UNIVERSIDADE DOS AÇORES

DEPARTAMENTO DE BIOLOGIA

MESTRADO EM BIODIVERSIDADE E BIOTECNOLOGIA VEGETAL

Estudo morfológico e filogenético das subespécies *Daucus carota* ssp. *azoricus* e *Daucus carota* ssp. *maritimus* na ilha de S. Miguel

Amélia Maria Matias Vaz

Orientadora: Doutora Mónica Maria Tavares de Moura

Co-orientadores: Doutor Luís Filipe Dias e Silva

Doutor Hanno Schaefer

Ponta Delgada, Outubro 2014

TABLE OF CONTENTS

TABLE OF FIGURES

ABBREVIATIONS

Arg. – Argentina

BI – Bayesian Inference

COS – Conserved Ortholog Set

DA – Discriminant Analysis

ITS – Internal Transcribed Spacer

Leb. – Lebanon

ML – Maximum Likelihood

MP – Maximum Parsimony

PCA – Principal Components Analysis

AGRADECIMENTOS

Gostaria de agradecer à Profª. Doutora Mónica Moura, ao Prof. Doutor Luís Silva e ao Prof. Doutor Hanno Schaefer, meus orientadores. Antes de mais, por terem aceitado sê-lo, e depois por todo o apoio dado no decurso desta tese, por terem tido sempre tempo para esclarecer as minhas dúvidas e por não se terem exasperado pelas minhas demoras. Aos meus colegas do MBBV, obrigada por terem tornado o percurso mais divertido e desafiante. Não teria sido o mesmo sem todos vocês! Aos docentes do MBBV, o meu profundo agradecimento por tudo o que me ensinaram. Aos membros da equipa do CIBIO com quem convivi mais de perto, a quem "chateei" (sem aspas, também) sempre que tive alguma dúvida, ou precisei de alguma coisa, o meu muito obrigada por todo o apoio.

Gostaria de agradecer também a todos os que me apoiaram, ajudaram e motivaram ao longo de toda esta odisseia. A lista é demasiado grande para mencionar e, portanto, e para garantir que não deixo ninguém de fora, limito-me a chamar-vos meus amigos. Obrigada a todos, por tudo!

Deixo um agradecimento final à minha família que, embora ainda hoje, ao fim de mais anos do que quero contar, não entende muito bem este meu apego a "coisas verdes e viscosas", não deixou nunca de me apoiar no meu amor por elas.

RESUMO

A família Apiaceae está bem representada nos Açores, com 13 géneros e 18 espécies (6 das quais endémicas). Entre estas, encontram-se duas subespécies de *Daucus*: *D. carota* L. ssp. *azoricus* Franco, uma subespécie endémica descrita em 1971 e *D. carota* L. ssp. *maritimus* (Lam.) Batt.. *Daucus carota*, a única espécie cultivada do género, existe tanto na forma domesticada como na forma selvagem e a reprodução cruzada entre estas duas é muito comum. As subespécies descritas são fenotipicamente diversas e não há consenso acerca do número total de subespécies, caracteres diagnosticantes ou distribuição.

O objectivo deste estudo foi caracterizar morfológica e geneticamente as populações de *D. carota* L. existentes na ilha de São Miguel (Açores). Foram recolhidos espécimes de dez populações espalhadas por toda a ilha. A análise estatística dos dados recolhidos na caracterização morfológica não foi capaz de elucidar as relações entre os diferentes locais de amostragem. A sobreposição dos caracteres morfológicos e a grande diversidade fenotípica observada dentro de cada um dos locais amostrados parecem indicar que não existem diferenças morfológicas marcantes entre todos eles.

Sequenciámos uma região nuclear COS (conserved ortholog set) e a região nuclear ribossomal ITS (internal transcribed spacer) para averiguar a presença de diferenças genéticas. As sequências ITS revelaram pouca variação entre si, com apenas dois haplótipos diferentes nas 20 sequências analisadas. As sequências COS mostraram-se mais variáveis, com cinco haplótipos diferentes entre as sete sequências produzidas. As análises de Máxima Parcimónia, Máxima Verosimilhança e Bayesiana dos dois alinhamentos incluindo outras acessões de *Daucus* retiradas do Genbank indicou uma relação próxima entre as nossas amostras e as demais subespécies de *D. carota*, embora uma inferência conclusiva acerca da classificação infraespecífica de *D. carota* em geral, e dos nossos espécimes de São Miguel em particular, continue a ser problemática.

