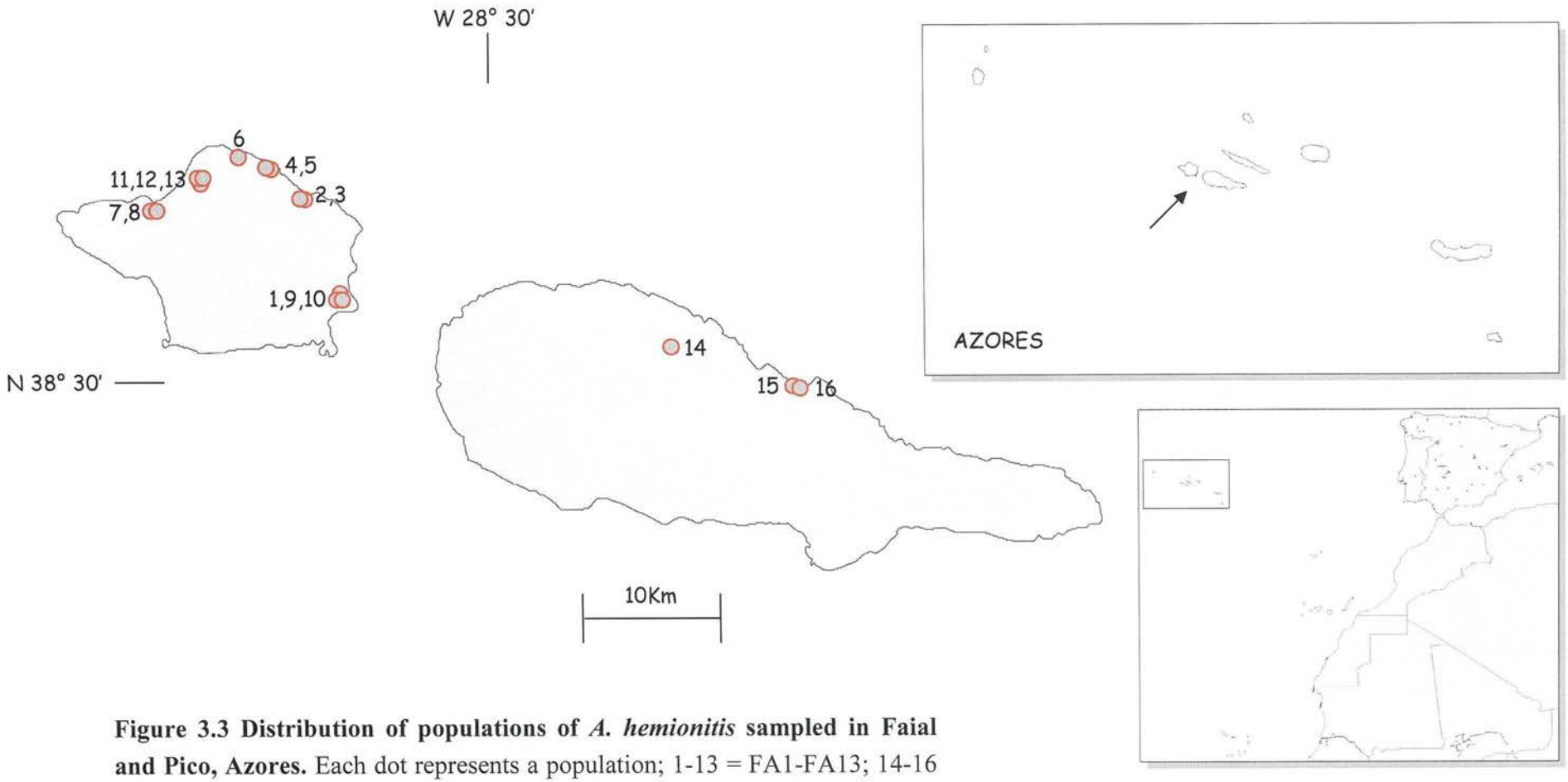


Figure 3.3 Distribution of populations of *A. hemionitis* sampled in Faial and Pico, Azores. Each dot represents a population; 1-13 = FA1-FA13; 14-16 = P1-P3 (references according to Table 3.2).

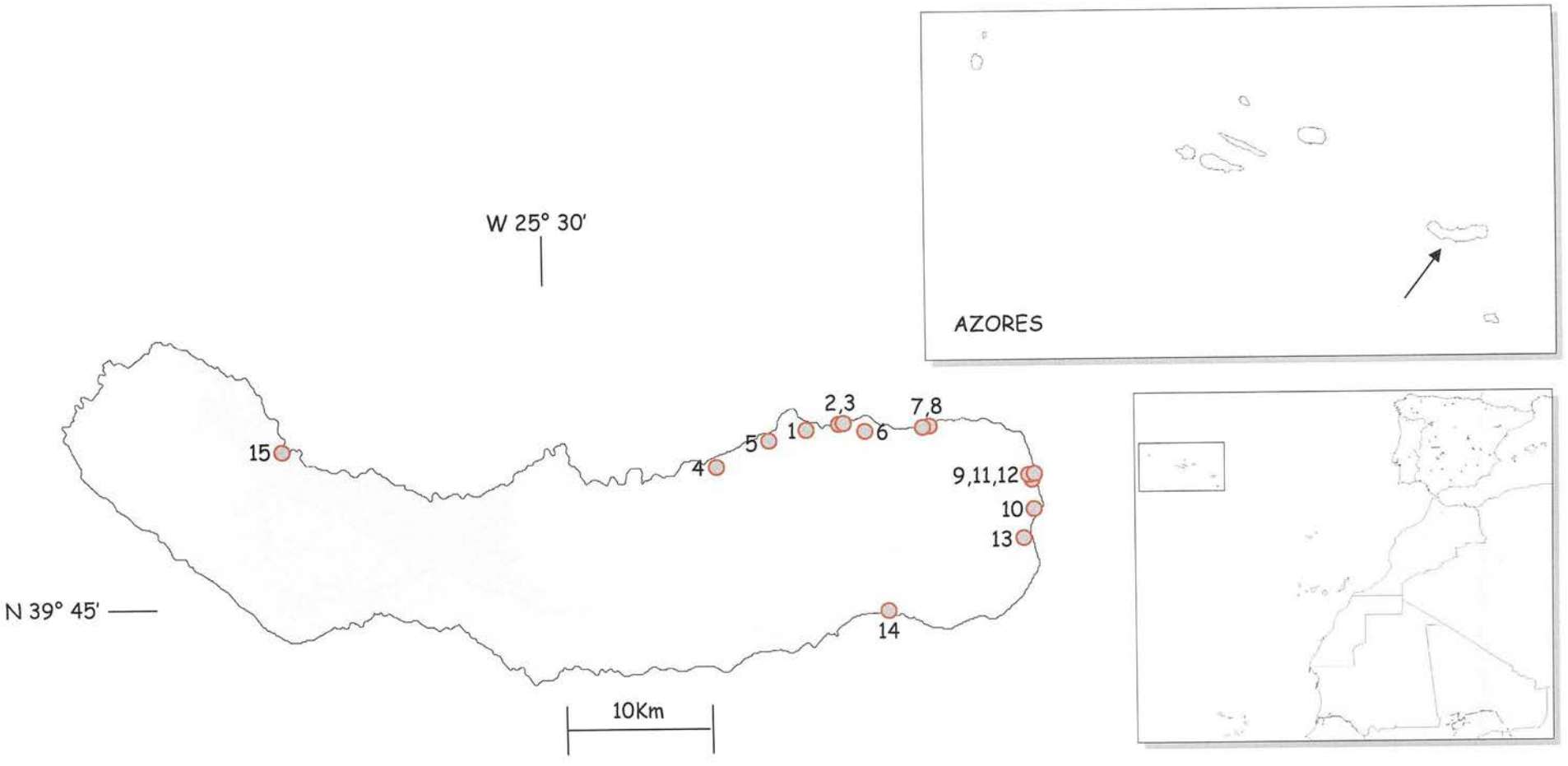


Figure 3.4 Distribution of populations of *A. hemionitis* sampled in São Miguel, Azores. Each dot represents a population; 1-15 = SM1-SM15 (references according to Table 3.2).

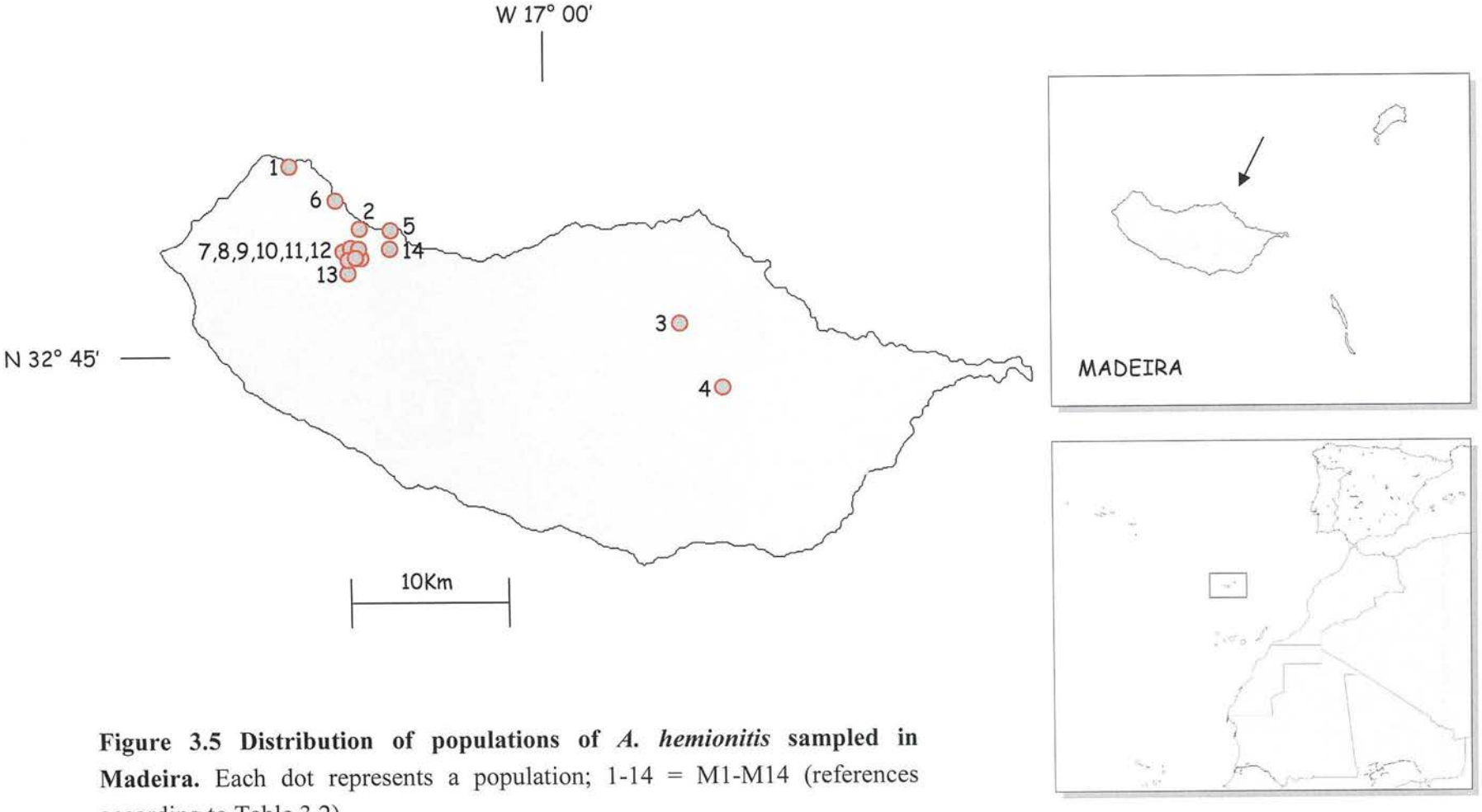


Figure 3.5 Distribution of populations of *A. hemionitis* sampled in Madeira. Each dot represents a population; 1-14 = M1-M14 (references according to Table 3.2).

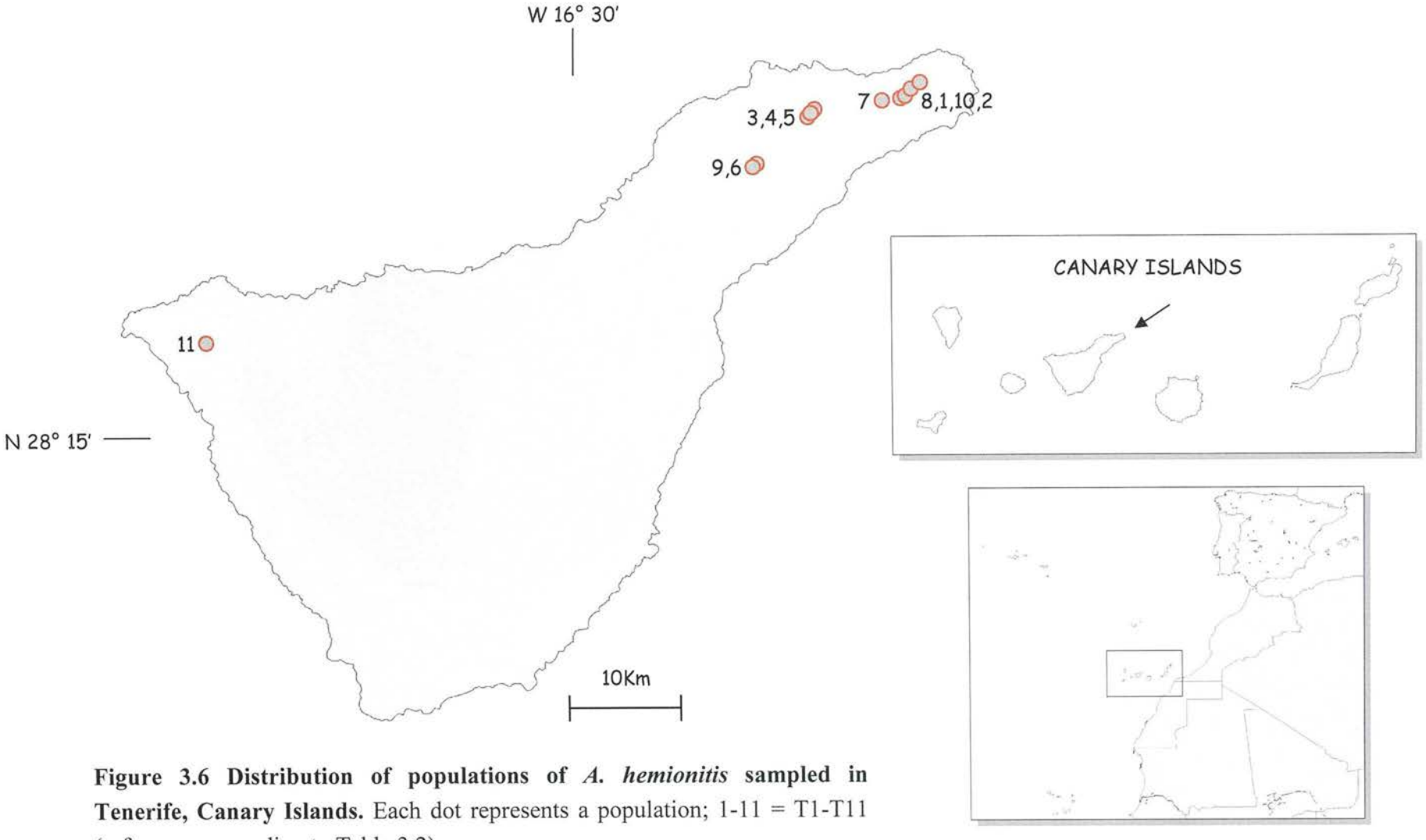


Figure 3.6 Distribution of populations of *A. hemionitis* sampled in Tenerife, Canary Islands. Each dot represents a population; 1-11 = T1-T11 (references according to Table 3.2).

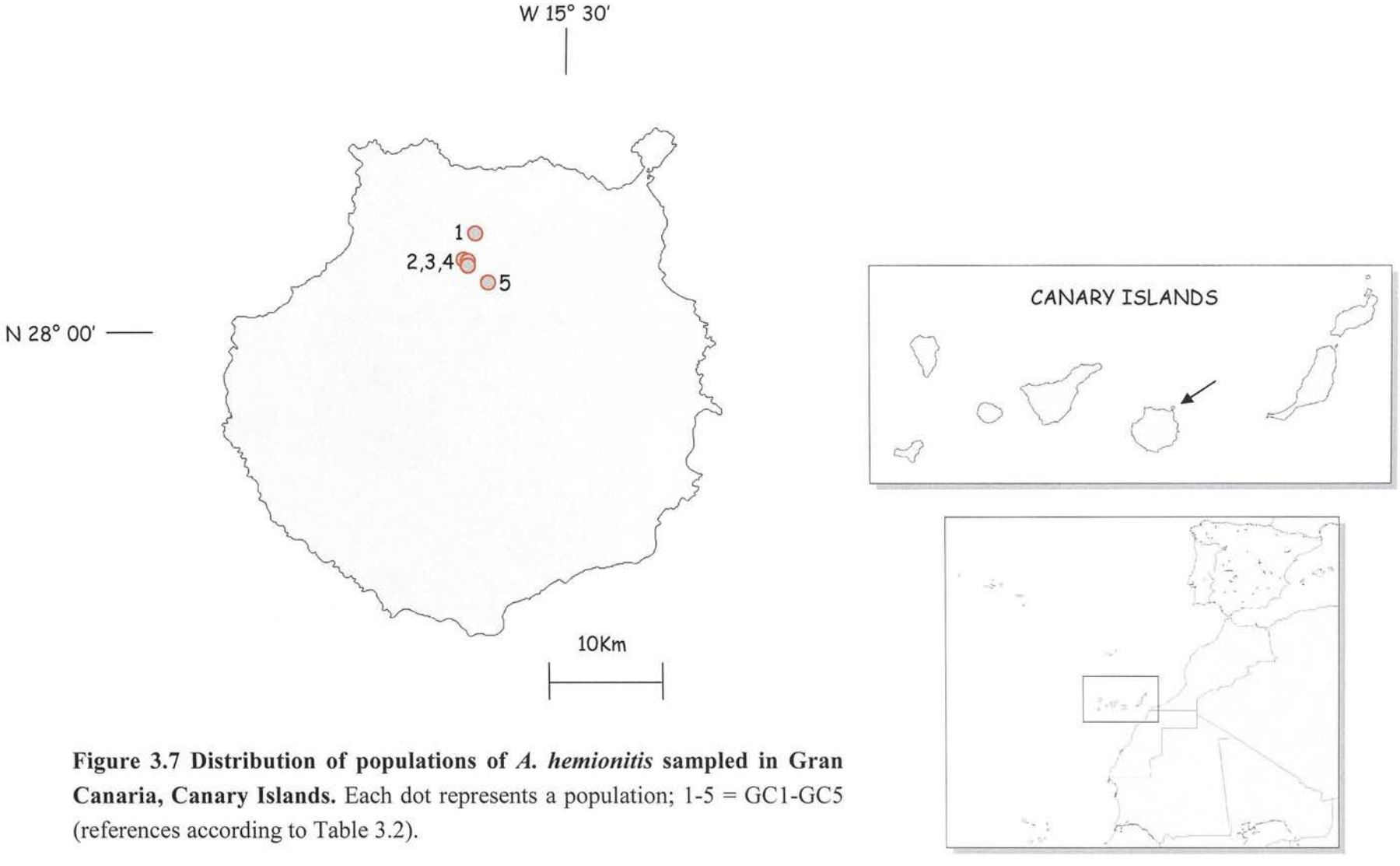


Figure 3.7 Distribution of populations of *A. hemionitis* sampled in Gran Canaria, Canary Islands. Each dot represents a population; 1-5 = GC1-GC5 (references according to Table 3.2).

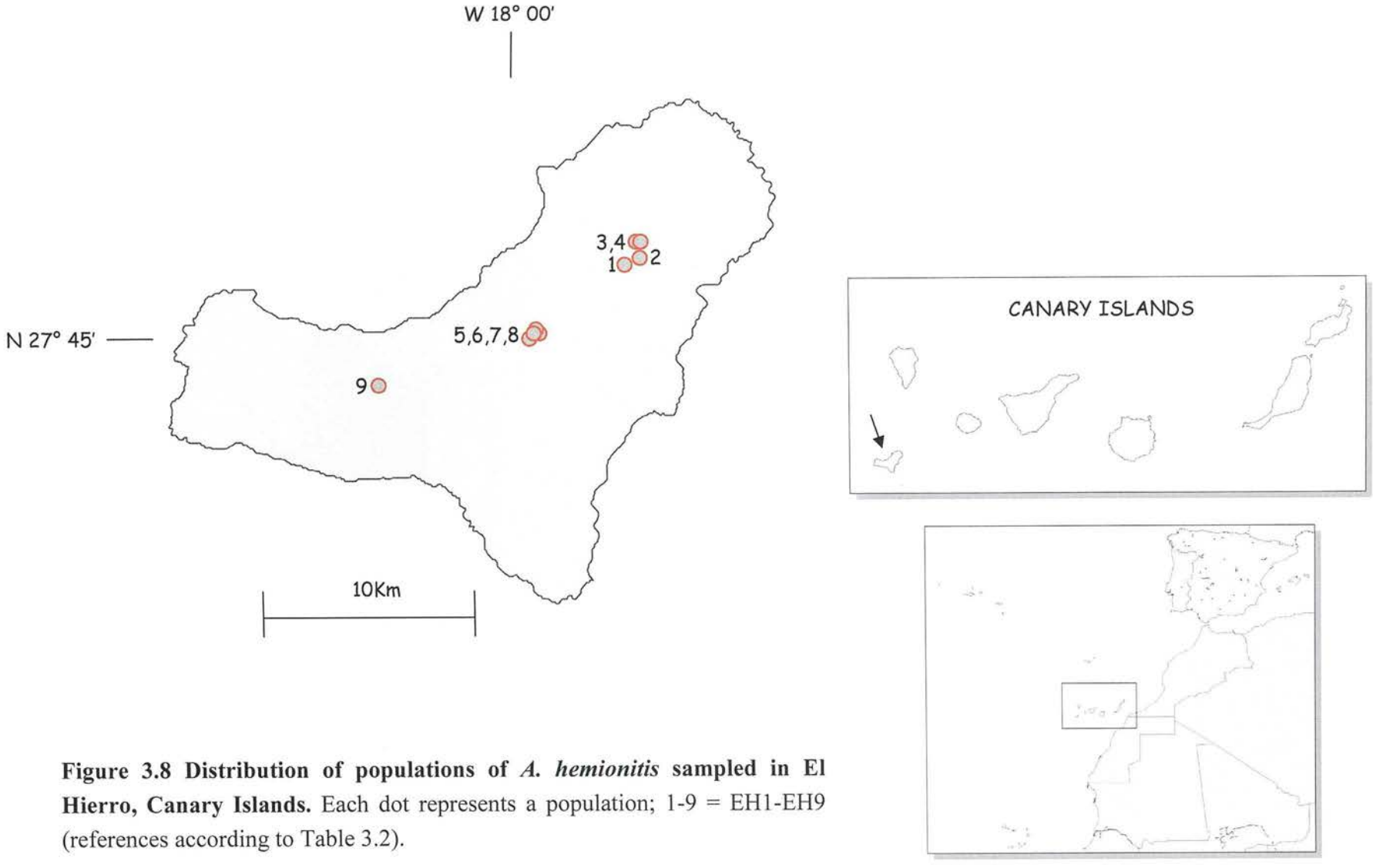


Figure 3.8 Distribution of populations of *A. hemionitis* sampled in El Hierro, Canary Islands. Each dot represents a population; 1-9 = EH1-EH9 (references according to Table 3.2).

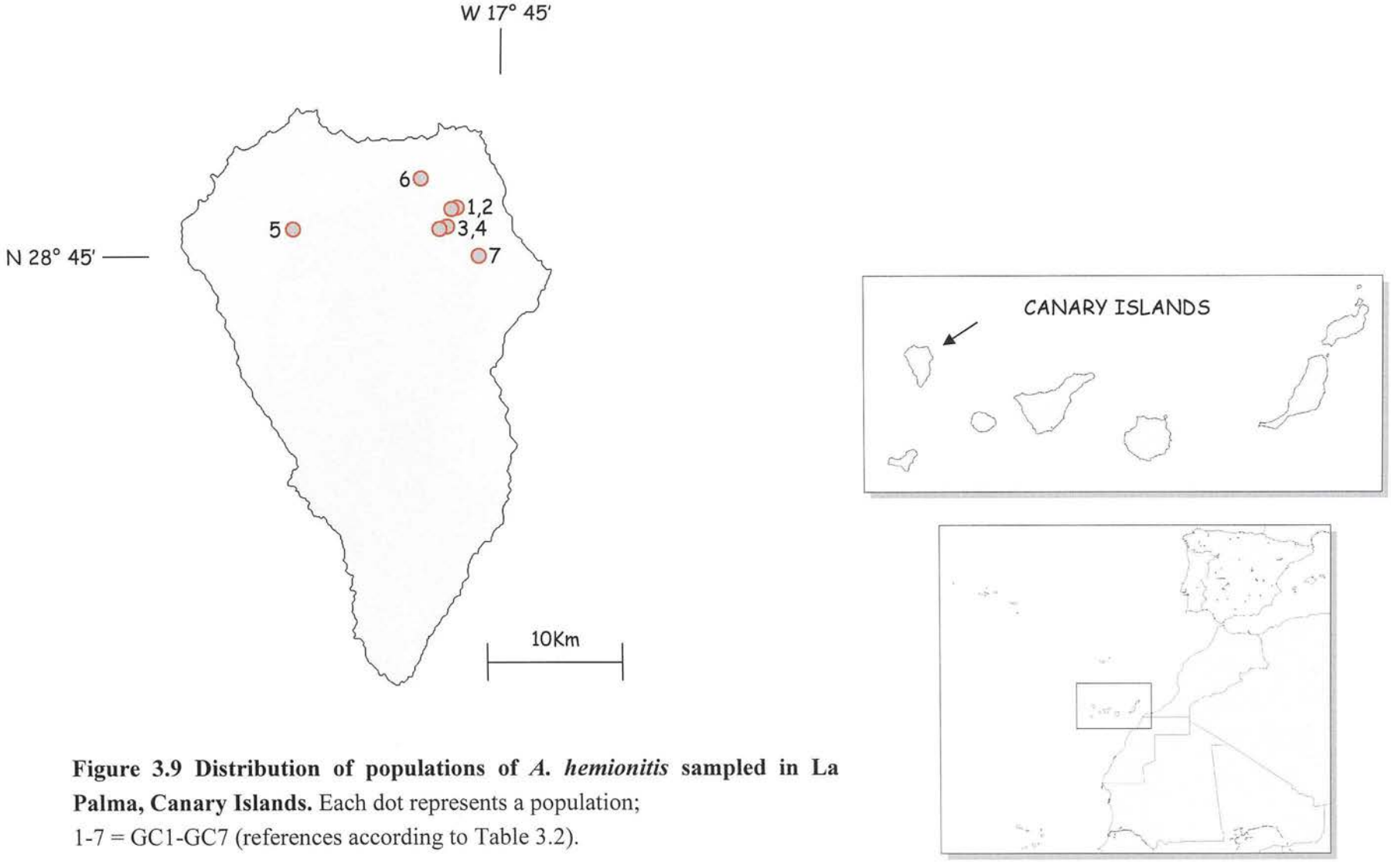


Figure 3.9 Distribution of populations of *A. hemionitis* sampled in La Palma, Canary Islands. Each dot represents a population;
1-7 = GC1-GC7 (references according to Table 3.2).

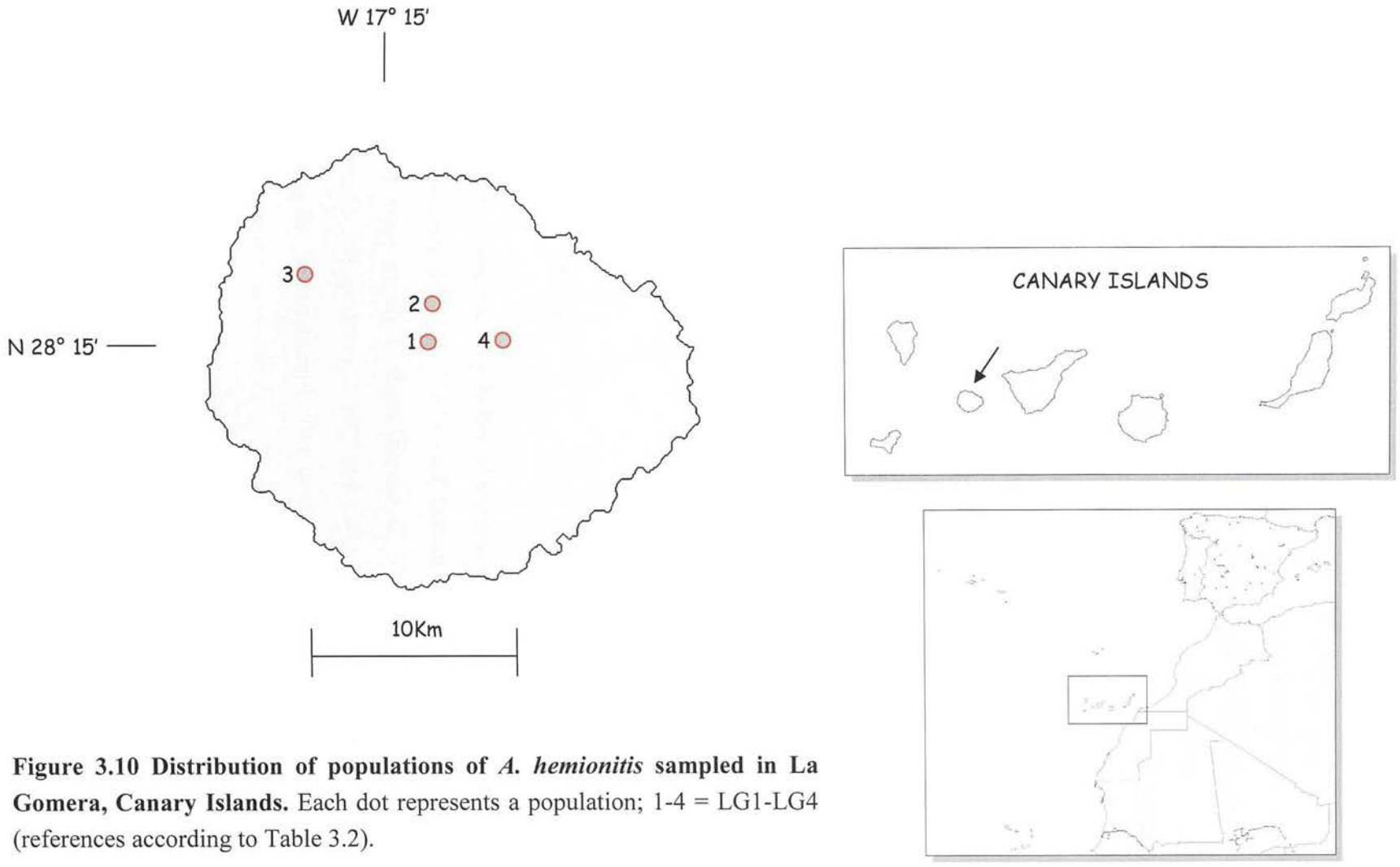


Figure 3.10 Distribution of populations of *A. hemionitis* sampled in La Gomera, Canary Islands. Each dot represents a population; 1-4 = LG1-LG4 (references according to Table 3.2).

3.2.2 Isozyme electrophoresis

3.2.2.1 *Enzyme extraction*

Approximately 50-100mg of plant material was ground on cold porcelain tiles with 12 wells, each individual in a separate well. The tissue was macerated with a pestle in 80µl of grinding buffer (Appendix I). Two filter paper wicks (2mm x 17mm) were soaked in the homogenate obtained and placed into 0.5ml microcentrifuge tubes. To prevent the degradation of enzymes, the tiles were placed on freezer packs and the grinding buffer and microcentrifuge tubes were kept on ice. Duplicates were made for small populations (≤12 plants). In this case, about 150mg of plant material was ground in 120µl and four wicks were used. Following a preliminary study of isozyme diversity in *A. hemionitis* from Sintra (Durães, 2000), multiple duplicate extractions were made to use as marker plants. Three plants were collected in Sintra (Sintra-PT, n=3) and one was collected on Flores, Azores (Flores-M, currently cultivated in the Chelsea Physic Garden, London). The enzyme extracts were placed in labelled plastic bags and stored in an ultra-cold freezer at -80°C until electrophoresis. Under these conditions, the enzyme activity can be maintained for up to five years.

3.2.2.2 *Starch gel preparation*

Starch gels were prepared the day before electrophoresis. Two buffer systems were used, morpholine-citrate (MC, pH 7.0-7.4) and lithium-borate (LiBo, pH 8.5). Gel and electrode buffers were modified from Wendel & Weeden (1990) and Soltis *et al.* (1983), respectively (Appendix I). MC and LiBo gels were ca. 13.7% and were prepared following the same protocol. Two thirds of gel buffer (310ml) were heated in a vacuum flask in a microwave oven until it reached boiling point; the remaining buffer (150ml) was mixed with 63g of hydrolysed potato starch (various suppliers). The starch solution was well mixed to avoid setting, poured into the boiling buffer and vigorously agitated; this solution was then brought to boil. After cooling for two minutes, bubbles were removed from the solution using a vacuum pump. The gel was then poured into

the acrylic moulds, covered with a glass plate when the surface turned opaque and left to set in a humid atmosphere overnight. Prior to sample loading, gels were placed at 4°C for 30 minutes to increase rigidity.

3.2.2.3 *Separation of enzymes by starch gel electrophoresis*

i) Sample loading

The gel was cut parallel to the cathodal end of the mould at about 3cm from the edge. Before loading, the tubes containing the wicks were placed on ice for 15 minutes to defrost. Between 40 and 45 wicks were then inserted into the gel cut spaced 1mm apart. In addition, marker samples were inserted in the middle of the gel. Two extra wicks soaked in bromophenol blue (0.04% in ethanol) were inserted on either side of the gel to indicate how far molecules have travelled. MC and LiBo gels were loaded in parallel, each receiving the same set of samples (one of the pair of wicks from an Eppendorf tube). Gels were covered with cellophane to prevent dehydration.

ii) Electrophoresis conditions

Gels were transferred into tanks containing the corresponding electrode buffer and placed at 4°C to prevent overheating. Electrophoresis started with a current of 35mA. After 30 minutes the wicks were removed and, as LiBo gels shrank slightly, a piece of plastic was inserted along the cathodal end of the gel. This prevented splitting along the cut, which would have interrupted the passage of the current. The current was then adjusted to 60mA for MC gels and 100mA for LiBo gels. Ice bags were placed on the top of the gels to provide further cooling. The overall running time was 6 hours.

3.2.2.4 *Visualisation of isozymes*

The isozymes were histochemically detected with specific stains. The mechanism of enzyme detection is based on precipitation of soluble markers (enzyme substrate and a dye) that become coloured when in contact with the active enzyme. The banding patterns obtained illustrate the mobility of the enzymes in the gel and are used as markers to reveal genetic variability. The majority of enzymes migrate to the anode but a few migrate cathodally, so both anodal and cathodal parts of the gel were stained.

i) Preparation of stains

Stock solutions of the stains were prepared freshly and kept at 4°C. Some chemicals (cofactors and light sensitive dyes) were added just before staining to allow a better banding pattern resolution. Stock solutions and protocols used in the preparation of the stains were modified from Wendel & Weeden (1990) and are provided in the Appendix I.

ii) Gel slicing and staining

Gels were sliced horizontally using a guitar string to obtain replicate slices. Each slice was then placed on a tray and stained for each enzyme system. Most enzyme systems were stained with wet stains (stain solution was poured over the gel) but a few required agarose overlays (stain was mixed with 0.9% agarose solution and poured over the gel). The overlay procedure helps highly diffusible reaction intermediates to be fixed *in situ* allowing a better band resolution. Thirteen enzyme systems were routinely stained: AAT, ACN, DIA, HEX, IDH, LAP, MDH, 6-PGD, PGI, PGM, SkDH, TPI and UGPP. Enzyme nomenclature and type of staining are indicated in Table 3.3. All gels were incubated in the dark at 37°C and checked regularly until bands appeared. This took between 10 minutes and one hour.

Table 3.3 Enzyme nomenclature according to the Enzyme Commission (EC, International Union of Biochemistry 1984), buffer systems and type of staining used to resolve the different enzymes and enzyme quaternary structure. Lithium-borate=L, morpholine-citrate =M, wet stain=W, overlay=O, monomer=M, dimer=D.

Abbreviation	Name	EC code	Buffer system used	Type of staining	Quaternary structure
AAT	aspartate aminotransferase	EC 2.6.1.1	L	W	D
ACN	aconitase	EC 4.2.1.3	M	O	M
DIA	diaphorase	EC 1.6.99	L	W	D
HEX	hexokinase	EC 2.7.1.1	M	W	M
IDH	isocitrate dehydrogenase	EC 1.1.1.42	M	W	D
LAP	leucine aminopeptidase	EC 3.4.11.1	L	W	M
MDH	malate dehydrogenase	EC 1.1.1.37	M	W	D
6-PGD	6-phosphogluconate dehydrogenase	EC 1.1.1.44	M	W	D
PGI	glucose-6-phosphate isomerase	EC 5.3.1.9	L	O	D
PGM	phosphoglucomutase	EC 5.4.2.2	M & L	O	M
SkDH	shikimate dehydrogenase	EC 1.1.1.25	M	W	M
TPI	triose-phosphate isomerase	EC 5.3.1.1	L	O	D
UGPP	UTP-glucose-1-phosphate uridylyltransferase	EC 2.7.7.9	M	O	M

3.2.2.5 *Data analysis*

i) Gel scoring and photographing

After staining, the gels were placed on glass plates to be photographed and scored. The banding patterns for each enzyme system were recorded on a score sheet. As the gels were not fixed for long-term storage, they were photographed using a digital camera.

ii) Interpretation of banding patterns

The banding pattern observed (electrophoretic phenotype) had to be translated into a genotype for the underlying genetic *loci*. Cytological data have demonstrated that *A. hemionitis* is diploid with 2n = 72 (Reichstein, 1981) and the isozyme data were interpreted accordingly.

The interpretation of the (sometimes) complex banding patterns benefited from comparison with data from a comprehensive survey of genetic diversity in *Asplenium* (Vogel *et al.*, 1998a, 1999a; Rumsey *et al.*, 1999; Suter *et al.*, 2000). Allelic variants within *loci* were distinguished from products of different *loci* by assuming that the enzymes in *A. hemionitis* conform to established models of organelle compartmentalisation (Gastony & Darrow, 1983; Weeden & Wendel, 1990).

For enzyme systems encoded by more than one *locus*, the most anodally (fastest) migrating isozyme was labelled '1' (e.g. AAT-1, MDH-1). The distance migrated by each allele from the origin was measured in millimetres and a mobility of '100' was given to a high frequency allele. The mobility of other alleles was calculated accordingly (e.g. 88, 110). The *locus* PGI-2 showed a small number of alleles that migrated cathodally and these were labelled as arbitrary small numbers (e.g. 20, 35). The nomenclature of alleles has no implications for the data analysis. Isozyme alleles are an unordered character and can be labelled in other ways (e.g. 1, 2, 3 or a, b, c). The absolute distance migrated by a given allele could differ between electrophoresis sessions due to variations in the length of electrophoresis or consistency of the starch

gels. However, the relative mobilities of alleles remained approximately constant and correct determination of band homologies was aided by the marker plants.

All isozymes investigated in this study show either monomeric or dimeric quaternary structure (Table 3.3 and Figure 3.11). Homozygous individuals show a single-banded phenotype in both cases. Heterozygotes for monomeric enzymes systems show a two-banded phenotype. Heterozygotes for dimeric enzyme systems show a three-banded phenotype. The exceptions were the *loci* HEX and MDH-2 for which homozygous individuals showed, respectively, a double and triple banding pattern. These loci may be showing some form of post-translational modification (Wendel & Weeden, 1990). This interpretation is consistent with that in studies of other *Asplenium* taxa conducted in the NHM laboratory (Vogel *et al.*, 1998a, 1999a). These patterns are illustrated in Figure 3.11.

Isozyme electrophoresis of thirteen enzyme systems revealed 22 putative *loci* that were all polymorphic. Seven of these *loci* could not be reliably interpreted due to overlapping migration of alleles of different *loci* (AAT-2, DIA-2, MDH-3, MDH-4 and PGM-1) or inconsistent resolution (ACN-1 and PGI-1).

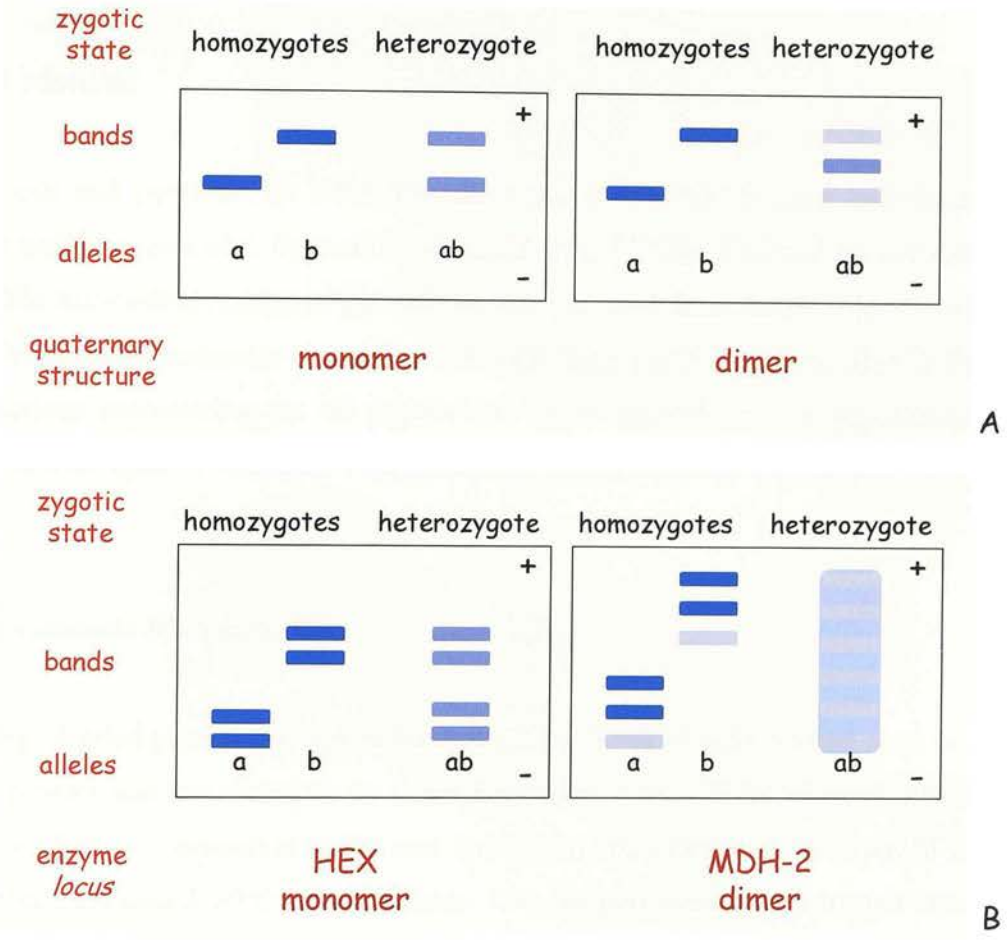


Figure 3.11 Schematic illustration of isozyme zymograms. Strength of bands is represented by different shades of blue. Symbols + and – represent, respectively, the anodal and cathodal ends of the gel. **A. Phenotypic patterns expected for monomeric and dimeric enzyme systems in diploid organisms.** Homozygous individuals show a single-banded phenotype in both cases. Heterozygotes from monomeric enzymes systems show a two-banded phenotype, in which the bands are of equal strength. Heterozygotes for dimeric enzyme systems show a three-banded phenotype. In this pattern the middle band is the strongest and represents the heterodimer between subunits encoded by the two different alleles. The weaker bands above and below this middle band represent the (homodimerised) alleles themselves. **B. Phenotypic patterns observed for HEX and MDH-2 *loci* in homozygous and heterozygous individuals of *A. hemionitis*.** Heterozygous individuals combine all parental bands and if the enzyme is dimeric, one or more bands in addition to the expected heterodimer may be observed.

3.3 Chloroplast DNA sequencing

3.3.1 Material

Methods and protocols for DNA extraction and amplification were optimised using fresh frond samples of *A. hemionitis* cultivated at the RBGE. Material for sequencing of specific non-coding regions of the cpDNA was obtained from fronds dried in silica gel and herbarium specimens. A total of 230 individual plants chosen randomly from all collections representing ca. 40 populations (3-10 individuals per population) were analysed for cpDNA variation. Populations and respective sites are listed in Table 3.4.

3.3.2 Genomic DNA extraction

100mg of dried plant tissue was macerated with a pinch of acid washed sand to a very fine powder and transferred to an 1.5ml Eppendorf tube. 500µl of extraction buffer, 50µl of Sarkosyl extraction buffer and 5µl of β-mercaptoethanol were added and the samples incubated at 60°C for 60 minutes. The samples were then extracted twice with SEVAC (24:1 chloroform:iso-Amyl alcohol) and the DNA precipitated with isopropanol. At this stage, the samples were left on ice for 60 minutes or in the freezer overnight. Pellets were washed in 180µl TE (Tris-hydrochloride EDTA buffer) and 20µl 3M NaOAc. A reprecipitation followed adding 600µl 100% ethanol and the tubes were placed on ice for 60 minutes. Pellets were washed in 70% ethanol, dissolved in 100µl TE and placed in the fridge overnight. The Tris-hydrochloride buffers the solution and the EDTA chelates any Mg^{2+} ions in the solution, protecting the DNA against degradation by nucleases, most of which require magnesium. All samples were stored in the freezer at -20°C. Buffer compositions are given in the Appendix I.

Table 3.4 Samples of *A. hemionitis* used for cpDNA sequence analysis. **n** is number of individuals sequenced in each population;* plants grown at RBGE.

