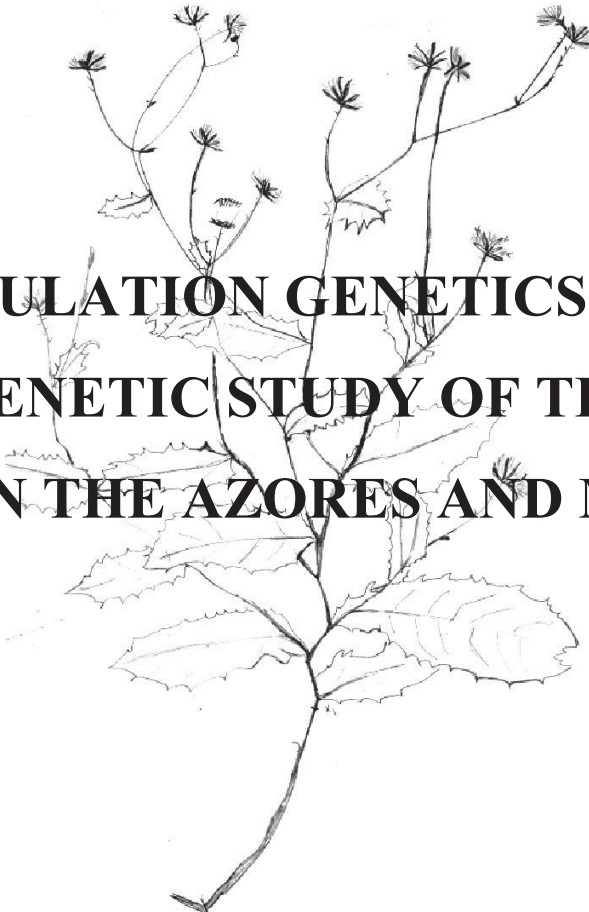


UNIVERSIDADE DOS AÇORES



DEPARTAMENTO DE BIOLOGIA

**MESTRADO EM BIODIVERSIDADE E BIOTECNOLOGIA
VEGETAL**



**POPULATION GENETICS AND
PHYLOGENETIC STUDY OF THE GENUS
TOLPIS IN THE AZORES AND MADEIRA**

ORIENTADORES

Professora Doutora MÓNICA MOURA
Doutora JULIE SARDOS

**Lurdes da Conceição Borges Silva
Ponta Delgada
Abril de 2012**

UNIVERSIDADE DOS AÇORES



DEPARTAMENTO DE BIOLOGIA

**MESTRADO EM BIODIVERSIDADE E BIOTECNOLOGIA
VEGETAL**

MASTER DISSERTATION INCLUDED IN DEMIURGO PROJECT
FINANCED BY FEDER, OF THE INICIATIVE MAC/1/C020- PROGRAMA
DE COOPERAÇÃO TRANSNACIONAL AÇORES-MADEIRA-CANÁRIAS
(MAC) 2007-2013

**União Europeia
Fundo Europeu de
Desenvolvimento Regional**

Investimos no seu futuro



ORIENTADORES

Professora Doutora MÓNICA MOURA
Doutora JULIE SARDOS

**Lurdes da Conceição Borges Silva
Ponta Delgada
Abril de 2012**

INDEX

Resumo

Acknowledgements

Abstract

Introduction	1
Material and Methods	9
1. Phylogenetic study	9
1.1. Taxon sampling	9
1.2. DNA Extraction and amplification	9
1.3. Alignment of sequence and analysis	10
2. Morphological study	11
2.1. Taxon sampling and characters measured	11
2.2. Statistical analysis	13
3. Population genetics study	14
3.1. Taxon sampling	14
3.2. Taxon extraction	18
3.3. Characterization of the microsatellites	19
3.4. Full scale genotyping	19
3.5 Data analysis	20
Results	22
1.1. <i>Tolpis</i> spp.	22
1.2. <i>Tolpis succulenta</i>	32
1.3. <i>Tolpis azorica</i>	37
1.4. Molecular phylogeny	46
1.5. Morphological variation	47
Discussion	51
Taxonomical treatment	54
Conclusion	57
Bibliographic references	58