Palavras-chave: endémico, morfometria, filogenia, ITS, COS

ABSTRACT

The Apiaceae family is well represented in the Azores, with 13 genera and 18 species (6 of them endemic). Among them are two subspecies of *Daucus*: *D. carota* L. ssp. *azoricus* Franco, an endemic subspecies described in 1971 and *D. carota* L. ssp. *maritimus* (Lam.) Batt. *Daucus carota*, the only cultivated species of the genus, exists in both domesticated and wild forms and interbreeding between these forms is very common. The described subspecies are phenotypically diverse and there is little consensus on their number, diagnostic characters and distribution.

The purpose of this study was to characterize morphologically and genetically the populations of *D. carota* occurring in São Miguel island (Azores). Specimens were collected from ten populations, distributed throughout the island. The statistical analysis of the data collected from the morphometric characterization did not help to resolve the relationships between the different sampling sites. The overlapping character states and the great phenotypic diversity we found within each of the sampled areas seem to indicate that there are no major morphological differences between all the sites.

We sequenced one nuclear conserved ortholog region (COS) and the nuclear ribosomal DNA internal transcribed spacer region (ITS) to test for genetic differences. Our ITS sequences revealed little variation among them, with only two different haplotypes emerging from the 20 sequences produced. The COS sequences obtained were more variable, with five haplotypes in the seven sequences analyzed. Maximum Likelihood, Maximum Parsimony and Bayesian analyses of both alignments including Genbank sequences of other *Daucus* accessions indicated a close relationship between our samples and other *D. carota* subspecies, although a conclusive inference of the infraspecific classification of *D. carota* in general, and of our S. Miguel specimens in particular, remains problematic.

Keywords: endemic, morphometrics, phylogenetics, ITS, COS

INTRODUCTION

The Apiaceae Lindl. family, commonly known as the carrot family, comprises ca. 434 genera and 3780 species (Stevens, 2001 onwards). Most members are easily identifiable by distinctive characters such as pinnately divided leaves with sheathing bases, small unspecialized flowers in compound umbels, and specialized dry schizocarp fruits (Hickey & King, 1997).

Despite being cosmopolitan, the majority of its diversity is concentrated in the northern hemisphere, with carrot (*Daucus carota* L. ssp. *sativus*) being its most well-known member, alongside some commonly used herbs and spices such as anise, coriander, cumin, dill or parsley (Spooner *et al.*, 2014).

This family is well represented in the Azores Archipelago, with several Azorean endemic species in its ranks: *Ammi trifoliatum* (H. C. Watson) Trel., *Ammi seubertianum* (H. C. Watson) Trel., *Angelica lignescens* Reduron & Danton, *Chaerophyllum azoricum* Trel., *Sanicula azorica* Guthn. ex Seub. and *Daucus carota* L. ssp. *azoricus* Franco (Silva *et al.*, 2010), the latter an endemic subspecies first described in 1971 (Franco, 1971). Another *D. carota* subspecies, *Daucus carota* L. ssp. *maritimus* (Lam.) Batt. is also reported as occurring in the Azores (Silva *et al.*, 2010).

The genus *Daucus* is most common in the Mediterranean region, although some species occur in other continents and in the southern hemisphere (Spooner *et al.*, 2013). *D. carota* L. ssp. *sativus*, the only cultivated species of this genus, exists in both domesticated and wild forms. More than 60 species have been proposed for variants within the "*D. carota* complex" alone, for which there are no or only poorly developed barriers to interbreeding between the wild and domesticated forms (Pujadas Salvà, 2002). The described subspecies are phenotypically diverse (see Spooner *et al.*, 2013 and Arbitzu *et al.* 2014b) and there is little consensus on the number of subspecies. Heywood (1968) describes 11 subspecies (*carota*, *maritimus*, *major*, *maximus*, *gummifer*, *commutatus*, *hispanicus*, *hispidus*, *gadecaei*, *drepanensis* and *rupestris*), Franco (1971) recognizes five (*hispidus*, *maritimus*, *sativus*, *maximus* and *azoricus*), Sáenz Laín

