SYSTEMATICS AND EVOLUTION IN THE TRIBE SCHIZOPETALAE (BRASSICACEAE): A MOLECULAR, MORPHOLOGICAL, AND ECOLOGICAL ANALYSIS OF THE DIVERSIFICATION OF AN ENDEMIC LINEAGE FROM THE ATACAMA DESERT (CHILE)

By
©2013
Oscar Fernando Toro Núñez

Submitted to the graduate degree program in Ecology and Evolutionary Biology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____
Chairperson Mark E Mort

_____
Daniel J. Crawford

_____
Craig C Freeman

_____
Jorge Soberón

_____
Rafe M. Brown

_____
Matthew J. Buechner

Date Defended: August 26, 2013

The Dissertation Committee for Oscar Fernando Toro Núñez
certifies that this is the approved version of the following dissertation:

SYSTEMATICS AND EVOLUTION IN THE TRIBE SCHIZOPETALAE (BRASSICACEAE): A MOLECULAR, MORPHOLOGICAL, AND ECOLOGICAL ANALYSIS OF THE DIVERSIFICATION OF AN ENDEMIC LINEAGE FROM THE ATACAMA DESERT (CHILE)

_____
Chairperson Mark E. Mort

Date approved: December 19, 2013

Abstract

As aridity has been identified as an active promoter of diversification in deserts, attempts to test organismal differentiation in the Atacama Desert have resulted particularly challenging. Most limitations are related to the recent origin of the extreme aridity in the Atacama Desert, which have stimulated a rapid process of diversification and obscured evidence of interspecific divergence. Based on its favorable biological attributes and high endemicity, genera from the tribe Schizopetalae (*Mathewsia* and *Schizopetalon*) emerge as a practical study group to conduct studies of diversification under rapid and recent diversification. The present dissertation focuses on exploring this issue, 1) solving the phylogenetic relationships in the tribe Schizopetalae, 2) describing patterns of interspecific divergence in a well-defined lineage of *Schizopetalon* from the Atacama Desert, and 3) searching and testing multiple highly variable nuclear loci for phylogenetic and phylogeographic purposes. The results confirmed the monophyletic status of the tribe Schizopetalae and genus *Schizopetalon*; nevertheless, genus *Mathewsia* requires to be redefined because the exclusion of *M. nivea*. Patterns of interspecific differentiation suggest a process of allopatric divergence promoted by ecological niche differentiation between the Andes and coastal ranges in the Atacama Desert. While this result is consistent with previous hypotheses of divergence by habitat differentiation, elements of hybridization, incomplete lineage sorting, and phenotypic plasticity obscured the identification of species limits and precluded a better inference of lineage isolation. The analysis of available genomic resources demonstrated the suitability of obtaining multiple low copy nuclear loci from already available genomic data in *Schizopetalon*. However, the use of these markers is yet limited, as the detection of multiple copies implies that further analyses are needed to discard paralogous copies. Overall, this dissertation sets the foundation for more elaborated studies, as more available genomic

resources and intricate pattern of divergence can result promising to explore the consequences of local patterns of extreme aridity in the diversification and evolution of species of Schizopetalae.

Keywords: Schizopetalae, *Schizopetalon*, *Mathewsia*, Atacama Desert, systematics, diversification, aridity.

Acknowledgements

I would like to acknowledge the following people, as contributors to this work and for helping in my formation in Biological sciences. I thank M. Mort, D. Crawford, C. Freeman, J. Soberón, and R. Brown as committee members for guiding my progress throughout my graduate studies. I would also like to thank J. Archibald, D. Martinez-Gordillo, N. Barve, L. Campbell, J. Oaks, and especially my friend A. Lira-Noriega, for their invaluable contributions to my knowledge and development of creative ideas to improve this dissertation. I thank the people who assisted me in obtaining herbarium and field collected specimens, including A. Marticorena (CONC), R. Rodriguez (CONC), M. Muñoz (SGO), E. Wood (GH), C. Morse (KANU), Diego Salariato (SI), Federico Luebert, Lira-Cuadra, T. S. Flores-Fernandez, J. Toro-Nuñez, and P. Ramirez-Toro. I also thank all people who contributed in forming my interest in science and the study of South American plants, including my undergraduate advisors E. Ruiz and M. A. Negritto; I. A. Al-Shehbaz for introducing me to the world of South American Brassicaceae: and G. Plunkett and A. Nicolas for allowing me to share field trips to Chile that inspired me to conduct research in the Atacama Desert. I thank all institutions that have provided funding for this research and my studies at KU, including The University of Kansas Graduate School, The University of Kansas Biodiversity Institute, The University of Kansas Department of Ecology and Evolutionary Biology, Sigma Xi, CONICYT, and Fulbright.

Finally, I would like to thank my parents (Johanna and Guillermo) for their constant support and encouragement during my studies and I want to dedicate this dissertation to my wife Carola and son Gaspar for their unceasing support and hours of company and joy during the hardships found in the development of this study.

Table of Contents

LIST OF TABLES

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER ONE

INTRODUCTION

Beginning in the Miocene (Stuessy & Taylor, 1995; Young, Ulloa Ulloa, Luteyn, & Knapp, 2002), dynamic geological processes and shifts in prevailing ocean currents dramatically changed the landscape of central-northern Chile and adjacent southern Peru and western Argentina. Within this region is found the Valdivian Rainforest Biodiversity Hotspot (Arroyo et al., 2004). This region is effectively a continental island between the Pacific Ocean to the west, the Andes Mountains to the east, and the extremely arid portions of the Atacama Desert to the north (Gengler-Nowak, 2002; Rauh, 1985). Here, the timing and levels of precipitation vary widely, from year-round rainfall in southern Chile to winter-rainfall in central-northern Chile, within the latter there exists a gradient from Mediterranean-like to progressively more arid conditions oriented in a south to north pattern (Gengler-Nowak, 2002; Rauh, 1985; Rundel et al., 1991). The variation in precipitation has been hypothesized to play an important role in the origin and diversification of the flora of Chile. The combined atmospheric and climatic attributes make the flora of the arid regions of Chile an ideal system to study plant diversification in harsh and unpredictable habitats (Arroyo, Squeo, Armesto, & Villagran, 1988; Stuessy & Taylor, 1995).

The diversity of abiotic factors (e.g., microclimatic variation, topography, and long term climatic stability) that govern the dynamics of the aridity in Atacama Desert have been used to explain the distribution and differentiation of the flora in the Atacama (i.e., Arroyo et al., 1988; Luebert, 2011; Marquet et al., 1998; Villagran, Arroyo, & Marticorena, 1983). In fact, based on the

prevailing source of precipitation, the Atacama Desert has been informally divided into two regions, the fog-influenced desert to the north and the fog-free desert to the south (see below). The presence of these conditions has promoted the study of the relationship between diversification and extreme aridity in several native desert taxa (Dillon et al., 2009; Gengler-Nowak, 2002, 2003; Guerrero et al., 2011a; Heibl & Renner, 2012; Hershkovitz, Arroyo, Bell, & Hinojosa, 2006; Jara, 2010; Katinas & Crisci, 2000; Luebert & Wen, 2008; Luebert, Wen, & Dillon, 2009; Perez, Arroyo, Medel, & Hershkovitz, 2006; Tu, Dillon, Sun, & Wen, 2008). As a result, there have been advances in taxonomic, systematic and biogeographic information linking aridity to the diversification of this flora (e.g., Guerrero, Durán, & Walter, 2011b; Guerrero, Rosas, Arroyo, & Wiens, 2013; Luebert, 2011). Despite these advances, a comprehensive understanding of the role of aridity in the diversification of the Atacama flora is mostly lacking.

To examine the potential role of aridity and local climatic conditions in shaping patterns of organismal diversification and morphological evolution, I will conduct a multifaceted study of the South American endemic tribe of Schizopetalae (Brassicales, Brassicaceae). The phylogenetic relationships within and among the two genera of Schizopetalae (*Mathewsia* and *Schizopetalon*; Fig. 1) remain unknown. This lineage possesses advantageous ecological and biogeographic characteristics to study the plant evolution, particularly in the fog-free area of the Atacama Desert. Therefore, focused studies in both *Mathewsia* and *Schizopetalon* present a good opportunity to explore the effect of aridity on diversification.

Tribe Schizopetalae

Mathewsia and *Schizopetalon* belong to Brasssicaceae, a well-known clade of about 338 genera and 3709 species distributed worldwide (Al-Shehbaz, Beilstein, & Kellogg, 2006; Warwick et

al., 2010). Currently, 48 tribes are recognized within Brassicaceae (German & Al-Shehbaz, 2010); *Schizopetalon* and *Mathewsia* have recently been placed together with strong support in the tribe Schizopetalae *s.s* (Warwick, Sauder, Mayer, & Al-Shehbaz, 2009).

The segregation of the former tribe Schizopetalae (*sensu* Al-Shehbaz et al., 2006; Beilstein, Al-Shehbaz, & Kellogg, 2006; Beilstein, Al-Shehbaz, Mathews, & Kellogg, 2008) to include only *Mathewsia* and *Schizopetalon* is congruent with recent phylogenies of the family and is also supported by geographic distribution and morphology (Warwick et al., 2010; Warwick et al., 2009). These genera are the only members of Brassicaceae that simultaneously occur in the coastal desert and scrub (i.e., matorral) of Chile. Both possess flowers that are presumably moth-pollinated, with a long corolla tube, white petals that open only during the night, and pollen grains with a coarse surface reticulum. *Mathewsia* and *Schizopetalon* are distinguished from one another by differences in habit, and petal and fruit shape.

Mathewsia comprises six or nine species of perennial suffruticose plants endemic to northern Chile and southern Peru, with species distributed from the Arequipa district of southern Peru (17° S) to the Valparaiso area (33° S) in Chile (Table 2). Species are distributed mainly along the coast where they occur in dry gullies (*M. biennis, M. collina, M. diversifolia, M. foliosa*, *M. incana,* and *M. peruviana*) or rocky slopes close to the absolute desert area (i.e., 0 mm precipitation; 26° S), but species extend up to 3000 meters (*M. auriculata*, and *M. nivea*) (Rollins, 1966). One previously recognized species, *M. matthioloides* (Rollins, 1966), is now placed within *Sibara*, a genus disjunctly distributed in North and South America (Al-Shehbaz, 2010a, 2010b). The current taxonomy identifies species of *Mathewsia* based on differences of the

pubescence of the fruits and leaves, and the shape of the leaves.

Schizopetalon is a small genus (~ 10 species) of annual herbs that occur primarily in Chile, but one species (*S. rupestre*) is native to both Chile and Argentina (Fig. 3). Species range from the southernmost limit of the absolute Atacama desert (24° S) to the Mediterranean scrub and semiarid area near Santiago and Valparaiso (34° S). Species of *Schizopetalon* can be divided into two groups, based on the ecoregion where they occur: those that inhabit sandy, coastal landscapes from the coast to 800 m (*S. biseriatum*, *S. tenuifolium*, *S. arcuatum*, *S. maritimum*, *S. corymbosum*, and *S. walkeri*), and those that occur along stony slopes of the Andes Mountains from 2000 to 3500 m (*S. bipinnatifidum*, *S. brachycarpum*, *S. dentatum*, and *S. rupestre*; Al-Shehbaz, 1989). The taxonomy of *Schizopetalon* species is based on variation in fruit trichomes, orientation of fruit pedicels, occurrence of flower bracts, seed shape, and structure of the cotyledons.

To date, the molecular systematics of *Mathewsia* and *Schizopetalon* has been limited to the study of tribal levels in Brassicaceae (e.g., Beilstein et al., 2008; Warwick et al., 2010; Warwick et al., 2009). At lower taxonomic levels, the only attempt to reconstruct the phylogenetic relationships in this group was made by Al-Shehbaz (1989) using morphology. Beyond of this point, further aspects of evolution and taxonomy have remained mostly unknown. This has represented a limitation not only for taxonomic studies of *Mathewsia* and *Schizopetalon*, but also precludes insights into the patterns of diversification within and among these genera, which could add significantly to the overall analyses of the evolution of the flora of the Atacama as a whole. Despite its high endemicity, recent accounts in the Atacama flora have been unable to determine

levels of conservation due to the uncertain taxonomic status of their species (Squeo, Arancio, & Gutierrez, 2008). With about a 30% of the species known only from their type specimens (see Chapter 2), conducting studies *Schizopetalon* and *Mathewsia* would not only contribute to the knowledge of Brassicaceae, but also would be an improvement for the study of the Atacama flora.

Chilean Atacama Desert

The Atacama Desert is relatively small (~55, 000 mi^2) compared to other well-studied systems (e.g., Sonoran or Nevi deserts), but is considered one of the most ancient and driest deserts on Earth (e.g., Clarke, 2005; Hartley, Chong, Houston, & Mather, 2005; Houston, 2006; Houston & Hartley, 2003). The geographic extent of the Atacama Desert has been debated (Luebert, 2011), but most consider it to be a subunit of the Coastal Peruvian-Chilean Desert (Rauh, 1985; Rundel et al., 1991). The most accepted biological limits for the Atacama Desert extends from the areas of the southern limit Peru with Chile to the Norte Chico Area (18°S and 30°S, respectively). Nevertheless, this desert is mostly identified for climatological purposes in the Norte Grande area of Chile (Latorre et al., 2007; Rauh, 1985; Rundel et al., 1991) (Fig. 4).

The extreme aridity of the Atacama is the result of the combined effects of the high Andes Mountains to the east, the high-pressure air mass of Pacific anticyclone in the west, and the cold northward Humboldt Current along the Pacific coast. The interaction of these factors promote strong desiccation and a rain shadow effect on the west side of the Andes (Garreaud, Molina, & Farias, 2010; Hartley, 2003; Hartley & Chong, 2002; Houston & Hartley, 2003; Julia, Montecinos, & Maldonado, 2008). The origin of these conditions can be traced back to at least

the Oligocene-Miocene (Dunai, González López, & Juez-Larré, 2005; Garreaud et al., 2010), but aridity *per se* has been suggested in the Atacama at least from Triasic or Jurasic times (Clarke, 2005; Hartley et al., 2005). While the precise timing of the development of aridity in the Atacama Desert is still debatable (reviewed by Latorre et al., 2007), there is evidence that glacial events during the Quaternary caused a shift in precipitation patterns that moved the winter rainfall boundary northward. This shift relaxed the effects of extreme aridity in southern habitats (Lamy, Klump, Hebbeln, & Wefer, 2000; Latorre et al., 2007; Maldonado & Villagrán, 2002; Veit, 1993) and may have allowed for sympatry in formerly isolated lineages. It has been suggested that climatic oscillations, both over geological time and seasonally (e.g., ENSO), can have a major effect the distribution plant lineages and thus might promote periods of isolation and reconnection (Gengler-Nowak, 2002; Jara, 2010; Luebert & Wen, 2008; Ossa, Perez, & Armesto, 2013; Viruel, Catalán, & Segarra-Moragues, 2011).

Based on patterns of aridity and precipitation regime, the Atacama Desert can be subdivided into two areas. One area is located in the North range, covering coastal and mid-altitude areas (0 to 1000 m) from the Azapa Valley to the Copiapo Valley (16°S – 27° S). This section is characterized by the presence of rare to null precipitation and has been suggested to barely harbor microbial life (e.g., Navarro-Gonzalez et al., 2003). The only influence of precipitation in this area is provided by very sporadic events from the oceanic influence at the southern coastal margin in the Copiapo River (27° S) and more frequent summer rainfall from tropical influence in the high Andes and Puna grassland area (di Castri & Hajek, 1976; Luebert & Pliscoff, 2006). Due to this condition, vegetation only occurs along steep coastal cliffs and gullies (Lomas vegetation) influenced by the constant coastal fog and in the Andes over ~ 3000 m (di Castri &

Hajek, 1976; Luebert & Pliscoff, 2006).

The second area is the southern Atacama, a fog-free area (26°S-30°S; Rauh, 1985; Rundel et al., 1991) that is variable in terms of topography and patterns of precipitation. Topographically, the region includes the Andes Mountains (up to 5000 m) in the east, inter-Andean dry valleys in the middle, and a long dry coastal region of extensive plains interrupted by steep cliffs in the west. This topographic heterogeneity interacts with precipitation to generate distinct patterns of vegetation over latitudinal and altitudinal gradients (Arroyo et al., 1988; Luebert & Pliscoff, 2006; Villagran et al., 1983). The southern Atacama is characterized by low and infrequent winter precipitation (Rauh, 1985; Rundel et al., 1991; Rundel & Mahn, 1976), with occasional winter droughts more frequent with decreasing latitude (Latorre et al., 2007). Precipitation is influenced by cold air masses that produce winter precipitation from the south (Julia et al., 2008). These conditions are associated with ENSO events (El Niño – La Niña), which create sporadic bursts of biological activity (i.e., seed germination and flowering) every 3 to 6 years in coastal areas (Armesto, Vidiella, & Gutierrez, 1993; Cereceda et al., 2000; Dillon & Rundel, 1990; Gutiérrez & Meserve, 2003; Holmgren et al., 2006; Jaksic, 1998; Jaksic, 2001).

Aridity and diversification of the Atacama flora

Areas influenced by aridity represent intriguing cases of plant evolution; in part due to their capacity to harbor plant lineages with high levels of diversity and endemism (Comes, 2004). Among the World's deserts, hot deserts, because they present challenging abiotic conditions, are particularly interesting for studies of plant diversification. It is generally accepted that geographic isolation and environmental diversity are factors that promote evolutionary

divergence in plants (Grant, 1981; Levin, 1993; Stebbins, 1950). In this sense, deserts, which are often considered fairly environmentally homogeneous, do not offer a diverse set of habitats to promote local adaptation (Evenari, 1985; Noy-Meir, 1973). Despite this, there are several examples of plant radiation in desert habitats (e.g., Cowling, Rundel, Desmet, & Esler, 1998; Klak, Reeves, & Hedderson, 2004; Luebert & Wen, 2008; Mummenhoff et al., 2005), which brings into question what factors promote the formation of such diversity.

One explanation for this pattern is that the aridity that characterizes deserts directly influences diversification, and possibly local adaptation. By definition, aridity is the imbalance between precipitation and ground evapotranspiration and it is the most important factor in the generation of desiccation in temperate regions (Evenari, 1985). Aridity is recognized as an important force promoting diversification of Angiosperms(Axelrod, 1967, 1972; Stebbins, 1952), especially in desert lineages (Shmida, 1985; Whitford, 2002). These lineages often have acquired reproductive attributes, such as rapid seed/fruit production or mechanisms to assure seed germination when water is available, that allow them to cope with arid conditions (Gutterman, 2002). In addition, water availability effectively dictates where plants can survive and serves to create geographic barriers, which can greatly limit gene flow and produce geographic structure and isolation among populations (e.g., Arafeh et al., 2002; Comes & Abbot, 1999).

The Atacama Desert combines diverse topography, gradients of temperature and precipitation, and temporal variability in precipitation that creates a complex, heterogeneous series of habitats over a small geographic region. These conditions have significant effects on the distribution of lineages, and can isolate and promote divergence at small spatial scales (Luebert, 2010). Despite

this combination of factors that putatively interact to promote diversification, there are few evolutionary studies focused at the population level in the Atacama Desert (Albrecht, Escobar, & Chetelat, 2010; Crawford, Sagastegui-Alva, Stuessy, & Sanchez-Vega, 1993; Gengler & Crawford, 2000; Graham, 2005). Moreover, plants that inhabit the southern, fog-free zone have almost no data related to population diversity (but see Ossa et al., 2013; Viruel et al., 2011), despite the attention that its flora has received for conservation issues (Squeo et al., 2008)

Considering the limited study of the evolutionary effects of aridity in the Atacama Desert, *Mathewsia* and *Schizopetalon* emerge as a suitable system to reconstruct and understand plant diversification in the fog-free Atacama. In this study I will address the question of how conditions of aridity in the Atacama Desert influence local diversification. For this purpose, I focus my study on the systematics of Schizopetalae, with special emphasis on the distribution of morphometric, genetic and climatic diversity of coastal and montane species of *Schizopetalon* from the southern fog-free Atacama. Species of *Schizopetalon* are widely distributed throughout this region and experience a range of arid conditions. I will exploit these characteristics of *Schizopetalon* to understand diversification under extreme aridity and use the Atacama subclade of *Schizopetalon* to explore patterns of interspecific differentiation in extremely arid habitats. Given the lack of information in Schizopetalae, I will:

1) Estimate phylogenetic relationships using rapidly-evolving DNA sequences to test the monophyly and taxonomic limits of *Mathewsia* and *Schizopetalon*, identify their subgroups, and determine a potential monophyletic study group of species of *Schizopetalon* from the Atacama Desert.

2) Analyze the influence of climate divergence and habitat isolation in the generation of interspecific variability, analyzing patterns of morphometric and ecological niche differentiation in the study group.
3) Explore the use of low copy nuclear genes for phylogenetic and phylogeographic studies in Schizopetalon, and explore taxonomic limits and phylogeographic scenarios for the study group.

Figures

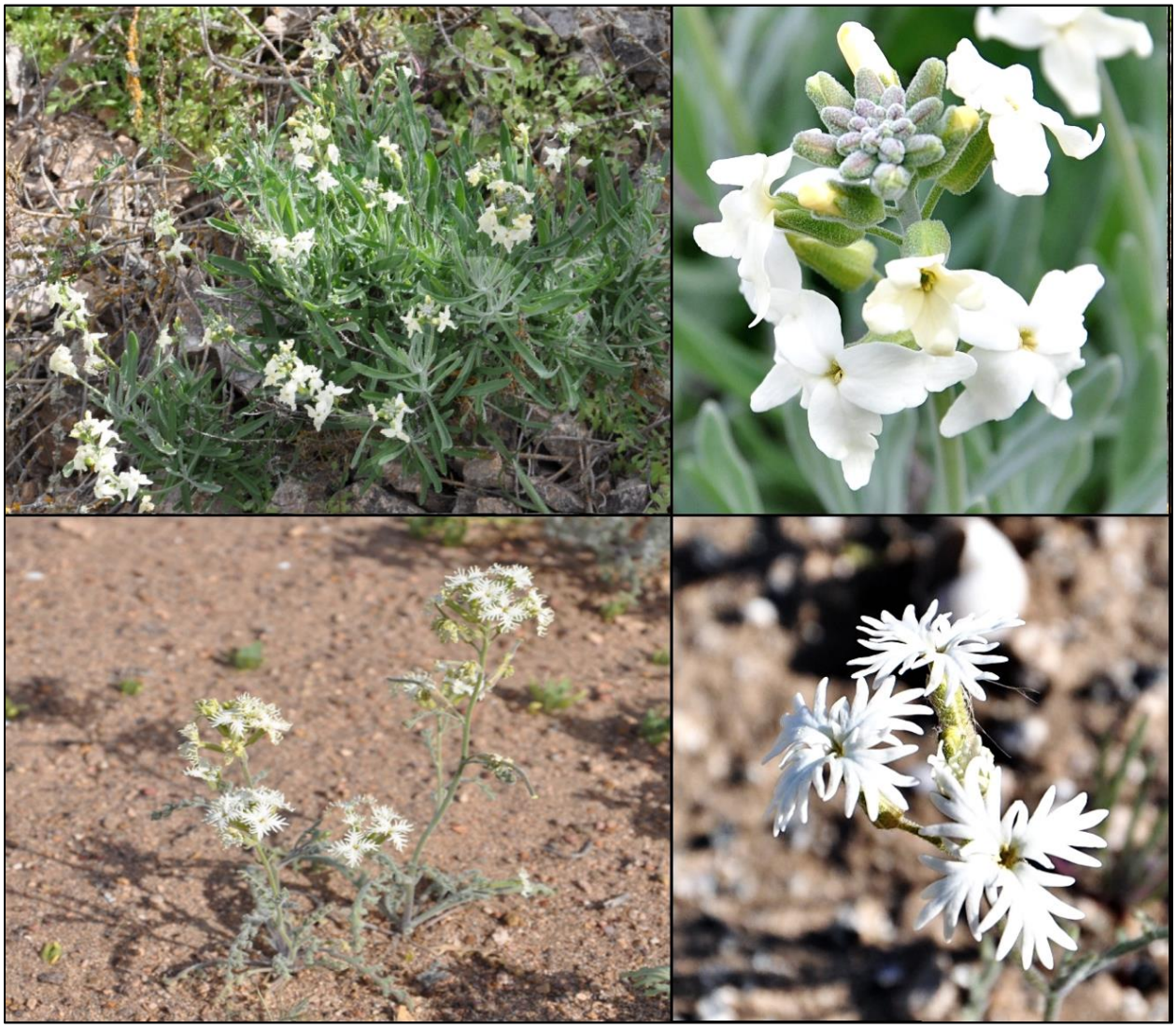


Fig. 1. *Mathewsia* and *Schizopetalon*. Upper pictures: *M. incana*, lower pictures: *S. tenuifolium*.

Pictures at the left: habit, at the right: flowers of *Mathewsia* and *Schizopetalon*, respectively.

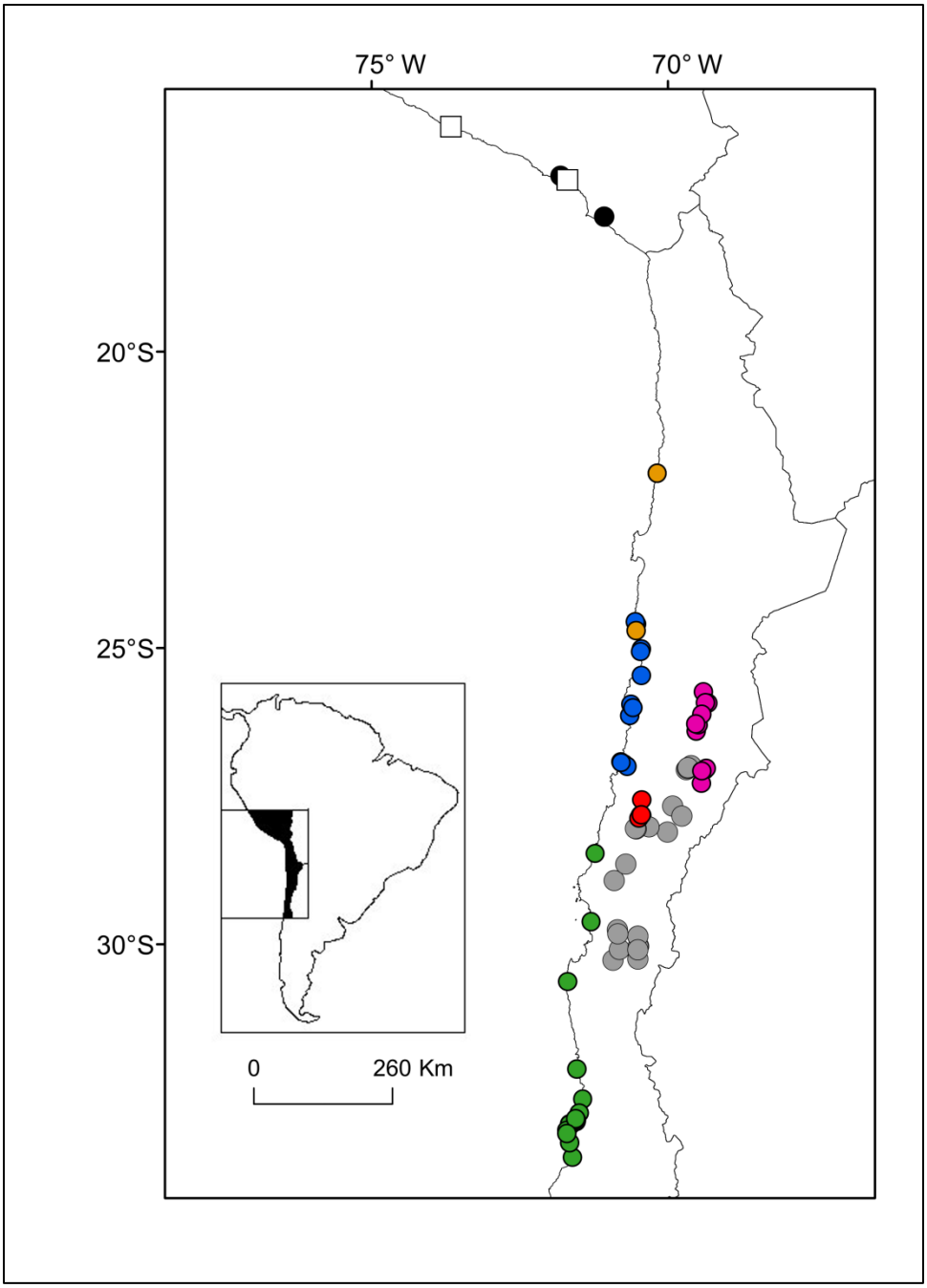


Fig. 2. Geographic distribution of *Mathewsia*: *M. auriculata* (gray dots), *M. biennis* (red dots), *M. colina* (orange dots), *M. diversifolia* (white squares), *M. foliosa* (green dots), *M. incana* (blue dots), *M. nivea* (purple dots), *M. peruviana* (black dots)

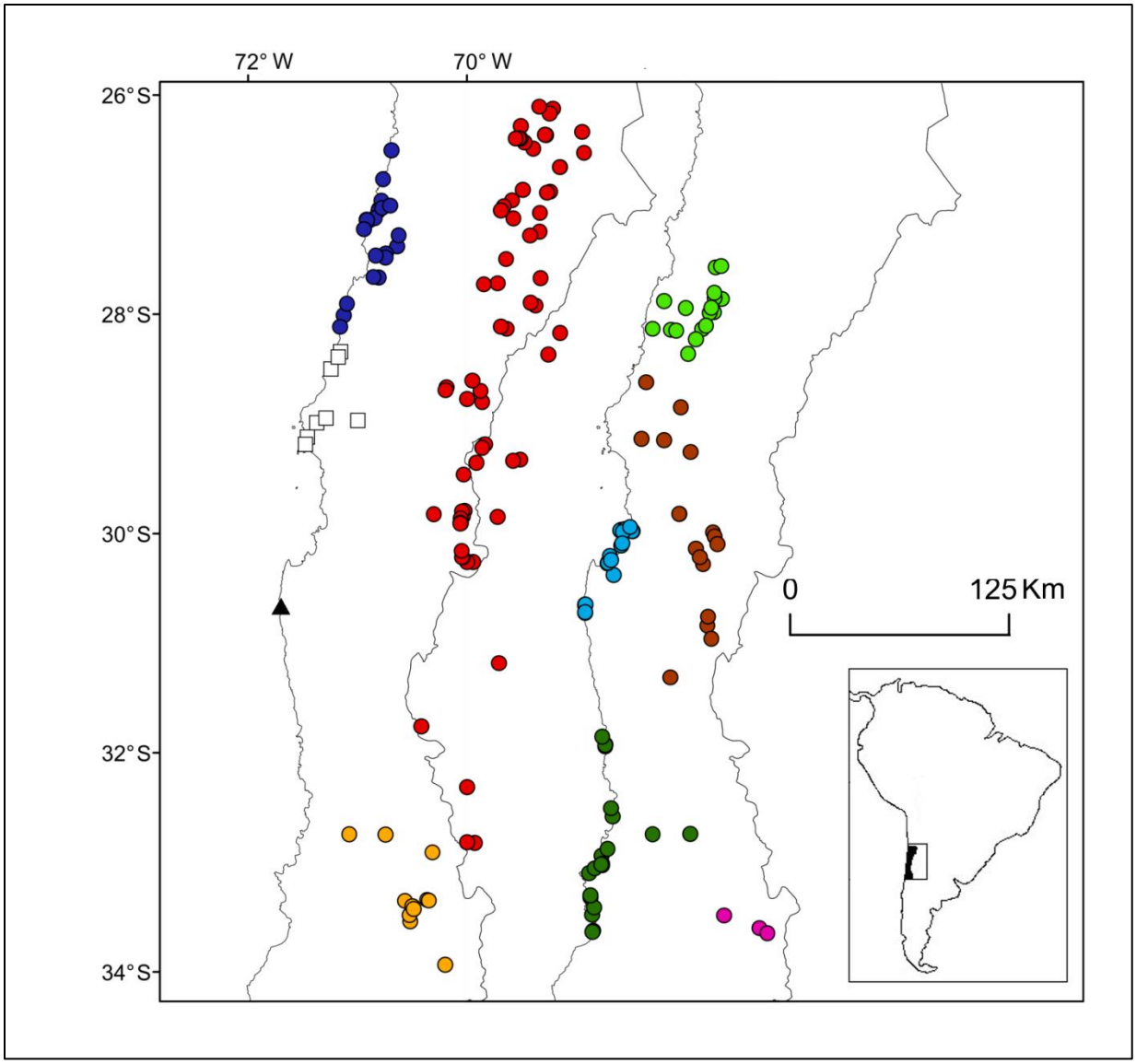


Fig. 3. Geographic distribution of *Schizopetalon*. Left map: *S. arcuatum* (white squares), *S. biseriatum* (blue dots), *S. corymbosum* (black triangle) *S. dentatum* (orange dots), and *S. rupestre* (red dots). Right map: *S. bipinnatifidum* (brown dots), *S. brachycarpum* (purple dots), *S. maritimum* (light blue dots), *S. tenuifolium* (light green dots), and *S. walkeri* (dark green dots).

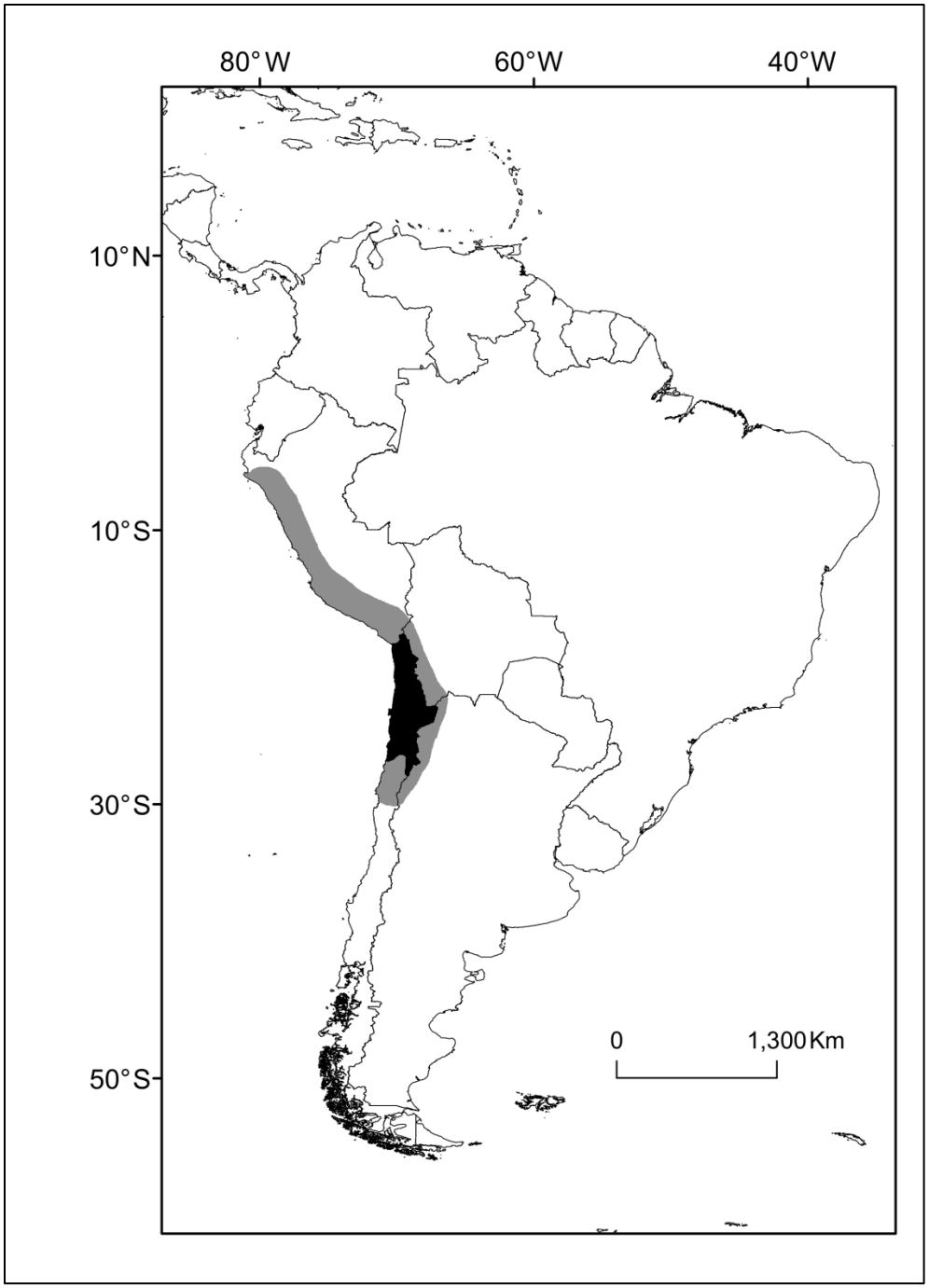


Fig. 4. Extent of the Peruvian-Chilean Desert (gray area) and the Atacama Desert within (black area).

CHAPTER TWO

PHYLOGENETIC RELATIONSHIPS OF *MATHEWSIA* AND *SCHIZOPETALON* (BRASSICACEAE) INFERRED FROM NRDNA AND CPDNA REGIONS: TAXONOMIC AND EVOLUTIONARY INSIGHTS FROM AN ATACAMA DESERT ENDEMIC LINEAGE [1]

Abstract

Mathewsia (6 - 9 spp.) and *Schizopetalon* (10 spp.) are sister genera of the tribe Schizopetaleae (Brassicaceae, Cruciferae). The genera are mostly endemic to Chile and neighboring Peru and Argentina; most species are in the Atacama Desert and adjacent areas. This distribution and the presence of different life forms in *Mathewsia* (perennial) and *Schizopetalon* (annual) provide the opportunity to compare patterns of diversification under extreme environmental pressure. In this paper, phylogenetic relationships are estimated for 17 species using a combination of two nuclear regions (rDNA ITS + ETS) and four plastid regions (*atpI-atpH*, *trnQ-rps16*, *trnH-psbA* spacers, and *rps16* intron) using maximum parsimony, maximum likelihood, and Bayesian approaches. Analyses of the combined cpDNA and nrDNA data result in topologies that are highly similar, although several instances of topological incongruence are noted. Results from all three criteria and all data sets support the monophyly of *Schizopetalon*, but not of *Mathewsia*. *Mathewsia* is rendered non-monophyletic because *M. nivea* is placed outside of Schizopetaleae (*Mathewsia* + *Schizopetalon*); a result that is also supported by morphology. The monophyly of some species is confirmed; however, several currently recognized species were not recovered as a clade and will require additional study. Further analyses are needed to resolve the placement of *M. nivea* and the origins of discordant nuclear and plastid signal.

[1] Chapter published: Toro & al. 2013. Taxon 62 (2):343-356

Introduction

During the last 20 years, the flora of the Atacama Desert (Chile, *sensu* Rundel et al., 1991) has received renewed attention, which has resulted in improved estimates of phylogenetic relationships, more robust classifications, and the formulation of biogeographic hypotheses for Atacama plant lineages (e.g., Dillon et al., 2009; Guerrero et al., 2011a; Heibl & Renner, 2012; Katinas & Crisci, 2000; Luebert & Wen, 2008; Luebert et al., 2009; Viruel et al., 2011). Nevertheless, much of the plant diversity of this desert remains to be the focus of phylogenetic studies (Luebert, 2011). One such lineage comprises *Mathewsia* and *Schizopetalon* (Brassicaceae), which has yet to be the focus of a comprehensive systematic study, despite the large number of endemic taxa and life-history attributes that would prove ideal for studying plant diversification in the Atacama Desert.

Mathewsia Hook. & Arn. and *Schizopetalon* Sims belong to Brassicaceae, a well-known monophyletic family of 49 tribes, about 322 genera, and 3660 species that are distributed worldwide (Al-Shehbaz, 2012a; Al-Shehbaz et al., 2011; German & Al-Shehbaz, 2010). Based on limited sampling, the two genera have recently been placed together with strong support (> 90% bootstrap) to form tribe Schizopetaleae s.str. (Warwick et al., 2010; Warwick et al., 2009). The re-circumscription of the former tribe Schizopetaleae (*sensu* Al-Shehbaz et al., 2006; Beilstein et al., 2006; Beilstein et al., 2008) to include only *Mathewsia* and *Schizopetalon* is congruent with recent phylogenies of the family and is also supported by geographic distribution and morphology (Warwick et al., 2010; Warwick et al., 2009). These genera are the only members of the former tribe Schizopetaleae whose species are primarily distributed across the southern Atacama Desert and the matorral of central Chile (26°–33° S), with a few species across

southern Peru and the neighboring montane areas of Chile and Argentina. Their morphology is characterized by the presence of a mixture of simple and dendritic trichomes (Beilstein et al., 2008; Warwick et al., 2009), and fragrant flowers with long corolla tubes, white petals, and pollen grains with a coarse surface reticulum (Al-Shehbaz, 1989), which are thus presumably moth-pollinated.

Mathewsia comprises six to nine species of perennial shrubs and subshrubs endemic to northern Chile and southern Peru, with species distributed from the Arequipa district of southern Peru (17° S) to the Valparaiso area (33° S) in Chile (Fig. 1; Appendix 1). Species are distributed mainly at low elevations along the coast, where they occur in dry gullies (*M. linearifolia* Turcz., *M. collina* I.M. Johnst., *M. densifolia* Rollins var. *densifolia*, *M. densifolia* var. *stylosa* Rollins, *M. foliosa* Hook. & Arn., *M. incana* Phil., *M. peruviana* O.E. Schulz); however, *M. auriculata* Phil. and *M. nivea* (Phil.) O.E. Schulz occur on rocky slopes in the Andes up to 3000 meters (Rollins, 1966). *Mathewsia* is recognized for its suffruticose habit, yellowish-white flowers with entire petals, and incumbent cotyledons. The current taxonomy (Rollins, 1966) distinguishes among the species of *Mathewsia* by differences in leaf shape and pubescence of the fruit valves and leaves.

Schizopetalon comprises ten species of annual herbs that occur primarily in Chile, although one species (*S. rupestre* (Barn.) Reiche) is native to both Chile and Argentina (Fig. 2; Appendix 1). Species range from the southernmost limit of the absolute Atacama Desert (26° S) to the matorral scrub and semiarid areas near Santiago and Valparaiso (34° S). Like *Mathewsia*, species of *Schizopetalon* can be divided into two groups based on the ecoregion where they occur. One

group inhabits sandy, coastal landscapes from the coast to 800 m altitude (*S. arcuatum* Al-Shehbaz, *S. biseriatum* Phil., *S. corymbosum* Al-Shehbaz, *S. maritimum* Barn., *S. tenuifolium* Phil., *S. walkeri* Sims), and the second group occurs along rocky slopes of the Andes from 2000 to approximately 3800 m (*S. bipinnatifidum* Phil., *S. brachycarpum* Al-Shehbaz *S. dentatum* Gilg & Mushl., *S. rupestre*; Al-Shehbaz, 1989). *Schizopetalon* is morphologically distinct from *Mathewsia* because of its annual habit, the presence of white flowers with pinnatifid petals, mostly uniseriate seeds, and filiform or incumbent cotyledons. The last character, along with the variation in fruit trichomes, orientation of fruiting pedicels, seed shape, and occurrence of flower bracts, are used in combination to distinguish the species of *Schizopetalon* (Al-Shehbaz, 1989).

Both *Mathewsia* and *Schizopetalon* have species that are present in the Atacama Desert, which make them good candidates for studies of plant diversification in hyperarid habitats. The Atacama Desert is characterized by its well-defined geographic boundaries, variable topography, rare and intermittent episodes of rainfall, and a marginal gradient of aridity as a function of latitude and altitude (e.g., Luebert & Pliscoff, 2006; Rauh, 1985; Rundel et al., 1991). Such characteristics have made this area attractive for studies of plant diversification (e.g., Dillon et al., 2009; Gengler-Nowak, 2002; Katinas & Crisci, 2000; Luebert & Wen, 2008; Luebert et al., 2009). In contrast to previously studied groups, *Mathewsia* and *Schizopetalon* offer a system to explore the diversification of two lineages that have different life-history strategies (perennial vs. annual), but share a recent common ancestor. However, a necessary first step is to test the current classification, the monophyly of these genera, and the relationships of their species by producing a robust estimate of phylogeny based on multiple markers.

The last taxonomic revisions of *Mathewsia* and *Schizopetalon* were produced by Rollins (1966) and Al-Shehbaz (1989), respectively. These revisions suggest that a complete knowledge of the total diversity of species is still far from being achieved, bringing into question the monophyly of some of the proposed species. A complicating factor to studying these genera is the paucity of herbarium specimens, which might reflect a lack of collecting efforts and/or the rare and scattered occurrence of populations in the field. For example, about 30% of the species of *Mathewsia* and *Schizopetalon* are only known from their type localities (e.g., *M. collina*, *M. densifolia*, *S. arcuatum*, *S. brachycarpum*, *S. corymbosum*). Similarly, the only attempt to reconstruct the phylogenetic relationships in the two genera was made in *Schizopetalon* based on an analysis of morphology (Al-Shehbaz, 1989). To date, no other studies have been conducted using any alternative source of characters.