Resumo

O género *Tolpis* é constituído por aproximadamente 12 espécies que se encontram distribuídas por África, Europa e pelas ilhas da Macaronésia. O género *Tolpis* está incluído na tribo cosmopolita Lactuceae Cass. pertencente à família Asteraceae. No arquipélago dos Açores, *T. azorica* (Nutt.) P. Silva, é considerada como endêmica e está atualmente indicada como presente em todas as ilhas do arquipélago com excepção de Santa Maria e Graciosa. *T. succulenta* (Dryander in Aiton) Lowe é indicada para todas as ilhas do arquipélago dos Açores, ocorrendo também no arquipélago da Madeira.

Para estimar a variabilidade intra e intergenética nas várias populações de *Tolpis* spp. nos Açores e na Madeira, foi utilizado um conjunto de 5 microsatélites (SSR) em 478 indivíduos de *Tolpis* spp. Para estes marcadores, *T. azorica* apresenta 59 alelos no total (média de 11,8), oscilando entre 5, para o marcador TA3B02, e 22 para o marcador TA3B05, apresentando um excesso global de homozigotia (Multilocus Fis=0,298, variando desde -0,05 para TA3B05 a 0,49 para TA2A01). Quanto a *T. succulenta*, foram obtidos 56 alelos no total (média de 11,2), variando entre 7 para o marcador TA2A02 e 18 para TA3B05, com um excesso global de homozigotia (Multilocus Fis=0.383, desde 0.06 para TA3B05 a 0.582 para TA2A01). As duas espécies têm um valor equivalente de alelos raros, nomeadamente 55.9% para *T. azorica* e 57.1% para *T. succulenta*. Após a realização de uma Análise de Componentes Principais (ACP) e análises Bayesianas confirmou-se a existência de uma estrutura populacional composta por três grupos bem definidos. Nesta estrutura *T. azorica* surge dividida em dois grupos, sendo o terceiro composto por *T. succulenta*. Uma análise mais detalhada a *T. succulenta* confirmou um agrupamento diferencial entre indivíduos pertencentes a populações dos Açores e da Madeira.

Análises morfológicas e dados moleculares obtidos a partir de sequências da região nuclear ITS foram utilizadas para esclarecer as relações filogenéticas entre as espécies de *Tolpis* endêmicas presentes nos arquipélagos dos Açores e da Madeira. Os resultados obtidos são congruentes com os obtidos através da análise da estrutura genética populacional e têm como resultado final a nomeação de duas novas espécies, que são aqui descritas pela primeira vez. Uma actualização da circunscrição geográfica das espécies de *Tolpis* endêmicas dos Açores é aqui também indicada.

Acknowledgments

This study would not have been possible without the financial support of the MAC-PCT program and FEDER. I'm grateful to Secretaria Regional do Ambiente and Direção dos Serviços Florestais for their support in the field.

I'm grateful to my advisers a special and most grateful thank you to everyone who shared their knowledge and provided me with priceless help. Dr. Miguel Sequeira for all the help with the morphological characters and collecting the specimens from Madeira. Dr. Luis Silva for the morphological statistical analyses. Prof. Dr. Frias Martins for all his patience in translating the Latin descriptions. Msc. Elisabete Dias for all her help and for being there during hard times, and finally, the lab technician Natália Cabral for her friendship.

I also want to dedicate this accomplishment to my deceased grandfather António Borges, I wouldn't be who I am and be able to pursue my dreams without his life example and his personal love of botany. For always supporting and being there a very special thank you to my boyfriend Lisardo, to my sister Paula that always has been my safe harbour, also to my mother Luísa and my grandmother Lurdes for all their support and worries. Finally to my nephews and the rest of the family.