Region	Population/Site	Code	n *trn*L	n *rps4*
Portugal	Sintra	PT-1	3	3
	Sintra	S13	10	10
	Sintra	S14	9	10
	Sintra	S15	1	1
	Sintra	S16	1	1
Azores	Flores	F5	9	9
	Flores	F6	3	3
	Corvo	C1	9	8
	Corvo	C2	4	4
	Faial	FA1	10	10
	Faial	FA5	1	1
	Faial	FA10	3	2
	Faial	FA13	1	1
	Pico	P2	3	3
	Pico	P4	10	10
	São Miguel	SM4	10	10
	São Miguel	SM15	3	2
	Terceira	TE1	1	0
	Terceira	TE2	1	0
	Terceira	TE3	1	0
	Terceira	TE4	1	0
	São Jorge	SJ1	1	0
Madeira	Madeira	M1	3	2
	Madeira	M2	4	4
	Madeira	M3	10	10
	Madeira	M9	10	10
	Madeira	M15	1	1
Canary Islands	Tenerife	T1	1	1
	Tenerife	T4	3	3
	Tenerife	T6	6	6
	Tenerife	T7	1	1
	Tenerife	T8	1	1
	Tenerife	T11	10	10
	Tenerife	T12	1	1
	Tenerife	1959*	1	1
	Gran Canaria	GC1	10	10
	Gran Canaria	GC3	3	3
	Gran Canaria	GC4	1	1
	Gran Canaria	GC5	1	1
	Gran Canaria	GC6	1	1
	La Gomera	LG2	3	3
	La Gomera	LG3	10	10
	La Gomera	1881*	1	0
	La Gomera	1883*	1	1
	La Gomera	1884*	1	1
	La Gomera	2072*	1	1
	La Palma	LP1	3	3
	La Palma	LP5	10	10
	La Palma	2127*	1	1
	El Hierro	EH3	10	10
	El Hierro	EH7	3	3
	El Hierro	1950*	1	1
	El Hierro	1947*	1	1
Cape Verde	São Nicolau	SN1	1	0
	São Nicolau	SN2	1	0
	São Nicolau	SN3	1	0
	Santo Antão	SA1	1	0
	Santo Antão	SA2	1	0
Morocco	Tanger	TA1	10	9
	Tanger	TA2	6	5
Algeria	Algiers	ALG1	1	0
Total			231	214

3.3.3 Chloroplast DNA amplification

The cpDNA regions of interest for sequence analysis were amplified using PCR. The targeted regions were the chloroplast gene fragments *trn*L-*trn*F and *rps*4. The primers Fern-1 (Trewick *et al.*, 2002) and F (Taberlet *et al.*, 1991) were used to amplify the *trn*L intron and the *trn*L-*trn*F intergenic spacer. Primers *rps*4F and *trn*AS (Schneider *et al.*, 2004a) were used to amplify the *rps*4 partial gene and *rps*4-*trn*S intergenic spacer region. These two cpDNA regions will be referred to as *trn*L and *rps*4, respectively.

i) Primer sequences

Fern 1(forward)	5'- GGC AGC CCC CA(AG) AAT CAG GG(AG) AAC C -3'
F (reverse)	5'- ATT TGA ACT GGT GAC ACG AG -3'
*rps*4 (forward)	5'- ATG TC(AC) CGT TA(CT) CGA GG(AG) CCT CGT -3'
*trn*AS (reverse)	5'- TAC CGA GGG TTC GAA TC -3'

ii) PCR components and thermal cycling conditions

PCR amplifications were carried out in thin walled 0.5ml tubes (final volume of 20µl) in a GeneAmp PCR system 9700 (PE Applied Biosystems). Amplification reactions contained: autoclaved deionised water, 1 x Magnesium-free PCR buffer (Bioline), 1.5mM $MgCl_2$ (Bioline), 0.2mM dNTP mix (Sigma), 0.2mg/ml bovine serum albumin (Promega), 1unit *Taq* polymerase (Bioline), 0.2mM of each primer (MWG-Biotech), 1.0µl 1:10 DNA dilution (approximately 50ng of genomic template).

Table 3.5 PCR thermal cycling conditions.

PCR steps	Temperature	Duration	No. of cycles
Hot start	94°C	2 min	1
Denaturation of DNA	94°C	30 s	
Annealing of primers	55°C (*rps*4); 50°C (*trn*L)	30 s	30
Extension	72°C	2 min	
Final extension	72°C	5 min	1
	4°C	until used	

3.3.4 Agarose gel electrophoresis

PCR products were checked for successful amplification on a 1% agarose gel. The required agarose was weighed and added to 1 x TAE buffer in a conical flask. The mixture was heated in a microwave oven for the minimum time for the agarose to completely dissolve, with frequent swirling. The mixture was cooled to approximately 60°C and poured in a casting tray. The gel was allowed to set for 30 minutes, then placed in a gel tank and covered with 1 x TAE buffer. The samples were mixed with 10 x loading buffer to a final concentration of 2 x, loaded in the wells and the appropriate voltage applied (typically 80V for 30minutes). After electrophoresis was complete the gel was stained in 5µg/ml ethidium bromide solution. The ethidium bromide intercalates between DNA base-pairs and fluoresces when activated with ultraviolet radiation allowing visualisation of DNA bands. The gel was placed on a UV light source and photographed.

Successful reactions were then cleaned from solution using a GFX™ PCR Purification kit (Amersham Biosciences), according to the manufacturer's protocol. Cleaned DNA was eluted from the GFX columns in 20µl H_2O and its concentration was assessed by running 5µl on a 1% agarose 1 x TAE gel. Buffers compositions are given in the Appendix I.

3.3.5 Cycle sequencing

Sequencing reactions, cleaning and subsequent electrophoresis of products were carried out by the staff of the sequencing facility at the Institute of Cell, Animal and Population Biology, University of Edinburgh. Products were electrophoresed on an ABI 3730 DNA Analyser (Applied Biosystems). Sequences were returned from the facility as Abi Trace Files (*.abl).

i) Cycle sequencing components and thermal cycling conditions

Sequence reactions were performed with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit Protocol (1/4 reactions). Each reaction was prepared in 0.5ml tubes and contained (final volume of 10µl): 4µl PCR product, 3µl 1 x Sequencing buffer, 1µl Primer (3.2pmolµl^{-1}), 2µl BigDye.

Table 3.6 Cycle sequencing thermal cycling conditions.

Steps	Temperature	Duration	No. of cycles
Step 1	95°C	30 s	25
Step 2	50°C	20 s	
Step 3	60°C	4 min	
Step 4	4°C	until used	

3.3.6 Chloroplast DNA sequence assembly and editing

Complementary 5'→3' and 3'→5' sequence strands were assembled into 'contigs' using SeqManII (Lasergene Navigator, DNAStar). Alignment of sequences was performed in MegAlign (DNAStar). Contigs and alignments were edited manually where appropriate.

3.4 Cytology

A small number of specimens was examined to confirm previous chromosome counts showing that *A. hemionitis* from Portugal (Lovis *in* Reichstein, 1981) and the Azores (Queirós & Ormonde, 1987) is diploid (2n=72) Fresh fronds from three plants growing at the RBGE were used as a source of sporangia. These plants were cultivated from spores collected in Tenerife (Canary Islands). Chromosome squash preparation followed standard procedures (Jong, 1997). Developing sporangia were dissected onto a microscope slide and covered with a drop of aceto-carmine solution. The preparation was heated gently over a spirit-lamp and removed just prior to boiling. It was then covered with a coverslip and two layers of filter paper. Firm pressure was applied with a thumb. Slides were examined at 40x or 100x magnification on a binocular microscope. All plants examined were diploid and showed the same number of chromosomes with 2n=72, therefore confirming previous results.

CHAPTER 4

INVESTIGATION OF BREEDING SYSTEMS AND FINE SCALE GENETIC STRUCTURE OF *Asplenium hemionitis*

4.1 Introduction

The breeding system plays a fundamental role in the evolutionary biology of a species, not only resulting in reproduction but also determining gene flow and the distribution of genetic variation within and among populations (Soltis & Soltis, 1990a; Keiper & McConchie, 2000). Most work on plant breeding systems and their genetic effects has been conducted on angiosperms (Brown, 1990). In this group, the gametophytic phase of the life cycle is represented by pollen and ovules that develop on the sporophyte. However, in pteridophytes, the gametophytic phase is free living and mating does not involve the sporophytic phase of the life cycle. It is therefore necessary to consider the life cycle in detail to understand the factors likely to affect the mating systems in pteridophytes.

4.1.1 Pteridophyte life cycle

Pteridophytes do not flower but reproduce sexually from spores. The life cycle of a pteridophyte has two independent generations (Figure 4.1). The gametophyte generation (n) begins with the production of haploid spores in the sporangia developed on the mature plant (sporophyte). In the young sporangium, the diploid spore mother cell undergoes meiosis and produces four haploid cells that will become the spores. The haploid spores are then dispersed, usually by wind. When spores land on a suitable (moist) substrate they germinate to form a threadlike protonema, which then develops into a gametophyte. This young gametophyte is photosynthetic and continues to grow via mitotic divisions, developing into a heart-shaped prothallus. It is on this structure that the male and female gametangia will develop. Sperm cells are formed in antheridia near the base of the prothallus. Archegonia containing a single egg cell cluster around the apex. Water is required to allow the motile sperm to travel to the opening of the archegonium. The two haploid cells fuse (fertilisation) to form a diploid zygote, initiating the sporophytic generation (2n). The zygote divides mitotically to form an embryo and eventually becomes an independent sporophyte. The sporophyte and

gametophyte are two independent living organisms that can have distinct ecological requirements and experience separate selective forces (Soltis & Soltis, 1990a).

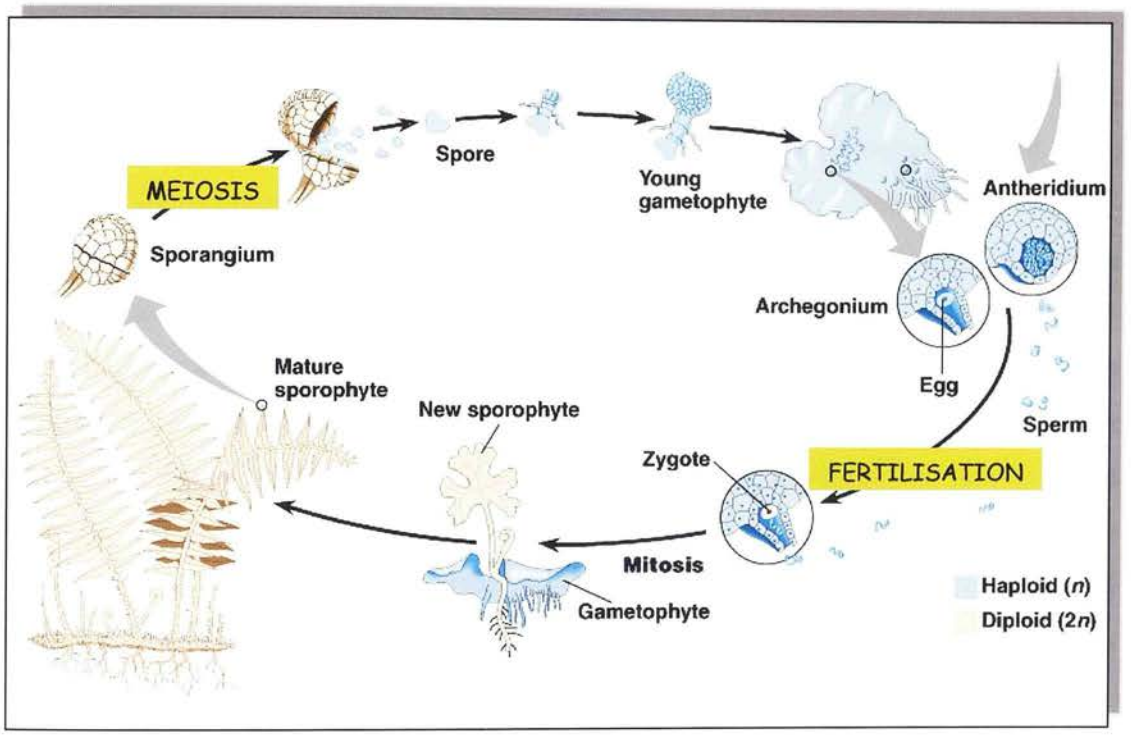


Figure 4.1 Pteridophyte life cycle. Note that although this diagram implies that intra-gametophytic selfing is the only means of fertilisation, inter-gametophytic selfing and outcrossing are also possible (see section 4.1.2). Modified from Pearson Education, Inc., publishing as Benjamin Cummings.

4.1.2 Breeding systems in homosporous pteridophytes

The breeding system of pteridophytes is primarily determined by the independent gametophyte generation, although certain aspects of the sporophyte generation (e.g. distribution, habitat and plant height) may also play a relevant role. However *A. hemionitis* is a homosporous pteridophyte, producing only one type of spore. Homosporous pteridophytes are unique among vascular plants because their gametophytes have the potential to become bisexual and self-fertilise. Consequently, three types of breeding systems are possible in homosporous ferns (Klekowski, 1979): 1) inter-gametophytic crossing may occur when two gametophytes derived from different sporophytes cross-fertilise (analogous to outcrossing in seed plants); 2) inter-gametophytic selfing may occur when two gametophytes derived from the same sporophyte cross-fertilise (analogous to selfing in seed plants) and 3) intra-gametophytic selfing is possible in some pteridophytes and involves the self-fertilisation of a single gametophyte. This produces a completely homozygous sporophyte in one generation (there is no analogue in seed plants). Intra-gametophytic selfing is an extreme form of inbreeding that allows for a single dispersed spore to be an effective coloniser. For outcrossing to occur, two gametophytes originating from spores of different sporophytes must establish at the same site and be sexually mature at the same time.

The potential ability of pteridophytes to self-fertilise led to the proposal (Klekowsky & Baker, 1966; Klekowsky, 1973, 1979) that intra-gametophytic selfing is the predominant mode of reproduction in natural populations of homosporous pteridophytes. However, genetic information acquired via enzyme electrophoresis has led to reassessments of the traditional view on pteridophyte mating systems (reviewed by Haufler, 1987; Soltis & Soltis, 1987, 1990a). It has been discovered that, in nature, outcrossing is the main breeding system in diploid homosporous pteridophytes, while inbreeding is the predominant breeding system in polyploids. Other reviews by Soltis & Soltis (1990b, c, 1992) report that not only do gametophytes of most homosporous pteridophytes fail to self-fertilise, but nearly all matings occur between gametophytes derived from different sporophytes. The general trend is that most species are highly

outcrossing, a few are nearly exclusively inbreeding and a small proportion has a mixed mating system (Soltis & Soltis, 1992).

4.1.3 Hardy-Weinberg equilibrium in homosporous pteridophytes

Analysis of conformance of a population to Hardy-Weinberg (HW) genotypic proportions can be informative when determining the prevailing breeding system. The mating system of a population can be inferred by the distribution of allelic variation into genotypes compared with genotypic proportions predicted using the HW equation. Deviation from HW expectations can be represented by the fixation index (Wright, 1965), or inbreeding coefficient (F), which ranges from -1 to 1. Negative F values indicate an excess of heterozygotes and positive values indicate a deficiency of heterozygotes, relative to HW expectations. A value of zero (or close to zero) signifies HW equilibrium (and random mating between gametophytes). If homosporous pteridophytes are highly inbred, populations exhibit values of F approaching unity. The use of the F value is not the optimal choice for evaluating mating systems because other factors, including drift, selection, migration, and perhaps, mutation, may contribute to deviations from HW equilibrium (Soltis & Soltis, 1990a). However, the mating system is typically regarded as the primary force determining F, and it is used here to estimate the relative roles of selfing and outcrossing in *A. hemionitis*.

Most of the studies on breeding systems of homosporous pteridophytes indicate little, if any, deviation from HW equilibrium (reviewed by Soltis & Soltis, 1990b, 1992). Only a few species (Werth *et al*, 1985; Soltis & Soltis, 1986; Suter, *et al.*, 2000) showed values of F approaching 1.0, signifying extreme inbreeding. For example, high F values were found in the tetraploid *A. trichomanes* L. subsps. *quadrivalens* D. E. Meyer ex Lovis (Suter *et al.*, 2000) and the diploid *A. trichomanes* L. subsps. *trichonames* (Vogel *et al.*, 1999b) indicating substantial inbreeding and the ability to colonise from single spores. These features were hypothesised to have contributed to the postglacial colonisation and widespread distribution of the species in Europe. In contrast, the diploid *A. trichomanes* L. subsps. *inexpectans* Lovis (Vogel *et al.*, 1999b), a species

with a more restricted distribution, showed *F* values near HW equilibrium, compatible with some outbreeding and imposing a restriction on its ability to colonise. *A. hemionitis* is a diploid *taxon* confined to the Macaronesian islands and hyper-oceanic enclaves in Portugal and Morocco. The aim is to determine the breeding system of this species, to establish whether it is capable of intra-gametophytic selfing, and to investigate variation in breeding system across its geographic range.

Electrophoretic analysis of isozymes has been proven to be a powerful technique for exploring the breeding systems in *Asplenium* (Vogel *et al.*, 1999a, 1999b; Suter *et al.*, 2000). The first objective of this study is to use isozyme markers to infer the breeding system of *A. hemionitis*. Isozyme genotype frequencies in each population will be used to calculate the inbreeding coefficient (*Fis*) and to establish whether this varies among populations.

Estimations of the inbreeding coefficient using isozyme markers do not distinguish between intra and inter-gametophytic selfing. The second objective is to use breeding experiments to determine the nature of any inbreeding previously detected with the isozyme data. Independent data on selfing can be obtained from controlled isolate and crossing experiments. These can demonstrate the ability of the species for self fertilisation (intra or inter-gametophytic selfing) and crossing. Therefore, controlled isolate experiments and crosses between gametophytes of known origin were initiated.

In order to determine whether inbreeding could occur within populations as a consequence of small scale genetic structuring of genotypes, a study was conducted of the spatial distribution of sporophytes within a single population. Inter-gametophytic selfing is likely to occur where pairs of neighbouring sporophyte genotypes, and the gametophytes derived from them, are more closely related than are pairs of genotypes located further apart. The objective of the study was therefore to look for an increase in the relatedness of sporophytes as distance between them decreases.

The information on the breeding system is a first step in understanding the distribution of genetic diversity within and among populations of *A. hemionitis*. These issues will be discussed in following chapters.

4.2 Materials and methods

4.2.1 Breeding experiments

Given the potential for gametophytes of homosporous pteridophytes to become bisexual and self-fertilise, experiments were carried out to determine whether sporophytes of *A. hemionitis* can be raised from individual gametophytes via intra-gametophytic selfing.

Experiments on the breeding system were set up as described in section 3.1 (Chapter 3). Cultivation of gametophytes (and fertilisation with development of sporophytes) can potentially take several months to years. Consequently, the individuals and populations of *A. hemionitis* selected for these experiments were randomly chosen from material available at the first stage of this study. Table 4.1 indicates the sites selected for each experimental treatment (isolate and pair experiments). The number of different treatments was set up according to the availability of germinated spores.

Table 4.1 Source of individuals of *A. hemionitis* used in the breeding experiments.

Experimental treatment	Source of individuals	Number of plates
Intra-gametophytic selfing (single gametophytes)	Sintra, Faial, Pico, São Miguel, Tenerife, Gran Canaria, Morocco	40
Inter-gametophytic selfing (pairs of gametophytes from the same individual)	Sintra, Faial, Tenerife, Gran Canaria	30
Outcrossing within populations (pairs of gametophytes from different individuals of the same population)	Faial, Pico, São Miguel, Tenerife, Gran Canaria	20
Outcrossing between populations (pairs of gametophytes from different populations)	Sintra, Faial, Pico, São Miguel, Madeira, Tenerife, Gran Canaria, Morocco	30

4.2.2 Estimation of the inbreeding coefficient using isozyme markers

4.2.2.1 *Material and isozyme electrophoresis*

A total of 105 populations of *A. hemionitis* comprising 2073 individuals were analysed for isozyme markers. The mean number of individuals scored per population was 20 (4–116). Data from thirteen enzyme systems encoded by 15 *loci* were used to estimate the inbreeding coefficient. Material collections, enzymes and methods used in this analysis are indicated in detail in section 3.2 (Chapter 3).

4.2.2.2 *Statistical analysis*

The inbreeding coefficient (*Fis*) was estimated for all isozyme *loci* in each population and departure from HW equilibrium was assessed using FSTAT 2.9.3.2 (Goudet, 2001). Tests for heterozygote deficiency (*HD*) and heterozygote excess (*HE*) were performed by randomising alleles among individuals. Calculation of mean *Fis* estimates assumes independence of *loci*. Genotypic linkage equilibrium between all polymorphic *loci* was tested using FSTAT 2.9.3.2.

Global inbreeding coefficient (*f*, Weir & Cockerham, 1984) for the species, each island and region was estimated over all isozyme *loci* using FSTAT 2.9.3.2. Homogeneity of *Fis* values between *loci* was tested by jackknifing and confidence intervals generated by bootstrapping over all *loci*.

Populations were grouped according to geographical isolation from other populations. This grouping resulted in eleven islands (Flores, Corvo, Pico, Faial, São Miguel, Madeira, Tenerife, Gran Canaria, La Gomera, La Palma and El Hierro) and two 'mainland islands' (Sintra and Tanger). 'Islands' were further grouped into five regions. These were Portugal, Morocco, Azores, Madeira and Canary Islands.

4.2.3 Analysis of fine scale genetic structure

Comparison of geographic neighbours within population was carried out in one population from Sintra (S14, n=116) growing on a manmade wall (c. 30m long). The position of all sporophytes on the wall was mapped, and the isozyme genotype of these individuals determined at eleven polymorphic *loci*. Analysis at the individual level was carried out by calculating a measure of genetic relatedness (kinship coefficient, Loiselle *et al.*, 1995) between all pairs of individuals within specific distance classes, using SPAGeDI 1.1b (Hardy & Vekemans, 2002). This kinship is computed as a correlation multilocus coefficient between allelic states. Physical distance (in cm) between pairs of individuals is calculated in SPAGeDI from the vertical and horizontal axis coordinates. Maximum distance between individuals (c. 3000cm) was divided in 100 distance classes. Upper distances for each distance class are automatically chosen in SPAGeDI in order to obtain intervals with the same number of pairwise comparisons (maximum distance for each distance class is given in Appendix II). Mean kinship estimations over all *loci* and per distance class were regressed against classes of physical distances. Gene dispersal distance (S) was estimated using the formula $S=(Nb/4\pi.D)^{1/2}$ (Fenster *et al.*, 2003), where D is the population density and Nb (neighbourhood size) is inferred as $Nb\approx-(1-K)/b_{log}$ (K is the kinship coefficient between adjacent individuals and b_{log} is the regression slope based on the logarithm of spatial distance).

4.3 Results

4.3.1 Breeding experiments

A. hemionitis exhibited good spore germination under controlled conditions. Many gametophytes showed very slow growth and were only at the appropriate stage for the experimental treatments after several months. In the intra-gametophytic selfing experiment, none of the isolated gametophytes produced a sporophyte. These gametophytes were allowed to grow on for a year, and eventually died or were lost due to algal and fungal growth. In the inter-gametophytic selfing and crossing experiments, sporophytes were observed after two to 12 months. Approximately 50% of the paired gametophytes derived from the same individual produced a sporophyte (Figure 4.3). The crossing experiments produced sporophytes, both in combinations of gametophytes from the same population and from different populations. The rate of sporophyte production was c. 35% and 53.3%, respectively. Sporophytes were observed in both gametophytes in seventeen pairs, equally in the inter-gametophytic selfing and crossing experiments. The paired gametophytes on which sporophytes were never observed, were maintained in culture and proliferated extensively, some growing up to 2cm long. Different stages of the development of gametophyte and sporophyte in *A. hemionitis* are illustrated on Figure 4.2.

The breeding experiments provide no evidence that intra-gametophytic selfing can occur in *A. hemionitis*, at least under laboratory conditions. However, selfing is still possible via inter-gametophytic selfing and took place in the experiment with a slightly lower probability than inter-gametophytic crossing (Figure 4.3).

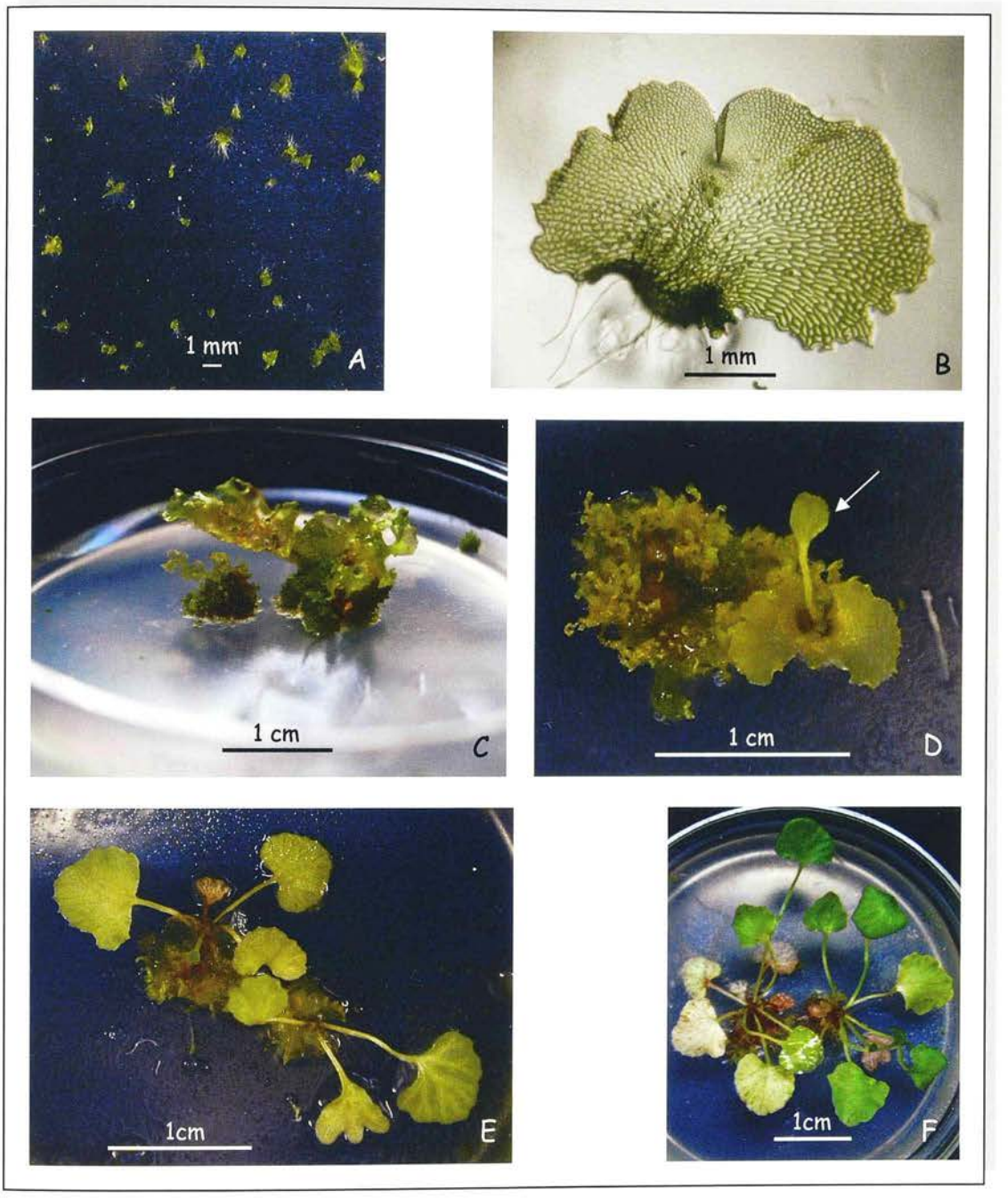


Figure 4.2 In vitro development of the gametophyte and sporophyte of *A. hemionitis*.
A. Three month old gametophytes growing on Phytagel; **B.** Five month old gametophyte; **C.** Two year old paired gametophytes that proliferated extensively without ever producing a sporophyte; **D.** Two month old sporophyte (indicated by arrow) produced in only one of the paired gametophytes; **E.** Two month old sporophytes produced in both of the paired gametophytes; **F.** Eight month old sporophytes (same as E).

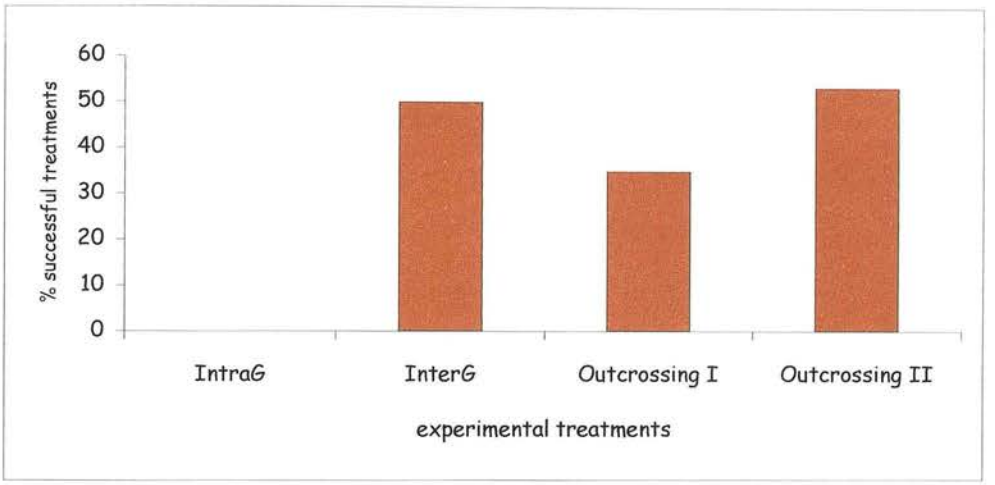


Figure 4.3 Percentage of treatments that produced sporophytes in the different breeding experiments. IntraG: Intra-gametophytic selfing; InterG: inter-gametophytic selfing; Outcrossing I: outcrossing within populations; Outcrossing II: outcrossing between populations. Note: some treatments produced several sporophytes due to regeneration of gametophytes.

4.3.2 Extent of inbreeding as estimated with *Fis*

Fis estimates for 105 populations of *A. hemionitis* (Tables 4.2 to 4.7) range between -0.385 (Sintra 6) and 0.450 (La Palma 6). Thirty one populations (29.5%) exhibit significantly positive *Fis* values indicating a deficit of heterozygotes (*P*-value <0.05). These populations are found throughout the main regions of distribution of *A. hemionitis* (three in Sintra, eleven in the Azores, three in Madeira and fourteen in the Canary Islands). Twenty eight populations (26.7%) exhibit negative *Fis* values but only one is significantly different from zero (-0.130, São Miguel 1). The mean *Fis* values differ between the thirteen 'islands' (Table 4.8), ranging from 0.203 (Flores) to -0.022 (Tanger). Sintra, Flores, Faial, São Miguel, Tenerife, La Palma and La Gomera exhibit *Fis* values that are significantly positive (*P*-value <0.001). At a regional scale (Table 4.9), Portugal and the archipelagos of Azores and Canary Islands show significantly positive *Fis* values (respectively 0.119, 0.092 and 0.082, *P*-value <0.001). The global inbreeding (*f*) estimated over all *loci* and populations is 0.084 and is significantly different from zero (*P*-value <0.001). Across all *loci*, significant genotypic association is only found for one pair of *loci* (SKDH and DIA, *P*-value=0.00048).

Intra-gametophytic selfing generates individuals that are homozygous at all *loci*. The proportion of individuals homozygous at all *loci* analysed is 1.5% (Table 4.10). The lowest occurrence of such individuals is in Madeira (0.43%) and the highest is in Faial (4.07%) and Flores (3.57%).

These results show that, although outcrossing is the predominant breeding system in *A. hemionitis*, there is evidence for inbreeding in some populations. The low level of individuals homozygous at all *loci* analysed suggests that this is not accounted for by intra-gametophytic selfing.

4.3.3 Estimate of fine scale genetic structure at Sintra

Figure 4.4 illustrates the change in relatedness between sporophytes on a wall at Sintra (Sintra 14) as distance between individuals changes. The analysis of fine scale genetic structure in this population indicates a significant positive autocorrelation among individuals located up to 2.8m (maximum distance for each distance class in Appendix II). Kinship values are not significantly different from zero beyond this distance, with the exception of a significantly positive correlation at about 4.3m, 10.6m and 22.0m. These values correspond to events likely to have occurred by chance. The correlation is significantly supported by the regression analysis with a coefficient of determination (r^2) of 0.472. Therefore, the correlogram shows that sporophytes growing closely together are more related to each other than to the more distant ones. This indicates a non-random distribution of related individuals on the wall and substantial substructure within the population. The approximate gene dispersal distance is 22cm, which represents an area of about $0.152m^2$ around each individual. These results suggest that most spores fall very close to the source sporophyte and dispersion is restricted.

Table 4.2 Inbreeding coefficient estimates (*Fis*) for each of 15 isozyme *loci* in populations of *A. hemionitis* from Sintra (S1-S14) and Tanger (TA1-TA2). Also shown are mean over all *loci Fis* estimates and *P*-values for heterozygote deficit (*HD*) and excess (*HE*). n: number of individuals per population; ns: non-significant at the 5% level; ***P*-value <0.01, ****P*-value <0.001; blank spaces indicate not calculated due to monomorphic *loci*.

	Sintra														Tanger	
locus	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**TA1**	**TA2**
n	6	11	22	19	17	5	7	9	29	17	15	17	65	116	39	11
AAT-1	.	.	.	.	.	.	.	0.000	.	.	.	.	.	.	-0.118	0.000
ACN-2	.	.	.	.	.	.	.	.	.	.	.	.	0.190	-0.303	0.223	0.000
DIA	0.000	.	-0.105	0.000	.	.	.	.	0.208	-0.032	0.000	0.050	0.552	1.000	0.323	0.412
HEX	-0.250	.	0.250	0.212	-0.043	.	0.143	-0.116	0.038	0.194	0.200	0.000	0.263	0.144	.	.
IDH	-0.154	-0.368	-0.183	-0.125	0.220	-0.333	0.750	-0.297	0.404	-0.181	-0.107	0.155	0.059	0.220	-0.096	-0.111
LAP	.	.	.	.	.	.	.	.	.	.	.	0.273	-0.014	0.606	0.002	.
MDH-1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.126	-0.053
MDH-2	-0.250	-0.053	0.344	-0.091	0.000	.	0.217	0.636	-0.143	0.614	0.200	0.000	-0.143	-0.018	-0.163	-0.250
6-PGD	0.070	.	-0.121	0.463	0.614	-0.600	-0.220	0.216	0.218	-0.168	-0.069	0.427	0.261	0.191	-0.027	.
PGI-2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-0.119	-0.008
PGM-2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-0.056	.
SkDH	0.318	0.070	0.176	0.169	-0.124	-0.143	0.273	0.111	-0.266	0.033	0.169	0.221	-0.019	0.261	-0.013	-0.250
TPI-1	.	.	.	.	.	.	.	.	.	.	.	.	0.663	.	-0.027	.
TPI-2	0.000	-0.111	0.000	.	.	.	0.000	1.000	.	.	.	0.458	-0.041	0.219	-0.245	-0.132
UGPP	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mean	**-0.005**	**-0.102**	**0.042**	**0.100**	**0.139**	**-0.385**	**0.197**	**0.135**	**0.073**	**0.046**	**0.058**	**0.203**	**0.139**	**0.164**	**-0.016**	**-0.044**
P-value *HD*	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	**	**	***	ns	ns
P-value *HE*	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

Table 4.3 Inbreeding coefficient estimates (*Fis*) for each of 15 isozyme *loci* in populations of *A. hemionitis* from the Azores (Flores F1-F6; Corvo C1-C2; Faial FA1-FA9). Also shown are mean over all *loci Fis* estimates and *P*-values for heterozygote deficit (*HD*) and excess (*HE*). n: number of individuals per population; ns: non-significant at the 5% level; **P*-value <0.05, ***P*-value <0.01; blank spaces indicate not calculated due to monomorphic *loci*.

	Flores						Corvo		Pico			Faial								
locus	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**	**C1**	**C2**	**P1**	**P2**	**P3**	**FA1**	**FA2**	**FA3**	**FA4**	**FA5**	**FA6**	**FA7**	**FA8**	**FA9**
n	10	23	16	11	14	9	13	10	30	15	4	34	14	11	28	10	30	26	14	27
AAT-1	0.000	.	-0.143	-0.048	.	.	-0.519	-0.047	0.066	.	.	-0.008	.	0.000	.	-0.231	-0.018	-0.067	-0.037	0.000
ACN-2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DIA	.	0.656	0.651	0.766	0.000	0.500	0.000	1.000	0.659	.	.	0.660	.	.	-0.080	0.640	0.788	-0.027	.	0.169
HEX	.	0.000	.	-0.100	.	0.000	-0.043	0.000	.	0.000	.	0.469	0.316	0.000	-0.050	.	.	-0.412	-0.043	0.085
IDH	0.386	0.200	0.380	-0.082	0.085	-0.333	0.065	-0.200	0.103	-0.077	-0.200	0.192	-0.300	0.000	-0.325	-0.047	0.120	-0.351	0.314	0.224
LAP	.	.	.	.	.	.	.	0.000	.	.	.	0.480	.	.	0.476	-0.200	-0.059	.	.	.
MDH-1	0.000	.	0.000	.	0.000	0.407	0.321	.	0.361	.	0.143	0.000	0.000	-0.176	-0.043	0.000	-0.168	.	-0.083	-0.053
MDH-2	1.000	-0.110	0.192	0.464	0.496	0.774	0.090	-0.200	0.097	.	-0.412	0.189	0.350	-0.228	0.288	0.274	-0.111	0.087	-0.144	-0.070
6-PGD	-0.200	.	0.000	-0.200	.	-0.231	0.294	0.077	-0.226	-0.235	-0.200	0.000	0.464	-0.026	0.053	-0.213	-0.081	-0.053	0.304	-0.040
PGI-2	0.000	.	.	.	0.649	.	.	.	.	.	.	0.000	.	.	.	.	-0.086	.	-0.040	.
PGM-2	.	.	.	.	.	.	-0.091	.	.	.	.	0.000	.	0.000	-0.038	-0.200	-0.018	-0.020	.	-0.040
SkDH	-0.207	0.082	0.200	1.000	0.378	.	-0.019	0.845	0.135	.	0.000	-0.051	0.217	0.138	0.190	-0.047	-0.025	0.144	0.023	0.456
TPI-1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TPI-2	-0.047	0.000	-0.200	-0.132	0.000	0.190	.	.	0.392	-0.120	-0.200	-0.015	.	-0.053	-0.259	0.043	-0.189	0.034	-0.110	0.000
UGPP	.	.	0.000	0.000	.	.	0.000	.	0.000	0.000	.	.	.	.	0.000	.	0.000	.	.	.
Mean	**0.155**	**0.100**	**0.142**	**0.148**	**0.352**	**0.219**	**-0.003**	**0.229**	**0.136**	**-0.137**	**-0.164**	**0.226**	**0.206**	**-0.047**	**0.031**	**0.005**	**-0.028**	**-0.068**	**0.020**	**0.104**
P-value *HD*	ns	ns	ns	ns	**	ns	ns	ns	*	ns	ns	**	*	ns	ns	ns	ns	ns	ns	ns
P-value *HE*	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

Table 4.4 Inbreeding coefficient estimates (*Fis*) for each of 15 isozyme *loci* in populations of *A. hemionitis* from the Azores (Faial FA10-FA13; São Miguel SM1-SM15). Also shown are mean over all *loci Fis* estimates and *P*-values for heterozygote deficit (*HD*) and excess (*HE*). n: number of individuals per population; ns: non-significant at the 5% level; **P*-value <0.05, ***P*-value <0.01, ****P*-value <0.001; blank spaces indicate not calculated due to monomorphic *loci*.

	Faial				São Miguel														
locus	**FA10**	**FA11**	**FA12**	**FA13**	**SM1**	**SM2**	**SM3**	**SM4**	**SM5**	**SM6**	**SM7**	**SM8**	**SM9**	**SM10**	**SM11**	**SM12**	**SM13**	**SM14**	**SM15**
n	17	30	24	5	26	10	9	30	24	15	21	8	29	21	11	10	26	7	20
AAT-1	.	.	-0.023	-0.067	-0.161	0.391	0.273	0.076	0.034	-0.176	0.013	-0.750	-0.043	0.231	-0.280	.	0.315	-0.231	0.118
ACN-2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DIA	-0.067	-0.043	1.000	0.600	0.580	-0.059	.	.	0.296	.	0.000	0.000	0.525	-0.053	.	0.561	-0.008	.	-0.042
HEX	0.050	.	0.657	.	-0.680	-0.636	-0.600	0.408	-0.426	0.200	0.452	0.368	-0.269	-0.187	0.474	0.043	0.355	-0.200	-0.606
IDH	0.000	0.291	0.352	-0.143	-0.313	-0.600	-1.000	-0.063	-0.234	0.364	-0.026	0.067	-0.191	0.724	0.636	-0.200	0.233	0.000	0.127
LAP	0.000	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MDH-1	0.211	0.057	0.121	0.273	0.000	-0.059	.	-0.074	.	0.000	0.212	-0.077	-0.302	0.341	-0.111	0.100	0.000	-0.091	.
MDH-2	0.299	0.352	0.171	-0.200	0.169	0.327	0.368	-0.216	0.535	0.364	0.333	0.000	-0.045	0.446	0.735	0.106	0.082	0.419	0.309
6-PGD	.	-0.067	-0.207	-0.333	-0.042	-0.071	0.407	0.360	0.081	0.047	-0.053	0.475	0.429	-0.043	0.123	0.405	-0.084	-0.091	-0.146
PGI-2	.	-0.074	-0.095	.	0.000	0.000	.	1.000	0.102	.	-0.026	-0.167	0.359	0.000	-0.053	-0.059	-0.088	.	.
PGM-2	0.000	-0.018	-0.023	0.600	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SkDH	0.319	0.316	0.265	0.000	-0.342	0.419	0.149	0.085	0.351	0.323	0.263	0.364	-0.053	0.703	0.057	-0.059	-0.072	0.793	0.247
TPI-1	.	.	0.000	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TPI-2	-0.032	-0.177	-0.095	0.273	0.000	.	.	-0.215	-0.160	0.012	-0.026	0.000	-0.077	-0.250	-0.053	-0.200	-0.116	.	0.324
UGPP	.	.	0.000	.	.	.	.	.	.	0.000	.	.	.	.	.	0.000	.	.	.
Mean	**0.159**	**0.135**	**0.127**	**0.145**	**-0.130**	**-0.004**	**-0.088**	**0.070**	**0.064**	**0.180**	**0.179**	**0.097**	**-0.001**	**0.251**	**0.239**	**0.117**	**0.103**	**0.180**	**0.069**
P-value *HD*	*	**	**	ns	ns	ns	ns	ns	ns	*	**	ns	ns	***	**	ns	ns	ns	ns
P-value *HE*	ns	ns	ns	ns	*	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

Table 4.5 Inbreeding coefficient estimates (*Fis*) for each of 15 isozyme *loci* in populations of *A. hemionitis* from Madeira (M1-M9). Also shown are mean over all *loci Fis* estimates and *P*-values for heterozygote deficit (*HD*) and excess (*HE*). n: number of individuals per population; ns: non-significant at the 5% level; **P*-value <0.05; blank spaces indicate not calculated due to monomorphic *loci*.

locus	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**	**M9**	**M10**	**M11**	**M12**	**M13**	**M14**
n	5	8	67	10	18	12	13	14	18	17	27	5	13	4
AAT-1	-0.091	.	0.053	-0.286	-0.071	-0.105	0.000	-0.330	-0.155	-0.086	-0.109	-0.143	0.486	-0.500
ACN-2	.	.	1.000	.	.	.	0.065	.	.	.	.	.	.	.
DIA	-0.333	-0.273	0.063	0.100	-0.081	0.148	-0.043	0.621	0.214	0.273	-0.270	0.000	0.117	-0.500
HEX	.	.	.	.	.	.	.	0.000	-0.030	.	0.000	.	0.000	.
IDH	-0.200	-0.200	0.043	.	-0.030	-0.146	-0.091	0.304	-0.079	-0.174	0.281	-0.333	0.007	0.571
LAP	-0.143	-0.105	-0.022	0.400	.	.	-0.043	.	0.000	0.154	0.658	0.000	0.000	.
MDH-1	.	.	0.000	.	.	.	.	.	.	.	.	.	.	.
MDH-2	.	.	-0.095	.	0.000	.	-0.043	0.464	.	0.048	.	0.000	-0.021	0.000
6-PGD	1.000	.	0.214	0.372	0.000	-0.023	.	0.500	-0.214	0.000	0.294	0.000	1.000	.
PGI-2	-0.091	0.164	-0.121	0.153	-0.158	0.006	-0.200	-0.033	0.302	-0.010	-0.003	0.077	-0.462	-0.200
PGM-2	-0.333	.	0.210	.	0.000	-0.111	.	0.188	0.329	0.000	-0.226	.	1.000	-0.200
SkDH	-0.185	-0.296	0.119	0.633	0.201	-0.071	-0.043	0.197	0.441	-0.075	0.008	0.200	0.318	-0.125
TPI-1	.	.	1.000	0.640	.	.	.	.	.	.	.	.	.	.
TPI-2	0.273	0.028	-0.068	-0.125	-0.157	-0.203	0.000	-0.013	-0.199	0.150	0.034	-0.280	-0.210	.
UGPP	-1.000	0.300	0.057	-0.080	0.469	0.154	0.000	-0.013	0.382	0.089	0.346	-0.333	0.063	0.143
Mean	**-0.062**	**-0.024**	**0.054**	**0.232**	**0.008**	**-0.019**	**-0.057**	**0.159**	**0.106**	**0.018**	**0.042**	**-0.073**	**0.110**	**-0.085**
P-value *HD*	ns	ns	ns	*	ns	ns	ns	*	*	ns	ns	ns	ns	ns
P-value *HE*	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

Table 4.6 Inbreeding coefficient estimates (*Fis*) for each of 15 isozyme *loci* in populations of *A. hemionitis* from the Canary Islands (Tenerife T1-T11; Gran Canaria GC1-GC5; El Hierro EL1-EL2). Also shown are mean over all *loci Fis* estimates and *P*-values for heterozygote deficit (*HD*) and excess (*HE*). n: number of individuals per population; ns: non-significant at the 5% level; **P*-value <0.05, ***P*-value <0.01, ****P*-value <0.001; blank spaces indicate not calculated due to monomorphic *loci*.

	Tenerife											Gran Canaria					El Hierro	
locus	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**	**T8**	**T9**	**T10**	**T11**	**GC1**	**GC2**	**GC3**	**GC4**	**GC5**	**EL1**	**EL2**
n	31	6	12	25	9	10	17	49	11	35	25	30	38	30	6	14	10	5
AAT-1	.	.	0.645	0.040	.	0.000	.	.	-0.047	.	0.161	-0.048	-0.211	-0.047	-0.250	-0.110	0.250	.
ACN-2	-0.132	-0.111	0.459	0.000	-0.077	-0.200	-0.103	.	.	.	.	-0.203	0.269	.	-0.143	.	.	.
DIA	0.202	0.615	0.000	.	.	.	0.000	0.472	.	0.391	0.000	-0.009	1.000	.	.	.	.	.
HEX	.	.	.	.	.	.	0.000	.	.	.	-0.175	.	.	0.000	.	.	-0.500	.
IDH	-0.146	-0.481	0.645	-0.026	-0.297	0.100	0.458	-0.134	-0.132	-0.188	0.137	-0.185	0.136	-0.318	-0.667	-0.113	0.217	0.143
LAP	0.380	1.000	-0.114	-0.043	.	.	0.212	0.306	.	0.210	.	0.132	-0.058	-0.058	.	-0.130	0.400	.
MDH-1	.	.	.	.	-0.231	0.000	.	-0.021	.	0.000	-0.021	-0.084	0.226	0.208	0.000	0.458	.	.
MDH-2	-0.122	-0.250	0.529	-0.047	0.351	-0.216	0.273	0.047	0.043	-0.333	-0.417	0.140	0.000	-0.149	-0.250	0.316	0.100	-0.091
6-PGD	0.061	-0.154	0.457	0.127	-0.067	-0.143	-0.155	0.198	0.459	0.265	0.495	-0.094	-0.039	0.000	-0.111	.	.	.
PGI-2	-0.008	.	.	-0.067	0.000	0.000	.	-0.011	-0.053	0.397	0.000	-0.051	0.490	-0.079	-0.111	-0.333	.	0.077
PGM-2	0.138	-0.053	0.000	-0.071	-0.164	-0.034	-0.110	0.143	0.302	0.447	0.276	0.000	0.205	-0.020	0.000	-0.091	-0.200	0.000
SkDH	0.421	0.333	0.132	0.109	0.262	0.303	-0.035	0.199	-0.186	0.269	0.210	-0.104	0.144	0.100	0.167	0.137	0.237	1.000
TPI-1	.	0.000	.	-0.067	.	1.000	.	0.306	1.000	0.000	.	.	-0.042	-0.018	.	.	-0.059	.
TPI-2	0.235	.	0.302	-0.021	.	-0.059	.	-0.117	-0.026	-0.015	0.000	-0.074	-0.042	-0.077	.	0.000	-0.080	.
UGPP	-0.056	.	-0.048	-0.020	.	0.000	0.626	-0.011	-0.053	0.000	0.027	-0.024	-0.057	.	-0.111	.	.	0.000
Mean	**0.142**	**0.081**	**0.303**	**0.016**	**0.024**	**0.074**	**0.139**	**0.151**	**0.078**	**0.133**	**0.104**	**-0.045**	**0.091**	**-0.072**	**-0.149**	**0.023**	**0.070**	**0.217**
P-value *HD*	**	ns	***	ns	ns	ns	*	***	ns	**	*	ns	*	ns	ns	ns	ns	ns
P-value *HE*	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

Table 4.7 Inbreeding coefficient estimates (*Fis*) for each of 15 isozyme *loci* in populations of *A. hemionitis* from the Canary Islands (El Hierro EH3-EH9; La Palma LP1-LP7; La Gomera LG1-LG4). Also shown are mean over all *loci Fis* estimates and *P*-values for heterozygote deficit (*HD*) and excess (*HE*). n: number of individuals per population; ns: non-significant at the 5% level; **P*-value <0.05, ***P*-value <0.01, ****P*-value <0.001; blank spaces indicate not calculated due to monomorphic *loci*.