This paper will present the first comprehensive study of the phylogenetic relationships within the tribe Schizopetaleae and its two genera *Mathewsia* and *Schizopetalon* using a combination of nrDNA and cpDNA regions. The main objectives of the study are to: (1) estimate phylogenetic relationships in *Mathewsia* and *Schizopetalon* using multiple DNA data sets and (2) assess the monophyly of currently recognized taxa.

Materials and Methods

Taxon sampling

Taxa were obtained from a combination of herbarium collections (CONC, MO, OS, GH) and specimens collected during fieldtrips in July 2010 and December 2011 (Appendix 1).

Approximately 95% of the species diversity of *Schizopetalon* and *Mathewsia* is represented in this sampling, including two individuals of each species if it was possible. For rare, geographically restricted species, DNA was extracted from type specimens and their sequences were compared to those obtained from vouchers collected in the proximity of the type localities. The only exception was *S. arcuatum,* which was only sampled from populations located in the original collecting site, corroborating their identities with morphology. Only two species were not included in the present study: *M. mathioloides* and *S. corymbosum*. The former species, which was recognized by Rollins (1966), was excluded because it is now believed to belong to *Sibara* Greene, a genus disjunctly distributed in North and South America (Al-Shehbaz, 2010a, 2010b, 2012b); *S. corymbosum* was excluded because no type or field-collected specimen was available for study.

Because the sister tribe to Schizopetaleae has not been clearly resolved, outgroups were selected based on broad phylogenetic studies of Brassicaceae(Al-Shehbaz et al., 2006; Beilstein et al., 2008; Warwick et al., 2010; Warwick et al., 2009). These include *Eudema nubigena* Bonpl. (Eudemeae), *Menonvillea chilensis* (Turcz.) B.D. Jacks. (Cremolobeae), *Weberbauera colchaguensis* (Barn.) Al-Shehbaz, and *Neuontobotrys tarapacana* (Phil.) Al-Shehbaz (Thelypodieae). All outgroup and ingroup taxa were rooted with *Lepidium angustissimum* Phil. (Lepidieae) as a representative taxon of the sister lineage of Schizopetaleae and its allied tribes in Brassicaceae (Lineage I; Franzke et al., 2011).

Molecular data and sequencing

A small amount of silica-dried material (for field-collected specimens) or mounted specimens

from herbarium collections was used to extract total DNA using a modified CTAB method (Doyle & Doyle, 1987; Mort et al., 2001). For samples with low DNA quality and/or older material (from the 1970s or earlier), the Archive Pure DNA plant purification kit (5 PRIME, Gaithersburg, Maryland, U.S.A.) was used. Sequence data from the nuclear ribosomal ITS and ETS and the chloroplast regions *trnH-psbA* (*psbA*), *atpI-atpH* (*atpH*), *rps16-trnQ* (*trnQ*) spacer regions, and the *rps16* intron were obtained because of their documented phylogenetic utility for studies at lower taxonomic levels (Mort et al., 2007; Shaw et al., 2005; Shaw, Lickey, Schilling, & Small, 2007). These target DNA regions were amplified using the primer combinations N-nc18/c26A (ITS; Wen & Zimmer, 1996), ETS-18S/ETS-9 (ETS; Wright et al., 2001), trnH(GUG)/psbA (*trnH-psbA* spacer; Hamilton, 1999), rps16F/rps16R (*rps16 intron*; Shaw et al., 2005), atpI/atpH (*atpI-atpH* spacer; Shaw et al., 2007), and trnQ(UUG)/5′rps16 (*trnQ-rps16* spacer; Shaw et al., 2007) following the amplification protocols published by the respective authors. The same primers were used for forward and reverse DNA sequencing reactions. All sequencing was performed at the KU core DNA sequencing facility or by Macrogen Inc. (Rockville, Maryland, U.S.A.). The resulting contigs were assembled and edited using Sequencher v.5.0 (Gene Codes Corporation, Ann Arbor,Michigan, U.S.A.).

Phylogenetic analyses

Samples were labeled in the phylogenetic analyses by their names and identified using morphology (Appendix 1). DNA sequences were aligned using MUSCLE v.3.8.31 (Edgar, 2004) and corrected visually to minimize apparent character state changes with Jalview v. 2.7 (Waterhouse et al., 2009). In order to assess topological congruence among different phylogenetic reconstruction criteria, relationships were compared using maximum parsimony

(MP), maximum likelihood (ML) and Bayesian (BI) methods. Phylogenetic analyses were conducted individually for each DNA region and combined as nrDNA and cpDNA partitions.

MP analyses were conducted using TNT v.1.1 (Goloboff, Farris, & Nixon, 2008) with characters equally weighted and using a heuristic search with TBR + RAS of 5000 replicates retaining two trees per iteration. Levels of support for branches were estimated with 1000 bootstrap replicates (Felsenstein, 1985). Additionally, the phylogenetic utility of indels was explored in MP in an attempt to detect additional support for the groups retrieved. Both simple and complex indels were coded using the methods described by Simmons & Ochoterena (2000), as implemented in SeqState v.1.32(Müller, 2005). These characters were added as additional unordered binary state data and significant differences in bootstrap support values and new retrieved clades were reported.

For ML and BI analyses, independent models of molecular evolution for each DNA partition were estimated using jModelTest v.0.1.1 with the corrected Akaike information criterion test (Guindon & Gauscuel, 2003; Posada, 2008). ML analyses were conducted with GARLI v.2.0 (Zwickl, 2006) using two independent search runs, with a maximum of five million generations each. BI analyses were conducted using MrBayes v.3.12 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003), using two independent runs with three hot chains and one cold chain through ten million generations and sampling every 100 samples. The first 1000 samples were discarded as burn-in after confirming stationarity of the Markov chains for all the model parameters using Tracer v.1.5 (Rambaut & Drummond, 2007). The convergence of posterior splits probabilities among runs was assessed using the program AWTY (Are we there yet?;

Nylander, Wilgenbusch, Warren, & Swofford, 2008). Support values in ML were calculated using bootstrap with 200 replications in GARLI and the posterior probabilities in the most credible tree for BI analysis. For convenience, bootstrap values greater than 80% were considered moderately supported and greater than 90% as highly supported for MP and ML. For Bayesian analysis, clades with posterior probabilities greater than 0.95 were considered significantly supported (Alfaro & Holder, 2006). Trees retrieved from the analysis of all different criteria (MP, ML, BI) were visualized and edited with TreeGraph 2 (Stöver & Müller, 2010) for presentation of the results.

Data set and partition congruence

To address levels of congruence among data partitions and their influence on combined data sets (nrDNA & cpDNA), incongruence tests were conducted at character and topological levels. Character disagreement was analyzed with an incongruence length difference (ILD) test (Farris, Källersjö, Kluge, & Bult, 1994) of 1000 permutations conducted in TNT (*ILD.run* script) among all pairwise combinations of DNA partitions for the cpDNA and nrDNA data sets. Incongruence at topological level was assessed through the direct observation of nodes with well supported incongruence (clades with bootstrap > 80% and posterior probability > 0.95 or "hard incongruence" (Wendel & Doyle, 1998) between the resulting topologies among partitions and data sets. Additionally, levels of support and disagreement of clade robustness for each partition within nrDNA and cpDNA data sets were measured with partitioned Bremer support (PBS; Gatesy, O'Grady, & Baker, 1999; Lambkin, Lee, Winterton, & Yeates, 2002) using the script *bpsup.run* (Peña et al., 2006) in TNT. PBS assigns an averaged Bremer index (Bremer, 1988, 1994) with positive and negative values according to the support and disagreement for a

particular node that each partition presents in a combined tree (Lambkin et al., 2002).

Results

Analysis of congruence

In general, all tree topologies from MP, ML and BI analyses were similar except for the strong topological incongruence detected in samples of *Schizopetalon bipinnatifidum* (Los Molles) and *S. brachycarpum* between the nrDNA and cpDNA data sets (ILD, $P < 0.001$). Also, a hard incongruence was detected for the grouping of *Mathewsia incana* in the ML and BI analyses. In general, analyses of the cpDNA and nrDNA data sets resulted in topologies that were highly congruent. Significant levels of character incongruence were only detected among the cpDNA regions (ILD, $P < 0.03$). Examination of the individual trees retrieved for each of these regions (trees not shown) revealed low phylogenetic resolution and no clades with hard incongruences. Only one weakly supported incongruence was detected (MP and ML bootstrap < 80%) in the grouping of a sample of *S. rupestre*, which was placed in different clades in the *atpH* and *rps16* data sets. PBS values for these partitions also supported this observation (see below). With the elimination of this sample, a significant increment in the ILD score for the cpDNA data set was obtained (ILD, $P > 0.08$). As no other evidence of topological incongruence was detected among any of the DNA regions that might impede the use of combined data sets, the results and discussion will be focused only on the separate analysis of the nrDNA and cpDNA data sets.

Sequences and data sets

Thirty-nine ingroup and outgroup samples were successfully amplified for the cpDNA and nrDNA regions. These included two samples of a population of *Schizopetalon bipinnatifidum*

(Los Molles), which were included because of the hard incongruence detected between nrDNA and cpDNA data sets. Two vouchers from the same area (Los Molles, Coquimbo, Chile), with different collection dates (*Garaventa 8229* at OSU from 1954 and *Jiles 3274* at CONC from 1974; Appendix 1), were sampled. No differences were detected in any of the DNA sequences sampled from nuclear and chloroplast regions between these two samples that could suggest a potential error of DNA amplification or sequencing. Only one specimen each of *Mathewsia peruviana*, *S. dentatum*, *S. maritimum*, and *S. tenuifolium* was available for study.

Most of the specimens amplified for all DNA regions. The ETS region was the most difficult to amplify for all samples, especially for the outgroup taxa *Eudema nubigena*, *Menonvillea chilensis*, and *Weberbauera colchaguensis* (Appendix 1). ETS was scored as missing data for these three taxa; analyses conducted both including and excluding these taxa revealed that their inclusion did not alter the resulting topologies. Independent BI analyses using different priors for branch length in each data set (brlenspr = Unconstrained Exp: (1), (10), (100)) did not reveal any effects associated with including taxa with missing data on the posterior probabilities of the resulting clades in the ingroup (Lemmon, Brown, Stanger-Hall, & Moriarty Lemmon, 2009; but see Wiens & Morrill, 2011) proposed in this study (data not shown). No major alterations were made in the data sets with the exception of a poly A-T section in the *trnH-psbA* region (positions from base pair 356 to base pair 373), which was discarded due to difficult alignment. Polymorphisms for nrDNA and large insertions in cpDNA were detected across samples. Those data were maintained as no influence was detected in the retrieved topologies from the different analyses.

The alignment of the nrDNA regions resulted in a data set of 1312 total characters, with 912 constant and 211 informative characters (16.08% of the total, 8.92% for the ingroup). The cpDNA data set resulted in a total of 2810 characters, with 2371 constant characters and 183 informative characters (6.51% of the total, 4.66% for the ingroup). The cpDNA regions contained a larger number of indels than nrDNA: 138 (53 ingroup) and33 (10 ingroup), respectively. All statistics for each region and data set are summarized in Table 1.

nrDNA analyses

MP analysis resulted in three trees of 565 steps each (Fig. 3A; CI = 0.829, RI = 0.890), and their consensus topology was highly congruent with the best ML tree (equal in both runs) and the BI most credible tree (Fig. 3B). ML and BI analyses resulted in trees with exactly the same topology, and differences were only detected in branch lengths. The inclusion of indels did not significantly alter the MP topologies. The only differences detected by including indels were the increase of support levels for two samples of *S. arcuatum* (i.e., *S. arcuatum* 2, *S. arcuatum* 3) in the cpDNA data set (Fig. 4A) and the weakly supported (bootstrap < 60) association of the Atacama with the Mediterranean subclade of *Schizopetalon* in the nrDNA data set (Fig. 3A).

The nrDNA data partitions under all optimality criteria supported the monophyly of *Schizopetalon*, but not of *Mathewsia* due to the placement of *M. nivea* outside the tribe Schizopetaleae (Fig. 3A, B). Within the *Mathewsia* clade, two latitudinal groups (Fig. 1) are retrieved with high support. Henceforth, we will refer to these subclades for both nrDNA and cpDNA (see below) data sets as: northern subclade (*M. collina*, *M. peruviana*, *M. densifolia*) and southern subclade (*M. auriculata*, *M. linearifolia*). The nuclear data suggested an association of

the northern subclade with *M. incana* while the southern subclade is associated with *M. foliosa*. High or moderate support is obtained for the monophyly of all species of *Mathewsia*, except *M. linearifolia* (only BI significantly supported), *M. densifolia*, and *M. incana* (neither monophyletic).

Within the *Schizopetalon* clade, two subclades were retrieved in most of the analyses with high support, which are named by their ecogeographic affinity (Fig. 1) for both nrDNA and cpDNA data sets: the Atacama subclade (*S. biseriatum*, *S. tenuifolium*, *S. rupestre*, *S. bipinnatifidum*) and the Matorral subclade (*S. dentatum*, *S. walkeri*, *S. maritimum*) (Figs. 3, 4). Whereas the Atacama subclade received high support in all analyses, the Matorral subclade was only highly supported by ML and BI (Fig. 3B). Additionally, a highly supported lineage of three populations of *S. arcuatum* was recovered by all analyses. Although each of these three lineages of *Schizopetalon* are well supported, there is no resolution among them based on analyses obtained of the nrDNA data set, except for a weakly supported relationship (bootstrap < 60%, not shown) between the Atacama and Matorral subclade indicated by MP analyses that include indels (Fig. 3A).

cpDNA analyses

A total of ten trees of 601 steps was recovered with MP (Fig. 4A; CI = 0.813, RI = 0.823). In general, the clades recovered by MP analyses (Fig. 4A) were similar to the best ML tree and the most credible BI tree. In this data set, *Schizopetalon* was recovered as monophyletic, but this was only significantly supported in the ML and BI analyses (Fig. 4B). However, consistent with the analysis of the nrDNA data set, *Mathewsia* was not recovered as monophyletic because of the position of *M. nivea* (Fig. 4A, B). Additionally, *M. foliosa* was placed as sister to *Schizopetalon*,

but only the BI analyses recovered this with significant support (Fig. 4A, B). Within the *Mathewsia* clade, northern and southern subclades were again recovered, with a highly supported association of *M. incana* to the southern subclade in ML and BI. For this analysis, none of the geographic subclades is supported by the MP analysis, and the northern subclade is supported only by BI. *Mathewsia collina* and *M. foliosa* were the only species recovered as monophyletic with high support (only BI significantly supported).

In *Schizopetalon*, the Atacama and Matorral subclades were recovered with varying levels of support, and *S. arcuatum* was placed as sister to the Atacama subclade with high support in all analyses. The monophyly of *S. bipinnatifidum*, and *S. walker* was confirmed with moderate to high support. The inclusion of indels did not recover any additional groups, but increased the support for clades of *S. arcuatum*, *S. brachycarpum*, and *S. bipinnatifidum*.

Discussion

This study represents the first attempt to reconstruct phylogenetic relationships in *Mathewsia* and *Schizopetalon*. In general, most of the retrieved groups (over 80% of the total clades) had moderate to high levels of support, and the presence of hard topological incongruence between the nrDNA and cpDNA data sets was detected in limited cases.

Monophyly of Mathewsia and Schizopetalon

Analyses of both cpDNA and nrDNA confirm the sister relationship of *Mathewsia* and *Schizopetalon* (Figs. 3, 4), thus supporting previous studies that placed them in the tribe Schizopetaleae (Warwick et al., 2010; Warwick et al., 2009). At the genus level, results provide

high or almost moderate support for the monophyly of *Schizopetalon*. However, the monophyly of Mathew*sia* is not supported; *M. nivea* is placed not only outside the genus but also outside of Schizopetalae (Figs. 3, 4). Although the taxonomic sampling outside of Schizopetaleae in the present study precludes robust placement of *M. nivea* in Brassicaceae, it is clear that *M. nivea* is not resolved within *Mathewsia*. Thus, the results of the current study highlight the necessity to redefine the boundaries of *Mathewsia*. Historically, *Mathewsia* has been defined as a genus of perennial subshrubs or shrubs that have entire petals (Garaventa, 1940; Rollins, 1966). However, several overlooked fruit morphological features make *M. nivea* distinctive among species of *Mathewsia*. For example, it is the only species that has slightly flattened, linear fruits with uniseriate seeds, compared with the strongly flattened, oblong to lanceolate fruits with biseriate seeds characteristic of the rest of the genus. While these fruit and seed characters were noted by previous students of Brassicaceae, they were not considered sufficient enough to place *M. nivea* in a different genus. In part, this was because the fruits of Brassicaceae were assumed to show high levels of interspecific variation (Al-Shehbaz, 1989; Rollins, 1966) and Brassicaceae are well known to show convergent evolution in many fruit characters (e.g., Al-Shehbaz et al., 2006; Franzke et al., 2011; Koch, Al-Shehbaz, & Mummenhoff, 2003). Another striking difference between *M. nivea* and the remainder of *Mathewsia* is flower color. The petals of *M. nivea* are intensely yellow whereas the rest of the genus has white or yellow-creamy petals. These differences in flower color cannot be seen in herbarium specimens (OTN, pers. obs.). The present results suggest that *Mathewsia* should be delimited to include only those species that possess white to yellow-creamy flowers and flattened, lanceolate to oblong fruits with biseriate seeds.

Except for the MP of the cpDNA data set, all analyses of both data partitions recovered *Schizopetalon* as a strongly supported, monophyletic genus. The results agree with the previous morphological circumscription of the genus that considered it peculiar in Brassicaceae for having pinnatifid petals (Al-Shehbaz, 1989; Franzke et al., 2011; Warwick et al., 2009). Despite the general agreement of previous taxonomic treatments (Al-Shehbaz, 1989; Gilg & Muschler, 1909; Reiche, 1895), the monophyly of *Schizopetalon* remained unclear. This is because alternative hypotheses (Barnéoud, 1845, 1846) suggested that the unique seed and cotyledon diversity of *Schizopetalon* might reflect two genera instead of one: *Schizopetalon* comprised of species with filiform cotyledons and *Perreymondia* Barnéoud with species that have incumbent cotyledons. Our results do not support that separation, as no monophyletic group was retrieved with the exclusive presence of one cotyledon type. The present data supports earlier studies that treated *Schizopetalon* as a single genus despite the differences in cotyledon morphology (Al-Shehbaz, 1989; Gilg & Muschler, 1909; Reiche, 1895).

Mathewsia subclades and relationships

The results suggest that *Mathewsia* can be divided into two geographical groups: the northern subclade, and the southern subclade plus *M. incana* and *M. foliosa* (Figs. 3, 4). The northern subclade includes *M. densifolia, M. peruviana*, and *M. collina*. Species in this clade are highly similar in their morphology, being distinguished only by differences in density of pubescence on the fruits, pedicel length, and the presence of bracteate racemes (Rollins, 1966). Of all *Mathewsia* species, the species of the northern subclade are probably the least known in terms of number of collected specimens. *Mathewsia collina* was recovered with strong support for a grouping of the type specimen (*Johnston 1599* at GH) and a specimen (*Quezada & Ruiz 263* at

CONC) erroneously identified as *M. biennis* (*M. biennis* = *M. laciniata*, Al-Shehbaz, pers. comm.). This result is of particular interest and significance because *M. collina* is known only from these two samples that were collected in 1925 and 1991, respectively. Our analyses recover a clade comprising the type specimens of the two suggested varieties of *M. densifolia* (var. *densifolia*– *Ferreyra 11678* and var. *stylosa* – *Vargas 8587*) and a sampleof *M. peruviana*. Although this lineage is well supported, there is a general lack of resolution among these taxa (Figs. 3, 4).

The relationships found among the species of the northern subclade are of special interest because of the biogeographic relevance of the area where they occur. They are distributed disjunctly along the coast of southern Peru (*M. densifolia*, *M. peruviana*) and the northern coast of Chile (*M. collina*), respectively (16°– 24° S; Fig. 1). The Chilean and Peruvian areas harbor very restricted plant communities (i.e., lomas) that are phytogeographically separated from each other by an area of hyperaridity (e.g., Dillon & Rundel, 1990; Gengler-Nowak, 2002; Luebert, 2011; Luebert & Pliscoff, 2006; Rauh, 1985; Rundel et al., 1991). Our results suggest a well-supported relationship of the northern clade species and are concordant with recent studies that reveal a close relationship between species from the two areas (e.g., Gengler-Nowak, 2002; Luebert, 2011; Luebert & Wen, 2008).

In contrast to the arid northern coast of the Atacama Desert, the southern Atacama area is characterized by the presence of a more gradual topological relief and higher levels and frequency of winter precipitation over a limited geographic range (Luebert & Pliscoff, 2006; Rauh, 1985; Rundel et al., 1991). The southern subclade comprises *M. auriculata* and *M.*

linearifolia and was recovered with significant support except for the MP analyses of the cpDNA data set (Fig. 4A). These species occur in the southern Atacama Desert (26°–30° S) along rocky hillsides in the Andes and inland areas (Fig. 1). In contrast to the species of the northern subclade, the morphological differentiation between and within *M. auriculata* and *M. linearifolia* is more conspicuous. This is based on the presence of entire and auriculate vs. lobed leaves, presence vs. absence of bracts, and glabrous vs. pubescent fruits, respectively. *Mathewsia auriculata* exhibits a high level of morphological variation in its leaves, which might be explained by its distribution in a large number of ecoregions across the inland and the Andes of the southern Atacama (Fig. 1). In contrast to this, *M. linearifolia* is restricted to a small locality at low to medium altitude (~500–1000 m) and exhibits lower morphological variation.

The two representative accessions of *M. incana* were not resolved as sister by either data partition or analytical method. The lack of monophyly might reflect paucity of molecular divergence, which precludes the establishment of unambiguous relationships to other species. Furthermore, analyses of the cpDNA and nrDNA data place these accessions in different positions; analyses of nrDNA place them in the northern subclade, whereas cpDNA data place them in the southern subclade (Figs. 3, 4). Evidence from geographic distribution suggests that this incongruence might be explained by considering *M. incana* to be of hybrid origin between parental species from the two subclades. *Mathewsia incana* is located in humid gullies along the dry coast of Pan de Azucar National Park and Caldera (24°–27° S; Fig. 1) on the periphery of the southern limit of *M. collina* (northern subclade) and the northern limit of *M. linearifolia* and *M. auriculata* (southern subclade). However, while the geographic distribution of the putative parental species makes the hybrid origin of *M. incana* feasible, our results do not offer

conclusive evidence and further studies are required to test this hypothesis.

In case of the placement of *Mathewsia foliosa*, a soft topological incongruence was detected (bootstrap < 70, e.g., Wendel & Doyle, 1998). The nuclear data set (Fig. 3) strongly allies *M. foliosa* with the southern subclade, which is compatible with the close geographic distribution of these taxa. This species is distributed across the semiarid and Mediterranean coast of Chile, from Huasco to Valparaiso (28°–33° S; Fig. 1). In contrast, the cpDNA data set (Fig. 4) suggests that *M. foliosa* is sister to *Schizopetalon*; although with weak support (except in the BI analysis). This result is surprising in that morphologically *M. foliosa* is much more similar to *Mathewsia* than to *Schizopetalon*. In addition, there is no evidence for different rates of molecular evolution between these lineages generating artifacts like long branch attraction (Bergsten, 2005; Wendel & Doyle, 1998). It seems possible that this grouping of haplotypes is due to the retention of a haplotype of the common ancestor of *M. foliosa* and *Schizopetalon*. However, testing this hypothesis will require additional sampling of both gene regions and taxa.

Schizopetalon subclades and relationships

The results of the analyses of both nrDNA and cpDNA data sets recover two subclades grouped by ecogeographic affinity (i.e., Atacama and Matorral, Fig. 1), and three subclades with different levels of support and topological congruence among data sets (i.e., *S. arcuatum*, *S. brachycarpum*, and *S. bipinnatifidum* Los Molles). The Atacama subclade comprises *S. biseriatum, S. bipinnatifidum, S. tenuifolium,* and *S. rupestre*. Similar to the southern clade of *Mathewsia*, these species occur along the coast and Andean ranges of the southern portion of the Atacama Desert and montane areas of Chile and Argentina (Fig. 2). The coastal species (i.e., *S. biseriatum* and *S. tenuifolium*) occur from Chañaral (26° S) to the north of Huasco (28° S), and

the montane species (i.e., *S. bipinnatifidum* and *S. rupestre*) occur from El Salvador (26° S) to the vicinity of Andes of Santiago (Chile) and Mendoza (Argentina) (33° S). Morphological characters partially support the delimitation of the Atacama subclade because of the common, but not exclusive, presence of ovate-oblong seeds and undivided cotyledons.

All relationships among species in the Atacama subclade were either poorly resolved or poorly supported, except for relationships suggested by the BI analysis of the chloroplast data (Fig. 4B). This analysis indicated a sister relationship between *S. bipinnatifidum* and a clade comprising *S. tenuifolium*, *S. biseriatum*, and *S. rupestre*. Additional evidence for the monophyly of *S. bipinnatifidum* (excluding Los Molles population because of topological incongruence between nrDNA and cpDNA data sets) is obtained from the MP analyses of cpDNA + indel characters, and for the monophyly of *S. biseriatum* from the ML of the nuclear data set (Figs. 3B, 4A).

No morphological characters known from the revision of *Schizopetalon* (Al-Shehbaz, 1989) support the inclusion of *S. bipinnatifidum* in this subclade. *Schizopetalon bipinnatifidum* is distinct based on the unique presence of flower bracts along the raceme, recurved fruiting pedicels, and filiform cotyledons. In contrast to this, *S. biseriatum*, *S. rupestre*, and *S. tenuifolium* show partial or total absence of flower bracts, and have straight or tortulose flower bracts and incumbent cotyledons. Despite these differences, if the sister relationship between *S. bipinnatifidum* and the rest of the Atacama clade, as detected in the cpDNA BI analysis (Fig. 4B), reflects evolutionary history, it would have important biogeographic implications. *Schizopetalon bipinnatifidum* is a mid-elevation species (1200–2400 m) of the Atacama and Coquimbo Regions (Al-Shehbaz, 1989) adjacent to the coastal species (i.e., *S. tenuifolium* and *S.*

biseriatum) and the species from the Andes (*S. rupestre*; Fig. 2). The phylogenetic results indicate that these species share a recent common ancestor with *S. bipinnatifidum*, suggesting an origin of this group in mid-elevation transitional areas where *S. bipinnatifidum* occurs. However, because of the lack of corroboration from the nuclear data and support from the MP and ML analysis, this biogeographic scenario must be viewed as hypothetical.

The Matorral subclade comprises *S. dentatum, S. maritimum*, and *S. walkeri*. These species are confined to the semiarid and Mediterranean coast from Coquimbo to Valparaiso (29°–33° S) and the inter-montane area of the Andes and Santiago (32°–33° S) (Fig. 2). Compared to the geographic area of the Atacama subclade, this region is characterized by a milder climate with higher annual precipitation, which results in sclerophyllous Mediterranean vegetation (Luebert & Pliscoff, 2006). This group is not supported by any exclusive combination of morphological characters. All analyses of the nrDNA data set and the MP analysis of the cpDNA data set recover a highly supported clade of coastal species (i.e., *S. maritimum* and *S. walkeri*), and the coastal-montane *S. dentatum* (Figs. 3, 4A). The coastal species are defined by their straight pedicels in fruit, filiform-twisted cotyledons, and globose seeds, compared to *S. dentatum*, which has recurved fruiting pedicels, oblong straight cotyledons, and oblong-plump seeds.

The monophyly of *Schizopetalon arcuatum* was supported in all analyses. This is noteworthy because, despite its distinctive morphology, *S. arcuatum* is often misidentified as *S. biseriatum* and *S. maritimum* in herbarium material (Al-Shehbaz, 1989). *Schizopetalon arcuatum* is distinguished by the presence of strongly arcuate fruits and infloresences that are bracteate throughout (Al-Shehbaz, 1989). Prior to the present study, *S. arcuatum* was known only from its

type locality. The additional populations included in this study were discovered during fieldwork in 2010. The strong molecular support for the monophyly of the *S. arcuatum* clade is consistent with field observations of morphology (Figs. 3, 4). The location of the specimens identified as *S. arcuatum* reveals a distribution (28°– 29° S) that covers no more than 100 km from the coastal bluffs north of Huasco to the sandy gullies of Chañarcillo (Fig. 2).

Such a distribution illustrates the occurrence of this species across a transition zone between the distribution of the coastal species of the Atacama and Matorral subclades. The chloroplast trees (Fig. 4) suggest a sister relationship of *S. arcuatum* with the Atacama subclade, which is supported also by the shared presence of ovate-oblong seeds with incumbent cotyledons. However, this relationship was not supported by analyses of the nuclear data (Fig. 3). In the nuclear trees resulting from analyses that included indel characters (trees not shown), a weakly supported sister relationship is suggested between *S. arcuatum* and the Atacama + Matorral subclades. While past hybridization may explain this incongruence, the results of the present study cannot rule out alternative hypotheses for the relationships of *S. arcuatum* and the rest of the subclades.

Another important result from the analysis of relationships within *Schizopetalon* is the recovery as closest relatives of two morphologically similar samples of *S. brachycarpum*. This was supported by indels in the chloroplast data set (Fig. 4). This species is only known from its type specimen (*Schlegel 926* at CONC) and one additional field-collected specimen (*Toro 56* at KANU). Both were collected in a small area near Santiago de Chile (70° W, 33° S) at more than 2000 m above sea level (Fig. 2). In addition to the reported results, the high morphological

similarity of the two specimens in critical characters (e.g., narrowly oblong fruits with no more than 4 or 6 seeds per locule) suggests that the species status of *S. brachycarpum* should be retained. However, considering the low number of known specimens, the inclusion of additional populations would be needed to confirm this hypothesis.

The two different hard topological incongruencies between the nuclear and the chloroplast data sets (*S. brachycarpum* and one population from Los Molles of *S. bipinnatifidum*; Figs. 3, 4) might suggest hybridization between species from the Atacama and Matorral subclades and/or incomplete lineage sorting. However, similar to the cases reported for *M. incana* and *S. arcuatum*, more exhaustive analyses are required to determine the nature of these incongruencies.

Acknowledgements

The authors thank the curators of CONC, GH, MO, and OS for their permission to study and sample DNA from their collections, and Francisco Lira-Cuadra, Jaime Toro and Tomas Flores-Fernandez for assistance in fieldwork. We kindly thank Emily Wood (GH), Alicia Marticorena, and Roberto Rodriguez (CONC) for granting access to the type specimens of *Mathewsia collina*, *M. densifolia* and *Schizopetalon brachycarpum* for DNA extraction; and Daniel J. Crawford and two anonymous reviewers for their thoughtful comments on earlier versions of this manuscript. This project was funded with the support of the Fratcher Botany Fellowship Scholarship (KU) and Sigma Xi Grant-in-Aid for Research Award. OTN also thanks the Fulbright- CONICYT Equal Opportunities Scholarship that supported his doctoral studies at KU and made this project possible.

Table 1. Statistics from the phylogenetic analyses of all regions and data sets. Indel scores are given in parentheses. AICc, Akaike information criterion corrected by size.

	ITS	ETS	*rps16*	*trnH-psbA*	*atpH-atpI*	*rps16-trnQ*	cpDNA	nrDNA
Total (+ Indels)	724 (743)	588 (602)	1092 (1125)	472 (527)	670 (699)	576 (596)	2810 (2946)	1312 (1345)
Constant characters	523	389	978	309	583	501	2371	912
Variable characters	201(220)	199 (213)	114 (147)	163 (208)	87 (116)	75 (95)	439 (575)	400 (432)
Informative characters	104 (111)	107 (115)	38 (56)	71 (92)	37 (52)	37 (46)	183 (245)	211 (226)
Percentage	14.36% (14.94%)	18.20% (19.10%)	3.48% (4.98%)	15.04% (17.46%)	5.52% (7.44%)	6.42% (7.72%)	6.51% (8.31%)	16.08% (16.80%)
Constant ingroup	652	461	1033	385	631	542	2591	1113
Variable ingroup	72	127	59	87	39	34	219	199
Informative ingroup	58	59	26	48	26	31	131	117
Percentage ingroup	8.01	10.03	2.38	10.17	3.88	5.38	4.66	8.92
Total indels	19	14	33	55	29	19	138	33
Ingroup indel informative	5	5	15	16	15	7	53	10
MPT							10 (4)	3 (3)
Tree length							601 (812)	565 (619)
CI							0.813 (0.766)	0.829 (0.832)
RI							0.823 (0.788)	0.890 (0.893)
Likelihood							-ln 7193.2460	-ln 4834.6911
Modeltest (AICc)	TVMef+G	TrN+G	TPM1uf+G	F81+G	TVM+G	TPM1uf+G		

Figures

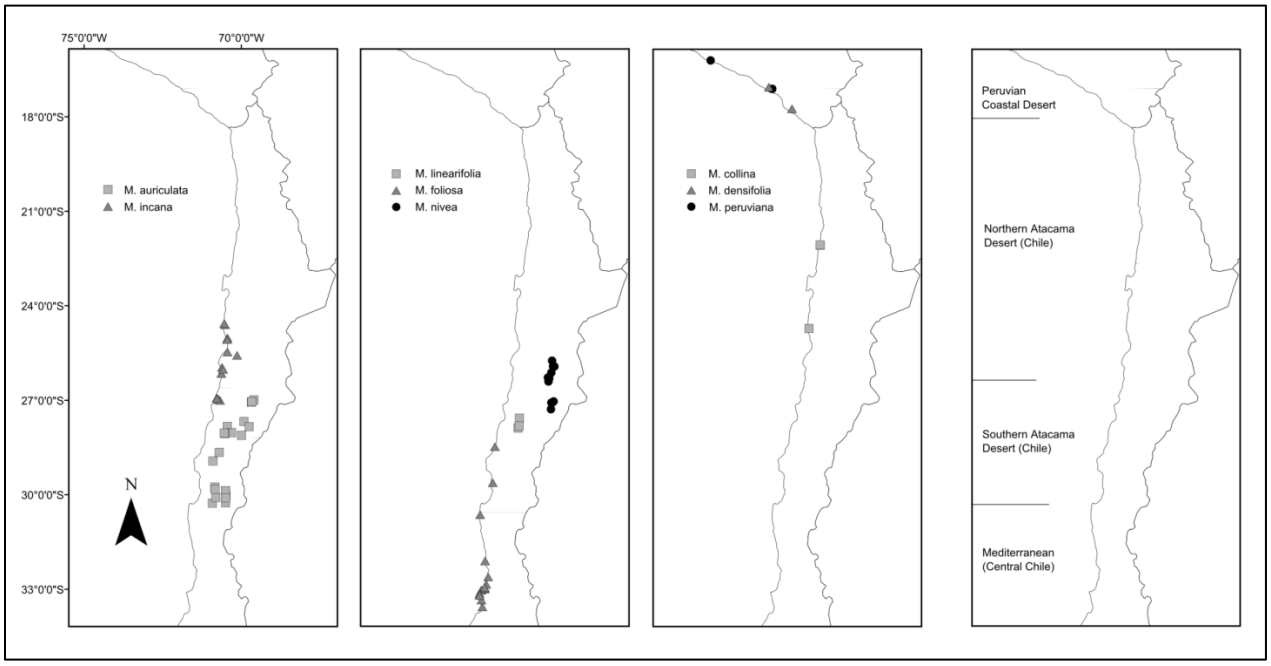


Fig. 1. Distribution of species of *Mathewsia* and the ecoregions of *Mathewsia* and *Schizopetalon* referred to in this study.

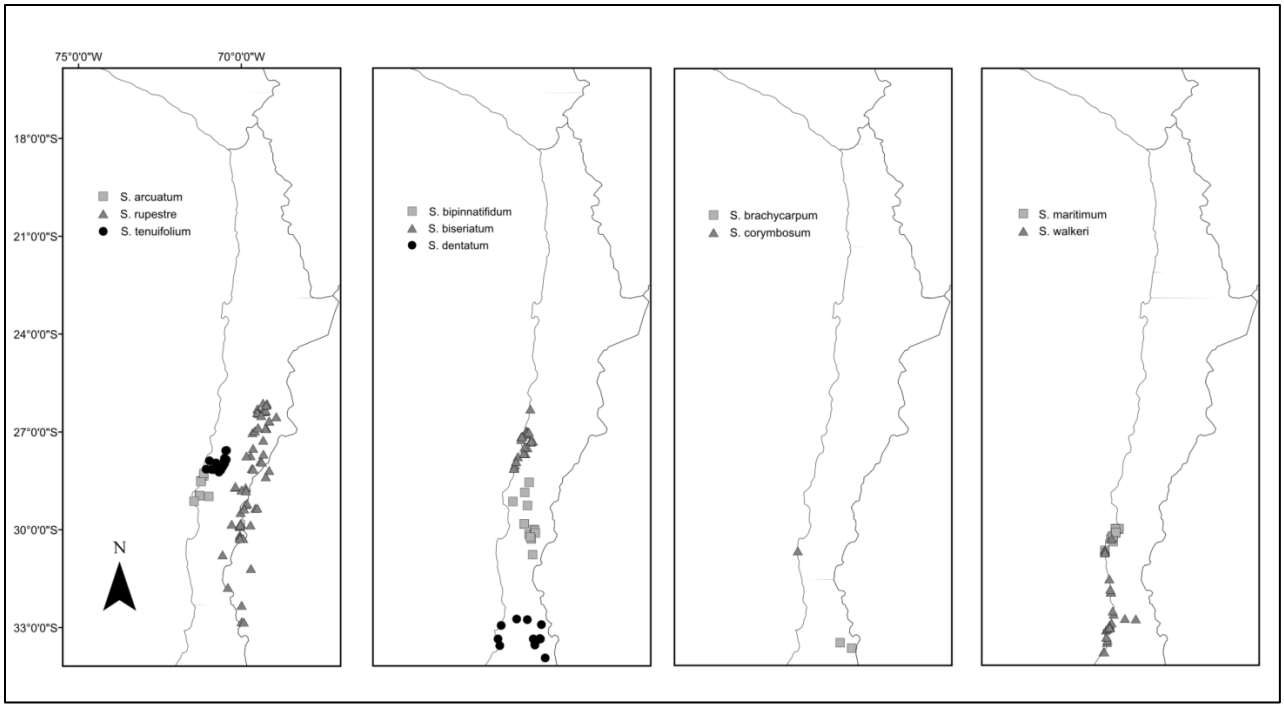
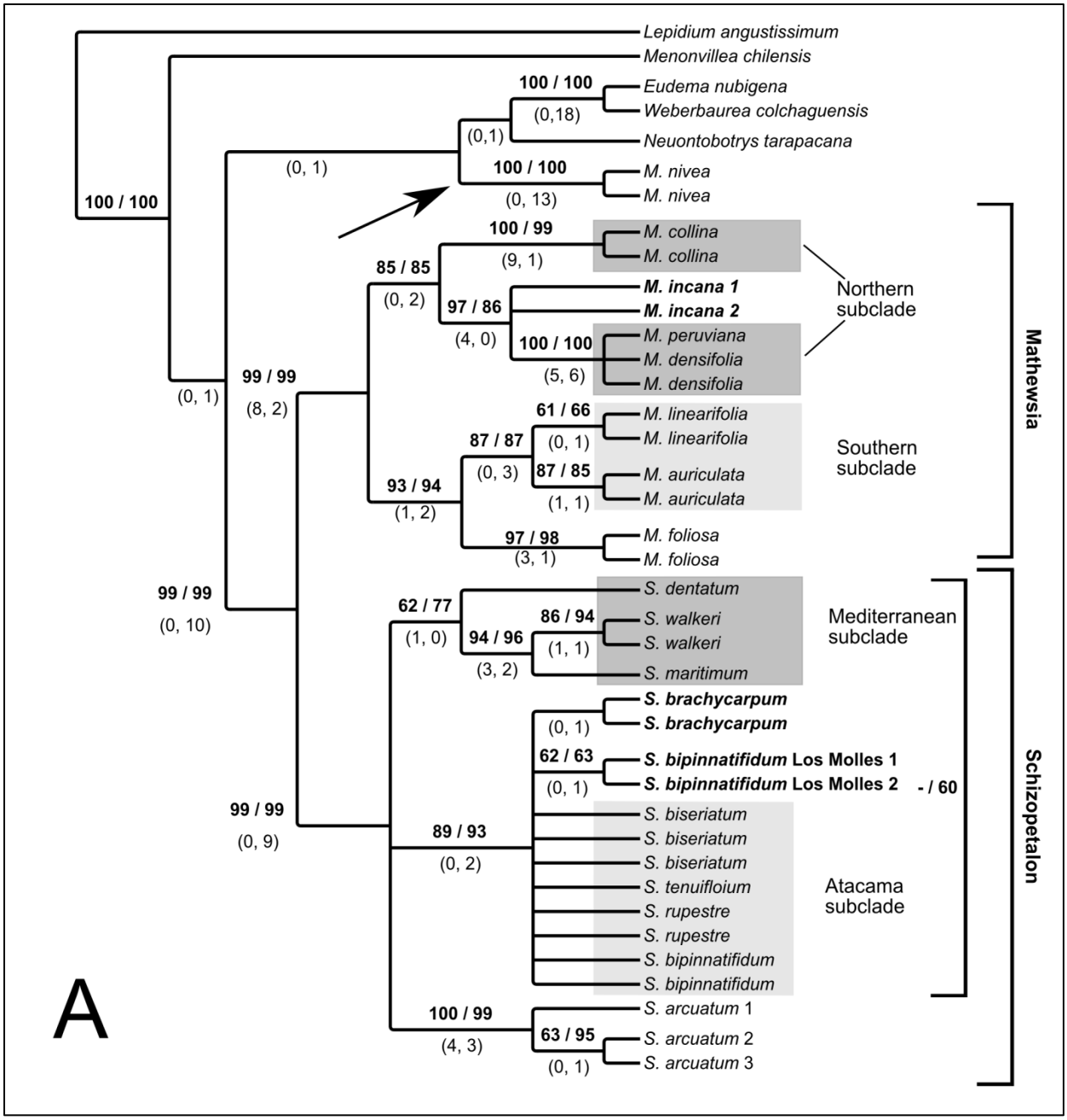


Fig. 2. Distribution of species of *Schizopetalon*. For ecoregions, see Fig. 1


Lepidium angustissimum
Menonvillea chilensis
100 / 100
Eudema nubigena
(0,18)
Weberbaurea colchaguensis
(0,1)
Neuontobotrys tarapacana
(0, 1)
100 / 100
M. nivea
100 / 100
(0, 13)
M. nivea
100 / 100
100 / 99
(9, 1)
M. collina
M. collina
85 / 85
(0, 2)
M. incana 1
97 / 86
M. incana 2
Northern subclade
(4, 0)
100 / 100
M. peruviana
(5, 6)
M. densifolia
M. densifolia
(0, 1)
99 / 99
(8, 2)
61 / 66
M. linearifolia
(0, 1)
M. linearifolia
87 / 87
(0, 3)
87 / 85
M. auriculata
Southern subclade
93 / 94
(1, 1)
M. auriculata
(1, 2)
97 / 98
M. foliosa
(3, 1)
M. foliosa
Mathewsia
99 / 99
(0, 10)
S. dentatum
62 / 77
(1, 0)
86 / 94
S. walkeri
Mediterranean subclade
94 / 96
(1, 1)
S. walkeri
(3, 2)
S. maritimum
S. brachycarpum
(0, 1)
S. brachycarpum
62 / 63
S. bipinnatifidum Los Molles 1
(0, 1)
S. bipinnatifidum Los Molles 2
- / 60
99 / 99
(0, 9)
S. biseriatum
89 / 93
S. biseriatum
(0, 2)
S. biseriatum
S. tenuifloium
Atacama subclade
S. rupestre
S. rupestre
S. bipinnatifidum
S. bipinnatifidum
Schizopetalon
A
100 / 99
S. arcuatum 1
(4, 3)
63 / 95
S. arcuatum 2
(0, 1)
S. arcuatum 3