Abstract

Tolpis consists of approximately 12 species distributed in Africa, Europe, and the islands of Macaronesia. The genus *Tolpis* is included in the cosmopolitan tribe Lactuceae Cass., within the Asteraceae family. In the Azores Islands, *T. azorica* (Nutt.) P. Silva, is an endemic specie to the archipelago and is currently listed as present in all the islands except Santa Maria and Graciosa. *T. succulenta* (Dryander in Aiton) Lowe is said to occur in all the Azores Islands and is shared with the Madeira archipelago.

To evaluate intra and intergenetic variability in several populations of Azorean and Madeiran *Tolpis* spp., a set of 5 microsatellite loci (SSR) were applied to 478 individuals of *Tolpis* spp. As main results, *T. azorica* exhibits 59 alleles in total (average of 11.8), ranging from 5 for marker TA3B02 to 22 for TA3B05, and an overall excess of homozygotes (Multilocus $F_{is}=0.298$, ranging from -0.05 for TA3B05 to 0.49 for TA2A01), while *T. succulenta*, exhibits 56 alleles in total (average of 11.2), ranging from 7 for marker TA2A02 to 18 for TA3B05, and an overall excess of homozygotes (Multilocus $F_{is}=0.383$, ranging from 0.06 for TA3B05 to 0.582 for TA2A01). The two species have an equivalent value of rare alleles, 55.9% for *T. azorica* and 57.1% for *T. succulenta*. A Principal Component Analysis (PCoA) and a Bayesian approach confirmed the existence of a population genetic structure with three well-confirmed groups, where *T. azorica* is divided in two groups, with the third one aggregating the *T. succulenta* individuals. A further detailed analysis to *T. succulenta* confirmed the occurrence of a differential grouping between individuals from Azores and Madeira populations.

Morphological analysis and molecular sequence data from the ITS nuclear region was used for elucidating the phylogenetic relationships of the endemic *Tolpis* spp. in the archipelagos of Azores and Madeira. The results are congruente with the population genetic structure analyses obtained, and as main results two new species are named and described here for the first time. An update of the geographic circumscription of the Azorean endemic *Tolpis* spp. is also indicated here.

Introduction

Island endemics represent some of the most unusual and rare plants in the world (Carlquist, 1974) and have long fascinated biologists because of their different morphologies and interest as systems for the study of plant evolution (Stuessy and Ono, 1998; Emerson, 2002).

The endemic floras of oceanic islands provide some of the clearest and most well known examples of adaptive radiation (Givnish & Systsma, 1997). High rates of endemism are due to biocontainment of such islands, which are usually colonized by only a very small number of continental species that evolve, leading frequently to speciation. This is common in herbaceous settlers which often radiate after colonizing new types of habitats (Carlquist 1974, Givnish 1998).

In this study we focus on the genetic variability of several *Tolpis* species which are part of the endemic flora of the Azores and Madeira archipelagos.

The Azores and Madeira archipelagos belong to the Macaronesian Region, composed by a set of North Atlantic islands with biological affinities, due to the colonization process. In addition to the Azores and Madeira, it includes the archipelagos of Cape Verde, Canary Islands and the Salvage Islands. There are many endemic species in the Macaronesian archipelagos, with a greater number of endemics in the Canary Islands (Moore et al., 2002).