	El Hierro							La Palma							La Gomera			
locus	EH3	EH4	EH5	EH6	EH7	EH8	EH9	LP1	LP2	LP3	LP4	LP5	LP6	LP7	LG1	LG2	LG3	LG4
n	35	7	8	5	27	7	23	26	12	36	9	30	10	13	29	19	30	11
AAT-1	.	.	.	.	-0.110	-0.091	0.282	-0.088	-0.048	-0.173	-0.032	0.000	0.455	-0.143	0.000	-0.059	0.155	.
ACN-2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DIA	0.000	.	.	-0.143	-0.087	-0.333	-0.011	.	0.302	-0.007	0.000	-0.062	1.000	.	-0.080	.	0.223	0.429
HEX	-0.154	.	0.000	.	-0.182	-0.200	0.055	.	.	0.185	.	.	.	.	.	-0.053	.	-0.053
IDH	-0.173	0.217	-0.400	-1.000	-0.457	-0.474	-0.143	-0.071	.	-0.007	.	-0.130	.	.	-0.064	0.469	0.576	0.652
LAP	.	.	.	.	.	0.000	.	-0.114	.	-0.114	.	0.314	.	.	0.357	.	.	.
MDH-1	.	.	.	.	0.000	0.143	.	0.000	.	0.000	.	.	.	.	.	-0.029	.	.
MDH-2	0.061	0.000	-0.429	0.000	-0.093	0.368	-0.144	0.362	0.479	-0.077	0.510	0.204	0.308	0.287	0.096	-0.333	-0.118	-0.429
6-PGD	0.000	.	.	.	0.350	0.000	-0.086	-0.097	0.429	-0.110	-0.123	-0.042	-0.029	-0.412	0.364	0.186	0.079	-0.077
PGI-2	0.292	-0.043	-0.077	0.600	-0.026	0.000	0.368	-0.121	-0.100	-0.042	.	0.035	0.000	-0.103	0.378	0.202	0.263	0.103
PGM-2	0.132	-0.500	-0.105	-0.143	0.339	-0.500	-0.026	-0.123	0.302	-0.078	0.442	-0.040	0.300	0.250	0.041	0.256	-0.094	-0.053
SkDH	-0.200	-0.667	0.475	0.250	-0.001	0.489	0.406	0.172	0.451	0.292	0.818	0.242	0.642	0.669	0.083	-0.180	0.213	0.346
TPI-1	0.133	0.143	.	.	0.658	.	.	0.297	.	-0.007	.	.	.	-0.043	.	.	-0.012	0.492
TPI-2	-0.100	.	.	0.000	-0.109	0.000	0.157	-0.074	-0.234	-0.119	-0.067	.	.	0.000	0.069	-0.014	.	.
UGPP	.	.	0.300	.	0.312	.	0.000	.	.	.	0.000	0.000	0.000	.	-0.042	.	.	.
Mean	**-0.018**	**-0.161**	**0.024**	**-0.034**	**0.025**	**-0.064**	**0.098**	**0.057**	**0.266**	**0.015**	**0.381**	**0.098**	**0.450**	**0.202**	**0.097**	**0.061**	**0.136**	**0.150**
P-value *HD*	ns	ns	ns	ns	ns	ns	ns	ns	**	ns	***	ns	***	*	*	ns	**	*
P-value *HE*	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

Table 4.8 Mean inbreeding coefficient estimates (*Fis*) and bootstrap confidence intervals over 15 isozyme *loci* in populations of *A. hemionitis* grouped by island; * significant at the 5% level; *P*-value<0.001.

Region	Site/island	No. populations	Mean *Fis*	Confidence intervals (95% level)
Portugal	Sintra	14	0.119*	0.056–0.184
Morocco	Tanger	2	-0.022	-0.113–0.083
Azores	Flores	6	0.203*	0.026–0.329
	Corvo	2	0.088	-0.092–0.313
	Pico	3	0.062*	0.040–0.181
	Faial	13	0.074	-0.001–0.130
	São Miguel	15	0.084	-0.006–0.170
Madeira	Madeira	14	0.050	-0.001–0.112
Canary Islands	Tenerife	11	0.123*	0.037–0.195
	Gran Canaria	5	-0.003	-0.056–0.048
	El Hierro	9	0.043	-0.048–0.144
	La Gomera	4	0.110*	0.024–0.195
	La Palma	7	0.127*	0.008–0.223

Table 4.9 Mean inbreeding coefficient estimates (*Fis*) and bootstrap confidence intervals over 15 isozyme *loci* in populations of *A. hemionitis* grouped by region. Also shown is the global inbreeding coefficient for *A. hemionitis*; * significant at the 5% level; *P*-value<0.001.

Regions	Mean *Fis*	Confidence intervals (95% level)
Portugal	0.119*	0.056–0.184
Morocco	-0.022	-0.113–0.083
Azores	0.092*	0.011–0.159
Madeira	0.050	-0.001–0.112
Canary Islands	0.082*	0.020–0.135
f	0.084*	0.039–0.124

Table 4.10 Proportion of completely homozygous individuals at all scored *loci* in *A. hemionitis* main distribution sites.

Region	Site/island	No. of individuals	% individuals homozygous at all scored *loci*
Portugal	Sintra	355	0.56
Morocco	Tanger	50	0.00
Azores	Flores	84	3.57
	Corvo	23	0.00
	Pico	49	2.00
	Faial	270	4.07
	São Miguel	285	2.12
Madeira	Madeira	231	0.43
Canary Islands	Tenerife	230	1.30
	Gran Canaria	118	0.00
	El Hierro	127	1.57
	La Gomera	89	1.12
	La Palma	162	0.61
All populations		2073	1.50

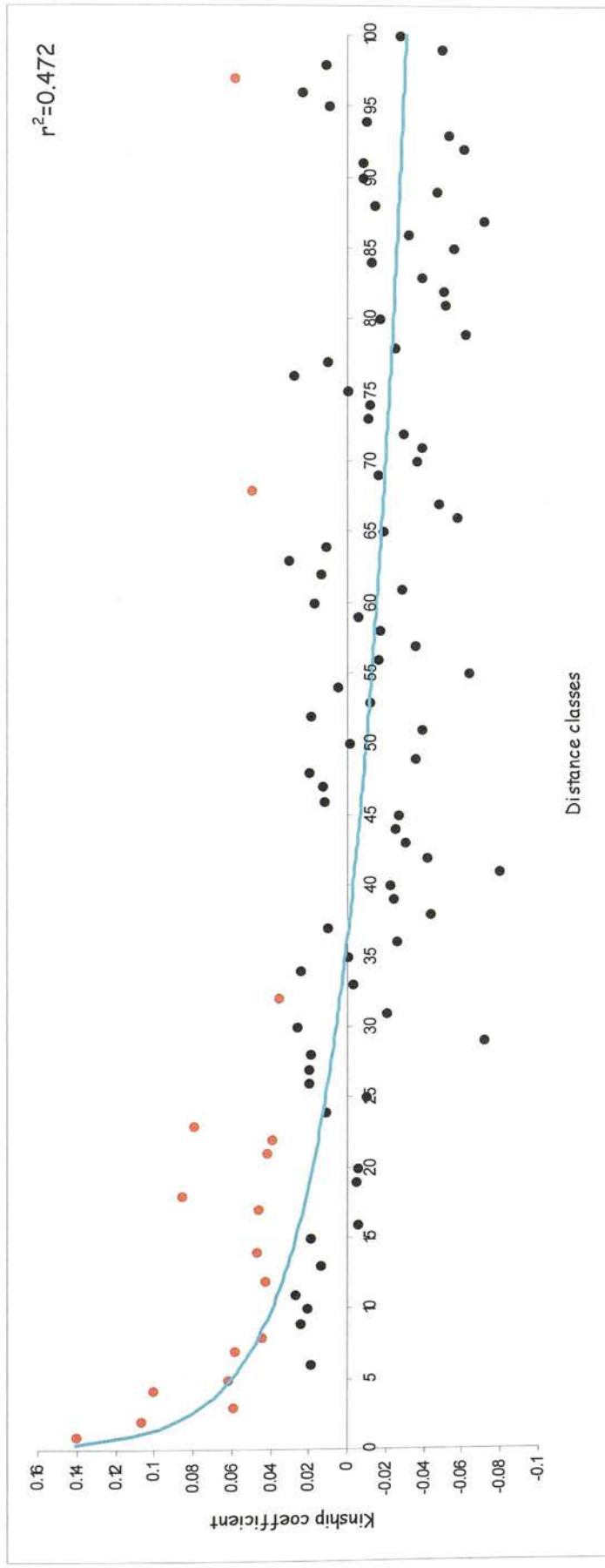


Figure 4.4 Correlogram of pairwise kinship coefficients in one population of *A. hemionitis* growing on a Man-made wall in Sintra (S14). Maximum distance between distance classes is indicated in Appendix II; red dots represent kinship coefficient values that are significantly different from zero (*P*-value<0.05); r^2 is the regression coefficient of determination.

4.4 Discussion

One of the main genetic effects of inbreeding within populations is to increase the levels of homozygosity relative to those expected under conditions of random mating. Estimates of the inbreeding coefficient (*Fis*) based on isozyme genotype frequencies revealed that, whereas outcrossing appears to be the predominant breeding system in populations of *A. hemionitis*, there is evidence for inbreeding. The amount of inbreeding is greater in some populations than others (positive *Fis* values vary from 0.005 to 0.450). The global inbreeding in *A. hemionitis* is 0.084, and it is within the range of other species of *Asplenium* that are considered to be outcrossing. Examples are the diploids *Asplenium ceterach* L. (*Fis*=0.048, J. Vogel, pers. comm.) and *Asplenium fontanum* (L.) Bernh. (*Fis*=0.090, H. Hunt, pers. comm.). Soltis & Soltis (1990a, 1992) reported several species of homosporous pteridophytes considered to be outbreeders with *Fis* values ranging from -0.141 to 0.163.

The breeding experiments did not show any evidence for an ability for intra-gametophytic selfing since none of the isolated gametophytes produced a sporophyte. Additionally, a direct count of completely homozygous individuals for all scored *loci* and over all populations showed a proportion of only 1.5%. These individuals could have been derived from intra- or inter-gametophytic selfing but also from crossing between closely related individuals. Although these results do not rule out the possibility of sporadic intra-gametophytic selfing events in nature, the main proportion of inbreeding verified must be a result of inter-gametophytic selfing. If inter-gametophytic selfing is the sole cause of inbreeding, then the over all inter-gametophytic selfing rate S [calculated as $S=2Fis/(1+Fis)$] in *A. hemionitis* would be 15.5%.

The mean *Fis* values differ between the thirteen 'islands' and are significantly positive in Sintra, Flores, Faial, São Miguel, Tenerife, La Palma and La Gomera. At a regional scale, Portugal and the archipelagos of Azores and Canary Islands show significantly positive *Fis* values. An increase in the rate of inbreeding can occur when related

gametophytes establish in close proximity due to limited spore dispersal. Variation in population rate of inbreeding may reflect varying ecological conditions among populations, such as sporophyte density and availability of suitable sites for spore germination and gametophyte establishment (Soltis & Soltis, 1990a). The availability of safe sites can be lower in patchy habitats, such as man-made walls and disturbed areas. Although *A. hemionitis* often grows in stable, climax communities (Canary Islands), it is also found in artificial and semi-natural habitats such as man-made walls (mainly Sintra), roadside slopes and verges (Azores) and abandoned cultivation terraces (e.g. Madeira and Morocco). The apparent population variability in mating system of this species may both reflect and allow its wide ecological amplitude and its great geographical range. In Sintra, where all populations grow on man-made walls, *A. hemionitis* has a global *Fis* that is the highest recorded for the species. However, mean *Fis* values for artificial, semi-natural and natural habitats are very similar (respectively, 0.067, 0.082 and 0.069) and other factors, like the range of spore dispersal, should be considered.

If limited spore dispersal is occurring, then it should lead to fine scale genetic structuring within populations. Evidence that this is the case comes from the assessment of spatial autocorrelation among individuals within a population of *A. hemionitis*. The analysis of relatedness of sporophytes growing on a Man-made wall in Sintra showed that individuals that grow more closely together are more related to each other than to the ones further away on the same wall. This substantial substructure of the sporophyte population can determine the structure of the gametophyte population. As a result, the probability of gametophytes growing close together being related is higher on a patchy habitat than on a homogeneous one. This aspect can increase the amount of inbreeding because of crossings between related gametophytes, and not only by selfing alone. This corresponds well with the significant *Fis* value observed in this population (0.164) and indicates a certain degree of isolation by distance within the population. This evidence suggests that most spores fall very close to the source sporophyte and dispersion is restricted. Peck *et al.* (1990) reports that the majority (>80%) of recorded spores disperse within 2m from the parent plant. These results are consistent with this observation, as shown by the approximate gene dispersal distance

calculated in this study of 22cm (which represents an area of about $0.152m^2$ around each individual). In conclusion, it seems that the amount of inbreeding found in *A. hemionitis* is greatly due to significant restriction of spore dispersal within populations and that the nature of the habitat may not be that significant.

Considering that *A. hemionitis* is mainly outcrossing in nature and any amount of inbreeding seems to derive from inter-gametophytic selfing, then there must be one or more effective mechanisms limiting intra-gametophytic selfing events. Variability in mating system among populations of homosporous pteridophytes can be influenced by a number of mechanisms affecting sexual expression in gametophytic populations (Soltis & Soltis, 1987a, b; Ranker, 1992). These include asynchronous maturation of gametangia on a single gametophyte, the action of male-inducing pheromones called antheridiogens and inbreeding depression caused by high levels of genetic load.

The ontogenic sequence of gametangial development determines whether gametophytes undergo a prolonged unisexual phase or whether they are predominantly bisexual in nature. A preliminary analysis of gametophyte development in *A. hemionitis* growing on Phytagel showed an extensive unisexual phase (males or females) of several months. However, the proportion of males and females is not known. After this phase, most gametophytes became bisexual, but nevertheless, isolated gametophytes were never seen to produce any sporophytes. This developmental sequence is consistent with observations from other outcrossing homosporous pteridophytes like *Gymnocarpium dryopteris* (L.) Newman subsp. *disjunctum* (Rupr.) Sarvela (Kirkpatrick *et al.*, 1990) and several species of *Pleopeltis* (Hooper & Haufler, 1997).

The hormone antheridiogen is considered a powerful mechanism promoting outcrossing in homosporous pteridophytes (Soltis & Soltis, 1990b). Antheridiogens are produced by maturing, fast growing gametophytes that become female, and are released on the surrounding substrate. Newly germinated gametophytes exposed to an antheridiogen-enriched substrate will become male only, whereas the more mature gametophytes become female, thus promoting inter-gametophytic mating events. This system is in accordance with the gametangial sequence of development referred to above

(Kirkpatrick *et al.*, 1990) and has been observed for several species of homosporous pteridophytes. Soltis & Soltis (1987a) stated that all species with an antheridiogen system (for which breeding system data were available) were outcrossers. Although antheridiogen influence has not been studied in *A. hemionitis*, this could be a potential mechanism regulating its breeding system.

Genetic load or inbreeding depression may also promote outcrossing in pteridophytes (Kirkpatrick *et al.*, 1990; Soltis & Soltis, 1992, Hooper & Haufler, 1997). It has been demonstrated in homosporous pteridophytes that the level of genetic load (the presence of deleterious recessive mutations) within a population can affect the relative success of selfing *versus* outcrossing mating events (Hooper & Haufler, 1997). Several species examined by Hedrick (1987) showed a bimodal distribution of genetic load. This author reports low levels of genetic load for inbreeding species and higher levels for species with a mainly outcrossing breeding system. The genetic load in pteridophytes in generally estimated by isolating gametophytes and counting the proportion that does not produce viable sporophytes in the laboratory (Hedrick, 1987). For a sporophyte to be produced by intra-gametophytic selfing, the gametophyte must be completely lethal-free. In the preliminary breeding experiments with isolated gametophytes, no sporophytes were produced. This might be indicative of high genetic load and inbreeding depression in *A. hemionitis*.

The contribution of different controlling factors for the breeding system of *A. hemionitis* may be complex, but the most important ones seem to be ecological conditions determining isolation within and between populations, structure of gametophyte populations and high genetic load. Although the inbreeding coefficient varies among populations, in many cases, the mating system of a single population cannot be viewed as representative of the species as a whole.

In conclusion, it has been shown in this study that *A. hemionitis* is a predominantly outcrossing fern which does not show any significant intra-gametophytic selfing. Inbreeding does occur as a consequence of inter-gametophytic selfing and mating between related gametophytes. Furthermore, it seems that spore dispersal within

populations is greatly restricted and gives rise to populations with significant genetic substructure. In this case, gene dispersal distance is about 22cm.

Measures of spore dispersal within populations have been shown to be limited. In the next chapter, the genetic structure among populations will be analysed. This will be done both within and between archipelagos, in order to determine the patterns of gene flow over small and large geographical scales. Ultimately, information on genetic structure can be valuable to deduce the colonisation history of *A. hemionitis*.

CHAPTER 5

ISOZYME DIVERSITY AND PATTERNS OF POPULATION DIFFERENTIATION IN *Asplenium hemionitis*

5.1 Introduction

This study focuses on the demographic processes that shaped the current genetic variation and its geographic distribution in the fern *A. hemionitis*.

The genetic variation present in a species has three main components: diversity (the amount of genetic variation), differentiation (the distribution of genetic variation) and distance (the amount of genetic variation between pairs of populations). Genetic variation is continually being created by mutation and at the same time eroded by selection and drift. Mutation rate is likely to be extremely low and estimated values will depend on the genome and marker investigated (Lowe *et al.*, 2004).

5.1.1 Genetic diversity

One of the main factors affecting genetic diversity is population size (Ellstrand & Elam, 1993; Widmer & Lexer, 2001). Large populations can maintain high levels of genetic diversity. The smaller the population, the more likely it is that random chance events change allele frequencies, a process called genetic drift. At its most extreme, drift can lead to extinction of alleles and loss of polymorphism such that a *locus* becomes fixed for a single allele (Hartl & Clark, 1997). Reductions in population size can occur through events such as colonisation of a new habitat by a small number of individuals (founder effect) or through processes such as habitat fragmentation where previously widespread populations become reduced (bottleneck). One of the expected outcomes of either founder effects or genetic bottlenecks is that the newly created population or the remaining population may have an allele frequency different from that of the source or original population (Lowe *et al.*, 2004). Theoretically, some degree of drift will always occur, unless the population is infinitely large, an extremely rare situation in nature. Molecular markers have been effectively used to detect reduction of diversity due to historical bottlenecks in the fern *Dryopteris cristata* (L.) A. Gray (Landergott *et al.*, 2001). In this investigation, a positive correlation was found between actual population

size and genetic diversity. Populations with recent historical bottlenecks showed a substantial and significant reduction in genetic variation compared with populations without bottlenecks; comparatively, high levels of genetic variation were still maintained in larger populations.

By examining the levels of genetic diversity in populations of *A. hemionitis* we can therefore make some inferences about the history of the populations throughout the geographic range. The level of variation detected will be dependent on the effective size of those populations over time. If these have remained large for a long period of time, high variability will be expected. However, if the populations have been founded from a limited number of spores, diversity will be reduced. Likewise, if populations have passed through recent bottlenecks, reduction in diversity will take place. Thus, by comparing genetic diversity levels in different areas of the range of *A. hemionitis*, it may be possible to identify those where populations have remained large over long periods of time, and distinguish these from areas that have suffered founder effects or recent bottlenecks.

A comparison of the number of private alleles may also be informative. Rare private alleles are likely to be lost most easily due to founder effects or bottlenecks. They contribute little to gene diversity measures however (Comps *et al.*, 2001; Widmer & Lexer, 2001). Thus, a comparison of private allele distribution across populations can give a further indication of which populations have retained large effective sizes over time.

5.1.2 Genetic structure

The movement of genes among populations is called gene flow. This is the major parameter that prevents populations from differentiating over time and can have profound impacts on the structure of genetic diversity (Lowe *et al.*, 2004). In seed plants, gene flow comprises the movement of gametes (pollen) and zygotes (seed). In pteridophytes, gene flow is mediated by dispersal of very small haploid spores (on

average 25-60μm) that can be easily transported by wind (Jermy, 1984). In order for gene flow to occur, spores must germinate, haploid gametophytes must establish and fertilisation must take place. The physiological and biotic requirements of spore germination and ecological conditions for gametophyte and sporophyte establishment determine where a species can grow and reproduce.

Two important factors affecting the dispersal capacity of spores are the distance that a viable spore can be carried by wind and the number of spores produced by the plant. There is ample evidence that pteridophyte spores can be transported by air currents for very long distances, as demonstrated by the extensive fern floras of isolated oceanic islands (Tryon, 1986). From the analysis of relationships of island fern floras, Tryon (1970) reported that 800km is only a slight barrier to the migration of pteridophytes, but dispersal across 1000km is a rare event.

Most pteridophytes produce large quantities of spores, on a scale of millions per year, and these are often viable for a year or longer. The dispersal curve is leptokurtic (Schneller & Schmid, 1982), the majority of recorded spores (>80%) dispersing only a short distance, within a radius of 25cm to 2m from the parent plant (Jermy, 1984; Peck *et al.*, 1990). However, if only a very small percentage of the annual production of a population is released for long-distance air transport, this is still a large number of spores (Jermy, 1984; Tryon, 1986). Tryon (1986) estimated an annual spore production for a pteridophyte with fronds c. 2-3cm long of about 100000. *A. hemionitis* fronds are sub-entire, on average 10-35 x 10-15cm and with a large area covered by spore-producing sori. There are no direct counts of the number of spores produced by *A. hemionitis*, but it would be reasonable to expect that the production of spores will be significantly higher (≈45x higher with c.4500000 spores per year). The size of these spores is on average 25μm long (Ormonde, 1996), at the lower range of average size for the majority of pteridophytes (25-60μm). The outer wall of *A. hemionitis* spores is ornamented with crests and perforations (Pangua & Prada, 1988; Ormonde, 1996) which might be an adaptation to movement by air currents. I would therefore anticipate that the dispersal capabilities of *A. hemionitis* are extremely high; this prediction may help interpret why this species has occupied the entire range of Macaronesia without

any speciation events. If no radiation within *A. hemionitis* has been detected, one can assume that there is genetic connectivity between populations maintained by a high level of gene flow. I have shown that there is restriction of gene flow within a population (Chapter 4), but as mentioned above, long-distance dispersal might still be effective due to production of large quantities of spores. To test these hypotheses we can use isozyme markers to measure the extent of genetic differentiation across the entire range.

An important feature of the distribution of *A. hemionitis* is that it occupies a series of archipelagos composed of different islands and occurs as a series of populations within islands. This provides an ideal opportunity to not only measure differentiation across the whole of the species range but also to estimate the hierarchical genetic structuring within islands, between islands within archipelagos, and between archipelagos. From these data it will be possible to determine whether restriction of gene flow among populations takes place principally among populations within islands, among islands within archipelagos, or among archipelagos.

5.1.3 Isolation by distance

In a situation where populations of a species are not completely panmictic (normally the case), there will be genetic differentiation over some spatial scale due to lack of gene flow. When there is lower gene flow between more distant populations, which consequently exhibit higher differentiation, this effect is termed isolation by distance (Lowe *et al.*, 2004).

If restricted spore flow is the only factor influencing the genetic structure of *A. hemionitis* populations, genetic differences between populations should follow an isolation by distance (IBD) model in which pairwise *Fst* values (a measure of genetic differentiation/distance) increase with the geographic distance. If restriction of gene flow is the same at all spatial scales, the IBD should be the same when short distances (within islands) are considered as when long distances (between archipelagos) are

analysed. This can be tested for by calculating IBD over different spatial scales. If IBD is not significant it is also informative to look for populations that do not conform to this pattern. This may indicate further barriers to spore flow that are not associated with geographic distance.

Therefore, to explore all these aspects, 105 populations were sampled in a hierarchical fashion. Data are analysed to determine patterns of variation in genetic diversity, the population structure over different scales, and the degree to which it conforms to an IBD model. Results are interpreted in light of population ecology and history of *A. hemionitis*. For example, low levels of diversity or lack of isolation by distance in Macaronesia might mean that the populations are the result of a recent colonisation. Conversely, high levels of genetic diversity might represent relictual populations. The relationship between these populations is further elucidated based on measures of genetic distance.

5.2 Materials and methods

5.2.1 Plant material and isozyme electrophoresis

A total of 105 populations of *A. hemionitis* comprising 2073 individuals were analysed for isozyme markers. The mean number of individuals scored per population was 20 (4–116). Data from thirteen enzyme systems encoded by 15 *loci* were used to estimate genetic diversity parameters and population differentiation. Material collections, enzymes and methods used in this analysis are indicated in detail in Chapter 3 (section 3.2).

5.2.2 Statistical analysis

In order to assess the genetic diversity and population structure at different spatial scales, populations were grouped according to geographical isolation from other populations. This grouping resulted in eleven islands (Flores, Corvo, Pico, Faial, São Miguel, Madeira, Tenerife, Gran Canaria, La Gomera, La Palma and El Hierro) and two 'mainland islands' (Sintra and Tanger). 'Islands' were further grouped into five regions. These were Portugal, Morocco, Azores, Madeira and Canary Islands.

5.2.2.1 *Genetic diversity*

Estimates of proportion of polymorphic *loci* at the 99% level (P_{99}; frequency of the commonest allele has to be ≤ 0.99 for a *locus* to be considered polymorphic), number of alleles per *locus* (*A*), number of alleles per polymorphic *locus* (*Ap*), Nei's (1973) gene diversity (*He*) and observed heterozygosity (*Ho*) were calculated for each population using GDA 1.0 (Lewis & Zaykin, 2001). Measures of allelic richness (*R*), which allow for correction in sample size, were calculated in FSTAT 2.9.3.2. using the rarefaction

method described by El Mousadik & Petit (1996). The same genetic diversity parameters were calculated over all *loci* at the 'island', region and species level.

5.2.2.2 *Genetic differentiation*

The genetic differentiation coefficient (Θ_n[1], Weir & Cockerham, 1984) was estimated over all *loci* within 'islands' and regions using FSTAT 2.9.3.2. Homogeneity of Θ_n values between *loci* was tested by jackknifing and confidence intervals generated by bootstrapping over all *loci*. Differentiation within 'islands' and regions was tested using a *G*-test (Goudet *et al.*, 1996) based on 10000 permutations of genotypes among populations. Pair-wise Θ_n for all 'islands' and regions were generated and tested for significance (not assuming HW equilibrium) in FSTAT 2.9.3.2.

To further assess the extent of geographic differentiation, hierarchical analyses based on Θ_n were also performed using GDA 1.0. Hierarchies were constructed according to the following population groupings:

Three-level hierarchies[2]

'Islands' (all populations grouped by 'islands') [[Sintra]:[Tanger]:[Madeira]:[Flores]:[Corvo]:[Pico]:[Faial]:[São Miguel]:[Tenerife]:[Gran Canaria]:[El Hierro]:[La Palma]:[La Gomera]]

Azores [[Flores]:[Corvo]:[Pico]:[Faial]:[São Miguel]]

Canary Islands [[Tenerife]:[Gran Canaria]:[El Hierro]:[La Palma]:[La Gomera]]

Four-level hierarchies[2]

All populations [[Mainland[Sintra]:[Tanger]]:[Azores[Flores]:[Corvo]:[Pico]:[Faial]:[São Miguel]]:[Canary Islands & Madeira[Madeira]:[Tenerife]:[Gran Canaria]:[El Hierro]:[La Palma]:[La Gomera]]]

[1] Θ_n is a notation for isozyme markers (nuclear) to distinguish from Θ_{cp} for chloroplast markers used in Chapter 6.

[2] Square brackets represent hierarchical levels. Ex: [Sintra] signifies all populations within Sintra.

Azores & Canary Islands [[Azores[Flores]:[Corvo]:[Pico]:[Faial]:[São Miguel]]: [Canary Islands [Tenerife]:[Gran Canaria]:[El Hierro]:[La Palma]:[La Gomera]]]

5.2.2.3 *Isolation by distance*

To test for isolation by distance, estimates of pair-wise genetic distance (Rousset, 1997) were correlated with pair-wise geographical distances using matrix correlation methods based on the Mantel test (10000 randomisations) in IBD 1.53 (Bohonak, 2002). Genetic distance estimates were obtained using the population differentiation coefficient Θ_n[3] (Weir & Cockerham, 1984) calculated from allele frequencies in FSTAT. Geographical distances between populations were calculated in Microsoft Excel from latitude and longitude coordinates using the method (Appendix IV) developed by Charles H. Pearson (http://www.cpearson.com/excel/latlong.htm, © Copyright 1997-2003).

The analysis was conducted for all populations (99) of *A. hemionitis* and all populations within the archipelagos of the Azores and within the Canary Islands. To assess isolation by distance at a smaller scale, one island within each archipelago was chosen for the analysis (São Miguel, Tenerife and Madeira). In this study, six populations were eliminated due to unavailability of coordinates (F1, F2, F6, P1, EH1 and LG1).

The strength of the isolation by distance relationship was determined with reduced major axis (RMA) regression and calculated in IBD 1.53.

5.2.2.4 *Genetic relationship among populations*

Genetic distances (Nei, 1978) among populations were calculated from the allele frequency data and used to construct a tree of genetic relationship using the neighbour-joining method in GDA 1.0.

[3] Θ_n substitutes Fst in Rousset's genetic distance Fst/(1-Fst)

5.3 Results

5.3.1 Genetic diversity

Isozyme electrophoresis of thirteen enzyme systems revealed 22 putative *loci* that were all polymorphic (AAT-1, AAT-2, ACN-1, ACN-2, DIA-1, DIA-2, HEX, IDH, LAP, MDH-1, MDH-2, MDH-3, MDH-4, 6-PGD, PGI-1, PGI-2, PGM-1, PGM-2, SkDH, TPI-1, TPI-2 and UGPP). Seven of these *loci* could not be reliably interpreted due to overlapping migration of alleles of different *loci* (AAT-2, DIA-2, MDH-3, MDH-4 and PGM-1) or inconsistent resolution (ACN-1 and PGI-1). The remaining fifteen *loci* were consistently scored for all populations.

Over all populations (2073 individuals), a total of 95 alleles were recorded (Table 5.5). Allele frequencies per *locus* for all populations and each 'island' are indicated in Appendix III (Tables I and II). The number of alleles at polymorphic *loci* varies between four (ACN-2, DIA-2, MDH-1 and TPI-1), five (HEX, LAP, TPI-2 and UGPP-2), six (AAT-1, IDH and 6-PGD), seven (PGM-2), eight (MDH-2), ten (PGI-2) and sixteen (SkDH). Alleles per *locus* and respective mobilities are indicated in Table 5.1.

5.3.1.1 *Levels of genetic diversity within populations*

In general, the populations of *A. hemionitis* show a high level of genetic diversity within populations. The data are summarised in Table 5.1. The mean proportion of polymorphic loci (P_{99}) is 0.648 with values ranging from 0.214 in a small population such as Sintra 6 (n=5) to 1.000 in a larger population such as El Hierro 7 (n=27). The mean of the mean number of alleles per *locus* (*A*) is 2.070 and per polymorphic *loci* (*Ap*) is 2.611. The lowest *A* and *Ap* values are, respectively, 1.214 in Sintra 6 (n=5) and 2.000 in Flores1 (n=10) and Madeira 7 (n=13). The highest *A* and *Ap* values are 3.467 and 4.083, respectively, both found in Tenerife 8 (n=49). The mean gene diversity (*He*) is 0.235 and ranges from 0.136 in Madeira 7 (n=13) to 0.353 in Madeira 10 (n=17) and El Hierro 7 (n=27). Considering the allelic richness (*R*), the lowest value is 1.086 in

Pico 2 (n=15) and the highest is 1.354 in Madeira 1 (n=5) and Madeira 10 (n=17); the mean value is 1.223.

Sintra and the Azorean islands exhibit the lowest average genetic diversity for individual populations (Table 5.2, values highlighted in blue). The highest values are found in the Canary Islands (Table 5.2, values highlighted in red). For example, the lowest average individual population gene diversity (*He*) is 0.164 in Flores (Azores) and the highest is 0.301 in La Gomera (Canary Islands). At a regional scale, Madeira and the Canary Islands are, on average, more diverse within individual populations than Portugal, Morocco and the Azores (Table 5.3).

5.3.1.2 *Total genetic diversity within' islands' and regions*

Total genetic diversity values for each 'island' are indicated in Table 5.2. Tenerife exhibits the highest P_{99}, *A*, *Ap* and *R* (respectively, 1.000, 5.133, 5.133 and 3.445) while La Gomera shows the highest *He* (0.319). Sintra exhibits the lowest P_{99} and *R* values (respectively, 0.600 and 2.057). Morocco has the lowest number of alleles per polymorphic *loci* (2.385) and Flores has the lowest gene diversity (0.174). Considering the regional scale (Table 5.3), all *loci* in the Canary Islands are polymorphic with a mean number of alleles per *locus* of 5.667. Sintra has the lowest percentage of polymorphic *loci* (60%) with an *Ap* of 2.333. In terms of gene diversity, once again the Canary Islands show the highest value (0.336) and Sintra shows the lowest value (0.214).

There is a clear trend of increasing genetic diversity with increasing number of sampled individuals (Table 5.4). The global gene diversity for *A. hemionitis* is 0.320 but it is still lower than the gene diversity in the Canary Islands (0.336).

Table 5.1 Genetic diversity within all populations of *A. hemionitis*. Genetic parameters shown are proportion of polymorphic *loci* (P_{99}), number of alleles per *locus* (*A*), number of alleles per polymorphic *locus* (*Ap*), gene diversity (*He*), observed heterozygosity (*Ho*) and allelic richness (*R*).

Region	Population	P_{99}	*A*	*Ap*	*He*	*Ho*	*R*
Sintra	Sintra 1	0.500	1.786	2.571	0.225	0.226	1.210
	Sintra 2	0.286	1.500	2.750	0.124	0.136	1.116
	Sintra 3	0.500	1.714	2.429	0.161	0.154	1.150
	Sintra 4	0.429	1.643	2.500	0.154	0.139	1.144
	Sintra 5	0.357	1.571	2.600	0.141	0.131	1.142
	Sintra 6	0.214	1.214	2.000	0.097	0.129	1.090
	Sintra 7	0.429	1.643	2.500	0.213	0.173	1.198
	Sintra 8	0.500	1.714	2.429	0.218	0.190	1.204
	Sintra 9	0.429	1.714	2.667	0.199	0.185	1.186
	Sintra 10	0.429	1.857	3.000	0.185	0.176	1.172
	Sintra 11	0.429	1.643	2.500	0.192	0.181	1.179
	Sintra 12	0.571	2.000	2.750	0.260	0.209	1.243
	Sintra 13	0.667	2.133	2.700	0.212	0.184	1.213
	Sintra 14	0.533	2.000	2.750	0.195	0.163	1.195
Azores	Flores 1	0.667	1.667	2.000	0.207	0.178	1.166
	Flores 2	0.400	1.533	2.333	0.108	0.100	1.111
	Flores 3	0.600	1.733	2.111	0.160	0.138	1.160
	Flores 4	0.600	1.667	2.111	0.190	0.169	1.197
	Flores 5	0.467	1.667	2.429	0.146	0.096	1.145
	Flores 6	0.467	1.600	2.286	0.171	0.135	1.171
	Corvo 1	0.667	1.867	2.300	0.236	0.238	1.237
	Corvo 2	0.533	1.800	2.500	0.205	0.160	1.205
	Pico 1	0.600	2.000	2.667	0.219	0.190	1.219
	Pico 2	0.417	1.583	2.400	0.108	0.122	1.086
	Pico 3	0.400	1.467	2.167	0.190	0.217	1.191
	Faial 1	0.800	2.000	2.250	0.133	0.103	1.133
	Faial 2	0.429	1.571	2.333	0.159	0.128	1.149
	Faial 3	0.643	1.857	2.333	0.174	0.182	1.163
	Faial 4	0.846	2.231	2.455	0.244	0.239	1.214
	Faial 5	0.714	1.929	2.300	0.290	0.289	1.271
	Faial 6	0.857	2.429	2.667	0.255	0.266	1.241
	Faial 7	0.600	2.000	2.667	0.201	0.218	1.205
	Faial 8	0.643	1.929	2.444	0.225	0.224	1.213
	Faial 9	0.714	1.857	2.200	0.202	0.181	1.189
	Faial 10	0.600	1.733	2.222	0.182	0.154	1.182
	Faial 11	0.600	2.000	2.667	0.211	0.187	1.215

Table 5.1 Continued.

Region	Population	P_{99}	*A*	*Ap*	*He*	*Ho*	*R*
Azores (Cont.)	Faial 12	0.929	2.214	2.308	0.231	0.205	1.218
	Faial 13	0.643	1.786	2.222	0.258	0.243	1.261
	São Miguel 1	0.667	2.000	2.500	0.207	0.248	1.221
	São Miguel 2	0.643	1.857	2.333	0.258	0.259	1.241
	São Miguel 3	0.400	1.533	2.333	0.189	0.213	1.198
	São Miguel 4	0.600	1.733	2.222	0.192	0.179	1.192
	São Miguel 5	0.600	2.067	2.778	0.219	0.205	1.218
	São Miguel 6	0.600	1.733	2.222	0.231	0.190	1.231
	São Miguel 7	0.714	2.071	2.500	0.219	0.183	1.207
	São Miguel 8	0.714	1.929	2.300	0.278	0.253	1.259
	São Miguel 9	0.667	1.867	2.300	0.218	0.219	1.218
	São Miguel 10	0.714	2.143	2.600	0.294	0.224	1.277
	São Miguel 11	0.600	1.933	2.556	0.252	0.199	1.259
	São Miguel 12	0.714	2.000	2.400	0.297	0.264	1.277
	São Miguel 13	0.667	2.067	2.600	0.212	0.191	1.213
	São Miguel 14	0.467	1.800	2.714	0.218	0.181	1.218
	São Miguel 15	0.533	1.800	2.500	0.237	0.223	1.239
Madeira	Madeira 1	0.714	1.929	2.300	0.379	0.400	1.354
	Madeira 2	0.538	1.846	2.571	0.258	0.264	1.224
	Madeira 3	0.867	3.200	3.462	0.285	0.270	1.285
	Madeira 4	0.600	1.933	2.556	0.257	0.200	1.257
	Madeira 5	0.667	2.000	2.500	0.217	0.215	1.217
	Madeira 6	0.600	2.000	2.667	0.229	0.245	1.240
	Madeira 7	0.667	1.667	2.000	0.136	0.144	1.136
	Madeira 8	0.733	2.200	2.636	0.276	0.233	1.276
	Madeira 9	0.733	2.267	2.727	0.287	0.257	1.287
	Madeira 10	0.733	2.467	3.000	0.353	0.348	1.354
	Madeira 11	0.733	2.267	2.727	0.282	0.272	1.284
	Madeira 12	0.667	2.000	2.500	0.301	0.320	1.301
	Madeira 13	0.800	2.467	2.833	0.275	0.246	1.275
	Madeira 14	0.533	1.600	2.125	0.264	0.283	1.264
Canary Islands	Tenerife 1	0.786	2.929	3.455	0.298	0.256	1.278
	Tenerife 2	0.643	2.071	2.667	0.254	0.250	1.252
	Tenerife 3	0.733	2.533	3.091	0.210	0.157	1.222
	Tenerife 4	0.800	2.267	2.583	0.262	0.258	1.262
	Tenerife 5	0.571	2.214	3.125	0.237	0.232	1.220
	Tenerife 6	0.857	2.500	2.750	0.261	0.243	1.244
	Tenerife 7	0.714	2.571	3.200	0.253	0.218	1.236
	Tenerife 8	0.800	3.467	4.083	0.257	0.219	1.257

Table 5.1 Continued.

Region	Population	P_{99}	*A*	*Ap*	*He*	*Ho*	*R*
Canary Islands (Cont.)	Tenerife 9	0.714	2.000	2.400	0.246	0.227	1.229
	Tenerife 10	0.800	2.800	3.250	0.222	0.193	1.222
	Tenerife 11	0.923	3.000	3.167	0.320	0.289	1.279
	Gran Canaria 1	0.867	2.800	3.077	0.289	0.304	1.291
	Gran Canaria 2	0.933	2.800	2.929	0.263	0.239	1.263
	Gran Canaria 3	0.786	2.643	3.000	0.214	0.241	1.210
	Gran Canaria 4	0.714	2.000	2.400	0.284	0.321	1.265
	Gran Canaria 5	0.643	1.857	2.333	0.268	0.263	1.251
	El Hierro 1	0.643	2.071	2.667	0.276	0.257	1.257
	El Hierro 2	0.462	1.538	2.167	0.200	0.162	1.173
	El Hierro 3	0.786	2.357	2.727	0.295	0.301	1.274
	El Hierro 4	0.500	1.583	2.167	0.206	0.236	1.165
	El Hierro 5	0.538	1.769	2.429	0.220	0.215	1.191
	El Hierro 6	0.583	1.750	2.286	0.247	0.254	1.198
	El Hierro 7	1.000	3.154	3.154	0.353	0.344	1.306
	El Hierro 8	0.857	2.000	2.167	0.321	0.340	1.300
	El Hierro 9	0.846	2.846	3.182	0.289	0.261	1.251
	La Palma 1	0.786	2.714	3.182	0.248	0.239	1.237
	La Palma 2	0.667	2.333	3.000	0.270	0.201	1.216
	La Palma 3	0.929	3.000	3.154	0.288	0.284	1.269
	La Palma 4	0.667	2.333	3.000	0.252	0.160	1.202
	La Palma 5	0.714	2.429	3.000	0.252	0.236	1.244
	La Palma 6	0.667	2.250	2.875	0.236	0.140	1.197
	La Palma 7	0.667	2.417	3.125	0.260	0.210	1.209
	La Gomera 1	0.786	2.786	3.273	0.250	0.226	1.233
	La Gomera 2	0.769	2.308	2.700	0.313	0.295	1.272
	La Gomera 3	0.692	2.231	2.778	0.328	0.284	1.285
	La Gomera 4	0.692	2.385	3.000	0.331	0.284	1.287
Morocco	Tanger 1	0.867	2.200	2.385	0.277	0.281	1.276
	Tanger 2	0.600	1.733	2.222	0.227	0.236	1.227
Over all 'within population' mean		**0.648**	**2.069**	**2.611**	**0.235**	**0.219**	**1.223**

Table 5.2 'Intra island' genetic diversity in *A. hemionitis*. Genetic parameters shown are proportion of polymorphic *loci* (P_{99}), number of alleles per *locus* (*A*), number of alleles per polymorphic *locus* (*Ap*), gene diversity (*He*) and allelic richness (*R*); standard error for *He* in brackets. Values in regular font refer to means across individual populations within 'islands'; values in bold refer to total diversity within 'islands' (merged populations); under each 'island' name it is indicated, in this order, number of populations, total number of individuals and mean number of individuals per population; lowest and highest values are highlighted, respectively, in blue and red.

Region	'Island'	P_{99}	*A*	*Ap*	*He*	*R*
Portugal	**Sintra**	0.448	1.724	2.582	0.184 (0.048)	1.175
	14/355/25	**0.600**	**2.533**	**3.333**	**0.214 (0.066)**	**2.057**
Morocco	**Tanger**	0.733	1.967	2.303	0.252 (0.065)	1.252
	2/50/25	**0.867**	**2.200**	**2.385**	**0.272 (0.070)**	**2.106**
Azores	**Flores**	0.533	1.644	2.212	0.164 (0.042)	1.158
	6/84/13	**0.733**	**2.200**	**2.636**	**0.174 (0.045)**	**2.074**
	Corvo	0.600	1.833	2.400	0.220 (0.057)	1.221
	2/23/11	**0.733**	**2.133**	**2.545**	**0.235 (0.061)**	**2.202**
	Pico	0.472	1.683	2.411	0.173 (0.045)	1.165
	3/49/16	**0.667**	**2.133**	**2.700**	**0.209 (0.054)**	**2.123**
	Faial	0.694	1.964	2.390	0.213 (0.055)	1.204
	13/270/20	**0.800**	**3.133**	**3.417**	**0.218 (0.056)**	**2.348**
	São Miguel	0.620	1.902	2.457	0.235 (0.061)	1.231
	15/285/18	**0.667**	**2.867**	**3.600**	**0.252 (0.065)**	**2.361**
Madeira	**Madeira**	0.685	2.132	2.615	0.271 (0.070)	1.268
	14/231/16	**0.933**	**3.867**	**4.000**	**0.295 (0.076)**	**2.824**
Canary Islands	**Tenerife**	0.758	2.577	3.070	0.256 (0.066)	1.246
	11/230/20	**1.000**	**5.133**	**5.133**	**0.287 (0.074)**	**3.445**
	Gran Canaria	0.789	2.420	2.748	0.263 (0.068)	1.256
	5/118/23	**0.933**	**3.933**	**4.071**	**0.278 (0.072)**	**2.842**
	El Hierro	0.619	2.007	2.559	0.240 (0.062)	1.233
	9/130/12	**0.933**	**4.000**	**4.214**	**0.316 (0.082)**	**3.275**
	La Palma	0.720	2.480	3.048	0.256 (0.066)	1.225
	7/162/21	**0.800**	**3.600**	**4.083**	**0.256 (0.066)**	**2.836**
	La Gomera	0.722	2.397	2.938	0.301 (0.078)	1.269
	4/89/21	**0.933**	**3.533**	**3.714**	**0.319 (0.082)**	**2.977**

Table 5.3 Genetic diversity in the five main regions of *A. hemionitis* distribution. Genetic parameters shown are proportion of polymorphic *loci* (P_{99}), number of alleles per *locus* (*A*), number of alleles per polymorphic *locus* (*Ap*), gene diversity (*He*) and allelic richness (*R*); standard error for *He* in brackets. Values in regular font refer to means across individual populations within regions; values in bold refer to total diversity within regions (merged populations); under each region name it is indicated, in this order, number of populations, total number of individuals and mean number of individuals per population; lowest and highest values are highlighted, respectively, in blue and red.

Region	P_{99}	*A*	*Ap*	*He*
Portugal	0.448	1.724	2.582	0.184 (0.048)
14/355/25	**0.600**	**2.533**	**2.333**	**0.214 (0.055)**
Morocco	0.733	1.967	2.303	0.252 (0.065)
2/50/25	**0.867**	**2.200**	**2.385**	**0.272 (0.062)**
Azores	0.619	1.863	2.391	0.211 (0.054)
39/711/17	**0.800**	**3.667**	**4.000**	**0.239 (0.070)**
Madeira	0.685	2.132	2.615	0.271 (0.70)
14/231/16	**0.933**	**3.867**	**4.000**	**0.295 (0.076)**
Canary Islands	0.737	2.411	2.879	0.266 (0.069)
36/729/19	**1.000**	**5.667**	**5.667**	**0.336 (0.087)**

Table 5.4 Total genetic diversity within increasing scales of sampling in *A. hemionitis*. Genetic parameters shown are proportion of polymorphic *loci* (P_{99}), number of alleles per *locus* (*A*), number of alleles per polymorphic *locus* (*Ap*), gene diversity (*He*) and allelic richness (*R*); standard error for *He* in brackets; n is the mean number of individuals per sampling scale.