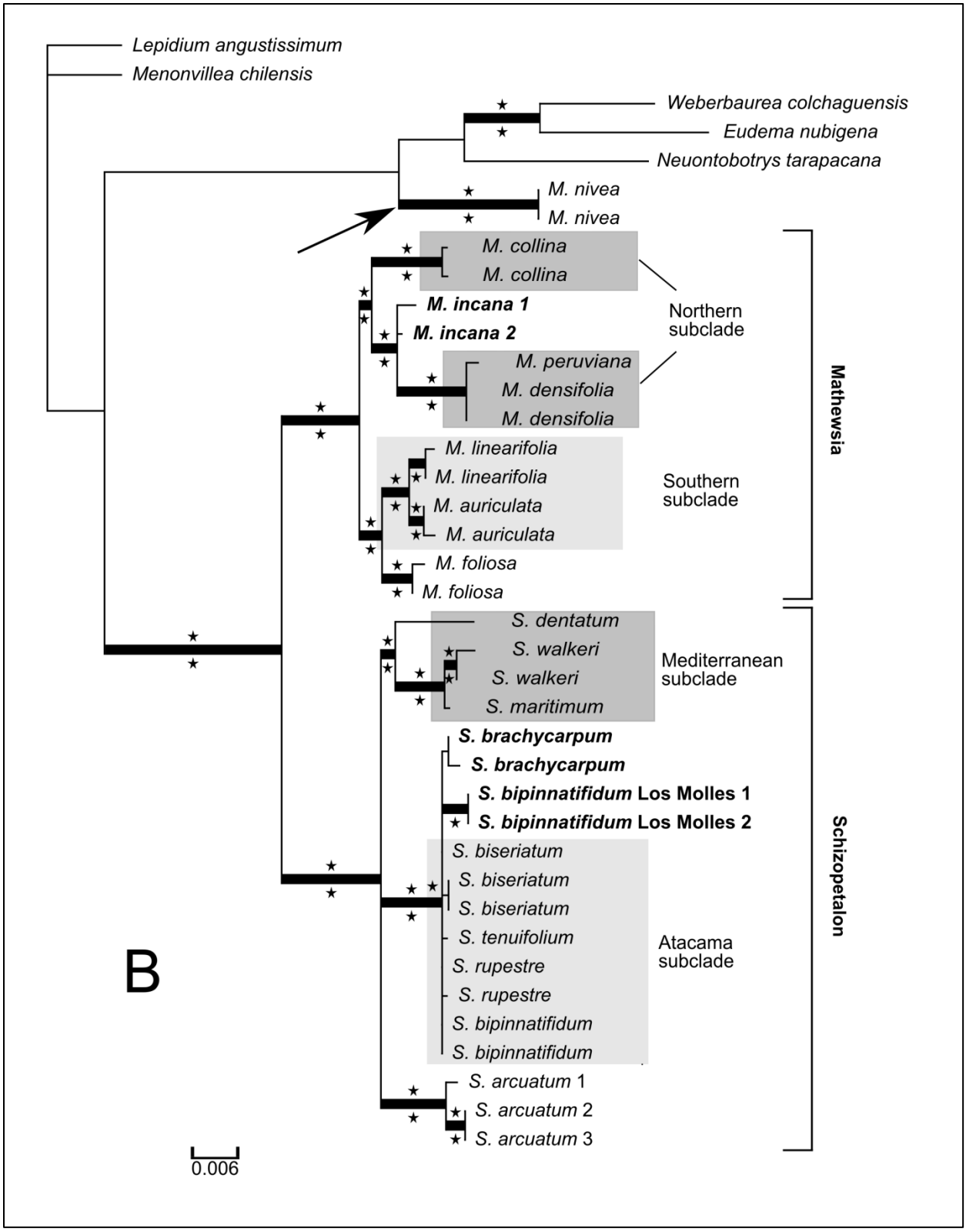
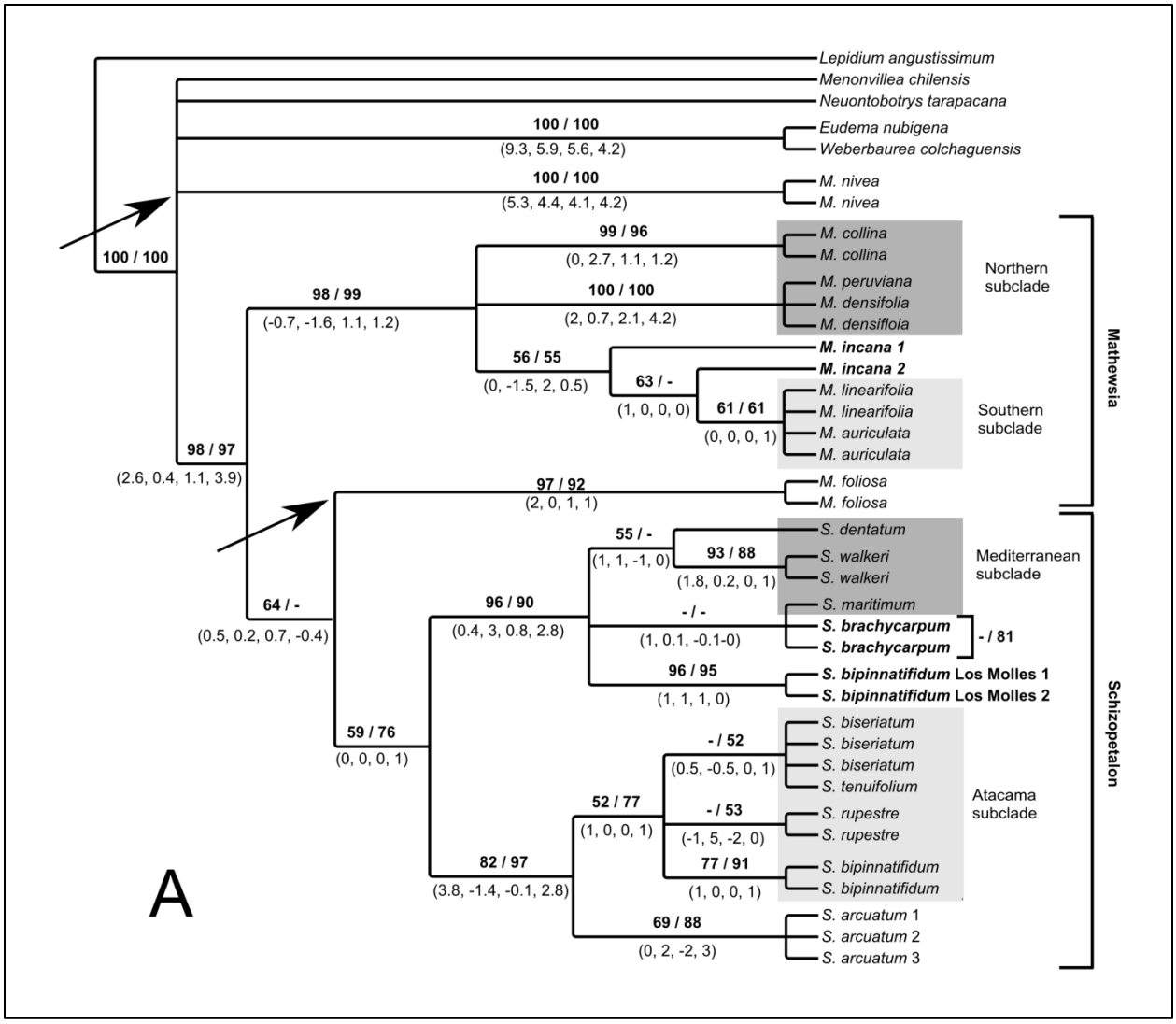


Fig. 3 A, B. Trees retrieved from the analysis of the nrDNA data set. Taxa in bold show strong

topological incongruence between the nrDNA and cpDNA data sets. The arrow indicates the position of *M. nivea* in the estimated trees and boxes illustrate the subclades found. **A,** MP consensus tree from three most parsimonious trees of 565 steps each (CI = 0.829, RI = 0.890). Bootstrap values are indicated in bold above branches (data set alone / indels included). Partitioned Bremer support values are shown below branches in parenthesis (ETS, ITS). **B,** Most likely tree retrieved with ML (−ln 4834.6911). Branches with high support are represented in bold and the source of support is shown with stars above (ML bootstrap) and below (BI posterior probabilities) branches. For convenience of presentation, the length of the branches of the outgroup taxa was reduced.


Lepidium angustissimum
Menonvillea chilensis
Neuontobotrys tarapacana
100 / 100
(9.3, 5.9, 5.6, 4.2)
Eudema nubigena
Weberbaurea colchaguensis
100 / 100
(5.3, 4.4, 4.1, 4.2)
M. nivea
M. nivea
99 / 96
(0, 2.7, 1.1, 1.2)
M. collina
M. collina
Northern subclade
100 / 100
98 / 99
(-0.7, -1.6, 1.1, 1.2)
100 / 100
(2, 0.7, 2.1, 4.2)
M. peruviana
M. densifolia
M. densifloia
Mathewsia
56 / 55
(0, -1.5, 2, 0.5)
M. incana 1
M. incana 2
63 / -
(1, 0, 0, 0)
61 / 61
(0, 0, 0, 1)
M. linearifolia
M. linearifolia
M. auriculata
M. auriculata
Southern subclade
98 / 97
(2.6, 0.4, 1.1, 3.9)
97 / 92
(2, 0, 1, 1)
M. foliosa
M. foliosa
55 / -
(1, 1, -1, 0)
93 / 88
(1.8, 0.2, 0, 1)
S. dentatum
S. walkeri
S. walkeri
Mediterranean subclade
64 / -
(0.5, 0.2, 0.7, -0.4)
96 / 90
(0.4, 3, 0.8, 2.8)
- / -
(1, 0.1, -0.1-0)
S. maritimum
S. brachycarpum
S. brachycarpum
- / 81
96 / 95
(1, 1, 1, 0)
S. bipinnatifidum Los Molles 1
S. bipinnatifidum Los Molles 2
Schizopetalon
59 / 76
(0, 0, 0, 1)
- / 52
(0.5, -0.5, 0, 1)
S. biseriatum
S. biseriatum
S. biseriatum
S. tenuifolium
52 / 77
(1, 0, 0, 1)
- / 53
(-1, 5, -2, 0)
S. rupestre
S. rupestre
Atacama subclade
82 / 97
(3.8, -1.4, -0.1, 2.8)
77 / 91
(1, 0, 0, 1)
S. bipinnatifidum
S. bipinnatifidum
69 / 88
(0, 2, -2, 3)
S. arcuatum 1
S. arcuatum 2
S. arcuatum 3
A

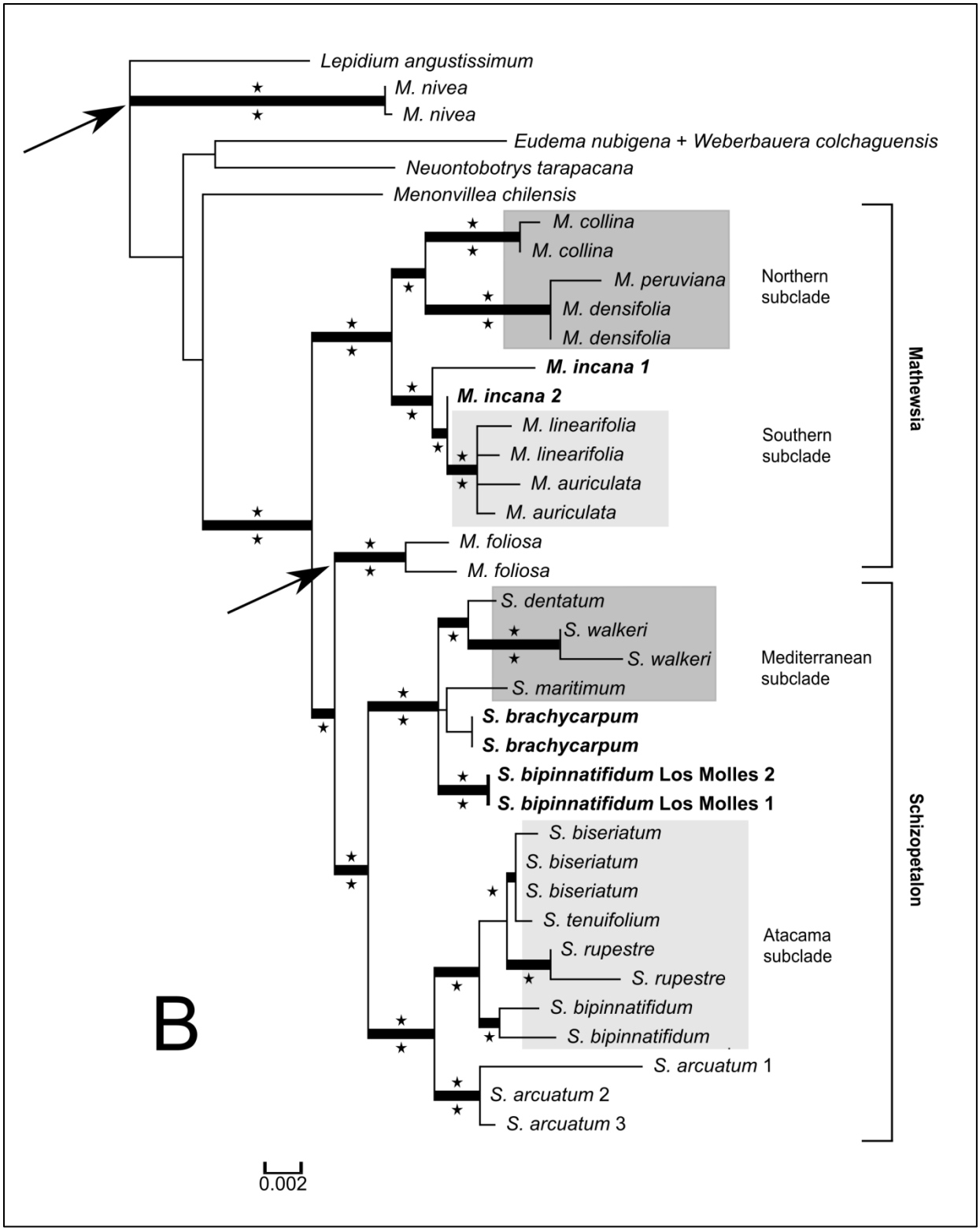


Fig. 4 A, B. Trees retrieved from the analysis of the cpDNA data set. Taxa in bold show strong topological incongruence between nrDNA and cpDNA. The arrows indicate the positions of *M.*

foliosa and *M. nivea* and boxes illustrate the subclades found. **A,** MP consensus tree from ten most parsimonious trees of 601 steps each (CI = 0.813, RI = 0.823). Bootstrap values are indicated in bold above branches (data set alone / indels included). Partitioned Bremer support values are shown below branches in parenthesis (*atpH*, *psbA*, *rps16*, *trnQ*). Additional clades retrieved from data set + indels and levels of bootstrap support are shown in brackets. **B,** most likely tree retrieved with ML (−ln 7193.2460). Branches with high support are represented in bold and the source of support is shown with stars above (ML bootstrap) and below (posterior probabilities) branches. For convenience of presentation, the length of the branches of the outgroup taxa was reduced.

CHAPTER THREE

UNRAVELING PATTERNS OF SPECIES DIVERGENCE IN THE ATACAMA SUBCLADE OF *SCHIZOPETALON* (BRASSICACEAE) USING MORPHOMETRIC, ECOLOGICAL, AND MOLECULAR APPROACHES

Abstract

Phylogenetic analyses of Tribe Schizopetaleae (Brassicaceae) resolve a well-supported clade of four species of *Schizopetalon* (i.e., *S. bipinnatifidum*, *S. biseriatum*, *S. rupestre*, and *S. tenuifolium*) that are endemic to the southern Atacama Desert and the Andes (i.e., the Atacama subclade). These species are allopatric and distributed across a range of environmental conditions, yet previous molecular and morphological studies provide only limited resolution among these taxa. This lack of divergence has made phylogenetic analyses and testing the limits of the species difficult, which precludes assessing the patterns of diversification or testing hypotheses that aridity may promote geographic isolation and genetic divergence among populations. Using a multifaceted approach, patterns of species differentiation were explored using morphometric, ecological niche modeling, and nuclear DNA sequence data. The latter data were analyzed with a general mixed Yule-coalescent (GMYC) model to identify shifts in rates from a Yule (speciation) to coalescent (population) process. Morphometric analyses only supported the differentiation of one of the Andean species (*S. rupestre*) from the rest of the Atacama species; however, ecological niche modeling and GMYC mostly support the current view that there are four recognizable species in the Atacama subclade. In addition, GMYC resolves a cluster, albeit with low support, that comprises haplotypes from multiple species, which suggests incomplete lineage sorting and/or past hybridization. The difference in the results obtained between morphometry and the rest of the data suggest that phenotypic plasticity,

hybridization, and incomplete lineage sorting may obscure the boundaries between these desert annuals.

Introduction

An important goal in the study of desert floras is to understand the role of climate in the process of diversification. Deserts are environmental mosaics, where the local temperature and levels and timing of precipitation can create an array of habitats over small geographic scales (Noy-Meir, 1973, 1985; Shmida, 1985). This property is capable of isolating lineages in habitats with different climatic conditions at local and regional scales (Axelrod, 1972; Stebbins, 1950, 1952). Given this pattern of isolation and putative local adaptation, it is possible that populations and species might accumulate differences in genotype and phenotype over a range of spatial and climatic conditions (e.g., Arafeh et al., 2002; Comes & Abbot, 1999; Nevo, Zohary, Brown, & Haber, 1979; Volis, Mendlinger, Turuspekov, & Esnazarov, 2002). Because of that, deserts represent ideal habitats to study local habitat differentiation, where climate is a dominate factor in the diversification of their floras (Gutterman, 2002; Whitford, 2002).

The Atacama Desert is considered one of the driest and oldest deserts in the world (Hartley et al., 2005; Houston & Hartley, 2003). The combination of diverse topography, environmental gradients of temperature and precipitation, and variability of precipitation creates a complex series of habitats over different spatial scales. These conditions affect the distribution of lineages, and can isolate and promote divergence among lineages, even at small spatial scales (Guerrero et al., 2011a; Guerrero et al., 2011b; Luebert, 2010). Such factors combined have led to the ideas that aridity might be important in promoting geographic isolation and divergence of endemic lineages (e.g., Arroyo et al., 1988; Gengler-Nowak, 2002; Luebert & Wen, 2008) and that local

climate can promote rapid diversification of the Atacama flora (Dillon et al., 2009; Gengler-Nowak, 2002; Heibl & Renner, 2012; Luebert & Wen, 2008). Nonetheless evidence for species divergence associated to these conditions is mostly lacking.

One approach to elucidate the factors promoting divergence in Atacama angiosperm lineages is to assess species limits by integrating multiple sources of data. The combined use of genetic, phenotypic, and ecological evidence might reveal congruent patterns that support climate-driven habitat isolation causing lineage divergence. In contrast, these same data might suggest other factors that explain the phylogeny and distribution of species (i.e., ecological filtering). The availability of GIS- (Elith & Leathwick, 2009; Peterson et al., 2011) related to lineage diversification based on ecological differentiation based data to trace environmental proxies provides the opportunity to evaluate species' environmental requirements –i.e. ecological niche- (e.g., Evans, Smith, Flynn, & Donoghue, 2009; Nakazato, Warren, & Moyle, 2010; Warren, Glor, & Turelli, 2008; Wiens & Graham, 2005). Similarly, given the influence of environmental variation on the phenotypic response in plants (Clausen, Keck, & Hisley, 1947; Pigliucci, 2005; Schlichting, 1986), patterns of quantitative morphometric differentiation for desert plants (i.e., leaves, shoots, etc.) might reflect local climatic variation in natural populations. In the Atacama Desert, if climate is a strong predictor of species divergence, a reasonable expectation is that trends of ecological, morphometric, and genetic differentiation should show similar patterns of divergence. Using a small lineage of Atacama-endemic mustards, a multifaceted systematic study was initiated to examine the limits of species under patterns of morphometric differentiation, genetic cohesion, and the role of ecological niche differentiation in the evolution of these taxa

Previous phylogenetic analyses of Tribe Schizopetaleae (Brassicaceae) resolve a well-supported clade of four species of annual plants that are endemic to the southern Atacama Desert (Toro-Núñez, Mort, Ruiz-Ponce, & Al-Shehbaz, 2013). This lineage (i.e., the Atacama subclade) comprises species that are distributed across a range of latitudes and altitudes. The species occur in different environmental conditions, ranging from sandy, coastal landscapes up to 800 m (e.g., *S. biseriatum* and *S. tenuifolium*) to ~ 4000 m across the ranges of intermontane and the Andes cordilleras (e.g., *S. bipinnatifidum* and *S. rupestre*). Although the four species, as currently recognized, are geographically isolated, previous studies reveal very low levels of interspecific differentiation in molecular (Toro-Núñez et al., 2013) and morphological characters (Al-Shehbaz, 1989). Thus additional analyses using more variable molecular markers are required to examine the patterns of morphological variation and ecological differentiation in this lineage.

Using the Atacama subclade as a focal group, the objectives of this study are to: 1) characterize patterns of morphometric divergence among the species, 2) infer and compare the ecological niche of the Atacama species, 3) examine species limits by conducting maximum likelihood (ML) analyses to infer species limits under a Yule (speciation) to coalescent (population) process, and 4) integrate the results of these lines of evidence to understand patterns of evolution in this lineage.

Materials and Methods

Morphological data acquisition and quality assessment

A total of 165 individuals representing the four subclade species were collected and deposited in KANU prior to morphometric analyses. Additionally, 403 individuals from herbarium specimens

obtained from CONC, OS, MO, UC and GH were added to the study. In total, the specimens come from 143 collection localities, which represent well the geographic range and morphological variation of the species of the Atacama subclade. For the present study, the current taxonomy of *Schizopetalon* was used to identify the sampled individuals (Al-Shehbaz, 1989). From these specimens, four linear measurements (obtained from all specimens) and four shape-derived measurements from leaf blades (obtained from 409 specimens) were recorded (Table 1). A high resolution digital photograph was captured for each individual (RAW format, 5200 x 5200 pixels) using a Nikon D90 with a 18 — 108 mm lens for linear measurements and a 60 mm macro lens for leaves. When available, two cauline or basal leaves were collected per individual; these leaves were rehydrated in distillated water and stretched prior to being photographed. Leaf blade photographs were transformed to simplified silhouettes to collect area, perimeter, and bounding rectangle measurements as surrogate of length and width (Fig. 2). All measurements, transformations, and calibrations of images were conducted using ImageJ v. 1.45s (Abramoff, Maghalaes, & Ram, 2004).

Linear and leaf blade shape measurements were averaged with the arithmetic mean per individual when two or more measurements per structure were available and log transformed for correction. Since allometric/isometric bias (i.e., the influence of body size on the shape/size of organs) can be a confounding factor for comparisons among individuals with attributes influenced by non-climatic factors, (e.g., age, reproductive stage, differential allocation of resources, etc.; Weiner, 2004), size corrections were applied before conducting morphometric comparisons. Stem and inflorescence internode lengths and leaf area and perimeter measurements were significantly correlated ($r > 0.5$, $P > 0.05$) with plant size and leaf length, respectively. To control for this effect, a linear regression against plant size and leaf length was

conducted on the characters, and their residuals were used as surrogates of size corrected measurements in the subsequent analyses.

Because of the complex pattern of leaf dissection among species of the Atacama subclade, an index of dissection based on the total perimeter over the square root of the area (Kinkaid & Schneider, 1983; McLellan, 1993; McLellan & Endler, 1998) was employed. This dimensionless index gives a simple estimate of how much more complex the outline of the leaf is compared to an idealized circle, thereby providing a reasonable estimate of the level of leaf dissection (McLellan & Endler, 1998). This estimate was favored over other descriptive techniques for leaf shape (landmark geometry, Fourier elliptics, etc; see Cope et al., 2012) due to the high level of variation among individuals and complexity detected for comparable interpretations in preliminary analyses. Finally, every uncorrected and corrected measurement was tested and confirmed for a nearly normal distribution using an Andersen-Darling Normality test as implemented in MINITAB v. 16.2.3 (MINITAB Inc., State College, PA, USA).

Interspecific morphometric comparison

To detect patterns of morphometric variation across species, a combination of multivariate and univariate analyses were conducted. Multivariate analyses using a principal component analysis (PCA) were conducted under a variance/covariance association matrix in order to determine if the ordination generated in the multivariate space represents concerted differences from the sum of the characters among species. A data set comprised of only individuals with leaf shape and linear measurements was used for this analysis, since fewer individuals were measured for leaf characters. Comparisons between the original and reduced data set in linear measurements

suggest no influence in the use of this reduced data set (data not shown). In order to evaluate the multivariate trend of variation in the data, a non-parametric one way MANOVA was used, since no homogeneity of variance and covariance and multivariate normality was achieved in the data set. This test is a multivariate analogue of Fisher's F-ratio, but it relies on symmetric distances or dissimilarity matrices to detect differences and permutations for establishing P-values (Anderson, 2001). This test was conducted using Mahlanobis distances and a posterior species pairwise comparison using a Bonferroni correction in linear and leaf shape characters. To detect individual character differences and trends among species, a Kruskal-Wallis test and a Dunn's posterior multiple comparison tests were conducted. Multiple comparisons were set in six pairwise species comparisons, at an individual Bonferroni correction of $P < 0.008$ using MINITAB with the script *Krusmc.MAC* (available in http://www.minitab.com/en-TW/support/macros/default.aspx). Because a potential bias can be attributed to the influence of the high number of individuals collected from one event in the data set (e.g., 2010 and 2011 field collected specimens, ~32% of the total of individuals), a Mann-Whitney U test was conducted to compare rank differences in every character between herbarium and field collected specimens. For these analyses, samples of *S. bipinnatifidum* were excluded since the paucity of field collected individuals (e.g., only five specimens). All multivariate analyses were conducted with PAST v. 2.15 (Hammer, Harper, & Ryan, 2001) and univariate analyses in MINITAB.

Species occurrence and environmental data

Species occurrence data was obtained from field collected specimens during 2010 and 2011 and herbarium records. These specimens were georeferenced to their nearest landmark referred to in their labels, using gazetteer information of geographic landmarks with Geolocate

(http://www.museum.tulane.edu/geolocate) or Google Earth. Their accuracy was mostly confirmed during fieldtrips in 2010 and 2011. Records with ambiguous or unconfirmed localities were removed (< 10 records total) and closely associated localities (< 1 km) were assigned to one coordinate only. As a result, a final data set of 114 unique confirmed coordinates was obtained (Appendix 2).

The estimate of the environmental suitability to infer the ecological niche for species of the Atacama subclade was conducted using presence-background estimates (Elith & Leathwick, 2009), using a combination of bioclimatic and remote sensed data. Fifteen bioclimatic variables (Appendix 3) were obtained from the Worldclim data base (Hijmans et al., 2005) at a spatial resolution of 30 arc seconds (~ 1 km). Since precipitation and temperature layers are expected to be redundant in the study area, climatic layers were reduced using a PCA based on standardized variables using *princomp* function in R. For subsequent analysis, the first six components were used, since they explained 99 % of their total data set variability (Appendix 3).

Previous studies have detected a lack of resolution and precision for estimating species' ranges from Worldclim database layers for the Atacama Desert (Luebert, 2010). For this reason, the Normalized Deviation Vegetation Index (NDVI) was also used to increase ENM precision. NDVI is an index based on remote detection of spectral measurements of red/infrared reflectance associated with the photosynthetic activity of a particular patch (Sellers et al., 1992). NDVI data sets were chosen since their proven utility as a surrogate of the precipitation and temperature conditions that trigger biological activity (e.g., seed germination) in other desert areas, such as Patagonia (e.g., Fabricante, Oesterheld, & Paruelo, 2009; Verón & Paruelo, 2010) and the

Atacama Desert itself (Nakazato et al., 2010). NDVI layers were constructed based on localities and collection time of each population record (August to September) in a series of 16 days during 2002 to 2011 from MODIS sensor on board the TERRA series of satellites (available at http://modis.gsfc.nasa.gov). Representative layers of the maximum and minimum mean values for rainy years (i.e., 2002, 2005, 2009, 2010, and 2011) and the maximum standard deviation values for the entire time-span of the data set were constructed for the ENM estimation. Categorization of rainy or dry years was determined using the mean precipitation values from climatic records in Chamonate - Copiapo station (27˚22'S - 70˚18'W; Dirección General de Aeronáutica Civil, Chile), which is close to the northernmost and driest boundary of the distribution of the study group. As a threshold value to identify rainy years, an estimate of 20 mm of mean year precipitation was used, based on previous germination experiments conducted in the area for *Schizopetalon* (Vidiella, Armesto, & J.R. Gutierrez, 1999). Manipulation of layers and PCA analysis was conducted in *R* using the package *raster* v. 2.031 (Hijmans & van Etten, 2012).

Ecological niche model estimation

The models of the species of the Atacama subclade were calibrated over two different, but overlapped, geographic training areas that hypothesize the potential distribution throughout the species life-span ("M" based on Peterson et al., 2011; Soberón & Peterson, 2005). The training area for *S. biseriatum*, *S. bipinnatifidum* and *S. tenuifolium* was defined using the presence of the highest altitude of the Andes as the eastern boundary, and covering the maximum known latitudinal extent (Atacama and Mediterranean areas; see Fig. 5) for the Chilean species of *Schizopetalon* (see Al-Shehbaz, 1989). A larger training area, which encompassed both the

Argentinean and Chilean sides of the Andes as the eastern boundary and limits from the absolute Desert (26° S) to the end of the Mediterranean vegetation (38° S), was considered for *S. rupestre*. Following Barve & al. (2011), this criterion of training range delimitation was applied because the risk of bias in the estimates of areas of occurrence given the notorious differences in the extent of the geographic distribution and potential for dispersal present between the coastal and Andean species.

Niche models of all species were inferred using maximum entropy method implemented in MaxEnt v. 3.3.3a (Phillips, Anderson, & R.E. Shapire, 2006). Models were estimated using the first six principal components of the bioclimatic variables and the three NDVI derived layers, using the default parameters in MaxEnt except for 30% of test points per species. The resulting models were converted to a binary prediction (i.e., suitable or unsuitable niche) using the minimum suitability value in the training data set (Phillips et al., 2006). This conservative approach was preferred over the usual raw outputs from MaxEnt, because the low probability values reported by this program can make it difficult and inaccurate to detect the overlap between areas possessing or lacking suitable conditions for each species (Peterson, Papeş, & Eaton, 2007).

Beyond estimating climatic niches based on inferred species distributions, the niche overlap between species pairs was explored using a test of background similarity (Warren et al., 2008). This test evaluates if the ecological niches of two species are more or less similar than would be expected by chance. This is achieved by testing if their ENM overlap scores (Hellinger's I and Shoener's D; Warren et al., 2008) are significantly different from a null distribution produced by

the randomized calculation of these indexes, generated from the common local environmental background where both species occur (Warren et al., 2008). For this purpose, null D and I distributions were generated from 200 niche estimates for each species using the same thresholds and parameters used in MaxEnt, but randomizing the sets of points selected from each species' background (Nakazato et al., 2010; Warren et al., 2008). The best estimates of ENM were reduced to binary data using the values of the minimum training presence as threshold and combined in training mask with the largest geographic extension (*S. rupestre*). A one tail Z-test was conducted to estimate the significance of the differences of D and I indexes from the null distribution for every species comparison. The process of randomization and calculation of niche overlap statistics was conducted using the package *phyloclim* v. 0.9-3 in R (Heibl & Calenge, 2013). Z-test statistics were conducted in *R* and model predictions were visualized using ARCMAP 9.3. (ESRI, Redlands, CA, USA).

Assessing species limits using GMYC

Delimitation of Atacama subclade species was conducted using a GMYC (general mixed Yule-coalescent; Pons et al., 2006) model on MMT nuclear haplotype data set (Toro-Núñez & al. unpublished; see Chapter 4). This analysis uses a ML approach on time-calibrated gene trees to delimit species with a single locus. GMYC estimates thresholds that identify shifts in branch rates from a Yule (speciation) to coalescent (population) process (Esselstyn et al., 2012; Pons et al., 2006). Both single and multiple threshold hypotheses are analyzed against a null model (i.e., only one species present) using a likelihood ratio test to determine possible shift point onto ultrametric trees (Esselstyn et al., 2012; Fujisawa & Barraclough, 2013; Pons et al., 2006).

The presence of multiple identical haplotypes, especially from homozygous biallelic parental haplotypes, can bias the estimates of thresholds and cause over-partitioning of the delimited groups (Barley, White, Diesmos, & Brown, 2013; Fujisawa & Barraclough, 2013; Reid & Carstens, 2012). To reduce this problem, identical haplotypes in the MMT data set were reduced to a single terminal using DNAsp v. 5.1 and phylogenetic inference was conducted only on unique haplotypes. Using Bayesian inference of phylogeny with BEAST v. 1.7 (Drummond, Suchard, Xie, & Rambaut, 2012), ultrametric trees were obtained using a strict clock and a constant coalescent growth model. A constraint on the molecular rate clock of 0.448 – 0.986 substitutions/site/year x 10^{-8} was employed in a lognormal prior distribution, based on average estimates from nuclear loci in *Arabidopsis* species (Huang et al., 2012). Coalescent sampling was conducted using four MCMC runs of 10 million replicates and sampling every 1000 iterations. Runs were analyzed using Tracer v. 1.5 (Rambaut & Drummond, 2007) after discarding the first 1000 trees as burn-in and convergence of posterior splits probabilities was confirmed among runs with AWTY (Are we there yet?; Nylander et al., 2008). Retrieved trees were summarized as a maximum clade credibility tree using median node ages, and branch lengths were placed as substitutions per site.

The GMYC model was used to delimit the species using the maximum credibility tree as the ultrametric input tree. Prior to the analysis, a haplotype sequence of *S. walkeri* was designated as the outgroup of the Atacama subclade (Toro-Núñez et al., 2013). Delimited species were considered only when both monophyletic groups from phylogenetic and GMYC analyses concurred. The GMYC analysis was conducted using the R (R Development Core Team, 2013) package SPLITS (Ezard, Fujisawa, & Barraclough, 2013), employing the default settings on both

single- and multiple-threshold models.

Results

Morphometric variability

PCA analysis (Appendix 4) on the reduced data set using the combination of linear and leaf characters distinguished only two clusters. The scatterplots for combinations of the first, second and third component revealed overlap among all species, except for *S. rupestre* (Fig. 3). The separation of this species from the rest is mostly explained by the first component, which primarily reflects variation of the area in leaf blades (0.7549). The other two remaining components represent variation in stem internodes (0.7761) and inflorescence internodes (0.812), respectively. The non-parametric MANOVA suggested significant differences in the data set ($F = 22.87$, $P < 0.001$) and significant differences in all the pairwise species comparisons (all $P < 0.001$). Similar results were obtained for the data set of linear measurements ($F = 20.02$, $P < 0.001$) and the leaf data set ($F = 35.62$, $P < 0.001$). Among the pairwise comparisons, all combinations were significant ($P < 0.001$), except for *S. bipinnatifidum – S. tenuifolium* based on leaf shape ($P = 0.08$).

For the univariate comparisons, the Kruskal – Wallis test showed significant differences ($P < 0.001$) in every comparison across characters (Table 2; Fig. 4). Dunn's post-hoc test suggested random differences in most of the pairwise species comparisons. Nevertheless, there were a few patterns of character differentiation. For linear characters, differences in sepal length ($H = 59.26$, $P < 0.008$) were detected in almost every species comparison, except for *S. bipinnatifidum* - *S. rupestre* ($Z = 1.57$, $P > 0.06$). The remaining characters showed differences in at least two

species combinations, but without revealing congruent differences among characters. For leaf characters, as suggested in the multivariate analyses, the area of the leaves provides significant differences only where *S. rupestre* was involved. Similar to the non-parametric MANOVA, all other leaf characters revealed differences in every combination, except for *S. tenuifolium* - *S. bipinnatifidum* (Table 2).

When the data were analyzed to elucidate congruence associated with the origin of the specimens (i.e., field only and herbarium only), similar results were found in most of the character comparisons (Kruskal-Wallis, $P < 0.005$). However, sepal length comparisons between field-collected ($H_{sepals} = 1.29$, $P > 0.35$) and herbarium material ($H_{sepals} = 54.01$, $P < 0.001$) were significantly different. For the Mann Whitney test, almost every character revealed significant differences between field and herbarium specimens per species, except the length/width ratio and the dissection index in the leaves (Table 3).

Ecological Niche Modeling

All inferred niche models had high predictability expressed by the AUC values in the training and test data (AUC > 0.95; Table 4) and successfully estimated species distributions. Predicting the distribution of *S. biseriatum* was problematic due to excessively lumped geographic localities that resulted in spatial autocorrelation. To correct for this problem, the occurrence points were rarefied by a lag distance of ~ 12 km to relax this spatial bias. MaxEnt jackknife scores suggest that the coastal (i.e., *S. biseriatum* and *S. tenuifolium*) and intermontane species (i.e., *S. bipinnatifidum*) had a high predictive contribution (30-60%) to model their distributions from the first principal component (Table 4). This vector is mostly explained by the moderate (~30%)

contribution of almost all precipitation estimates plus annual temperature range and seasonality (see Supplement 2). The minimal and maximal mean estimates of NDVI had the second highest contribution (36.7%) in *S. tenuifolium* and *S. bipinnatifidum* (36.7%). In contrast, the distribution model for *S. rupestre* received a greater contribution from the first (32.6%), fourth (28.1) and second (17%) principal components. Similar to the other set of principal components, the first and second vectors were also explained by the moderate contribution of almost all variables. The fourth and second factors explained the variability of precipitation seasonality (Bio15). In this case, no significant contribution was detected by any of the NDVI layers (Table 4).

A visual examination suggests differences in the size and location of the inferred suitable areas of species distribution. First, areas of occurrence were suggested to be highly localized in coastal species compared to montane species; *S. rupestre* has the largest range and *S. tenuifolium* has the smallest, with differences of more than 16 – fold between them. There is very little overlap among the distributions of the species and visual inspection reveals that overlap is limited to the boundaries of their distributions (Fig. 5) with limited contact between species. In terms of their ecological niches, marginal overlap was detected, with I and D indexes with magnitudes below 0.1 (Table 4). When compared to the null distribution of the background similarity test, all species pairs had highly significant evidence ($P < 0.0001$; Table 5) to reject the hypothesis that the ecological niche of the species of the Atacama subclade was more similar than random for every species combination treated.

Assessing species limits using GMYC

Phylogenetic analysis of unique haplotypes resulted in the same five groups that were previously

recovered with distance, parsimony network and ML analyses (see Chapter 4; Fig. 7). Single and multiple thresholds from GMYC suggest the presence of multiple species, as both models were favored over the null model ($P<0.005$) of only one species (only coalescent rates; Appendix 5). The single-threshold inferred supports the existence of five species (groups A, B, C, D, and E; Fig 7). The multiple-threshold model suggested the presence of eight species. Both models suggested a common threshold for delimitation, whereas the multiple-threshold model suggested at least two additional potential thresholds within groups D and E (Fig. 7). Regardless of the threshold model analyzed, groups A, B, and C were supported with posterior probabilities greater than 0.95 by GMYC; group D was recovered with lower than 0.95.

Discussion

Are species of the Atacama clade distinguished by morphology?

Morphometric analyses provide an effective source of information to detect differentiation in plants, especially when climate plays a role in differentiation. For desert plants, morphometric variability can provide insights into the influence of climate on diversification, because it can function as a proxy of the local environment, habitat adaptation, and evolutionary divergence (e.g., Arafeh et al., 2002; Ehlringer, 1988; Housman, Price, & Redak, 2002; Knight & Ackerly, 2003). In the context of the Atacama Desert, the Atacama subclade of *Schizopetalon* provides an opportunity to examine patterns of morphological variation under the assumption that differences in local climatic conditions and resulting habitat isolation can result in lineage divergence.

A total of eight morphometric characters (Table 1), four linear and four shape-derived, were measured for a total of 568 specimens of the Atacama subclade of *Schizopetalon*. The sampled

individuals are from 143 collection localities representing well the distribution of these four species. Multivariate PCA analyses from a reduced data set including only leaf shape and linear measurements recover two clusters (Fig. 3), one representing *S. rupstre* and the other including individuals from *S. biseriatum*, *S. tenuifolium*, and *S. bipinnatifidum*. The high degree of overlap in morphometric characters for the latter three taxa is not completely surprising. As Al-Shehbaz (1989) noted, the lower elevation species in the Atacama subclade are difficult to differentiate based on vegetative morphology alone.

Although the morphological features analyzed here do not provide separation among three of the species, these data can still provide insights about the process(es) of divergence in the Atacama subclade. For example, there are discernible levels of morphometric variation in leaf characters, such as area, perimeter, length/width, and index of dissection in both the univariate and multivariate analysis (Fig. 3; Table 2). The analyses of index of dissection discriminate well *S. rupestre*, which is the only species in the Atacama subclade that has reduced leaf margins; the three lower elevation species (*S. biseriatum*, *S. tenuifolium*, and *S. bipinnatifidum*) have complex patterns of leaf dissection and are more difficult to discriminate (Fig. 6).

Among all leaf measurement, only differences in leaf area resulted in statistically significant interspecific differentiation (Table 2 & 3). A plausible explanation for this trend is that desiccation directly influences the area of leaves and smaller leaves reduce water lost via transpiration (De Micco & Aronne, 2012). It is possible that the similarity in leaf area among the lower altitude species is related to their occurrence in habitats that have similar levels of water availability compared to the high Andes. This might also explain why *S. rupestre* leaves with

larger areas are the only distinctive ones in this subclade, as this species is confined to high altitudes (> 2700 m) where lower temperatures reduce transpiration compared to lower altitude arid habitats.

The limited morphometric variability in linear and leaf characters might be explained by a combination of factors. The most likely is the influence of stochastic climate and its potential effects on the phenotypic plasticity of the analyzed characters. Annual desert plants are often highly plastic in morphology in response to stochastic environments (e.g., Angert, Horst, Huxman, & Venable, 2010; Magyar, Kun, Oborny, & Stuefer, 2007) and it is possible that the overlapping morphology between recognized species is the result of phenotypic plasticity. An initial test of this idea is possible by comparing morphometric variation temporally. The individuals sampled for the present study were derived from both field-collected and herbarium specimens (165 and 403 individuals, respectively). When the data were partitioned into field versus herbarium individuals, significant differences are noted based on the origin of the specimens. An example of this effect could be noted in specimens of *S. biseriatum*, which displayed significant differences in almost every character to the point of altering the results of the statistical analyses (e.g., sepal length; Table 3). Assuming that the collection and pressing of the specimens did not alter morphology, these differences can be attributed to climatic conditions during the year of collecting. Given that the collection year (2010) was particularly rainy and humid, it is likely that the favorable conditions during sampling resulted in the measurements falling outside the range of variation noted for specimens measured from collections made during more arid growing seasons.

Do species of the Atacama clade have divergent ecological niches?

The climate of the Atacama Desert is extreme and dynamic and is a major factor in shaping the distribution of species, creating geographic isolation, and promoting allopatric speciation (Guerrero et al., 2011a; Luebert, 2010). The results of the present study support this pattern of diversification for the Atacama subclade, given the differences detected in the ecological niches of all the analyzed species.

At first glance, the results from niche modeling revealed differences in the projected suitable localities for species (Fig. 4), which suggests niche differentiation. The predicted distributions for the coastal (*S. biseriatum* and *S. tenuifolium*) and montane (*S. bipinnatifidum* and *S. rupestre*) species were consistent with the habitats described by Al-Shehbaz (1989). This projection suggests ecological differentiation based on the differences in the climate of the coastal desert and Andes. The coastal area has stable climatic conditions dominated by extended periods of drought and temperatures that oscillate during diurnal and seasonal periods (Julia et al., 2008; Luebert & Pliscoff, 2006). In this region, rare winter precipitation provides the moisture needed for seed germination (e.g., Armesto et al., 1993; Gutiérrez, Arancio, & Jaksic, 2000; Jaksic, 1998; Vidiella & Armesto, 1989). In contrast, montane species are mostly affected by differences in seasonal temperature, which are influenced by an altitudinal gradient (Julia et al., 2008; Luebert & Pliscoff, 2006). Even though our analyses did not reveal a dominant factor in the prediction of ecological niche models of the analyzed species (Appendix 3), differences in the range of seasonal precipitation and temperature seem to play an important role in shaping species' distributions. Another relevant element detected in our analysis was the high contribution in the PCA of the maximum NDVI values for coastal species and *S. bipinnatifidum*

(Table 4). Unlike the interpolated climatic layers used for ecological niche modeling (Worldclim), NDVI offers greater accuracy for inferring distributions because it is a surrogate of vegetation presence. In other words, NDVI provides precise estimates of species occurrence because it discerns areas with vegetation present during years with contrasting patterns of precipitation (Fabricante et al., 2009; Verón & Paruelo, 2010).

Another result revealed from ENM analysis is the contrasting properties of the ecological niche and the range size between the species of the Atacama subclade. Our analyses confirmed that the coastal species occur in very restricted segments of the coast and inland dry plains of the southern Atacama Desert. In contrast, the montane species inhabit larger areas of the Andes over an extended latitudinal range (Fig. 4). Given that the area delimitation was conducted using a conservative approach (i.e., minimum training threshold), this result indicates that the ecological niche of the coastal species is confined to a much reduced portion of the Atacama Desert. In contrast, our analyses illustrate larger predicted occurrence of montane species in the Andes and adjacent areas. This result suggests that the climatic conditions at high altitudes in the Andes (> 2700 m) are more homogeneous and widespread, resulting in larger geographic ranges for optimal ecological niches in Andean species compared to their lower altitude sister species.