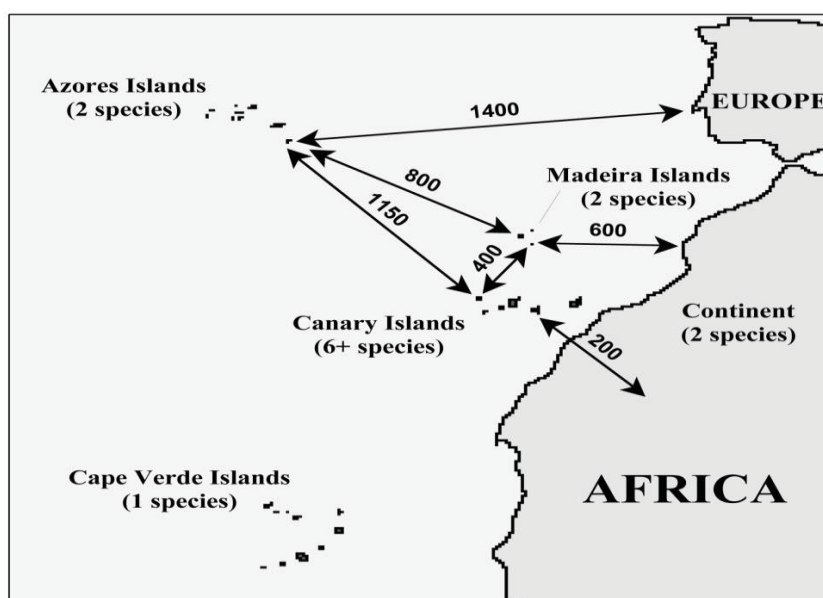


Fig. 1. Distribution of *Tolpis* across Macaronesia and the adjacent continents of Africa and Europe, along with distances used in computing minimum distance (D) scores. Adapted from Morre et al, 2002

The Azores are located in the North Atlantic Ocean, within a range bounded by the parallel 36 ° 55 '43" and 39 ° 43' 02" N and the meridians 24 ° 46 '15" and 31 ° 16' 02" W. The position it occupies manifests itself in a strong geographic isolation, since it is roughly about 1430 km from the European continent and more than 3900 km from North America. The archipelago, consisting of nine islands, is divided into three distinct groups (Western, Central and Eastern) with several islets and has a direction WNW-ESSE, due to regional tectonic phenomena. The maximum spacing between the islands exceeds 340 nautical miles (630 km), and corresponds to the distance between Corvo and Santa Maria (Azevedo, 1996).

The surface of the islands (2334 km²) is about 2.6% of the country (88,797 km²). However, the islands show very unequal size: the largest, São Miguel (745.8 km²), Pico (448.4 km²) and Terceira (403.4 km²), representing 70% of the total area, São Jorge (245, 9 km²), Faial (173.8 km²) and Flores (141.6 km²) have an intermediate size, Santa Maria (97.1 km²), Graciosa (61.2 km²) and Corvo (17.2 km²) are the smallest. Given the criteria of UNESCO, which defines "small islands" with an area less than 1000 km², all Azores islands are included in this classification (Azevedo, 1996).

The maximum elevation of the islands is highly variable, ranging from 402 m in Graciosa and 2351 m at Pico Mountain, the highest point of Portugal. Pico Island is the most eccentric in terms altimetric with 16% of the area above 800m.

The Madeira archipelago is located in the Atlantic Ocean southwest of the Iberian Peninsula, roughly between latitudes 30 °01'N and 33 °31'N and longitudes 15°17'30"W 51'W of Greenwich. The archipelago has a total area of approximately 796.8 km², and is formed by the island of Madeira with 736 km², a length of about 58 km in the direction E – W and a width of 23 km in direction N - S (Ribeiro, 1949), and the island of Porto Santo with 42.26 km², situated in the extreme NE of the archipelago, and, therefore, closer to the European and African continents. Near Madeira two groups of uninhabited islands occur, the Desertas and the Salvage Islands, which are nature reserves in the archipelago. The first covers an area of 14.2 km² and include the islands of the Deserta Grande, Bugio and Chão. The latter with an area of 3.6 km², comprising the islands of Selvagem Grande, Selvagem Pequena and the islet of Fora. Together, these constitute an archipelago individualized, reaching 250 km SSE of the eastern extremity of Madeira.

Tolpis is characterized by a suite of morphological characters, including usually annual or perennial herbaceous habit, mostly rosulate leaves, capitula arranged in

corymbs or panicles yellow florets, styles short, cypselas ribbed and glabrous, and a pappus of few to many long scabrid bristles interspersed with minute scales (Park et al. 2001). The genus belongs to the Asteraceae, which is one of the largest botanical families, consisting of approximately 110 genera and 2500 species. There have been several systematic approaches on the Asteraceae family (Cronquist, 1955; Poljakov, 1967; Carlquist, 1976; Wagenitz, 1976; Jeffrey, 1978; Robinson, 1981, 1983; Thorne, 1983; Bremer, 1987) and most of these studies acknowledged that there are two subfamilies, Lactucoideae and Asteroideae, with *Tolpis* nested in the sub-family Lactucoideae.