Sampling scales	n	P_{99}	*A*	*Ap*	*He*
Populations	20	0.648	2.069	2.611	0.235 (0.061)
'Islands'	148	0.830	3.214	3.520	0.261 (0.067)
Regions	381	0.840	3.573	3.864	0.271 (0.070)
Global	2073	1.000	6.333	6.333	0.320 (0.083)

5.3.2 Geographic distribution of variation and population differentiation

5.3.2.1 *Distribution of alleles*

From a total of 95 alleles recorded in *A. hemionitis*, 40% are present in Portugal, 35.7% in Morocco, 61.1% in Madeira, 57.9% in the Azores and 85.3% in the Canary Islands (Table 5.5). Tenerife contains more alleles than any other 'island' (81.1% of the total). The 'islands' with fewer alleles are Flores, Corvo, Pico (Azores) and Tanger (average 34.2%), but these are also the ones where fewer populations were sampled.

The highest number of private alleles (Table 5.5) is found in the Canary Islands (20.0% of the species total). The Azores and Madeira contain much fewer but equal numbers of private alleles (4.2%). On the mainland, Sintra contains three private alleles (3.2%) and Morocco contains none. When considering the individual 'islands', La Palma and La Gomera (Canary Islands), and the islands of the Azores do not contain any private alleles. The Canary Islands, more specifically Tenerife, are clearly the centre of diversity in terms of alleles. With the exception of the alleles MDH-1 94 and MDH-2 122 in the Azores (frequency of 0.098 and 0.128, respectively), all other private alleles occur at very low frequencies (≤ 0.05). Frequency of all private alleles is indicated in Table 5.6.

5.3.2.2 *Intra and inter-'island' differentiation*

Estimates of Θ_n range from 0.040 in La Palma to 0.180 in La Gomera (Table 5.7). For all 'islands', Θ_n is not high but it is significant (*P*-values<0.01). In general, populations within 'islands' were differentiated to a similar extent. These results show that the majority of genetic variation is found within populations.

Pair-wise Θ_n among all 'islands' (Table 5.10) show that Madeira diverges highly from the other 'islands'. Estimates range from 0.403 (Portugal) to 0.292 (La Gomera). The less divergent 'islands' are within the archipelagos of the Azores and Canary Islands.

5.3.2.3 *Intra and inter-regional differentiation*

Within the archipelagos of the Azores and Canary Islands, Θ_n is respectively, 0.149 and 0.185 (Table 5.8). Both estimates are significant (*P*-value<0.001) and higher than in the other regions, but they comprise many more populations that are distributed in different islands and more far apart.

Pair-wise Θ_n among all regions (Table 5.11) shows that Madeira is highly divergent from all the other regions but more from the mainland. The less divergent regions are the Canary Islands and Portugal, and the former and the Azores.

5.3.2.4 *Hierarchical analysis of differentiation*

Table 5.9 summarises estimations of differentiation coefficients for different hierarchies of geographic distribution. Within the Azores, 6.4% (Θ_nP=0.064) of interpopulation variation is partitioned between the islands while 16.7% (Θ_nS=0.167) is partitioned between populations within islands. Within the Canary Islands, the same estimates are, respectively 12.7% (Θ_nP=0.127) and 20.8% (Θ_nS=0.208). Once again, this is evidence for higher differentiation within the Canary Islands than within the Azores. A hierarchical analysis grouping all populations within 'islands' shows that 21.5% (Θ_nP=0.215) of variation is partitioned between 'islands' and 29.1% (Θ_nS=0.291) is partitioned between populations within 'islands'. A four-level hierarchy including only the archipelagos of the Azores and Canary Islands shows that only 6.4% (Θ_nP=0.064) of variation is partitioned between these two regions, while 24.6% (Θ_nSS=0.246) is partitioned between populations within islands. A four-level hierarchy including all populations, grouped according to the degree of proximity between archipelagos and the two continental regions, shows that only 3.4% (Θ_nP=0.034) of variation is partitioned between regions. This value increases greatly to 22.3% (Θ_nS=0.223) between 'islands' within regions and 29.8% (Θ_nSS=0.298) between populations within 'islands' within regions. Over all, it seems that interpopulation variation is largely due to differences between populations within each local area rather than due to geographic structuring between regions.

Table 5.5 Geographical distribution of isozyme alleles scored for *A. hemionitis*. Presence of alleles is indicated with the symbol +; yellow, violet and blue shaded areas represent alleles private to, respectively, individual 'islands', the Azores and Canary Islands; ns is not scored; number of alleles per *locus* in brackets.

				Azores						Canary Islands				
loci	allele mobility	Sintra	Tanger	Flores	Corvo	Pico	Faial	São Miguel	Madeira	Tenerife	Gran Canaria	El Hierro	La Palma	La Gomera
AAT-1	74	+		+	+	+	+	+	+	+	+	+	+	+
(6)	88								+					
	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	115											+		
	118		+				+	+	+	+	+	+	+	+
	124						+	+		+				
ACN-2	78								+		+			
(4)	91	+	+							+	+	ns	ns	ns
	100	+	+	+	+	+	+	+	+	+	+			
	110								+	+				
DIA-1	83			+			+	+	+	+	+	+	+	+
(4)	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	123	+	+	+	+	+	+	+	+	+	+	+	+	+
	141								+					
HEX	77	+												
(5)	100	+	+	+	+	+	+	+		+	+	+	+	+
	111	+						+	+			+	+	+
	118	+		+	+	+	+	+	+	+	+	+	+	+
	126	+												
IDH	57								+					
(6)	77								+					
	91	+												
	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	115	+	+	+	+	+	+	+	+	+	+	+	+	+
	127	+	+						+	+	+	+		+
LAP	93								+	+				
(5)	97						+		+	+	+	+	+	+
	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	102	+	+				+		+	+	+	+		
	105	+			+					+	+	+	+	
MDH-1	87		+						+	+	+			
(4)	94			+	+	+	+	+						
	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	109				+		+			+	+	+	+	+
MDH-2	58	+				+			+	+				
(8)	83							+	+	+	+			
	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	104						+	+						
	106		+	+		+	+			+	+		+	
	108						+	+					+	
	114		+	+	+	+	+	+	+	+	+	+	+	+
	122			+	+	+	+	+						
6-PGD	71	+							+	+	+	+	+	+
(6)	88			+	+	+	+	+	+	+		+	+	+
	95	+			+	+	+		+	+	+			
	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	107	+	+		+	+	+	+	+	+	+	+	+	+
	110							+	+	+				+

Table 5.5 Continued.

loci	allele mobility	Sintra	Tanger	Azores					Madeira	Canary Islands				
				Flores	Corvo	Pico	Faial	São Miguel		Tenerife	Gran Canaria	El Hierro	La Palma	La Gomera
PGI-2	20									+				
(10)	35									+	+		+	
	60		+					+	+	+	+	+	+	+
	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	115								+	+	+	+		+
	125						+	+	+	+				
	130		+	+			+	+	+	+	+	+	+	+
	150									+		+		+
	155							+	+	+	+	+		+
	185								+	+				
PGM-2	56									+	+	+	+	+
(7)	84									+	+	+		
	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	119		+				+		+	+	+	+	+	+
	122				+		+							
	129									+		+		
	134									+				
SkDH	43									+				
(16)	52									+			+	
	59	+	+						+	+	+	+	+	+
	60									+	+	+	+	
	69	+								+	+	+	+	+
	72									+	+	+	+	+
	76	+							+	+	+	+	+	+
	83	+	+				+	+		+	+	+	+	+
	86									+			+	+
	90	+		+	+	+	+	+	+	+	+	+	+	+
	96			+	+	+	+	+		+		+	+	+
	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	104								+	+	+		+	
	110			+	+	+	+	+	+	+	+	+	+	+
	117									+		+	+	
	124									+				+
TPI-1	81						+		+	+	+	+	+	+
(4)	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	114	+	+						+	+	+	+	+	
	118											+		
TPI-2	65											+		
(5)	69		+	+			+		+	+	+	+	+	+
	87									+				
	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	123	+	+	+		+	+	+	+	+	+	+	+	+
UGPP	79			+						+	+			
(5)	92			+	+		+	+	+	+		+	+	+
	100	+	+	+	+	+	+	+	+	+	+	+	+	+
	105										+			
	108					+	+	+	+	+	+	+		+
Total	95	38	33	33	32	32	47	43	58	77	59	59	54	53
				Azores: 55						Canary Islands: 81				
Private alleles		3	0	0	0	0	0	0	4	4	1	3	0	0
				Azores: 4						Canary Islands: 19				

Table 5.6 Frequency of the private alleles present in the different regions of *A. hemionitis* range. Names of the alleles according to Table 5.5.

Region	Private alleles	Frequency
Portugal	HEX 77	0.017
	HEX 126	0.016
	IDH 91	0.051
Azores	MDH-1 94	0.098
	MDH-2 104	0.024
	MDH-2 122	0.128
	PGM-2 122	0.005
Madeira	AAT-1 188	0.018
	IDH 57	0.007
	IDH 77	0.013
	DIA-1 141	0.030
Canary Islands	AAT-1 115	0.001
	PGI-2 20	0.001
	PGI-2 35	0.002
	PGI-2 150	0.003
	PGM-2 56	0.034
	PGM-2 84	0.016
	PGM-2 129	0.006
	PGM-2 134	0.001
	SkDH 43	0.002
	SkDH 52	0.004
	SkDH 60	0.008
	SkDH 72	0.031
	SkDH 86	0.003
	SkDH 117	0.009
	SkDH 124	0.012
	TPI-1 118	0.014
	TPI-2 65	0.003
	TPI-2 87	0.001
	UGPP 105	0.001

Table 5.7 Mean genetic differentiation coefficient (Θ_n) estimated for populations of *A. hemionitis* grouped by 'island'. Test of population differentiation not assuming random mating: * *P*-value<0.01, ** *P*-value<0.001.

Region	'Island'	Θ_n	Confidence intervals (95% level)
Portugal	Sintra	0.086**	0.055 – 0.113
Morocco	Tanger	0.067**	0.030 – 0.100
Azores	Flores	0.151**	0.045 – 0.223
	Corvo	0.096**	0.020 – 0.187
	Pico	0.109*	0.000 – 0.195
	Faial	0.102**	0.077 – 0.122
	São Miguel	0.110**	0.072 – 0.139
Madeira	Madeira	0.087**	0.075 – 0.105
Canary Islands	Tenerife	0.076**	0.042 – 0.119
	Gran Canaria	0.072*	0.037 – 0.104
	El Hierro	0.152**	0.087 – 0.222
	La Palma	0.040*	0.021 – 0.060
	La Gomera	0.180**	0.079 – 0.313

Table 5.8 Mean genetic differentiation coefficient (Θ_n) estimated for populations of *A. hemionitis* grouped by region. Also shown is the global Θ_n for the species (obtained over all populations). Test of population differentiation not assuming random mating: * *P*-value<0.001.

Region	Θ_n	Confidence intervals (95% level)
Portugal	0.086*	0.055 – 0.113
Morocco	0.067*	0.030 – 0.100
Azores	0.149*	0.119 – 0.177
Madeira	0.087*	0.075 – 0.105
Canary Islands	0.185*	0.095 – 0.339
Global	0.276*	0.198– 0.387

Table 5.9 Hierarchical analysis of population structure for different geographic units in *A. hemionitis* sampling range. $\Theta_n P$, $\Theta_n S$ and $\Theta_n SS$ are Weir & Cockerham's (1984) estimators of Wright's (1965) F-statistics ($\Theta_n SS$ is only estimated for 4-level hierarchies); 95% confidence intervals are shown in brackets; n pops is the number of populations in each hierarchy.

Hierarchies	n pops	$\Theta_n P$	$\Theta_n S$	$\Theta_n SS$
3-level hierarchy[1]				
'Islands'	105	**0.215** (0.127 – 0.329)	**0.291** (0.208 – 0.403)	—
Azores	39	**0.064** (0.014 – 0.124)	**0.167** (0.127 – 0.208)	—
Canary Islands	36	**0.127** (0.034 – 0.276)	**0.208** (0.103 – 0.374)	—
4-level hierarchy[2]				
All populations	105	**0.034** (-0.010 – 0.072)	**0.223** (0.137 – 0.331)	**0.298**(0.215 – 0.405)
Azores & Canary Islands	75	**0.064** (0.021 – 0.110)	**0.162** (0.073 – 0.266)	**0.246** (0.157 – 0.352)

[1] **'islands'** [[Sintra]:[Tanger]:[Madeira]:[Flores]:[Corvo]:[Pico]:[Faial]:[São Miguel]:[Tenerife]:[Gran Canaria]:[El Hierro]:[La Palma]:[La Gomera]]
Azores [[Flores]:[Corvo]:[Pico]:[Faial]:[São Miguel]]
Canary Islands [[Tenerife]:[Gran Canaria]:[El Hierro]:[La Palma]:[La Gomera]]

[2] **All populations** [[mainland[Sintra]:[Tanger]]:[Azores[Flores]:[Corvo]:[Pico]:[Faial]:[São Miguel]]:[Canary Islands & Madeira[Madeira]:[Tenerife]:[Gran Canaria]:[El Hierro]:[La Palma]:[La Gomera]]]
Azores & Canary Islands [[Azores[Flores]:[Corvo]:[Pico]:[Faial]:[São Miguel]]:[Canary Islands [Tenerife]:[Gran Canaria]:[El Hierro]:[La Palma]:[La Gomera]]]

Table 5.10 Pair-wise Θ_n for all 'islands' in *A. hemionitis* sampling range. Lowest and highest values are highlighted, respectively, in blue and red. All estimations are significant at the 5% level (Bonferroni adjusted *P*-value=0.00064), except values marked *ns*.

	Sintra	Flores	Corvo	Pico	Faial	São Miguel	Tenerife	Gran Canaria	El Hierro	La Palma	La Gomera	Tanger	Madeira
Sintra	0												
Flores	0.298	0											
Corvo	0.254	0.080	0										
Pico	0.301	0.035	0.071	0									
Faial	0.215	0.035	0.063	0.047	0								
São Miguel	0.152	0.104	0.052	0.108	0.090	0							
Tenerife	0.128	0.266	0.203	0.275	0.234	0.139	0						
Gran Canaria	0.174	0.279	0.217	0.278	0.231	0.153	0.074	0					
El Hierro	0.147	0.156	0.131	0.164	0.120	0.138	0.113	0.148	0				
La Palma	0.219	0.126	0.128	0.137	0.096	0.164	0.198	0.229	0.038^{ns}	0			
La Gomera	0.260	0.154	0.139	0.164	0.140	0.161	0.199	0.240	0.060	0.071^{ns}	0		
Tanger	0.208	0.367	0.320	0.355	0.280	0.237	0.141	0.079	0.171	0.271	0.276	0	
Madeira	0.403	0.385	0.365	0.388	0.387	0.331	0.352	0.388	0.301	0.369	0.292	0.371	0

Table 5.11 Pair-wise Θ_n for all regions in *A. hemionitis* sampling range. Lowest and highest values are highlighted, respectively, in blue and red. All estimations are significant at the 5% level (Bonferroni adjusted *P*-value=0.005).

	Portugal	Azores	Canary Islands	Morocco	Madeira
Portugal	0				
Azores	0.178	0			
Canary Islands	0.090	0.093	0		
Morocco	0.202	0.292	0.137	0	
Madeira	0.396	0.361	0.296	0.388	0

5.3.3 Isolation by distance (IBD)

Significant IBD (Mantel test, r =0.365, *P*-value<0.01) was found for the entire population data set (Figure 5.1), although the regression analysis only explains 13.3% of the variance in *A. hemionitis* distribution range (as indicated by the regression trendline coefficient of determination r^2). Due to the fact that Madeira shows a high genetic distance from all the other regions, the IBD analysis was repeated excluding all populations from this region. The results indicate a slightly greater IBD (Mantel test, r =0.431, *P*-value<0.001), consistent with the larger distances between the nearest archipelagos.

Within the archipelagos of the Azores and Canary Islands (Figure 5.2), the relationship between genetic and geographical distance is stronger than for the whole set of populations, with the regression explaining, respectively, 16.7% and 19.8% of the variance among populations. In both archipelagos, IBD is significant as shown by the Mantel test and *P*-value<0.0001 (r =0.408 for the Azores and 0.445 for the Canary Islands).

At a smaller geographical scale (within 'islands'), there is still significant IBD. In Tenerife (Canary Islands), 25.8% of the variance is explained by the regression analysis, while in São Miguel (Azores) it is only 14.6%. In Madeira, the strength of the

correlation between genetic and geographic distance is even lower (r^2=0.12). Significance of IBD is indicated by the Mantel test with *P*-value<0.002, 0.0007 and 0.021 for Tenerife, São Miguel and Madeira, respectively. Regression analyses are illustrated in Figure 5.3. In Sintra, which represents an even smaller geographic scale (up to 1.2km), no IBD was detected ($r^2=2x10^{-6}$).

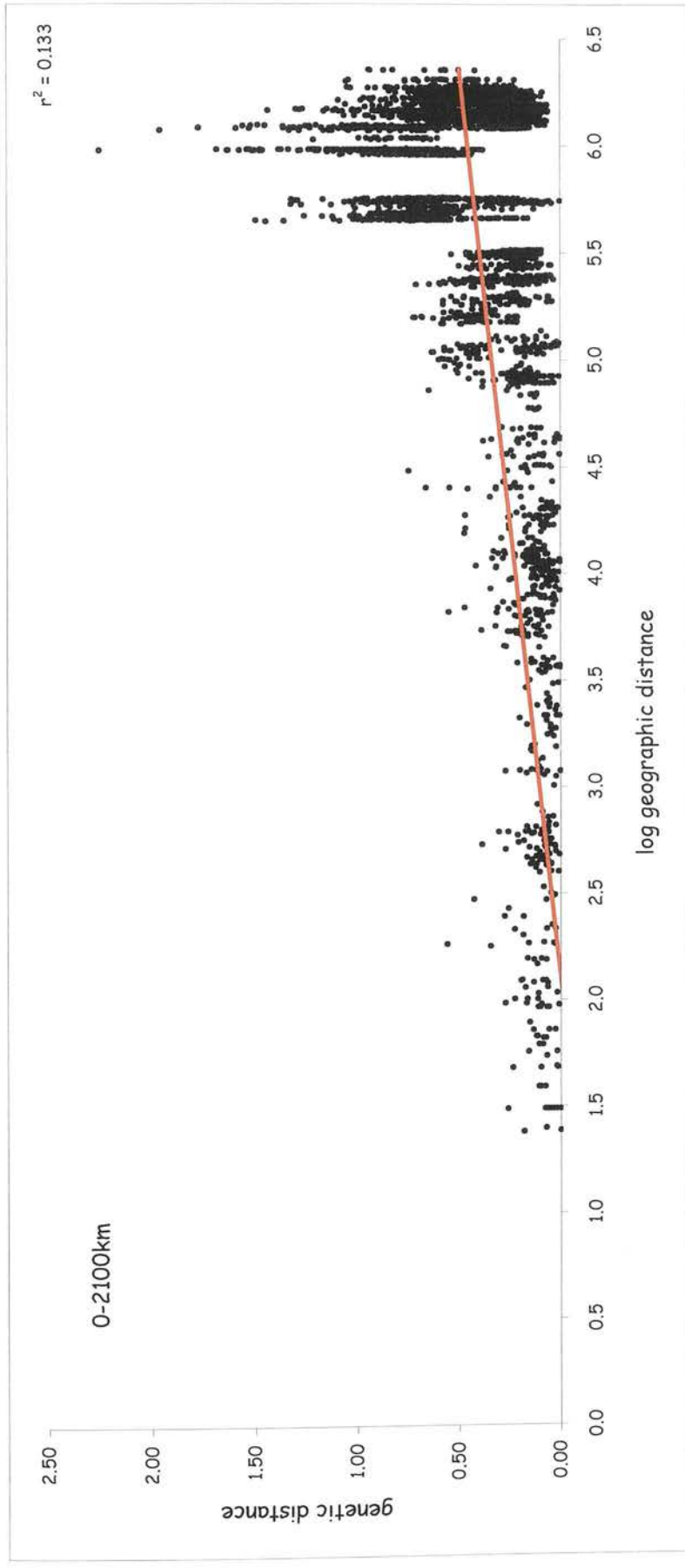
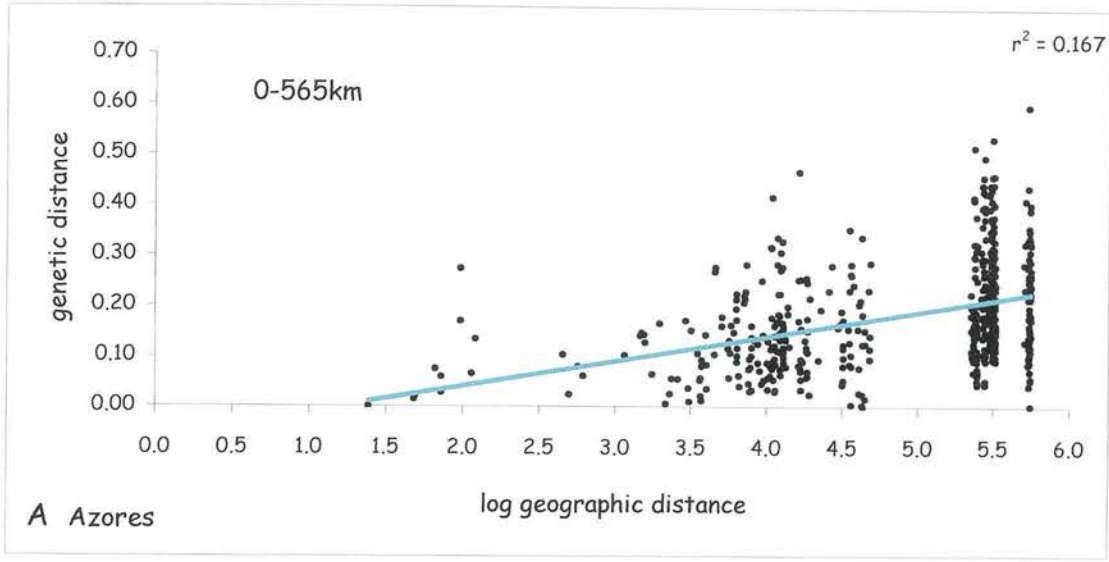


Figure 5.1 Correlation of pair-wise genetic and geographical distance for all *A. hemionitis* populations. Also indicated is the regression coefficient of determination (r^2) and the geographical distance range.



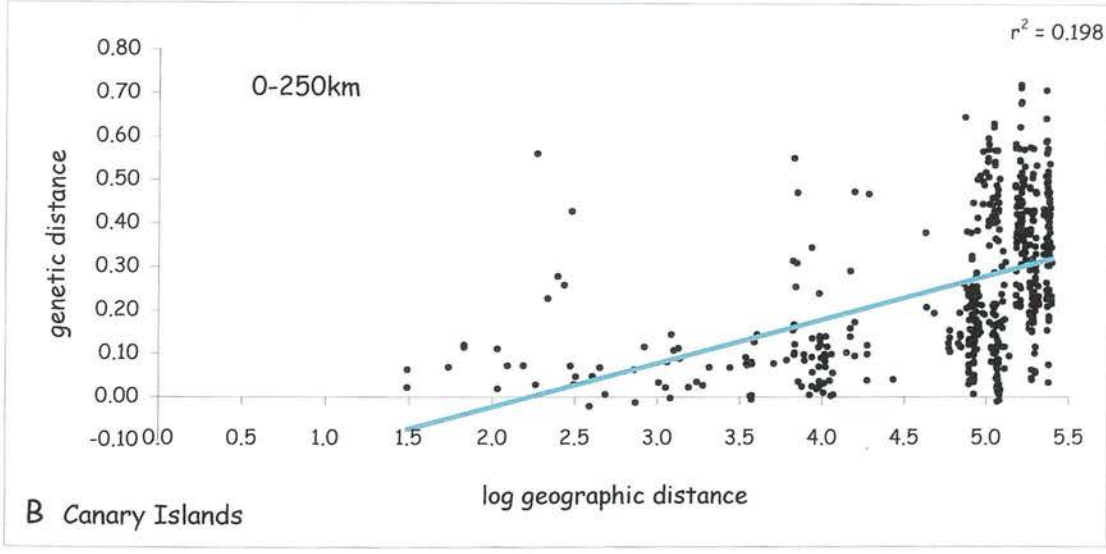
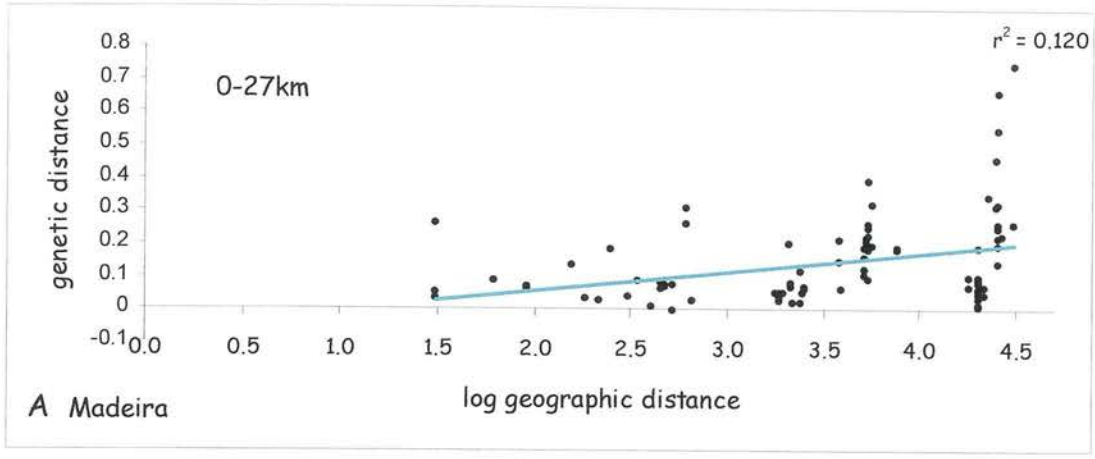
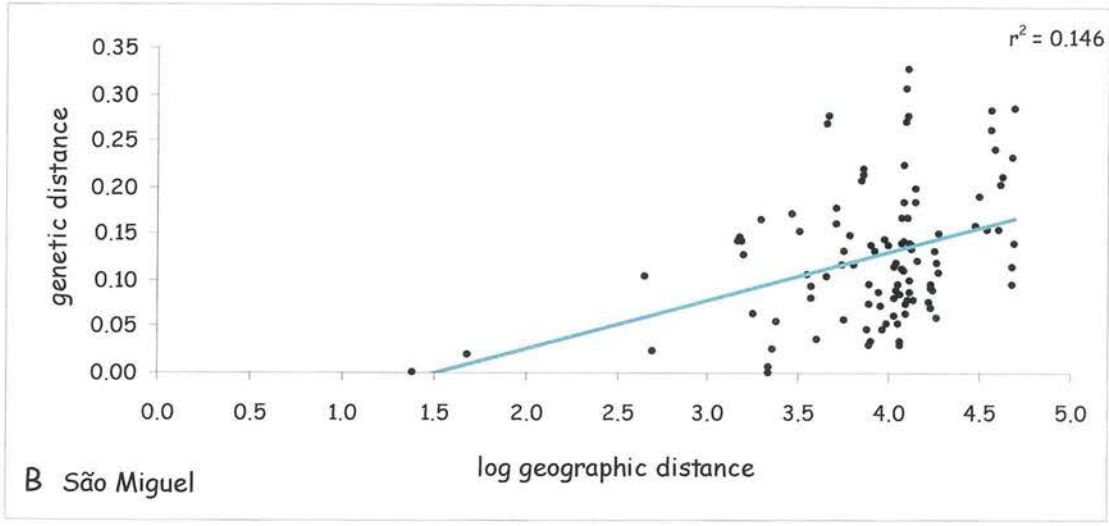


Figure 5.2 Correlation of pair-wise genetic and geographical distance for *A. hemionitis* populations within the archipelagos of the Azores (A) and Canary Islands (B). Also indicated is the regression coefficient of determination (r^2) and the geographical distance range. Note the difference in geographical scale for each archipelago.





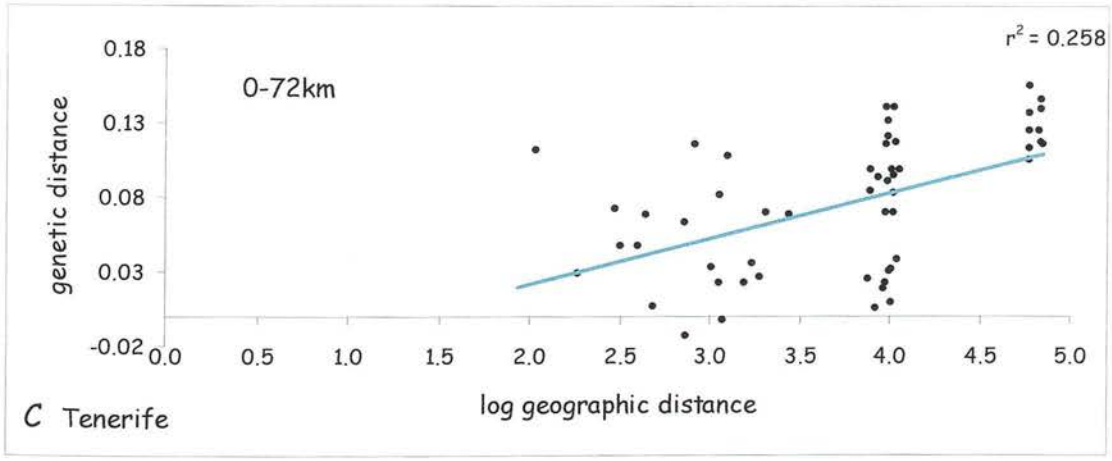


Figure 5.3 Correlation of pair-wise genetic and geographical distance for *A. hemionitis* populations within the islands of Madeira (A), São Miguel (B) and Tenerife (C). Also indicated is the regression coefficient of determination (r^2) and the geographical distance range. Note the difference in genetic and geographical distance scales for each island.

5.3.4 Genetic relationship between populations

The neighbour-joining analysis based on Nei's (1978) genetic distance revealed three major clusters (Figure 5.4). The left cluster (in green) includes all populations from Madeira, four from the Canary Islands (LG2, LG4, EH2 and EH5) and one from the Azores (F1). The bottom cluster (in pink) includes the populations from Morocco, Portugal, Tenerife, Gran Canaria and one from El Hierro (EH6). A third cluster (in blue) includes the remaining populations from the Azores and a subcluster (dashed line) with the populations from La Palma, El Hierro and La Gomera. This subcluster also includes the most western population in Tenerife (T11).

The most significant outcome from the neighbour-joining tree is the differentiation of Madeira and the Azores, and clustering of the Canary Islands with the mainland. Also significant is the clustering of the three most western Canary Islands (La Palma, El Hierro and La Gomera) within the Azores.

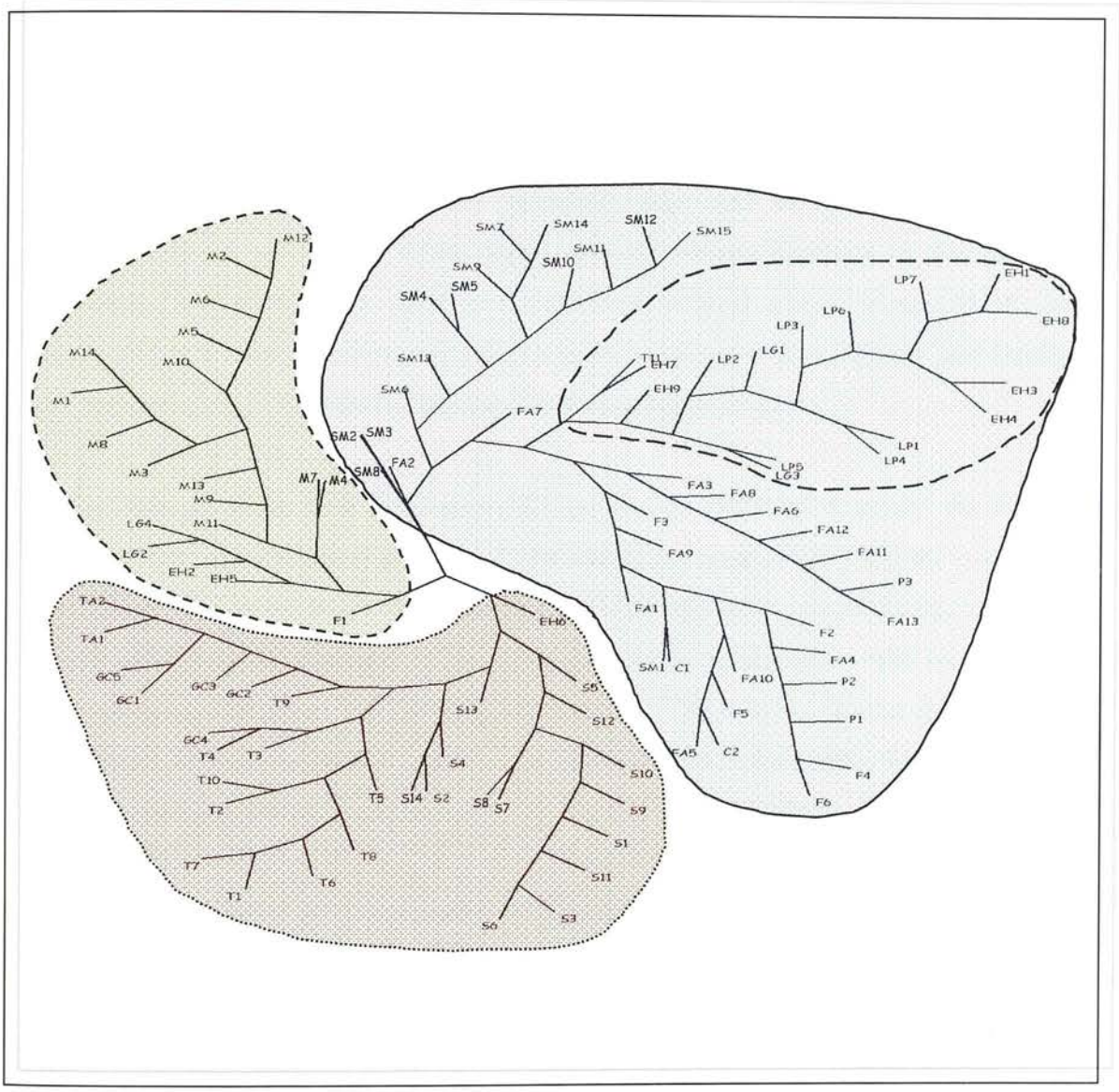


Figure 5.4 Neighbour-joining tree illustrating the relationship among 105 populations of *A. hemionitis*. Length of branches is not proportional to genetic distance between populations; abbreviations of populations according to table 3.2 (Chapter 3); green, blue and pink shaded areas represent three distinctive population clusters (Madeira; Azores and W Canary Islands; and central Canary Islands, Morocco and Sintra, respectively); dashed line within the blue area represents a subcluster including populations from La Palma, El Hierro and La Gomera (Canary Islands) within the Azores cluster.

5.4 Discussion

5.4.1 Genetic diversity

This study revealed that *A. hemionitis* is highly variable for isozyme markers. Allelic variation is present throughout the sampling range and all populations showed intrapopulational variation. When considering the genetic diversity of the fern *A. hemionitis*, it is necessary to consider a number of different scales. Firstly, it will be considered at a level within populations, secondly at a level within 'islands' and finally at a regional scale, i.e. within archipelagos and continental areas.

At the 'island' scale, both for mean within population and total diversity, the lowest gene diversity (*He*) is found in Sintra (0.184 and 0.214, respectively) and the islands of the Azores, in particular the islands of Flores (0.164 and 0.174, respectively) and Pico (0.173 and 0.209, respectively). This is also reflected at the regional scale, but between these regions, Sintra has the lowest gene diversity. The trend is similar for the other parameters of genetic diversity, like mean proportion of polymorphic *loci* (*P*), mean number of alleles per *locus* (*A*) and polymorphic *loci* (*Ap*), and allelic richness (*R*).

There are a number of reasons why the lowest levels of genetic diversity may be recorded in Sintra. Sintra has the smallest spatial scale of any area in this study. Populations may be separated by a maximum distance of up to 1.2km, compared with typical values of hundreds of km for the other populations. This reduced spatial scale will mean that there is less potential for establishment of individuals from distant populations. Also, the diversity may be reduced by a decline in total number of populations in the region, i.e. Sintra comprises fewer individuals than the other areas when considering it as a region. One factor which may be instrumental in the reduced geographical range in Sintra is the Man-induced habitat loss. *A. hemionitis* is confined to artificial walls that are at least 150 years old. However, earlier collections from this area in Portugal, dating back to the 17th century (Durães, 2000), suggest that it has also been growing here at natural sites. This suggests that these populations are secondary

colonisations from older populations, whose natural sites may have been lost due to habitat alterations and that some genetic diversity may have been lost in the process.

Furthermore, *A. hemionitis* might have had a wider range in continental Europe in the past. Assuming that it was once further widespread, climate alterations like the last glaciation events could have shaped its ecological range, greatly reducing the area where it could establish. This is the case for many plants in Europe (Hewitt, 2004) because the continent has been affected by repeated glaciation events over the last 2.4My. The significance of this is further discussed in the next Chapter (6) based on data from cpDNA sequencing.

The Azores may have lower genetic diversity because it is the most remote region (in terms of both distances from the mainland and nearest archipelagos, Chapter 1). This large distance may have resulted in a reduced number of spores colonising the islands, i.e. populations may be the result of a founder effect. The Azores is also the youngest archipelago (Chapter 1), and this will have allowed less time in which diversity could have accumulated. The higher diversity found in the island of São Miguel when compared with the other islands of the Azores also supports this hypothesis since it is older than the other islands.

The islands of Madeira and Canary Islands contain the highest levels of genetic diversity, both within population and total diversity. This is true for all parameters of genetic diversity. For example, in Tenerife all *loci* are polymorphic (P=1.0) and it has the highest number of alleles per *locus* (5.133). The highest *He* is found in La Gomera, both within populations (mean=0.301) and total gene diversity (0.319). In Madeira, the total diversity is slightly lower than these maximum values (P, A, Ap and *He* are 0.933, 3.867, 4.000 and 0.295, respectively), but it is still higher than in Sintra and the Azores, and even some of the Canary Islands. As a region, the Canary Islands contain more diversity than any of the other regions (P, A, Ap and *He* are 1.000, 5.667, 5.667 and 0.336, respectively). Madeira and the Canary Islands (with the exception of El Hierro and La Palma) are the oldest islands in the considered range of distribution. Also, in these areas the natural habitat is more diverse and better preserved (Chapter 1). Another

important reason for higher diversity is that these islands are closer to each other and the mainland than the Azores, allowing for more gene exchange.

In general, there are no dramatic differences in total genetic diversity between the various regions. This is maybe due to the outcrossing breeding system (Chapter 4), which ultimately leads to the accumulation and maintenance of genetic diversity. The differences observed and explained above appear slight when the genetic diversity in other pteridophytes is considered.

The global genetic diversity in *A. hemionitis* is high when placed in the context of other pteridophytes (*P*, *A* and *He* are 1.000, 6.333 and 0.320, respectively). Li & Haufler (1999), in a comparison of 37 fern species based on isozyme data, reported a variation in gene diversity from zero to 0.33, with an average of 0.109; the average values for *P* and *A* are respectively, 0.36 and 1.65. Higher gene diversities are found in species of the genus *Asplenium* in Europe. J. Vogel (pers. com.) found gene diversities of 0.310 in *A. ceterach* and 0.230 in *A. trichomanes* subsp. *inexpectans*, both diploid species with European distributions. *A. onopteris* L. is also a diploid and it is present in the Macaronesian and Mediterranean regions. This species also shows high gene diversity (0.285, C. Durães, unpublished data). Like *A. hemionitis*, these species of *Asplenium* are considered to be mainly outcrossing (J. Vogel, pers. com.) and possibly highly capable of effective dispersal. In contrast, low genetic diversities based on isozyme data, have been reported for *A. trichomanes* subsp. *quadrivalens* (Suter *et al.*, 2000) with *P*, *A* and *He* of 0.076, 1.086 and 0.018, respectively. This species is considered to be highly inbreeding (Suter *et al.*, 2000). Identical outcomes from isozyme data are reported for *A. septentrionale* (L.) Hoffm. and *A. ruta-muraria* L. (Schneller & Holderegger, 1996). High genetic diversities in *A. hemionitis* might also be due to large effective population sizes and population stability for long periods of time.

An alternative explanation for the high genetic diversity found in *A. hemionitis* is the large number of samples analysed when compared with the data sets for other pteridophytes. In fact, as shown in this study, the genetic diversity increases with increasing number of sampled individuals. Other evidence is, however, contradictory.

A smaller data set within *A. hemionitis*, for example the Canary Islands, which is comparable to data sets analysed for other pteridophytes, still shows high genetic diversity (*P*, *A* and *He* are 1.000, 5.667, 5.667 and 0.336, respectively), comparable to the global diversity for *A. hemionitis* (*P*, *A* and *He* are 1.000, 6.333 and 0.320, respectively).

An important component of *A. hemionitis* genetic diversity is the alleles that are private to specific areas. The highest number of private alleles is found in the Canary Islands (20% of the species total number of alleles). The regions of Sintra, Azores and Madeira contain considerably fewer private alleles but in a similar proportion (3.2, 4.2, and 4.2%, respectively). The presence of private alleles is, in itself, an indication for some degree of independent evolution. The high number of private alleles in the Canary Islands suggests that the populations have been large for a long period of time, so that genetic drift did not obliterate those alleles. The frequency of these private alleles is, however, very low ($\leq$0.05). Exceptionally, the Azores contain two private alleles at higher frequencies (0.098 and 0.128). This suggests that the Azores is a more isolated area that has been evolving independently. Dispersal (gene flow) out of the Azores may be less effective than dispersal into the Azores, so that the frequency of its private alleles increases and these are not dispersed to other regions. The hypothesis of higher isolation and independent evolution of the Azores (and the other regions) can be tested using DNA sequence data, a method that allows determination of relatedness between variants, and it is explored in the next Chapter (6).

5.4.2 Genetic structure and relationships among populations

The amount of gene flow between both populations and different areas affects the genetic structure and population differentiation of a plant species. The genetic differentiation in *A. hemionitis* was assessed at different geographical scales using a hierarchical approach. At the 'island' scale, estimates of genetic differentiation are not high but significant. The equivalent values obtained for most of the 'islands' show that populations within 'islands' have differentiated to a similar extent. At this scale the

majority of genetic variation can be found within populations as opposed to among populations. The islands in which populations have differentiated, shown by their Θ_n values, to a greater degree, are La Gomera (0.180), El Hierro (0.152), both in the Canary Islands, and Flores in the Azores (0.151). The sites where the least differentiation has occurred are Sintra (0.086), Tanger (0.067), Gran Canaria (0.072) and La Palma (0.040). In terms of intraregional differentiation, the islands of the Azores are less differentiated (0.149) than the Canary Islands (0.185).

Previous results (Chapter 4) have demonstrated that at the smallest scale (within populations) there is some restriction of gene flow. The fact that there is significant genetic differentiation among populations within islands shows that there is also restriction of gene flow at a higher geographical scale. This is likely to be the result of a barrier blocking effective spore dispersal. This may be because the majority of detected spores fall close to the populations and long distance dispersal is relatively infrequent. The higher population differentiation found, for example, in La Gomera and El Hierro, suggests that the spore dispersal barrier is more effective in these islands, or maybe that its populations have been separated for a longer period of time.

When the population structure is considered at the regional scale, the highest levels of differentiation are found within the archipelagos of the Azores (Θ_n=0.149) and Canary Islands (Θ_n=0.185). One factor which may be responsible for this elevated level of differentiation is that these populations are structured differently. The material sampled here comprised a greater number of populations and these were distributed across a wider geographical range. These two factors could both contribute to the existence of well differentiated populations by maintaining some degree of independence and isolation.

A hierarchical analysis of genetic variance showed that whether we consider hierarchies within archipelagos (e.g. the Azores or the Canary Islands) or all regions as a whole, the level of differentiation increases when we consider smaller geographical scales. For example, a four-level hierarchy including all regions showed that only 3.4% of genetic variation is present between regions, 22.3% among 'islands' and 29.8% among

populations within 'islands'. This is additional evidence for restriction of gene flow within 'islands', but also for the fact that, although most likely an infrequent event, sufficient spores reach distant archipelagos and prevent high differentiation between regions.

Analysis of differentiation between all different pairwise combinations of 'islands' revealed that Madeira has diverged significantly from the others (all pairwise $\Theta_n \geq 0.292$, average=0.361), but more from Sintra (Θ_n=0.403). The least divergence between pairs of 'islands' is found within the Canary Islands (Θ_n=0.038–0.240, average=0.137) and the Azores (Θ_n=0.035–0.108, average=0.068). This is consistent with the expectations based on proximity of populations within these archipelagos. At the regional scale, Madeira is genetically closer to the Canary Islands, although differentiation is still significant (Θ_n=0.296), and more differentiated from the mainland (Θ_n with Portugal=0.396 and with Tanger=0.388). Differentiation of Madeira from the Azores (Θ_n=0.361) falls between the Canary Islands and the mainland.

These results show that Madeira is highly divergent from the other regions, and especially so from the mainland populations, a pattern also noticeable in the neighbour-joining (NJ) analysis. The reason why this is so, is not clear. A similar phenomenon has been recorded for the liverwort *Porella canariensis* (F. Weber) Underw. (Freitas & Brehm, 2001), which has a similar distribution range to *A. hemionitis* and also disperses by spores. The authors explain the high degree of genetic variation there as an adaptation to a greater variety of ecological niches, successive colonisation waves from different origins or older colonisation, although the reasons for the greater genetic differentiation relative to the other regions have not been explained to date. Arroyo-García *et al.* (2001) investigated the genetic similarities among laurel populations, including *Laurus nobilis* L. from the Mediterranean region and Iberian Peninsula, and *L. azorica* from Macaronesia. In this study, closer genetic affinities were found between populations from the Canary Islands and Sintra, and between populations from Madeira and the south of Spain. The authors do not put forward an explanation for this pattern, possibly because the main interest of their investigation concerned the Iberian Peninsula/Macaronesia and Mediterranean area dichotomy, associated, respectively,

with *L. azorica* and *L. nobilis*. Our study has a similar outcome, although a closer association between Madeira and elsewhere on the mainland has not been found. Madeira may be related to populations of *A. hemionitis* on the mainland that have not been sampled (Chapter 1), like the ones in Alger (Algeria) and Casablanca (Morocco), or even populations in southern Spain that have become extinct. Other relationship patterns between populations, including the Azores in relation to the other regions, and the position of the western Canary Islands within the Azores cluster (evident in the NJ tree) will be discussed with data from DNA sequencing (Chapter 6), as the isozyme information does not allow for conclusive speculations.

5.4.3 Isolation by distance

Under an IBD model, it is expected that Θ_n, a measure of genetic differentiation among populations, would increase as populations become more widely separated. Thus, significant values of Θ_n indicate IBD. The analysis of autocorrelation of genetic and physical distance between all populations of *A. hemionitis* indicates little but significant IBD (r^2=0.133). When the data are analysed removing all the populations from Madeira, the IBD correlation is stronger (r^2=0.186). This is not surprising given that larger distances exist between the nearest archipelagos, which is more congruent with the IBD model. This also agrees with the greater divergence between Madeira and the other regions. At smaller geographical scales, both within archipelagos (up to 565km) and within islands (up to72km), significant IBD was also detected. These results add up to the evidence that there is restriction of gene flow at small geographical scales.

The overall conclusions of this study are that *A. hemionitis* is highly genetically variable, both within the different regions of its geographical range and when compared with homosporous pteridophytes in general. There is no evidence for dramatic historic events like genetic bottlenecks or strong founder effects. On the contrary, it seems that most populations have managed to maintain an adequate effective size that overcomes

loss of diversity due to genetic drift. The high genetic diversity may also be due to an old origin of *A. hemionitis* with subsequent accumulation of variation.