The results from ecological niche background similarity tests indicate significant divergence among the ecological niches of the Atacama subclade species of *Schizopetalon*. Niche overlap tests provide evidence that the ecological niches of all species analyzed are less similar than expected by chance ($P < 0.001$; Table 4). These differences were found in similar magnitudes for the index of similarity used (Shoener's D or I; Table 4). Overall, all analyses suggest that species'

niches are highly predetermined and different due to the local conditions of where they occur.

Are species of the Atacama subclade differentiated by genetic – coalescent methods?

GMYC analyses offer an approach to detect a shift from a Yule vs coalescent process on gene trees, which is useful for detecting the limits of species. Although this approach is not as well established as other multilocus approaches (Carstens, Pelletier, Reid, & Satler, 2013), there is a growing literature for using this method in a variety of lineages (e.g., Barley et al., 2013; Esselstyn et al., 2012; Fujisawa & Barraclough, 2013). In the current study, GMYC analyses were used in conjunction with results obtained from ecological niche data to examine the limits of taxa as well as patterns of diversification in the Atacama clade of *Schizopetalon*.

The GMYC analysis delimited five entities, which mostly reflect the currently accepted four species in the Atacama subclade. Two groups, representing northern and southern lineages, were delimited for *S. rupestre* (groups A and B; Fig 7). These geographic clades were not previously suggested based on analyses of morphology or phylogenetic analyses of DNA sequences (Al-Shehbaz, 1989; Toro-Núñez et al., 2013). The present results suggest genetic substructuring in *S. rupestre* and the possible existence of cryptic diversity. In contrast, the delimitation of *S. biseriatum* (group C) and *S. tenuifolium* (group D) resulted in more complex patterns in the GMYC analysis (Fig. 7). This is due to the presence of mixed haplotypes from coastal and Andean species in *S. tenuifolium* (group D; Fig 7), and in the outgroup (group E) which would suggest the presence of incomplete lineage sorting and/or hybridization (chapter 4).

The only species recognized by Al Shebaz (1989) that was not delimited by GMYC analyses was

S. bipinnatifidum; this species grouped with the southern clade of *S. rupestre* (group B). Although this result might suggest similar patterns of incomplete lineage sorting and hybridization found in the delimitation of *S. tenuifolium* (group D), this conclusion should be considered only tentative until additional analyses with improved sampling of *S. bipinnatifidum* is conducted (see chapter 4).

Can habitat isolation and climatic differentiation explain patterns of morphometric, ecological niche, and genetic differentiation in the Atacama subclade?

The Atacama Desert has been the subject of recurrent studies trying to determine the effects of extreme aridity on the generation of patterns of divergence (Luebert, 2011). Thus far, only a few studies in Atacama angiosperms have examined interspecific divergence, mainly focusing on the effects of genetic divergence due to local aridity (Albrecht et al., 2010; Crawford et al., 1993; Gengler & Crawford, 2000; Viruel et al., 2011). In contrast to these molecular-oriented studies, the present study represents the first attempt to demonstrate interspecific divergence using multiple sources of evidence in an endemic lineage from the Atacama Desert.

The results indicate that climatic differentiation among habitats may be a strong factor of in the divergence of the lineages in the Atacama subclade of *Schizopetalon*. Nevertheless, the lack of differentiation from molecular and morphological data also implies that the limits of all analyzed species are not clear. These observations suggest a pattern of allopatric divergence driven by local environmental differences, which seems to be obscured due to the contribution of several processes like phenotypic plasticity, hybridization, and/or lack of coalescence. This scenario seems likely, considering the presence of local climatic variability in the area (e.g., Gengler-

Nowak, 2002; Luebert & Wen, 2008; Ossa et al., 2013; and Chapter 4), the recent origin of hyperaridity in the Atacama flora (e.g., Miocene; Latorre et al., 2007), and the strong selection for morphological features to survive under such conditions (e.g., Angert et al., 2010).

While the previous hypothesis is framed on the assumption that divergence is driven by the local conditions of the Atacama Desert, our data cannot reject an equally likely scenario of ex-situ diversification and subsequent ecological filtering over the altitudinal-ecological gradient (e.g., Graham, Parra, Rahbek, & McGuire, 2009; Parra, Rahbek, McGuire, & Graham, 2011). This hypothesis implies divergence outside the boundaries or restricted inside to small areas of the areas of distribution of the Atacama Desert, with a subsequent colonization after periods of local adaptation or formation of the new ecological habitats (e.g., Heibl & Renner, 2012).

Finally, while population differentiation driven by ecological conditions is plausible for lineages in the Atacama Desert, additional considerations are required given the nature of the biological system under study and the type of analyses conducted in this study. Even though our morphometric analyses revealed significant interspecific overlap, other attributes under more selective pressure, but not measurable without in situ observation, could reveal clearer patterns of phenotypic differentiation. For example, given that annual desert plants are strongly subject to interspecific selection in reproductive features of life history (i.e., self-fertilization, pollen limitation, germination times and rates, etc; Gutterman 2002), measuring these characters could provide additional evidence either for or against interspecific phenotypic overlap. Another consideration is the possibility overestimating allopatric divergence in the ecological niche overlap test due to the influence of climatic gradients. This issue is related to high probability of finding unique ecological combinations across a distribution, which can artificially increase the

estimates ecological divergence (Broenninmann et al., 2012; Godsoe, 2010; Warren, Glor, & Turelli, 2010; and Appendix 6). To determine if these limitations affect the inferences of the current study will require further scrutiny. In this sense, we suggest that field observations using alternative approaches (e.g., Kearney & Porter, 2009) might provide an effective test of our results, and assist in differentiating the effects of local aridity in the promotion of lineage diversification in the Atacama Desert versus ecological filtering.

Table 1. List of morphological characters analyzed, their code, and the number of individuals measured per species.

		Number of individuals				
Character	Code	*S.biseriatum*	*S.tenuifolium*	*S.rupestre*	*S.bipinnatifidum*	Total
Average length of lowermost stem internodes (size corrected - cm)	StIn c	137	100	260	66	563
Average length of lowermost inflorescence internodes (size corrected - cm)	InIn c	135	93	251	63	542
Average length of fruit-flower pedicel (cm)	FPe	138	101	261	66	566
Average sepal length (cm)	SeLe	130	100	261	63	554
Average leaf blade area (size corrected - mm^2)	A c	92	78	188	51	409
Average leaf blade perimeter (size corrected - mm)	P c	92	78	188	51	409
Average length/width ratio	L/W	92	78	188	51	409
Dissection Index	DI	92	78	188	51	409

Table 2. Results obtained from univariate character comparisons among species (Kruskal – Wallis test, H) and between pair of species (Dunn's multiple comparison test, Z). Asterisks and bold letters represent significant differences for comparison tests (Kruskal Wallis, $P < 0.05$; Dunn's test at 6 comparisons Bonferroni corrected, $P < 0.008$).

	Linear				Leaf			
	SeLe	FPe	StIn c	InIn c	A c	P c	L/W	DI
Kruskal-Wallis	**59.26(*)**	**75.86(*)**	**85.38(*)**	**56.90(*)**	**264.16(*)**	**115.21(*)**	**128.74(*)**	**263.45(*)**
S. biseriatum - *S. tenuifolium*	**3.012 (*)**	**5.409(*)**	**4.705(*)**	1.301	0.797	**3.559(*)**	**3.213(*)**	**3.179(*)**
S. biseriatum - *S. bipinnatifidum*	**6.229 (*)**	2.566	**6.123(*)**	**4.123(*)**	0.768	**4.163(*)**	**4.028(*)**	2.494
S. biseriatum - *S. rupestre*	**6.709 (*)**	**8.067(*)**	**8.687(*)**	**5.020(*)**	**12.746(*)**	**4.708(*)**	**10.807(*)**	**10.229(*)**
S. tenuifolium - *S. bipinnatifidum*	**3.453 (*)**	2.049	1.886	**4.924(*)**	1.426	0.993	1.158	0.299
S. tenuifolium - *S. rupestre*	**2.786 (*)**	1.318	**2.641(*)**	**5.849(*)**	**12.527(*)**	**8.516(*)**	**6.537(*)**	**13.296(*)**
S. bipinnatidfidum - *S. rupestre*	1.57	**3.454(*)**	0.233	0.512	**9.245(*)**	**8.398(*)**	**4.255(*)**	**11.001(*)**

Table 3. Descriptive statistics for each character per species. Med: Median, SD: Standard Deviation, M-W: Mann Whitney test. Statistical significant differences are represented in bold and asterisks represent their P-values (* = P < 0.05). The results of the Kruskal-Wallis tests did not consider *S. bipinnatifidum* specimens (see text).

		SeLe	FPe	StIn c	InIn c	A c	P c	L/W	DI
S.biseriatum	Mean	-0.211	-0.041	1.1	0.95	0.886	1.007	0.47	1.073
	SD	0.055	0.126	0.162	0.177	0.086	0.072	0.063	0.081
	Med	-0.213	-0.051	1.095	0.937	0.888	1	0.466	1.081
	M-W	**3985** (*)	**4071** (*)	**4166** (*)	**4802** (*)	**1398** (*)	**1356** (*)	1987	1663
S.bipinnatifidum	Mean	-0.156	-0.099	0.96	1.051	0.898	1.077	0.408	1.144
	SD	0.058	0.108	0.113	0.137	0.13	0.081	0.094	0.107
	Med	-0.157	-0.104	0.953	1.071	0.899	1.09	0.407	1.133
S.tenuifolium	Mean	-0.189	-0.141	1.004	0.918	0.861	1.059	0.424	1.146
	SD	0.049	0.123	0.13	0.151	0.104	0.076	0.086	0.09
	Med	-0.19	-0.124	1.003	0.932	0.86	1.066	0.425	1.147
	M-W	1999	**2525** (*)	**23277** (*)	1427	**10388** (*)	11830	12686	12531
S.rupestre	Mean	-0.174	-0.166	0.956	1.044	1.142	0.951	0.344	0.893
	SD	0.058	0.129	0.135	0.162	0.101	0.085	0.08	0.091
	Med	-0.17	-0.165	0.955	1.046	1.149	0.942	0.343	0.892
	M-W	26711.5	**27897** (*)	**2065** (*)	24343	**1945** (*)	**1764** (*)	1538	1468
Kruskal-Wallis	Herbarium Only	**54.01** (*)	**13.91** (*)	**54.01** (*)	**49.82** (*)	**135.8** (*)	**31.3** (*)	**60.87** (*)	**120.79** (*)
	Field Only	1.29	**80.52** (*)	**42.59** (*)	**7.17** (*)	**101.24** (*)	**41.68** (*)	**69.03** (*)	**94.66** (*)

Table 4. Summary of variable contribution in the estimates of the ecological niche models for the study group. Lower numbers represent the contribution of each variable in simple and the jackknifed models (parenthesis).

	S. bipinnatifidum	*S. biseriatum*	*S. rupestre*	*S. tenuifolium*
Variable 1	PC 4	PC 3	PC 1	PC 1
	46 (9.8)	38.1 (0)	52.9 (32.6)	40.2 (57.1)
Variable 2	PC1	PC 2	PC 4	NDVI min
	29.1 (31.3)	32.7 (29.5)	27.7 (28.1)	22.2 (20.3)
Variable 3	NDVI max	PC 1	PC 2	PC 3
	13.3 (36.7)	12.1 (60.5)	6.1 (17)	10.8 (6.8)
Variable 4	PC3	NDVI min	PC 3	PC 4
	5.1 (7.9)	7.8 (0)	5.3 (2.5)	7.7 (1.7)
Variable 5	NDVI min	PC 4	NDVI sd	PC 2
	3.5 (13.5)	6.7 (8.5)	3.5 (2.1)	6.8 (1)
Variable 6	PC2	PC 5	PC 5	PC 5
	2.8 (0.2)	2.4 (1.5)	3.2 (5.7)	5.1 (2.7)
Variable 7	PC 6	NDVI sd	NDVI max	NDVI max
	0.3 (0.7)	0.2 (0)	0.9 (11.4)	3.3 (7.7)
Variable 8	NDVI sd	PC 6	PC 6	PC 6
	0 (0)	0 (0)	0.4 (0.4)	2.5 (2.2)
Variable 9	PC5	NDVI max	NDVI min	NDVI sd
	0 (0)	0 (0)	0 (0)	1.4 (0.5)
AUC training	0.967	0.993	0.979	0.989
AUC test	0.957	0.982	0.963	0.992

Table 5. Indices of ecological niche overlap (D and I) and background similarity test (Z test).

Species for observed distribution	Background species	D	I	More or less similar than expected (Z-test)
S. biseriatum	*S. tenuifolium*	0.0930	0.0835	P<001
	S. bipinnatifidum	0.0023	0.0010	P<001
	S. rupestre	0.0004	0.0001	P<001
S. tenuifolium	*S. biseriatum*	0.0930	0.0835	P<001
	S. bipinnatifidum	0.0011	0.0004	P<001
	S. rupestre	1.11E-15	9.99E-16	P<001
S. bipinnatifidum	*S. biseriatum*	0.0023	0.0010	P<001
	S. tenuifolium	0.0011	0.02486	P<001
	S. rupestre	0.0384	0.0004	P<001
S. rupestre	*S. biseriatum*	0.0004	0.0001	P<001
	S. tenuifolium	1.11E-15	9.99E-16	P<001
	S. bipinnatifidum	0.0384	0.02486	P<001

Figures

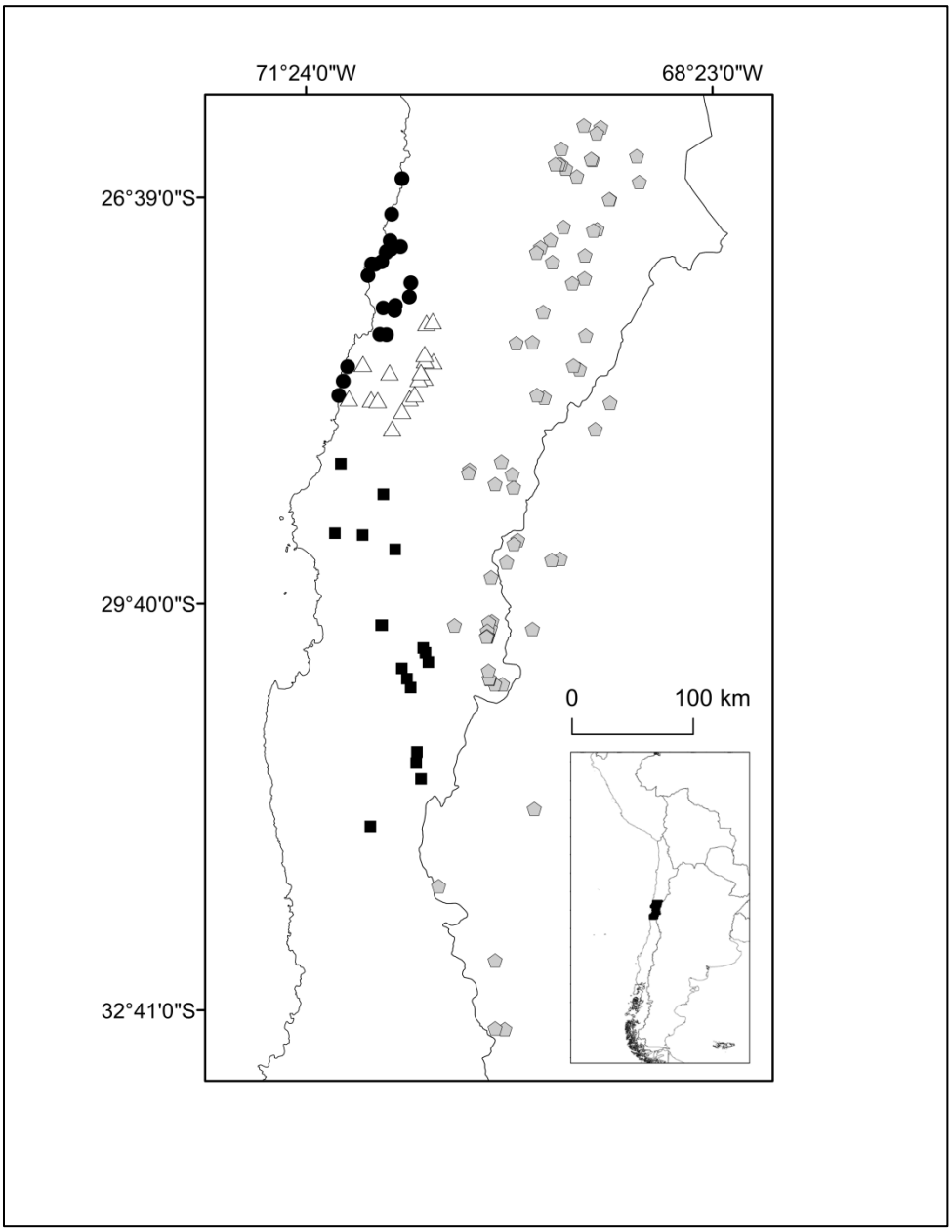


Fig 1. Distribution of the species of the Atacama subclade of *Schizopetalon*; black dots (*S. biseriatum*), hollow triangles (*S. tenuifolium*), black squares (*S. bipinnatifidum*) and grey hexagons (*S. rupestre*).

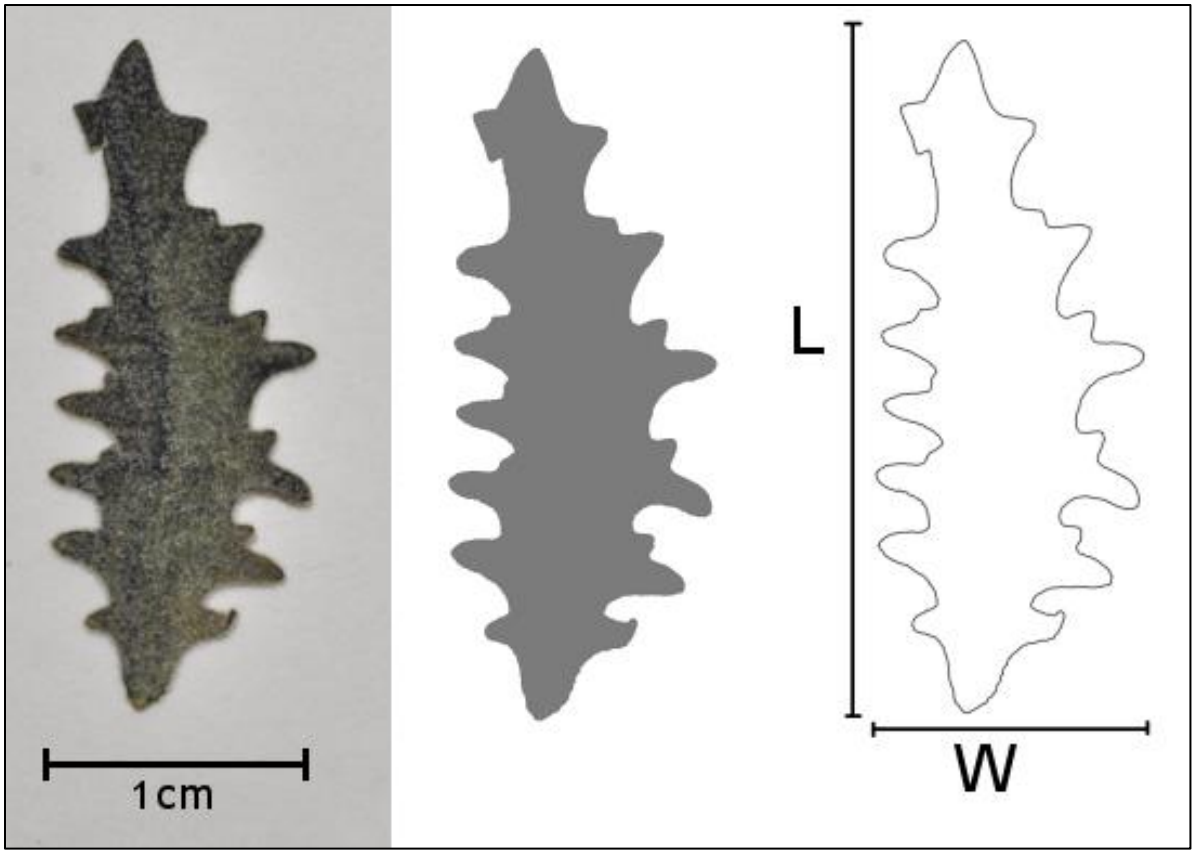


Fig. 2. Treatment applied to pictures of leaves for measurement of area, perimeter, and length/width ratio. From left to right: original, shaded lamina representing binary picture and estimated area, and squared boundaries with vertical and horizontal lines representing length (L) and width (W), respectively.

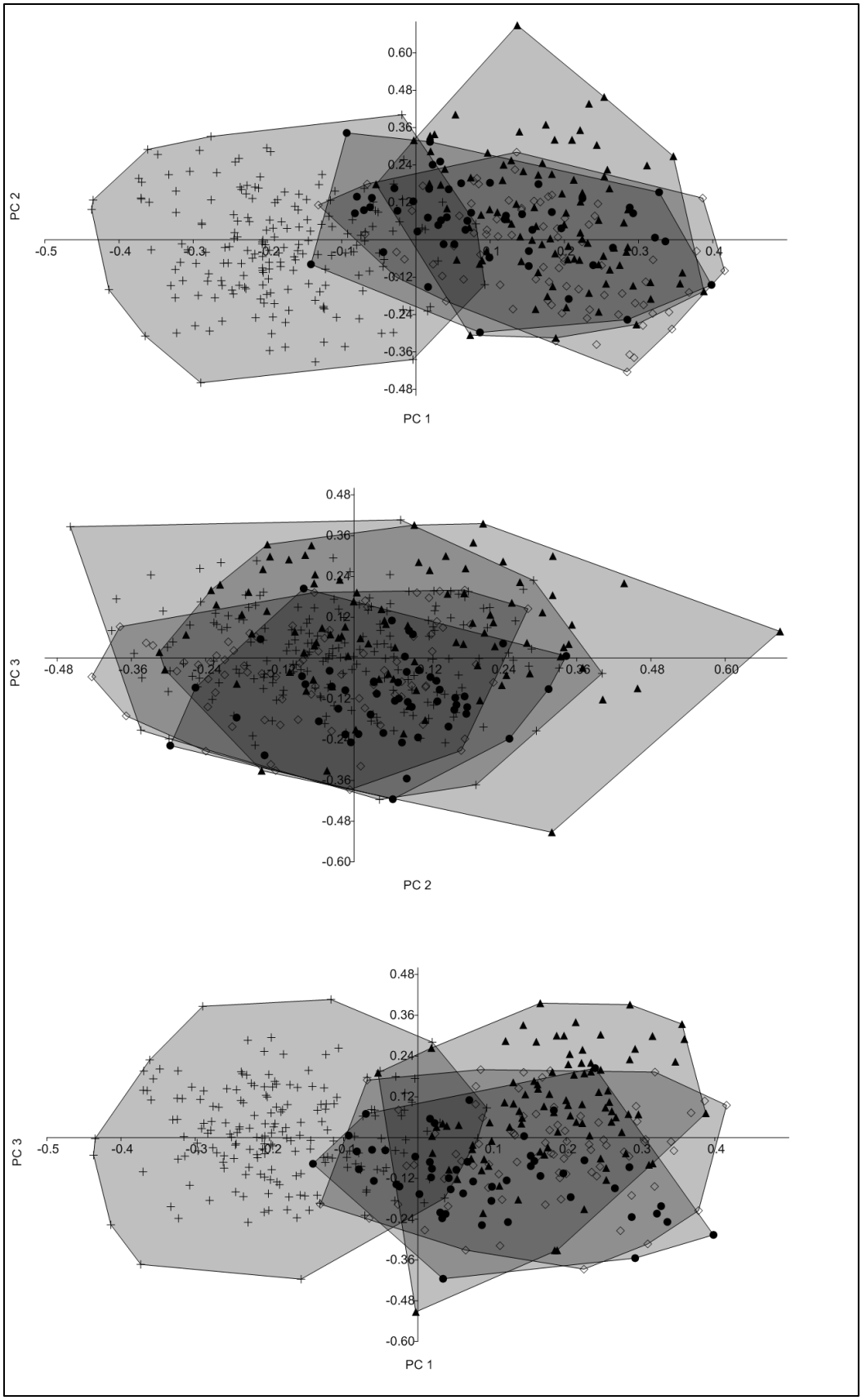

PC 2
PC 1
PC 3
PC 2
PC 3
PC 1

Fig. 3. Principal component analysis based on four linear and four leaf character. The three first axes accounts for ~ 78 % of total variation. Symbols represent different species: triangles (*S. biseriatum*), empty diamonds (*S. tenuifolium*), black dots (*S. bipinnatifidum*) and crosses (*S. rupestre*).

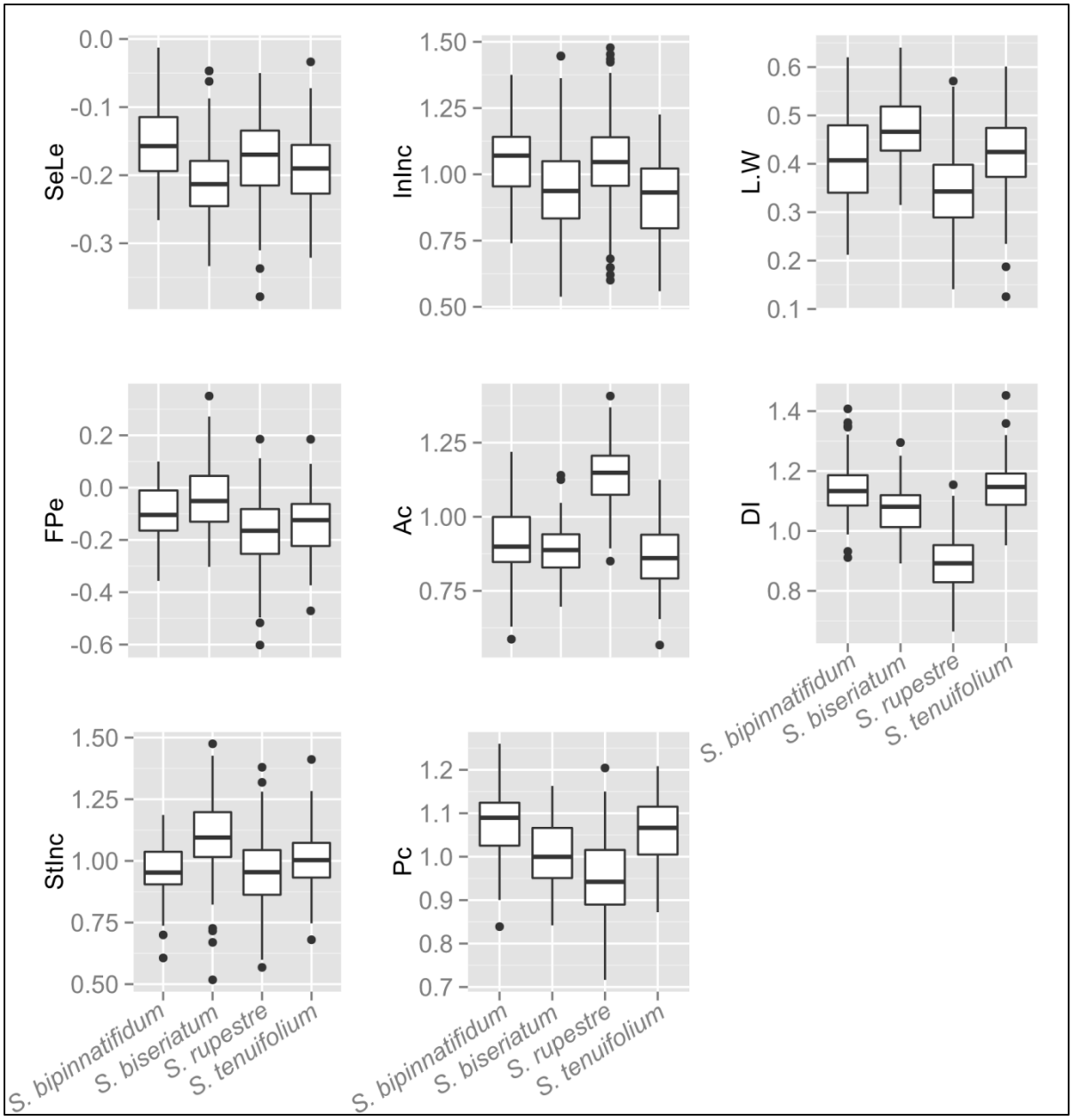


Fig. 4. Boxplot graphs of morphometric characters individual species. Black dots represent outliers.

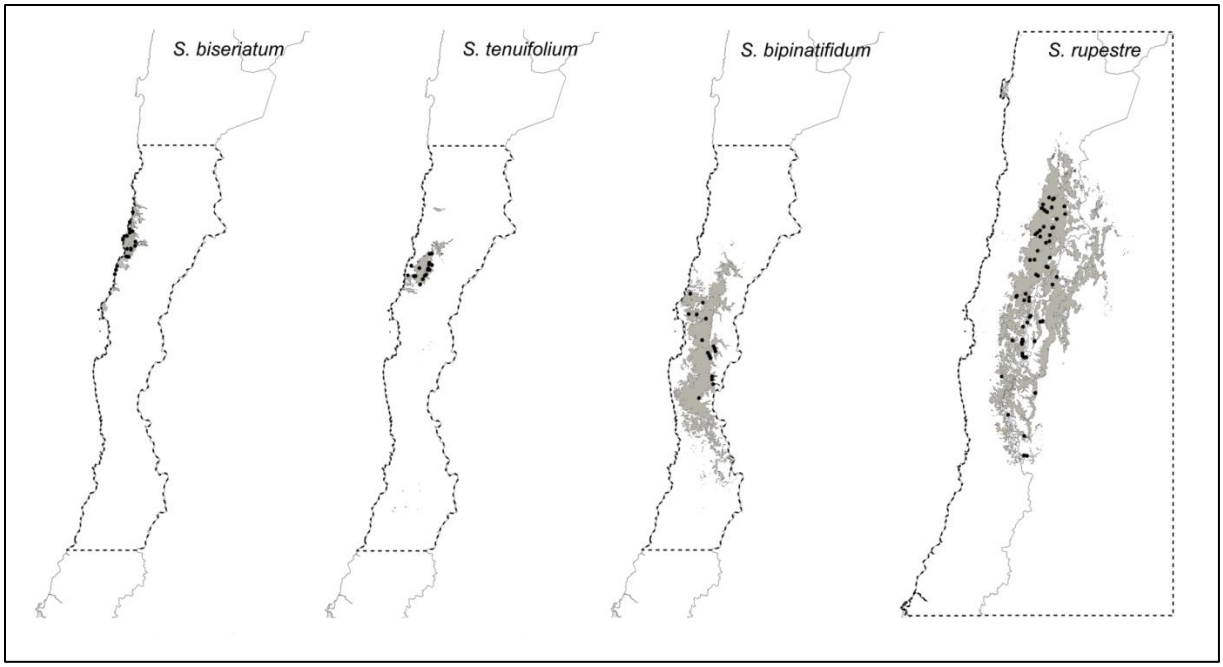


Fig. 5. ENM-derived potential distribution of the species of the Atacama subclade inferred from bioclimatic (6 components) and NDVI data. Shaded area represents the estimated presence areas based on the minimum training presence value points. Points of only presence are shown in black dots and training masks are delimited in dashed lines.

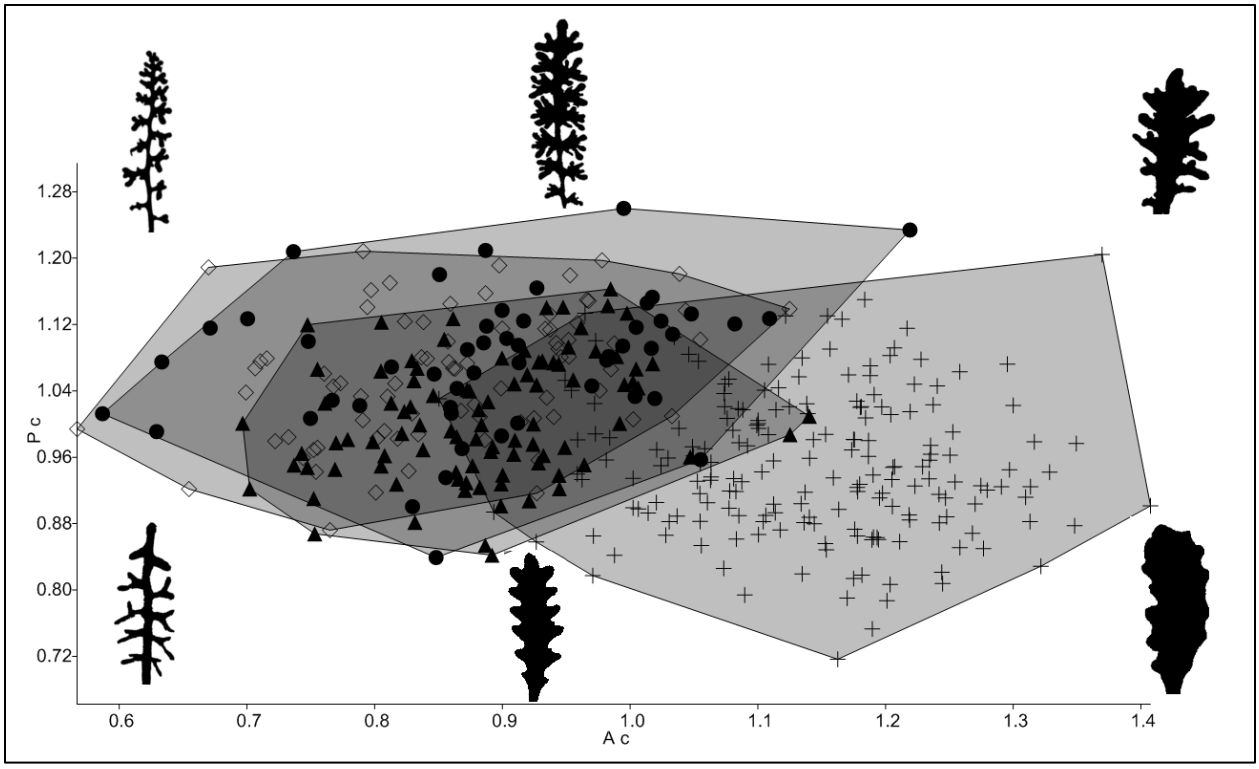


Fig. 6. Representation of leaf variation in the morphospace created by area (A c) and perimeter (P c). Figures represent examples of dissection and shape inside the area created by A c and P c. Symbols represent different species: triangles (*S. biseriatum*), empty diamonds (*S. tenuifolium*), black dots (*S. bipinnatifidum*) and crosses (*S. rupestre*).

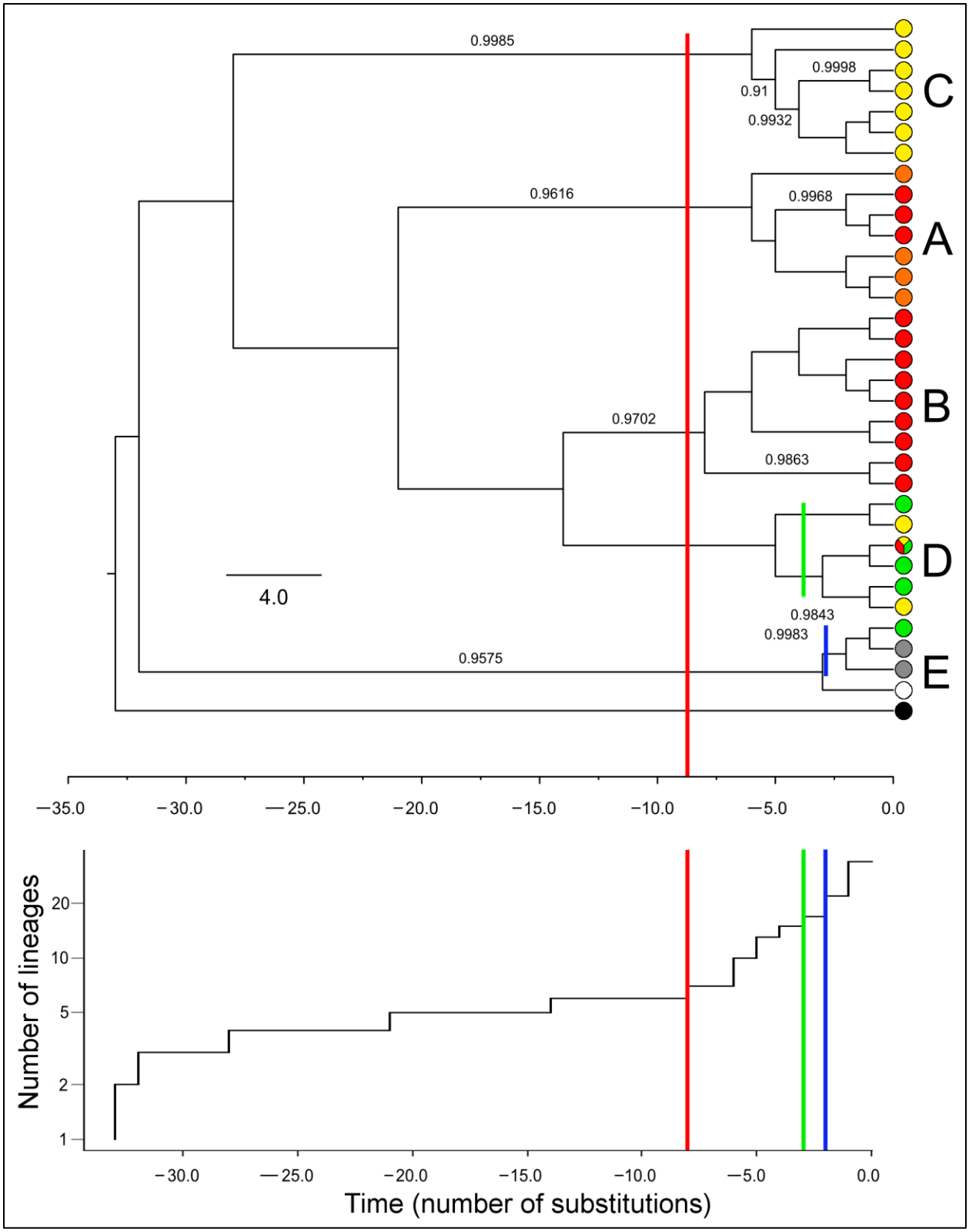

0.9985
0.9998
0.91
0.9932
C
0.9616
0.9968
A
B
0.9702
0.9863
D
4.0
0.9843
0.9983
0.9575
E
−35.0
−30.0
−25.0
−20.0
−15.0
−10.0
−5.0
0.0
Number of lineages
20
10
5
2
1
−30.0
−25.0
−20.0
−15.0
−10.0
−5.0
0.0
Time (number of substitutions)

Fig 7. Maximum clade probability ultrametric tree (above) and the GMYC plot of lineages over time of the Atacama subclade (below), inferred with MMT haplotypes. Colored circles at the tips of the ultrametric tree represent the identity of the unique or shared reduced haplotypes (*S. arcuatum*: gray, *S. biseriatum*: yellow, *S. bipinnatifidum*: orange, *S. rupestre*: red, *S. tenuifolium*: green, and *S. walkeri*: black) and numbers above braches represent significant posterior probability (> 0.95) of each delimited group. The vertical red bar shows the inferred single-threshold point from Yule and coalescent process; the green and blue vertical bars represent the points of transitions in the case of the multiple-threshold model.

CHAPTER FOUR

DETERMINING PATTERNS OF INTERSPECIFIC DIVERSIFICATION IN THE ATACAMA SUBCLADE OF *SCHIZOPETALON* (BRASSICACEAE) USING LOW COPY NUCLEAR GENES

Abstract

The present study illustrates how to obtain low copy nuclear gene sequences (LCNG) in species of the Atacama subclade of *Schizopetalon* to study patterns of interspecific diversification in the Atacama Desert. The use of LCNG can greatly enhance resolving the patterns of evolution in this group, especially to study phylogeography and population divergence. Using available genomic databases in Brassicaceae, a set of 21 intron regions were screened and tested for phylogenetic utility. Most of the difficulties encountered in obtaining LCNGs in *Schizopetalon* were not related to difficulties in the transferability of primers, but, as commonly found in other angiosperms, trying to find sequences with limited copy variation. Two low copy nuclear genes (ELF8 and MMT) were examined to determine patterns of evolution in the Atacama subclade. Sequences of MMT revealed patterns of local geographic isolation in populations, whereas ELF8 had limited utility due to apparent recombination. Evidence of shared haplotypes among the main geographic groups suggests a process of hybridization and/or incomplete lineage sorting in the northern range of distribution, between coastal and Andean lineages. The hypothesis of past hybridization is favored because ecological niche modeling suggests climatic oscillations that altered patterns of aridity that changed the distribution of populations during the Quaternary.

Introduction

The Atacama Desert is an intriguing area to study plant diversification. Due, in part to its hyper-aridity and the uniqueness of the flora that inhabits the region, there has been an increased number of phylogenetic studies of angiosperm lineages (Luebert, 2011). These studies have resulted in the formulation of hypotheses related to modes of speciation, most of which suggest rapid divergence of local lineages driven by extreme aridity (Dillon et al., 2009; Gengler-Nowak, 2002; Guerrero et al., 2011a; Guerrero et al., 2013; Heibl & Renner, 2012; Luebert & Wen, 2008; Luebert et al., 2009). Although the connection between climate and diversification in the Atacama Desert is generally accepted (Luebert, 2011), there has yet to be comprehensive studies of patterns of inter and intraspecific diversification within endemic plant lineages. One potential issue associated with studying plant diversification at lower levels in the Atacama Desert is that, like oceanic islands, the rapid and recent divergence (Guerrero et al., 2013)in most Atacama and Andean flowering plant lineages results in low levels of molecular divergence and thus, limited resolution of interspecific relationships.

A source of molecular markers for such studies is low copy nuclear genes (LCNG), which are nuclear regions that are variable enough to be phylogenetically informative at lower taxonomic levels. Their utility resides in their combination of exons and introns; the former useful for primer design because they are relatively well conserved and the latter for their higher rate of sequence evolution making them useful at lower phylogenetic levels (Sang, 2002; Small, Cronn, & Wendel, 2004). Compared to traditional molecular markers used in plants (nuclear ribosomal and chloroplast DNA; Mort et al., 2007). In sum, LCNGs have higher levels of molecular divergence, biparental inheritance, and serve as putatively unlinked molecular markers (Álvarez,

Costa, & Felliner, 2008; Sang, 2002; Small et al., 2004).

Because of that, LCNG have proven useful to infer patterns of evolution at lower taxonomic levels, especially to resolve challenging taxonomic, systematic, and evolutionary scenarios at the species and/or population level in plants. Analyses of LCNG have helped to solve complex scenarios of speciation such as rapid radiations (e.g., Cronn, Small, Haselkonr, & Wendel, 2002; Malcomber, 2002; Naumann et al., 2011), hybridization (e.g., Clarkson et al., 2010; Howarth & Baum, 2005), and polyploidy (e.g., Clarkson et al., 2010; Kim, Sulton, & Donoghue, 2008; Marcussen et al., 2012; Smedmark, Erikson, Evans, & Campbell, 2003). These data are a key tool for studying plant phylogeography (Schaal et al., 1998) and comprehensive analyses of speciation. For example, Barret and Freudenstein (2011) illustrated the utility of LCNG in *Corallorhiza* species (Orchidaceae). Their study used analyses of LCGN in conjunction with morphology and geography to determine the limits of species and test discordant hypotheses of species differentiation. Given the documented utility of these markers for lower level analyses, LCNG could provide an important source of data to disentangle the patterns of diversification in recently derived lineages of the Atacama flora.

The Atacama subclade of *Schizopetalon* represents a suitable candidate to analyze patterns of interspecific diversification through the use of LCNG. This is a small lineage of four species of Brassicaceae (i.e., *S. bipinnatifidum*, *S. biseriatum*, *S. rupestre* and *S. tenuifolium*) that are endemic to South America. All species occur in the southern Atacama Desert, from the boundaries of the absolute desert (~0 mm 27S) to the semi-arid area of Coquimbo (31° S). According to previous phylogenetic studies using morphology (Al-Shehbaz, 1989), these species

are located in different ecoregions: *S. biseriatum* + *S. tenuifolium* (coast), *S. bipinnatifidum* (intermediate altitude) and *S. rupestre* (Andes)**.** Molecular analyses of a combined nuclear ribosomal (ITS and ETS) and chloroplast (*trnH-psbA* + *rps16* + *atpI-atpH* +*trnQ-rps16*) data set recovered a well-supported Atacama subclade (Toro-Núñez et al., 2013). However, relationships within this clade were largely unresolved.