The genus *Tolpis* is included in the cosmopolitan tribe Lactuceae Cass. It is currently considered a tribe within the Asteraceae family comprising 98 genera and over 1550 species (Bremer, 1994). A phylogenetic study conducted in the Asteraceae based on morphological (Karis et al., 1992) and molecular (Jansen et al., 1991, Him et al., 1992) data indicated that Lactuceae is monophyletic, and their closest relatives Liabeae Rydb Cass. and Vernoniae Cass.

The classification of infratribal Lactuceae has been the focus of several morphological and molecular studies to find the right place of *Tolpis* (Stebbins 1953; Jeffrey, 1966; Tomb, 1975, 1977; Baagoe, 1980; Blackmore, 1981; Whitton, 1981; Bremer, 1994). Stebbins (1953) placed *Tolpis* in subtribe Cichoriinae with seven other genera and suggested that *Tolpis* was most closely allied to *Arnoseris* and *Hispidella* (Park et al., 2001). Jeffrey (1966) placed *Tolpis* in an informal *Tolpis* subgroup and allied it to *Arnoseris* based on several morphological characters. Pollen investigations of Blackmore (1981) suggested that the small pollen grains with double rows of spines on the equatorial ridges support a close relationship between *Arnoseris*, *Hispidella* and *Tolpis*.

Bremer's (1994) subtribal treatment, which was based on a morphological cladistic analysis of 22 genera representing the major monophyletic groups of Lactuceae, classified *Tolpis* in the Hieraciinae with six other genera. Jarvis (1980) suggested that morphological similarities of the pappus and cypselas and chromosome number suggest a close relationship between *Tolpis* and *Hieracium*.

Based on studies of Park *et al.* (2001), their analysis of the Lactuceae using the chloroplast gene *ndhF* found *Tolpis* to be monophyletic and nested within the Lactuceae as a sister group to a large clade containing such genera as *Lactuca* L., *Crepis* L., *Reichardia* Roth, *Dendroseris* and *Hypochoeridinae*. However the morphological

cladogram did not resolve satisfactorily the intergeneric relationships in the Lactuceae (Park et al. (2001).

Tolpis consists of approximately 12 species distributed in Africa, Europe, and the islands of Macaronesia (Jarvis, 1980; Park et al., 2001; Moore et al., 2002). Eleven species are Macaronesian endemics, with at least one species native to each of the four major island group (Jarvis, 1980).

In the Azores Islands, *T. azorica* (Nutt.) P. Silva, has been listed as the only species of the genus endemic to the archipelago (Silva et al., 2011), and considered present in all the islands except Santa Maria and Graciosa. *T. succulenta* (Dryander in Aiton) Lowe, said to occur in all the Azorean islands, is shared with the Madeira archipelago. Another *Tolpis* endemic occurring in the Madeira archipelago is *T. macrorhiza* Lowe.

T. azorica grows in the mountainous, very humid, inland of the islands while *T. succulenta* in the Azores and Madeira archipelagos is found on the dryer coastal cliffs, rocks and rocky banks up to 1000 m, in Madeira also occasionally up to 1500 mm in open habitats. Despite this ecological cleavage in their habitats, the presence in some islands of both of these endemics raises the question of putative gene flows between the two species.

An extensive synonymy revision of *T. azorica* and *T. succulenta* was conducted by Lack in 1981 (see “Taxonomic treatment” in Discussion). The holotype of *T. azorica* is a specimen from Pico collected by the Hochstetter (father and son) in 1838, and held at TUB herbarium. All other types of different synonymy are from the central or western groups of islands. The holotype of *T. succulenta* can be found at BM and was cultivated at Kew from material obtained at the Madeira Island. All synonymy types are equally originated from Madeira.