The genetic structure and IBD analysis imply that there is some restriction of gene flow, although enough spores disperse long distances and reach isolated archipelagos. This evidence suggests that two distinct types of gene flow might be present. One involves the spores that fall near the parent plant, within the population, giving rise to genetic structure within populations and islands, and the other involves the spores that are caught up in the air and travel very long distances reaching isolated archipelagos. The latter is certainly responsible for the low variation found among regions. Another issue to consider is that some populations might be receiving or transmitting more gene flow than others, depending on their environmental conditions. For example, larger populations or populations located in exposed windy areas have a higher capacity for dispersing their spores. In addition, once a population is established in an area, most of the spores it receives come from that population and not from outside.

The populations of the different regions seem to be interconnected by some gene flow. In contrast, the presence of private alleles in specific regions shows that those areas might have been isolated to a certain degree, evolving independently. These aspects are explored in the next Chapter using data from cpDNA sequences. This information allows us to follow the direction of evolution and determine if certain areas have, or have not, been evolving independently.

CHAPTER 6

CHLOROPLAST DNA DIVERSITY AND PHYLOGEOGRAPHY OF *Asplenium hemionitis*

6.1 Introduction

The genetic structuring of plant populations and the establishment of independent evolutionary lineages are the result of both current patterns of interpopulation genetic exchange and common ancestry (Schaal *et al.*, 1998). In the previous chapter (5), the patterns of genetic variation caused by gene flow were analysed using isozyme markers. Isozymes function well as nuclear species-specific markers in the quantification of genetic diversity and differentiation. However, since the patterns of genetic variation derived from this method are exclusively based on allele frequencies, it is not possible to associate a phylogenetic signal with those patterns (Lowe *et al.*, 2004), identify independent evolutionary lineages and assess genetic change over time (Schaal & Olsen, 2000). The alternative is to employ phylogeographic methods which examine the distribution of allele genealogies in an explicit geographical context (Schaal & Olsen, 2000). Allele genealogies are constructed with neutral discrete genetic information and are arranged as a tree of character relatedness.

6.1.1 Phylogeographic methods

Phylogeographic methods based on characterisation of DNA base substitutions and insertions/deletions (e.g. RFLPs and DNA sequencing) provide a means of examining the relatedness of variants within and among populations (Ennos *et al.*, 1999). This phylogenetic approach, in which the genealogical pattern of relationships among alleles is easily inferred, has the potential to distinguish biogeographic patterns of genetic variation caused by gene flow from those caused by ancestral polymorphism (Schaal *et al.*, 1998). For both plants and animals, those methods make use of genetic variation present in the nuclear (n) and mitochondrial (mt) genomes. In plants, additional genetic information is present in the chloroplast (cp) genome (Chapter 2). Plant organelle genomes differ from the nuclear genome because, in most cases, they behave as a single haploid, non-recombinant gene that has low mutation rates (Ennos *et al.*, 1999). These

attributes determine their potential as effective genetic markers and will be considered in the context of this study.

The vast majority of phylogeographic studies have focused on animal systems, and most of these have relied on the rapidly evolving regions of the mtDNA as a source of genetic variation (Schaal & Olsen, 2000). In contrast, plant mtDNA, which has a slow rate of sequence evolution (Chapter 2), has rarely been used in phylogeographic studies, unlike cpDNA (Lowe *et al.*, 2004). Also, the plant mitochondrial genome is more complex than the chloroplast genome (Backert *et al.*, 1997) and is prone to extensive intramolecular recombination that causes heteroplasmy (Palmer, 1992). These mtDNA features make cpDNA largely preferable for phylogeographic studies (Lowe *et al.*, 2004). More recently, nDNA has been suggested as suitable for the analysis of phylogeographic structure (Schaal *et al.*, 1998; Schaal & Olsen, 2000), since it is possible to analyse numerous *loci* and test complex evolutionary hypotheses (Lowe *et al.*, 2004). However, a significant drawback with nDNA *loci* is the problem of paralogy (Lowe *et al.*, 2004), i.e. divergence between two homologous characters or *loci* following genome or *locus* duplication. Another problem with nDNA is finding appropriate variation (Schaal *et al.*, 1998). A potential candidate in the nuclear genome for phylogeographic studies is the ITS region, widely used in plant systematics. However, this region generally contains low intraspecific variation and, because it is part of a multicopy gene family, interpretation of polymorphism can be confounding (Schaal *et al.*, 1998). Additionally, the low-copy nuclear gene glyceraldehyde 3-phosphate dehydrogenase (G-3-PDh) has successfully been used in phylogeographic studies (Olsen & Schaal, 1999). As part of the phylogeographic study in *A. hemionitis*, I have conducted PCR trials for ITS and G-3-PDh, but both regions failed to amplify. The best alternative was to use cpDNA sequencing as a source of ordered genetic variation.

6.1.2 Detecting phylogeographic structure in *A. hemionitis* using chloroplast DNA

The detection of phylogenetically informative intraspecific variation is probably the single most difficult problem affecting phylogeographic studies (Schaal *et al.*, 1998). Gene diversity increases with both population size and mutation rate, and for an equivalent population size and mutation rate, gene diversity is always lower for haploid (like cpDNA) than for diploid markers (Ennos *et al.*, 1999). In addition, mutation rates in the organelle genomes are lower than in the nuclear genome (Wolfe *et al.*, 1987). Therefore, due to both the haploid nature and lower mutation rates, lower levels of gene diversity are expected in organelle genomes (cpDNA in the case of this study) than in diploid nuclear genomes.

Chloroplast DNA has been widely applied to studies of population history of temperate seed plants of Europe that have been affected by glaciations (e.g. Grivet & Petit, 2002; Petit *et al.*, 2002; Rendell & Ennos, 2002, 2003). The time scale of such events is only of thousands of years and the rate of mutation of the cpDNA is too slow to provide for a phylogenetic tree (Ennos *et al.*, 1999). Over the time scale relevant for *A. hemionitis* (i.e. millions of years as opposed to thousands), I anticipate that there has been enough time for the accumulation of cpDNA differences, therefore allowing the detection of independent evolutionary lineages at the metapopulation level. Also, those studies of seed plants rely on the fact that the cpDNA is uniparentally inherited (mostly maternally; see Chapter 2) and thus has the potential to give a clearer picture of past migration history than would nuclear markers. This is because seed and pollen have different dispersal abilities thus generating discrete patterns of chloroplast and nuclear genome variation. *A. hemionitis* disperses only by haploid spores that have equal potential to transmit their cpDNA information to the next generation. Therefore, the significance of uniparental inheritance of the cpDNA in *A. hemionitis*, as in all pteridophytes, is only of relevance in hybridisation studies.

The objective of this study is to use cpDNA sequence data to investigate the genetic structuring of *A. hemionitis* that has been shaped by its evolutionary history. DNA sequencing is considered the most informative tool to investigate the demographic

histories of populations (Emerson *et al.*, 2001; Schlötterer, 2004). Therefore, I targeted non-coding regions of the cpDNA of *A. hemionitis*, specifically the *trn*L-F and *rps*4 fragments. Previous surveys demonstrated that these regions were likely to reveal an informative amount of diversity in *Asplenium* (Vogel *et al.*, 1996; Trewick *et al.*, 2002; Schneider *et al.*, 2004a). These studies employ newly-designed primers that are more specific for ferns, the genus *Asplenium* in particular, since the universal primers previously published have proven unreliable in obtaining clear, well resolved sequences.

Due to the fact that information on DNA base substitutions and insertions-deletions shows the relatedness between different haplotypes, it is possible to determine the direction in which evolution took place. This time frame is very important for the understanding of the haplotype relationships between different regions. *A. hemionitis* has a vast distribution range and it is interesting to determine how much variation is shared by distant isolated areas or, in fact, if these have been evolving independently.

Differentiation between populations and geographical regions is here assessed by estimating the parameters of genetic differentiation *Gst* and *Nst* proposed by Pons & Petit (1995, 1996) for haploid markers. *Gst* is based on variation in haplotype frequencies alone, whereas *Nst* also takes into account genetic similarities between haplotypes. A comparison between these measures can be used to detect phylogeographic structure due to independent evolution of different populations. For example, when *Nst* is significantly higher than *Gst*, there is evidence that closely related haplotypes are more likely to be found in the same rather than different populations (Lowe *et al.*, 2004). Differentiation among populations is further assessed at different hierarchical levels with an analysis of molecular variance (AMOVA). The hierarchy is based on the distribution of *A. hemionitis* (design equivalent to that used for isozyme data in the previous chapter) in different archipelagos that comprise a series of islands with several individual populations.

Finally, the ancestral or derived nature of different haplotypes and their relationships is assessed by constructing haplotype genealogies. In addition, a phylogenetic tree using

other *taxa* from the genus *Asplenium* may help in determining the most ancestral haplotypes.

In contrast to many phylogeographic studies in which only one or two individuals per population are analysed, I sampled several individuals per population. This will allow the assessment of the phylogeographic structure at the population level (since more than one haplotype may be present in a single population) and estimate the coefficients of differentiation referred above. This sampling strategy, although with fewer individuals sampled, is equivalent to the one used for the study of isozyme diversity. Whereas comparisons of diversity and differentiation parameters between chloroplast and nuclear markers are highly desirable, I also aim to improve the information and hypotheses derived simply from the isozyme study.

6.2 Materials and methods

6.2.1 Plant material and screening of cpDNA polymorphism

The targeted regions for sequencing were the cpDNA fragments *trn*L intron + *trn*L-*trn*F intergenic spacer and *rps*4 gene + *rps*4-*trn*S intergenic spacer. These two cpDNA regions will be referred to as *trn*L and *rps*4, respectively. The sequenced individuals (c.230) represent c. 30 populations (3-10 individuals per population, average of 6.6). Some isolated individuals used for method optimisation are also included in the results. Populations, respective sites and sequencing method details are described in Chapter 3.

6.2.2 Data analysis

6.2.2.1 *Scoring of cpDNA polymorphisms*

Alignment of sequences was performed in MegAlign (DNAStar). Contigs and alignments were edited manually where appropriate. Polymorphisms were scored as unordered binary (point mutations) or multistate (insertion/deletion) characters. Final haplotypes obtained are a combination of the *trn*L and *rps*4 sequences.

Relatedness between haplotypes, represented by statistical parsimony networks, was generated with the program TCS (v.1.18, Clement *et al.*, 2000). All mutations were weighted equally. Highest outgroup (most ancestral haplotype) probability was also calculated with TCS.

6.2.2.2 *Chloroplast haplotype diversity and differentiation*

Diversity and differentiation parameters for the chloroplast genome were estimated following Pons & Petit (1995, 1996) using the programs HAPLODIV[1] and

HAPLONST[1]. Diversity parameters include mean within-population gene diversity (*hs*) and the total gene diversity (*ht*). Genetic differentiation was assessed estimating the coefficients *Gst*, *Nst* (Pons & Petit, 1995, 1996) and Θ_{cp} (Weir & Cockerham, 1984). *Gst* is based on variation in haplotype frequencies alone (unordered alleles), whereas *Nst* also takes into account genetic similarities between haplotypes (ordered alleles). Θ_{cp} weights each population according to its sample size while in *Gst* all populations receive equal weights. Diversity parameters for ordered alleles are *vs* and *vt*. *Gst* and *Nst* were compared by a permutation test using 2000 permutations (Burban *et al.*, 1999) in PERMUT[1]. Only those populations with more than two sampled individuals and regions with more than two populations were included in the analysis. All measures of diversity and differentiation were calculated for Portugal, Azores, Madeira and the Canary Islands (Morocco, with only two populations, was excluded).

6.2.2.3 *Analysis of molecular variance (AMOVA)*

To further assess the extent of genetic differentiation within and among groups of populations, a hierarchical analysis of molecular variance was carried out using ARLEQUIN (v.2000, Schneider *et al.*, 2000). Significance of variance estimates was tested using 1000 permutations.

Hierarchy (five regions):

[[Portugal[Sintra PT-1, Sintra 13, Sintra 14]]:[Morocco[Tanger 1, Tanger 2]:[Azores[Flores 5, Flores 6, Corvo 1, Corvo 2, Pico 2, Pico 4, Faial 1, São Miguel 4, São Miguel 15]]:[Madeira[Madeira 2, Madeira 3, Madeira 9]]:[Canary Islands[Tenerife 4, Tenerife 6, Tenerife 11, Gran Canaria 1, Gran Canaria 3, El Hierro 3, El Hierro 7, La Palma 1, La Palma 5, La Gomera 2, La Gomera 3]]]

[1] Software available at http://www.pierroton.inra.fr/genetics/labo/Software/

6.2.2.4 *Intraspecific phylogeny*

The 28 haplotype sequences (*trn*L and *rps*4 combined) were used to build an intraspecific phylogeny. Six diploid species of the genus *Asplenium* were included as outgroup. These species were considered appropriate as an outgroup based on recent phylogenetic studies of the genus *Asplenium* (Pintér *et al.*, 2002; Van den Heede *et al.*, 2003; Schneider *et al.*, 2004b) and are the following: *A. viride* Huds. (Austria), *A. trichomanes* subsp. *trichomanes* (Germany), *A. fontanum* (Germany), *A. septentrionale* (L.) Hoffm. (Turkey), *A. petrarchae* (Guerin) D. C. subsp. *bivalens* D. E. Meyer (Majorca) and *A. jahandiezii* (Litard.) Rouy (France). *trn*L and *rps*4 fragments (1898 characters) for all haplotypes and outgroup were assembled in EditSeq (DNAStar) and aligned with MegAlign (DNAStar). Parsimony analysis with equally weighted characters were performed for the combined data set and trees were generated with PAUP* 4.0b10 (Swofford, 2000), with the following options (Schneider *et al.*, 2004b): heuristic search with 100 random-sequence-addition replicates, tree bisection-recognition branch swapping, 'MULtrees' option on and collapse zero-length branches on. Tree branch support was estimated in PAUP by bootstrap analysis with 1000 replicates. A strict consensus based upon five most parsimonious trees was obtained (with 50% majority-rule bootstrap values). Bootstrapping provides a statistical estimate of the reliability of groupings. This involves taking a subsample of data (residues from alignment) and creating trees based on those subsamples. Bootstrap values obtained represent the percentage of bootstrapped trees in which the group is preserved. Generally 90% bootstrap values inspire high levels of confidence in the clade, whereas levels below 50% are considered poorly supported.

6.3 Results

6.3.1 Chloroplast haplotypes

A total of 231 individual plants was sequenced for *trn*L. The rate of successful sequencing was lower for *rps*4 and in total, only 214 sequences were obtained.

The *rps*4 region showed considerably more variation than *trn*L. The number of haplotypes observed for the *trn*L (c.900bp) and *rps*4 (c.1050bp) fragments is, respectively, eleven (A-K, Table 6.1) and eighteen (A-R, Table 6.1). The variation found in the *trn*L fragment results from a $[T]_n$ microsatellite varying in length between eight and nine bases, a $[C]_n$ microsatellite varying in length between seven and eight bases, four transitions (substitutions between purines or between pyrimidines) and five transversions (substitutions between purines and pyrimidines). The variation found in the *rps*4 fragment results from an $[A]_n$ microsatellite varying in length between six and seven bases, two deletions (one and 11 bases), two repetitions (four and six bases), ten transitions and two transversions. Distribution of *trn*L and *rps*4 haplotypes is illustrated in Figures 6.1 and 6.2, respectively.

The combination of *trn*L and *rps*4 sequences from the same individual (202 individuals) resulted in 28 haplotypes (Table 6.1). Haplotype K in *trn*L is not combined due to unavailability of the corresponding *rps*4 sequence (failed amplification). The number of changes between haplotypes varies between one and five.

The minimum spanning network (MSN) representing relatedness between haplotypes (Figure 6.4) is fully resolved, with the exception of two hypothetical intermediate haplotypes not detected. These hypothetical haplotypes may have been missed during population sampling or may never have existed at all (if mutations did not accumulate in single steps). In the given MSN, haplotype 1 has the highest probability for most ancestral (1.6 x higher than haplotype 18; 4.3 x higher than haplotypes 13, 8, 22, 16, 9 and 17; 26.5 x higher than the remaining haplotypes).

Relatedness between haplotypes present in the different regions is illustrated in Figure 6.5. It is quite clear from these haplotype MSN representations that most lineages are confined to particular regions. The exceptions concern haplotypes 1 and 18, widespread across *A. hemionitis* range, and haplotypes 21 and 13, shared, respectively by the Azores and Madeira, and the Azores and the Canary Islands. Also evident in the haplotype MSNs is the presence of a lineage in the Azores (haplotypes 17, 8 and 7) derived from a haplotype (16) that is not present there. Haplotype 16 has only been detected in Madeira, suggesting that it has originated there and subsequently dispersed to the Azores. Therefore, either we failed to detect haplotype 16 in the Azores or it has become extinct there.

Table 6.1 Description of the chloroplast haplotypes detected by sequencing of the *trn*L and *rps*4 cpDNA fragments. Letters A-K and A-R represent, respectively, the haplotypes detected in *trn*L and *rps*4, while numbers 1-29 represent combinations of both. Only character state differences shown (total number of bases is c. 1900); ? represents missing data.

Combined haplotypes	*trn*L											*rps*4																	Haplotype
	Haplotype																												
1	A	G	T	A	A	C	A	–	G	G	G	G	C	T	C	C	C	A	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	A
2	B	G	–	A	A	C	A	–	G	G	G	G	C	T	C	C	C	A	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	A
3	C	G	T	C	A	C	A	–	G	G	G	G	C	T	C	C	C	A	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	A
4	D	G	T	A	G	C	A	–	G	G	G	G	C	T	C	C	C	A	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	A
5	E	G	T	A	A	T	A	–	G	C	G	G	C	T	T	T	C	G	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	B
6	F	G	T	A	A	C	T	–	G	G	G	G	C	T	C	C	C	A	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	A
7	G	G	T	A	A	C	A	C	G	G	G	G	C	C	T	C	C	G	-----	T	C	A	T	----	ACAGAATATAG	C	G	–	C
8	G	G	T	A	A	C	A	C	G	G	G	G	C	T	T	C	C	G	-----	T	C	A	T	----	ACAGAATATAG	C	G	–	D
9	G	G	T	A	A	C	A	C	G	G	G	G	C	T	T	T	C	G	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	B
10	G	G	T	A	A	C	A	C	G	G	G	G	C	T	T	T	T	G	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	E
11	H	G	T	A	A	C	A	–	A	G	G	G	C	T	T	T	C	G	-----	T	C	A	T	----	----------	C	G	T	F
12	I	G	T	A	A	C	A	–	G	G	C	G	C	T	C	C	C	A	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	A
13	J	G	T	A	A	C	A	–	G	G	G	T	C	T	C	C	C	A	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	A
14	J	G	T	A	A	C	A	–	G	G	G	T	C	T	C	C	C	A	-----	T	C	A	T	AAGC	ACAGAATATAG	C	G	T	G
15	A	G	T	A	A	C	A	–	G	G	G	G	T	T	T	T	C	G	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	H
16	A	G	T	A	A	C	A	–	G	G	G	G	C	T	T	C	C	G	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	I
17	A	G	T	A	A	C	A	–	G	G	G	G	C	T	T	C	C	G	-----	T	C	A	T	----	ACAGAATATAG	C	G	–	D
18	A	G	T	A	A	C	A	–	G	G	G	G	C	T	T	T	C	G	-----	T	C	A	T	----	ACAGAATATAG	C	G	T	B
19	A	G	T	A	A	C	A	–	G	G	G	G	C	T	T	T	C	G	-----	T	T	A	T	----	ACAGAATATAG	C	G	T	J
20	A	G	T	A	A	C	A	–	G	G	G	G	C	T	T	T	C	G	-----	T	C	–	T	----	ACAGAATATAG	C	G	T	K
21	A	G	T	A	A	C	A	–	G	G	G	G	C	T	T	T	C	G	-----	T	C	A	G	----	ACAGAATATAG	C	G	T	L
22	A	G	T	A	A	C	A	–	G	G	G	G	C	T	T	T	C	G	-----	T	C	A	T	----	----------	C	G	T	F
23	A	G	T	A	A	C	A	–	G	G	G	G	C	T	C	C	C	A	AATAT	T	C	A	T	----	ACAGAATATAG	C	G	T	M
24	A	G	T	A	A	C	A	–	G	G	G	G	C	T	C	C	C	A	-----	C	C	A	T	----	ACAGAATATAG	C	G	T	N
25	A	G	T	A	A	C	A	–	G	G	G	G	C	T	C	C	C	A	-----	T	T	A	T	----	ACAGAATATAG	C	G	T	O
26	A	G	T	A	A	C	A	–	G	G	G	G	C	T	C	C	C	A	-----	T	C	C	T	----	ACAGAATATAG	C	G	T	P
27	A	G	T	A	A	C	A	–	G	G	G	G	C	T	C	C	C	A	-----	T	C	A	T	----	ACAGAATATAG	T	G	T	Q
28	A	G	T	A	A	C	A	–	G	G	G	G	C	T	C	C	C	A	-----	T	C	A	T	----	ACAGAATATAG	C	A	T	R
29	K	A	T	A	A	C	A	–	G	G	G	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?

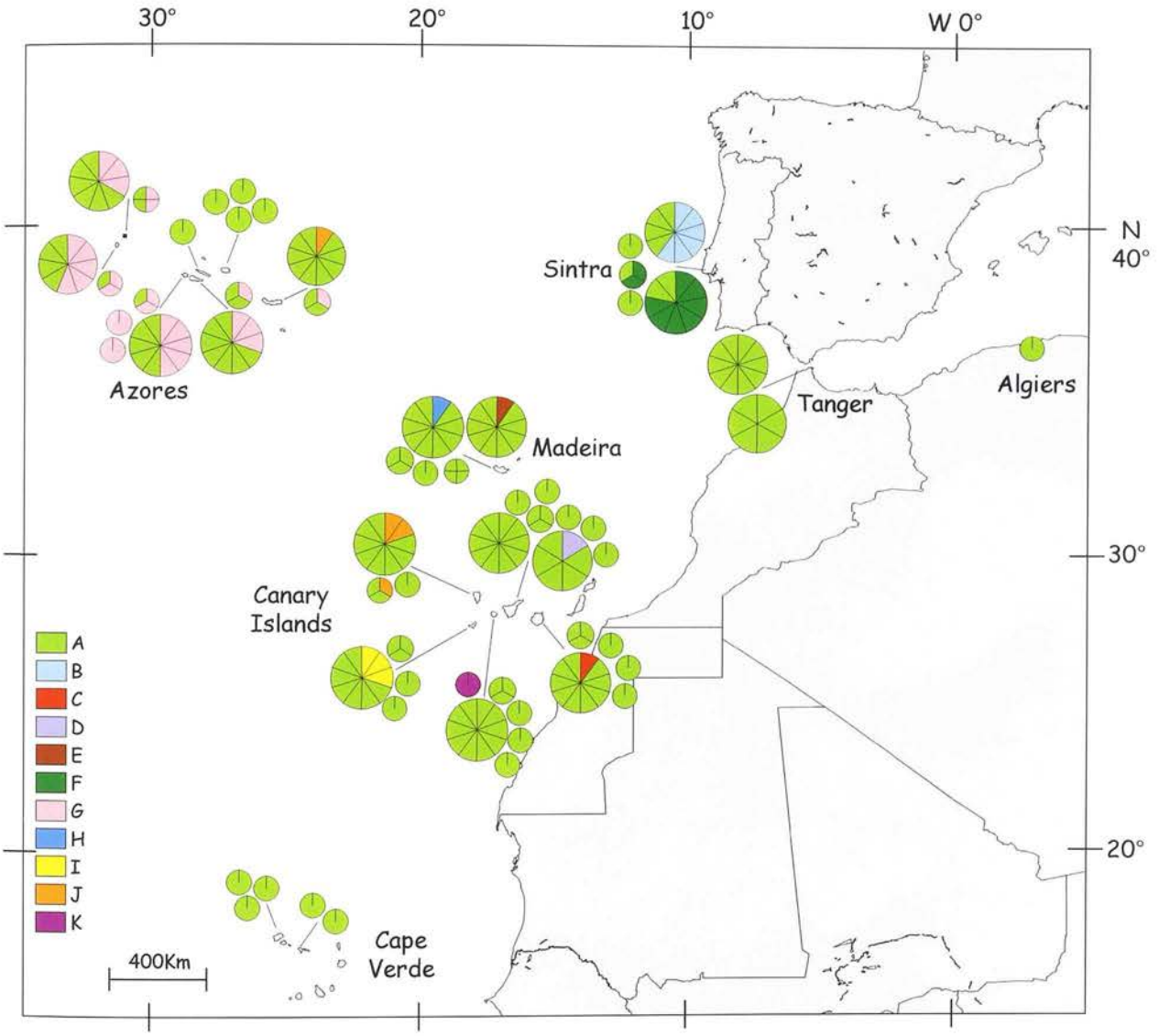


Figure 6.1 Geographical distribution of 11 *trn*L chloroplast haplotypes detected in 231 *A. hemionitis* individuals. Each colour represents one haplotype (A-K); each section in the pie charts represents one individual.

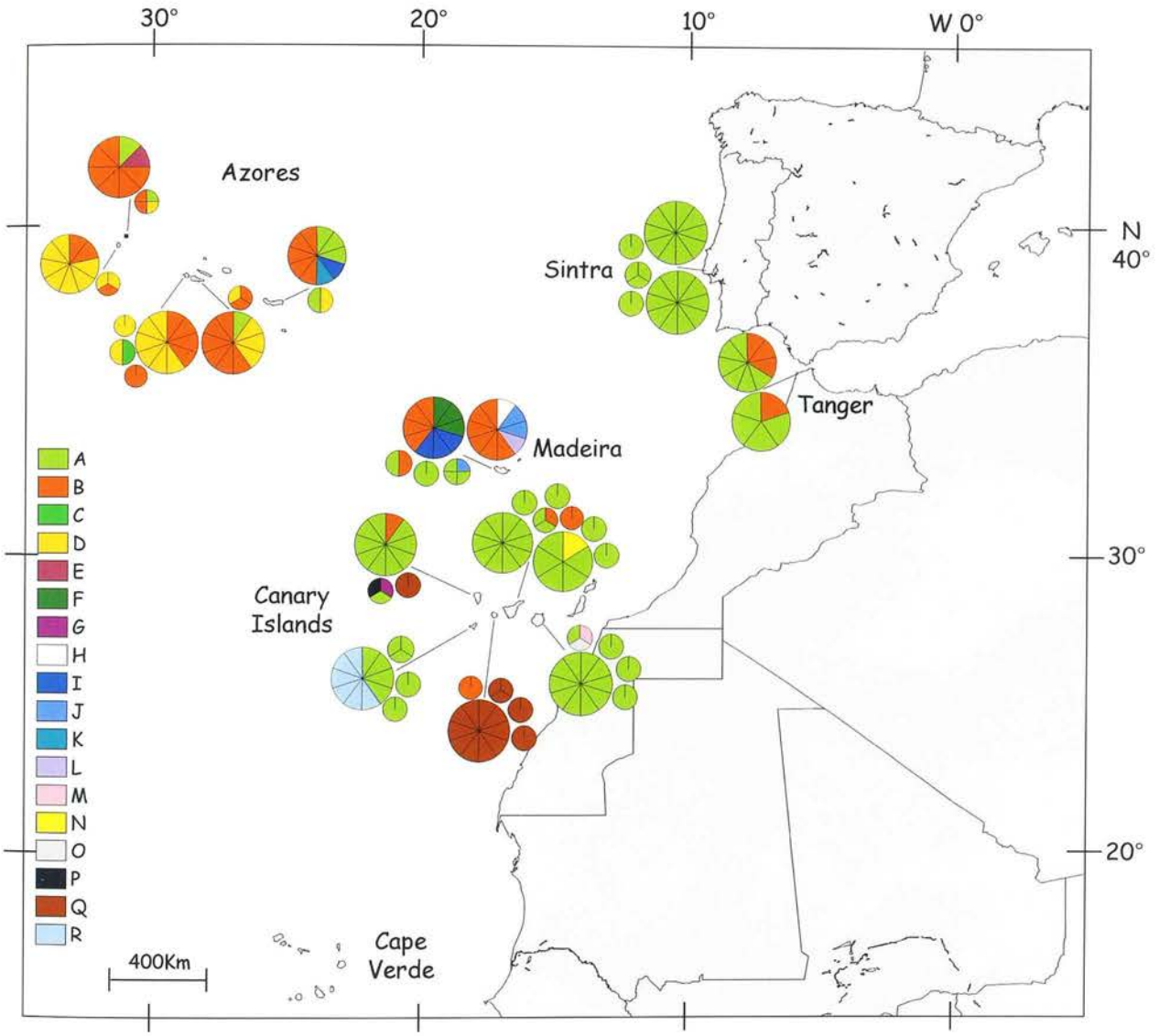


Figure 6.2 Geographical distribution of 18 *rps*4 chloroplast haplotypes detected in 214 *A. hemionitis* individuals. Each colour represents one haplotype (A-R); each section in the pie charts represents one individual.

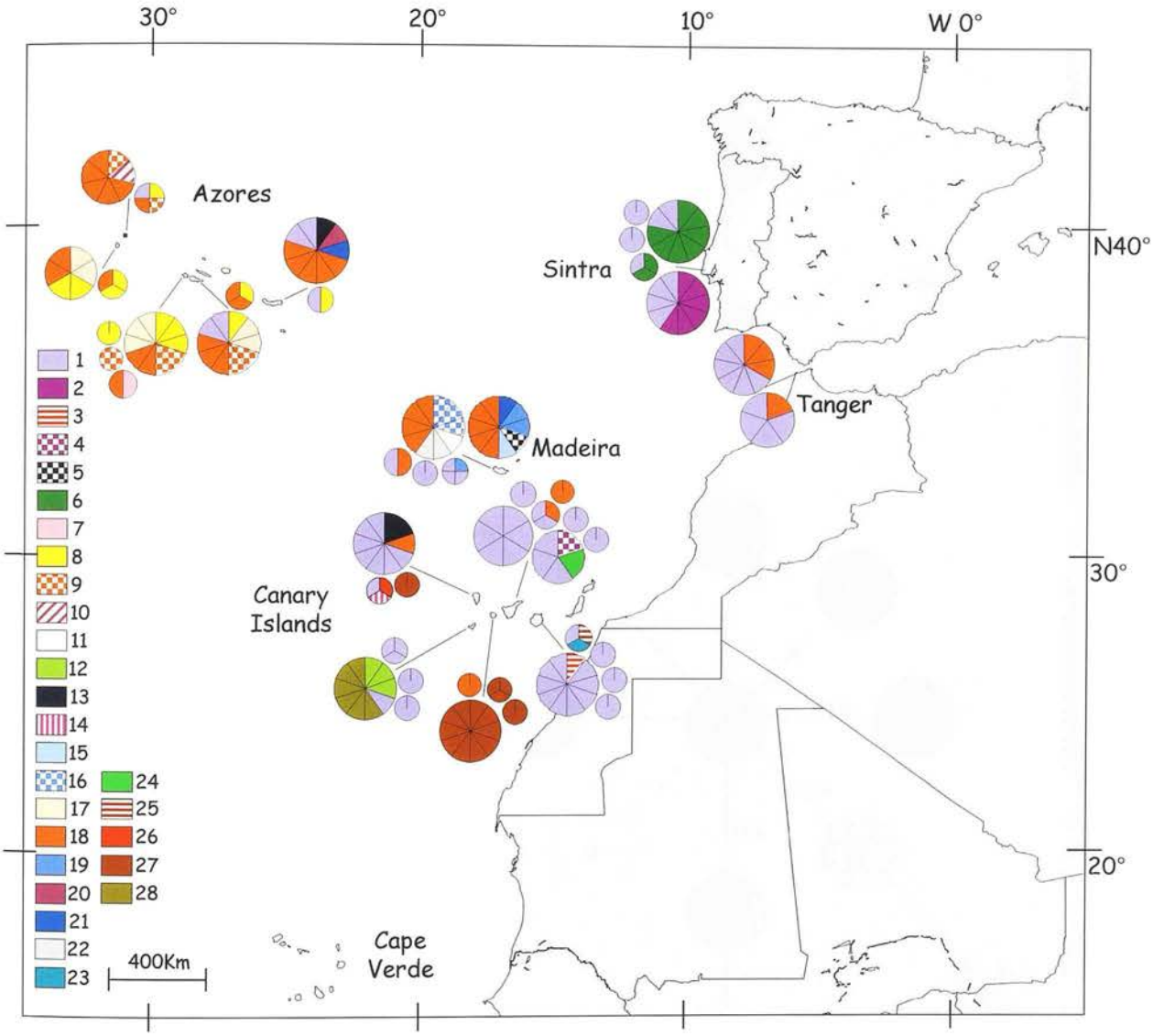


Figure 6.3 Geographical distribution of 28 chloroplast haplotypes (*trn*L and *rps*4 combined) detected in *A. hemionitis* (202 individuals). Each colour represents one haplotype (1-28); each section in the pie charts represents one individual.

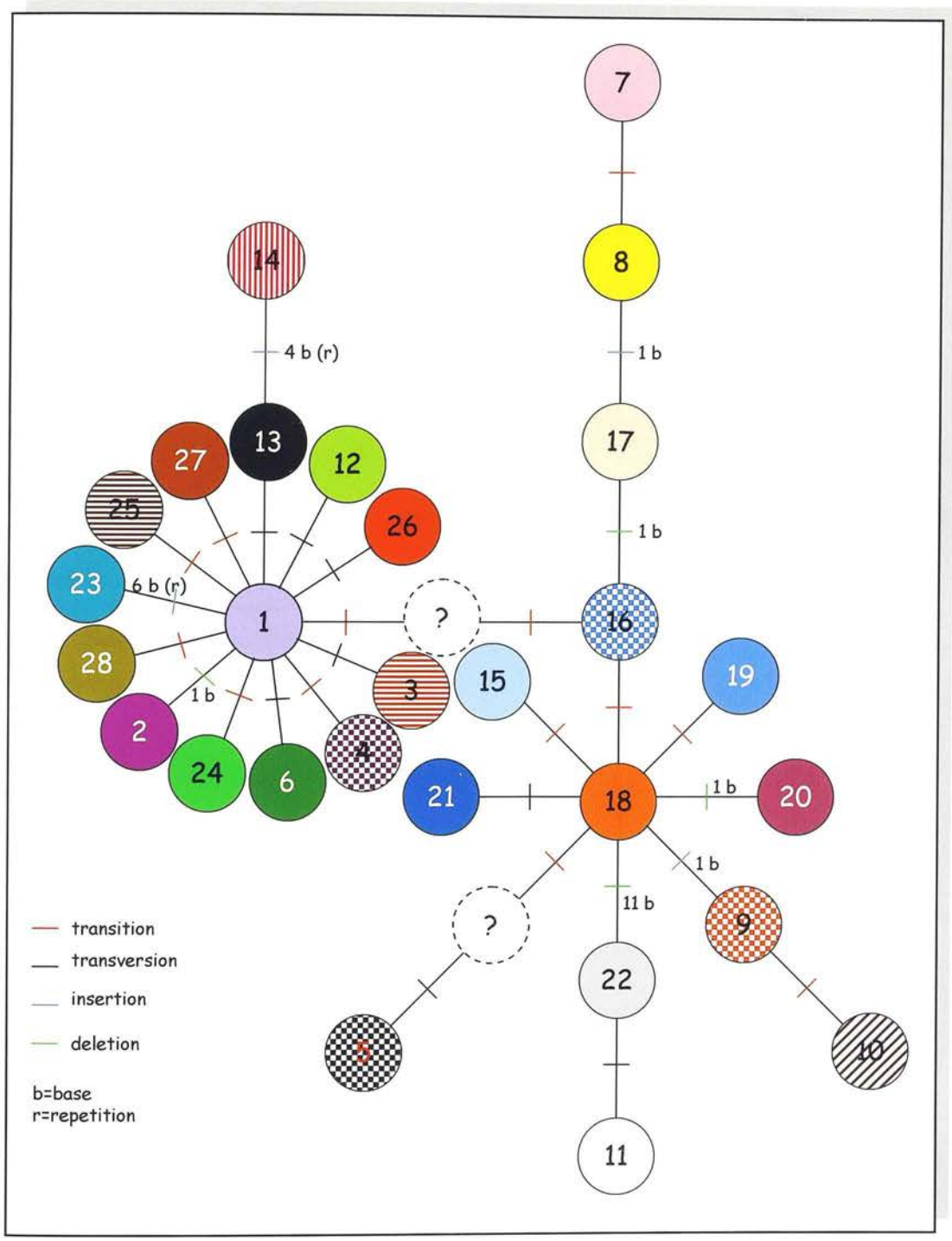


Figure 6.4 Minimum spanning network representing relatedness between 28 chloroplast haplotypes (*trn*L and *rps*4 combined) in *A. hemionitis*. Each coloured circle represents one haplotype (1-28); dashed circles represent hypothetical haplotypes not detected.

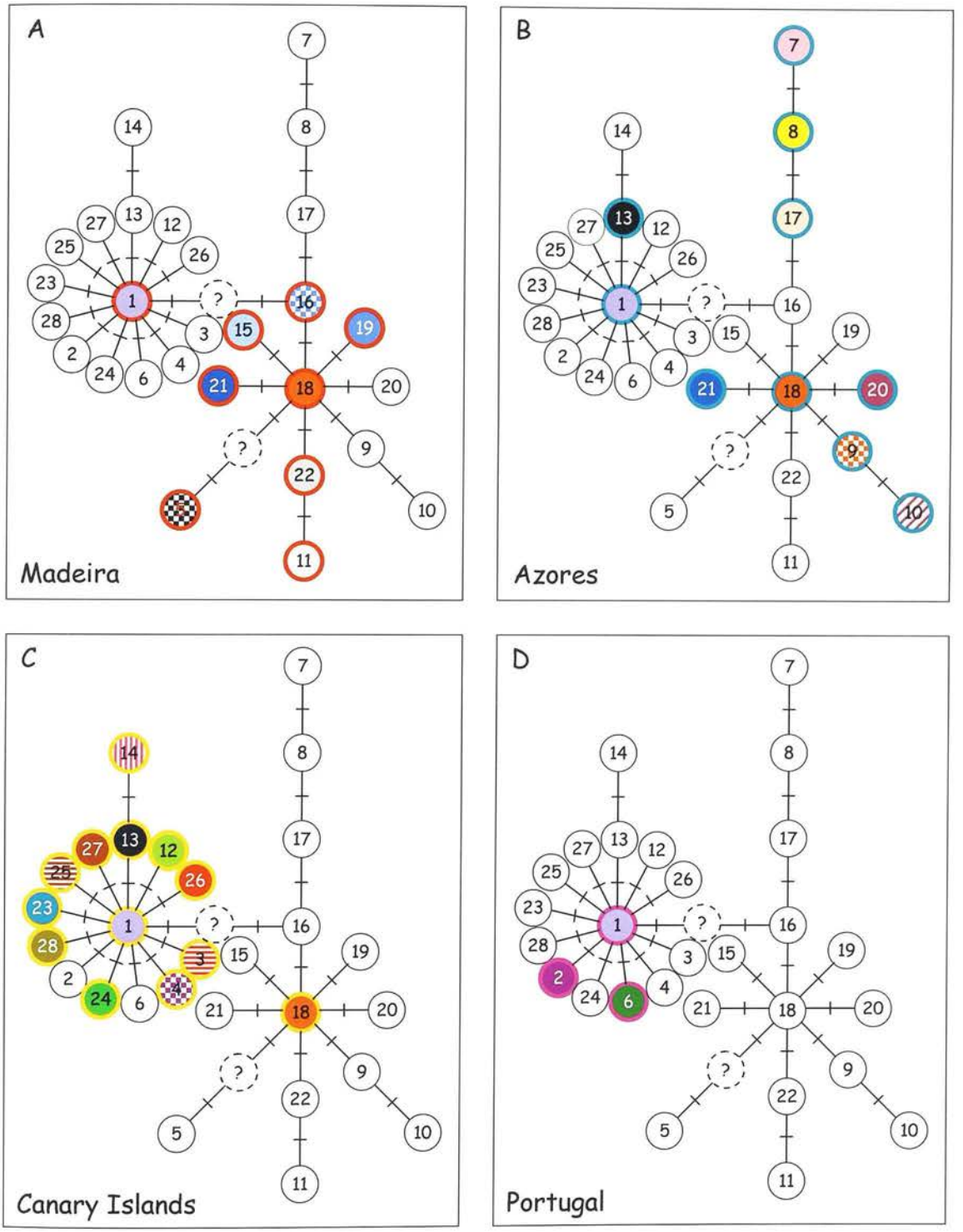


Figure 6.5 Minimum spanning networks representing relatedness between chloroplast haplotypes (*trn*L and *rps*4 combined) in *A. hemionitis* within the regions of Madeira (A), Azores (B), Canary Islands (C) and Portugal (D). Each coloured circle represents one haplotype present in the region; blank circles represent haplotypes not present; dashed circles represent hypothetical haplotypes not detected. Changes between haplotypes as in Figure 6.4. Morocco not illustrated with only haplotypes 1 and 18 present.

6.3.2 Chloroplast diversity within populations and regions

Distribution of the 28 haplotypes detected is illustrated in Figure 6.3. Twenty six of 30 populations (≥2 individuals) are polymorphic, containing a maximum of five haplotypes. Two haplotypes (1 and 18) are widely distributed. Seventeen haplotypes (60.7%) are private to single populations; of these, thirteen were detected only once. All regions, except Morocco, contain private haplotypes (Table 6.2.); the Canary Islands contain the highest number, although most at very low frequency.

Estimates of diversity are based on 27 populations (≥3 individuals). The total gene diversity (Table 6.2) is similar for each region, ranging from 0.723 in the Azores to 0.837 in Madeira. The mean within-population gene diversity ranges from 0.416 in the Canary Islands to 0.769 in the Azores. Therefore, although the Azores contains the lowest total gene diversity, it shows the highest mean within-population gene diversity. The global gene diversity for *A. hemionitis* is 0.834 and the mean within-population gene diversity is 0.564. Considering ordered alleles, Madeira shows the highest total gene diversity (0.897) and the Canary Islands the lowest (0.485).

Table 6.2 Diversity of the chloroplast genome in *A. hemionitis*. Parameters shown are mean within-population gene diversity (*hs*) and the total gene diversity (*ht*) in the main distribution areas; *vs* and *vt* are the corresponding parameters for ordered alleles; values calculated for populations ≥3 individuals and regions ≥3 populations; standard error indicated in brackets; h: no. of haplotypes; ph: private haplotypes; nc: not calculated.

		unordered alleles		ordered alleles	
Region	**h/ph**	***hs***	***ht***	***vs***	***vt***
Portugal	3/2	0.530 (0.080)	0.728 (0.126)	0.397 (0.060)	0.763 (0.167)
Morocco	2/0	0.453*	0.426*	nc	nc
Azores	10/6	0.769(0.059)	0.723 (0.037)	0.715 (0.064)	0.729 (0.062)
Madeira	9/6	0.692 (0.072)	0.837 (0.095)	0.630 (0.089)	0.897 (0.059)
Canary Islands	13/10	0.416 (0.119)	0.726 (0.091)	0.297 (0.084)	0.485 (0.082)
All populations	28	0.564 (0.056)	0.834 (0.037)	0.326 (0.039)	0.683 (0.053)

* Calculated in FSTAT

6.3.3 Genetic differentiation and variance

Over all populations, 32.3 % of the detected variation is due to differences between populations (*Gst*=0.323). Weir & Cockerham's Θ_{cp}, which takes into account the sample sizes, is higher (0.387) than *Gst*, calculated with the method of Pons & Petit. The distribution of haplotypes showed strong phylogeographic structuring. The coefficient of differentiation *Nst* (0.523) is significantly higher than *Gst* (0.323), based on the permutation test (0% of permuted *Nst* values greater than the observed value).

When all populations are merged within each region, again *Nst* (0.447) is significantly higher than *Gst* (0.193; Θ_{cp}=0.200), with 0% probability of happening by chance. Considering each region independently, all regions except the Azores (*Gst*=-0.014) show genetic differentiation between populations, with *Gst* ranging from 0.174 in Madeira to 0.427 in the Canary Islands (Table 6.3). In all regions, except the Canary Islands, *Nst* is higher than *Gst*, but it is not significant, as revealed by the permutation tests. This suggests that most of the phylogeographic structuring is present among regions, and that these contain haplotype lineages that are evolving independently.

Further evidence for phylogeographic structuring among regions was obtained from the hierarchical analysis based on a division of populations into five main areas (Table 6.4). This analysis shows that 39.80% of variation is partitioned between regions (populations not merged), while only 14.21% is partitioned between populations within regions. A great portion of the variation is contained within populations (45.99%). All estimates are significant with *P*-value<0.00001.

6.3.4 Intraspecific phylogeny

The phylogenetic tree obtained (Figure 6.6) shows *A. hemionitis* as a monophyletic group of haplotypes (bootstrap support of 100%). The basal position of haplotype 6 (only present is Sintra) in relation to all other haplotypes suggests that it is possibly the most ancestral haplotype from which the others have derived (bootstrap support of 70%).

Table 6.3 Differentiation of the chloroplast genome in *A. hemionitis* main distribution regions. Parameters shown are the coefficient of differentiation for unordered alleles (*Gst*) and ordered alleles (*Nst*); Θ_{cp} is Weir & Cockerham's (1984) coefficient of differentiation; values calculated for populations ≥3 individuals and regions ≥3 populations; standard error indicated in brackets; per: permuted values; obs: observed values; nc: not calculated.

Region	Θ_{cp}	*Gst*	*Nst*	% per≥obs
Portugal	0.394	0.273 (0.179)	0.479 (0.152)	71.8
Morocco	-0.133*	nc	nc	nc
Azores	0.004	-0.014 (0.065)	0.046 (0.103)	67.6
Madeira	0.153	0.174 (0.074)	0.297 (0.094)	5.8
Canary Islands	0.532	0.427 (0.157)	0.388 (0.147)	100.0
Among regions	0.200	0.193 (0.036)	0.447 (0.038)	0.0
All populations	0.387	0.323 (0.064)	0.523 (0.054)	0.0

* Calculated in FSTAT

Table 6.4 Hierarchical analysis of differentiation of the chloroplast genome in *A. hemionitis*. Parameter shown is percentage of variance; regions are Portugal, Morocco, Azores, Madeira and the Canary Islands; **P*-value<0.00001.