One of the main limitations in obtaining LCNG is the lack of sequence information to design primers to amplify orthologous sequences (Schlüter, Stuessy, & Paulus, 2008). Nevertheless, newly available data from next generation sequencing has facilitated the widespread use of these markers across different plant lineages (e.g., Zhang, Zeng, Shan, & Ma, 2012; Zimmer & Wen, 2012). Even though the possibility of using these databases to design primers to amplify orthologous genes, their use for phylogenetic and phylogeographic studies is still a difficult process. The necessary criteria for selecting appropriate markers include: 1) consistent amplification across taxa, 2) single or at least low copy number, 3) few alleles per sequence, and 4) sufficient levels of variation for phylogenetic and phylogeographic analyses. While such information is still limited in some lineages, genomic information from model systems (e.g., *Arabidopsis thaliana* and *Brassica napus*) might be applied to non-model organisms, especially in well studied families, such as Brassicaceae. The present study will explore the use of LCNG in the Atacama subclade of *Schizopetalon* using available genomic resources for Brassicaceae and applying bioinformatic approaches to inferring alleles. The main objectives of this study are 1) screen available sources of nuclear data in Brassicaceae, 2) use statistical approaches to infer alleles in an attempt to lower the costs incurred from PCR subcloning, 3) identify potential LCNG for phylogenetic or phylogeographic use, 4) test their utility for the Atacama subcalde of

Schizopetalon, and 5) infer patterns of diversification.

Materials and Methods

Specimen and DNA sampling

A total of 29 populations were sampled from field locations during 2010 and 2011; these populations represent most of the geographic range of the target species (Fig.1; Appendix 7). From these populations, three individuals were randomly selected for DNA extraction and PCR amplification. Sampling efforts for DNA analysis in *S. rupestre* populations were focused only on the northern Chilean range of its distribution, because it was not possible to reach populations in the Argentinean range. Despite this, the populations collected in Chile covered the entire northern geographic range of *S. rupestre*. Samples were identified using both morphology and previous molecular studies (Al-Shehbaz, 1989; Toro-Núñez et al., 2013) and voucher specimens from each population were deposited at the McGregor Herbarium (KANU) at The University of Kansas.

To find markers for the study group, LCNG known in Brassicaceae (e.g., Koch, Haubold, & Mitchell-Olds, 2000; Kuittinen et al., 2002; Slotte et al., 2006; Wright, Lauga, & Charlesworth, 2003) were tested, but none consistently amplified (i.e., absence of single bands or readable sequences). For this reason, a *de novo* approach to primer design for LCNG was conducted by searching from expressed sequence tag sequences (EST) in the genomic comparison between *Arabidopsis thaliana* (TAIR; Lamesch et al., 2011) and *Brassica napus* (BRAD; Cheng et al., 2011). The IntrEST database (Ilut & Doyle, 2012) was used with search preferences focused on nuclear regions with a range of 1196 to 2804 aligned bp, 90% minimum identity threshold, and

introns no larger than 500 bp. Using the *A. thaliana* reference genome available in GenBank (Benson et al., 2010), a search of the initial set of ESTs was conducted to identify regions with enough physical separation to decrease the chance of recombination and linkage disequilibrium among loci. DNA regions were chosen based on their mapped presence on different chromosomes or using a minimum physical distance of 3kb or 332 to 575 cM, based on the report of 50 % decreased chance of linkage disequilibrium in syntenic genes in *A. thaliana* (Kim et al., 2007). Primer design of the selected nuclear regions was conducted using Primer3 (Rozen & Skaletsky, 2000) on completely congruent regions of *A. thaliana* and *B. napus*, preferring primers of 20 to 25 bp of length and melting temperatures ranging from 53 to 60 C. Finally, primer sets for 22 candidate genes were obtained (Table1; Appendix 8).

A small amount of silica-dried material was used to extract total DNA using a modified CTAB method (Doyle & Doyle, 1987; Mort et al., 2001). Polymerase chain reactions (PCR) were conducted using an initial denaturation of 94 C for 5 min, 35 cycles of 95 C for denaturation, 50 C of annealing, and 72 of extension and a final extension of 72 C for 10 min. PCR reactions were conducted with a negative control and a sample of *Arabidopsis thaliana* as a positive control. Purified amplicons were sequenced by Macrogen Inc. (Rockville, Maryland, U.S.A.) with the same forward and reverse primers used for PCR. The resulting contigs were edited and assembled using GENEIOUS v. 6 (Biomatters. available from http://www.geneious.com/). Heterozygous positions were detected as double peaks using the 40% overlap settings in GENEIOUS, discarding regions that had more than two overlapped peaks at any given site per individual. The resulting sequences were confirmed via BLAST searches and using the *A. thaliana* reference genome. Since the inherent risk of amplification of paralogous sequences, an

additional search of sequences was conducted in BLAST on known conserved domain sequences (CDS) of *A. thaliana* and *B. napus*. This step was conducted in order to corroborate that amplified sequences did not belong to known duplicated genes or regions. An arbitrary threshold of e-values above $1x10^{-10}$ similarity was used to determine the known presence of paralog copies. Sequence alignment was conducted using Jalview v. 2.8 (Waterhouse et al., 2009)

Molecular analyses

All candidate regions were selected based on the following criteria (Fig. 2): single bands in PCR, absence of overlapped reads per sequence, limited number of heterozygous positions (>30%), and appropriate levels of variability among sequences (> 5 divergent nucleotide sites). The selected sequences were screened using eight to twenty samples from different geographic localities and species, with a combination of samples from the Atacama subclade (i.e., *S. biseriatum*, *S. bipinnatifidum*, *S. tenuifolium*, and *S. rupestre*), the Mediterranean subclade (i.e., *S. maritimum* and *S. walkeri*), and *S. arcuatum*. Levels and distribution of molecular variability were explored using a neighbor-net network based on uncorrected p-distances on all the candidate genes using SPLITSTREE v. 4 (Huson & Bryant, 2006).

In order to assess the potential utility of the selected regions for phylogenetic and phylogeographic analyses, two regions, ELF8 (fourth intron) and MMT (ninth and tenth intron) were tested further for gene tree estimation. Prior to conducting analyses, the loci were phased to infer the most likely haplotypes for heterozygous individuals. Due to the detection of length-variable DNA in sequence readings from several samples, a preliminary allele size correction was conducted using CHAMPURU v. 1.0 (Flot, 2007; Flot, Tillier, Samadi, & Tillier, 2007).

Subsequently, all sequences were formatted with SEQPHASE (Flot, 2010) and were phased with PHASE v. 2.1.1 (Stephens & Donnelly, 2003; Stephens, Smith, & Donnelly, 2001) using the recombination model (-MR0; Li & Stephens, 2003) on two independent runs of 1000 iterations (burnin = 100, thinning interval = 100) each were conducted. Obtained haplotypes were determined with a margin of 60% of posterior probability to include potential missing haplotypes (Garrick, Sunnucks, & Dyer, 2010). Summary statistics of variability (e.g., number of segregating sites, number of haplotypes, haplotype diversity, nucleotide diversity per site and Watterson index) were obtained from the aligned sequences. In addition, a recombination test (Hudson's four gamete test) and neutrality test (Tajima D and Fu's F) were conducted using DNAsp v. 5 (Librado & Rozas, 2009).

A ML analysis was conducted using GARLI v. 2.0 (Zwickl, 2006), performing 10 replicate runs and 2 million of iterations of each. To determine branch support, 100 bootstrap replicates were conducted. Data sets were rooted with sequences obtained from *S. arcuatum*, *S. walkeri,* and *S. maritimum* in both data sets. Models of molecular substitution for ML analyses were inferred using Modeltest v. 2.1.1 (Darriba, Taboada, Doallo, & Posada, 2012; Guindon & Gauscuel, 2003). Additionally, display of haplotype network was conducted using statistical parsimony in TCS v. 1.0 (Clement, Posada, & Crandall, 2000) using a threshold connection of 95 %, and a neighbor-net network using SPLITSTREE v. 4. Indels were not considered for the reconstruction of haplotype networks and parsimony networks.

Results

Screening potential low copy nuclear regions

Positive PCR signal of the 21 candidate loci was detected in the ingroup (*S. biseriatum*, *S. bipinnatifidum*, *S. tenuifolium*, and *S. rupestre*) and outgroup (*S. arcuatum*, *S. maritimum*, and *S. walkeri*) species of *Schizopetalon*. From these samples, following the criteria for nuclear locus selection (Fig. 2), the search was narrowed to four regions (ALDH108A, CH4, ELF8, and MMT; Table 1). These loci were determined to be associated with metabolic and developmental functions, with varied chromosome position and intron numbers (Fig. 3) in *Arabidopsis*. The first region corresponded to the aldehyde dehydrogenase 108A (ALDH108A) locus, located on chromosome 1 and comprised three flanked introns: five (101 bp), sixth (80 bp), and seventh (103 bp). The second region corresponded to the trans-cimmate 4-monooxygenase (CH4) locus, which was located on chromosome 2 and included two introns: first (104 bp) and second (255 bp). Also on chromosome 2, the protein early flowering 8 (ELF8) locus that included a single intron was identified (485 bp). Finally, the locus methionine S-methyltransferase (MMT) locus which is found on chromosome 5 included two introns - the ninth (144 bp) and tenth (367 bp).

The resulting topologies generated from these four loci resulted in congruent topologies and in most cases separated ingroup and outgroup sequences in the neighbornet network (Fig. 4). All four loci separated samples of *S. arcuatum* from the remaining species of the Atacama subclade. In the Atacama subclade, none of the taxonomically recognized species were resolved within a single clade. Different levels of resolution were detected among loci and ELF 8 and MMT were the most stable in terms of number of cycles (i.e., incongruences) produced in the networks. Given this difference, these two regions were selected for expanded population sampling.

Genetic variation and gene trees in ELF8 and MMT introns

The fourth intron of ELF8 after phasing was 384 bp and had 46 heterozygous positions in the ingroup (~ 11.7%). A total of 20 insertion and deletion events (indels) were detected, which ranged in size from 1 to 14 bp. Most of these indels appeared to not be phylogenetically informative among individuals and populations, except for an indel of heterogeneous size (13 to 15 bp) shared across all species in the ingroup and outgroup in positions 170 to 183 (Appendix 9). The ninth and tenth intron of the MMT locus had a combined length of 647 bp with 41 heterozygous positions in the ingroup (~ 4.79%). A total of four indels were detected, ranging in size from 1 to 213 bp. Two of these indels correspond to one individual from population 30 of *S. tenuifolium*. The other two indels correspond to a large insertion shared among all individuals from populations 110 and 112 of *S. rupestre* (Appendix 8). A total of 122 sequences of ELF8 were obtained from 61 individuals. Thirty two sequences were homozygous for this locus. A total of 146 sequences were obtained for MMT, representing a total of 82 individuals. Fifty seven sequences were homozygous.

Haplotype diversity was high with similar levels of variability in ELF8 ($h = 0.814$) and MMT ($h = 0.735$); nevertheless, there was almost a ten-fold difference between these loci for nucleotide diversity ($ELF\pi = 0.0123$, $MMT\pi = 0.00471$). At the species level, haplotype and nucleotide differences were also found between loci. In MMT, *S. biseriatum* was the most variable species in terms of haplotype number and nucleotide diversity ($h = 0.666$, $\pi = 0.00813$), whereas *S. tenuifolium* ($h = 0.452$, $\pi = 0.00283$) and *S. rupestre* ($h = 0.656$, $\pi = 0.00042$) had lower levels of variability (Table 1). For ELF8, all species had similar magnitudes of haplotype number and nuclear variability- *S. biseriatum* ($h = 0.844$, $\pi = 0.01387$), *S. tenuifolium* ($h = 0.835$, $\pi = 0.01484$), and *S. rupestre* ($h = 0.828$, $\pi = 0.0123$). Since only one population of *S. bipinnatifidum*

was sampled, molecular variability is not reported for this species. Neither Tajima's D and Fu's F neutrality tests revealed evidence to reject the null hypothesis of neutral evolution in all species (Table 2), with the exception of Fu's F test for MMT in *S. rupestre* ($F = -4.578$, $P < 0.005$).

Haplotype network analyses revealed differences between loci. The levels of incongruence in the estimated distance networks were excessively high in the ELF8 fourth intron, especially because of the presence of multiple cycles and the total absence of grouping among the estimated haplotypes. Results from recombination tests suggest a high number of events per species (from 2 to 4) for this locus. Given that intragenic recombination affects the assumption of independence among nucleotide sites and regions, this locus was eliminated from further sequencing efforts. In MMT, two recombination points were detected involving positions 431, 510 and 600 (Table 2). The common break point (510) reveals two haplotypes from different alleles and species (populations 16 and 115, individuals 20 and 21 respectively) with dissimilar sequences; eliminating the involved sequences or breaking apart the matrix retaining the longest section did not affect any of the results. Therefore, no alterations were made to this matrix and the remaining results were based on analyses of this locus.

Gene trees inferred from parsimony network, network analysis, and ML reveal very similar patterns for the MMT locus. The haplotype parsimony network identified 30 haplotypes, excluding indels (Table 3). The organization of these haplotypes suggests at least five distinct groups (Fig. 5) and some of these reflect geography rather than taxonomy. The first group (A) includes haplotypes of *S. bipinnatifidum* (population 115) and populations 110 and 112 of *S. rupestre*. A second group (B) comprises haplotypes from populations representing the rest of the

geographic range of *S. rupestre,* located to the north of populations 110 and 112. This group also includes the haplotypes of population 42 of *S. tenuifolium*. A third group (C) comprises haplotypes from *S. biseriatum*, which was represented by all alleles from populations 35 and 39, and alleles of one parent from populations 20, 21, and 22. The fourth group (D) includes a collection of alleles from one and two parents' alleles of *S. tenuifolium*, *S. biseriatum*, and *S. rupestre*, which grouped most of the haplotypes (Fig. 5). The final group (E) represents the haplotypes of the outgroup taxa (i.e., *S. walkeri*, *S. maritimum*, and *S. arcuatum*), plus three haplotypes from individuals of population 42 of *S. tenuifolium*. The distance network (Fig. 6) shows similar patterns, separating the two groups of *S. rupestre* (A + B), *S. biseriatum* (C), and the group with the haplotypes from all species (D) plus the outgroup (E). The rooted ML tree (Fig. 7), estimated with a HKY model of nucleotide substitution, revealed the same configuration as the parsimony and the distance network. Haplotypes of *S. arcuatum* and *S. maritimum* were highly supported (99% bootstrap) as sister to the ingroup (99% bootstrap). The ingroup had the same four basic groups of haplotypes with different levels of support. The highest support was retrieved in group C (*S. biseriatum*; 84% bootstrap), followed by moderate to low support to the haplotypes of populations 110 and 112 (64% bootstrap) and the rest of the haplotypes of *S. rupestre* (59% bootstrap). Other small groups were detected in haplotypes from partial alleles of *S. tenuifolium* (populations 28, 33, and 22; 67% bootstrap) and the haplotypes of population 61 of *S. rupestre* (61% bootstrap). In this case, no support was obtained for the grouping of haplotypes of *S. bipinnatifidum* or the group D, which was unresolved.

Discussion

LCNG as alternative source to characterize molecular variability in Schizopetalon

Given the challenging task of unraveling patterns of evolution and diversification in Brassicaceae, the use of genomic resources described from *A. thaliana* has become a routine for evolutionary studies in this family (e.g., Bailey & Doyle, 1999; Baum, Yoon, & Oldham, 2005; Galloway, Malmberg, & Price, 1998; Koch et al., 2000; Kuittinen et al., 2002). In this context, LCNG have emerged as a reliable source of nuclear markers to unravel complicated evolutionary patterns, ranging from the characterizing lineage diversification by genome duplication (e.g., Barker, Vogel, & Schranz, 2009; Franzke et al., 2011), allopolyploid speciation, (Lee, Mummenhoff, & Bowman, 2002; Slotte, Huang, Lascoux, & Ceplitis, 2008), or the dynamics of population demography and recent speciation (Guo et al., 2009; St. Onge et al., 2011). As a result, a vast wealth of molecular resources have become available for less studied lineages in Brassicaceae, which accompanied with the emergence of statistical tools for allele discrimination, offer a practical and cost-effective solution for the search and use of LCNG. The results of this study suggest a high reliability in the search and detection of orthologous LCNG; nevertheless, with different levels of success in their applicability for phylogenetic and phylogeographic studies given the chosen strategy for testing (i.e., avoiding PCR subcloning).

In the present study, the viability of LCNG for exploring alternative sources of molecular markers in *Schizopetalon* was, at least, partially demonstrated using the available resources for Brassicaceae. There was a high success rate for PCR amplification indicating conservation in the flanking exons permitting primer development for the species from both the Atacama and Mediterranean subclades. These results contrast with preconceived ideas about the high specificity of LCNG primers in plants that preclude their widespread use at interspecific levels (Hughes, Eastwood, & Bailey, 2006; Mort & Crawford, 2004; Sang, 2002; Small et al., 2004).

Given that the primers were designed from conserved loci of *A. thaliana* and *B. napus*, two taxonomically distant genera (Lineage I and Lineage III; Franzke et al., 2011), suggests that they will be highly transferable within the family. In fact, most of these primers were useful in genera closely related to *Schizopetalon* (i.e., *Neuontobotrys*, *Menonvillea*, and *Mathewsia*; data not shown), supporting their utility in other Brassicaceae genera. However, a more difficult issue was detected in the high incidence of multiple bands and/or reads from direct sequencing (Table 1), which is fairly common with these markers (e.g., Hughes, Eastwood, & Bailey, 2006; Pillon et al., 2013). For Brassicaceae, this result is not surprising given the high incidence of polyploidy in the evolutionary history of the family (Franzke et al., 2011; Marhold & Lihova´, 2006).

From the set of 21 candidate genes, only four qualified as potential regions for phylogenetic and phylogeographic use. The patterns of variability revealed in the distance network are congruent with previous estimates of phylogenetic relationships (Toro-Núñez et al., 2013). Results revealed separate ingroup and outgroup clusters in the proposed regions (Fig. 3), which suggests two subclades in *Schizopetalon* (Atacama subclade and Mediterranean subclade; Toro-Núñez et al., 2013). Differences detected in the complexity of the networks suggest processes such as reticulation, recombination, and/or incomplete lineage sorting; nevertheless, it was possible to recognize some distinct groups. The phased haplotypes of ELF8 were of limited utility due to the presence of intraspecific recombination events (Table 2). When phased sequences of MMT were analyzed for haplotype diversity, the presence of alleles of different sizes was detected. Despite the successful use of bioinformatic techniques to resolve heterozygous positions (PHASE and CHAMPURU), there was still some uncertainty in the reconstruction of haplotypes. Considering these limitations, the use of more direct techniques would be beneficial (e.g., PCR subcloning) or

direct PCR multiplexing from multiple loci across individuals and populations (e.g., next generation sequencing; Puritz, Addison, & Toonen, 2012).

Evolutionary and phylogeographic scenarios in the Atacama subclade

Despite the described limitations, it is possible to infer phylogeographic patterns for the Atacama subclade. The main organization of the haplotypes suggests at least four discernible groups, which were defined by the presence of exclusive and mixed haplotypes. All the results suggest that the main group (D) is represented by a collection of haplotypes from three of the species analyzed (*S. biseriatum*, *S. tenuifolium*, and *S. rupestre*), whereas groups A, B (*S. rupestre, S. bipinnatifidum*), and C (*S. biseriatum*) represent clusters with unique haplotypes per species.

S. rupestre (groups A, B)

For *S. rupestre*, the organization of most of haplotypes in groups A and B suggests genetic differentiation of this species from the rest of the Atacama subclade (Figs. 2, 3, and 4). Within *S. rupestre,* the two haplotype groups correspond to a latitudinal break in the genetic continuity of the species. Although other sources of data, such as morphology, do not reflect this biogeographic pattern, it is possible to suggest at least two hypotheses. First, *S. rupestre* is the only montane species in the Atacama subclade and has the largest latitudinal distribution of the entire genus (Al-Shehbaz, 1989; Toro-Núñez et al., 2013). This distribution suggests that genetic differentiation between the two haplotype clusters could be due to isolation during the Quaternary, when the Andes become a more effective barrier to migration due to colder conditions and expansion of the perpetual snow-line to lower altitudes (e.g., Ammann, Jenny, Kammer, & Messerli, 2001; Veit, 1993). The second hypothesis is the differentiation of these

southern populations due to geographic isolation and genetic bottlenecks. In either case, in order to advance in the study of these hypotheses, it is necessary to sample populations from Argentina to have completed the entire geographic range of this species.

S. biseriatum (group C)

For *S. biseriatum*, the presence of exclusive haplotypes was detected in group C. The populations identified in this group correspond to the southern portion of the distribution (populations 35 and 39). This pattern is intriguing, given their close occurrence to northern populations of *S. arcuatum* in the coast of Huasco (28 °27' S 71°13'W). This is area is a narrow strip of coastal dunes, delimited by low cliffs where populations of both species occur. These populations are separated by no more than 5 km, but there is a distinctive break in the morphology between the species (Toro, unpubl.). This result suggests high genetic isolation, resulting in a lack of gene flow and reproductive isolation of both species.

S. tenuifolium (group D)

Among all the groups detected in this study, the most intriguing is group D, which represents haplotypes from all species of the Atacama subclade (except *S. bipinnatifidum*), with the largest contribution of haplotypes being from *S. tenuifolium*. The reasons for the presence of a high proportion of haplotypes in this group can be explained by two hypotheses. The first is the retention of ancestral polymorphisms due to the recent divergence and incomplete lineage sorting. In this sense, haplotypes from *S. tenuifolium* could represent ancestral copies from which new haplotypes were derived through range expansion and geographic isolation. A second hypothesis is hybridization of *S. tenuifolium* with *S. biseriatum* and *S. rupestre*. Given that

nuclear loci are comprised of two parental alleles (Sang, 2002; Small et al., 2004), the presence of haplotypes in different groups could be interpreted as an indication of retention of ancestral copies from past or present hybridization events.

While southern populations of *S. biseriatum* (populations 35 and 39) are differentiated and have unique alleles in group C, the northern populations (20, 21and, 22) share alleles with groups C and D. Hybridization might be the cause of this pattern and is likely supported by the presence of contact zones in the delta of Rio Copiapo (27 19S, 70 54W; Fig 8), where the coastal hills are depressed. The presence of other populations having alleles present from both *S. biseriatum* and *S. tenuifolium* in the same area also supports this hypothesis (e.g., populations 12, 13, 15, 22; Fig. 8). Modern hybridization between *S. rupestre* (populations 61, 78, 85, and 101) and *S. tenuifolium* is unlikely, given climatic, geographic, and phenological differences between these taxa. Populations from this area are separated by more than 2000 m of altitude and differ in phenology with *S. tenuifolium* flowering in August to October and *S. rupestre* in December to February. The genetic connection between these two groups might be explained by climatic oscillations during the Quaternary when the Atacama Desert had a more humid and mesic climate (e.g., Contreras et al., 2010; Gayo et al., 2011; Lamy et al., 2000; Latorre et al., 2007). The changes in climate during the Quaternary have been suggested to have profound influence in evolution of lineages, by modifying the ranges of populations and promoting genetic isolation or hybridization (Gengler-Nowak, 2002; Luebert & Wen, 2008). Thus, although *S. tenuifolium* and *S. rupestre* are currently isolated geographically, it is possible that in the past their ranges could have overlapped allowing for hybridization with the retention of shared haplotypes during the return of drier conditions in the Holocene.

The context of past hybridization in the Atacama subclade inferred from MMT reveals an intriguing pattern, which might add new evidence for the dynamic role of aridity in endemic lineages of the Atacama Desert. Notwithstanding the paucity of intraspecific studies in the Atacama Desert, the known phylogeographic patterns plant lineages suggest the expansion, contraction, and fragmentation of populations due to the oscillating climatic conditions (Albrecht et al., 2010; Gengler & Crawford, 2000; Luebert et al., 2009; Ossa et al., 2013; Viruel et al., 2011). For the Atacama subclade of *Schizopetalon*, the presence of shared haplotypes among Andean and coastal populations suggests modification in the configuration of the climatic ecoregions in the southern Atacama Desert (25° -32°), which is a scenario not considered in previous studies or only hypothesized from geological and phytogeographic evidence (e.g., Latorre et al., 2007). Additional inspection of potential habitat suitability using ecological niche modeling (see Appendix 10 for details) during the Quaternary supports the idea of expansion of Andean-like habitats to coastal lower lands (Fig. 9). Under all climatic scenarios (Miroc and CCS3) habitat overlap is inferred in northern distribution for *S. rupestre* (Andes), *S. biseriatum*, and *S. tenuifolium* (coast) (Fig. 9), which could permit past gene flow between ecoregions. This scenario is also suggested from the test of neutrality for the MMT sequences in *S. rupestre* (Table 2), especially under an hypothetical case of expansion given in high negative Fu's F values present in this species (Ramos-Onsins & Rozas, 2002). Nevertheless, this evidence should be considered only tentative until more extensive haplotype sampling of individuals per populations is conducted to confirm these results. Despite this, the results suggest that the influence of climatic divergence between Andean and coast habitats previously suggested (Chapter 2) is rather fragile and susceptible to change during the Quaternary, making plausible the modification

of distribution of species of the Atacama subclade with potential evolutionary consequences for their diversification, such as hybridization.

Table 1 Set of primers for scanned LCNG with positive PCR reaction, their melting temperature, and the result from the screening process. The regions selected and analyzed in the present study are in asterisk (*)

.	Sequence Primer	Melting T (°C)	Single Band	Single Read	Molecular variability (>5%)
ALDH10A8 F	TGT AAC CAC TTC AAT GCA AA	50.3	YES	YES	YES
ALDH10A8 R	TCA CTT GCTTCA CAC TCA AG	52.8			
ATEX070A1 F	TGC AGA AGT TAC AGT GAG CA	54	NO	NO	NO
ATEX070A1 R	GAA CAG CAG TCT TTG TAG CA	52.8			
ATGR R	AAGTTCAACACCAACAGCTT	54.6	YES	NO	NO
ATGR1 F	CTG TGG AGT TTG CAT CAA TA	54.9			
ATprot F	GCT CTG AAG CAC AAG AT	52.7	YES	YES	NO
ATprot R	GTG GCT GCA GTT ACT CTGA	54.6			
BRI1 F	CAA CAC AGC CTG GAG ACT ATA A	57	YES	YES	NO
BRI1 R	CAA ATC TTT CGA ATC CAT CC	57			
CC1 F	GAA TTA TGG GTG GTT CAA AT	48.5	NO	NO	NO
CC1 R	GGG GAG AAT ATC ATC ATT TG	48.1			
CH4 F	TGT TCC GTA TCA TGT TCG AT	51.3	YES	YES	YES
CH4 R	TCA GGA TGG TTC ACT AGC TC	53.7			
ELF8 F (*)	GCA ATG AGA AAA GAA ATC CA	45.7	YES	YES	YES
ELF8 R (*)	TGT CTT TTT GCA CTC TTG C	51.3			
EIFE3F	TWA GGT TCA CAA TCC ATC TC		YES	NO	NO
EIFE3R	GGC ATG TGT GTT TGT CAA	51.4			
FAB1 F	ATG GGG CCA AAC TAT TCT AT	51	YES	NO	NO
FAB1 R	AGG CTC GGT CAT GTG ATA G	53.9			
FASS1 F	TGT CTG CAG GAA CTA ATG GA	56.8	YES	YES	NO
FASS1 R	CCT CTC CCT TGG AGA TCA A	57.7			
GCN3 F	CGT TCA CTT GAT TCA AAA GC	50.2	YES	YES	NO
GCN3 R	TTC ATG CTT ATA CTG CTG AGG	52.2			
MMT R (*)	GAT GTC GAC TTT TCT TGG AT	50.1	YES	YES	YES
MMT R (*)	AAA GAT TGC AGCCTT AAC AG	50.9			
MFP2 F	GTT CCT GGA GTT ACT GAT CG	52.4	NO	NO	NO
MFP2 R	TCT TCC CAA CAT CTA ACA GG	51.6			
PEX5 F	CCA ATG GGT TAA TGA GTT TT	48.5	NO	NO	NO
PEX5 R	CTG CTT CAC TCA GAA GTC CT	54.1			
PGM F	CTG AGA TTA AGG TGG CAG AG	55.1	NO	NO	NO
PGM R	ATT TCT GTC ACC ATC ACC AT	55.1			
PHS2 F	GAA TAT GAA TTG GCT GCA CA	57.1	NO	NO	NO
PHS2 R	KAC ACA TAA GTT AGC CAT TCT C	55.7			
PYD4 F	GTC CAA CTG GAA ATT YAA CG	50.7	NO	NO	NO
PYD4 R	GGC TYT CCT CAC GAT GTC	54.1			
RGD3 F	CTG TGG TCC CTT TTT GAC TT	56.7	NO	NO	NO
RGD3 R	AGA ACA AGA AGC GGA TGA CT	56.5			
THFS F	CAC TAC GAT CTC TAT GGC AAG	55.1	NO	NO	NO
THFS R	GCG AAG ACA AGT AAC AAC CT	54.6			
VHA F	ACA ATT GAC AAT GAC GTT GC	57	NO	NO	NO
VHA R	TGT TGA TGA AAT CTG GAT CG	57			

Table 2. Summary statistics of molecular variation detected in haplotypes obtained from the fourth intron of ELF8 and the ninth and tenth intron from MMT locus. In bold are highlighted statistically significant results in neutrality tests. In parenthesis are described the locations for recombination, for those cases with evidence for intra sequence recombination.

MMT	*S. biseriatum*	*S. tenuifolium*	*S. rupestre*	All (no outgroups)
Number of sequences	52	38	58	146
Number of nucleotides	491	491	625	647
Number of segregating sites (S)	22	12	11	37
Number of haplotypes (h)	11	6	11	27
Haplotype diversity (Hd)	0.666	0.452	0.656	0.735
Nucleotide diversity per site (π)	0.00813	0.00283	0.00042	0.00471
Watterson Index per site (θ)	0.01127	0.00661	0.0055	0.01541
Tajima's D	-0.89420 (P>0.1)	-1.77805 (0.1>P>0.05)	-1.23984 (P>0.1)	
Fu's F	-0.2558(P>0.1)	-0.748(P>0.1)	**-4.578 (P<0.05)**	
Recombination four gamete's test	NO	NO	NO	YES (431,510 - 510,600)
ELF8	*S. biseriatum*	*S. tenuifolium*	*S. rupestre*	All (no outgroups)
Number of sequences	42	22	42	122
Number of nucleotides	369	384	339	384
Number of segregating sites (S)	35	17	18	31
Number of haplotypes (h)	20	12	12	34
Haplotype diversity (Hd)	0.844	0.835	0.828	0.782
Nucleotide diversity per site (π)	0.01387	0.01484	0.01411	0.00162
Watterson Index per site (θ)	0.0251	0.0138	0.01234	0.01967
Tajima's D	-1.55933 (P>0.1)	0.27746 (P>0.1)	0.46333 (P>0.1)	
Fu's F	-1.22459 (P>0.1)	0.58702 (P>0.1)	1.46553 (0.1>P>0.005)	
Recombination four gamete's test	YES (4-188, 188-191, 246-256, 256-260)	YES (204-239, 264-310)	YES (19-48, 48-201, 280-285)	YES (19-48, 48-264, 201-264, 269-273, 310-333)

Table 3. Summary of polymorphism in the ninth and tenth intron of MMT locus in *Schizopetalon* (Atacama subclade + outgroups).

Haplotype																																									
A	C	C	G	A	C	T	G	C	C	G	A	T	C	C	G	C	G	T	G	A	G	G	T	C	G	A	T	T	A	C	G	C	A	A	C	C	G	C	C	C	T
B	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.
C	.	.	.	C	.	.	.	.	.	.	.	C	.	.	A	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.
D	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	T	.	.
E	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.
F	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	A	.	.	C	.	.	.	T	.	A	.
G	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	.	A	.
H	.	.	.	C	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.
I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
J	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.	T	.	C	.	.	.	T	.	.	.
K	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
L	.	.	A	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	G	.	C	.	.	.	.	G	.	.	A	.	.	.	C	.	G	.	T	.	.	.
M	.	.	A	C	.	.	.	.	.	.	.	.	.	T	A	.	.	.	.	G	.	C	.	.	.	.	G	.	.	A	.	.	.	C	.	G	.	T	.	.	.
N	T	.	A	C	A	.	A	.	.	.	.	.	.	.	A	.	.	.	.	G	.	C	.	.	.	.	G	.	.	A	.	.	.	C	.	G	A	T	.	.	.
O	.	.	A	C	A	.	A	.	.	.	.	.	.	.	A	.	.	.	.	G	.	C	.	.	.	.	G	.	.	A	.	.	.	C	.	G	.	T	.	.	.
P	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	A	.	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.
Q	.	.	A	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	G	A	C	.	.	.	.	G	A	.	A	.	.	.	C	.	G	.	T	.	.	.
R	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	G	.	.	.	.	.	.	G	.	.	A	.	.	.	C	.	G	.	T	.	.	.
S	.	.	A	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	A	G	.	.	.	.	.	.	G	.	T	A	.	.	.	C	.	G	.	T	.	.	.
T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	.	.	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.
U	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.	.	.	C	T	.	.	T	.	.	.
V	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	A	.	.	.	.	.	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.
W	.	T	.	C	.	.	.	.	.	T	.	.	.	.	A	A	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.
X	.	T	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.
Y	.	T	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	G	.	T	A	.	.	.	C	.	.	.	T	.	.	C
Z	.	T	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.	.	T	C	.	.	.	T	.	.	.
AA	.	T	.	C	.	A	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.
BB	.	T	.	C	.	A	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	G	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.
CC	.	.	.	C	.	.	.	.	.	.	C	.	.	.	A	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.
DD	.	.	.	C	.	.	.	.	A	.	.	.	.	.	A	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	.	.	.	C	.	.	.	T	.	.	.

Table 4 Haplotype identity obtained from parsimony network analysis. Numbers represent haplotypes detected per group and in parenthesis is number of individuals representing the haplotypes per group.

	Haplogroups				
Populations	A	B	C	D	E
11				6(3)	
12				8(4)	
13				6(3)	
15				2(2)	
16				2(1)	
20			1(1)	3(5)	
21			2(3)	1(1)	
22			1(1)	3(5)	
28				1(1)	
29				6(3)	
30				6(6)	
33				2(1)	
35			6(3)		
39			6(3)		
41				5(4)	2(1)
42		3(2)		5(3)	
48					2(1)
49					2(1)
51					2(1)
61		7(4)		1(1)	
78				2(2)	
83		2(2)			
85		1(1)		3(2)	
87		2(2)			
93		1(1)			
98		3(3)			
101		5(3)		1(1)	
104		6(3)			
105		2(2)			
110	4(2)				
112	6(3)				
115	6(3)				
s252					2(1)
s244-24					2(1)

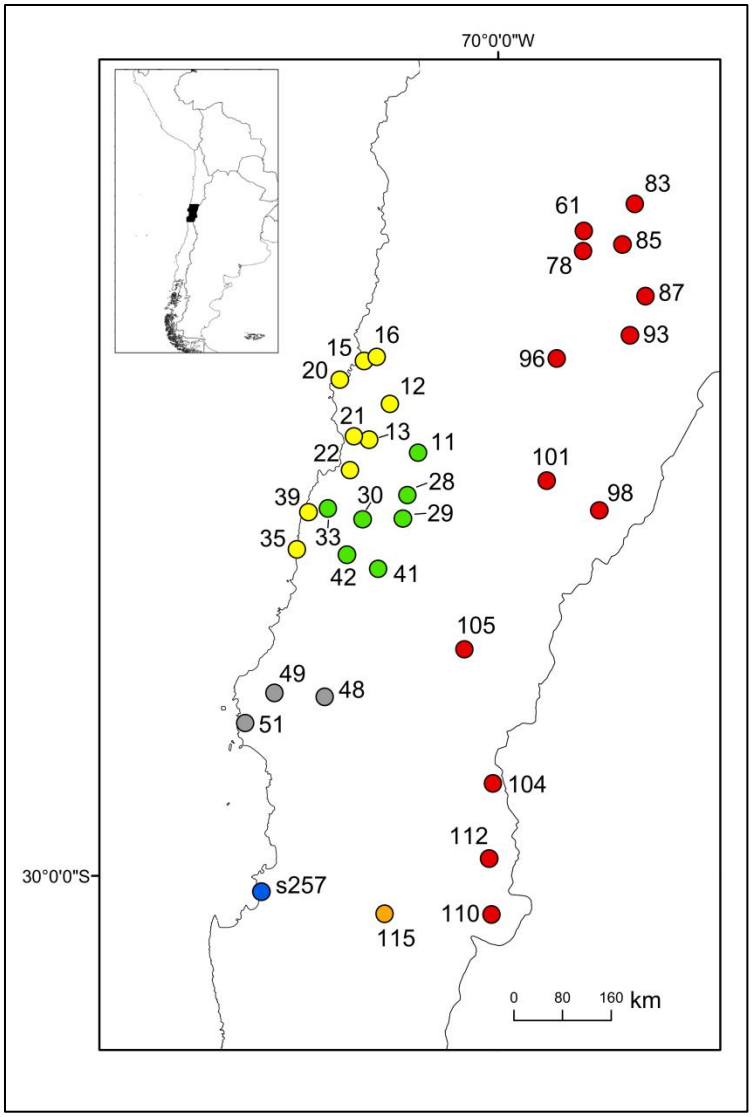


Fig. 1. Distribution of sampled populations used for the analysis of LCNG in *Schizopetalon*: *S. arcuatum* (grey dots), *S. biseriatum* (yellow dots), *S. bipinnatifidum* (orange dots), *S. maritimum* (blue dots), *S. tenuifolioum* (green dots), and *S. rupestre* (red dots). Samples of *S. walkeri* are not represented in this map because uncertainty about its specific locality.

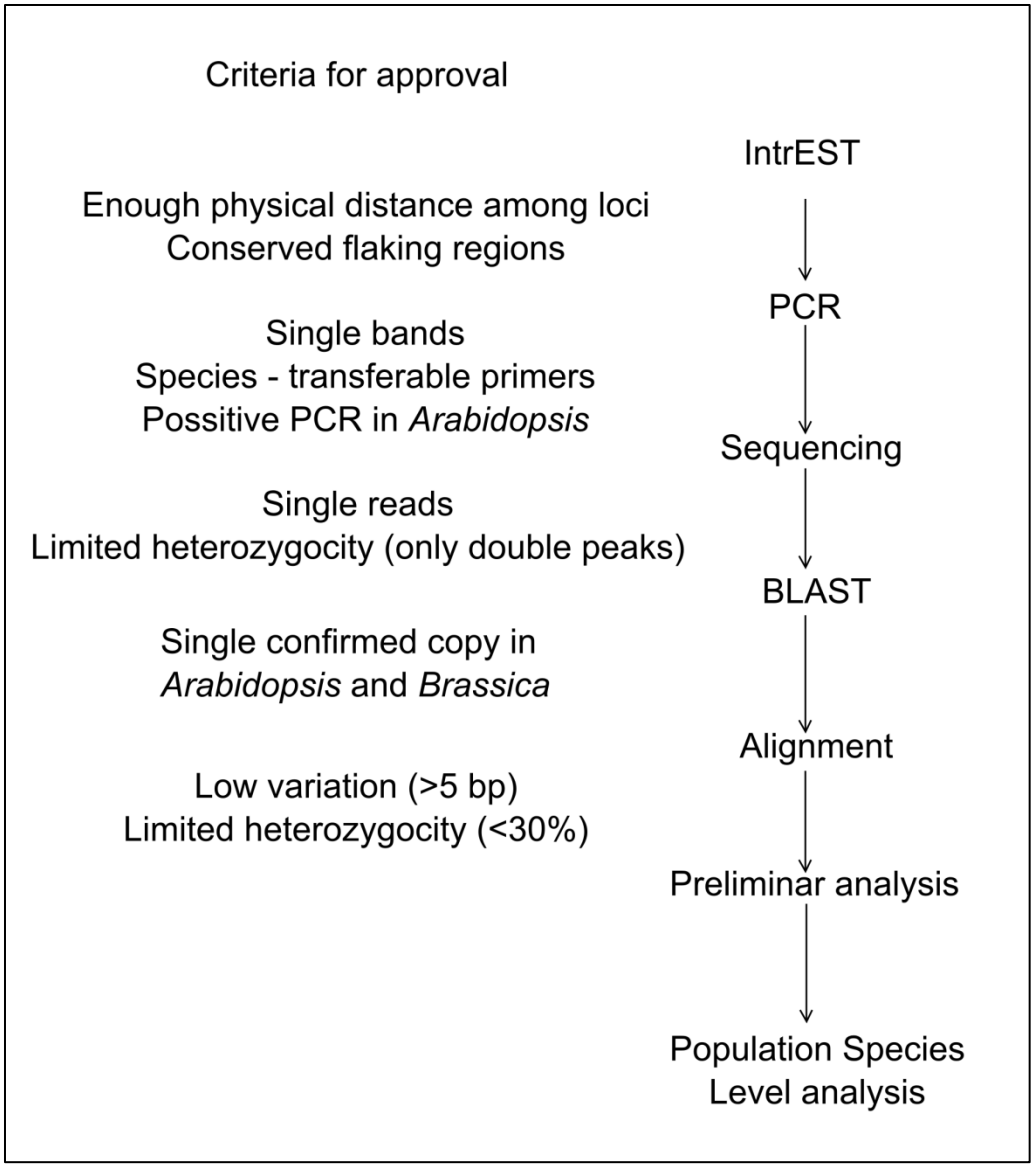


Fig. 2. Flow chart summarizing the step-by-step procedure for detection and testing of orthologous LCNG in *Schizopetalon*. Steps in the analysis are depicted in the right and criteria for passing from one step into another in the left.

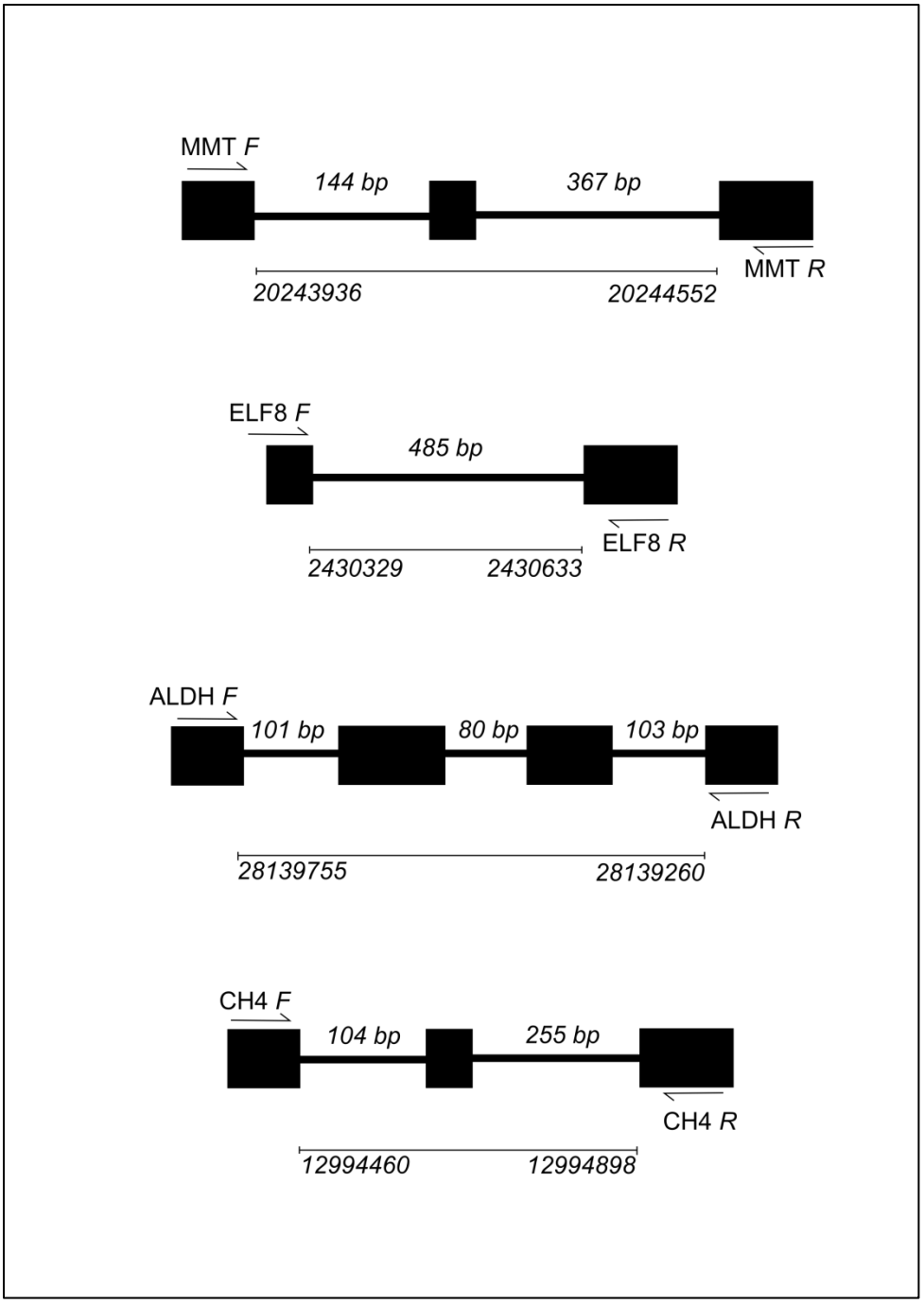


Fig. 3. Diagram of selected LCNG with phylogenetic and phylogeographic potential. Boxes and sticks represent exons and introns, respectively. Numbers above sticks represent intron size in *Schizopetalon* (ingroup and outgroup) and below lines and numbers represent specific initial and end position of introns in the *A. thaliana* genome presented in Genbank.