Lack (1981), observed that unusual plants of *T. azorica* have been collected in São Miguel and Santa Maria islands with long hairs at the base of the achenes and somehow odd distribution of the leaves along the stem, indicating that the systematic position of these conspicuous variants needed further analyses. Furthermore, recent molecular studies showed that there is some variability between individuals of *T. succulenta* from Madeira and the Azores (Moore et al., 2002). Other molecular studies conducted recently in the Azores with different species have been unravelling the occurrence of intra-archipelago patterns of diversity in the endemic flora that were

previously unaccounted for, producing evidence that the current taxonomic groups may not be accurately defined (Carine & Schafer, 2010, Moura *et al.*, 2010a).

Morphological and molecular data analysis has been helpful to understand the phylogenetic relationships among plants. Nuclear and chloroplastial genomic DNA sequences have been widely used to achieve this goal in plants. The selection of the marker depends of the purpose and objectives of the study and the relationship level among species (Mayer & Soltis, 1999).

In this investigation, we used the information contained in the nuclear region of the internal transcribed spacer ITS (ITS1-5.8S-ITS2), to propose a molecular phylogenetic hypothesis for the position of *Tolpis* spp. in the archipelagos of Azores and Madeira. The ITS region has often been used as a tool to study genetic diversity (Schmichl *et al.*, 2010) and to develop phylogenetic studies for comparisons by genus and other higher taxonomical categories (Baldwin *et al.*, 1995; Schmichl *et al.*, 2010). This application has been extended with success to Macaronesian (Vargas *et al.*, 1999; Valcárcel *et al.*, 2003) and Azorean endemic taxa (Moura, 2005; Carine & Schaefer, 2010; Moura *et al.*, 2010a,b). ITS sequences have excellent advantages for phylogenetic studies, namely their universality, biparental inheritance, simplicity, intragenomic uniformity, intergenomic variability and low functional constraint (Mort *et. al.*, 2007).

Although the use of ITS sequence data is very common in phylogenetic studies, characteristics of nrDNA evolution may result in comparison of paralogs if sufficient caution is not exercised (Alvarez and Wendel, 2003). Bailey *et al.* (2003) investigated the phylogenetic implications of paralogy resulting from the tandem existence of both nrDNA pseudogenes and functional copies within an individual. While the presence of pseudogenes is not necessarily confounding to phylogenetic analysis if all copies have been sampled from each individual, rates of evolution among functional and nonfunctional paralogs may be sufficiently different to result in long-branch artifacts (Bailey *et al.*, 2003).

Alvarez and Wendel (2003) discussed other problems: the effect of secondary structure on base substitution resulting in nonindependence of characters, difficulties with alignment, problems with contamination resulting from universal primers, generally high levels of homoplasy due to rapid evolution, and difficulties with amplification that arise from secondary structure and the existence of multiple rDNA arrays. Although several of these concerns are common to most rapidly evolving regions, they did emphasize the need for caution in molecular phylogenetic studies.

In order to deal with the above mentioned problems related to ITS, and further support our phylogenetic analysis of *Tolpis*, morphological data analysis was added to the study. Morphological characters were gathered from literature and personal observations, subject to a preliminary screening for adequacy and then measured in several herbarium specimens of *T. azorica*, *T. succulenta* (from Azores and Madeira) and *T. macrorhiza*.

The second part of our study includes the evaluation of intra and intergenetic variability in several populations of Azorean and Madeiran *Tolpis* using microsatellites loci (SSR). In a recent review, Bussell et al. (2005) stated that SSR markers might be useful for phylogenetic analyses involving species that are closely related and represent recent radiations. The set of microsatellites markers obtained in this study were used to elucidate the relationship between the populations within and between species at the scale of the archipelago and to compare the species micro-evolution within the genus.