Hierarchical groups	% variance
Among regions	39.80*
Among populations within regions	14.21*
Within populations	45.99*

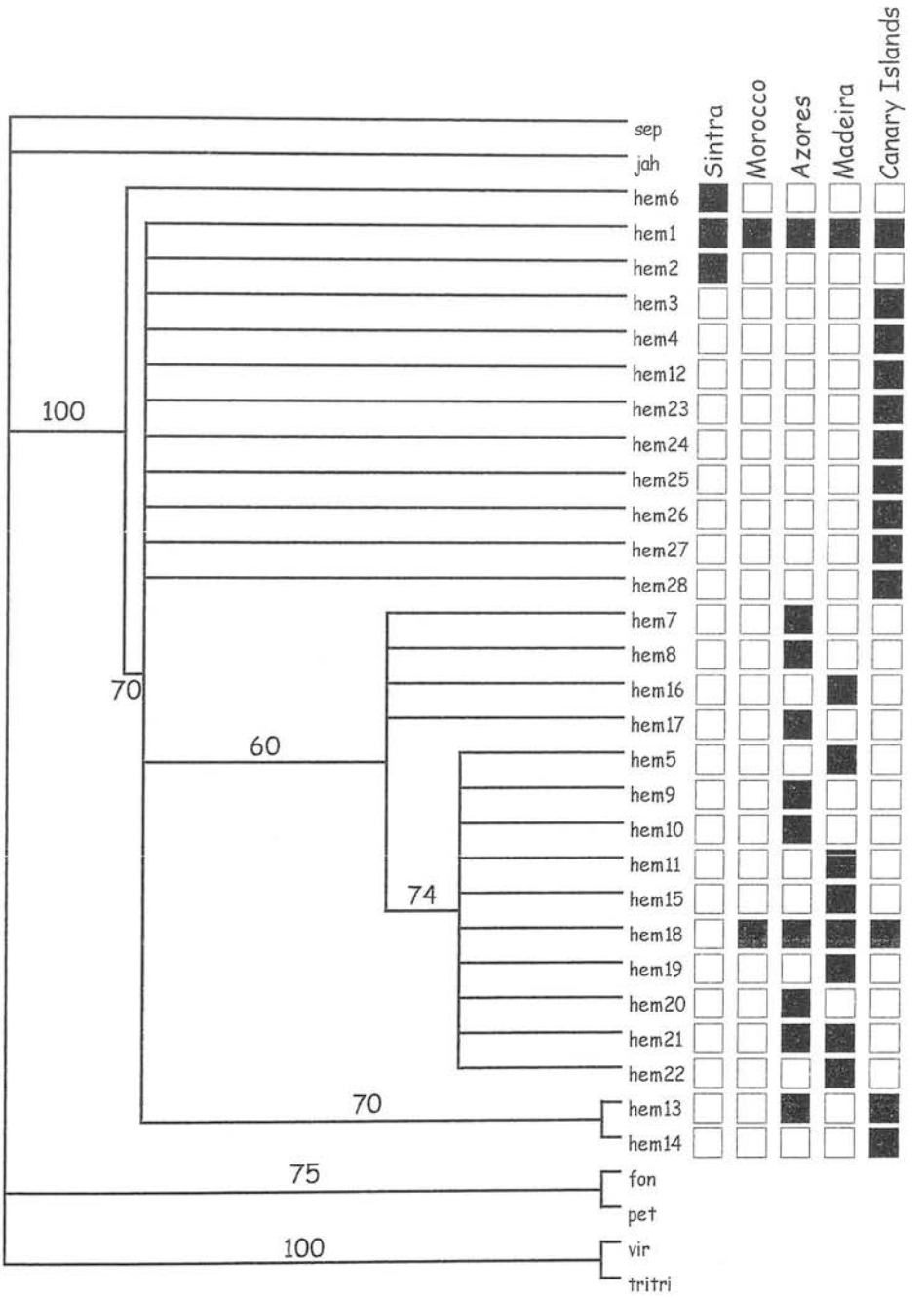


Figure 6.6 Phylogeny of *A. hemionitis* cpDNA haplotypes. A strict consensus based upon the five most parsimonious trees is shown. 50% majority-rule bootstrap values are shown above nodes. Distribution of haplotypes is represented on the right-hand side of the tree. hem: *A. hemionitis*; vir: *A. viride*; tritri: *A. trichomanes* subsp. *trichomanes*; fon: *A. fontanum*; sep: *A. septentrionale*; pet: *A. petrarchae* subsp. *bivalens*; jah: *A. jahandiezii*.

6.4 Discussion

6.4.1 Chloroplast haplotype diversity and distribution

In this investigation, two regions of the cpDNA (*trn*L and *rps*4) have been sequenced in an effort to detect enough polymorphism for the phylogeographic study. A total of 231 individual sequences were obtained for the *trn*L fragment. The rate of successful sequencing was lower for *rps*4 and in total, only 214 sequences were obtained. Despite this, more haplotypes were detected for the *rps*4 than for the *trn*L fragment (18 and 11, respectively). The maximum sequence divergence is also higher in *rps*4 (0.38%) than in *trn*L (0.22%). In a comprehensive study of the *A. ceterach* polyploid complex in Europe, using 331 individuals sampled over a broad geographic scale (Trewick *et al.*, 2002), nine haplotypes were found for the *trn*L fragment, but the maximum divergence between haplotypes was slightly higher than in *A. hemionitis*.

Integration of the *trn*L and *rps*4 data allowed the resolution of 28 haplotypes, distributed over 30 populations and 202 individuals. The nucleotide variation detected is caused by three microsatellites, two deletions, two repetitions and 21 substitutions, equivalent to a 0.26% sequence divergence.

As part of the sampling strategy, several individuals per population were sampled. The simple fact that 90% of the populations were found to be polymorphic justifies the adoption of this sampling policy. This effort was also rewarded by the detection of low frequency haplotypes within populations, an important part of the chloroplast variation (46% of haplotypes were detected only once) in *A. hemionitis*.

Two haplotypes, 1 and 18, show widespread distribution and are present at high frequencies (0.35 and 0.19, respectively). These two haplotypes are at the origin of several haplotype lineages and are related by a single haplotype (16) which is private to the region of Madeira. In this analysis, haplotype 1 showed the highest probability for a basal position in the MSN and seems to be the best candidate for the most ancestral haplotype. The remaining haplotypes, with the exception of 21 and 13 which are shared by two regions, are confined to particular regions. This evidence in itself suggests a

certain degree of genetic isolation among the different regions and will be discussed next in conjunction with results from population differentiation analysis.

Estimates of diversity for the chloroplast genome showed that the total gene diversity (*ht*) and mean within-population gene diversity (*hs*) in *A. hemionitis* are high (0.834 and 0.564, respectively). These values appear high empirically, but the lack of similar studies in Pteridophytes does not allow for comparisons. However, *ht* is equivalent to that found in angiosperms like *Calluna vulgaris* (L.) Hull (0.842; Rendell & Ennos, 2002), *Hedera* sp. (0.878; Grivet & Petit, 2002), *Quercus* sp. (mean for six species =0.828; Petit *et. al.*, 2002) and *Ilex aquifolium* L. (0.622; Rendell & Ennos, 2003). In contrast, *hs* is higher in *A. hemionitis* when compared with the same angiosperms. Although comparisons with angiosperms are the best alternative, it should be taken into account that those species have different life and demographic histories. For example, *Hedera* sp. represents a complex of four species with different ploidy levels and potential for more chloroplast diversity. The high levels of intrapopulational diversity in *A. hemionitis* are also related to its mode of dispersal (via haploid spores), since migration of cpDNA genes is not confined to the seed like in angiosperms, where large variation between and small variation within populations is expected. Furthermore, it is difficult to contrast gene diversity in *A. hemionitis* with that found in those studies because the amount of diversity detected is dependent on the genetic marker used and the amount of DNA bases screened. While those studies analysed larger fractions of the cpDNA using molecular markers like RFLPs and microsatellites, I analysed fewer DNA bases in different genes but used the most sensitive method (sequencing), detecting all variation.

It is legitimate, however, to compare gene diversity within *A. hemionitis* by examining the different regions of its range. The total gene diversity (*ht*) within each region is high (0.723-0.837). While differences in *ht* between the different regions are only subtle, more significant differences have been found in the mean within-population gene diversity (*hs*). The most diverse populations are found in the Azores (*hs*=0.769) and the least diverse are found in the Canary Islands (*hs*=0.416). Therefore, although the Azores contains the lowest total gene diversity, it shows the highest mean within-population gene diversity. This is because measures of gene diversity for a given

sample of individuals take into account not only the number of haplotypes but also the evenness of haplotype frequencies in that sample. Accordingly, although there are more haplotypes in the Canary Islands that contribute to the higher total gene diversity verified, many haplotypes are rare and appear at uneven frequencies among populations. In the Azores, the haplotypes are distributed more evenly within populations, which contributes to a higher *hs*. The regions of Sintra, Morocco and Madeira contain intermediate *hs* values, but fewer populations were analysed for these regions. When taking into account the genetic distances between haplotypes, the regional total gene diversity values (*vt*) are higher than *ht*, meaning that the considered haplotypes represent diverse lineages. In the Canary Islands *vt* is lower (0.485) than *ht* (0.726); although several haplotypes are present in this region, many occur at very low frequencies, contributing little to the total diversity. The evenness of haplotype distribution is reflected in the genetic differentiation statistics, and will be discussed further by comparing the coefficients of differentiation *Gst* and *Nst*.

6.4.2 Population differentiation and phylogeographic structure

Substantial population differentiation over all populations (Gst=0.323) was detected in the chloroplast genome of *A. hemionitis*. This differentiation is only 1.2 times higher (1.4 times higher when considering Θ_{cp}=0.387) than that revealed by isozyme markers (Θ_n=0.276) at an equivalent geographic scale (Chapter 5). For most angiosperms, in which the chloroplast is maternally inherited, much higher differentiation is expected for the chloroplast than for the nuclear genome (Ennos *et al.*, 1999). This is because differentiation due to drift is more marked for haploid maternally inherited organelle genes than for diploid nuclear genes. Additionally, gene flow is less substantial for maternally inherited organelle genes which are dispersed only via seed, than for biparentally inherited nuclear genes, which are dispersed via both pollen and seed (Rendell & Ennos, 2003). Given that ferns disperse only by haploid spores, the same predictions cannot be made. Differentiation in the chloroplast genome of *A. hemionitis* is not much higher than in the nuclear genome because both are being dispersed in the same mode; on the other hand, the chloroplast genome is still haploid and twice as affected by drift as a diploid one, which may account for the small level of elevated

differentiation in relation to the nuclear genome. Therefore, discrepancies between differentiation of the nuclear and chloroplast genomes in pteridophytes are expected to be lower than in angiosperms, given that dispersal between populations is potentially higher and the chloroplast genome is carried by all spores.

The coefficient of differentiation *Nst* over all populations (0.523) is significantly higher than *Gst* (0.323), meaning that there is strong phylogeographic structuring in the distribution of related haplotypes. The same pattern is found among regions, with *Nst* (0.447) still significantly higher than *Gst* (0.193). This evidence, coupled with the high number of private haplotypes detected within the archipelagos of the Azores, Madeira and Canary Islands, suggests that groups of related haplotypes (lineages) are somehow restricted to particular geographic areas that have been evolving independently. A similar pattern has been found, for example, in *I. aquifolium* (Rendell & Ennos, 2003) and several species of *Quercus* in Europe (Dumolin-Lapègue *et al.*, 1997a; Ferris *et al.*, 1998; Petit *et al.*, 2002), for which the distribution range has been affected by the climate changes of the Quaternary. In these *taxa*, the phylogeographic structure is explained by effective isolation of temperate *taxa* in southern refugia over numerous ice ages. Although mixing of lineages takes place in interglacial periods, there is no migration of these mixed lineages back to the refugia at the onset of the next glaciation (Ferris *et al.*, 1999). In contrast, *C. vulgaris* (Rendell & Ennos, 2002) showed no phylogeographic structure, a result explained by the higher capacity for seed dispersal and cold-tolerance with mixing of lineages from different refugia after each glaciation. Yet, the genetic differentiation in *C. vulgaris* was higher in the southern populations, and as for the *taxa* referred above, this is explained by extra isolation of populations or refugia and restriction of gene flow.

The hierarchical analysis based on a division of *A. hemionitis* populations into five regions revealed that most of the variation is contained both among regions (39.80%) and within populations (45.99%), while much less variation is partitioned between populations within regions (14.21%). As for the isozyme markers, a large fraction of the variation is found within populations. The isozyme markers showed much less differentiation among regions, but significant isolation by distance. This is not surprising given that isozymes (nuclear markers) correspond to an effective population

size twice as big as cpDNA markers. It is concluded that gene flow among regions is restricted, but nevertheless enough to create shared diversity between those regions. The partition of variation found in the chloroplast genome also shows that there is restriction of gene flow among regions; however, a few haplotypes are shared by different regions, providing evidence again for the existence of some gene flow. This is the case of haplotype 13 which is present both in São Miguel (Azores) and La Palma (Canary Islands) and haplotype 21, shared by São Miguel and Madeira. In both cases, it is not possible to establish the direction of the migration, since those haplotypes could have been originated in any of the regions involved. Moreover, two haplotypes (1 and 18) which are at the centre of several lineages, are widespread, meaning that gene flow has at least been effective in the past.

The pattern of genetic differentiation and phylogeographic structure within each region differs from that found in *A. hemionitis* as a whole, especially so as each region can be identified with a distinctive pattern. All regions, except the Azores (*Gst*= -0.014), show genetic differentiation, with *Gst* ranging from 0.174 (Madeira) to 0.427 (Canary Islands). Morocco did not show genetic differentiation, but because only two populations were analysed, comparisons between *Gst* and *Nst* are not possible and it will not be considered in this discussion.

In both regions of Sintra and Madeira, *Nst* is higher than *Gst*, but not significantly higher. However, in Madeira that difference is close to significance, indicating a stronger phylogenetic structure than in Sintra. In the Canary Islands, the differentiation has the highest value (*Gst*=0.427) but it is not phylogenetically structured (*Gst* > *Nst*), indicating that there is a mix of unrelated haplotypes. Therefore, in these three regions, especially in the Canary Islands, there seems to be some restriction of gene flow between populations. The region of the Azores shows the lowest genetic differentiation (*Gst*= -0.014), and although *Nst* is higher (0.046), it is not significant. This pattern suggests that within the archipelago of the Azores, although the populations are separated by larger distances than in the other regions, there is less restriction in gene flow. However, using isozyme markers, significant population differentiation was found within the Azores. This might indicate that isozymes are more suitable than cpDNA, which behaves as a single gene, to estimate genetic differentiation, especially

so if many isozyme *loci* are scored. A higher population differentiation within the Canary Islands was also found using isozyme markers; a plausible explanation is the fact that, in this case at least, *A. hemionitis* grows in more closed areas, within forests and gullies, which might restrict the opportunity for long-distance spore dispersal.

6.4.3 Inferences on colonisation routes

The analysis of genetic variation using markers with a phylogenetic signal provides a temporal dimension for the genetic structure. Consequently, it is possible to determine the direction of DNA changes and determine which haplotypes (or lineages) are the most ancestral. Based both on the haplotype minimum spanning networks (MSN) and the phylogenetic tree, some hypothetical migration routes will be put forward. According to Schaal *et al.* (1998, and references therein), the most ancient haplotypes should be located at the centre of the gene tree and be geographically widespread, whereas the most recent haplotypes should be at the tips of the gene tree and be localised geographically. This basic pattern can be, however, modified by interpopulation gene flow and persistence of polymorphisms that are differentially sorted following divergence of populations (Schaal *et al.*, 1998). The pattern of haplotype distribution in *A. hemionitis* is a good example of conformity with the basic corollary referred to above (Figure 6.4). However, it also shows some incongruence between haplotype position on the MSN and geographic distribution, and it is further discussed next.

Haplotypes 1 and 18 are geographically widespread and at the centre of several lineages. According to the MSN and the probability test performed with TCS, haplotype 1 has the highest probability of being the most ancestral. Contrastingly, the phylogenetic tree recovers haplotype 6 in a basal position in relation to the other haplotypes (bootstrap support of 70%). Haplotype 6, only present in Sintra, could have originated haplotype 1 (one base change), which then spread across the *A. hemionitis* range. Once established in the different regions, this haplotype would have originated several haplotypes that are private to the different regions (e.g. haplotypes 3, 4, 12, 23, 24, 25, 26, 27 and 28 in the Canary Islands and 16 in Madeira). Haplotype 18 is also at

the centre of several other lineages, which are independent from haplotype 1. It has diversified in the Azores and Madeira, although more extensively in the latter. Haplotype 16, derived from haplotype 1, is only present in Madeira, but it has originated a lineage that occurs only in the Azores (haplotypes 17, 8 and 7). It is possible that it has appeared firstly in Madeira and then colonised the Azores where it evolved; in this case, either I failed to detect haplotype 16 in the Azores or it has become extinct there. There is also evidence for migration routes between the Azores and the Canary Islands (both share haplotype 13) and between the Azores and Madeira (both share haplotype 21). It is difficult to infer in which region haplotypes 21 and 13 initially appeared because their direct ancestral haplotypes (18 and 1, respectively) are present in both regions. It is plausible, however, that haplotype 18 originated in Macaronesia, and then colonised the mainland, as shown by its presence in Morocco. The presence of private haplotypes in Sintra (2 and 6) gives strong evidence that its populations are relictual, and therefore it is unlikely that the site has been colonised, or even recolonised, from Macaronesia. This view is also supported by the presence of private isozyme alleles in Sintra (Chapter 5).

Several authors (e.g. Sunding, 1979; Rodrigo Pérez, 1992) consider *A. hemionitis* as an element of the European Tertiary flora (Chapter 1). There is at least one reference for the presence of *A. hemionitis* in the fossil record of southern France (Saporta, 1865). The stratigraphic layers where it has been found are placed in a Miocene context (23.5 to 5.3My BP). The nomenclature used to classify the fossil is somewhat confounding, but it is most likely to be *A. hemionitis* to which the author is referring. In conjunction with the existing evidence for the presence of many other typically Macaronesian flora elements in the fossil record of Tertiary Europe (Chapter 1), it is possible that *A. hemionitis* was part of that flora.

A plausible scenario is that of a metapopulation present in southern Europe, and maybe even northern Africa, at the end of the Tertiary (c. 10-2My BP). This metapopulation was characterised by effective long distance spore dispersal and colonised the Atlantic islands, which began emerging c. 20My BP. During the glaciations (c. 2My-10000y BP), the populations in Europe became reduced in size, losing the capacity for long distance dispersal. Long term geographic isolation, counteracted by some gene flow,

and recent events such as habitat loss and consequent population fragmentation led to the phylogeographic structuring observed today.

The study of cpDNA has brought further insights into the genetic structure and colonisation history of *A. hemionitis*. Some of these, like the direction of evolution, could not be determined using isozymes alone. Substantial variation was detected in the chloroplast genome of *A. hemionitis*, but less than expected for a species that has existed for such a long period of time. This is in accordance with the lack of diversification (in morphology) verified across the different regions of its range, and might indicate a slow divergence rate. The variation found shows, however, a strong phylogeographic structure, indicating that all regions are relatively genetically isolated and gene flow is restricted. In general, an evolutionary rate of 0.1% divergence per My is accepted for cpDNA in higher plants (Ferris *et al.*, 1999). The Pleistocene ice-ages date back to about 2.5 My (Hewitt, 2004) and thus would be represented by a cpDNA sequence divergence of about 0.25%. This is extremely close to the 0.26% divergence found for *A. hemionitis* haplotypes. However, this value can not be applied to divergence of populations or regions since many populations exhibit a mix of haplotypes with different divergence rates. The 0.26% divergence rate in *A. hemionitis* is equivalent to 2.6My. It is conceivable that the populations were initially large and connected by effective long-distance spore dispersal. These populations began diverging more markedly during the glaciations due to fragmentation, reduction in size and consequent decline of the potential for the establishment after long-distance dispersal. Another outcome from this study is that taking into account the age of the islands to explain, for example, diversity can be misleading. A good example is the case of the Azores, which, although comprising the youngest islands in Macaronesia, shows considerable private cpDNA diversity. Additionaly, São Miguel is the oldest of the sampled islands in the Azores and the one with the most haplotypes. This is in agreement with the hypothesis of older age/higher diversity. Many questions regarding the colonisation history of *A. hemionitis*, like the reduced relationship between Madeira and the mainland or the partition of diversity within the Canary Islands, still remain. These questions will be further addressed in the next chapter (7, General discussion) and discussed in the context of possible additional research.

CHAPTER 7

GENERAL DISCUSSION AND CONCLUSIONS

7.1 General discussion

7.1.1 Introduction

Compared with seed plants, relatively few studies of levels and distribution of genetic variation within and between populations have been published in pteridophytes and those have mostly investigated species with continental distributions (e.g. Soltis & Soltis, 1990c; Ranker, 1992). The scarce number of investigations on insular pteridophytes is nearly exclusively dedicated to species from the Hawaiian archipelago (e.g. Ranker, 1994; Ranker *et al.*, 1996; Li & Haufler, 1999). The Macaronesian islands have received far less attention and the few published studies concern cytological investigations on polyploid and hybrid complexes within a single genus (*Dryopteris* sp., Gibby, 1979) and preliminary results on the genetic variation of *Trichomanes speciosum* Willd. using PCR-RFPLs (Rumsey *et al.*, 1996). To my knowledge, this is the first comprehensive investigation of a pteridophyte using a variety of molecular marker approaches to analyse levels and patterns of genetic diversity, breeding system, gene flow, phylogeography and colonisation history.

A. hemionitis presents a set of attributes which make it an excellent subject for the study of plant colonisation and evolution on oceanic islands. Firstly, *A. hemionitis* is a diploid fern, presently recorded for 23 islands in the four main archipelagos of the Macaronesian region and a few restricted sites on the European and African W coasts. These areas on the mainland are effectively isolated by ecological and habitat range constraint and can be regarded as islands. The Macaronesian archipelagos have different degrees of isolation from each other and from the mainland, which presents a strong potential for independent evolution and radiation. This phenomenon, however, has not been observed in *A. hemionitis*, suggesting a certain degree of connectivity between distant populations. This hypothesis was tested in this investigation by employing isozyme markers in order to determine the levels of population differentiation within islands and within and among archipelagos. In conjunction, the levels of gene flow and isolation by distance were examined for different distribution

hierarchies, i.e. within populations, within islands and within and between archipelagos. The breeding system was also investigated given that it is one of the major factors affecting population structure.

Secondly, the natural habitat of *A. hemionitis* (shady slopes and gullies in the *Laurisilva* forest) has been greatly reduced, in some areas (e.g. the Azores and Cape Verde) virtually lost. Consequently, its ecological range has been extended to semi-natural (e.g. road verges) and artificial (e.g. Man-made walls) habitats, a process that has been taking place in the last 500y since the archipelagos were first colonised and the forests have gradually disappeared. A reduction of the habitat is most prominent on the mainland where the climatic cyclic alterations of the Pleistocene, in conjunction with Man's influence, have restricted many plant and animal species to refugia. Hypotheses about loss or reduction of diversity due to bottlenecks were tested, using isozyme markers, by examining the levels of intrapopulational diversity and absolute diversity within islands and archipelagos. Additionally, reduction of diversity due to founder events in distant islands was also tested.

Thirdly, the Macaronesian populations of *A. hemionitis*, like other Macaronesian *taxa*, have been proposed to represent ancient continental lineages that have survived in that region. Understanding the spatial direction in which colonisation events have been taking place requires a phylogeographic approach. Using this approach, it is possible to test hypotheses concerning independent evolution on isolated areas, detect relictual populations and infer colonisation routes. For this purpose, cpDNA sequences were used, not only to examine the levels of chloroplast diversity and differentiation, but also to identify ancestral and derived lineages and their distribution.

The major findings in this investigation are summarised next, along with a discussion of their significance in the context of the principles of colonisation and evolution on oceanic islands identified for angiosperms.

7.1.2 Breeding system

Homosporous pteridophytes, like angiosperms, are capable of outcrossing and inbreeding (inter-gametophytic selfing in the case of pteridophytes); additionally, in pteridophytes, an extreme form of inbreeding can occur in which a completely homozygous individual is produced (intra-gametophytic selfing). An outcrossing breeding system helps maintaining genetic diversity, whereas inbreeding leads to the partition of that diversity among populations. Therefore, determining the breeding system of *A. hemionitis* is the first step to understand its population genetics.

In this investigation, the breeding system of *A. hemionitis* was determined by analysis of the inbreeding coefficient using isozyme markers; however, because this method does allow for distinction between intra and inter-gametophytic selfing, breeding experiments were also conducted. Estimates of the inbreeding coefficient (*Fis*) revealed that, whereas outcrossing appears to be the predominant breeding system in populations of *A. hemionitis*, there is evidence for inbreeding. The breeding experiments did not show any evidence for intra-gametophytic selfing since none of the isolated gametophytes produced a sporophyte. Additionally, only 1.5% of 2100 individuals were completely homozygous for all scored *loci*. These individuals could have derived from intra- or inter-gametophytic selfing but also from crossing between closely related individuals.

The amount of inbreeding was more noticeable in some populations, islands and regions than others. This may reflect varying ecological conditions among populations, such as sporophyte density and availability of suitable sites for spore germination and gametophyte establishment. It has also been argued that patchiness of habitat can lead to increased rates of inbreeding (Soltis & Soltis, 1990a), for example on Man-made walls. In Sintra, where all populations grow on Man-made walls, *A. hemionitis* has a global *Fis* that is the highest recorded for the species. However, mean *Fis* values for artificial, semi-natural and natural habitats are very similar and other factors, like the range of spore dispersal and population density should be considered. In more exposed locations, spores can be carried away by wind more effectively, entering different

populations and thus promoting outcrossing. Additionally, in populations with higher density, the probability of crossing between non-sib gametophytes is higher.

These findings are in agreement with other isozyme studies on pteridophytes which have shown that outcrossing is common in diploids; on the other hand, inbreeding is reported to be the predominant breeding system in polyploids (Soltis & Soltis, 1990a, b, 1992; Masuyama & Watamo, 1990). Based on a comprehensive study of European *Asplenium* using isozyme data, Vogel *et al.* (1999b) have reported a higher density of diploids in Mediterranean areas which have served as refugia during the ice-ages (Iberian Peninsula, Balkans, Greece, Italy and maritime Alps) than in more northern areas, where polyploids prevail. Those diploid *taxa* were found to be mainly outbreeding whereas their more widespread polyploid derivatives and some diploids at more northern latitudes are mainly inbreeding. In complete agreement with this, *A. hemionitis* is a diploid pteridophyte that is confined to refugial areas in Macaronesia and W Portugal and NW Africa, and which shows an outcrossing breeding system.

7.1.3 Population genetic effects of isolation and habitat reduction

A. hemionitis has been found to be highly diverse, both in terms of nuclear and chloroplast genomes. This high variation is present at different sampling scales, i.e. within populations, islands and regions. In general, there are no dramatic differences in total genetic diversity between the various regions. However, some differences are present and require further discussion. The Canary Islands and Madeira exhibit the highest level of genetic diversity, both in terms of isozymes and number of cpDNA haplotypes. Both the Azores and Sintra showed lower levels of diversity in terms of isozyme markers and cpDNA.

The Canary Islands are clearly the centre of genetic diversity, probably because these islands still harbour adequate extents of natural areas suitable for *A. hemionitis*. The same is true for Madeira, and although none of the populations were sampled from genuine natural sites, some are probably present in more inaccessible areas of the

Laurisilva. The Azores contain less diversity, probably due to a combination of factors, like more recent emergence of the islands when compared with the other Macaronesian archipelagos, greater isolation and scarcity of natural habitats. The populations on the mainland, specifically in Portugal, maintain less diversity than the insular populations, very likely a consequence of range constraints during the Pleistocene glaciations and recent habitat loss and subsequent colonisation of artificial walls. Additionally, the populations in Portugal suffer from elevated isolation in relation to the other areas of the distribution range and occupy a smaller spatial scale.

All regions exhibit a certain level of private variation, both in terms of isozyme alleles and cpDNA haplotypes. As for diversity values, private variation is higher in Madeira and the Canary Islands and lower in the Azores and the mainland. The presence of private variation is probably the consequence of a certain degree of isolation, which is not surprising due to the large distances between all regions. The islands of the Azores and Sintra contain less private variation than the other regions, perhaps for the same reasons already appointed for lower diversity in general.

A. hemionitis has clearly been highly successful in travelling large distances since it is currently found in islands as far apart as 2000km. However, despite travelling long distances in the process of colonisation, *A. hemionitis* does not display the clear loss of diversity that is often associated with long-distance dispersal and founder events. Also, there is no support for the hypothesis that reduction of habitat has been accompanied by loss of diversity due to genetic bottlenecks. It is possible, however, that *A. hemionitis* has been through genetic bottlenecks not recent enough to be detected, from which it has recovered by means of effective gene flow between distant areas.

It seems that *A. hemionitis* has been able to retain similar levels of diversity in all areas of its distribution range by outcrossing and maintaining an effective gene flow between those areas. Additionally, an ancient rather than recent colonisation of the different range areas provided enough time for accumulation of mutations, thus increasing diversity. These two hypotheses are discussed next, respectively, by looking at the

patterns of gene flow and by inferring the origin and colonisation history of *A. hemionitis*.

7.1.4 Patterns of contemporary gene flow and isolation by distance

The establishment of new individuals in populations other than their source (gene flow) counteracts the effects of genetic drift and differentiation due to isolation. By analysing the population genetic structure at different levels, it is possible to recognise patterns of differentiation and, consequently, gene flow.

The pattern of gene flow within populations was analysed by examining the fine-structuring of individuals growing on a wall. The physical distance between individuals on the wall was regressed against their genetic relatedness (based on their multi-*locus* phenotype), producing a correlogram of relatedness of individuals within the population. Substantial structuring and isolation by distance was found within that population, showing that there is intrapopulational restriction of gene flow. The gene (spore-mediated) dispersal was determined to be approximately 22cm from the mother plant, which is in accordance with empirical values reported for other pteridophytes (Jermy, 1984). This finding has an important significance in terms of partition of genetic diversity among populations and regions. Accordingly, high levels of population differentiation are expected because gene flow is restricted to such short distances.

An analysis of genetic differentiation at different hierarchical distribution levels (populations, islands and regions) using F-statistics, revealed a somewhat contrasting pattern of gene flow. Some population differentiation was found at all scales of distribution, but much less than expected. The majority of the genetic variation was found within rather than among populations, islands or regions. Additionally, differentiation increased when considering smaller scales. These evidences show that there is in fact restriction of gene flow, but although likely an infrequent event, long-distance dispersal does occur and is sufficient to prevent high levels of differentiation

between islands and regions. These assumptions on the existence of short and long-distance gene flow are further supported by a small but significant degree of isolation by distance detected both at the regional and island levels. Connectivity between distant populations is also shown by the capacity of outcrossing between individuals from different archipelagos, as demonstrated with the breeding experiments.

7.1.5 Origin, evolution and colonisation history

Along with many other Macaronesian pteridophytes, *A. hemionitis* has been considered to have had a continental origin with subsequent colonisation of the Atlantic islands. During the Pleistocene glaciations, the continental populations became restricted to refugial pockets but the Macaronesian populations remained unaffected due to the buffering effect of the Atlantic Ocean. Over this long period of time, gene flow between the Macaronesian archipelagos and the mainland was restricted and the different regions became genetically divergent due to isolation and independent evolution. This hypothesis was tested using cpDNA sequence data of several individuals per population from different regions.

Substantial differentiation was detected in the chloroplast genome of *A. hemionitis*. This differentiation is associated with phylogeographic structuring, meaning that groups of related haplotypes are restricted to particular regions that have been evolving independently. This is further supported by the fact that all areas, with the exception of Morocco (only two populations sampled), exhibit a relatively high number of private haplotypes. Based both on isozyme and cpDNA evidence, it appears that *A. hemionitis* is relatively isolated in the various regions and has been evolving independently. However, some gene flow between distant areas still occurs and it is the major force maintaining species cohesion and preventing diversification.

The populations in Sintra maintain relatively high genetic diversity and some private variation. Most certainly, those populations have been subsisting there for a long period of time and can, therefore, be considered relictual. Both the MSN and the phylogenetic

tree do not contradict an origin on the mainland for *A. hemionitis*. In fact, the most ancestral haplotypes (1 and 6) are only present together in Sintra (1 is widespread across all regions, 6 is private to Sintra). From the haplotype distribution, there is evidence of multiple colonisation events between the Canary Islands, Azores and Madeira. There is also evidence for back-colonisation of Morocco from Macaronesia.

The isozyme data showed a strong relationship between Sintra, Morocco and the Canary Islands, a result not supported by the cpDNA data. Also, the position of the Azores and Madeira in relation to the other regions still remains rather unclear. There is some evidence for a closer genetic relationship between Madeira and the Canary Islands/Azores than with the mainland. There is also some evidence for a closer relationship between the Azores and the most western Canary Islands. In a preliminary phylogeographic investigation of the fern *T. speciosum* from Macaronesia and continental Europe (Rumsey *et al.*, 1996), a clear N/S dichotomy was found. The authors report the presence of two chloroplast haplotypes using PCR-RFLPs. The populations of the Azores, central and northern Europe share a common haplotype, while the populations from the Iberian Peninsula (N and S), Italy, Madeira and the Canary Islands share a different haplotype. In Britain, Ireland and NW France, the two haplotypes are present and mixed in the same populations. The authors suggest that this mix of haplotypes may reflect the retreat and expansion of *T. speciosum* populations during the Pleistocene ice-ages. Overall, it seems that the Azores has affinities with northern and central Europe while Madeira and the Canary Islands are more related to the Iberian Peninsula and Italy. Accordingly, it is also possible that the Azorean populations of *A. hemionitis* are derived from continental populations that existed at more northern latitudes before the glaciations. As for the Madeiran populations, they might be derived from populations in S Spain or S France that are now extinct. It is clear that the genetic relationships between all Macaronesian archipelagos and the mainland are complex, and that any patterns will be further confounded by multiple colonisation events and the active geologic history of the islands.

A possible way of achieving a clearer picture is by carrying out further sampling. For example, sampling populations from N Algeria, the W coast of Morocco and the island

of Porto Santo in the Madeiran archipelago, could give further insights into the relationship of Madeira with the mainland. Porto Santo (c. 14 My) is much older than Madeira (c. 4.6My) and was possibly colonised first. Additional sampling on the western side of Tenerife may help in a better understanding of the relationship between this island and the western Canary Islands, and maybe even with the Azores. The Canary islands of Lanzarote and Fuerteventura, closer to the African coast and where a few populations of *A. hemionitis* still subsist could hold further insights into the relationships with the mainland. Ideally, populations from the archipelago of Cape Verde should also be sampled. Furthermore, in terms of methodology, sequencing of single-copy nuclear genes, potentially more variable than cpDNA, could bring new perspectives into the phylogeography of *A. hemionitis*.

7.2 Conclusions

In the beginning of this thesis, a number of principles regarding the colonisation and evolution on oceanic islands was presented. Those principles are predominantely based on studies of angiosperms and their value is arguable in the context of this study. Each principle will be considered next and discussed how it holds for *A. hemionitis*.

a) *The success of long-distance dispersal is inversely proportional to island remoteness.* *A. hemionitis* colonised all archipelagos of the Macaronesian region, from Cape Verde and the Canary Islands to the more remote islands of the Azores, showing that it has been relatively successful in long-distance dispersal. More distant areas have not been colonised, possibly due to ecological constraint.

b) *Introductions to distant archipelagos take place once.* Contrastingly, evidence was gathered in this investigation that only a multiple colonisation scenario fits the distribution patterns of diversity found in *A. hemionitis*.

c) *Geographic isolation reduces gene flow between populations and over time, colonial populations become genetically divergent from their parent population due to mutation, genetic drift and/or natural selection.* Gene flow was found to be restricted, but still present, between the various regions of *A. hemionitis* range. The result is that all regions have become relatively genetically divergent but cohesion of the species has been maintained. This is further supported by the absence of noticeable diversification within *A. hemionitis*.

d) *Return to continents appears to be rare.* Some evidence for back-colonisation of the mainland (at least in Morocco) was found in *A. hemionitis*.

Overall, it seems that the elevated potential of pteridophytes for long-distance dispersal in relation to most angiosperms, have made them more efficient colonisers and more able to maintain species cohesion. The principles referred to above should then be adequately modified to accommodate exceptions gathered from investigations on pteridophytes and other plant groups with identical modes of dispersal (e.g. bryophytes and lichens).

In the specific context of the Macaronesian region, it would be interesting to carry out similar investigations in pteridophytes with distribution patterns identical to those of *A. hemionitis*. Good candidates are *Woodwardia radicans* (L.) Sm., *Culcita macrocarpa* C. Presl., *Davallia canariensis* (L.) Sm. and *Dryopteris guanchica* Gibby & Jermy. Their populations on the mainland are also considered to be relictual representatives of the European Tertiary subtropical flora.

REFERENCES

Abratis, M., Schmincke, H.-U. & Hansteen, T. H. (2002). Composition and evolution of submarine volcanic rocks from the central and western Canary Islands. *International Journal of Earth Sciences* **91**: 562-582.

Aldridge, A. E. (1979). Evolution within a single genus: *Sonchus* in Macaronesia *in* Bramwell, D. (ed.), *Plants and Islands*: 279-291. Academic Press, London.

Anguita, F. & Hernán, F. (2000). The Canary Islands origin: a unifying model. *Journal of Volcanology and Geothermal Research* **103**:1-26.

Arroyo-García, R., Martínez-Zapater, J. M., Fernández Prieto, J. A. & Álvarez-Arbesú, R. (2001). AFLP evaluation of genetic similarity among laurel populations (*Laurus* L.). *Euphytica* **122**: 155-164.

Aubreville, A. (1976). Centre terciaire d'origine, radiation et migration des flores angiospermiques tropicales. *Adansonia* **3**: 297-354.

Avise, J. C. (1998). The history and purview of phylogeography: a personal reflection. *Molecular Ecology* **7**: 371-379.

Axelrod, D. I. (1975). Evolution and biogeography of Madrean-Tethyan sclerophyll vegetation. *Annals of the Missouri Botanical Garden* **62**: 280-334.

Backert, S., Nielsen, B. L. & Börner, T. (1997). The mystery of the rings: structure and replication of mitochondrial genomes from higher plants. *Trends in Plant Science* **2**: 477-483.

Baker, A. J. (2000). *Molecular Methods in Ecology.* Blackwell Science, Oxford.

Baldwin, B. G., Crawford, D. J., Francisco-Ortega, J., Kim, S. C., Sang, T., & Stuessy, T. F. (1998). Molecular phylogenetic insights on the origin and evolution of oceanic island plants *in* Soltis, D. E., Soltis, P. S. & Doyle J. J. (eds.), *Molecular Systematics of Plants II. DNA Sequencing*: 410-441. Kluwer Academic Publishers, London.

Benl, G. (1969). Zur Variabilität der Blattgestalt bei *Asplenium hemionitis* L. *Berichte der Naturwissenschaftlichen Gesellschaft zu Bayreuth* **13**: 63-68.

Böhle, U.-R., Hilger, H. H., Martin, W. F. (1996). Island colonisation and evolution of the insular woody habit in *Echium* L. (Boraginaceae). *Proceedings of the National Academy of Sciences USA* **93**: 11740-11745.

Bohonak, A. J. (2002). IBD (Isolation By Distance): a program for analyses of isolation by distance. *Journal of Heredity* **93**: 153-154.

Bolle, C. (1864). Die Standorte der Farn auf der Canarischen Inseln, II. *Zeitschr. Allg. Erdk. Berlin* **17**: 249-282.

Bormüller, J. (1904). Ergebnisse zweier botanisher Reisen nach Madeira und den Canarischen Inseln. *Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie* **33**: 387-492.

Bouza, N., Caujapé-Castells, J., González-Pérez, M. A., Batista, F. & Sosa, P. A. (2002). Population structure and genetic diversity of two endangered endemic species of the Canarian laurel forest: *Dorycnium spectabile* (Fabaceae) and *Isoplexis chalcantha* (Scrophulariaceae). *International Journal of Plant Sciences* **163**: 619-630.

Bramwell, D. (1972). Endemism in the flora of the Canary Islands *in* Valentine, D. H. (ed.), *Taxonomy, phytogeography and evolution*: 141-159. Academic Press, London.

Bramwell, D. (1976). The endemic flora of the Canary Islands *in* Kunkel, K. (ed.), *Biogeography and Ecology in the Canary Islands*: 207-240. Junk, The Hague.

Bramwell, D. & Bramwell, Z. (2001). *Wild flowers of the Canary Islands.* Editorial Rueda, S. L., Madrid.

Bramwell, D. & Richardson, I. B. K. (1973). Floristic connections between Macaronesia and the East Mediterranean region. *Monographiae Biologicae Canariensis* **4**: 118-125.

Brown, A. H. D. (1990). Genetic characterization of plant mating systems *in* Brown, A. H. D., Clegg, M. T., Kahler, A. L. & Weir, B. S. (eds.), *Plant population genetics, breeding and genetic resources*: 145-162. Sinauer Associates, Sunderland, MA.

Brown, T. A. (1999). *Genomes.* Bios Scientific Publishers, Oxford.

Burban, C., Petit, R. J., Carcreff, E. & Jactel, H. (1999). Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudii* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. *Molecular Ecology* **8**: 1593-1602.

Carine, M. A., Russell, S. J., Santos-Guerra, A. & Francisco-Ortega, J. (2004). Relationships of the Macaronesian and Mediterranean floras: molecular evidence for multiple colonisations into Macaronesia and back-colonisation of the continent in *Convolvulus* (Convolvulaceae). *American Journal of Botany* **9**: 1070-1085.

Carter, R. E. (2000). DNA fingerprinting using minisatellites probes *in* Baker, A. J. (ed.), *Molecular Methods in Ecology*: 113-135. Blackwell Science, Oxford.

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

Comps, B., Gömöry, D., Letouzey, J., Thiébaut, B. & Petit, R. J. (2001). Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics* **157**: 389-397.

Cronk, Q. C. B. (1992). Relict floras of Atlantic islands: patterns assessed. *Biological Journal of the Linnean Society* **46**: 91-103.

Demesure, B., Sodzi, N. & Petit, R. J. (1995). A set of universal primers for the amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* **4**: 129-131.

Dias, E. M. (1996). *Vegetação natural dos Açores. Ecologia e sintaxonomia das florestas naturais.* PhD Thesis. Universidade dos Açores, Angra do Heroísmo.

Dray, A. M. (1985). *Plantas a proteger em Portugal continental.* Serviço Nacional de Parques, Reservas e Conservação da Natureza., Lisboa.

Dumolin-Lapègue, S., Demesure, B., Fineschi, S., Le Corre, V. & Petit, R. J. (1997a). Phylogeographic structure of white oaks throughout the European continent. *Genetics* **146**: 1475-1487.

Dumolin-Lapègue, S., Pemonge, M.–H. & Petit, R. J. (1997b). An enlarged set of consensus primers for study of organelle DNA in plants. *Molecular Ecology* **6**: 393-397.

Durães, C. (2000). Estudo da variabilidade de *Asplenium hemionitis* L. (Aspleniaceae) em Portugal. MSc Thesis. Universidade de Coimbra, Coimbra.

Dyer, A. F. (1979). The culture of fern gametophytes for experimental investigation *in* Dyer, A. F. (ed.), *The experimental biology of ferns*: 253-305. Academic Press, London.

El Mousadik, A. & Petit, R. J. (1996). Chloroplast DNA phylogeography of the argan tree of Morocco. *Molecular Ecology* **5**: 547-555.

Ellstrand, N. C. & Elam, D. R. (1993). Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics* **24**: 217-242.

Emerson, B. C. (2002). Evolution on oceanic islands: molecular approaches to understanding pattern and process. *Molecular Ecology* **11**: 951-966.

Emerson, B. C., Oromi, P. & Hewitt, G. M. (2000). Tracking colonization and diversification of insect lineages on islands: mitochondrial DNA phylogeography of *Tarphius canariensis* (Coleoptera: Colydiidae) on the Canary Islands. *Proceedings of the Royal Society London B* **267**: 2199-2205.

Emerson, B. C., Paradis, E. & Thébaud, C. (2001). Revealing the demographic histories of species using DNA sequences. *Trends in Ecology and Evolution* **16**: 707-716.

Ennos, R. A., Sinclair, W. T., Hu, X-S. & Langdon, A. (1999). Using organelle markers to elucidate the history, ecology and evolution of plant populations *in* Hollingsworth, P. M., Bateman, R. M. & Gornall, R. J. (eds.), *Molecular Systematics and Plant Evolution*: 1-19. Taylor & Francis, London.

Fairbanks, R. G. (1989). A 17,000-year glacio-eustatic sea level record: influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation. *Nature* **342**: 637-642.

Fennane, M. & Tattou, M. I. (1998). Catalogue des plantes vasculaires rares, menacées ou endémiques du Maroc. *Bocconea* **8**: 1-252.

Fenster, C. B., Vekemans, X. & Hardy, O. J. (2003). Quantifying gene flow from spatial genetic structure data in a metapopulation of *Chamaecrista fasciculata* (Leguminosae). *Evolution* **57**: 995-1007.

Féraud, G., Schincke, H.-U., Lietz, J., Gostaud, J., Pritchard, G. & Bleil, J. (1984). New K-Ar ages, chemical analysis and magnetic data of rocks from the islands of Santa Maria (Azores), Porto Santo and Madeira (Madeira archipelago) and Gran Canaria (Canary Islands). *Arquipélago* **5**: 213-240.

Ferguson, D. K. (1974). On the taxonomy of recent and fossil species of *Laurus* (Lauraceae). *Botanical Journal of the Linnean Society* **68**: 51-72.

Ferris, C., King, R. A. & Hewitt, G. M. (1999). Isolation within species and the history of glacial refugia *in* Hollingsworth, P. M., Bateman, R. M. & Gornall, R. J. (eds.), *Molecular Systematics and Plant Evolution*: 20-34. Taylor & Francis, London.

Ferris, C., King, R. A., Väinölä, R. & Hewitt, G. M. (1998). Chloroplast DNA recognizes three refugial sources of European oaks and suggests independent eastern and western immigrations to Finland. *Heredity* **80**: 584-593.

Francisco-Ortega, J., Jansen, R. K. & Santos-Guerra, A. (1996a). Chloroplast DNA evidence of colonisation, adaptive radiation, and hybridisation in the evolution of the Macaronesia flora. *Proceedings of the National Academy of Sciences USA* **93**: 4085-4090.

Francisco-Ortega, J., Santos-Guerra, A. & Carvalho, J.A. (1996b). Isozyme differentiation in the endemic genus *Argyranthemum* (Asteraceae: Anthemideae) in the Macaronesian Islands. *Plant Systematics and Evolution* **202**: 137-152.

Francisco-Ortega, J., Santos-Guerra, A., Hines, A. & Jansen, R. K. (1997). Molecular evidence for a Mediterranean origin of the Macaronesian endemic genus *Argyranthemum* (Asteraceae). *American Journal of Botany* **84**: 1595-1613.

Freitas, H. & Brehm, A. (2001). Genetic diversity of the Macaronesian leafy liverwort *Porella canariensis* inferred from RAPD markers. *The Journal of Heredity* **92**: 339-345.

Furnes, H. & Stillman, C. J. (1987). The geochemistry and petrology of an alkaline lamprophyre sheet intrusion complex on Maio, Cape Verde Republic. *Journal of the Geological Society, London* **144**: 227-241.

García-Talavera, F. (1987). Las Canarias orientales y la vecina costa africana en el Holoceno. *Eres* **7**: 55-63.

García-Talavera, F. (1999). La Macaronesia. Consideraciones geológicas, biogeográficas y palaeoecológicas *in* Fernández-Palacios, J. M, Bacallado, J. J. & Belmonte, J. A. (eds.), *Ecología y cultura en Canarias*: 39-63. Múseo de la Ciencia, Cabildo Insular de Tenerife, Santa Cruz de Tenerife, Canary Islands, Spain.

Gastony, G. J. & Darrow, D. C. (1983). Chloroplastic and cytosolic isozymes of the homosporous fern *Athyrium filix-femina* L. *American Journal of Botany* **70**: 1409-1415.

Geldmacher, J., Hoernle, K., van den Bogaard, P., Zankl, G. & Garbe-Schönberg, D. (2001). Earlier history of the ≥70-Ma-old Canary hotspot based on the temporal and geochemical evolution of the Selvagen Archipelago and neighbouring seamounts in the eastern North Atlantic. *Journal of Volcanology and Geothermal Research* **111**: 55-87.

Geldmacher, J., van der Bogaard, P., Hoernle, K. & Schmincke, H.-U. (2000). The $^{40}Ar/^{39}Ar$ age dating of Madeira archipelago and hotspot track (eastern North Atlantic). *Geochemistry, Geophysics and Geosystems* **1**.

Gibby, M. (1979). Palaeoendemism and evolution in Macaronesian *Dryopteris in* Bramwell, D. (ed.), *Plants and Islands*: 347-358. Academic Press, London.

Goudet, J. (2002). *FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2).* Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

Grivet, D. & Petit, R. J. (2002). Phylogeography of the common ivy (*Hedera* sp.) in Europe: genetic differentiation through space and time. *Molecular Ecology* **11**: 1351-1362.