(1981) also five (*maximus, maritimus, gummifer, carota* and *hispanicus*), and Pujadas Salvà (2002; 2003) describing ten subspecies in 2002 (*carota, maximus, sativus, cantabricus, hispanicus, gummifer, majoricus, fontanesii, commutatus* and *halophilus*) , and nine in the following year (*carota, maximus, sativus, cantabricus, hispanicus, gummifer, majoricus, commutatus, halophilus*), just for the Iberian Peninsula.

Of the two subspecies believed to occur in the Azores Archipelago, only *D. carota* L. ssp. *maritimus* (Lam.) Batt. is acknowledged and discussed in the more recent studies (Sáenz Laín, 1981; Pujadas Salvà, 2002). The reported Azorean endemic subspecies *D. carota* L. ssp. *azoricus* Franco wasn't taken into consideration in these works.

Furthermore, the status of *D. carota* L. ssp. *maritimus* (Lam.) Batt. doesn't gather consensus, for while Franco (1971) and Sáenz Laín (1981) consider it a subspecies, Pujadas Salvà (2002) more recently, considered it, and its reported occurrence in Portugal, to be most likely the result of misidentification of late-season specimens from the typical subspecies and variety (*D. carota* L. ssp. *carota* var. *carota*).

It is a well-known fact that the genus *Daucus* is in need of extensive modern taxonomic and monographic studies, as shown by the increase in published articles on the subject in recent years and the fact that, despite this, a definitive taxonomy has not yet been presented. The most recent monograph concerning the taxonomy of *Daucus* and using a morphoanatomical classification was published in 1981, by Sáenz Laín and, although it is still one the most widely used references for taxonomic studies of this genus, as several authors have pointed out (Spooner *et al.*, 2013; Arbizu *et al.*, 2014b), it is based solely in herbarium specimens observations, citing few specimens, and lacking complete synonymies, distribution maps and phylogenetic data.

Practical identifications have long been based on local floristic treatments, relying on Floras and other region-specific publications such as those from Algeria (Quezel & Santa, 1963), Europe (Heywood, 1968), the Iberian Peninsula and Balearic Islands (Pujadas Salvà, 2003) or Portugal (Franco, 1971).

This presents a serious problem of congruence, since more often than not, the descriptions and taxonomic keys in these publications use different characters to identify the same species and subspecies, present different and often overlapping sets of character states for the same species, have incomplete synonymies that make it very hard to compare the taxonomic concepts presented and often have little information about distribution and geographic ranges of the addressed species (Arbizu *et al.*, 2014b).

Identifications are often problematic because there is no accepted standard to quantify and describe the vast range of variation found in *Daucus*, nor has anyone made a single compilation of type specimens complete with all the plant parts necessary for an unambiguous and definitive identification (Arbizu *et al.*, 2014b; Spooner *et al.*, 2014).

Both Arbizu *et al.* (2014b) and Spooner *et al.* (2014) report the morphological character states traditionally used as species identifiers, as being overlapping. Most of the studied characters showed tremendous variation within some *taxa* and overlapping ranges across *taxa*, demonstrating that the majority of *Daucus* species are distinguished by size and meristic variation, not the possession of unique traits.

The analysis of representative *Daucus carota* subspecies accessions conducted by Spooner *et al.* (2014) emphasized the great morphological similarity among those *taxa* and suggests that for wild *D. carota* specimens, only two subgroups may be separated morphologically: *D. carota* ssp. *gummifer* (plants with a relatively short stature, thick, broad leaf segments, usually flat or convex fruiting umbels, found coastal regions) and *D. carota* ssp. *carota sensu lato* (taller plants with thinner narrower leaf segments, fruiting umbels usually curved upward and closed in a characteristic "bird's nest" form, found in coastal regions, but also inland), corresponding to the two species or the two "species aggregates" previously suggested by other authors (see Spooner *et al.*, 2014).