In previous population studies, ISSR markers proved to be useful in *Tolpis* for grouping individuals into populations and species (Archibald, 2006). According to Jarvis (1980), the lack of clustering showed by members of *Tolpis* may be the result of the presence of a suite of ancestral polymorphic markers segregating within populations referable to more than one species in the current taxonomy of the genus. Gene flow between populations could also serve to slow or prevent the sorting of particular loci or combinations of loci within populations. These processes would result in more ISSR loci segregating within populations than there are markers or arrays of markers that define populations or groups of populations (Archibald, 2006).

In a study by Crawford et al. (2006), the allozyme variation among species of *Tolpis* in the Canary Islands was generally similar to other insular endemics in having low genetic diversity within species, a relatively high proportion of the diversity among populations, and low divergence among species. *Tolpis* differs from some other endemics in the lower genetic diversity within and higher proportion of diversity among populations of self-compatible species compared to self-incompatible species.

As the genus *Tolpis* comprises recent colonization/speciation events within the genus, the transferability of the SSR developed from the Azorean species was tested on *T. macrorhiza* and the homonymous but distinct *T. succulenta* (Moore et al. 2002) that are both endemic to Madeira Islands.

Compared with other classes of markers, microsatellites or Simple Sequence Repeats (SSRs), are highly polymorphic due to its repetitive nature and have been

widely used to study the effect of genetic drift and gene which determine population genetic diversity (Wright and Bentzen, 1994; Collevatti et al., 1999; Daynandan et al., 1997).

These markers have a number of advantages, since they are abundant and extensively cover the genome, have multiallelic nature, require small amounts of DNA for analysis, are easy to detect by PCR (polymerase chain reaction), have Mendelian inheritance type and are expressed as co-dominant alleles (Schlotterer & Pemberton, 1994; Ellegren, 2004).

The most efficient measure to assess population structure is based on Wright's *F*-statistics (1951), Wright's inbreeding coefficient (*FST*, also called θ) being particularly useful for analysing microsatellite markers because it is able to discriminate between alleles, especially that rare ones, although *FST* produced using such markers can sometimes be overestimates of the true value.

Microsatellite markers include loci with a large number of alleles, but one question that should be asked is whether a large number of loci or a large number of alleles is more important in genetic assessment. Working on the relationship between the allele number and the coefficient of variation of four genetic distances, Kalinowski (2002) used simulated data to show that highly polymorphic loci provided better estimates of genetic distance than less polymorphic loci and that increased allele number was associated with a decrease in the coefficient of variation of each of the four genetic distances studied. These results show that there is no requirement to examine either highly polymorphic loci or large numbers of loci, the only requirement being that a sufficient number of alleles are examined.

However, the high mutation rate of microsatellites can also invalidate many assumptions used in some conventional population structure analysis because different populations may share homoplasic alleles at frequencies that depend on both the rate and the details of the mutation process (Estoup *et al.*, 2002). When such effects are ignored the rate of gene flow or genetic introgression can be overestimated (Balloux *et al.* 2000). Slatkin (1995) developed the *RST* statistic (analogous to *FST*) to take into account the effects of mutation, but although *RST* performs better than *FST* in some circumstances it can also be sensitive to details of the mutation process (Balloux and Goudet, 2002). Since mutation rate varies widely between loci within species (Di Rienzo *et al.* 1998) one advantage of loci with a high mutation rate is that genetic

differentiation reaches equilibrium faster, offering the possibility of obtaining estimates from larger and more widely spaced populations.

The objectives of our study were: (i) increase the sampling of *Tolpis* spp. populations in order to complete the already existing collection at AZB; (ii) to understand the intra and interpopulational genetic variability existing in *Tolpis* spp. populations and their relationships, with an emphasis in S. Miguel Island; (iii) identify cases of populations with very low genetic variability; (iv) determine the correspondence between the currently accepted distribution of *Tolpis* spp. in the Azores and the main groups obtained; (v) confirm the occurrence of other taxonomic groups besides those currently listed.

This research was part of the Demiurge Project, financed by FEDER, the initiative MAC/1/C020-Transnational Cooperation Programme Madeira-Açores-Canarias (MAC) 2007-2013, which has as one of its aims to collect data from genetic populations of several species of the endemic flora of Macaronesia.