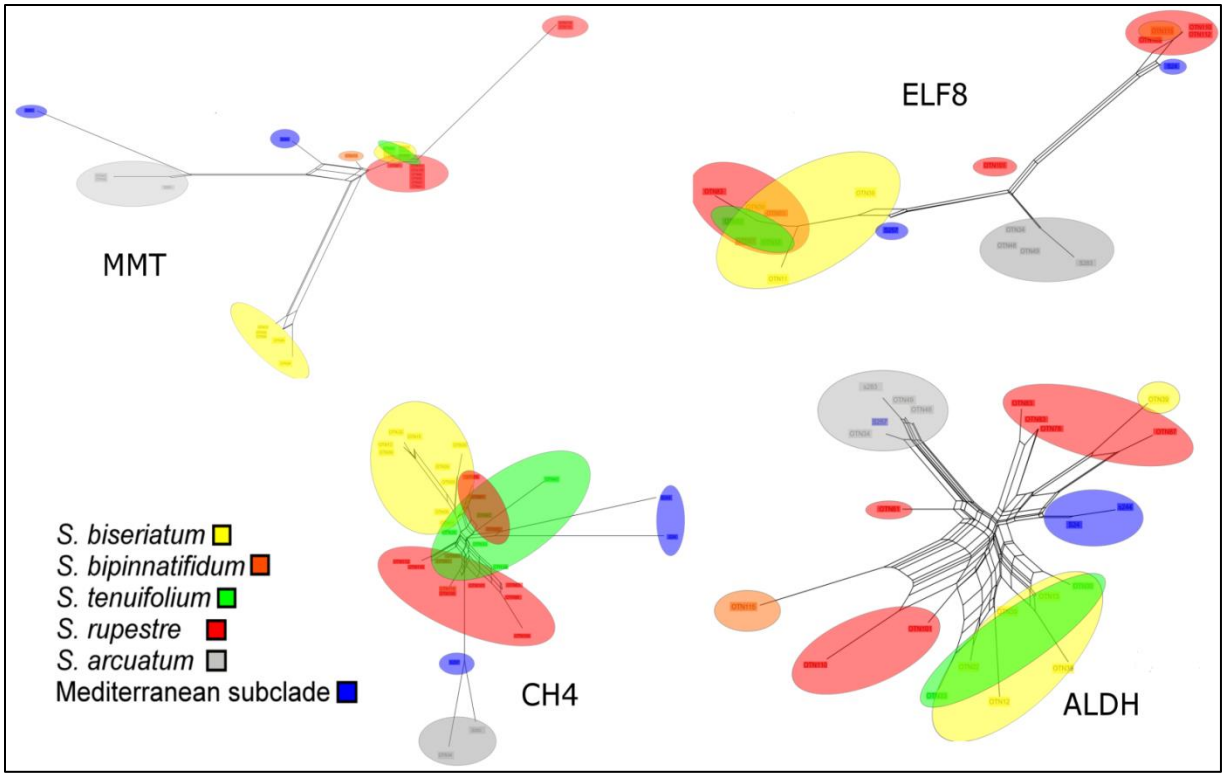


Fig. 4. Neighbor-net networks obtained with unphased sequences from the four selected LCNG. Colored tips represent sequences and colored areas depict suggested associations based on the species color-code. Networks were calculated using uncorrected p-distances.

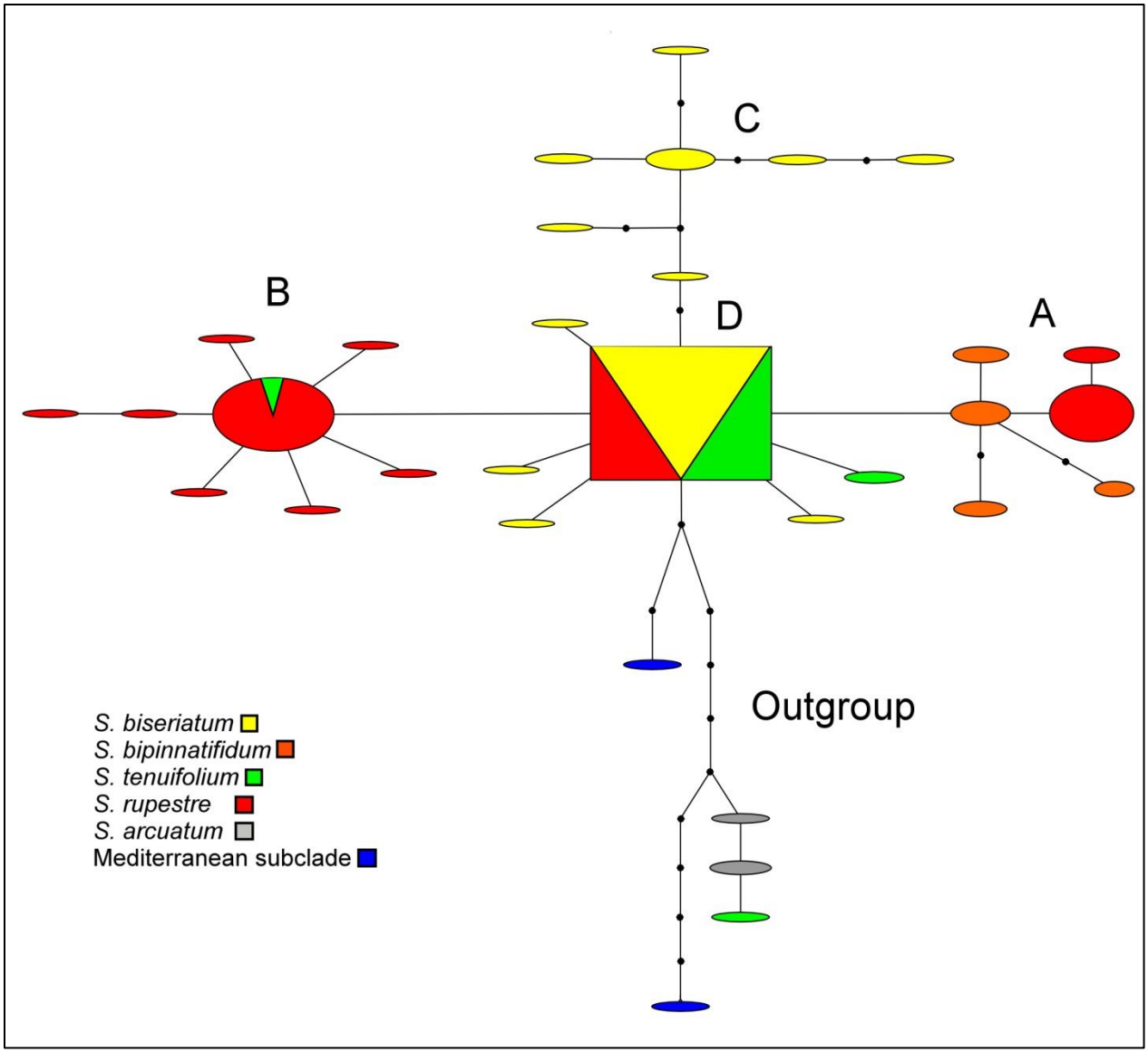


Fig. 5. Parsimony network obtained with the combined ninth and tenth intron from MMT region. Colored circles and square represent haplogroups and their size is proportional to the number of haplotypes that they represent. Small dots depict the necessary number of steps to connect different haplogroups.

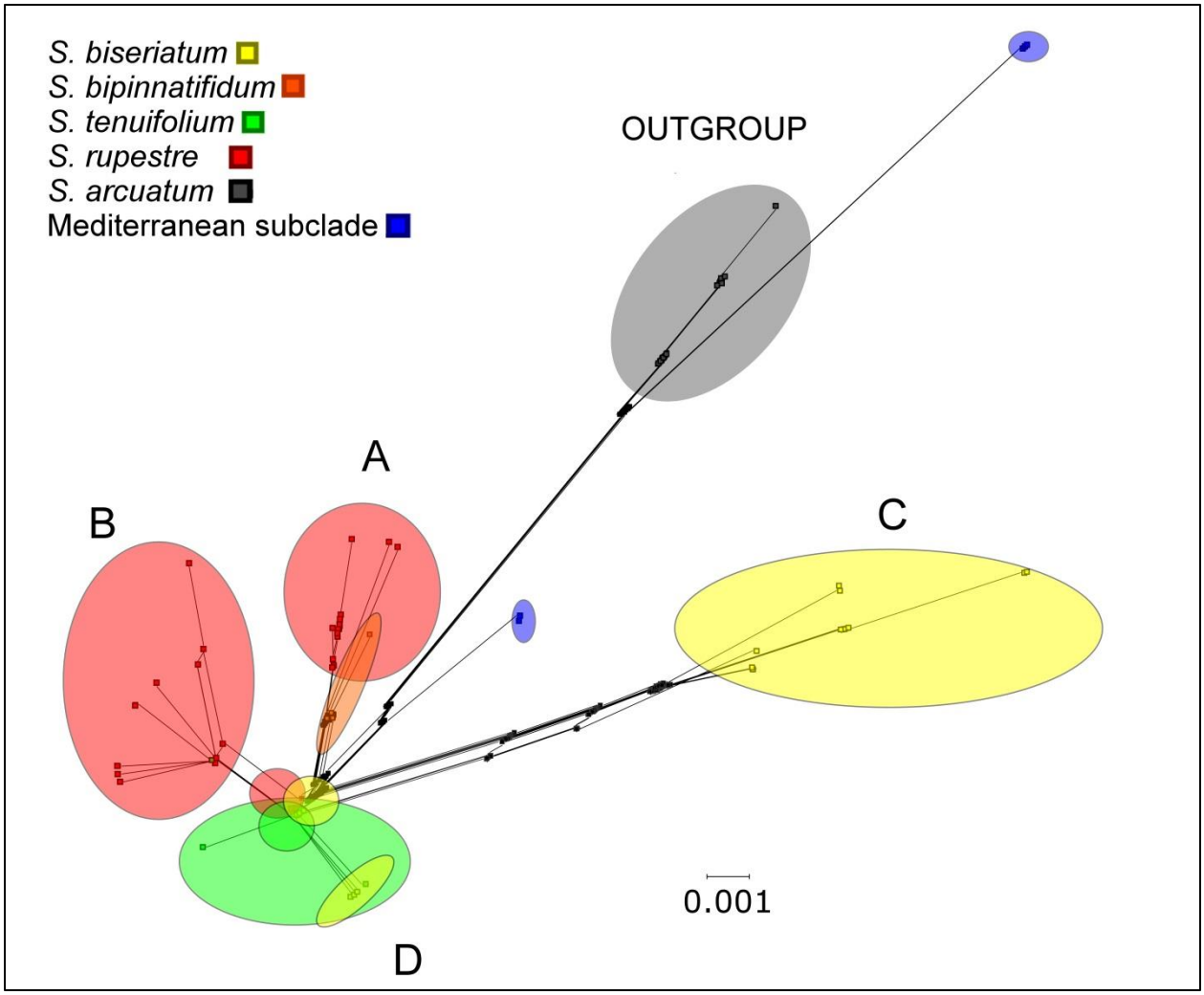


Fig. 6. Neighbor-net network obtained with uncorrected p-distance from MMT haplotypes.

Colored tips depict individual haplotypes and colored areas represent suggested haplogroups.

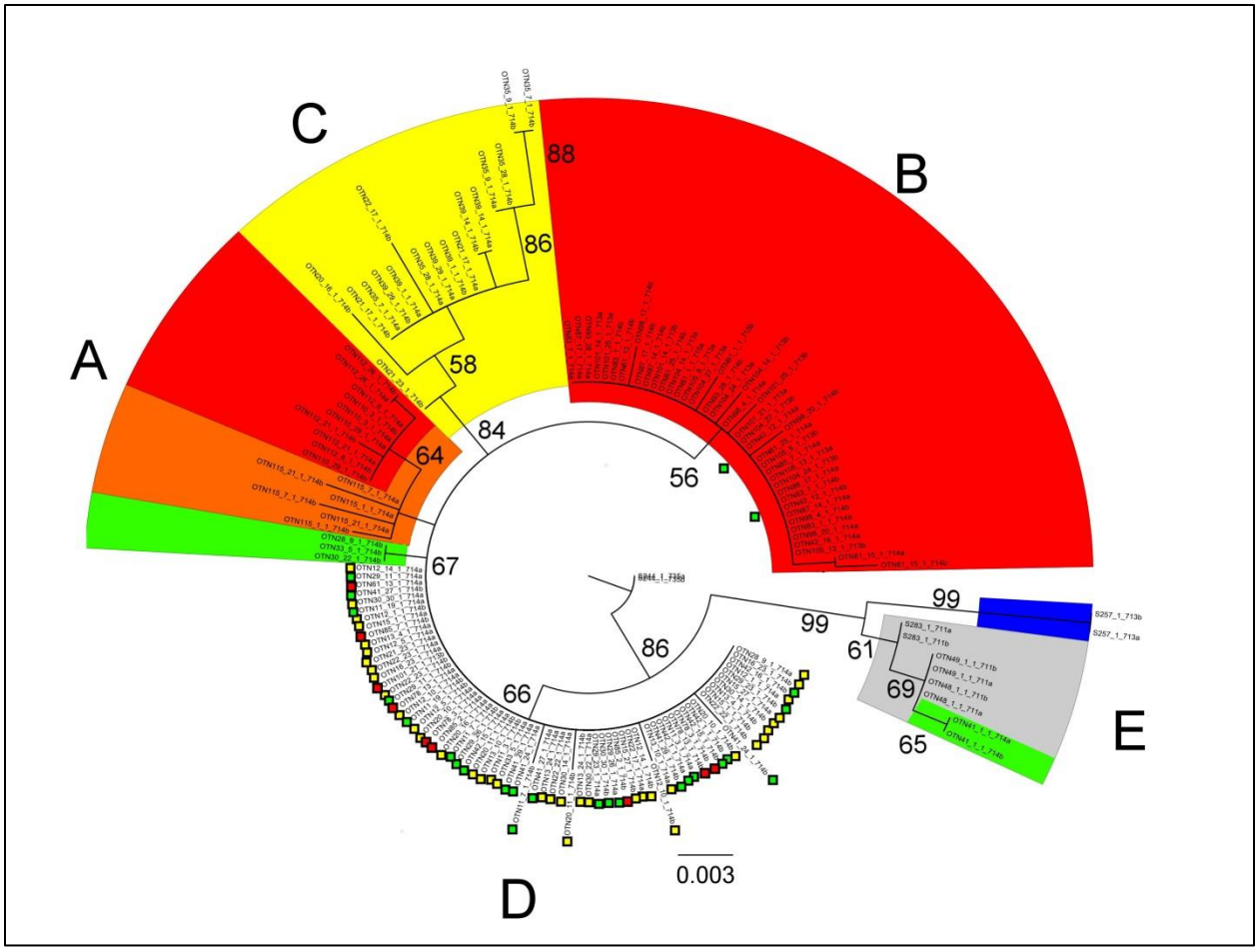


Fig. 7. Circular ML tree obtained with MMT haplotypes. Colored branches are based on species code used in previous figures, *S. biseriatum* (yellow), *S. bipinnatifidum* (orange), *S. tenuifolium* (green), and *S. rupestre* (red). Haplotypes from group D are individually colored. Bootstrap support is represented in numbers above branches.

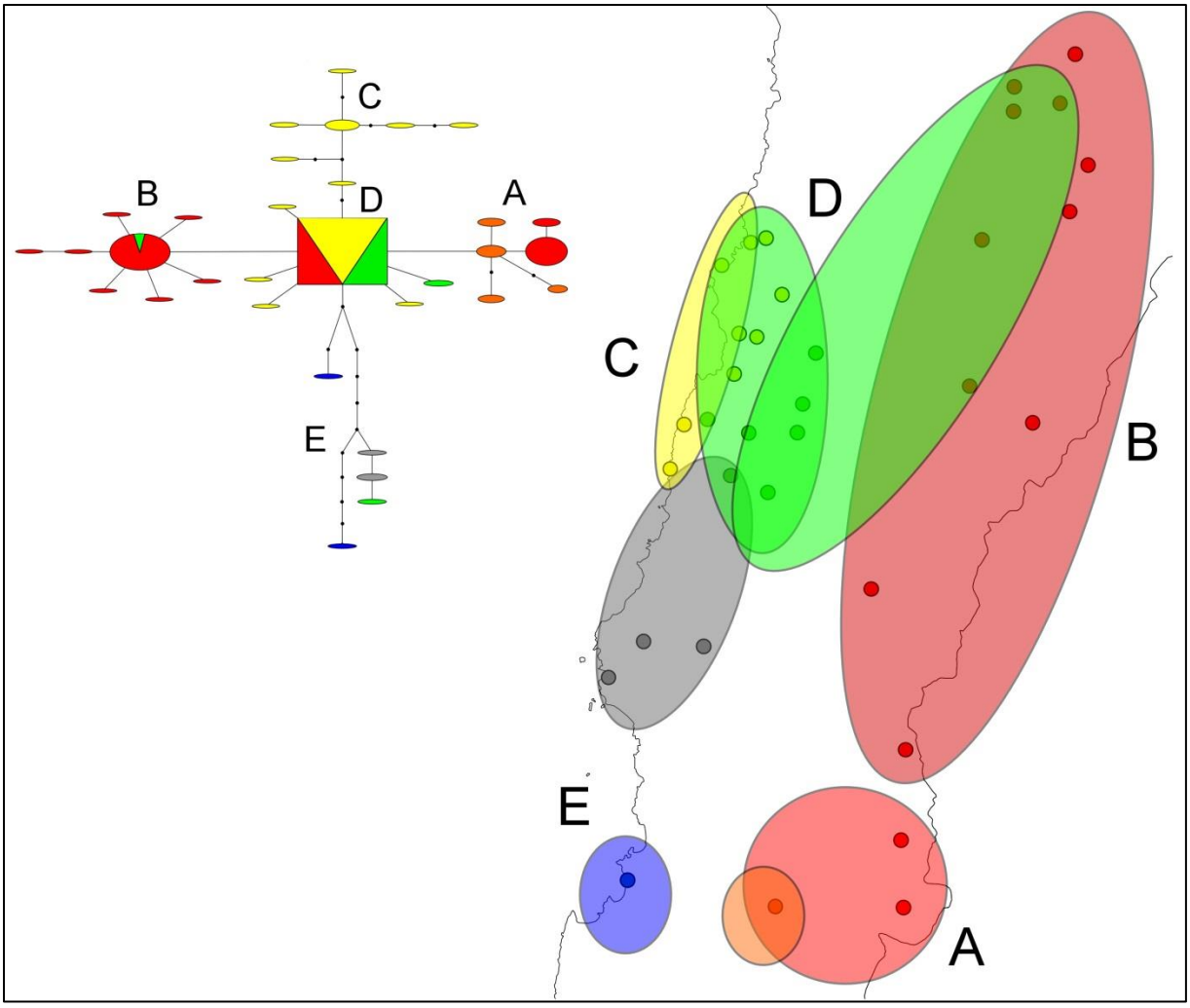


Fig. 8. Geographic distribution of suggested haplogroups, represented in colored ellipses. Populations with shared haplogroups are depicted with overlapped ellipses.

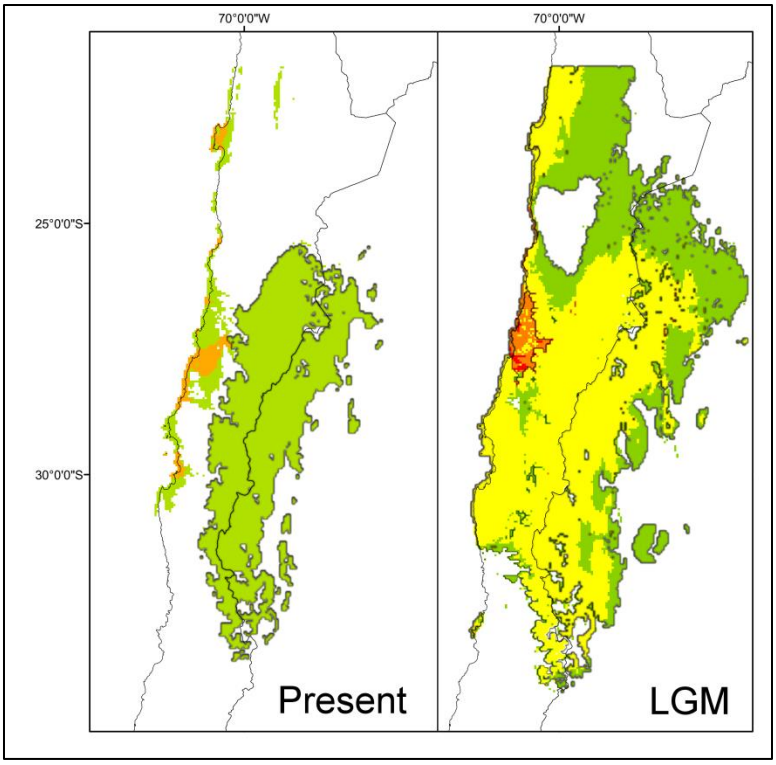


Fig. 9. Estimate of suitable areas using ENM under present and Last Glacial Maximum (LGM) climatic conditions for coastal (*S. biseriatum* + *S. tenuifolium*) and Andean species (*S. rupestre*). Models of distribution were transformed to binary data (absence and presence) using a minimum threshold training probability (see Appendix 10). Colored areas represent overlap among species models. For the present model, orange depicts overlap between coastal species and green areas without overlap. In the LGM model, overlap is colored given the proposals of both past models (MIROC + CCSM): yellow (*S. rupestre*; MIROC + CCSM), orange (*S. biseriatum* + *S. rupestre*; all models), and red (all species; all models).

Bibliography

Abramoff, M. D., Maghalaes, P. J., & Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics International, 11*(7), 36-42.

Al-Shehbaz, I. A. (1989). Systematics and phylogeny of *Schizopetalon* (Brassicaceae). *Harvard Papers in Botany, 1*(1), 10-46.

Al-Shehbaz, I. A. (2010a). Proposal to conserve the name *Sibara* against *Machaerophorus* (Cruciferae). *Taxon, 59*(4), 1287-1287.

Al-Shehbaz, I. A. (2010b). A synopsis of the genus *Sibara* (Brassicaceae). *Harvard Papers in Botany, 15*(1), 139-147.

Al-Shehbaz, I. A. (2012a). A generic and tribal synopsis of the Brassicaceae (Cruciferae). *Taxon, 61*(5), 931-954.

Al-Shehbaz, I. A. (2012b). *Machaerophorus* is united with the genus *Sibara* (Brassicaceae). *Harvard Papers in Botany, 17*, 1-2.

Al-Shehbaz, I. A., Beilstein, M. A., & Kellogg, E. A. (2006). Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. *Plant Systematics and Evolution, 259*, 89-120.

Al-Shehbaz, I. A., German, D. A., Karl, R., Jordon-Thaden, I., & Koch, M. (2011). Nomenclatural adjustmens in the tribe Arabidae (Brassicaceae). *Plant Diversity and Evolution, 129*(1), 71-76.

Albrecht, E., Escobar, M., & Chetelat, R. T. (2010). Genetic diversity and population structure in the tomato-like nightshades *Solanum lycopersicoides* and *S. sitiens*. *Annals of Botany, 105*(4), 535-554.

Alfaro, M. E., & Holder, M. T. (2006). The posterior and the prior in Bayesian phylogenetics. *Annual Review of Ecology, Evolution, and Systematics, 37*, 19-42.

Álvarez, I., Costa, A., & Felliner, G. N. (2008). Selecting single-copy nuclear genes for plant phylogenetics: a preliminary analysis for the Senecioneae (Asteraceae). *Journal of Molecular Evolution, 66*(3), 276-291.

Ammann, C., Jenny, B., Kammer, K., & Messerli, B. (2001). Late Quaternary Glacier response to humidity changes in the arid Andes of Chile (18 - 29°S). *Palaeogeography, Palaeoclimatology, Palaeoecology, 172*, 313-326.

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology, 26*, 32-46.

Angert, A. L., Horst, J. L., Huxman, T. E., & Venable, D. L. (2010). Phenotypic plasticity and precipitation response in Sonoran Desert winter annuals. *American Journal of Botany, 97*(3), 405-411.

Arafeh, R. M. H., Sapir, Y., Shmida, A., Iraki, N., Fragman, O., & Comes, H. P. (2002). Patterns of genetic and phenotypic variation in *Iris haynei* and *I. atrofusca* (Iris sect. *Onocyclus* = the royal irises) along an ecogeographical gradient in Israel and the West Bank *Molecular Ecology, 11*, 39-53.

Armesto, J. J., Vidiella, P. E., & Gutierrez, J. R. (1993). Plant communities of the fog-free coastal desert of Chile: plant strategies in a fluctuating environment. *Revista Chilena de Historia Natural, 66*, 271-282.

Arroyo, M. T. K., Marquet, P., Marticorena, C., Simonetti, J., Cavieres, L., Squeo, F. A., & Rozzi, R. (2004). Chilean winter rainfall-valdivian forest. In Mittermeier R.A., Robles Gil P., Hoffmann M., Pilgrim J., Brooks T., Goettsch C., Lamoreaux J. & G. A. B. Fonseca (Eds.), *Hotspot Revisited* (pp. 99-103). CEMEX, Mexico.

Arroyo, M. T. K., Squeo, F. A., Armesto, J. J., & Villagran, C. (1988). Effects of aridity on plant

diversity in the Northern Chilean Andes: results of a natural experiment. *Annals of the Missouri Botanical Garden, 75*(1), 55-78.

Axelrod, D. I. (1967). Drought, diastrphism, and quantum evolution. *Evolution, 21*(2), 201-209.

Axelrod, D. I. (1972). Edaphic aridity as a factor in Angiosperm evolution. *The American Naturalist, 106*(949), 311-320.

Bailey, C. D., & Doyle, J. D. (1999). Potential phylogenetic utility of the low-copy nuclear gene *PISTILLATA* in dicotyledonous plants: comaprison to nrDNA ITS and trnL intron in *Sphaerocardamum* and other Brassicaceae. *Molecular Phylogenetics and Evolution, 13*(1), 20-30.

Barker, M. S., Vogel, H., & Schranz, M. E. (2009). Paleopolyploidy in the Brassicales: analyses of the *Cleome* transcriptome elucidate the history of genome duplications in Arabidopsis and other Brassicales. *Genome biology and evolution, 1*, 391-399.

Barley, A. J., White, J., Diesmos, A. C., & Brown, R. (2013). The challenge of species delimitation at the extremes: diversification without morphological change in philippine sun skins. *Evolution, Advanced Publication.*

Barnéoud, M. J. M. (1845). Observation sur le groupe des Schizopetalées. *Annales des Sciences Naturelles (a) Botanique. (b) Zoologie. Paris., 3*, 165-168.

Barnéoud, M. J. M. (1846). *Schizopetalon* and *Perreymondia*. *Historia fisica y politica de Chile, Botanica, 1*, 144-150.

Barrett, C. F., & Freudenstein, J. V. (2011). An integrative approach to delimiting species in a rare but widespread mycoheterotrophic orchid. *Molecular Ecology, 20*(13), 2771-2786.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., . . . Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling

and species distribution modeling. *Ecological Modelling, 222*, 1810-1819.

Baum, D. A., Yoon, H.-S., & Oldham, R. L. (2005). Molecular evolution of the transcription factor *LEAFY* in Brassicaceae. *Molecular Phylogenetics and Evolution, 37*, 1-14.

Beilstein, M. A., Al-Shehbaz, I. A., & Kellogg, E. A. (2006). Brassicaceae phylogeny and trichome evolution. *American Journal of Botany, 93*(4), 607-619.

Beilstein, M. A., Al-Shehbaz, I. A., Mathews, S., & Kellogg, E. A. (2008). Brassicaceae phylogeny inferred from *phytochrome A* and *ndhF* sequence data: tribes and trichomes revisited. *American Journal of Botany, 95*(10), 1307-1327.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2010). GenBank. *Nucleid Acid Research, 38*, D46-D51.

Bergsten, J. (2005). A review of long-branch attraction. *Cladistics, 21*, 163-193.

Bremer, K. (1988). The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution, 42*, 795-803.

Bremer, K. (1994). Branch support and tree stability. *Cladistics, 10*, 295-304.

Broenninmann, O., Fitzpatrick, M. C., Pearman, P. B., Petitpierre, B., Pellissier, L., Yoccoz, N. G., . . . Guisan, A. (2012). Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography, 21*, 481-497.

Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology, 22*, 4369-4383.

Cereceda, P., Larrain, H., Osses, P., Lazaro, P., Garcia, J., & Hernández, V. (2000). El factor clima en la floración del desierto en los años “El Niño” 1991 y 1997. *Revista de Geografía del Norte Grande, 27*, 37-52.

Cheng, F., Liu, S., Wu, J., Fang, L., Sun, S., Liu, B., . . . Wang, X. (2011). BRAD, the genetics

and genomics database for *Brassica* plants. *BMC Plant Biology, 11*(1), 136.

Clarke, J. D. E. (2005). Antiquity of aridity in the Chilean Atacama Desert. *Geomorphology, 73*, 101.114.

Clarkson, J. J., Kelly, L. J., Leicht, A. R., Knapp, S., & Chase, M. W. (2010). Nuclear glutamine synthetase evolution in *Nicotiana*: phylogenetics and the origins of allotetraploids and homoploid (diploid) hybrids. *Molecular Phylogenetics and Evolution, 55*, 99-112.

Clausen, J., Keck, D. D., & Hisley, W. M. (1947). Heredity of geographically and ecologically isolated races. *The American Naturalist, 81*(797), 114-133.

Clement, M. D., Posada, D., & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology, 9*(10), 1657-1660.

Comes, H. P. (2004). The Mediterranean region – a hotspot for plant biogeographic research. *New Phytologist, 164*(1), 11-14.

Comes, H. P., & Abbot, R. J. (1999). Population structure and gene flow across arid versus mesic environments: a comparative study of two parapatric *Senecio* species from the Near East *Evolution, 53*(1), 36-54.

Contreras, S., Lange, C. B., Pantoja, S., Lavik, G., Rincón-Martínez, D., & Kuypers, M. M. M. (2010). A rainy northern Atacama Desert during the last interglacial. *Geophysical Research Letters, 37*, L23612.

Cope, J. S., Corney, D., Clark, J. Y., Remagnino, P., & Wilkin, P. (2012). Plant species identification using digital morphometrics. *Expert Systems with Ampplications, 39*, 7562-7373.

Cowling, R. M., Rundel, P. W., Desmet, P. G., & Esler, K. J. (1998). Extraordinary high regional-scale plant diversity in southern African arid lands: subcontinental and global

comparisons. *Diversity and Distributions, 4*(1), 27-36.

Crawford, D. J., Sagastegui-Alva, A., Stuessy, T. F., & Sanchez-Vega, I. (1993). Variacion aloenzimatica en la rara especie endemica peruana *Chuquiraga oblongifolia* (Asteraceae). *Arnaldoa, 1*, 73-76.

Cronn, R. C., Small, R. L., Haselkonr, T., & Wendel, J. F. (2002). Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. *American Journal of Botany, 89*(4), 707-725.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModeltest 2: more models, new heuristics and parallel computing. *Nature Methods, 9*(8), 772.

De Micco, V., & Aronne, G. (2012). Morpho-anatomical traits for plant adaptation to drought. In R. Aroca (Ed.), *Plant Resposnes to Drought Stress* (pp. 37 - 61). Berlin Heidelberg: Springer-Verlag.

di Castri, F., & Hajek, E. R. (1976). *Bioclimatologia de Chile*. Santiago, Chile: Ediciones Universidad Católica de Chile.

Dillon, M. O., & Rundel, P. W. (1990). The botanical response of the Atacama and Peruvian Desert floras to the 1982-1983 El Niño event. In Glynn P.W. (Ed.), *Global Ecological Consequences of the 1982-1983 El Nino-Southern Oscillation* (pp. 487-504). Amserdam; New York: Elsiever.

Dillon, M. O., Tu, T., Xie, L., Quipuscoa Silvestre, V., & Wen, J. (2009). Biogeographic diversification in *Nolana* (Solanaceae), a ubiquitous member of the Atacama and Peruvian Deserts along the western coast of South America. *Journal of Systematics and Evolution, 47*(5), 457-476.

Doyle, J. J., & Doyle, J. D. (1987). A rapid DNA isolation procedure for small quantities of

freash leaf tissue. *Phytochemical Bulletin, 19*, 11-15.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution, .Advance online publication.*

Dunai, T. J., González López, G. a., & Juez-Larré, J. (2005). Oligocene–Miocene age of aridity in the Atacama Desert revealed by exposure dating of erosion-sensitive landforms. *Geology, 33*(4), 321-324.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research, 35*(5), 1792-1797.

Ehlringer, J. R. (1988). Changes in leaf characteristics of species along elevational gradients in Weatch Front, Utah. *American Journal of Botany, 75*(5), 680-689.

Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review in Ecology, Evolution, and Systematics, 40*, 677-697.

Esselstyn, J. A., Evans, B. J., Sedlock, J. L., Anwarali Khan, F. A., & Heaney, L. R. (2012). Single-locus species delimitation: a test of the mixed Yule-coalescent model, with an empirical application to Philippine round-leaf bats. *Proceedings of the Royal Society B: Biological Sciences.* Advance online publication

Evans, M. E. K., Smith, S. a., Flynn, R. S., & Donoghue, M. J. (2009). Climate, niche evolution, and diversification of the "bird-cage" evening primroses (Oenothera, sections Anogra and Kleinia). *The American Naturalist, 173*(2), 225-240.

Evenari, M. (1985). The Desert Environment. In M. Evenari, I. Noy-Meir & D. W. Goodall (Eds.), *Hot Deserts and Arid Shrublands - Ecosystems of the World* (Vol. 12A, pp. 1-22).

Amsterdam: Elsevier Science Publishers.

Ezard, T., Fujisawa, T., & Barraclough, T. C. (2013). splits: Species' Limits by Threshold Statistics. R package version 1.0-18/45. http://R-Forge.R-project.org/projects/splits/.

Fabricante, I., Oesterheld, M., & Paruelo, J. M. (2009). Annual and seasonal variation of NDVI explained by current and previous precipitation across Northern Patagonia. *Journal of Arid Environments, 73*(8), 745-753.

Farris, J. S., Källersjö, M., Kluge, A. G., & Bult, C. (1994). Testing significance of incongruence. *Cladistics, 10*(3), 315-319.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution, 39*, 783-791.

Flot, J.-F. (2007). CHAMPURU 1.0: a computer software for unraveling mixtures of two DNA sequences of unequal lengths. *Molecular Ecology Resources, 7*(6), 974-977.

Flot, J.-F. (2010). SEQPHASE: a web tool for interconverting PHASE input/output files and FASTA sequence alignments. *Molecular Ecology Resources, 10*, 162-166.

Flot, J.-F., Tillier, A., Samadi, S., & Tillier, S. (2007). Phase determination from direct sequencing od length-variable DNA regions. *Molecular Ecology Resources, 6*(3), 627-630.

Franzke, A., Lysak, M. A., Al-Shehbaz, I. A., Koch, M. A., & Mummenhoff, K. (2011). Cabbage family affairs: the evolutionary history of Brassicaceae. *Trends in Plant Science, 16*(2), 108-116.

Fujisawa, T., & Barraclough, T. C. (2013). Delimiting species using single-locus and the generalized mixed Yule coalescent approach: a revised method and evaluation on simulated data sets. *Systematic Biology. Advance online publication.*

Galloway, G. L., Malmberg, R. L., & Price, R. A. (1998). Phylogenetic utility of the nuclear gene Arginine Decarboxylase: An example from Brassicaceae. *Molecular Biology and Evolution, 15*(10), 1312-1320.

Garaventa, A. (1940). El genero *Mathewsia* en Chile. *Revista Universitaria, 25*, 255-267.

Garreaud, R. D., Molina, A., & Farias, M. (2010). Andean uplift, ocean cooling and Atacama hyperaridity: A climate modeling perspective. *Earth and Planetary Science Letters, 292*(1-2), 39-50.

Garrick, R. C., Sunnucks, P., & Dyer, R. J. (2010). Nuclear gene phylogeography using PHASE: dealing with unresolved genotypes, lost alleles, and systematic bias in parameter estimation. *BMC Evolutionary Biology, 10*, 118.

Gatesy, J., O'Grady, P., & Baker, R. H. (1999). Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level Artiodactyl taxa. *Cladistics, 15*, 271-313.

Gayo, E. M., Latorre, C., Santoro, C. M., Maldonado, A., & De Pol-Holz, R. (2011). Hydroclimate variability in the low-elevation Atacama Desert over the last 2500 years. *Climate of the Past Discussions, 7*, 3165-3202.

Gengler-Nowak, K. M. (2002). Reconstruction of the biogeographical history of Malesherbiaceae. *The Botanical Review, 68*, 171-188.

Gengler-Nowak, K. M. (2003). Molecular phylogeny and taxonomy of Malesherbiaceae. *Systematic Botany, 28*(2), 333-344.

Gengler, K. M., & Crawford, D. J. (2000). Genetic diversities of four little-known species of *Malesherbia* (Malesherbiaceae) endemic to the arid inter-Andean valleys of Peru. *Brittonia, 52*(4), 303-310.

German, D. A., & Al-Shehbaz, I. A. (2010). Nomenclatural novelties in miscellaneous Asian Brassicaceae. *Nordic Journal of Botany, 28*, 646-651.

Gilg, E., & Muschler, R. (1909). Aufzählung aller zur Zeit bekannten Südamerikanischen Cruciferen. *Botanische Jahrbücher fur Systematik, Pflanzengeschichte und Pflanzengeographie, 42*, 437-487.

Godsoe, W. (2010). Regional variation exaggerates ecological divergence in niche models. *Systematic Biology, 59*(3), 298-306.

Goloboff, P. A., Farris, J. S., & Nixon, K. C. (2008). TNT, a free program for phylogenetic analysis. *Cladistics, 24*(5), 774-786.

Graham, C. H., Parra, J. L., Rahbek, C., & McGuire, J. A. (2009). Phylogenetic structure in tropical hummingbird communities. *Proceedings of the National Academy of Sciences, 106*(Supplement 2), 19673-19678.

Graham, E. B. (2005). Genetic diversity and crossing relationships of *Lycopersicon chilense*, PhD Dissertation. University of California, Davis, USA.

Grant, V. (1981). *Plant Speciation*. New York: Columbia University Press.

Guerrero, P. C., Arroyo, M. T. K., Bustamante, R. O., Duarte, M., Hagemann, T. K., & Walter, H. E. (2011a). Phylogenetics and predictive distribution modeling provide insigths in the the geographic divergence of *Eriocyse* subgen. *Neoporteria* (Cactaceae). *Plant Systematics and Evolution, 297*, 113-128.

Guerrero, P. C., Durán, A. P., & Walter, H. E. (2011b). Latitudinal and altitudinal patterns of the endemic cacti from the Atacama desert to Mediterranean Chile. *Journal of Arid Environments, 75*(11), 991-997.

Guerrero, P. C., Rosas, M., Arroyo, M. T. K., & Wiens, J. J. (2013). Evolutionary lag times and

recent origin of the biota of an ancient desert (Atacama-Sechura). *Proceedings of the National Academy of Sciences*. Advance online publication.

Guindon, S., & Gauscuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology, 52*, 696-704.

Guo, Y.-L., Bechsgaard, J. S., Slotte, T., Neuffer, B., Lascoux, M., Weigel, D., & Schierup, M. H. (2009). Recent speciation of *Capsella rubella* from *Capsella grandiflora*, associated with loss of self-incompatibility and an extreme bottleneck. *Proceedings of the National Academy of Sciences of the United States of America, 106*(13), 5246-5251.

Gutiérrez, J. R., Arancio, G., & Jaksic, F. M. (2000). Variation in vegetation and seed bank in a Chilean semi-arid community affected by ENSO 1997. *Journal of Vegetation Science, 11*(5), 641-648.

Gutiérrez, J. R., & Meserve, P. L. (2003). El Niño effects on soil seed bank dynamics in north-central Chile. *Oecologia, 134*(4), 511-517.

Gutterman, Y. (2002). *Survival Strategies of Annual Desert Plants*. Berlin Heidelberg: Springer-Verlag.

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

Hammer, Ø., Harper, D. A. T., & Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica, 4*(1), 1-9.

Hartley, A. J. (2003). Andean uplift and climate change. *Journal of the Geoloical Society, 160*, 7-10.

Hartley, A. J., & Chong, G. (2002). Late Pliocene age for the Atacama Desert: Implications for the Desertification of western South America. *Geology, 30*(1), 43-46.

Hartley, A. J., Chong, G., Houston, J., & Mather, A. E. (2005). 150 millions of climatic stability, evidence from Atacama Desert. *Journal of Geological Society of London, 162*, 421-424.

Heibl, C., & Calenge, C. (2013). Phyloclim: integrating phylogenetics and climatic niche modeling. R package version 0.9-3. http://CRAN.R-project.org/package=phyloclim.

Heibl, C., & Renner, S. S. (2012). Distribution models and a dated phylogeny for Chilean *Oxalis* species reveal occupation of new habitats by different lineages, not rapid adaptive radiation. *Systematic Biology, 61*(5), 823-834.

Hershkovitz, M. A., Arroyo, M. T. K., Bell, C., & Hinojosa, L. F. (2006). Phylogeny of *Chaetanthera* (Asteraceae: Mutisieae) reveals both ancient and recent origins of the high elevation lineages. *Molecular Phylogenetics and Evolution, 41*, 594-605.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology, 25*, 1965-1978.

Hijmans, R. J., & van Etten, J. (2012). Raster: geographic analysis and modeling with raster data. R package version 2. 031. http://CRAN.R-project.org/package=raster. Retrieved from http://CRAN.R-project.org/package=raster

Holmgren, M., Stapp, P., Dickman, C. R., Gracia, C., Graham, S., Gutierrez, J. R., . . . Squeo, F. A. (2006). Extreme climatic events shape arid and semiarid ecosystems. *Frontiers in Ecology and the Environment, 4*(2), 87-95.

Housman, D. C., Price, M. V., & Redak, R. A. (2002). Architecture of coastal and desert *Encelia farinosa* (Asteraceae): consequences of plastic and heritable variation in leaf characters. *American Journal of Botany, 89*(8), 1303-1310.

Houston, J. (2006). Variability of precipitation in the Atacama Desert: its causes and hydrological

impact. *International Journal of Climatology, 26*, 2181-2198.

Houston, J., & Hartley, A. J. (2003). The central Andean west-slope rainshadow and its potential contribution to the origin of hyper-aridity in the Atacama Desert. *International Journal of Climatology, 23*, 1453-1464.

Howarth, D. G., & Baum, D. A. (2005). Genealogical evidence of homoploid hybrid speciation in an adaptive radiation of *Scaveola* (Goodeniaceae) in the Hawaiian islands. *Evolution, 59*(5), 948-961.

Huang, C.-C., Hung, K.-H., Wang, W.-K., Ho, C.-W., Huang, C.-L., Hsu, T.-W., . . . Chiang, T.-Y. (2012). Evolutionary rates of commonly used nuclear and organelle markers of Arabidopsis relatives (Brassicaceae). *Gene, 499*(1), 194-201.

Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogeny. *Bioinformatics, 17*, 754-755.

Hughes, C. E., Eastwood, R. J., & Bailey, C. D. (2006). From famine to feast? Selecting nuclear DNA sequence loci for plant species-level phylogeny reconstruction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 361*, 211-225.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution, 23*(2), 254-267.

Ilut, D. C., & Doyle, J. J. (2012). Selecting nuclear sequences for fine detail molecular phylogenetic studies in plants: A computational approach and sequence repository. *Systematic Botany, 37*(1), 7-14.