Hamilton, M. B. (1999). Four primers pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* **8**: 521-523.

Hamrick, J. L. & Godt, M. J. (1990). Allozyme diversity in plant species *in* Brown, A. H. D., Clegg, M. T., Kahler, A. L. & Weir, B. S. (eds.), *Population genetics, breeding and genetic resources*: 43-63. Sinauer Associates, Sunderland, MA.

Hamrick, J. L., Godt, M. J. W. & Sherman-Broyles, S. L. (1992). Factors influencing levels of genetic diversity in woody plant species. *New Forests* **6**: 95-124.

Hansen, A. & Sunding, P. (1993). Flora of Macaronesia. Checklist of vascular plants. IV. Revised Edition. *Sommerfeltia* **17**: 1-298.

Hardy, O. J. & Vekemans, X. (2002). SPAGeDI: a versatile computer program to analyse spatial genetic structure at the individuals or population levels. *Molecular Ecology Notes* **2**: 618-620.

Harris, S. A. (1999). RAPDs in sytematics – A useful methodology? *in* Hollingsworth, P. M., Bateman, R. M. & Gornall, R. J. (eds.), *Molecular Systematics and Plant Evolution*: 211-228. Taylor & Francis, London.

Harris, S. A. & Ingram, R. (1991). Chloroplast DNA and biosystematics: the effects of intraspecific diversity and plastid transmission. *Taxon* **40**: 393-412.

Hartl, D. L. & Clark, A. G. (1997). *Principles of population genetics*. Sinauer Associates, Sunderland.

Haufler, C. H. (1987). Electrophoresis is modifying our concepts of evolution in homosporous pteridophytes. *American Journal of Botany* **74**: 953-966.

Hedrick, P. W. (1987). Genetic load and mating system in homosporous ferns. *Evolution* **41**: 1282-1289.

Helfgott, D. M., Francisco-Ortega, J. Santos-Guerra, A., Jansen, R. K. & Simpson B. B. (2000). Biogeography and breeding system evolution of the woody *Bencomia* alliance (Rosaceae) in Macaronesia based on ITS sequence data. *Systematic Botany* **25**: 82-97.

Hess, J., Kadereit, J. W. & Vargas, P. (2000). The colonization history of *Olea europaea* L. in Macaronesia based on internal transcribed spacer 1 (ITS-1) sequences, randomly amplified polymorphic DNAs (RAPD), and intersimple sequence repeats (ISSR). *Molecular Ecology* **9**: 857-868.

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

Hewitt, G.M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences* **359**: 183–195.

Hooper, E. A. & Haufler, C. H. (1997). Genetic diversity and breeding system in a group of neotropical epiphytic ferns (*Pleopeltis*; Polypodiaceae). *American Journal of Botany* **84**: 1664-1674.

Humphries, C. J. (1979). Endemism and evolution in Macaronesia *in* Bramwell, D. (ed.), *Plants and Islands*: 171-199. Academic Press, London.

Hunter, R. L. & Markert, C. L. (1957). Histochemical demonstration of enzymes separated by zone electrophoresis in starch gels. *Science* **125**: 1294-1295.

Jardim, R. & Fontinha, S. (2000). *The Laurel forest, a relic of Madeira Island.* Casa do Romeiro, Centro Social e Paroquial do Bom Jesus, Ponta Delgada, Madeira.

Jarne, P. & Lagode, P. J. L. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution* **11**: 424-429.

Jermy, A. C. (1984). Origin and distribution of Pteridophytes in the Mediterranean area. *Webbia* **38**: 397-416.

Ji, J., Russell, S. J., Gibby, M., Vogel, J. C., Sheffield, E. & Peng, H.-Z. (1994). A technique for the investigation of genetic variation in filamentous gametophytes of *Trichomanes* (Hymenophyllaceae). *American Fern Journal* **84**: 126-130.

Jones, C. J., Edwards, K. J., Castaglione, S., Winfield, W. O., Sala, f., van de Wiel, C., Bredmeijer, G., Vosman, B., Matthes, M., Daly, A., Brettschneider, R., Bettini, P., Buiatti, M., Maestri, E., Malcevschii, A., Marmiroli, N., Aert, R., Volckaert, G., Rueda, J., Linacero, R., Vazquez, A. & Karp, A. (1997). Reproducibility testing RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding* **3**: 381-390.

Jong, K. (1997). *Laboratory manual of plant cytological techniques*. Royal Botanic Garden, Edinburgh.

Juan, C., Emerson, B. C., Oromí, P. & Hewitt, M. G. (2000). Colonisation and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology and Evolution* **15**: 104-109.

Keiper, F. J. & McConchie, R. (2000). An analysis of genetic variation in natural populations of *Sticherus flabellatus* [R. Br. (St John)] using amplified fragment length polymorphim (AFLP) markers. *Molecular Ecology* **9**: 571-581.

Kephart, S. R. (1990). Starch electrophoresis of plant isozymes: a comparative analysis of techniques. *American Journal of Botany* **77**: 693-712.

Kirkpatrick, R. E. B., Soltis, P. S. & Soltis, D. E. (1990). Mating system and distribution of genetic variation in *Gymnocarpium dryopteris* ssp. *disjunctum*. *American Journal of Botany* **77**: 1101-1110.

Klekowsky, E. J. & Baker, H. G. (1966). Evolutionary significance of polyploidy in the Pteridophyta. *Science* **135**: 305-307.

Klekowsky, E. J. (1973).Sexual and subsexual systems in the homosporous ferns: a new hypothesis. *American Journal of Botany* **60**: 535-544.

Klekowsky, E. J. (1979). The genetics and reproductive biology of ferns p. 133-170 *in* Dyer, A. F. (ed.), *The experimental biology of ferns*. Academic Press, London.

Korpelainen, H. (1996). Relationships between life history characteristics and genetic structure in populations of *Equisetum* p.523 *in* Camus, J. M., Gibby, M. & Johns, R. J. (eds.), *Pteridology in perspective*. Royal Botanic Gardens, Kew.

Kreitman, M.E. & Akashi, H. (1995). Molecular evidence for natural selection. *Annual Review of Ecology and Systematics* **26**: 403-422.

Kunkel, G. (1965). Enumeratión de los helechos (Pteridofitos) de Lanzarote y notas sobre su distribución geográfica. *Revista del Museo Canario* **26**: 7-17.

Kunkel, G. (1966). Eine monströse form von *Asplenium hemionitis* L.: f. *cristatum*. *Nova Hedwigia* **11**: 351-352.

Kunkel, G. (1978). An excursion through my herbarium II. *Vieraea* **8**: 337-364.

Landergott, U., Holderegger, R., Kozlowski, G. & Schneller, J. J. (2001). Historical bottlenecks decrease genetic diversity in natural populations of *Dryopteris cristata*. *Heredity* **87**: 344-355.

Leberg, P.L. (1992). Effects of population bottlenecks on genetic diversity as measured by allozyme electrophoresis. *Evolution* **46**: 477-494.

Lewis, P. O., & Zaykin, D. (2001). *Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c).* Free program distributed by the authors over the internet from http://lewis.eeb.uconn.edu/lewishome/software.html

Li, J. & Haufler, C. H. (1999). Genetic variation, breeding systems, and patterns of diversification in Hawaiian *Polypodium* (Polypodiaceae). *Systematic Botany* **24**: 339-355.

Lobin, W., Fisher, E. & Ormonde, J. (1998). The fern and fern-allies (Pteridophyta) of the Cape Verde Islands, West Africa. *Nova Hedwigia* **115**: 1-112.

Loiselle, B. A., Sork, V. L., Nason, J. & Graham, C. (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany* **82**: 1420-1425.

Lowe, A., Harris, S. & Ashton, P. (2004). *Ecological genetics: design, analysis and application.* Blackwell Publishing, Oxford.

Mai, D. H. (1995). *Tertiäre vegetationsgeschichte Europas.* Fisher, Jena.

Masuyama, S. & Watano, Y. (1990). Trends for inbreeding in polyploid pteridophytes. *Plant Species Biology* **5**: 13-17.

Muñoz, J., Felicísimo, A. M., Cabezas, F., Burgaz, A. R. & Martínez, I. (2004). Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science* **304**: 1144-1147.

Murphy, R. W., Sites, J. W., Buth, D. G. & Haufler, C. H. (1996). Proteins I: Isozyme electrophoresis p. 51-120 *in* Hillis, D. M., Moritz, C. & Mable, B. (eds.), *Molecular Systematics*. Sinauer Associates Inc., Sunderland.

Nei, M. (1973). Analysis of genetic diversity in subdivided populations. *Proceedings of the National Academy of Sciences USA* **70**: 3321-3323.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583-590.

Nei, M. (1987). *Molecular Evolutionary Genetics.* Columbia University Press, New York.

Nei, M. (1990). DNA polymorphism and adaptive evolution p. 128-142 *in* Brown, A. H. D., Clegg, M. T., Kahler, A. L. & Weir, B. S. (eds.), *Plant population genetics, breeding and genetic resources*. Sinauer Associates, Sunderland, MA.

Newton, A. C., Allnutt, T. R., Gillies, A. C. M., Lowe, A.J. & Ennos, R. A. (1999). Molecular phylogeography, intraspecific variation and the conservation of tree species. *Trends in Ecology and Evolution* **14**: 140-145.

Nogueira, I. & Ormonde, J. (1986) *Asplenium in* Castroviejo, S., Laínz, M., López González, G., Monserrat, P., Muñoz Garmendia, F., Paiva, J. & Villar, L. (eds.), *Flora iberica: plantas vasculares de la Península Ibérica e Islas Baleares* **1**: 92. Rela Jardín Botánico, C. S. I. C., Madrid.

Olsen, K. M. & Schaal, B. A. (1999).). Evidence on the origin of cassava: phylogeography of *Manihot esculenta. Proceedings of the National Academy of Sciences USA* **96**: 5586-5591.

Ormonde, J. (1990). Pteridófitos endémicos, raros ou ameaçados das ilhas macaronésicas. *Fontqueria* **28**: 5-12.

Ormonde, J. (1991). Pteridófitas macaronésicas endémicas, raras ou em vias de extinção I. Aspleniaceas in Dias, E., Carretas, J. P. & Cordeiro, P. (eds.), *Comunicações Apresentadas nas 1as Jornadas Atlânticas de Protecção do Meio Ambiente, Angra do Heroísmo, 25 de Janeiro-1 de Fevereiro 1988*: 215-242. Angra do Heroísmo.

Ormonde, J. (1996). Estudos palinológicos em *Asplenium hemionitis* L. (Aspleniaceae) das ilhas Macaronésicas. *Anales Jardín Botánico de Madrid* **54**: 99-105.

Ouborg, N. J., Piquot, Y. & van Groenendael, J. M. (1999). Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology* **87**:551-568.

Page, R. D. M. & Holmes, E. C. (1998). *Molecular evolution: a phylogenetic approach.* Blackwell Science Ltd., Oxford.

Palmer, D. D. (2002). *Hawaii's ferns and fern allies.* University of Hawaii Press, Hawaii.

Palmer, J. D. (1992). Chloroplast and mitochondrial genome evolution in land plants p. 99-133 *in* Herrmann, R. G. (ed.), *Cell Organelles*. Springer-Verlag, Berlin.

Panero, J. L., Francisco-Ortega, J., Jansen, R. K. *et* Santos-Guerra, A. (1999). Molecular evidence for multiple origins of woodiness and a New World biogeographic connection of the Macaronesian island endemic *Pericalis* (Asteraceae: Senecioneae). *Proceedings of the National Academy of Sciences USA* **96**: 13886-13891.

Pangua, E. & Prada, C. (1988). Tipos esporales en Aspleniaceas Ibericas. *Lagascalia* **15** (Extra): 157-167.

Peck, J. H., Peck, C. J. & Farrar, D. R. (1990). Influences of life history attributes on formation of local and distant fern populations. *American Fern Journal* **80**: 126-142.

Peltier, J. P. (1973). Endémiques macaronésiennes au Maroc: inventaire bibliographique et problèmes taxionomiques. *Monographiae Biologicae Canariensis* **4**: 134-142.

Petit, R. J., Csaikl, U. M., Bordács, S., Burg, K., Coart, E., Cottrell, J., van Dam, B., Deans, J. D., Dumolin-Lapègue, S., Fineschi, S., Finkeldey, R., Gillies, A., Glaz, I., Goicoechea, P. G., Jensen, J. S., König, A. O., Lowe, A. J., Madsen, S. F., Mátyás, G., Munro, R. C., Olalde, M., Pemonge, M. H., Popescu, F., Slade, D., Tabbener, H., Taurchini, D., de Vries, S. G. M., Ziegenhagen, B. & Kremer, A. (2002). Chloroplast DNA variation in European white oaks. Phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management* 156: 5-26.

Pintér, I., Bakker, F., Barrett, J., Cox, C., Gibby, Henderson, S., Morgan-Richards, M., Rumsey, F., Russell, S., Trewick, S., M., Schneider, H. & Vogel, J. (2002). Phylogenetic and biosystematic relationships in four highly disjunct polyploid complexes in the subgenera *Ceterach* and *Phyllitis* in *Asplenium* (Aspleniaceae). *Organisms, Diversity and Evolution* **2**: 299-311.

Pons, O. & Petit, R. J. (1995). Estimation, variance and optimal sampling of gene diversity I. Haploid *locus*. *Theoretical and Applied Genetics* **90**: 462-470.

Pons, O. & Petit R. J. (1996). Measuring and testing genetic differentiation with ordered *versus* unordered alleles. *Genetics* **144**: 1237-1245.

Pryor, K. V., Young, J. E., Rumsey, F. J., Edwards, K. J., Bruford, M. W. & Rogers, H. J. (2001). Diversity, genetic structure and evidence of outcrossing in British populations of the rock fern *Adiantum capillus-veneris* using microsatellites. *Molecular Ecology* **10**: 1881-1894.

Queirós, M. & Ormonde, J. (1987). Contribuição para o conhecimento citotaxonómico da flora dos Açores. III. *Revista de Biologia da Universidade de Aveiro* **1**: 31-46.

Quézel, P. (1978). Analysis of the flora of Mediterranean and Saharan Africa. *Annals of the Missouri Botanical Garden* **65**: 479-534.

Ranker, T. A. (1992). Genetic diversity, mating systems, and interpopulation gene flow in neotropical *Hemionitis palmata* L. (Adiantaceae). *Heredity* **69**: 175-183.

Ranker, T. A., Floyd, S. K. & Trapp, P. G. (1994). Multiple colonisations of *Asplenium adiantum-nigrum* onto the Hawaiian archipelago. *Evolution* **48**: 1364-1370.

Ranker, T. A., Gemmill, C. E. C., Trapp, P. G., Hambleton, A. & Ha, K. (1996). Population genetics and reproductive biology of lava-flow colonising species of Hawaiian *Sadleria* (Blechnaceae) p. 581-597 *in* Camus, J. M., Gibby, M. & Johns, R. J. (eds.), *Pteridology in perspective*. Royal Botanic Gardens, Kew.

Reichstein, T. (1981). Hybrids in European *Aspleniaceae* (Pteridophyta). *Botanica Helvetica* **91**: 89-139.

Rendell, S. & Ennos, R. A. (2002). Chloroplast DNA diversity in *Calluna vulgaris* (heather) populations in Europe. *Molecular Ecology* **11**: 69-78.

Rendell, S. & Ennos, R. A. (2003). Chloroplast DNA diversity of the dioecious European tree *Ilex aquifolium* L. (English holly). *Molecular Ecology* **12**: 2681-2688.

Rivas-Martínez, S., Wildpret, W. de la T., Días-Gonzalez, T. E., Perez de Paz, P. L., Arco-Aguilar, M. del, Rodriguez-Delgado, O. (1993). Excursion guide. Outline vegetation of Tenerife Island (Canary Islands). *Itinera Geobotanica* 7:5-197.

Rodrigo Pérez, J. (1992). Origen de la flora Canaria in G. Kunkel (ed.), *Flora y vegetación del Archipiélago Canario* **1**: 33-54. Edirca, Madrid.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145**: 1219-1228.

Rumsey, F. J., Russell, S. J., Ji, J., Barrett, J. A. & Gibby, M. (1996). Genetic variation of the endangered filmy fern *Trichomanes speciosum* Willd. p. 161-165 *in* Camus,

J. M., Gibby, M. & Johns, R. J. (eds.), *Pteridology in perspective*. Royal Botanic Gardens, Kew.

Rumsey, F. J., Vogel, J. C., Russell, S. J., Barrett, J. A. & Gibby, M. (1999). Population structure and conservation biology of the endangered fern *Trichomanes speciosum* Willd. (Hymenophyllaceae) at its northern distribution limit. *Biological Journal of the Linnean Society* **66**: 333-244.

Salvo, A. E. & Ouyahya, A. (1999). Pteridophyta *in* Fennane, M., Ibu Tattou, M., Mathez, J., Ouyahya, A. & El Oualidi, J. (eds.), *Flore pratique du Maroc: manuel de détermination des plantes vasculaires* **1**: 46. Institute Scientifique, Université Mohammed V, Agdal, Rabat.

Santos-Guerra, A. (1999). Origen y evolución de la flora de Canarias p. 107-129 *in* Fernández-Palacios, J. M, Bacallado, J. J. & Belmonte, J. A. (eds.), *Ecología y cultura en Canarias*. Múseo de la Ciencia, Cabildo Insular de Tenerife, Santa Cruz de Tenerife, Canary Islands, Spain.

Saporta, G. de (1865). Études sur la végétation du sud-est de la France a l'époque tertiaire. XII partie. *Annales des Sciences Naturelles, Botanique,* série **5** (IV): 5-264.

Schaal, B. A. & Olsen, K. M. (2000). Gene genealogies and population variation in plants. *Proceedings of the National Academy of Sciences USA* **97**: 7024-7029.

Schaal, B. A., Hayworth, D. A., Olsen, K. M., Rauscher, J. T. & Smith, W. A. (1998). Phylogeographic studies in plants: problems and prospects. *Molecular Ecology* **7**: 465-474.

Schäfer, H. (2003). *Chorology and diversity of the Azorean flora*. Gebrüder Borntraeger, Berlin.

Scheidegger, A. E. (2002). Morphometric analysis and its relation to tectonics in Macaronesia. *Geomorphology* **46**:95-115.

Schlötterer, C. (2004). The evolution of molecular markers – just a matter of fashion? *Nature Reviews Genetics* **5**: 63-69.

Schneider, G. (1892). *Book of choice ferns* **1**: 582-587. London.

Schneider, H., Ranker, T. A., Russell, S. J., Cranfill, R., Geiger, J. M. O., Aguraiuja, R., Wood, K. R., Grundmann, M., Kloberdanz, K. & Vogel, J. C. (2004a). Origin of the endemic fern genus *Diellia* coincides with the renewal of

Hawaiian terrestrial life in the Miocene. *Proceedings of the Royal Society B* (doi: 10.1098/rspb.2004.2965).

Schneider, H., Russell, S. J., Cox, C. J., Bakker, F., Henderson, S., Rumsey, F., Barrett, J., Gibby, M., & Vogel, J. C. (2004b). Chloroplast phylogeny of Asplenioid ferns based on *rbcL* and *trn-L* spacer sequences (Polypodiidae, Aspleniaceae) and its implications for biogeography. *Systematic Botany* **29**: 260-274.

Schneider, S., Roessli, D. & Excoffier, L. (2000). *ARLEQUIN, Version 2.000: A software for population genetics data analysis.* Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.

Schneller, J. J. & Holderegger, R. (1996). Genetic variation in small, isolated fern populations. *Journal of Vegetation Science* **7**: 113-120.

Schneller, J. J. & Schmid, B. W. (1982). Investigations on the intraspecific variability in *Athyrium filix-femina* (L.) Roth. *Adansonia* **3-4**: 215-228.

Schneller, J., Holderegger, R., Gugerli, F., Eichenberger, K. & Lutz, E. (1998). Patterns of genetic variation detected by RAPDs suggest a single origin with subsequent mutations and long-distance dispersal in the apomictic fern *Dryopteris remota* (Dryopteridaceae). *American Journal of Botany* **85**: 1038-1042.

Silva, L. C., Bas, M. J. le & Robertson, A. H. F. (1981). An oceanic carbonite volcano on Santiago, Cape Verde Islands. *Nature* **294**: 644-645.

Silva, L., Tavares, J. & Smith, C. W. (2000). Biogeography of Azorean plant invaders. *Arquipélago*. Life and Marine Sciences, Supplement **2** (Part A): 19-27.

Sjögren, E. (1973). Plant communities of the natural vegetation of Madeira and the Azores. *Monographiae Biologicae Canariensis* **4**: 107-111.

Sjögren, E. (2000). Aspects of the biogeography of Macaronesia from a botanical point of view. *Arquipélago*. Life and Marine Sciences. Supplement **2** (Part A): 1-9.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science* **236**: 787-792.

Soltis, D. E. & Soltis, P. S. (1986). Electrophoretic evidence for inbreeding in the fern *Botrychium virginianum* (Ophioglossaceae). *American Journal of Botany* **73**: 588-592.

Soltis, D. E. & Soltis, P. S. (1987a). Polyploidy and breeding systems in homosporous Pteridophyta: a reevaluation. *American Naturalist* **130**: 219-232.

Soltis, P. S. & Soltis, D. E. (1987b). Population structure and estimates of gene flow in the homosporous fern *Polystichum munitum*. *Evolution* **41**: 620-629.

Soltis, D. E., Haufler, C. H., Darrow, D. E. & Gastony, G. J. (1983). Starch gel electrophoresis in ferns: a compilation of grinding buffers, gel and electrode buffers, and staining schedules. *American Fern Journal* **73**: 9-27.

Soltis, D. E., Soltis, P. S. (1990a). Polyploidy, breeding systems and genetic differentiation in homosporous pteridophytes p. 241-258 *in* Soltis, D. E. & Soltis, P. S. (eds.), *Isozymes in Plant Biology*. Chapman and Hall, London.

Soltis, P. S. & Soltis, D. E. (1990b). Evolution of inbreeding and outcrossing in ferns and fern-allies. *Plant Species Biology* **5**: 1-11.

Soltis, P. S. & Soltis, D. E. (1990c). Genetic variation within and among populations of ferns. *American Fern Journal* **80**:161-172.

Soltis, D. E., Soltis, P. S. (1992). The distribution of selfing rates in homosporous ferns. *American Journal of Botany* **79**: 97-100.

Sunding, P. (1970). Elementer i Kanariøyenes flora, og teorier til forklaring av floraens opprinnelse. *Blyttia* **28**: 229-259.

Sunding, P. (1979). Origins of the Macaronesian flora p. 13-40 *in* Bramwell, D. (ed.), *Plants and Islands*. Academic Press, London.

Suter, M., Schneller, J. J. & Vogel, J. C. (2000). Investigations into the genetic variation, population structure and breeding systems of the fern *Asplenium trichomanes* subsp. *quadrivalens*. *International Journal of Plant Sciences* **161**: 233-244.

Swofford, D. L. (2000). *PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4.* Sinauer Associates, Sunderland.

Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**: 1105-1109.

Tiffney, B. H. & Manchester, S. R. (2001). The use of geological and paleontological evidence in evaluating plant phylogeographic hypothesis in the Northern Hemisphere Tertiary. *International Journal of Plant Sciences* **162** (6 Suppl.): S3-S17.

Trewick, S. A., Morgan-Richards, M., Russell, S. J., Henderson, S., Rumsey, F. J., Pintér, I., Barrett, J. A., Gibby, M. & Vogel, J. C. (2002). Polyploidy, phylogeography and Pleistocene refugia of the rockfern *Asplenium ceterach*: evidence from chloroplast DNA. *Molecular Ecology* **11**: 2003-2012.

Tryon, R. M. (1970). Development and evolution of fern floras of oceanic islands. *Biotropica* **2**: 76-84.

Tryon, R. M. (1986). The biogeography of species, with special reference to ferns. *Botanical Review* **52**: 117-152.

Van der Heede, C. J., Viane, R. L. L. & Chase, M. W. (2003). Phylogenetic analysis of *Asplenium* subgenus *Ceterach* (Pteridophyta: Aspleniaceae) based on plastid and nuclear ribosomal ITS DNA sequences. *American Journal of Botany* **90**: 481-495.

Vogel, J. C., Rumsey, F. J., Russell, S. J., Holmes, J. S., Cox, C., Bujnoch, W., Stark, C., Barrett, J. A. & Gibby, M. (1999a). Genetic structure, reproductive biology and ecology in isolated populations of *Asplenium csikii* (Aspleniaceae, Pteridophyta). *Heredity* **83**: 604-612.

Vogel, J. C., Rumsey, F. J., Schneller, J. J., Barrett, J. A. & Gibby, M. (1999b). Where are the glacial refugia in Europe? Evidence from pteridophytes. *Biological Journal of the Linnean Society* **66**: 23-37.

Vogel, J. C., Rumsey, F. J., Schneller, J. J., Russell, S. J., Holmes, J. S., Barrett, J. A. & Gibby, M. (1998a). The origin, status and distribution of *Asplenium presolanense* spec. nov. (Aspleniaceae, Pteridophyta). *Botanica Helvetica* **108**: 269-288.

Vogel, J. C., Russell, S. J., Rumsey, F. J., Barrett, J. A. & Gibby, M. (1998b). Evidence for maternal transmission of chloroplast DNA in the genus *Asplenium* (Aspleniaceae, Pteridophyta). *Botanica Acta* **111**: 247-249.

Vogel, J. C., Russell, S. J., Barrett, J. A. & Gibby, M. (1996). A non-coding region of chloroplast DNA as a tool to investigate reticulate evolution in European *Asplenium* p. 313-329 *in* Camus, J. M., Gibby, M. & Johns, R. J. (eds.), *Pteridology in perspective*. Royal Botanic Gardens, Kew.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. & Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**: 4407-4414.

Wagner, W. L., Herbst, D. R. & Sohmer, S. H. (1990). *Manual of the flowering plants of the Hawaii* **1**. Columbia University Press, New York.

Wall, D. P. (2002). Use of the nuclear gene *glyceraldehyde 3-phosphate dehydrogenase* for phylogeny reconstruction of recently diverged lineages in *Mitthyridium* (Musci: Calymperaceae). *Molecular Phylogenetics and Evolution* **25**: 10-26.

Weeden, N. F. & Wendel, J. F. (1990). Genetics of plant isozymes p. 46-72 *in* Soltis, D. E. & Soltis, P. S. (eds.), *Isozymes in Plant Biology*. Chapman and Hall, London.

Weir, B. S. & Cockerham, C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**: 1358-1370.

Wendel, J. F. & Weeden, N. F. (1990). Visualization and interpretation of plant isozymes p. 5-45 *in* Soltis, D. E. & Soltis, P. S. (eds.) *Isozymes in plant biology*. Chapman and Hall, London.

Werth, C. R., Guttman, S. I. & Eshbaugh, W. H. (1985). Electrophoretic evidence of reticulate evolution in the Appalachian *Asplenium* complex. *Systematic Botany* **10**: 184-192.

Widmer, A. & Lexer, C. (2001). Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. *Trends in Ecology and Evolution* **16**: 267-269.

Wolfe, A. D. & Liston, A. (1998). Contributions of PCR-based methods to plant systematics and evolutionary biology p. 43-86 *in* Soltis, D. E., Soltis, P. S. & Doyle J. J. (eds.), *Molecular Systematics of Plants II. DNA Sequencing*. Kluwer Academic Publishers, London.

Wolfe, K. H., Li, W.-H. & Sharpe, P. M. (1987). Rates of nucleotide substitution vary greatly among plant mitochondria, chloroplast and nuclear DNAs. *Proceedings of the National Academy of Sciences USA* **84**: 9054-9058.

Wright, S. (1965). The interpretation of population structure by *F*-statistics with special regard to systems of mating. *Evolution* 19: 395-420.

APPENDICES

APPENDIX I

A. Breeding experiments

Composition of nutrient stock solutions (Dyer, 1979):

A.	$MgSO_4 \cdot 7H_2O$	$51gl^{-1}$
	KNO_3	$12gl^{-1}$
B.	$FeCl_3 \cdot 6H_2O$	$1.7gl^{-1}$
C.	$Ca(NO_3)_2 \cdot H_2O$	$144gl^{-1}$
D.	K2HPO$_4$	$25gl^{-1}$
E.	Mycostatin	10000 units ml^{-1}

B. Isozyme electrophoresis

i) Buffers

Protein extraction buffer: 100mM TRIS-HCl, 1mM EDTA, 10mM KCl, 10mM $MgCl_2$, 4% PVP 40 (stored at 4°C)

Grinding buffer: prepared freshly before protein extraction adding 10% DMSO and 0.5% β-mercaptoethanol to the extraction buffer (kept on ice during maceration)

Morpholine-citrate system, pH 7.0-7.4

Electrode buffer: 40mM Citric acid monohydrate, 1.4% N-(3-amino-propyl)morpholine

Gel buffer: 1 in 13 dilution of the electrode buffer

Lithium-borate system, pH 8.5

Electrode buffer: 43mM Lithium hydroxide monohydrate, 263mM boric acid

Gel buffer: 38mM Lithium hydroxide monohydrate, 45mM Citric acid monohydrate, 330mM TRIS, 300mM Boric acid (pH ca.8.0); dilute 1 in 10 before use

ii) Stains composition

Quantities shown here stain two gels in one tray.

AAT (aspartate aminotransferase)

buffer (80mM Na_2HPO_4, 1.3mM EDTA disodium, 0.5% PVP 40)	80ml
L-Aspartic acid	0.11g
α-Ketoglutaric acid	0.03g
Fast blue	0.06g

To add just before staining: fast blue and another 0.015g α-ketoglutaric acid

ACN (aconitase)

buffer (1M TRIS HCl pH 8.5)	11ml
buffer (2M TRIS HCl pH 11)	1ml
3M $MgCl_2 \cdot 6\ H_2O$	220µl
NADP	0.01g
Isocitrate dehydrogenase	0.012g
cis-Aconitic acid	0.06g
PMS (Phenazine methosulfate, 6.5mM)	1ml
MTT (Dimethylthiazol diphenyltetrazolium bromide, 17mM)	2ml

To add just before staining: isocitrate dehydrogenase, cis-aconitic acid, PMS and MTT

DIA (diaphorase)

buffer (5mM TRIS HCl pH 8.0)	100ml
menadione	0.049g
NBT (Nitro blue tetrazolium)	0.019g

To add just before staining: menadione and NBT

HEX (hexokinase)

buffer (0.1M TRIS HCl pH 8.0)	65ml
3M $MgCl_2 \cdot 6\ H_2O$	1,65ml
glucose	0.066g
EDTA (Ethylenediaminetetraacetic acid, tetrasodium)	0.033g
G-6-PDH (Glucose-6-phosphate dehydrogenase, 2000U/0.8ml))	9 µl
ATP (Adenosine triphosphate)	0.05g
NADP	0.01g
PMS	2ml
MTT	3ml

To add just before staining: G-6-PDH, PMS and MTT

IDH (isocitrate dehydrogenase)

buffer (0.1M TRIS HCl pH 8.0)	65ml
3M $MgCl_2 \cdot 6\ H_2O$	180µl
Isocitric acid	0.11g
NADP	0.011g
PMS	2ml
MTT	3ml

To add just before staining: PMS and MTT

LAP (leucine aminopeptidase)

buffer (0.1M Phosphate pH 6.0)	80ml
3M $MgCl_2 \cdot 6\ H_2O$	220µl
Fast black	0.1g
L-Leucine-β-naphthyl-amide	0.055g

To add just before staining: Fast black

MDH (malate dehydrogenase)

buffer (0.46M DL-Malic acid, 0.1M Trizma base, 0.67M NaOH, pH 8.0)	100ml
NAD	0.01g
PMS	2ml
MTT	3ml

To add just before staining: PMS and MTT

6-PGD (6-phosphogluconate dehydrogenase)

buffer (0.1M TRIS HCl pH 8.0)	65ml
3M $MgCl_2 \cdot 6\ H_2O$	320µl
6-Phosphogluconic acid	0.04g
NADP	0.01g
PMS	2ml
MTT	3ml

To add just before staining: PMS and MTT

PGI (glucose-6-phosphate isomerase)

buffer (0.1M TRIS HCl pH 8.0)	12ml
3M $MgCl_2 \cdot 6\ H_2O$	180µl
G-6-PDH	3.5µl
D-fructose 6-phosphate	0.01g
NADP	0.09g

PMS	1ml
MTT	2ml

To add just before staining: G-6-PDH, PMS and MTT

PGM (phosphoglucomutase)

buffer (0.1M TRIS HCl pH 8.0)	12ml
3M $MgCl_2 \cdot 6\ H_2O$	480µl
α-D-glucose 1-phosphate	0.115g
NADP	0.03g
G-6-PDH	7µl
PMS	1ml
MTT	2ml

To add just before staining: G-6-PDH, PMS and MTT

SkDH (shikimate dehydrogenase)

buffer (0.1M TRIS HCl pH 8.5)	65ml
Shikimic acid	0.055g
NADP	0.09g
PMS	2ml
MTT	3ml

To add just before staining: PMS and MTT

TPI (triose-phosphate isomerase)

buffer (0.1M TRIS, 0.024g EDTA tetrasodium, 2.6g arsenic acid)	180ml
G-3-PDH (Glyceraldehyde-3-phosphate dehydrogenase)	62.5µl
NAD	0.015g
Dihydroxyacetone phosphate	0.004g
PMS	1ml
MTT	2ml

To add just before staining: G-3-PDH, Dihydroxyacetone phosphate, PMS and MTT

UGPP (utp-glucose-1-phosphate uridyltransferase)

buffer (0.1M TRIS HCl pH 8.0)	12ml
3M $MgCl_2 \cdot 6\ H_2O$	300µl
Disodium pyrophosphate	0.012g
Phosphoglucomutase	87.5µl
G-6-PDH	5µl
Uridine-5-diphosphoglucose	0.02g

α-D-Glucose 1,6-diphosphate	0.002g
NAD	0.015g
PMS	1ml
MTT	2ml

To add just before staining: Phosphoglucomutase, G-6-PDH, PMS and MTT

C. cpDNA sequence analysis

i) Buffers

DNA extraction buffer: 2% CTAB, 100mM Tris-HCl pH 8, 1.4M NaCl, 20mM EDTA

Sarkosyl extraction buffer: 10% N-lauryl sarcosyne, 100mM Tris-HCl pH 8, 20mM EDTA

TE buffer: 10mM Tris-HCl pH 8.0, 10mM EDTA

Elution buffer: 10mM Tris-HCl pH 8.5

1 x TAE buffer: 4mM Tris-borate, 1mM EDTA

10 x loading buffer: 0.25% Bromophenol blue, 0.25% xylene cyanol FF, 15% Ficoll type 400

Appendix II

Maximum distance for each distance class used in SPAGeDI to calculate pairwise kinship coefficients.

Distance classes	Maximum distance	Distance classes	Maximum distance	Distance classes	Maximum distance	Distance classes	Maximum distance
1	10.20	26	330.01	51	735.30	76	1290.01
2	20.25	27	346.01	52	752.04	77	1319.00
3	30.07	28	360.20	53	766.01	78	1345.21
4	39.01	29	372.00	54	780.23	79	1381.03
5	48.17	30	395.01	55	795.01	80	1433.00
6	57.01	31	407.01	56	814.02	81	1474.02
7	65.46	32	425.04	57	828.04	82	1513.00
8	75.54	33	443.60	58	844.09	83	1556.00
9	84.17	34	457.28	59	860.02	84	1595.01
10	95.01	35	475.13	60	876.01	85	1617.18
11	107.30	36	491.26	61	893.04	86	1645.49
12	118.61	37	504.12	62	909.00	87	1664.00
13	132.37	38	517.03	63	928.19	88	1685.22
14	145.00	39	535.30	64	953.00	89	1715.17
15	155.59	40	547.57	65	975.13	90	1750.10
16	171.01	41	564.01	66	999.01	91	1795.05
17	184.00	42	580.01	67	1032.00	92	1856.00
18	198.30	43	598.37	68	1056.27	93	1944.01
19	216.11	44	618.43	69	1091.01	94	2011.01
20	228.05	45	638.15	70	1121.03	95	2075.04
21	247.07	46	657.25	71	1147.90	96	2144.01
22	265.64	47	677.00	72	1177.15	97	2202.02
23	283.57	48	691.21	73	1200.03	98	2349.03
24	295.68	49	706.03	74	1229.01	99	2487.00
25	315.00	50	720.00	75	1260.00	100	2990.31

Appendix III

Isozyme allelic frequencies in all populations of *A. hemionitis* (Table I) and all populations grouped within islands (Table II).

Table I. Isozyme allele frequencies in all populations of *A. hemionitis*.

locus	allelic mobility	Portugal														Morocco		Azores																			
		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	TA1	TA2	F1	F2	F3	F4	F5	F6	C1	C2	P1	P2	P3	FA1	FA2	FA3	FA4	FA5	FA6	FA7	FA8
AAT-1	n	6	11	22	19	17	5	7	9	29	17	15	17	63	116	39	11	10	23	13	12	14	9	13	10	28	15	4	34	14	11	0	9	30	25	8
	74	.	.	.	.	.	.	.	0.06	.	.	.	.	.	.	.	.	0.05	.	0.15	0.08	.	.	0.46	0.35	0.25	.	.	0.02	.	0.05		0.22	.	0.08	0.06
	88	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.
	100	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.94	1.00	1.00	1.00	1.00	1.00	1.00	0.89	0.96	0.95	1.00	0.85	0.92	1.00	1.00	0.54	0.65	0.75	1.00	1.00	0.97	1.00	0.96		0.78	0.97	0.92	0.88
	115	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.
	118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.12	0.05	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	0.03	.	0.06
	124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.02	.	.		.	.	.	.
ACN-2	n	0	0	0	0	0	0	0	0	0	0	0	0	65	76	38	11	0	14	5	12	7	9	13	7	20	0	4	15	0	0	0	0	0	6	0
	78													.	.	.	.																		.	
	91													0.19	0.30	0.49	0.68																			
	100													0.82	0.70	0.51	0.32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00		1.00	1.00						1.00	
	110																																			
DIA-1	n	6	11	22	19	17	5	7	9	29	17	15	17	64	114	39	11	10	23	16	12	13	9	13	10	30	15	4	34	14	11	28	10	30	26	13
	83																						0.06												0.02	
	100	0.92	1.00	0.89	0.97	1.00	1.00	1.00	1.00	0.74	0.94	0.97	0.77	0.95	0.99	0.87	0.68	1.00	0.94	0.91	0.79	0.96	0.89	0.96	0.80	0.95	1.00	1.00	0.96	1.00	1.00	0.91	0.85	0.92	0.94	1.00
	123	0.08	.	0.11	0.03	.	.	.	.	0.26	0.06	0.03	0.24	0.06	0.01	0.13	0.32	.	0.07	0.09	0.21	0.04	0.06	0.04	0.20	0.05	.	.	0.04	.	.	0.09	0.15	0.08	0.04	.
	141																																			
HEX	n	6	11	22	19	17	5	7	9	27	17	15	16	65	115	39	11	0	23	13	12	14	9	13	10	30	15	4	34	14	11	22	10	30	25	13
	77							0.14	0.11	0.13	0.03	.																									
	100	0.75	1.00	0.86	0.74	0.44	1.00	0.43	0.33	0.63	0.35	0.60	0.34	0.55	0.82	1.00	1.00		0.02	.	0.13	.	0.06	0.31	0.05	.	0.03	.	0.16	0.54	0.05	0.07	.	.	0.30	0.08
	111	.	.	.	.	.	.	.	.	.	.	.	0.31	.	.	.	.																			
	118	0.25	.	0.14	0.26	0.56	.	0.43	0.56	0.24	0.62	0.40	0.34	0.37	0.18	.	.		0.98	1.00	0.88	1.00	0.94	0.69	0.95	1.00	0.97	1.00	0.84	0.46	0.96	0.93	1.00	1.00	0.70	0.92
	126													0.09	.																					
IDH	n	6	11	22	19	17	5	7	9	29	17	15	17	65	115	38	11	10	23	16	12	14	7	13	10	29	15	4	34	14	11	23	10	28	26	13
	57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	77	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	91	.	.	.	.	.	.	.	.	.	.	.	.	0.02	0.15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.75	0.59	0.77	0.47	0.82	0.70	0.50	0.44	0.66	0.62	0.57	0.56	0.35	0.40	0.80	0.86	0.35	0.30	0.41	0.33	0.25	0.29	0.39	0.20	0.43	0.30	0.75	0.37	0.75	0.68	0.41	0.35	0.61	0.73	0.69
	115	0.17	0.36	0.11	0.47	0.18	0.30	0.50	0.56	0.35	0.27	0.23	0.38	0.59	0.45	0.05	.	0.65	0.70	0.59	0.67	0.75	0.71	0.62	0.80	0.57	0.70	0.25	0.63	0.25	0.32	0.59	0.65	0.39	0.27	0.31
	127	0.08	0.05	0.11	0.05	.	.	.	.	.	0.12	0.20	0.06	0.05	0.01	0.15	0.14																			
LAP	n	6	11	22	19	17	5	7	9	29	17	15	17	38	116	39	11	10	22	16	12	14	9	13	10	30	15	4	34	14	11	28	10	10	23	14
	93	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.06	.	.	0.07	.	0.10	.	.
	100	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.74	0.97	0.91	0.77	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.95	1.00	1.00	1.00	0.94	1.00	1.00	0.93	0.80	0.90	1.00	1.00
	102	.	.	.	.	.	.	.	.	.	.	.	.	0.03	0.01	0.23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.20	.	.	.
	105	.	.	.	.	.	.	.	.	.	.	.	0.27	.	0.08	.	.	.	.	.	.	.	.	.	0.05	.	.	.	.	.	.	.	.	.	.	.	.
MDH-1	n	6	11	22	19	17	5	7	9	29	17	15	17	65	116	39	11	10	22	16	12	14	9	13	10	30	15	4	32	14	11	25	10	30	25	14
	87	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.28	0.09	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	94	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.05	.	0.03	.	0.04	0.22	0.15	.	0.15	.	0.50	0.02	0.04	0.18	0.06	.	0.15	.	0.11
	100	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.72	0.91	0.95	1.00	0.97	1.00	0.96	0.78	0.81	1.00	0.85	1.00	0.50	0.98	0.96	0.82	0.94	0.95	0.82	1.00	0.89
	109	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.04	.	.	.	.	.	.	.	.	0.05	0.03	.	.
MDH-2	n	6	11	22	19	17	5	7	9	29	17	15	17	65	116	38	11	5	14	8	11	14	8	13	10	18	0	4	31	14	11	25	10	29	26	14
	58	0.50	0.09	0.11	0.11	0.03	.	0.50	0.17	0.14	0.18	0.20	0.03	0.13	0.02	.	.	.	.	.	.	.	.	.	.	0.03	.	.	.	.	.	.	.	.	.	.
	83	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.50	0.91	0.89	0.90	0.97	1.00	0.50	0.83	0.86	0.82	0.80	0.97	0.87	0.98	0.49	0.77	0.80	.	0.69	0.09	0.21	.	0.42	0.20	0.06	.	0.38	0.29	0.61	0.68	0.30	.	0.35	0.46	0.36
	104	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.07	.	.	.	.	.	.
	106	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.04	.	.	.	.	.	0.07	0.56	.	.	0.50	.	.	.	.	.	.	0.25	.	.	.
	108	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.09	0.04	.	.	0.02	.
	114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.47	0.23	0.20	0.68	0.31	0.68	0.46	0.44	0.23	0.40	0.25	.	0.25	0.34	0.32	0.14	0.54	0.30	0.53	0.50	0.32
	122	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.32	.	0.23	0.25	.	0.35	0.40	0.17	.	0.38	0.37	.	0.09	0.12	0.45	0.12	0.02	0.32
6-PGD	n	6	10	21	19	14	5	7	9	29	17	15	17	63	106	39	11	10	23	16	10	14	9	13	10	29	15	4	34	14	11	27	10	30	23	14
	71	.	.	.	.	.	.	.	.	.	.	.	.	0.01	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	88	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.20	.	0.03	0.60	.	0.22	0.19	0.05	0.19	0.17	0.25	.	0.11	0.05	0.13	0.05	0.30	0.04	.
	95	0.25	.	.	.	0.07	.	0.29	0.33	0.03	0.06	0.20	0.21	0.22	0.30	.	.	.	.	.	.	.	.	.	0.05	.	0.07	.	.	.	.	.	0.10	.	.	.
	100	0.33	1.00	0.52	0.90	0.61	0.60	0.36	0.39	0.40	0.29	0.27	0.56	0.48	0.59	0.96	1.00	0.80	1.00	0.97	0.40	1.00	0.78	0.81	0.75	0.78	0.70	0.75	0.99	0.86	0.91	0.70	0.70	0.47	0.74	0.82
	107	0.42	.	0.48	0.11	0.32	0.40	0.36	0.28	0.57	0.65	0.53	0.24	0.29	0.10	0.04	.	.	.	.	.	.	.	.	0.15	0.03	0.07	.	0.02	0.04	0.05	0.17	0.15	0.23	0.22	0.18
	110	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PGI-2	n	6	11	22	19	17	5	7	9	29	17	15	17	65	116	39	11	9	23	16	12	14	9	13	10	30	15	4	34	14	11	28	10	30	26	14
	20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	35	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	60	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.01	0.05	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.87	0.59	0.94	1.00	1.00	1.00	0.89	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	0.88	1.00	0.93
	115	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	125	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.08	.	.
	130	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.12	0.36	0.06	.	.	.	0.11	.	.	.	.	.	.	0.02	.	.	.	.	0.03	.	0.07
	150	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	185	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PGM-2	n	6	11	22	19	17	5	7	9	29	17	15	17	65	116	39	11	10	23	16	12	14	9	13	10	30	15	4	33	5	11	28	10	30	26	14
	56	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	84	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.89	1.00	1.00	1.00	1.00	0.99	1.00	0.96	0.95	0.80	0.97	0.96	1.00
	119	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.06	.	.	.	.	.	.	.	.	.	.	.	.	0.02	.	0.05	0.05	0.20	0.03	0.04	.
	122	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.12	.	.	.	.	.	.	.	.	.	.	.	.
	129	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	134	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SkDH	n	6	11	22	19	15	5	7	9	29	17	15	17	65	116	39	11	8	23	15	12	14	9	12	10	30	0	4	34	14	11	27	10	30	26	14
	43	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	52	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	59	.	.	.	.	.	.	.	.	.	0.06	.	.	.	.	0.10	0.36	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	60	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	69	0.17	0.23	0.23	0.24	0.23	.	0.07	0.39	0.50	0.03	.	0.21	0.25	0.18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	72	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	76	.	.	.	.	.	.	.	.	.	.	.	0.06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	83	0.25	0.32	0.50	0.34	0.40	0.80	0.79	0.50	0.16	0.53	0.60	0.41	0.15	0.49	0.53	0.55																	0.03	.	.
	86	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	90	0.08	0.23	0.05	0.18	0.07	.	.	.	0.02	0.03	0.10	0.12	0.07	0.03	.	.	0.44	0.24	0.20	0.08	0.25	.	0.21	0.15	0.17	.	.	0.02	0.32	.	0.06	0.35	0.13	0.27	0.21
	96	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.02	.	.	0.36	.	.	0.55	0.13	.	0.13	0.06	.	0.50	0.28	0.65	0.10	0.15	0.14
	100	0.50	0.23	0.23	0.24	0.30	0.20	0.14	0.11	0.33	0.35	0.30	0.21	0.53	0.30	0.37	0.09	0.56	0.61	0.80	0.92	0.39	1.00	0.75	0.30	0.62	.	0.88	0.93	0.68	0.50	0.57	.	0.70	0.58	0.61
	104	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	110	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.13	.	.	.	.	0.04	.	0.08	.	.	.	.	.	0.09	.	0.03	.	0.04
	117	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TPI-1	n	6	11	22	19	17	5	7	9	29	17	15	17	65	116	39	11	10	23	16	12	14	9	13	10	30	15	4	34	14	11	28	10	30	26	14
	81	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	114	.	.	.	.	.	.	.	.	.	.	.	.	0.02	.	0.04	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TPI-2	n	6	11	22	19	17	5	7	9	29	17	15	17	65	116	39	11	10	23	16	12	14	9	13	10	30	15	4	34	14	11	18	10	30	26	14
	65	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	69	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.01	.	.	0.02	0.13	.	.	0.06	.	.	.	.	.	.	.	.	.	.	0.07	0.06	.
	87	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.92	0.86	0.98	1.00	1.00	1.00	0.93	0.89	1.00	1.00	1.00	0.88	0.95	0.91	0.64	0.64	0.65	0.98	0.75	0.46	0.96	0.61	1.00	1.00	0.75	0.87	0.75	0.97	1.00	0.91	0.78	0.55	0.78	0.79	0.68
	123	0.08	0.14	0.02	.	.	.	0.07	0.11	.	.	.	0.12	0.05	0.09	0.35	0.36	0.35	.	0.13	0.54	0.04	0.33	.	.	0.25	0.13	0.25	0.03	.	0.09	0.22	0.45	0.15	0.15	0.32
UGPP	n	6	11	22	19	17	5	7	9	29	17	15	17	65	116	39	11	0	14	8	12	14	9	13	10	20	15	4	34	14	11	28	10	30	26	14
	79	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	92	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.04	.	.	0.04	.	.	.	.	.	.	.	.	.	.	.	.
	100	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00		1.00	0.94	0.96	1.00	1.00	0.96	1.00	0.98	0.97	1.00	1.00	1.00	1.00	0.98	1.00	0.98	1.00	1.00
	105	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	108	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.03	0.03	.	.	.	.	0.02	.	0.02	.	.