Also noteworthy is the study by Tavares *et al.* (2014), were DNA content analysis, morphological features of the fruits and chemical characterization of essential oils were used in an attempt to establish a clearer distinction between the four subspecies of *D. carota* native to Portugal.

In the last 25 years or so, there have been quite a few molecular studies focused in examining and elucidating the phylogenetic relationships within Apiaceae (Downie & Katz-Downie, 1996; Downie *et al.*, 1998, 2000a, 2000b, 2008; Lee & Downie, 1999, 2000; Zhou *et al.*, 2008) and, in particular, within *Daucus* (Vivek *et al.*, 1998, 1999; Martinez-Flores *et al.*, 2011; Iorizzo *et al.*, 2013; Spooner *et al.*, 2013; Lee & Park, 2014; Rong *et al.*, 2014; Arbizu *et al.*, 2014a).

Several molecular approaches have been used by the different authors: Vivek (1998, 1999) used restriction fragment length polymorphisms (RFLPs) of nuclear and plastid DNA; Downie & Katz-Downie (1996), Downie *et al.* (1998, 2000a, 2000b, 2008), Lee & Downie (1999, 2000), Zhou *et al.* (2008) and Lee & Park (2014) used a combination of nuclear ribosomal DNA internal transcribed spacer (ITS) sequences, plastid rpoC1 and rpl16 intron sequences, plastid matK coding sequences and plastid DNA restriction sites. More recently, single nucleotide polymorphisms (SNPs) were used by Iorizzo *et al.* (2013) and Rong *et al.* (2014), and conserved ortholog set (COS) markers where used, first by Spooner *et al.* (2013) and subsequently by Arbizu *et al.* (2014a), the latter using next-generation sequencing technology on their analysis.

Spooner *et al.* (2013) demonstrated the usefulness of COS markers to present well-resolved phylogenies in the *Daucus* clade. Their analysis placed two species from other genera within *Daucus*, supporting previous plastid and ribosomal DNA evidence pointing to the paraphyly of *Daucus* and the basic topological groups found by them were also congruent with previously published works. Another significant result was the close relationship of the infraspecific *taxa* of *Daucus carota* and two other species (*D. capillifolius* and *D. sahariensis*).

Arbizu *et al.* (2014a) built upon the study of Spooner *et al.* (2013) and attempted to use next-generation sequencing technology (phylogenomics) to resolve the phylogenetic relationships of *Daucus*. This study used multiple accessions per species and, much like its predecessor, the results obtained were concordant with previous studies, and it wasn't possible to clearly distinguish the subspecies of *D. carota*. Although they managed to group wild *Daucus carota* accessions collected in Tunisia and Libya and in Portugal and Spain together, a result that partially matches that of Iorizzo *et al.* (2013), who grouped *D. carota* ssp. *carota* and *D. capillifolius* from northern Africa, separate from *D. carota* from Europe, they were unable to separate *D. carota* ssp. *carota* from ssp. *gummifer*, a separation that was found by Iorizzo *et al.* (2013).

The molecular studies of Spooner *et al.* (2013), Iorizzo *et al.* (2013) and Arbizu *et al.* (2014a) and the morphological studies of Spooner *et al.* (2014) demonstrate the difficulty of producing a clearly defined infraspecific classification of *D. carota*, suggesting either a relatively recent divergence of populations of *D. carota* (Spooner *et al.*, 2013) or the possibility that there may be less intraspecific taxa than are currently accepted (Spooner *et al.*, 2014).

Although many of the currently accepted *Daucus* species (such as *D. aureus*, *D. capillifolius*, *D. crinitus*, *D. glochidiatus*, *D. involucratus*, *D. littoralis*, *D. muricatus* or *D. pusillus*) are indeed supported by molecular data and can be separated by morphology alone, for some species and, particularly, subspecies, the taxonomy of *Daucus* remains unclear, be it by lack of sufficient germplasm for definitive morphological and molecular analyses, of comprehensive herbarium studies to associate names to type specimens, or due to unsettled generic affiliations and undefined species boundaries (Arbizu *et al.*, 2014b).