Jaksic, F. M. (1998). The multiple facets of El Niño/Southern Oscillation in Chile. *Revista Chilena de Historia Natural, 71*, 121-131.

Jaksic, F. M. (2001). Ecological effects of El Niño in terrestrial ecosystems of Western South

America. *Oikos, 24*(3), 241-250.

Jara, P. A. (2010). Reconstruccion filogenetica del género endémico *Leucocoryne* (Alliaceae) y su correspondencia biogeografica con la aridización de la zona mediterránea y semiárida de Chile. Doctoral Dissertation. Universidad de Chile, Santiago, Chile.

Julia, C., Montecinos, S., & Maldonado, A. (2008). Caracteristicas climaticas de la region de Atacama. In F. A. Squeo, F. M. Arancio & J. R. Gutierrez (Eds.), *Libro rojo de la flora nativa y de los sitios prioritarios para su conservacion: región de Atacama* (pp. 25-42). La Serena, Chile: Universidad de la Serena.

Katinas, L., & Crisci, J. V. (2000). Cladistic and biogeographic analyses of the genera *Moscharia* and *Polyachyrus* (Asteraceae, Mutisieae). *Systematic Botany, 25*(1), 33-46.

Kearney, M., & Porter, W. (2009). Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters, 12*(4), 334-350.

Kim, S., Plagnol, V., Hu, T. T., Toomajian, C., Clark, R. M., Ossowski, S., . . . Nordborg, M. (2007). Recombination and linkage disequilibrium in Arabidopsis thaliana. *Nature Genetics, 39*(9), 1151-1155.

Kim, T. K., Sulton, S. E., & Donoghue, M. J. (2008). Allopolyploid speciation in *Persicaria* (Polygonaceae): insights from a low-copy nuclear region. *Proceedings of the National Academy of Sciences, 105*(34), 12370–12375.

Kinkaid, D. T., & Schneider, R. B. (1983). quantification of leaf shape with a microcomputer and Fourier transformation. *Canadian Journal of Botany, 61*, 2333-2342.

Klak, C., Reeves, G., & Hedderson, T. (2004). Unmatched tempo of evolution in Southern African semi-desert ice plants. *Nature, 427*, 63-65.

Knight, C. A., & Ackerly, D. D. (2003). Evolution and plasticity of photosynthetic thermal

tolerance, specific leaf area and leaf size. *New Phytologist, 160*, 337-347.

Koch, M., Al-Shehbaz, I. A., & Mummenhoff, K. (2003). Molecular systematics, evolution and population biology in the Mustard family (Brassicaceae). *Annals of the Missouri Botanical Garden, 90*(2), 151-171.

Koch, M. A., Haubold, B., & Mitchell-Olds, T. (2000). Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). *Molecular Biology and Evolution, 17*(10), 1483-1498.

Kuittinen, H., Aguadé, M., Charlesworth, D., Haan, A. D. E., Lauga, B., Mitchell-Olds, T., . . . Savolainen, O. (2002). Primers for 22 candidate genes for ecological adaptations in Brassicaceae. *Molecular Ecology Notes, 2*(3), 258-262.

Lambkin, C. L., Lee, M. S. Y., Winterton, S. L., & Yeates, D. K. (2002). Partitioned Bremer support and multiple trees. *Cladistics, 18*, 436-444.

Lamesch, P., Berandini, T. Z., Li, D., Swarbreck, D., Wilkis, C., Sasidharan, R., . . . Huala, E. (2011). The *Arabidopsis* information resource (TAIR): improved gene annotation and new tools. *Nucleic acids research, 40*, D1202-D1210.

Lamy, F., Klump, J., Hebbeln, D., & Wefer, G. (2000). Late Quaternary rapid climate change in northern Chile. *Terra Nova, 12*, 8-13.

Latorre, C., Moreno, P. I., Vargas, G., Maldonado, A., Villa-Martínez, R., Armesto, J. J., . . . Grosjean, M. (2007). Late Quaternary environments and palaeoclimate. In W. Gibbons & T. Moreno (Eds.), *The Geology of Chile* (pp. 311-330). London: London Geological Society.

Lee, J.-Y., Mummenhoff, K., & Bowman, J. L. (2002). Allopolyploidization and evolution of species with reduced floral structures in *Lepidium* L. (Brassicaceae). *Proceedings of the*

National Academy of Sciences, 99(26), 16835-16840.

Lemmon, A. R., Brown, J. M., Stanger-Hall, K., & Moriarty Lemmon, E. (2009). The Effect of Ambiguous Data on Phylogenetic Estimates Obtained by Maximum Likelihood and Bayesian Inference. *Systematic Biology, 58*(1), 130-145.

Levin, D. A. (1993). Local speciation in plants: the rule not the exception. *Systematic Botany, 18*(2), 197-208.

Li, N., & Stephens, M. (2003). Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics, 165*, 2213 - 2233.

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics, 25*, 1451-1452.

Luebert, F. (2010). *Systematics, ecology and evolution of Heliotropium sect. Cochranea (Heliotropiaceae) and the biogeography of the Atacama Desert.* Doctor rerum naturalim, Freire Universität, Berlin, Germany Retrieved from http://edocs.fu-berlin.de/diss/receive/FUDISS_thesis_000000020418

Luebert, F. (2011). Hacia una fitogeografía histórica del Desierto de Atacama. *Revista de Geografía del Norte Grande, 50*, 105-133.

Luebert, F., & Pliscoff, P. (2006). *Sinopsis bioclimatica y vegetacional de Chile*. Santiago de Chile: Editorial Universitaria.

Luebert, F., & Wen, J. (2008). Phylogenetic analysis and evolutionary diversification of *Heliotropium* sect . *Cochranea* (Heliotropiaceae) in the Atacama Desert. *Systematic Botany, 33*(2), 390-402.

Luebert, F., Wen, J., & Dillon, M. O. (2009). Systematic placement and biogeographical relationships of the monotypic genera *Gypothamnium* and *Oxyphyllum* (Asteraceae:

Mutisioideae) from the Atacama Desert. *Botanical Journal of the Linnean Society, 159*, 32-51.

Magyar, G., Kun, A., Oborny, B., & Stuefer, J. F. (2007). Importance of plasticity and decision making strategies for plant resource acquisition in spatio-temporally variable environments. *Plant Phytologist, 174*, 182-193.

Malcomber, S. T. (2002). Phylogeny of *Gaertnera* Lam. (Rubiaceae) based on multiple DNA markers: evidence of a rapid radiation in a widespread morphologically diverse genus. *Evolution, 56*(1), 42-57.

Maldonado, A., & Villagrán, C. (2002). Paleoenvironmental changes in the semiarid coast of Chile (~32°S) during the last 6200 cal years inferred from a swamp–forest pollen record. *Quaternary Research, 58*, 130-138.

Marcussen, T., Jakobsen, K. S., Danihelka, J., Ballard, H. E., Blaxland, K., Brysting, A. K., & Oxelman, B. (2012). Inferring species networks from gene trees in high-polyploid North American and Hawaian violets (*Viola*, Violaceae). *Systematic Biology, 61*(1), 107-126.

Marhold, K., & Lihova´, J. (2006). Polyploidy, hybridization and reticulate evolution: lessons from the Brassicaceae. *Plant Systematic and Evolution, 259*, 143-174.

Marquet, P. A., Bozinovic, F., Bradshaw, G. A., Cornelius, C., Gonzalez, H., Gutierrez, J. R., . . . Jaksic, F. M. (1998). Los ecosistemas del desierto de Atacama y area andina adyacente en el norte de Chile. *Revista Chilena de Historia Natural, 71*, 593-617.

McLellan, T. (1993). The roles of heterochrony and heteroblasty in the diversification of leaf shapes in *Begonia dregei* (Begoniaceae). *American Journal of Botany, 80*, 796-804.

McLellan, T., & Endler, J. A. (1998). The relative success of some methods for measuring and describing the shape of complex objects. *Systematic Biology, 47*(2), 264-281.

Mort, M. E., Archibald, J. K., Randle, C. P., Levsen, N. D., O'Leary, R. T., Topalov, K., . . . Crawford, D. J. (2007). Inferring phylogeny at low taxonomic levels: utility of rapidely evolving cpDNA and nuclear ITS locus. *American Journal of Botany, 94*(2), 173-183.

Mort, M. E., & Crawford, D. J. (2004). The continuing search: low-copy nuclear sequences for lower-level plant molecular phylogenetic studies. *Taxon, 53*(2), 257-261.

Mort, M. E., Soltis, D. E., Soltis, P. S., Francisco-Ortega, J., & Santos-Guerra, A. (2001). Phylogenetic relationships and evolution of Crassulaceae inferred from matK sequence data. *American Journal of Botany, 88*(1), 76-91.

Müller, K. F. (2005). SeqState - primer design and sequence statistics for phylogenetic DNA data sets. *Applied Bioinformatics, 4*, 65-69.

Mummenhoff, K., Al-shehbaz, I. A., Bakker, F. T., Linder, H. P., Bakkerf, F. T., & Miihlhausen, A. (2005). Morphological evolution, and speciation of endemic Brassicaceae genera in the Cape flora of southern Africa. *Missouri Botanical Garden, 92*(3), 400-424.

Nakazato, T., Warren, D. L., & Moyle, L. C. (2010). Ecological and geographic modes of species divergence in wild tomato. *American Journal of Botany, 97*(4), 680-693.

Naumann, J., Symmank, L., Samain, M.-S., Müller, K. F., Neinhuis, C., dePamphilis, C. W., & Wanke, S. (2011). Chasing the hare - evaluating the phylogenetic utility of a single copy region at and below species levels within the species rich group *Peperomia* (Piperaceae). *BMC Evolutionary Biology, 11*, 357.

Navarro-Gonzalez, R., Rainey, F. A., Molina, P., Bagaley, D. R., Hollen, B. J., de la Rosa, J., . . . McKay, C. P. (2003). Mars-like soils in the Atacama Desert, Chile, and the dry limit of microbial life. *Nature, 302*, 1018-1021.

Nevo, E., Zohary, D., Brown, A. H. D., & Haber, M. (1979). Genetic diversity and environmental

associations of wild barley, *Hordeum spontaneun*, in Israel. *Evolution, 33*(3), 815-833.

Noy-Meir, I. (1973). Desert ecosystems: environments and producers. *Annual Review of Ecology, Evolution, and Systematics, 4*, 25-51.

Noy-Meir, I. (1985). Desert Ecosystem Structure and Function. In E. M., N.-M. I. & D. W. Goodall (Eds.), *Hot Deserts and Arid Shrublands - Ecosystems of the World* (Vol. 12A, pp. 93-103). Amsterdam Elsevier Science Publishers.

Nylander, J. A. A., Wilgenbusch, J. C., Warren, D. L., & Swofford, D. L. (2008). AWTY (are we there yet ?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics, 24*(4), 581-583.

Ossa, P. G., Perez, F., & Armesto, J. J. (2013). Phylogeography of two closely related species of *Nolana* from the coastal Atacama Desert of Chile: post-glacial population expansions in response to climatic fluctuations. *Journal of Biogeography, 40*(11), 2191-2203.

Parra, J. L., Rahbek, C., McGuire, J. A., & Graham, C. H. (2011). Contrasting patterns of phylogenetic assemblage structure along the elevational gradient for major hummingbird clades. *Journal of Biogeography, 38*(12), 2350-2361.

Peña, C., Wahlberg, N., Weingartner, E., Kodandaramaiah, U., Nylin, S., Freitas, A. V. L., & Brower, A. V. Z. (2006). Higher level phylogeny of Satyrinae butterflies (Lepidoptera: Nymphalidae) based on DNA sequence data. *Molecular Phylogenetics and Evolution, 40*, 29-49.

Perez, F., Arroyo, M. T. K., Medel, R., & Hershkovitz, M. A. (2006). Ancestral reconstruction of flower morphology and pollination systems in *Schizanthus* (Solanaceae). *American Journal of Botany, 93*(7), 1029-1038.

Peterson, A. T., Papeş, M., & Eaton, M. (2007). Transferability and model evaluation in

ecological niche modeling: a comparison of GARP and Maxent. *Ecography, 30*, 550-560.

Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M., & Bastos-Araújo, M. (2011). *Ecological Niches and Geographic Distributions*. Princeton and Oxford: Princeton University Press.

Phillips, S. J., Anderson, R. P., & R.E. Shapire. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling, 190*, 231-259.

Pigliucci, M. (2005). Evolution of phenotypic plasticity: where are we going now? *Trends in Ecology & Evolution, 20*(5), 481-486.

Pillon, Y., Johansen, J., Sakishima, T., Chamala, S., Barbazuk, B. W., Roalson, E. H., . . . Stacy, E. A. (2013). Potential use of low-copy nuclear genes in DNA barcoding: a comparison with plastid genes in two Hawaiian plant radiations. *BMC Evolutionary Biology, 13*, 35.

Pons, J., Barraclough, T. C., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., . . . Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed species. *Systematic Biology, 55*(4), 595-609.

Posada, D. (2008). jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution, 25*, 1253-1256.

Puritz, J. B., Addison, J. A., & Toonen, R. J. (2012). Next-generation phylogeography: a targeted approach for multilocus sequencing of non-model organisms. *PLoS ONE, 7*(3), e34241.

R Development Core Team. (2013). R: A language and environment for statistical computing, Vienna, Austria. http://www.R-project.org: R Fundation for Statistical Computing.

Rambaut, A., & Drummond, A. J. (2007). Tracer v1.5. Available in http://beast.bio.ed.ac.uk/Tracer.

Ramos-Onsins, S. E., & Rozas, J. (2002). Statistical properties of new neutrality tests against

population growth. *Molecular Biology and Evolution, 19*(12), 2092-2100.

Rauh, W. (1985). The Peruvian-Chilean Deserts. In M. Evenari, I. Noy-Meir & D. W. Goodall (Eds.), *Hot Deserts and Arid Shrublands - Ecosystems of the World* (Vol. 12A, pp. 239-267). Amsterdam: Elsevier Science Publishers.

Reiche, C. (1895). Schizopetaléas. In C. Reiche (Ed.), *Flora de Chile* (pp. 104-109). Santiago: Imprenta Cervanntes.

Reid, N., & Carstens, B. (2012). Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evolutionary Biology, 12*(1), 196.

Rollins, R. C. (1966). The genus *Mathewsia* (Cruciferae). *Acta Botanica Neerlandica, 15*(1), 102-116.

Ronquist, F., & Huelsenbeck, J. P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics, 19*, 1572-1574.

Rozen, S., & Skaletsky, H. (2000). Primer 3 on the WWW for general users and for biologist programmers. In S. Krawetz & S. Misener (Eds.), *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (pp. 365-386). Totowa, New Jersey: Humana Press.

Rundel, P. W., Dillon, M. O., Palma, B., Mooney, H. A., Gulmon, S. L., & Ehleringer, J. R. (1991). Phytogeography and ecology of the coastal atacama and peruvian desert. *Aliso, 13*, 1-50.

Rundel, P. W., & Mahn, M. (1976). Community structure and diversity in a coastal fog desert in northern Chile. *Flora, 165*, 493-505.

Sang, T. (2002). Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry and Molecula Biology, 37*(3), 121-147.

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

Schlichting, C. D. (1986). The evolution of phenotypic plasticity in plants. *Annual Review in Ecology, Evolution, and Systematics, 17*, 667-693.

Schlüter, P. M., Stuessy, T., & Paulus, H. F. (2008). Making the first step: practical considerations for the isolation of low-copy nuclear sequence markers. *Taxon, 54*, 766-770.

Sellers, P. J., Berry, J. A., Collatz, G. J., Field, C. B., & Hall, E. G. (1992). Canopy reflectance , photosynthesis , and transpiration . III . A reanalysis using improved leaf models and a new canopy integration scheme. *Remote Sensing of Environment, 42*, 187-216.

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., . . . Small, R. L. (2005). The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany, 92*(1), 142-166.

Shaw, J., Lickey, E. B., Schilling, E. E., & Small, R. L. (2007). Comparison of whole chloroplast genome sequence to choose noncoding regions for phylogenetic studies in Angiosperms: The tortoise and the hare III. *American Journal of Botany, 94*(3), 275-288.

Shmida, A. (1985). Biogeography of the Desert Flora. In M. Evenari, I. Noy-Meir & D. W. Goodall (Eds.), *Hot Deserts and Arid Shrublands - Ecosystems of the World* (Vol. 12A, pp. 23-77). Amsterdam: Elsevier Science Publisher.

Simmons, M. P., & Ocheterena, H. (2000). Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology, 49*(2), 369-381.

Slotte, T., Ceplitis, A., Neuffer, B., Hurka, H., & Lascoux, M. (2006). Intrageneric phylogeny of *Capsella* (Brassicaceae) and the origin of the tetraploid *C. bursa-pastoris* based on

chloroplast and nuclear DNA sequences. *American Journal of Botany, 93*(11), 1714-1724.

Slotte, T., Huang, H., Lascoux, M., & Ceplitis, A. (2008). Polyploid speciation did not confer instant reproductive isolation in *Capsella* (Brassicaceae). *Molecular Biology and Evolution, 25*(7), 1472-1481.

Small, R. L., Cronn, R. C., & Wendel, J. F. (2004). Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany, 17*(2), 145-170.

Smedmark, J. E. E., Erikson, T., Evans, R. C., & Campbell, C. S. (2003). Ancient allopolyploid speciation in Geineae (Rosaceae): evidence from nuclear granule-bound starch synthase (GBSSI) gene sequences. *Systematic Biology, 52*(3), 374-385.

Soberón, J., & Peterson, A. T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics, 2*, 1-10.

Squeo, F. A., Arancio, F. M., & Gutierrez, J. R. (2008). *Libro Rojo de la Flora Nativa y de los Sitios Prioritarios para su Conservacion: Region de Atacama.* La Serena, Chile: Ediciones Universidad de la Serena.

St. Onge, K. R., Källman, T., Slotte, T., Lascoux, M., & A. Palmé. (2011). Contrasting demographic history and population structure in *Capsella rubella* and *Capsella grandiflora*, two closely related species with different mating systems. *Molecular Ecology, 20*, 3306-3320.

Stebbins, G. L. (1950). *Variation and evolution in plants*. New York: Columbia Unversity Press.

Stebbins, G. L. (1952). Aridity as a stimulus to plant evolution. *Evolution, 86*(826), 33-44.

Stephens, M., & Donnelly, P. (2003). A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics, 73*,

1162-1169.

Stephens, M., Smith, N., & Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics, 68*(978-989).

Stöver, B. C., & Müller, K. F. (2010). TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics, 11*(7).

Stuessy, T. F., & Taylor, C. (1995). Evolucion de la Flora Chilena. In Marticorena C. & R. Rodriguez (Eds.), *Flora de Chile* (Vol. 1, pp. 85-118). Concepcion: Universidad de Concepcion.

Toro-Núñez, O., Mort, M. E., Ruiz-Ponce, E., & Al-Shehbaz, I. A. (2013). Phylogenetic relationships of *Mathewsia* and *Schizopetalon* (Brassicaceae) inferred from nrDNA and cpDNA regions: taxonomic and evolutionary insights from an Atacama Desert lineage. *Taxon, 62*(2), 343-356.

Tu, T., Dillon, M. O., Sun, D., & Wen, J. (2008). Phylogeny of *Nolana* (Solanaceae) of the Atacama and Peruvian deserts inferred from sequences of four plastid markers and the nuclear LEAFY second intron. *Molecular Phylogenetics and Evolution, 49*, 561-573.

Veit, H. (1993). Upper Quaternary Landscape and Climate Evolution in the Norte Chico (Northern Chile): An Overview. *Mountain Research and Development, 13*(2), 139-144.

Verón, S. R., & Paruelo, J. M. (2010). Desertification alters the response of vegetation to changes in precipitation. *Journal of Applied Ecology, 47*(6), 1233-1241.

Vidiella, P. E., & Armesto, J. J. (1989). Emergence of ephemeral plant species from soil samples of the Chilean coastal desert in response to experimental irrigation. *Revista Chilena de Historia Natural, 62*, 99-107.

Vidiella, P. E., Armesto, J. J., & J.R. Gutierrez. (1999). Variation changes and sequential

flowering after rain in the southern Atacama Desert. *Journal of Arid Environments, 43*, 449-458.

Villagran, C., Arroyo, M. T. K., & Marticorena, C. (1983). Efectos de la desertizacion en la distribucion de la flora andina de Chile. *Revista Chilena de Historia Natural, 56*, 137-157.

Viruel, J., Catalán, P., & Segarra-Moragues, J. G. (2011). Disrupted phylogeographical microsatellite and chloroplast DNA patterns indicate a vicariance rather than long-distance dispersal origin for the disjunct distribution of the Chilean endemic *Dioscorea biloba* (Dioscoreaceae) around the Atacama Desert. *Journal of Biogeography, 39*(6), 1073-1085.

Volis, S., Mendlinger, S., Turuspekov, Y., & Esnazarov, U. (2002). Phenotypic and allozyme variation in Mediterranean and Desert populations of wild barley, *Hordeum spontaneum* Koch. *Evolution, 56*(7), 1403-1415.

Warren, D. L., Glor, R. E., & Turelli, M. (2008). Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution, 62*(11), 2868-2883.

Warren, D. L., Glor, R. E., & Turelli, M. (2010). ENMtools: a toolbox for comparative studies of environmental niche models. *Ecography, 33*, 607-611.

Warwick, S. I., Mummenhoff, K., Sauder, C. A., Koch, M., & Al-Shehbaz, I. A. (2010). Closing the gaps: phylogenetic relationship in the Brassicaceae based on DNA sequence data of nuclear ribosomal ITS region. *Plant Systematic and Evolution, 285*, 209-232.

Warwick, S. I., Sauder, C. A., Mayer, M. S., & Al-Shehbaz, I. A. (2009). Phylogenetic relationships in the tribes Schizopetaleae and Thelypodieae (Brassicaceae) based on nuclear ribosomal ITS region and plastid *ndhF* DNA sequences. *Botany, 87*(10), 961-985.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview version 2: A multiple sequence alignment and analyst workbench. *Bioinformatics, 25*(9), 1189-1191.

Weiner, J. (2004). Allocation, plasticity and allometry in plants. *Perspectives in Plant Ecology, Evolution and Systematics, 6*(4), 2078-2215.

Wen, J., & Zimmer, E. A. (1996). Phylogeny and biogeography of *Panax* L. (the ginseng genus, Araliaceae): inferences from ITS sequences of the nuclear ribosomal DNA. *Molecular Phylogenetics and Evolution, 6*, 167-177.

Wendel, J. F., & Doyle, J. D. (1998). Phylogenetic incongruences: window into genome history and molecular evolution. In D. E. Soltis, P. S. Soltis & J. D. Doyle (Eds.), *Molecular Systematics of plants II: DNA sequencing* (pp. 1-42). Boston, Dordrecht, London: Kluwer Academic Publishers.

Whitford, W. G. (2002). *Ecology of desert systems*: Academic Press, London.

Wiens, J. J., & Graham, C. H. (2005). Niche Conservatism: integrating evolution, ecology, and conservation biology. *Annual Review in Ecology, Evolution, and Systematics, 36*, 519-539.

Wiens, J. J., & Morrill, M. C. (2011). Missing Data in Phylogenetic Analysis: Reconciling Results from Simulations and Empirical Data. *Systematic Biology, 60*(5), 719-731.

Wright, S. D., Yong, C. G., Wichman, S. R., Dawson, J. W., & Gardner, R. C. (2001). Steeping stones to Hawaii: a trnas-equatorial dispersal pathway for *Metrosideros* (Myrtaceae) inferred from nrDNA (ITS+ETS). *Journal of Biogeography, 28*, 769-774.

Wright, S. I., Lauga, B., & Charlesworth, D. (2003). Subdivision and haplotype structure in natural populations of *Arabidopsis lyrata*. *Molecular Ecology, 12*(5), 1247-1263.

Young, K. R., Ulloa Ulloa, C., Luteyn, J. L., & Knapp, S. (2002). Plant evolution and endemism in Andean South America: An Introduction. *The Botanical Review, 68*(1), 4-21.

Zhang, N., Zeng, L., Shan, H., & Ma, H. (2012). Highly conserved low-copy nuclear genes as effective markers for phylogenetic analyses in angiosperms. *New Phytologist, 195*, 923-937.

Zimmer, E. A., & Wen, J. (2012). Using nuclear gene data for plant phylogenetics: Progress and prospects. *Molecular Phylogenetics and Evolution, 65*(2), 774-785.

Zwickl, D. J. (2006). Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. *Doctoral dissertation*.

Appendices

Appendix 1

Taxon and voucher information, and accession numbers of sequence data.

Lepidium angustissimum Phil., Chile, Región de Atacama, *Toro & Lira 36* (KANU) - [KC174335, KC174369, KC174408, KC174447, KC174484, KC174520]; ***Neuontobotrys tarapacana*** (Phil.) Al-Shehbaz, Chile, Región de Antofagasta, *Rosas 5574* (CONC) - [KC174336, KC174372, KC174411, KC174450, KC174487, KC174523]; ***Eudema nubigena*** Bonpl. Ecuador, *Holm-Nielsen 20875* MO - [–, KC174370, KC174409, KC174448, KC174485, KC17452]; ***Menonvillea chilensis*** Turcz., Chile, Región de Atacama, *Toro & Lira 10* KANU - [–, KC174371, KC174410, KC174449, KC174486, KC174522]; ***Weberbauera colchaguensis*** Barnéoud Al-Shehbaz, Chile, Región del Maule, *Ruthsatz 7848* MO - [–, KC174373, KC174412, KC174451, KC174488, KC174524]; ***Mathewsia auriculata*** Phil., Chile, Región de Atacama, *Clarke 22-02* CONC - [KC174347, KC174385, KC174424, KC174462, KC174499, KC174535], *Tellier 4137* SGO - [KC174348, KC174386, KC174425, –, KC174500, KC174500, KC174536]; ***Mathewsia linearifolia*** Turcz, Chile, Región de Atacama, *Zoellner 11658* MO* - [KC174339, KC174376, KC174415, KC174453, KC174491, KC174527]; *Zoellner 9799* MO - [KC174341, KC174378, KC174417, KC174455, –, KC174528]; ***Mathewsia collina*** I.M.Johnst., Chile, Región de Antofagasta, *Johnston 3599* GH** - [KC174342, KC174379, KC174418, KC174456, KC174493, KC174529], *Quezada & Ruiz 263* CONC - [KC174340, KC174377, KC174416, KC174454, KC174492, –]; ***Mathewsia densifolia*** Rollins, Peru, Departamento de Arequipa, *Ferreyra 11678* GH** - [KC174345, KC174383, KC174422, KC174460, KC174497, KC174533]; ***Mathewsia densifolia*** var. ***stylosa*** Rollins, Peru, Departamento de Moquegua, *Vargas C. 8587* GH** - [KC174346, KC174384, KC174423, KC174461, KC174498, KC174534]; ***Mathewsia foliosa*** Hook. & Arn., Chile, Región de Coquimbo, *Landrum & Landrum 9837* SGO* - [KC174349, KC174387, KC174426, KC174463, KC174501, KC174537]; Chile, Región de Valparaiso, *Luebert & Kritzner 1886* SGO - [KC174350, KC174388, KC174427, KC174464, KC174502, KC174538]; ***Mathewsia incana*** Phil., Chile, Región de Antofagasta, *Quezada & Ruiz 190* CONC - *M.incana1* - [KC174337, KC174374, KC174413, KC174452, KC174489, KC174525]; Chile, Región de Atacama, *Toro & Lira 17* KANU - *M.incana2* - [KC174338, KC174375, KC174414, –, KC174490, KC174526]; ***Mathewsia nivea*** Phil. O.E. Schulz, Chile, Region de Atacama, *Arancio & Squeo 10092* CONC - [–, KC174380, KC174419, KC174457, KC174494, KC174530]; *Arancio & Squeo 10165* CONC - [KC174343, KC174381, KC174420, KC174458, KC174495, KC174531]; ***Mathewsia peruviana*** O.E. Schulz, Peru, Departamento de Arequipa, without collector *739* MO* - [KC174344, KC174382, KC174421, KC174459, KC174496, KC174532]; ***Schizopetalon arcuatum*** Al-Shehbaz, Chile, Region de Atacama, *Toro & Lira 48* KANU - *S.arcuatum2* - [KC174365, KC174404, KC174443, KC174480, KC174516, KC174554]; *Toro & Lira 51* KANU - *S.arcuatum3* - [KC174366, KC174405, KC174444, KC174481, KC174517, KC174555]; *Toro & Lira 43* KANU - *S.arcuatum1* - [KC174360, KC174399, KC174438, KC174475, KC174511, KC174549]; ***Schizopetalon bipinnatifidum*** Phil., Chile, Region de Coquimbo, *Zöllner 12026* MO - [–, KC174390, KC174429, KC174466, KC174503, KC174540]; *Muñoz 4231* SGO - [KC174358, KC174397, KC174436, KC174473, KC174510, KC174547]; *Garaventa 8229* CONC - *LosMolles2* - [KC174352, KC174391, KC174430,

KC174467, KC174504, KC174541]; *Jiles 3274* OSU - *LosMolles1* - [KC174353, KC174392, KC174431, KC174468, KC174505, KC174542]; ***Schizopetalon biseriatum*** Phil., Chile, Region de Atacama, *Zöllner 11915* MO - [KC174351, KC174389, KC174428, KC174465, –, KC174539]; *Taylor & al. 10698* CONC - [KC174355, KC174394, KC174433, KC174470, KC174507, KC174544]; *Taylor & al. 10685* MO* - [KC174354, KC174393, KC174432, KC174469, KC174506, KC174543]; ***Schizopetalon brachycarpum*** Al-Shehbaz, Chile, Region Metropolitana, *Schlegel 926* CONC** - [KC174367, KC174406, KC174445, KC174482, KC174518, KC174556]; *Toro 56* KANU - [KC174368, KC174407, KC174446, KC174483, KC174519, KC174557]; ***Schizopetalon dentatum*** Gilg & Mushl., Chile, Region Metropolitana, *Arroyo & Humaña 995316* CONC - [KC174357, KC174396, KC174435, KC174472, KC174509, KC174546]; ***Schizopetalon maritimum*** Barnéoud, Chile, Region de Coquimbo, *Toro & Lira 53* KANU - [KC174361, KC174400, KC174439, KC174476, KC174512, KC174550]; ***Schizopetalon rupestre*** Reiche, Chile, Region de Atacama, *Tellier & Delaunoy 5576* CONC - [KC174362, KC174401, KC174440, KC174477, KC174513, KC174551]; Chile, Region de Atacama, *Letelier & Reyes 1140* CONC - [KC174363, KC174402, KC174441, KC174478, KC174514, KC174552]; ***Schizopetalon tenuifolium*** Phil., Chile, Region de Atacama, *Tellier 423* CONC - [KC174356, KC174395, KC174434, KC174471, KC174508, KC174545]; ***Schizopetalon walkeri*** Sims, Chile, Region de Valparaiso, *Luebert & Kritzner 1927* SGO - [KC174364, KC174403, KC174442, KC174479, KC174479, KC174553]; *Zöllner 8217* MO* - [KC174359, KC174398, KC174437, KC174474, –, KC174548]

" * " resequenced voucher (only for ITS region); " ** " sequences from type (isotype or holotype) specimens; "–" denotes absent sequence. Taxon, country, administrative region, *collector*, *collector number* (herbarium) - *special legend* (when cited in the analysis) - GenBank accession codes [ETS, ITS, *atpH-atpI* intergenic spacer, *trnH-psbA* intergenic spacer, *rps16* intron, *rps16-trnQ* intergenic spacer]

Appendix 2

Coordinate list from confirmed specimens for ENM analysis

Species	Collector	Herbarium	Locality (Label)	Longitude	Latitude
S.biseriatum	*Anonymous*	SGO	Caldera	-70.80588	-27.0577
S.biseriatum	Barros 2776	CONC	Chañaral	-70.6884	-26.509
S.biseriatum	Brinck & Munoz s.n.	SGO	Carrizal a Tororal	-71.12417	-28.0118
S.biseriatum	Gigoux sn, Jiles 5303	CONC, GH	Vecinity of caldera	-70.83802	-27.1247
S.biseriatum	Munoz & Johnson 1931	SGO	Mina Candelaria, camino a bahia salada	-70.80246	-27.6667
S.biseriatum	Munoz 4578	SGO	Km 917 S Obispito	-70.7637	-26.7706
S.biseriatum	Munoz 4686	SGO	Camino a Puerto Viejo	-70.93973	-27.2281
S.biseriatum	Munoz 4692	SGO	Llanos arenosos antes de quebrada honda	-71.12417	-28.0118
S.biseriatum	Ricardi 2261	CONC	Chañaral, en quebradas	-70.6884	-26.509
S.biseriatum	Saa s.n.	CONC	Longitudinal a Caldera	-70.63409	-27.3849
S.biseriatum	Taylor & al. 10685	MO	Bahia Inglesa	-70.88815	-27.146
S.biseriatum	Taylor & al. 10689	CONC	N of Caldera	-70.80181	-27.0479
S.biseriatum	Taylor & al. 10698	MO	N of Caldera, 10 km. By a large roadside shrine	-70.7866	-26.9863
S.biseriatum	Toro & Lira 12	KANU	Highway Copiapo - Caldera	-70.62078	-27.2818
S.biseriatum	Toro & Lira 13	KANU	Secondary road to Barranquilla	-70.74228	-27.487
S.biseriatum	Toro & Lira 15	KANU	Exit N of Caldera	-70.7706	-27.0332
S.biseriatum	Toro & Lira 16	KANU	Secondary road from Caldera to Quebrada Leon	-70.69777	-27.012
S.biseriatum	Toro & Lira 20	KANU	Dunes close to Morro Copiapo	-70.91055	-27.1424
S.biseriatum	Toro & Lira 21	KANU	Road to Barranquilla and Quebrada Honda	-70.82785	-27.4688
S.biseriatum	Toro & Lira 22	KANU	Road Bahia Blanca – Copiapo	-70.85103	-27.6636
S.biseriatum	Toro & Lira 35	KANU	Road Carrizal Bajo - Huasco, limits of Llanos del Challe reserve park	-71.15725	-28.1199
S.biseriatum	Toro & Lira 39	KANU	Road Caleta Totoral and Carrizal Bajo	-71.09163	-27.9057
S.biseriatum	Wendermann 443	MO, UC	Monte Amargo	-70.73692	-27.4491
S.biseriatum	Worth & Morrison 16193	UC	Morro Copiapo, WSW Caldera	-70.91384	-27.1418
S. tenuifolium	Tellier 855	CONC, MO	km 750 cruce Channarcillo	-70.51847	-27.8541
S. tenuifolium	Tellier 423	CONC, MO, SGO	Punta de Diaz	-70.59561	-28.1063

Species	Collector	Herbarium	Locality (Label)	Longitude	Latitude
S. tenuifolium	Jiles 2162	CONC	Algarrobal llanura	-70.91845	-28.1466
S. tenuifolium	Shlegel 3889	CONC	Llano Travesia	-70.50791	-27.5774
S. tenuifolium	Zoellner 9822, Koeler 131	CONC	Bandurrias (Cerro Blanco)	-70.52344	-27.9848
S. tenuifolium	Geisse s.n., Barros s.n.	CONC, SGO	Cerro Bandurrias	-70.4134	-27.8618
S. tenuifolium	Behn s.n.	CONC	Carrizal Bajo	-71.0831	-28.1365
S. tenuifolium	Toro & Lira 11	KANU	Highway Copiapo - Vallenar, 500 m from main road	-70.4594	-27.5637
S. tenuifolium	Toro & Lira 28	KANU	Road to Bahia Salada	-70.52107	-27.8075
S. tenuifolium	Toro & Lira 29	KANU	Highway Copiapo - Vallenar, crossroad to Chañarcillo	-70.54792	-27.9437
S. tenuifolium	Toro & Lira 30	KANU	W Road to El Totoral	-70.78002	-27.9467
S. tenuifolium	Toro & Lira 33	KANU	Secondary E Road to El Totoral	-70.97933	-27.8829
S. tenuifolium	Toro & Lira 41	KANU	Over sides of Copiapo - Vallenar Highway	-70.689	-28.2319
S. tenuifolium	Toro & Lira 42	KANU	Road from Canto del Agua to Carrizal Alto	-70.86875	-28.151
S. tenuifolium	Shlegel 3889, Garaventa 8228	OS	Llano Travesia	-70.50791	-27.5774
S. tenuifolium	*Anonymous*	SGO	Chanarcillo	-70.4543	-27.8635
S. tenuifolium	Munoz & al. 1011	SGO	Quebradita desde Copiapo a Vallenar	-70.56344	-27.9917
S. tenuifolium	Muñoz & al. 1062	SGO	3 km separacion camino a los colorados a Carrizal Bajo	-70.76058	-28.3627
S.bipinnatifidum	Toro & Lira 115	KANU	Portezuelo Tres Cruces	-30.21923	-70.6523
S.bipinnatifidum	Jiles 3203	OS	El Chañar, cordillera de Ovalle	-30.28466	-70.6242
S.bipinnatifidum	*Anonymous*	CONC	Rio Torca, Ovalle	-30.96175	-70.5651
S.bipinnatifidum	Munoz 1836	SGO	Cuesta Pajonales	-29.15201	-70.9811
S.bipinnatifidum	Munoz 4209, 4239, Wagenknecht 18550, Behn s.n., Ricardi & Marticorena 4453/938, Marticorena & Matthei 336	CONC, MO, UC	Cuesta la viñita - Camino Paihuano a Rivadavia	-29.82219	-70.8371
S.bipinnatifidum	Behn sn	CONC	Vallenar - Estancia La Totora	-28.62308	-71.1428
S.bipinnatifidum	Ricardi & Marticorena 4454/839	CONC	30 Km S de Vallenar	-29.992	-70.5322
S.bipinnatifidum	Jiles 1545, 1544	CONC	La Hualtata	-31.31783	-70.9251

Species	Collector	Herbarium	Locality (Label)	Longitude	Latitude
S.bipinnatifidum	Marticorena, Rodriguez & Weldt 1717	CONC	Mineral Los Cristales - Quebrada Las Salinas	-31.31783	-70.9251
S.bipinnatifidum	Jiles 3540	CONC	Mina Mantos Grandes	-29.13895	-71.1874
S.bipinnatifidum	Zoellner 8225	CONC	Monte Grande	-30.85803	-70.5619
S.bipinnatifidum	Zoellner 12026	MO	Coquimbo, on pass Pajonales (cuesta?)	-29.15201	-70.9811
S.bipinnatifidum	Muñoz 4321	SGO	A mitad de Cuesta La Viñita	-29.82219	-70.8371
S.rupestre	Gigoroux s.n.	GH	Quebrada Donna Ines Chica	-69.33793	-26.1125
S.rupestre	Toro & Flores 83	KANU	South slope Cerro Doña Ines	-69.21172	-26.1285
S.rupestre	Letelier & Reyes 1140	SGO	Ladera S cerro Doña Ines	-69.24243	-26.1704
S.rupestre	Toro & Flores 61	KANU	Close to El Salvador	-69.50573	-26.2856
S.rupestre	Toro & Flores 85	KANU	Close to Quebrada Pedernales, road to Montadon Mine	-69.28427	-26.364
S.rupestre	Zollner 3954	CONC	Campamento Pedernales	-69.27529	-26.37
S.rupestre	Niemayer s.n.	CONC	Vega de Agua Helada	-68.94678	-26.3425
S.rupestre	Munoz & al. 2757	SGO	Potrerillos - cuesta los patos	-69.52425	-26.3974
S.rupestre	Johnston 4722	GH	Potrerillos	-69.47278	-26.4348
S.rupestre	Zollner 18548, 18339	MO	Quebrada Potrerillos Atacama	-69.55014	-26.4015
S.rupestre	Toro & Flores 78	KANU	Road to Potrerillos, over Cuesta Los Patos	-69.50922	-26.4026
S.rupestre	Tellier 3707	SGO	Potrerillos, mina el hueso - quebrada agua de la falda	-69.39098	-26.492
S.rupestre	Arancio 92214	ULS	Quebrada el Salitral	-68.92691	-26.5329
S.rupestre	Toro & Flores 87	KANU	Road to Nueva Esperanza Mine	-69.14905	-26.6613
S.rupestre	Letelier & Squeo 1166	CONC	Campamento Mina Esperanza	-69.14905	-26.6613
S.rupestre	Toro & Flores 93	KANU	Cuesta Codoceo	-69.23995	-26.8855
S.rupestre	Marticorena & al. 9845	CONC	Copiapo a Tinogasta - Quebfrada Codoceo	-69.48848	-26.8673
S.rupestre	Ricardi & al. 595, 573	CONC	Salar de Maricunga	-69.26603	-26.8938
S.rupestre	Toro & Flores 96, Villagran & Arroyo 4646	KANU, CONC	Camino Internacional Copiapo - Tinogasta Cuesta El Salto	-69.66023	-27.0194
S.rupestre	Marticorena & al. 532, 548	CONC	Quebrada Chinches	-69.33026	-27.0793
S.rupestre	Niemayer s.n.	CONC	Vega del Obispo	-69.57261	-27.1282

Species	Collector	Herbarium	Locality (Label)	Longitude	Latitude
S.rupestre	Billiet & Jadin 5575	MO	Mina La Pepa	-69.33236	-27.2501
S.rupestre	Marticorena & al. 83578, 83579, Ricardi & al. 639	CONC, OS	Quebrada Vizcachas Atacama (26 51'S - 69 45'W)	-69.42223	-27.2846
S.rupestre	Johnston 4941	GH	Quebrada Tolar, Sierra San Miguel	-69.63974	-27.4996
S.rupestre	San Roman 2112, 1828	SGO	Quebrada de Serna	-69.84126	-27.7304
S.rupestre	Rosas 507	OS	Quebrada Yeguas Heladas	-69.32407	-27.6725
S.rupestre	Toro & Flores 101	KANU	Cuesta Los Castaños, road to El Juncal	-69.71893	-27.7229
S.rupestre	Toro & Flores 98	KANU	Road to Pircas Negras, close to Chilean Custom Office	-69.41585	-27.8965
S.rupestre	Wendermann 947, 948, 949, 950	CONC, GH, MO, UC	Rio Turbio, Atacama	-69.36995	-27.9233
S.rupestre	Niemayer s.n.	CONC	Cordillera de Pulido, Vegas de los Helades	-69.68657	-28.1167
S.rupestre	Toro J. s.n.	KANU	Campamento Las Breas, sector Mina Caserones	-69.63295	-28.1353
S.rupestre	Biurrun s.n.	SI	Entre Laguna Brava y Paso Pircas Negras, alrededor del Refugio de Pastillos-Argentina	-69.1446	-28.1743
S.rupestre	Krapovickas & Hunzkier 5655	BAA	Valle del Rio Blanco-Argentina	-69.25388	-28.3685
S.rupestre	Tellier 4989	CONC	Quebrada Piuquenes	-69.94881	-28.6099
S.rupestre	Tellier & Delaunoy 5576	CONC	Quebrada La Torora	-70.18447	-28.6687
S.rupestre	Toro & Flores 105	KANU	Road to Mine Los Morros	-70.19302	-28.6956
S.rupestre	Tellier 4969	CONC	Cuenca el Transito - Quebrada Cantaritos	-69.87127	-28.7035
S.rupestre	Arroyo 81518, Marticorena & al. 83388, 83424	CONC	Rio Laguna Grande	-69.99536	-28.7757
S.rupestre	Johnston 5979	GH	Laguna Chica	-69.85806	-28.8027
S.rupestre	Castellanos s.n.	P	Cajon Los Tambillos-Argentina	-69.82924	-29.1897
S.rupestre	Johnston 6088	GH	Pueblo Tambillo - Argentina	-69.85593	-29.2214
S.rupestre	Kelsing 9500	SI	Rio de la Placa, Quebrada Honda-Argentina	-69.51272	-29.3303
S.rupestre	Johnston 6157	GH	Rio Tagua & Rio Cura - Argentina	-69.57687	-29.3397
S.rupestre	Pujalte 200	BACP, SI	Valle Rio de las Taguas-Argentina	-69.91268	-29.3568

Species	Collector	Herbarium	Locality (Label)	Longitude	Latitude
S.rupestre	Toro & Flores 104	KANU	Road from Mine Barrick Security Post, through the beginning of Quebrada San Felix	-70.02743	-29.4672
S.rupestre	Squeo 88042, 88043	CONC, MO, OS	Camino al Indio	-70.0204	-29.7956
S.rupestre	Squeo 88002, 88003	CONC, MO	Cordillera Doña Ana	-70.0433	-29.8026
S.rupestre	Arancio 12775	ULS	Cordillera de Punilla	-70.29862	-29.8247
S.rupestre	Osorio s.n.	SGO	Valle del Elqui, Vega Piuquenes de Baños del Toro	-70.03257	-29.8454
S.rupestre	Moreau s.n.	CONC	Junta del Arroyo - Argentina	-69.71902	-29.8518
S.rupestre	Toro & Flores 112	KANU	Road to El Indio Mine / Banos del Toro	-70.05018	-29.9005
S.rupestre	Weckenmann sn, Beckett & al. 4657, Wagenknecht 18498, 18499, 18500, 18501, Zollner 10273, 12655, 12626, Wendermann 192, 193, 194, 195, Ricardi, Marticorena & Matthei 1127, Zoellner 10413, Squeo 88086, Jiles 6466, Stuessy & Ruiz 12787, Wagenknecht 18498	CONC, KANU, MO, SGO, UC	Baños del Toro	-70.05018	-29.9005
S.rupestre	Arroyo 81038, 81039, 81127	CONC	Cerro Tapado - La Laguna	-70.04636	-30.1608
S.rupestre	Toro & Flores 110, Wagenknecht 276, 277	CONC, KANU	Road to Paso Agua Negra, over slope slides of La Laguna	-70.03703	-30.2239
S.rupestre	Ricardi & al. 1780, 1781, 1727	CONC	Embalse La Laguna - Camino a	-69.94261	-30.2627
S.rupestre	Kielsing & al. 8100	SI	El Pachon-Argentina	-70.41595	-31.7623
S.rupestre	Kurtz s.n.	SI	Cordillera del Espinazito, Valle del Portillo-Argentina	-69.99604	-32.3135
S.rupestre	Bocher & al. 2180	BAA, MO, SGO	Between Puente del Inca and Las Cuevas-Argentina	-69.99727	-32.8159

Appendix 3

Results from PCA of climatic layers

	eigenvalue	% variance	% cumulated variance
PC1	2.563964193	0.438260826	0.438261
PC2	2.472160557	0.407438521	0.845699
PC3	1.073506763	0.076827785	0.922527
PC4	0.70437289	0.033076078	0.955603
PC5	0.623726779	0.025935673	0.981539
PC6	0.496442314	0.016430331	0.997969

Table A. Obtained eigenvalues, their percentage of variance and cumulated variance for climatic data obtained with the coastal mask (*S. biseriatum*, *S. tenuifolium*, and *S. bipinnatifidum*).

	eigenvalue	% variance	% cumulated variance
PC1	2.522991484	0.424365735	0.424366
PC2	2.052421309	0.280828882	0.705195
PC3	1.551630545	0.160503823	0.865698
PC4	1.167559563	0.090879689	0.956578
PC5	0.629559226	0.026422988	0.983001
PC6	0.469830976	0.014716076	0.997717

Table B. Obtained eigenvalues, their percentage of variance and cumulated variance for climatic data obtained with the Andean mask (*S. rupestre*).