Table I. Continued.

locus	allelic mobility	Azores																				Madeira														
		FA 9	FA 10	FA 11	FA 12	FA 13	SM 1	SM 2	SM 3	SM 4	SM 5	SM 6	SM 7	SM 8	SM 9	SM 10	SM 11	SM 12	SM 13	SM 14	SM 15	M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13	M 14
AAT-1	n	23	17	21	23	5	26	8	9	23	24	11	20	8	13	13	5	0	26	7	16	5	8	67	10	18	8	13	14	18	17	27	5	13	4
	74	0.02	.	.	0.04	0.10	0.29	0.25	0.06	0.30	0.19	.	0.20	0.44	0.08	0.31	0.50		0.08	0.21	0.78	.	.	0.12	0.25	0.08	0.13	.	0.07	0.06	0.29	0.09	.	0.19	0.38
	88	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.50	.	0.02	.	.	.	.	.	.	.	.	.	.	.
	100	0.98	1.00	1.00	0.96	0.80	0.60	0.75	0.72	0.63	0.79	0.82	0.75	0.56	0.92	0.65	0.40		0.81	0.71	0.22	0.50	1.00	0.80	0.75	0.89	0.81	0.96	0.68	0.81	0.56	0.85	0.80	0.73	.
	115	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	118	.	.	.	.	0.10	0.04	.	.	0.07	.	.	0.05	.	.	.	.	.	0.12	0.07	.	.	.	0.06	.	0.03	0.06	0.04	0.25	0.14	0.15	0.06	0.20	0.08	0.63
	124						0.08	.	0.22	.	0.02	0.18	.	.	.	0.04	0.10																			
ACN-2	n	0	17	20	0	5	5	0	8	30	30	9	0	0	7	0	5	4	6	7	20	0	0	65	10	18	12	13	14	18	17	27	5	13	4
	78	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.39	.	.	.	.	.	.	.
	91	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100		1.00	1.00		1.00	1.00		1.00	1.00	1.00	1.00			1.00		1.00	1.00	1.00	1.00	1.00			0.99	1.00	1.00	1.00	0.62	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	110																							0.02											
DIA-1	n	26	17	28	24	5	26	10	9	19	30	15	21	8	29	21	11	10	32	7	26	5	8	67	10	18	12	13	14	18	17	27	5	13	4
	83	.	.	0.02	.	.	.	0.10	.	.	.	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	0.04	0.03	.	.	0.30	.	.
	100	0.65	0.91	0.93	0.96	0.70	0.77	0.90	1.00	1.00	0.83	1.00	0.98	0.94	0.88	0.93	1.00	0.70	0.97	1.00	0.94	0.70	0.25	0.50	0.70	0.19	0.38	0.92	0.46	0.42	0.27	0.57	0.10	0.35	0.38
	123	0.35	0.09	0.05	0.04	0.30	0.23	.	.	.	0.17	.	0.02	0.06	0.12	0.07	.	0.30	0.02	.	0.06	0.30	0.75	0.48	0.30	0.61	0.54	0.08	0.50	0.53	0.74	0.43	0.60	0.62	0.63
	141																							0.02		0.19	0.08			0.03				0.04	
HEX	n	27	17	30	24	5	25	10	5	30	30	15	18	7	28	20	11	10	32	7	26	5	8	67	10	18	12	13	14	18	17	24	5	13	4
	77	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.17	0.24	.	0.06	.	0.46	0.40	0.60	0.55	0.58	0.40	0.58	0.29	0.50	0.55	0.41	0.55	0.28	0.79	0.52	.													
	111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.09	.	.	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.96	0.94	1.00	0.98	1.00	0.96	1.00
	118	0.83	0.77	1.00	0.94	1.00	0.54	0.60	0.40	0.45	0.42	0.60	0.42	0.71	0.50	0.45	0.59	0.45	0.63	0.21	0.48	.	.	.	.	.	.	.	0.04	0.06	.	0.02	.	0.04	.
	126																																		
IDH	n	27	17	29	24	5	22	9	9	30	30	15	21	8	29	21	11	10	31	7	25	5	7	65	10	18	12	13	14	18	17	27	5	13	4
	57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.13	.	.	.	.	.	.	.	.
	77	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.01	.	.	.	.	.	.	0.12	.	.	0.04	.
	91	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.37	0.35	0.83	0.75	0.80	0.25	0.61	0.50	0.22	0.20	0.50	0.05	0.50	0.28	0.55	0.36	0.40	0.29	0.07	0.16	0.20	0.21	0.42	.	0.06	0.08	0.12	0.18	0.28	0.29	0.19	0.30	0.27	0.38
	115	0.63	0.65	0.17	0.25	0.20	0.75	0.39	0.50	0.78	0.80	0.50	0.95	0.50	0.72	0.45	0.64	0.60	0.71	0.93	0.84	0.70	0.79	0.52	1.00	0.94	0.79	0.89	0.82	0.72	0.56	0.82	0.70	0.69	0.63
	127																					0.10		0.06							0.03				
LAP	n	27	17	30	24	0	26	10	9	29	30	6	6	8	22	21	10	10	32	7	26	5	8	67	10	18	12	13	14	18	17	26	5	13	4
	93	.	.	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.13	0.02	0.20	.	.	.	.	.	0.03	0.06	.	.	.
	97	.	0.03	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.20	0.06	0.02	0.30	.	.	0.08	.	0.03	0.24	.	.	0.04	.
	100	1.00	0.97	1.00	1.00		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.80	0.81	0.96	0.50	1.00	1.00	0.92	1.00	0.97	0.74	0.94	0.90	0.96	1.00
	102	.	.	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.10	.	.
	105	.	.	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MDH-1	n	27	16	30	24	5	25	10	7	30	30	15	19	8	29	21	11	10	32	7	26	5	8	67	10	18	12	13	14	18	17	27	5	13	4
	87	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.01	.	.	.	.	.	.	.	.	.	.	.
	94	0.22	0.19	0.17	0.17	0.40	0.02	0.10	.	0.08	.	0.03	0.26	0.13	0.24	0.12	0.14	0.30	0.02	0.14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.78	0.81	0.83	0.83	0.60	0.98	0.90	1.00	0.92	1.00	0.97	0.74	0.88	0.76	0.88	0.86	0.70	0.98	0.86	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	109	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MDH-2	n	27	17	29	23	5	25	10	7	30	29	15	19	8	29	21	11	10	29	7	26	5	8	67	10	18	12	13	14	18	17	27	5	13	4
	58	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.09	.	.	.	0.08	0.11	.	0.15	.	.	0.04	0.13
	83	.	.	.	.	.	.	.	.	0.03	.	.	.	.	.	.	.	.	.	.	.	.	.	0.01	.	.	.	.	0.04	.	0.03	.	0.10	.	.
	100	0.52	0.27	0.33	0.24	0.20	0.26	0.60	0.71	0.78	0.78	0.50	0.53	0.94	0.52	0.38	0.50	0.40	0.48	0.43	0.33	1.00	1.00	0.89	1.00	0.97	1.00	0.92	0.86	1.00	0.82	1.00	0.90	0.92	0.88
	104	.	0.06	.	0.13	.	.	.	.	.	0.03	.	.	.	.	0.24	0.32	0.10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	106	.	.	0.03	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	108	.	.	.	.	.	.	.	.	.	.	0.50	0.11	.	0.21	0.12	.	0.20	0.03	0.14	0.08	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	114	0.28	0.18	0.40	0.39	0.70	0.36	0.15	.	0.18	0.19	.	0.32	0.06	0.17	0.17	0.09	0.20	0.41	0.07	0.25	.	.	0.02	.	0.03	.	.	.	.	.	.	.	0.04	.
	122	0.20	0.50	0.24	0.24	0.10	0.38	0.25	0.29	.	.	.	0.05	.	0.10	0.10	0.09	0.10	0.07	0.36	0.35														
6-PGD	n	27	17	30	24	5	26	10	9	30	30	15	21	8	29	21	11	10	26	7	26	5	8	62	10	18	12	13	14	18	17	26	5	13	4
	71	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.20	.	0.11	0.25	.	0.04	.	0.04	0.17	0.12	0.14	0.10	.	.
	88	0.06	.	0.25	0.29	0.30	0.06	0.05	.	0.08	0.17	0.37	0.33	0.44	0.62	0.43	0.32	0.25	0.58	0.43	0.29	0.60	.	.	.	.	.	.	.	.	.	.	.	.	.
	95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.02	.	.	.	.	0.04	.	.	.	.	.	.	.
	100	0.94	1.00	0.63	0.56	0.70	0.94	0.70	0.78	0.92	0.73	0.57	0.55	0.31	0.38	0.41	0.64	0.50	0.42	0.43	0.67	0.20	1.00	0.86	0.70	0.97	0.92	1.00	0.93	0.75	0.71	0.85	0.90	0.92	1.00
	107	.	.	0.12	0.15	.	.	0.25	0.22	.	0.02	0.07	0.12	0.25	.	0.17	0.05	0.25	.	0.14	0.04	.	.	0.02	0.05	.	0.04	.	.	0.08	.	.	.	.	.
	110										0.08															0.03					0.18	0.02		0.08	
PGI-2	n	27	17	30	24	4	24	10	9	30	30	15	21	8	29	21	11	10	32	7	26	5	8	63	10	18	12	13	14	17	17	27	5	13	4
	20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	35	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	60	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.06	.	.	.	.	.	.	.	.	0.02	.	.	.
	100	1.00	1.00	0.92	0.90	1.00	0.98	0.95	1.00	0.97	0.85	1.00	0.95	0.81	0.91	0.98	0.91	0.90	0.91	1.00	1.00	0.50	0.50	0.44	0.70	0.25	0.46	0.54	0.18	0.27	0.29	0.32	0.50	0.42	0.75
	115	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	.
	125	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.10	.	.	.	.	.	0.18	0.05	0.36	0.29	0.46	0.07	0.44	0.27	0.35	0.10	0.15	.
	130	.	.	0.08	0.10	.	.	.	.	0.03	0.15	.	0.05	0.19	0.09	0.02	0.09	.	0.09	.	.	0.50	0.44	0.30	0.25	0.39	0.25	.	0.75	0.29	0.29	0.30	0.40	0.42	0.25
	150	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	155	.	.	.	.	.	.	0.05	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.06	.	.	.	.	.	.	.	0.02	.	.	.
	185	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.15	.	.	.	.
PGM-2	n	27	17	29	23	5	24	10	9	30	30	15	21	8	29	21	11	10	32	7	26	5	0	67	10	18	11	13	14	18	17	27	5	13	4
	56	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	84	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.94	0.97	0.97	0.96	0.70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.30		0.86	1.00	0.97	0.86	1.00	0.79	0.86	0.97	0.70	1.00	0.92	0.75
	119	0.06	0.03	.	.	0.30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.70	.	0.14	.	0.03	0.14	.	0.21	0.14	0.03	0.30	.	0.08	0.25
	122	.	.	0.03	0.04	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	129	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	134	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SkDH	n	27	17	30	24	5	23	10	9	30	30	14	21	8	29	20	11	10	32	7	25	5	8	67	10	18	11	13	14	18	16	27	5	13	4
	43	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	52	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	59	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	0.08	.
	60	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	69	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	72	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	76	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.04	.	.	.	.	.	.	.	.	.	0.04	.
	83	.	.	.	.	.	.	.	.	.	0.08	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	86	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	90	0.02	.	0.33	0.23	0.10	0.41	0.40	0.50	0.23	0.10	0.36	0.02	0.25	0.05	0.23	0.27	0.10	0.08	0.36	0.18	0.50	0.69	0.55	0.50	0.69	0.86	0.92	0.39	0.69	0.41	0.39	0.50	0.54	0.50
	96	0.19	0.47	0.22	0.06	.	0.07	0.40	0.17	.	0.02	0.04	.	0.25	0.03	0.10	0.05	.	0.02	.	.															
	100	0.80	0.53	0.42	0.71	0.90	0.50	.	.	0.77	0.68	0.61	0.86	0.38	0.91	0.68	0.68	0.90	0.91	0.50	0.48	0.20	0.06	0.17	0.50	0.17	0.09	.	0.07	0.11	0.34	0.46	0.10	0.23	0.13
	104	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.10	.	.	.	.	0.18	0.17	0.06	0.06	0.20	0.04	.
	110	.	.	0.03	.	.	0.02	0.20	0.33	.	0.12	.	0.12	0.13	.	.	.	.	.	0.14	0.34	0.30	0.25	0.13	.	0.14	0.05	0.08	0.36	0.03	0.19	0.09	0.20	0.08	0.38
	117	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TPI-1	n	27	17	30	24	5	25	10	9	30	30	15	21	8	29	21	10	10	32	7	26	5	8	67	10	18	12	13	14	18	17	27	5	13	4
	81	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	.
	100	1.00	1.00	1.00	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.85	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.15	.	.	.	.	.	.	.	.	.	.
	118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TPI-2	n	27	17	30	24	5	25	10	9	30	30	15	21	8	29	21	11	10	32	7	26	5	8	67	10	18	12	13	14	18	17	27	5	13	4
	65	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	69	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.13	0.02	0.10	0.22	.	.	.	0.19	0.12	0.02	0.10	0.04	.
	87	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.98	0.94	0.73	0.90	0.60	0.98	1.00	1.00	0.72	0.85	0.73	0.95	0.94	0.91	0.79	0.91	0.80	0.77	1.00	0.52	0.40	0.44	0.07	0.10	0.25	0.38	0.04	0.29	0.36	0.27	0.37	0.40	0.19	
	123	0.02	0.06	0.27	0.10	0.40	0.02	.	.	0.28	0.15	0.27	0.05	0.06	0.09	0.21	0.09	0.20	0.23	.	0.48	0.60	0.44	0.92	0.80	0.53	0.63	0.96	0.71	0.44	0.62	0.61	0.50	0.77	1.00
UGPP	n	27	17	30	24	5	26	10	9	30	30	15	21	8	29	21	11	10	32	7	26	5	8	67	10	18	12	13	14	18	17	27	5	13	4
	79	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	92	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	.	.	0.05	.	.	.	.	.	0.02	0.05	.	.	.	.	.	.	.	.	.	.
	100	1.00	1.00	1.00	0.98	1.00	1.00	1.00	1.00	1.00	1.00	0.97	1.00	1.00	1.00	1.00	1.00	0.95	1.00	1.00	1.00	0.50	0.56	0.53	0.85	0.72	0.38	0.96	0.71	0.78	0.68	0.54	0.30	0.73	0.50
	105	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	108	.	.	.	.	.	.	.	.	.	.	0.03	.	.	.	.	.	.	.	.	.	0.50	0.44	0.45	0.10	0.28	0.63	0.04	0.29	0.22	0.32	0.46	0.70	0.27	0.50

		Canary Islands																																	
locus	*allelic mobility*	T 1	T 2	T 3	T 4	T 5	T 6	T 7	T 8	T 9	T 10	T 11	GC 1	GC 2	GC 3	GC 4	GC 5	EH 1	EH 2	EH 3	EH 4	EH 5	EH 6	EH 7	EH 8	EH 9	LP 1	LP 2	LP 3	LP 4	LP 5	LP 6	LP 7	LG 1	LG 2	LG 3	LG 4
	n	0	0	12	25	1	10	0	14	10	18	24	23	24	30	6	14	10	5	35	7	8	5	27	7	23	38	12	36	9	44	10	13	29	19	28	9
AAT-1	74			.	.	.	.		.	.	.	0.10	0.07	0.19	0.18	0.25	0.32	.	.	.	.	.	.	.	.	0.07	0.04	.	0.03	0.06	.	0.10	0.15	.	.	0.05	.
	88			.	.	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100			0.88	0.66	1.00	0.95		1.00	0.65	1.00	0.83	0.94	0.81	0.77	0.75	0.68	0.75	1.00	1.00	1.00	1.00	1.00	0.87	0.86	0.87	0.88	0.92	0.82	0.89	0.99	0.80	0.85	0.98	0.92	0.80	1.00
	115			.	.	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	.	.	.
	118			0.13	.	.	0.05		.	0.35	.	0.06	.	.	0.05	.	.	0.25	.	.	.	.	.	0.11	0.14	0.07	0.08	0.08	0.15	0.06	0.01	0.10	.	0.02	0.08	0.14	.
	124			.	0.34	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	n	31	6	11	7	8	10	17	14	0	15	0	30	25	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ACN-2	78	.	.	.	.	.	.	.	.		.		0.02	.		0.08																					
	91	0.13	0.17	0.09	0.07	0.13	0.20	0.12	.		.		0.23	0.20		0.33																					
	100	0.87	0.83	0.82	0.93	0.88	0.80	0.88	1.00		1.00		0.75	0.80		0.58																					
	110	.	.	0.09	.	.	.	.	.		.		.	.		.																					
	n	30	6	12	25	9	10	17	49	10	35	25	30	38	30	6	14	9	5	35	7	8	5	27	7	23	38	12	36	9	44	6	13	29	19	30	11
DIA-1	83	0.03	.	.	.	.	.	.	0.06	.	0.03	.	0.02	.	.	.	.	.	.	.	.	.	0.20	0.09	.	0.02	.	0.08	0.01	0.06	0.07	.	.	0.09	.	0.05	.
	100	0.83	0.75	0.96	1.00	1.00	1.00	0.97	0.80	1.00	0.87	0.98	0.97	0.97	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	0.80	0.89	0.71	0.96	1.00	0.88	0.97	0.94	0.93	0.83	1.00	0.90	1.00	0.93	0.82
	123	0.13	0.25	0.04	.	.	.	0.03	0.14	.	0.10	0.02	0.02	0.03	.	.	.	.	.	0.01	.	.	.	0.02	0.29	0.02	.	0.04	0.01	.	.	0.17	.	0.02	.	0.02	0.18
	141	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	n	31	6	11	25	9	10	17	49	11	34	20	30	38	30	6	14	10	5	31	0	8	0	27	7	23	38	12	36	9	44	10	13	11	19	30	11
HEX	77	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.		.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	1.00	1.00	1.00	1.00	1.00	1.00	0.97	1.00	1.00	1.00	0.50	1.00	1.00	0.98	1.00	1.00	0.35	.	0.39		0.06		0.28	0.21	0.20	.	.	0.14	.	.	.	.	.	.	.	0.09
	111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1.00	0.32		0.94		.	.	.	.	.	0.08	.	.	.	.	.	0.42	.	0.91
	118	.	.	.	.	.	.	0.03	.	.	.	0.50	.	.	0.02	.	.	0.65	.	0.29		.		0.72	0.79	0.80	1.00	1.00	0.78	1.00	1.00	1.00	1.00	1.00	0.58	1.00	.
	126	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.		.	.	.	.	.	.	.	.	.	.	.	.	.	.
	n	31	5	12	25	9	10	17	46	11	34	25	30	38	29	6	13	10	4	32	7	8	5	19	7	21	31	0	34	0	14	0	0	29	18	26	9
IDH	57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.		.			.	.	.	.
	77	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.		.			.	.	.	.
	91	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.		.			.	.	.	.
	100	0.48	0.40	0.13	0.46	0.44	0.30	0.29	0.34	0.36	0.38	0.16	0.57	0.37	0.52	0.42	0.27	0.60	0.50	0.52	0.50	0.69	0.50	0.47	0.29	0.57	0.61		0.43		0.57			0.71	0.28	0.33	0.11
	115	0.52	0.40	0.88	0.54	0.56	0.70	0.71	0.64	0.64	0.56	0.78	0.18	0.45	0.35	0.58	0.19	0.40	0.50	0.48	0.50	0.31	0.50	0.53	0.43	0.43	0.39		0.57		0.43			0.29	0.72	0.67	0.83
	127	.	0.20	.	.	.	.	.	0.02	.	0.06	0.06	0.25	0.18	0.14	.	0.54	.	.	.	.	.	.	.	0.29	.	.		.		.			.	.	.	0.06
	n	31	6	12	25	0	0	17	49	11	35	0	30	14	30	0	14	10	0	0	7	0	5	0	7	0	12	0	12	0	13	0	0	28	0	0	0
LAP	93	.	.	.	.			.	.	.	0.09		.	.	.		.	.			.		.		.		.		.		.			.			
	97	0.08	0.17	0.04	.			0.06	0.08	.	0.09		0.05	0.11	0.07		.	0.15			.		.		.		0.04		0.04		0.31			0.09			
	100	0.36	0.83	0.83	0.94			0.50	0.66	1.00	0.71		0.68	0.71	0.92		0.57	0.70			1.00		1.00		0.93		0.83		0.83		0.69			0.91			
	102	0.44	.	.	0.06			0.41	0.25	.	0.11		0.08	0.14	.		.	0.15			.		.		.		.		.		.			.			
	105	0.13	.	0.13	.			0.03	0.01	.	.		0.18	0.04	0.02		0.43	.			.		.		0.07		0.13		0.13		.			.			
	n	31	6	12	25	9	10	17	49	11	35	25	30	38	30	6	14	10	5	35	7	8	5	27	7	23	38	12	36	9	44	10	13	29	19	30	11
MDH-1	87	.	.	.	.	0.22	.	.	0.03	.	0.01	.	0.45	0.09	0.12	0.08	0.25	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	94	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	1.00	1.00	1.00	1.00	0.78	0.95	1.00	0.97	1.00	0.99	0.96	0.50	0.91	0.88	0.92	0.75	1.00	1.00	1.00	1.00	1.00	1.00	0.98	0.64	1.00	0.99	1.00	0.99	1.00	1.00	1.00	1.00	1.00	0.95	1.00	1.00
	109	.	.	.	.	.	0.05	.	.	.	.	0.04	0.05	.	.	.	.	.	.	.	.	.	.	0.02	0.36	.	0.01	.	0.01	.	.	.	.	.	0.05	.	.
	n	31	6	12	25	9	10	17	49	10	35	24	30	38	30	6	14	10	5	33	7	6	4	27	7	23	35	12	33	9	41	10	13	26	19	30	11
MDH-2	58	.	.	.	.	.	.	.	0.03	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.04	.	.	.
	83	.	.	.	.	.	0.05	.	.	.	.	.	0.05	.	0.02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.69	0.75	0.46	0.58	0.61	0.75	0.74	0.71	0.55	0.74	0.60	0.15	0.67	0.47	0.25	0.46	0.70	0.50	0.67	0.93	0.33	0.88	0.37	0.71	0.59	0.44	0.42	0.61	0.50	0.48	0.60	0.46	0.62	0.37	0.48	0.32
	104	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	106	.	.	.	.	.	.	.	.	.	.	0.02	0.33	.	.	.	.	.	.	.	.	.	.	.	.	.	0.03	.	0.03	.	.	.	.	.	.	.	.
	108	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.21	0.13	.	0.22	0.09	0.10	0.15	.	.	.	.
	114	0.31	0.25	0.54	0.42	0.39	0.20	0.27	0.26	0.45	0.26	0.38	0.47	0.33	0.52	0.75	0.54	0.30	0.50	0.33	0.07	0.67	0.13	0.63	0.29	0.41	0.31	0.46	0.36	0.28	0.44	0.30	0.39	0.35	0.63	0.52	0.68
	122	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	n	31	6	12	25	9	10	17	49	11	31	22	30	38	30	6	14	10	4	12	0	2	0	27	3	21	35	11	33	9	42	10	13	26	19	30	11
6-PGD	71	0.03	.	.	.	0.06	0.15	.	0.01	0.09	0.02	0.21	0.10	0.04	0.02	.	.	.	.	.		.		.	.	0.05	.	.	0.02	0.06	.	0.05	.	.	.	0.15	0.09
	88	0.26	0.17	0.13	0.10	0.06	.	0.12	0.10	.	0.19	0.02	.	.	.	.	.	.	.	0.04		.		.	.	0.02	.	.	0.02	0.06	0.02	0.05	.	0.02	0.29	0.30	0.59
	95	.	.	.	.	.	.	.	0.03	.	.	.	.	.	.	0.17	.	.	.	.		.		.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.68	0.75	0.83	0.82	0.83	0.80	0.79	0.75	0.82	0.76	0.55	0.90	0.93	0.98	0.83	1.00	1.00	1.00	0.96		1.00		0.50	0.83	0.86	0.90	0.82	0.86	0.78	0.81	0.90	0.69	0.90	0.58	0.45	0.27
	107	0.02	.	.	0.02	0.06	.	0.09	0.05	0.09	0.02	0.23	.	0.03	.	.	.	.	.	.		.		0.50	0.17	0.07	0.10	0.18	0.11	0.11	0.17	.	0.31	.	0.13	0.10	0.05
	110	0.02	0.08	0.04	0.06	.	0.05	.	0.06	.	0.02	.	.	.	.	.	.	.	.	.		.		.	.	.	.	.	.	.	.	.	.	0.08	.	.	.
	n	31	6	12	25	9	10	17	49	11	35	25	28	38	30	6	13	10	5	34	7	8	5	27	7	23	38	12	36	9	43	10	13	29	19	30	9
PGI-2	20	0.02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	35	0.02	.	.	.	.	.	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.04	.	.	.	.
	60	.	.	.	0.08	.	0.05	.	.	.	.	.	0.13	0.03	0.05	.	0.27	.	0.50	0.04	.	.	.	0.02	.	0.02	.	0.08	0.04	.	0.06	.	.	0.03	.	.	.
	100	0.97	1.00	1.00	0.92	0.94	0.95	1.00	0.98	0.91	0.87	0.98	0.70	0.95	0.87	0.83	0.73	1.00	0.40	0.81	0.86	0.88	0.70	0.94	0.93	0.89	0.88	0.83	0.93	1.00	0.80	0.95	0.85	0.85	0.29	0.50	0.28
	115	.	.	.	.	.	.	.	.	.	0.06	.	.	.	0.02	.	.	.	.	.	0.07	.	.	.	.	.	.	.	.	.	.	.	.	.	0.16	.	0.11
	125	.	.	.	.	0.06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	130	.	.	.	.	.	.	.	.	.	0.07	.	0.18	0.03	0.05	.	.	.	0.10	0.13	0.07	0.13	0.30	0.04	0.07	0.07	0.12	0.08	0.03	.	0.14	0.05	0.12	0.12	0.50	0.47	0.44
	150	.	.	.	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.03	0.06
	155	.	.	.	.	.	.	.	.	0.09	.	.	.	.	.	0.17	.	.	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	0.05	.	0.11
	185	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	n	31	6	12	25	9	10	17	49	11	35	24	30	38	26	6	13	10	5	34	7	8	5	24	7	23	38	11	34	9	39	8	10	29	17	30	11
PGM-2	56	0.07	0.08	.	0.02	0.06	0.10	0.03	0.05	0.05	.	0.15	.	0.03	.	.	.	.	.	.	.	0.13	.	0.02	.	0.02	0.01	0.05	.	0.22	0.15	.	.	0.02	.	.	.
	84	.	.	.	.	.	0.20	0.09	0.06	.	.	0.08	.	0.05	.	.	.	.	.	.	.	.	.	0.04	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.76	0.83	0.96	0.90	0.78	0.70	0.82	0.77	0.77	0.89	0.63	0.98	0.86	0.96	0.92	0.89	0.20	0.90	0.49	0.36	0.81	0.80	0.67	0.36	0.76	0.68	0.77	0.59	0.61	0.68	0.56	0.25	0.71	0.65	0.70	0.55
	119	0.11	0.08	0.04	0.08	0.17	.	0.06	0.08	0.18	0.11	0.15	0.02	0.07	0.04	0.08	0.12	0.80	0.10	0.52	0.64	0.06	0.20	0.25	0.64	0.22	0.30	0.18	0.41	0.17	0.17	0.44	0.75	0.28	0.35	0.30	0.46
	122	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	129	0.07	.	.	.	.	.	.	0.03	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	.	.	.
	134	.	.	.	.	.	.	.	0.01	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	n	31	6	12	22	9	10	17	47	11	34	24	29	38	30	6	14	10	4	31	6	8	5	27	7	22	34	11	35	7	41	10	10	28	19	29	11
SkDH	43	.	.	.	.	.	.	.	0.03	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	52	.	.	.	.	.	.	.	.	.	0.04	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.07	0.01	.	0.05	.	.	.	.
	59	0.08	0.08	.	0.25	0.06	0.15	0.15	0.07	.	0.02	0.15	0.05	0.03	0.12	0.08	0.11	0.10	.	0.34	.	0.06	.	0.07	.	.	0.06	.	0.04	.	0.13	.	0.05	0.04	.	.	.
	60	.	.	0.04	.	.	.	.	0.03	.	.	0.02	.	.	0.02	.	.	.	.	.	.	0.13	.	.	.	.	.	.	0.04	.	.	.	.	.	.	.	.
	69	0.02	.	0.08	.	0.17	.	.	0.01	.	0.02	.	.	0.03	.	.	.	.	.	.	.	.	.	0.09	.	.	0.02	.	.	.	.	.	.	0.16	0.03	.	.
	72	.	.	0.17	0.11	.	.	0.06	0.13	.	0.07	.	0.02	0.01	.	.	.	.	.	.	.	.	.	.	.	0.11	.	0.09	.	.	.	0.05	.	0.09	.	.	.
	76	0.39	0.50	0.04	0.27	0.22	0.10	0.03	0.38	0.05	0.49	0.40	0.02	0.12	0.03	0.42	0.04	0.25	0.75	0.16	.	.	.	0.15	.	0.34	0.10	0.27	0.36	0.14	0.11	0.25	0.10	0.07	0.32	0.35	0.50
	83	0.36	0.08	0.08	0.09	0.11	0.30	0.38	0.15	0.36	0.10	.	0.26	0.33	0.32	.	0.25	0.55	.	0.40	0.58	.	0.30	0.24	0.64	0.21	0.52	0.27	0.23	0.50	0.29	0.15	0.10	0.16	0.03	.	.
	86	.	.	0.04	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.01	.	.	.	.	.	0.05
	90	0.03	0.25	0.04	0.07	0.11	.	0.18	0.03	.	0.04	0.04	0.02	0.04	.	0.17	.	.	.	0.02	.	0.44	0.10	0.04	.	0.02	0.04	0.14	0.07	0.14	0.20	0.10	0.25	0.23	0.53	0.40	0.27
	96	0.07	.	0.04	.	.	0.25	0.03	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.04	.	.	.	.	.	.	0.01	.	0.20	.	.	0.09	.
	100	0.07	0.08	0.13	0.21	0.22	0.05	0.15	0.14	0.59	0.18	0.31	0.64	0.42	0.52	0.25	0.61	0.10	0.25	0.08	0.42	0.38	0.40	0.24	0.07	0.30	0.24	0.23	0.24	0.14	0.23	0.35	0.20	0.20	0.11	0.17	0.14
	104	.	.	.	.	.	0.05	.	.	.	.	.	.	.	.	0.08	.	.	.	.	.	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.
	110	.	.	.	.	0.06	0.05	0.03	0.01	.	.	0.02	.	0.03	.	.	.	.	.	.	.	.	.	0.09	.	.	0.02	.	.	.	.	0.10	.	0.04	.	.	0.05
	117	.	.	.	.	.	0.05	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.20	0.04	0.29	0.02	.	.	0.01	.	.	.	0.05	.	.	.	.
	124	.	.	0.33	.	0.06	.	.	0.01	.	0.04	0.04	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.02	.	.	.
	n	31	6	12	25	9	10	17	49	11	35	25	30	38	30	6	14	10	5	34	7	8	5	27	7	23	38	12	36	9	44	10	13	29	19	30	11
TPI-1	81	.	.	.	0.08	.	.	.	0.01	.	.	.	.	.	0.03	.	.	0.05	.	0.07	0.36	.	.	0.06	.	.	0.07	.	0.01	.	.	.	.	.	.	0.42	0.50
	100	1.00	0.92	1.00	0.92	1.00	0.90	1.00	0.94	0.91	0.99	1.00	1.00	0.95	0.97	1.00	1.00	0.85	1.00	0.52	0.64	1.00	1.00	0.94	1.00	1.00	0.92	1.00	0.97	1.00	1.00	1.00	0.92	1.00	1.00	0.58	0.50
	114	.	0.08	.	.	.	0.10	.	0.05	0.09	0.01	.	.	0.05	.	.	.	0.05	.	0.13	.	.	.	.	.	.	0.01	.	0.01	.	.	.	0.08	.	.	.	.
	118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.05	.	0.28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	n	31	6	12	25	9	10	17	49	11	35	25	30	38	30	6	14	10	5	34	7	8	5	27	7	23	38	12	36	9	44	10	13	29	19	30	11
TPI-2	65	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.06	.	0.04	.	.	.	.	.	.	.	.	.	.	.
	69	0.05	.	0.04	0.04	.	0.10	.	0.11	0.05	0.03	0.02	.	.	0.02	.	.	0.05	.	.	.	.	.	0.09	0.07	0.09	0.04	0.04	0.06	.	.	.	0.04	0.02	0.03	.	.
	87	.	.	.	.	.	.	.	.	0.05	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.89	1.00	0.88	0.96	1.00	0.90	1.00	0.88	0.91	0.97	0.98	0.92	0.95	0.90	1.00	0.96	0.85	1.00	0.90	1.00	1.00	0.90	0.85	0.93	0.72	0.90	0.75	0.85	0.89	1.00	1.00	0.96	0.83	0.95	1.00	1.00
	123	0.07	.	0.08	.	.	.	.	0.01	.	.	.	0.08	0.05	0.08	.	0.04	0.10	.	0.10	.	.	0.10	.	.	0.15	0.07	0.21	0.10	0.11	.	.	.	0.16	0.03	.	.
	n	31	5	12	25	9	10	17	49	11	35	25	29	38	30	6	14	10	5	35	7	8	5	27	7	23	38	12	36	9	44	10	13	29	19	30	11
UGPP	79	.	.	.	.	.	0.05	.	.	.	.	.	0.02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	92	0.07	.	.	.	.	.	0.03	.	.	.	0.06	.	.	.	.	.	.	0.10	.	.	.	.	0.15	.	.	.	.	.	0.06	0.01	0.05	.	0.02	.	.	.
	100	0.92	1.00	0.92	0.50	1.00	0.95	0.82	0.98	0.91	0.99	0.78	0.95	0.93	1.00	0.83	1.00	1.00	0.90	1.00	1.00	0.44	1.00	0.78	1.00	0.98	1.00	1.00	1.00	0.94	0.99	0.95	1.00	0.93	1.00	1.00	1.00
	105	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	108	0.02	.	0.08	0.50	.	.	0.15	0.02	0.09	0.01	0.16	0.03	0.07	.	.	.	.	.	.	.	0.56	.	0.07	.	0.02	.	.	.	.	.	.	.	0.05	.	.	.

Table II. Isozyme allele frequencies in all populations of *A. hemionitis* grouped by islands.

				Azores						Canary Islands				
locus	*allelic mobility*	Sintra	Morocco	Flores	Corvo	Pico	Faial	São Miguel	Madeira	Tenerife	Gran Canaria	El Hierro	La Palma	La Gomera
	n	353	50	81	23	47	220	209	227	114	97	127	162	85
AAT-1	74	0.001	.	0.043	0.413	0.149	0.034	0.251	0.117	0.022	0.180	0.012	0.037	0.018
	88	.	.	.	.	.	.	.	0.018	.	.	.	.	.
	100	0.999	0.900	0.957	0.587	0.851	0.955	0.679	0.778	0.842	0.804	0.921	0.892	0.912
	115	.	.	.	.	.	.	.	.	.	.	0.004	.	.
	118	.	0.100	.	.	.	0.009	0.033	0.088	0.061	0.015	0.063	0.071	0.071
	124	.	.	.	.	.	0.002	0.036	.	0.075	.	.	.	.
	n	141	49	47	20	24	63	131	216	119	61	0	0	0
ACN-2	78	.	.	.	.	.	.	.	0.023	.	0.016	na	na	na
	91	0.248	0.531	.	.	.	.	.	.	0.097	0.230	na	na	na
	100	0.752	0.469	1.000	1.000	1.000	1.000	1.000	0.972	0.895	0.754	na	na	na
	110	.	.	.	.	.	.	.	0.005	0.008	.	na	na	na
	n	352	50	83	23	49	266	274	231	228	118	126	158	89
DIA-1	83	.	.	0.006	.	.	0.004	0.005	0.011	0.022	0.004	0.032	0.032	0.045
	100	0.940	0.830	0.916	0.891	0.969	0.908	0.914	0.461	0.901	0.983	0.940	0.956	0.921
	123	0.060	0.170	0.078	0.109	0.031	0.088	0.080	0.498	0.077	0.013	0.028	0.013	0.034
	141	.	.	.	.	.	.	.	0.030	.	.	.	.	.
	n	351	50	71	23	49	262	274	228	223	118	111	162	71
HEX	77	0.017	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.668	1.000	0.035	0.196	0.010	0.128	0.489	.	0.953	0.996	0.266	0.031	0.014
	111	0.014	.	.	.	.	.	0.011	0.989	.	.	0.203	0.019	0.254
	118	0.285	.	0.965	0.804	0.990	0.872	0.500	0.011	0.047	0.004	0.532	0.951	0.732
	126	0.016	.	.	.	.	.	.	.	.	.	.	.	.
	n	354	49	82	23	48	261	278	228	225	116	113	79	82
IDH	57	.	.	.	.	.	.	.	0.007	.	.	.	.	.
	77	.	.	.	.	.	.	.	0.013	.	.	.	.	.
	91	0.051	.	.	.	.	.	.	.	.	.	.	.	.
	100	0.501	0.816	0.323	0.304	0.417	0.577	0.293	0.250	0.349	0.448	0.522	0.525	0.427
	115	0.405	0.041	0.677	0.696	0.583	0.423	0.707	0.708	0.627	0.332	0.460	0.475	0.567
	127	0.042	0.143	.	.	.	.	.	0.022	0.024	0.220	0.018	.	0.006
	n	328	50	83	23	49	242	252	230	186	88	29	37	28
LAP	93	.	.	.	.	.	.	.	0.026	0.016	.	.	.	.
	97	.	.	.	.	.	0.023	.	0.052	0.065	0.057	0.052	0.135	0.089
	100	0.950	0.820	1.000	0.978	1.000	0.969	1.000	0.920	0.680	0.750	0.879	0.784	0.911
	102	0.008	0.180	.	.	.	0.008	.	0.002	0.204	0.051	0.052	.	.
	105	0.043	.	.	0.022	.	.	.	.	0.035	0.142	0.017	0.081	.
	n	355	50	83	23	49	263	280	231	230	118	127	162	89
MDH-1	87	.	0.240	.	.	.	.	.	0.002	0.017	0.208	.	.	.
	94	.	.	0.042	0.087	0.133	0.116	0.093	.	.	.	.	.	.
	100	1.000	0.760	0.958	0.891	0.867	0.878	0.907	0.998	0.976	0.780	0.976	0.994	0.989
	109	.	.	.	0.022	.	0.006	.	.	0.007	0.013	0.024	0.006	0.011
	n	355	49	60	23	22	261	276	231	228	118	122	153	86
MDH-2	58	0.100	.	.	.	0.023	.	.	0.052	0.007	.	.	.	0.012
	83	.	.	.	.	.	.	0.004	0.009	0.002	0.017	.	.	.
	100	0.900	0.551	0.225	0.326	0.114	0.360	0.531	0.931	0.669	0.441	0.590	0.500	0.477
	104	.	.	.	.	.	0.019	0.038	.	.	.	.	.	.
	106	.	0.031	0.092	.	0.409	0.013	.	.	0.002	0.085	.	0.013	.
	108	.	.	.	.	.	0.010	0.087	.	.	.	.	0.114	.
	114	.	0.418	0.508	0.304	0.250	0.385	0.212	0.009	0.320	0.458	0.410	0.373	0.512
	122	.	.	0.175	0.370	0.205	0.213	0.129	.	.	.	.	.	.
	n	338	50	82	23	48	266	279	225	223	118	79	153	86
6-PGD	71	0.001	.	.	.	.	.	.	0.089	0.043	0.042	0.013	0.010	0.064
	88	.	.	0.128	0.130	0.188	0.126	0.301	0.013	0.121	.	0.013	0.016	0.250
	95	0.183	.	.	0.022	0.021	0.004	.	0.007	0.007	0.008	.	.	.
	100	0.534	0.970	0.872	0.783	0.750	0.761	0.622	0.853	0.744	0.941	0.778	0.837	0.593
	107	0.281	0.030	.	0.065	0.042	0.109	0.068	0.016	0.054	0.008	0.196	0.137	0.070
	110	.	.	.	.	.	.	0.009	0.022	0.031	.	.	.	0.023
	n	355	50	83	23	49	269	283	226	230	115	126	161	87
PGI-2	20	.	.	.	.	.	.	.	.	0.002	.	.	.	.
	35	.	.	.	.	.	.	.	.	0.002	0.004	.	0.003	.
	60	.	0.020	.	.	.	.	0.002	0.004	0.011	0.083	0.040	0.031	0.011
	100	1.000	0.810	0.976	1.000	1.000	0.963	0.940	0.398	0.952	0.835	0.861	0.876	0.546
	115	.	.	.	.	.	.	.	0.004	0.009	0.004	0.004	.	0.046
	125	.	.	.	.	.	0.009	0.004	0.235	0.002	.	.	.	.
	130	.	0.170	0.024	.	.	0.028	0.053	0.330	0.011	0.065	0.087	0.090	0.356
	150	.	.	.	.	.	.	.	.	0.002	.	0.004	.	0.017
	155	.	.	.	.	.	.	0.002	0.018	0.004	0.009	0.004	.	0.023
	185	.	.	.	.	.	.	.	0.011	0.004	.	.	.	.
	n	355	50	84	23	49	258	283	222	229	113	123	149	87
PGM-2	56	.	.	.	.	.	.	.	.	0.050	0.009	0.016	0.060	0.006
	84	.	.	.	.	.	.	.	.	0.037	0.018	0.008	.	.
	100	1.000	0.950	1.000	0.935	1.000	0.953	1.000	0.860	0.797	0.920	0.585	0.628	0.672
	119	.	0.050	.	.	.	0.039	.	0.140	0.098	0.053	0.386	0.312	0.322
	122	.	.	.	0.065	.	0.008	.	.	.	.	.	.	.
	129	.	.	.	.	.	.	.	.	0.015	.	0.004	.	.
	134	.	.	.	.	.	.	.	.	0.002	.	.	.	.

Table II. Continued.

locus	*allelic mobility*	Sintra	Morocco	Azores					Madeira	Canary Islands				
				Flores	Corvo	Pico	Faial	São Miguel		Tenerife	Gran Canaria	El Hierro	La Palma	La Gomera
	n	353	50	81	22	34	269	279	229	223	117	120	148	87
SkDH	43	.	.	.	.	.	.	.	.	0.007	.	.	.	.
	52	.	.	.	.	.	.	.	.	0.007	.	.	0.010	.
	59	0.003	0.160	.	.	.	.	.	0.009	0.092	0.068	0.117	0.064	0.011
	60	.	.	.	.	.	.	.	.	0.011	0.004	0.008	0.010	.
	69	0.215	.	.	.	.	.	.	.	0.018	0.009	0.021	0.003	0.057
	72	.	.	.	.	.	.	.	.	0.063	0.009	0.021	0.010	0.029
	76	0.003	.	.	.	.	.	.	0.013	0.312	0.077	0.183	0.189	0.270
	83	0.392	0.530	.	.	.	0.004	0.009	.	0.177	0.282	0.321	0.314	0.057
	86	.	.	.	.	.	.	.	.	0.004	.	.	0.003	0.006
	90	0.059	.	0.204	0.182	0.147	0.151	0.197	0.570	0.056	0.026	0.050	0.122	0.356
	96	.	.	0.068	0.250	0.132	0.191	0.050	.	0.025	.	0.008	0.017	0.029
	100	0.327	0.310	0.691	0.545	0.647	0.636	0.665	0.205	0.179	0.513	0.213	0.236	0.161
	104	.	.	.	.	.	.	.	0.070	0.002	0.004	.	0.003	.
	110	.	.	0.037	0.023	0.074	0.019	0.079	0.133	0.011	0.009	0.021	0.010	0.017
	117	.	.	.	.	.	.	.	.	0.002	.	0.038	0.007	.
	124	.	.	.	.	.	.	.	.	0.034	.	.	.	0.006
	n	355	50	84	23	49	270	283	231	230	118	126	162	89
TPI-1	81	.	.	.	.	.	0.002	.	0.004	0.011	0.008	0.056	0.019	0.202
	100	0.996	0.970	1.000	1.000	1.000	0.998	1.000	0.989	0.965	0.975	0.825	0.969	0.798
	114	0.004	0.030	.	.	.	.	.	0.006	0.024	0.017	0.040	0.012	.
	118	.	.	.	.	.	.	.	.	.	.	0.079	.	.
	n	355	50	84	23	49	260	284	231	230	118	126	162	89
TPI-2	65	.	.	.	.	.	.	.	.	.	.	0.020	.	.
	69	.	0.010	0.036	.	.	0.013	.	0.061	0.050	0.004	0.044	0.028	0.011
	87	.	.	.	.	.	.	.	.	0.002	.	.	.	.
	100	0.945	0.640	0.780	1.000	0.786	0.842	0.827	0.216	0.933	0.932	0.869	0.914	0.933
	123	0.055	0.350	0.185	.	0.214	0.144	0.173	0.723	0.015	0.064	0.067	0.059	0.056
	n	355	50	57	23	39	270	285	231	229	117	127	162	89
UGPP	79	.	.	0.009	.	.	.	.	.	0.002	0.004	.	.	.
	92	.	.	0.009	0.022	.	0.002	0.002	0.009	0.017	.	0.035	0.009	0.006
	100	1.000	1.000	0.982	0.978	0.974	0.994	0.996	0.623	0.880	0.957	0.909	0.991	0.978
	105	.	.	.	.	.	.	.	.	.	0.009	.	.	.
	108	.	.	.	.	0.026	0.004	0.002	0.368	0.100	0.030	0.055	.	0.017

Appendix IV

Formula to calculate geographical distance between two points from latitude and longitude coordinates in Excel:

=6377*ACOS(COS(RADIANS(90-(LAT1*24)))*COS(RADIANS(90-(LAT2*24)))+SIN(RADIANS(90-(LAT1*24)))*SIN(RADIANS(90-(LAT2*24)))*COS(RADIANS(24*(LONG1-LONG2))))

where 6377 is the radius of the Earth, LAT1 is the latitude of point 1, LAT2 is the latitude of point 2, LONG1 is the longitude of point 1 and LONG2 is the longitude of point 2.