	PC1	PC2	PC3	PC4	PC5	PC6
Bio1	-0.04087007	0.396445911	0.064522	-0.19743	-0.02094	0.123493
Bio2	0.172309058	-0.15627389	0.750229	-0.09188	-0.00335	0.078557
Bio3	-0.25456404	-0.20647722	0.404417	-0.41321	-0.23773	-0.24461
Bio4	0.36302334	0.075271501	0.011958	0.318043	0.297609	0.199149
Bio5	0.102777974	0.382023525	0.143884	-0.11856	0.062715	0.154803
Bio6	-0.08230712	0.389848026	-0.05329	-0.2091	-0.04143	0.044322
Bio7	0.348636208	0.006314843	0.373285	0.16135	0.196435	0.2132
Bio10	0.031792405	0.399630032	0.053132	-0.12173	0.037972	0.155619
Bio11	-0.10272482	0.382802628	0.051043	-0.2365	-0.07252	0.089052
Bio12	0.376832236	0.049764686	-0.07058	-0.17572	-0.00508	-0.34179
Bio13	0.368523917	0.083731065	-0.06512	-0.14737	0.037836	-0.42868
Bio14	0.308649362	-0.16829562	-0.18246	-0.2106	-0.49308	0.403715
Bio15	0.020348782	0.310791612	0.18494	0.607377	-0.65914	-0.25044
Bio16	0.370246846	0.079381162	-0.06229	-0.16166	0.039354	-0.41154
Bio17	0.333209641	-0.15874977	-0.14279	-0.18475	-0.35061	0.28267

Table C. Obtained loadings of components (axis) for climatic data obtained with the coastal mask (*S. biseriatum*, *S. tenuifolium*, and *S. bipinnatifidum*).

	PC1	PC2	PC3	PC4	PC5	PC6
Bio1	-0.29733251	0.317787485	-0.02204	0.022853	-0.14554	-0.02897
Bio2	-0.08338944	-0.13754681	0.485668	-0.39932	-0.4167	-0.30166
Bio3	0.360518698	0.01892445	-0.06371	-0.05974	-0.56519	-0.34357
Bio4	-0.33769528	-0.08177309	0.288564	-0.06491	0.297115	0.074923
Bio5	-0.36077348	0.189425991	0.077406	-0.02275	-0.09646	-0.04661
Bio6	-0.21231821	0.341248596	-0.2628	0.167446	-0.19595	-0.01031
Bio7	-0.29003273	-0.09168007	0.389636	-0.21029	0.070233	-0.05565
Bio10	-0.34069456	0.246252707	0.040178	0.021027	-0.04257	-0.00431
Bio11	-0.22989172	0.375700955	-0.12591	0.064663	-0.24239	-0.05252
Bio12	-0.24197429	-0.28592310	-0.32166	-0.13273	-0.02837	-0.16354
Bio13	-0.2082691	-0.24178790	-0.39436	-0.25667	0.070124	-0.19522
Bio14	-0.18391997	-0.37320568	0.016239	0.278735	-0.36334	0.394714
Bio15	0.103599282	0.176023098	-0.15405	-0.68501	-0.15263	0.656057
Bio16	-0.21837641	-0.24524985	-0.38595	-0.23045	0.051746	-0.21252
Bio17	-0.1953818	-0.37401850	0.005448	0.26703	-0.3472	0.283068

Table D. Obtained loadings of components (axis) for climatic data obtained with the coastal mask (*S. rupestre*).

Variable code	Variable name
BIO1	Annual Mean Temperature
BIO2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO3	Isothermality (BIO2/BIO7) (* 100)
BIO4	Temperature Seasonality (standard deviation *100)
BIO5	Max Temperature of Warmest Month
BIO6	Min Temperature of Coldest Month
BIO7	Temperature Annual Range (BIO5-BIO6)
BIO10	Mean Temperature of Warmest Quarter
BIO11	Mean Temperature of Coldest Quarter
BIO12	Annual Precipitation
BIO13	Precipitation of Wettest Month
BIO14	Precipitation of Driest Month
BIO15	Precipitation Seasonality (Coefficient of Variation)
BIO16	Precipitation of Wettest Quarter
BIO17	Precipitation of Driest Quarter

Table E. List of Worldclim layers used in this study.

Appendix 4

Results of interspecific PCA of morphological data.

PC	Eigenvalue	% variance	% Cumulated variance
1	0.0411255	32.673	32.673
2	0.030227	24.014	56.687
3	0.0230864	18.341	74.028
4	0.0139991	11.122	86.15
5	0.0123764	9.8326	95.9826
6	0.00274707	2.1825	98.1651
7	0.00225208	1.7892	99.9543
8	5.71968E-05	0.045441	99.999741

Table A. Obtained eigenvalues, their percentage of variance and cumulated variance for morphological data.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
StIn c	0.2549	0.3141	0.812	-0.4084	0.06982	-0.00656	-0.07262	0.002248
InIn c	-0.3218	0.7761	-0.3849	-0.3776	-0.03837	-0.04024	0.01398	0.002871
SeLe	-0.05032	0.09907	0.02157	0.1166	-0.03244	0.9456	-0.2797	-0.00905
FPe	0.1753	0.5255	0.1462	0.8025	-0.05027	-0.1567	-0.02379	0.009153
A c	-0.7549	-0.0506	0.3637	0.1639	0.1471	0.06835	0.4201	0.2562
P c	0.1832	0.07508	-0.09901	0.02104	0.7327	0.1462	0.4193	-0.4653
L/W	0.3022	0.05069	-0.03075	-0.06259	-0.5443	0.2212	0.7456	0.01532
DI	0.3215	0.04755	-0.1661	-0.04193	0.368	0.06761	0.08722	0.847

Table B. Obtained loadings of components (axis) for morphological data. In bold are highlighted the characters that contribute the highest variability (>0.5) in the first three components.

Appendix 5

Estimates of species delimitation in the Atacama subclade, using haplotypes inferred from MMT.

Single threshold		Multiple threshold	
GMYC species	Haplotypes	GMYC species	Haplotypes
1	Hap_20	1	Hap_20
1	Hap_18	1	Hap_18
1	Hap_24	1	Hap_24
1	Hap_23	1	Hap_23
1	Hap_21	1	Hap_21
1	Hap_19	1	Hap_19
1	Hap_25	1	Hap_25
2	Hap_15	2	Hap_15
2	Hap_11	2	Hap_11
2	Hap_10	2	Hap_10
2	Hap_9	2	Hap_9
2	Hap_14	2	Hap_14
2	Hap_12	2	Hap_12
2	Hap_13	2	Hap_13
3	Hap_4	3	Hap_4
3	Hap_29	3	Hap_29
3	Hap_33	3	Hap_33
3	Hap_34	3	Hap_34
3	Hap_30	3	Hap_30
3	Hap_7	3	Hap_7
3	Hap_6	3	Hap_6
3	Hap_31	3	Hap_31
3	Hap_32	3	Hap_32
4	Hap_27	4	Hap_27
4	Hap_17	4	Hap_17
4	Hap_5	5	Hap_5
4	Hap_22	5	Hap_22
4	Hap_8	5	Hap_8
4	Hap_16	5	Hap_16
5	Hap_28	6	Hap_28
5	Hap_26	6	Hap_26
5	Hap_3	6	Hap_3
5	Hap_2	7	Hap_2
6	Hap_1	8	Hap_1

Table A. List of species inferred from MMT haplotypes

Results of GMYC delimitation		
Method	Single	Multiple
Likelihood of null model	-4.705519	-4.705519
ML of GMYC model	-0.6928173	0.293443
Likelihood ratio	8.025404	9.997924
LR test (null vs alternative)	**0.01808447(***)**	**0.006744944(***)**
Number of ML clusters	5	6
Confidence interval	1-6	4-11
Number of ML entities	5	8
Confidence interval	2-7	6-19
Threshold time (arbitrary)	-8	-8 -3 -2

Table B. Results of GMYC analyses inferred with single and multiple thresholds. Statistical significant differences ($P<0.05$; in bold) between null (one species) and alternative (multiple species) are inferred with a likelihood ratio test (LR).

Appendix 6

Results of multivariate species by species comparison (post hoc Dunn's test) at every climatic factor used in ENM analysis. In bold are marked significant results (P<0.008, Bonferroni corrected). Values of differences, lower and upper levels are referred to a 99.9% confidence intervals for a Kruskal –Wallis test. Results suggest a lack of differentiation in key climatic variables like maximum of NDVI and PC1 for some interspecific comparisons.

NDVI max

	Diff	Lower	Upper	Adj. P-Value
S. bipinnatifidum - S. biseriatum	53815.7594	51749.1058	55882.41297	0
S. bipinnatifidum - S. rupestre	53235.4277	52272.022	54198.83344	0
S. biseriatum - S. rupestre	**-580.33165**	**-2552.98172**	**1392.32**	**1**
S. bipinnatifidum - S. tenuifolium	46342.6307	44073.5797	48611.68169	0
S. biseriatum - S. tenuifolium	-7473.1287	-10321.2491	- 4625.00822	0
S. rupestre - S. tenuifolium	-6892.797	-9076.57438	- 4709.01966	0

NDVI min

	Diff	Lower	Upper	Adj. P-Value
S. bipinnatifidum - S. biseriatum	36547.2556	34480.602	38613.9	0
S. bipinnatifidum - S. rupestre	1029.33668	65.93095	1992.74	0.000346
S. biseriatum - S. rupestre	-35517.919	-37490.569	-33545	0
S. bipinnatifidum - S. tenuifolium	33484.962	31215.911	35754	0
S. biseriatum - S. tenuifolium	-3062.2936	-5910.41407	-214.17	0.00031
S. rupestre - S. tenuifolium	32455.6253	30271.8479	34639.4	0

NDVI sd

	Diff	Lower	Upper D	Adj. P-Value
S. bipinnatifidum - S. biseriatum	39511.3322	37444.67858 4	1577.98578	0
S. bipinnatifidum - S. rupestre	34827.6034	33864.19766 3	5791.00911	0
S. biseriatum - S. rupestre	-4683.7288	- 6656.37886 -	2711.07873	0
S. bipinnatifidum - S. tenuifolium	58756.4899	56487.43895 6	1025.54093	0
S. biseriatum - S. tenuifolium	19245.1578	16397.03732 2	2093.27821	0
S. rupestre - S. tenuifolium	23928.8866	21745.1092 2	6112.66392	0

PCA 1

	Diff	Lower	Upper	Adj. P-Value
S. bipinnatifidum - S. biseriatum	26440.4483	24373.7947	28507.1	0
S. bipinnatifidum - S. rupestre	-56229.764	-57193.1698	-55266	0
S. biseriatum - S. rupestre	-82670.212	-84642.8625	-80698	0
S. bipinnatifidum - S. tenuifolium	24366.0423	22096.9913	26635.1	0
S. biseriatum - S. tenuifolium	**-2074.406**	**-4922.52645**	**773.714**	**0.036631**
S. rupestre - S. tenuifolium	80595.8064	78412.029	82779.6	0

PCA 2

	Diff	Lower	Upper	Adj. P-Value
S. bipinnatifidum - S. biseriatum	-42499.337	-44565.991	-40433	0
S. bipinnatifidum - S. rupestre	47799.5276	46836.1218	48762.9	0
S. biseriatum - S. rupestre	90298.8649	88326.2149	92271.5	0
S. bipinnatifidum - S. tenuifolium	-36295.991	-38565.0418	-34027	0
S. biseriatum - S. tenuifolium	6203.34655	3355.2261	9051.47	0
S. rupestre - S. tenuifolium	-84095.518	-86279.2957	-81912	0

PCA 3

	Diff	Lower	Upper	Adj. P-Value
S. bipinnatifidum - S. biseriatum	40791.1188	38724.4652	42857.8	0
S. bipinnatifidum - S. rupestre	-48791.127	-49754.5328	-47828	0
S. biseriatum - S. rupestre	-89582.246	-91554.896	-87610	0
S. bipinnatifidum - S. tenuifolium	21362.6719	19093.621	23631.7	0
S. biseriatum - S. tenuifolium	-19428.447	-22276.5673	-16580	0
S. rupestre - S. tenuifolium	70153.799	67970.0217	72337.6	0

PCA 4

	Diff	Lower	Upper	Adj. P-Value
S. bipinnatifidum - S. biseriatum	-75558.463	-77625.1167	-73492	0
S. bipinnatifidum - S. rupestre	-64036.145	-64999.5508	-63073	0
S. biseriatum - S. rupestre	11522.318	9549.66793	13495	0
S. bipinnatifidum - S. tenuifolium	-46684.973	-48954.024	-44416	0
S. biseriatum - S. tenuifolium	28873.4901	26025.3697	31721.6	0
S. rupestre - S. tenuifolium	17351.1721	15167.3948	19534.9	0

PCA 5

	Diff	Lower	Upper	Adj. P - value
S. bipinnatifidum - S. biseriatum	-28817.54	-30884.1935	-26751	0
S. bipinnatifidum - S. rupestre	-11938.208	-12901.6135	-10975	0
S. biseriatum - S. rupestre	16879.3321	14906.682	18852	0
S. bipinnatifidum - S. tenuifolium	29898.272	27629.221	32167.3	0
S. biseriatum - S. tenuifolium	58715.8119	55867.6914	61563.9	0
S. rupestre - S. tenuifolium	41836.4798	39652.7024	44020.3	0

PCA 6

	Diff	Lower	Upper	Adj. P-Value
S. bipinnatifidum - S. biseriatum	-38169.652	-40236.3061	-36103	0
S. bipinnatifidum - S. rupestre	54506.4116	53543.0058	55469.8	0
S. biseriatum - S. rupestre	92676.064	90703.414	94648.7	0
S. bipinnatifidum - S. tenuifolium	-25876.652	-28145.7031	-23608	0
S. biseriatum - S. tenuifolium	12293.0004	9444.87995	15141.1	0
S. rupestre - S. tenuifolium	-80383.064	-82566.841	-78199	0

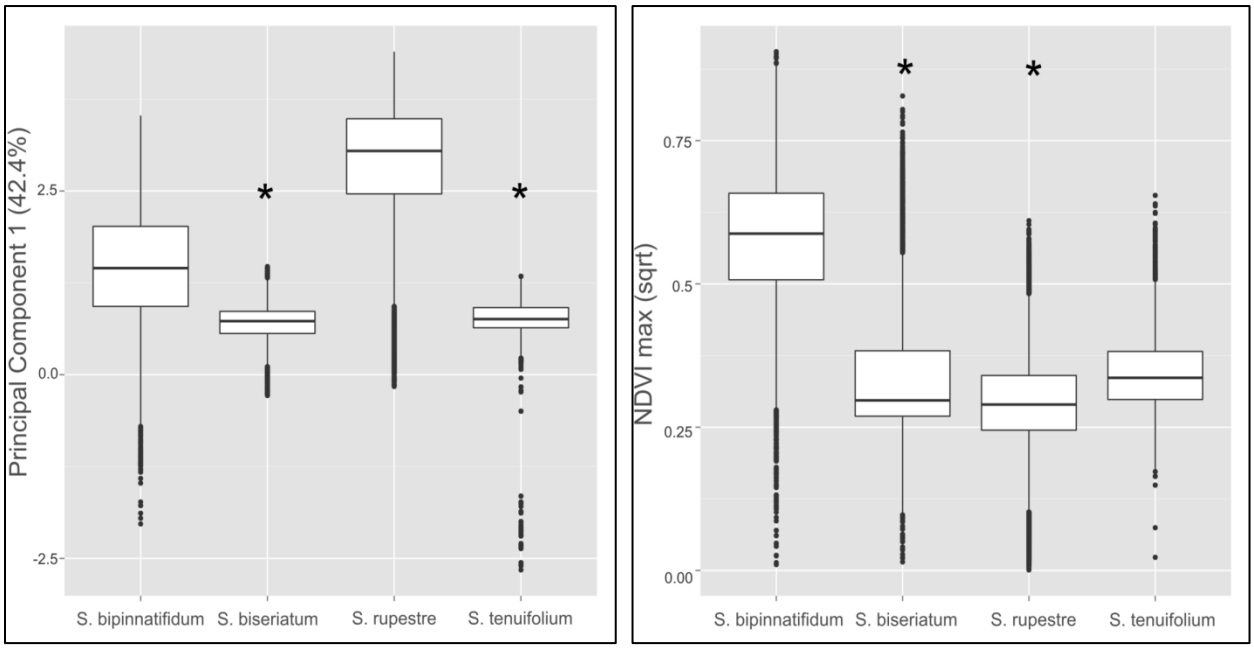


Fig. A. Boxplot of climatic variables (Principal Component 1: left, NDVI max: right) with statistical significant differences sampled from the areas of occurrence predicted per species in the ENM. Analyses were conducted using a Kruskal – Wallis and Dunn test using R. Significant pair comparisons are indicated in asterisk (*) and they was detected using a Bonferroni correction at $P < 0.003$. NDVI values were corrected with square root to maintain normality distribution.

Appendix 7

Populations sampled and the number of individuals sampled per population and region.

Population	Species	Locality	Longitude	Latitude	Altitude (m)	ELF8	MMT	CH4	ALDH
11	*S. tenuifolium*	5 N Highway, close to Copiapo	-70.45940	-27.56367	609	1	3		1
12	*S. biseriatum*	Road Copiapo to Caldera	-70.62078	-27.28177	450	2	4	1	
13	*S. biseriatum*	Road to Barranquilla	-70.74228	-27.48703	195	3	3	1	1
15	*S. biseriatum*	North exit of Caldera	-70.77060	-27.03323	127	3	2		1
16	*S. biseriatum*	Road to Quebrada El Leon	-70.69777	-27.01200	345	2	1		
20	*S. biseriatum*	Over dunes on coastal road from Caldera	-70.91055	-27.14238	55	3	3		
21	*S. biseriatum*	Road to Barranquilla and Quebrada Honda	-70.82785	-27.46878	122	3	2		
22	*S. biseriatum*	Crossroad Bahia Blanca - Copiapo	-70.85103	-27.66355	123		3	1	1
28	*S. tenuifolium*	Road to Bahia Salada, from 5 N Highway	-70.52107	-27.80750	464	2	1		1
29	*S. tenuifolium*	5 N Highway, in croassroad Chañarcillo - El Totoral	-70.54792	-27.94373	459	1	3		1
30	*S. tenuifolium*	Road to El Totoral	-70.78002	-27.94668	271	2	3	1	1
33	*S. tenuifolium*	Road to El Totoral from Route 1	-70.97933	-27.88290	111	1	3	1	1
35	*S. biseriatum*	Road Huasco - Carrizal Bajo	-71.15725	-28.11988	40	3	3	1	1
39	*S. biseriatum*	Between Caleta Totoral and Carrizal Bajo	-71.09163	-27.90572	151	2	3	3	2
41	*S. tenuifolium*	5 N highway	-70.68900	-28.23188	443	3	4		1
42	*S. tenuifolium*	Road to Canto del Agua and Carrizal Alto	-70.86875	-28.15097	270	1	4		1
48	*S. arcuatum*	Road to Carrizalillo, close to 5 N Highway	-70.99720	-28.97025	629	1	1	1	
49	*S. arcuatum*	Road to Carrizalillo	-71.28602	-28.94740	316		1	1	
51	*S. arcuatum*	Road to Carrizalillo and Quebrada Honda	-71.45478	-29.12052	35	1	1	1	1

Population	Species	Locality	Longitude	Latitude	Altitude (m)	ELF8	MMT	CH4	ALDH
78	*S. rupestre*	Road to Potrerillos and Cuesta Los Patos	-69.50922	-26.40263	2289	3	2	1	1
83	*S. rupestre*	Cerro Doña Ines	-69.21172	-26.12853	3826	1	2	2	2
85	*S. rupestre*	Close to Quebrada Pedernales	-69.28427	-26.36398	3599	1	2	2	
87	*S. rupestre*	Mining project Nueva Esperanza	-69.14905	-26.66127	3792	2	2	1	1
93	*S. rupestre*	Cuesta Codoceo	-69.23995	-26.88545	3652	2	1	2	1
98	*S. rupestre*	Road to Pircas Negras	-69.41585	-27.89652	3149	1	3		1
101	*S. rupestre*	Cuesta los Castaños	-69.71893	-27.72287	2940	2	3	1	1
104	*S. rupestre*	Secondary road from Barrick mine	-70.02743	-29.46715	3621	2	3		1
105	*S. rupestre*	Road to mining project Los Morros	-70.19302	-28.69558	2423	2	2		1
110	*S. rupestre*	Road to Paso Aguas Negras, La Laguna	-70.03703	-30.22390	2622		2	1	
112	*S. rupestre*	Juntas del Toro	-70.05018	-29.90052	2800	3	3		1
115	*S. bipinnatifidum*	Portezuelo 3 Cruces	-70.65225	-30.21923	2015	2	3	1	
s252	*S. maritimum*	Dunas de sector residencial de las Tacas	-71.36272	-30.09220	58	1	1	1	1
s244-24	*S. walkeri*	Dunas de Con Con (SGO 152008)	-32.93990	-71.54927	10	1	1	2	2

Appendix 8

List of screened Low Copy Nuclear Genes from IntrEST database

Region abbreviation	Locus' name	Chromosome location (*Arabidopsis*)	BRAD reference number	TAIR reference number	Flanked introns
ALDH108A	Aldehyde Dehydrogenase	Chromosome 1	TC64382	AT1G74920.1	5 and 6
ATEX070A1	Peroxisomal import receptor	Chromosome 5	TC63380	AT5G56290.1	2
ATGR1	Gluthiatone - disulfide reductase	Chromosome 3	TC64408	AT3G24170.3	4
Atprot	Unknown	Chromosome 3	TC88375	AT3G44330	7
BRL1	Kelch repeat - containing serine / threoine phosphatase family	Chromosome 4	TC80792	AT4G03080	4 and 5
CCC1	Cation-Chloride Co-Transporter	Chromosome 1	TC78524	AT1G30450.3	2
CH4	Cinnamate 4-Hydroxylase	Chromosome 2	TC63468	AT2G30490.1	1 and 2
ELF8	Early flowering 8	Chromosome 2	TC82509	AT2G06210.1	4
EIFE3	Eukaryotic Translation Initiation Factor	Chromosome 3	TC64270	AT3G57290	2 and 3
FAB1	Fatty Acid Biosynthesis	Chromosome 1	TC63446	AT1G74960.2	3 and 4
FASS1	Protein tonneau	Chromosome 5	TC77040	AT5G18580.1	4 and 5
GCN3	General Control Non-Represible	Chromosome 1	TC95073	AT1G64550.1	1, 2 and 3
MMT	Methionine S-methyltransfeerase	Chromosome 5	TC67627	AT5G49810.1	9 and 10
MFP2	Multifunctional Protein 2	Chromosome 3		AT3G06860.1	4 and 5
PEX5	Peroxisome Matrix Targeting signal -1	Chromosome 5	TC63380	AT5G56290.1	1
PGM	Phosphoglucomutase	Chromosome 5	TC63398	AT5G51820.1	4 and 5
PHS2	Alpha Glucan Phosphorilase	Chromosome 3	TC75018	AT3G46970.1	1 and 2
PYD4	Pyrimidine 4	Chromosome 3	TC70587	AT3G08860.1	undetermined
RGD3	Root growth defective 3	Chromosome 3	TC99814	AT3G54280	2 and 3

Region abbreviation	Locus' name	Chromosome location (*Arabidopsis*)	BRAD reference number	TAIR reference number	Flanked introns
THFS	10 - Formyletetrahydrofolase synthetase	Chromosome 1	TC63707	AT1G50480.1	1
VHA	Vacuolar ATP synthase subunit A	Chromosome 1	TC77040	AT1G78900.1	6 and 7

BRAD: The Brassica Database
TAIR: The Arabidopsis Information Resource

Appendix 9

Scheme of indels detected in ELF8 and MMT regions

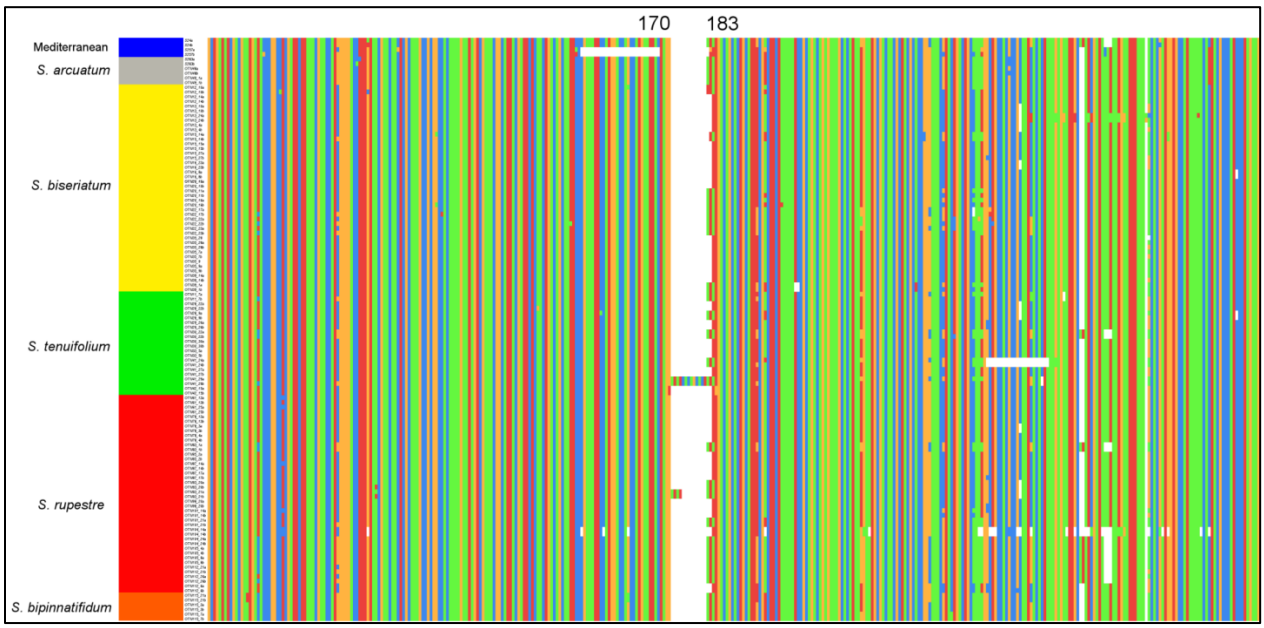


Fig. A. Indel scheme of in ELF8 regions organized by species. Indels are represented by blank spaces and nucleotide bases in colors: red: adenine, green: thymine, blue: cytosine, and orange: guanine. Large shared deletion (positions 170 to 183) is indicated above the matrix.

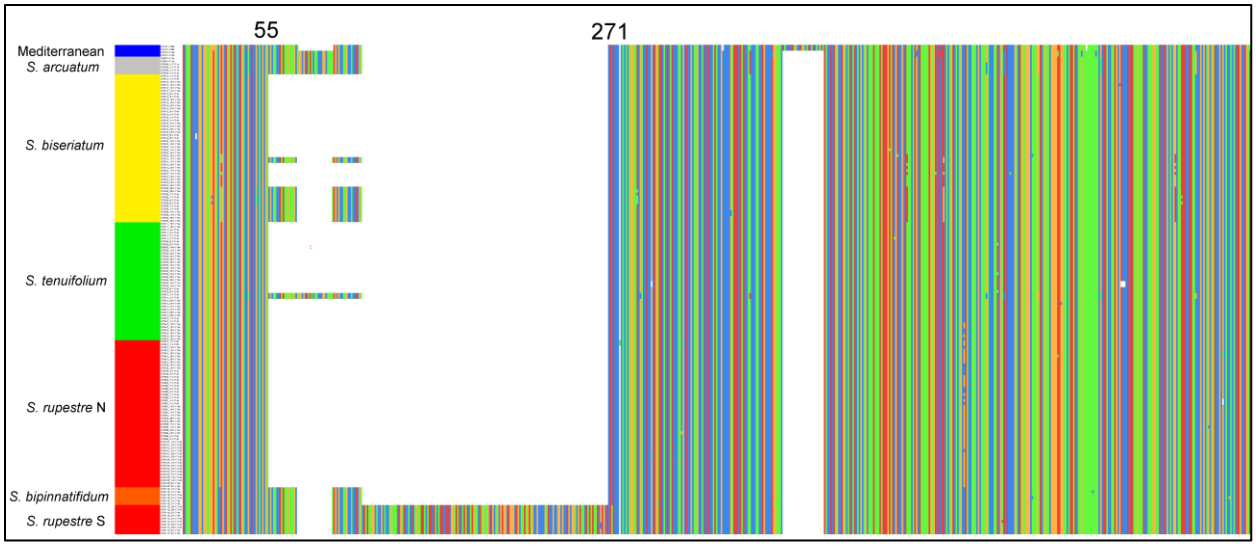


Fig. B. Indel scheme of in ELF8 regions organized by species and haplogroups in *S. rupestre*. Indels are represented by blank spaces and nucleotide bases in colors: red: adenine, green: thymine, blue: cytosine, and orange: guanine. Large shared indel (positions 55 to 271) is indicated above the matrix.

Appendix 10

Estimate of species distribution based on habitat suitability for present and Quaternary times using ENM

The present analysis was conducted employing MaxEnt v. 3.3.3a (Phillips & al., 2006) to infer the potential geographic distribution of the species of the Atacama subclade during present and Quaternary times (120 mya). The estimate was based on paleo-climatic models from the Last Glacial Maximum (21 mya) from CCSM and MIROC models, generated by the Paleoclimate Modelling Intercoparson Project Phase (PIMP2). All bioclimatic layers (present and last glaciar maximum) employed in this analysis were obtained from Worldclim database (Hijmans & al., 2005) at 2.5 arcsec of resolution.

A training area was determined using a larger spatial range than currently known for *Schizopetalon* species. This area encompassed the absolute desert (o mm; 20°00'S) to the end of the Mediterranean climate (Concepcion; 37°16'S) and covered a part of Argentina beyond the area of influence of the Andes (66°05'W). Highly correlated bioclimatic layers ($r > 0.8$) were removed using a moment-product correlation analysis. Eight variables were selected from the original 19. From the temperature variables, bio1 (annual mean temperature), bio2 (mean diurnal temperature range), bio3 (isothermality), bio4 (maximum temperature of the warmest month), and bio9 (mean temperature of the coldest month) were chosen. From precipitation variables, bio13 (precipitation of the wettest month), bio14 (precipitation of the driest month), bio15 (precipitation seasonality), bio16 (precipitation of the wettest quarter), and bio18 (precipitation of the warmest quarter) were chosen. A total of 114 confirmed points from *S. biseriatum*, *S.*

tenuifolium, and *S. rupestre* were used to conduct this analysis (see Appendix 2). Results were projected onto binary maps that represented an absence or presence under a 10% minimum training presence to detect spatial overlap. All analyses were conducted in R (R core development group) with the package *raster* v. 2.35 (Hijmans & van Etten, 2012).

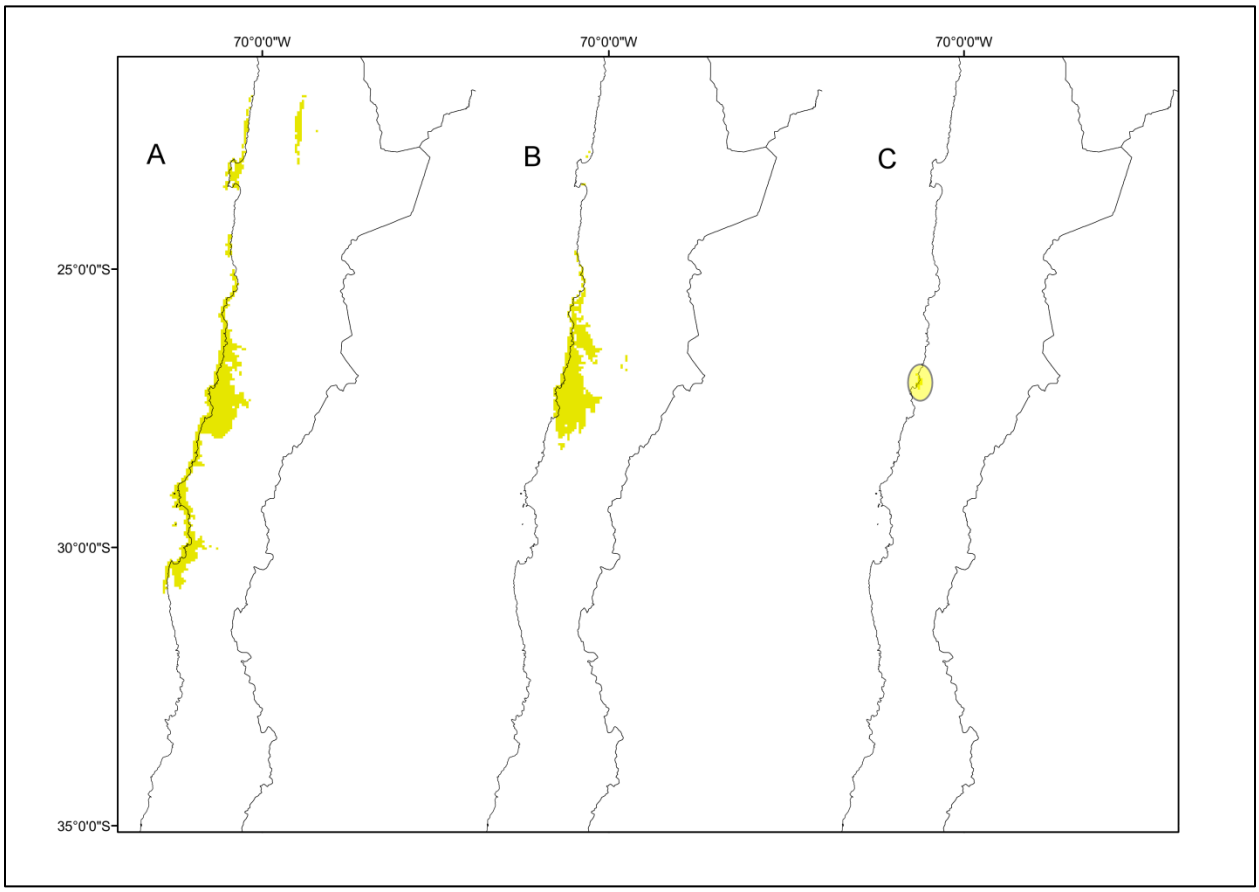


Fig. A. Model projection of suitable areas during the present (A) and Last Glacial Maximum (B: CCSM, C: MIROC) in *S. biseriatum* (yellow). Circle represents small presence area detected in the Miroc model.

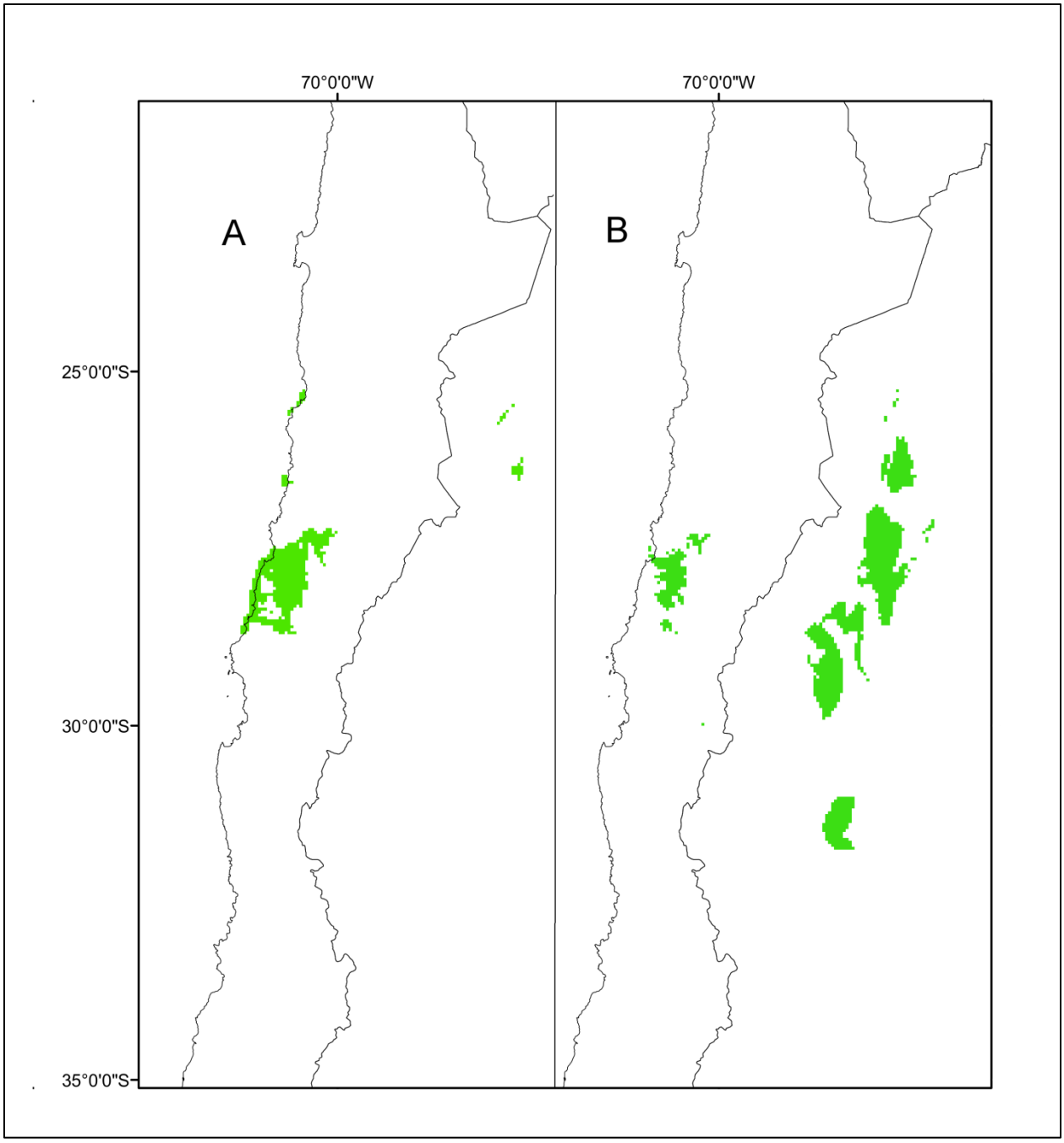


Fig. B. Model projection of suitable areas during the present (A) and Last Glacial Maximum (B: CCSM) in *S. tenuifolium*. The projection of MIROC model was omitted due to the null presence of suitable habitats in the training area.

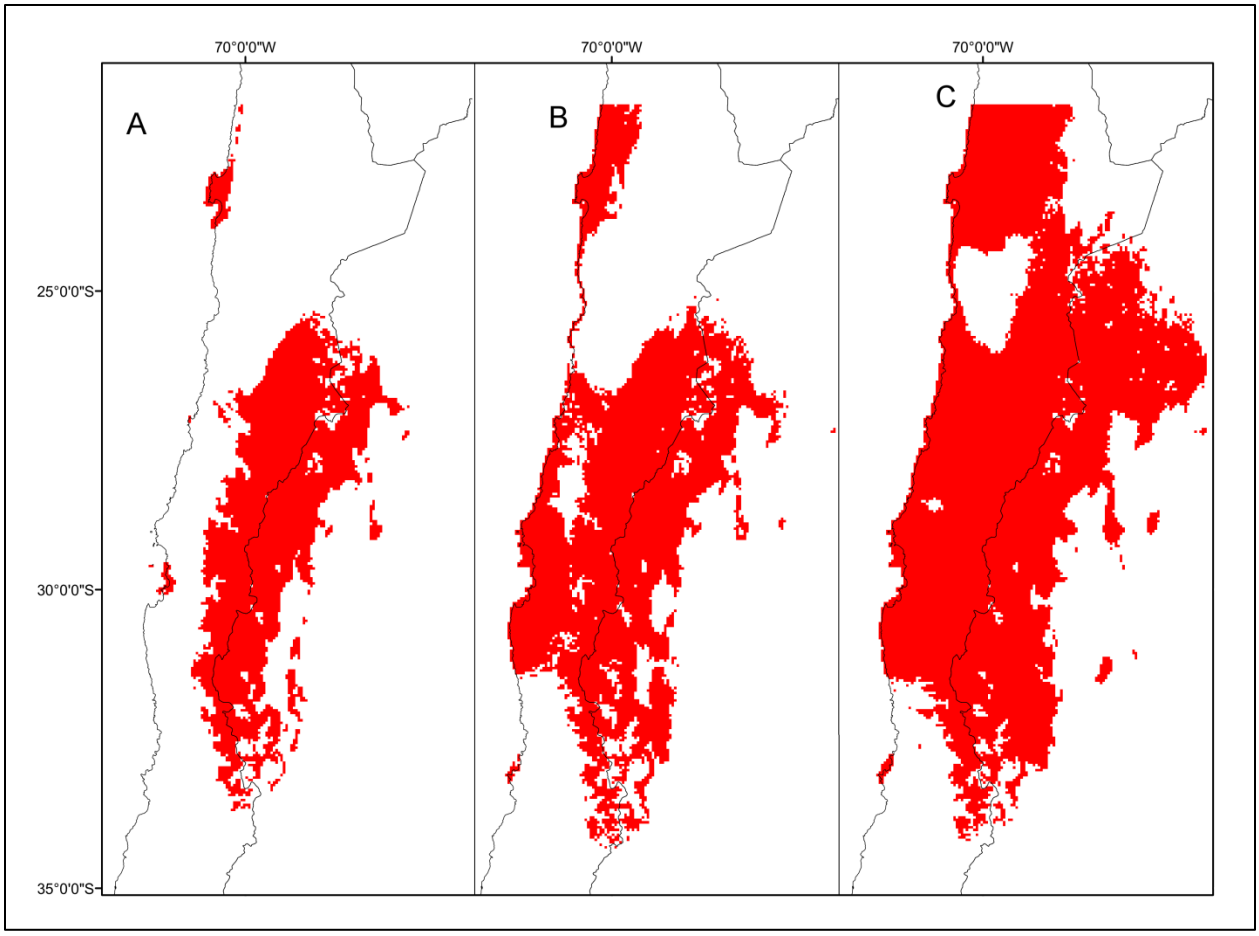


Fig. C. Model projection of suitable areas during the present (A) and Last Glacial Maximum (B: CCSM, C: MIROC) in *S. rupstre*.