ABSTRACT

Title of dissertation: A PHYLOGENETIC AND BIOGEOGRAPHIC ANALYSIS OF SANGUISORBEAE (ROSACEAE), WITH EMPHASIS ON THE PLEISTOCENE RADIATION OF THE HIGH ANDEAN GENUS *POLYLEPIS*.

Malin Sofia Kerr, Doctor of Philosophy, 2004

Dissertation directed by: Dr. Charles F. Delwiche
Dr. James L. Reveal

A phylogenetic and biogeographic analysis of the tribe Sanguisorbeae (Rosaceae) was conducted with emphasis on the radiation of the Andean tree *Polylepis*. Phylogenetic analyses of coding and non-coding nuclear markers reveal a complex evolutionary history of the tribe including ancient and recent allopolyploid hybridization. *Sanguisorba sensu lato* is shown to be paraphyletic and split between the allopolyploid hybrid *Sanguisorba* and the non-hybrid *Poterium* and *Poteridium*. A monophyletic origin of the southern hemispheric subtribe Sanguisorbinae is supported, and this clade is given a phylogentic taxon name (*Verruchaena*). Dating analyses using the penalized likelihood method suggest that this taxon originated in the late Miocene. A biogeographic hypothesis is presented in which *Verruchaena* originated in the New World with subsequent transoceanic dispersals to southern Africa and Australasia. The paramo genus *Polylepis* most likely arose from hybridization between two Andean ancestors supporting a "vertical" rather than "horizontal" origin of this taxon.

A PHYLOGENETIC AND BIOGEOGRAPHIC ANALYSIS OF SANGUISORBEAE (ROSACEAE), WITH EMPHASIS ON THE PLEISTOCENE RADIATION OF THE HIGH ANDEAN GENUS *POLYLEPIS*.

by

Malin Sofia Kerr

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2004

Advisory Committee:

Associate Professor Charles F. Delwiche, Chair/Advisor
Research Associate Torsten Eriksson
Associate Professor Irwin N. Forseth, Jr.
Associate Professor Patrick S. Herendeen
Professor Charles Mitter

AKNOWLEDGEMENTS

I would like to thank Dr. James L. Reveal, who was my primary advisor for two years before his retirement, for enthusiastically sharing his immense expertise in classical plant taxonomy and nomenclature. The educational foundation provided by Dr. Reveal was perfectly complemented by the guidance and support I received from Dr. Charles F. Delwiche during the following four years, primarily in the field of molecular phylogenetics. I am grateful to Dr. Delwiche for giving me the confidence and freedom to allow my research to develop in new directions. My dissertation research benefited greatly from the daily interactions I had in the Delwiche lab with students and staff. I am particularly grateful to Tsetso Bachvaroff, Mathew Cimino, John Hall, Ken Karol, Jeff Lewandowski, Tanya Marushak and Maria Virginia Sanchez Puerta for support and inspiration.

I would also like to thank Dr. Torsten Eriksson for introducing me to my study group, the Sanguisorbeae, and for countless inspiring discussions on Rosaceae evolution and on systematics in general. I am also greatly indebted to several people who generously shared rare plant material with me, in particular Michael Kessler (Universität Göttingen), Christopher Whitehouse (University of Cape Town), and John Clark (George Washington University). In addition I would like to acknowledge Daniel Larrea and Mónica Moraes R. (National Herbarium of Bolivia) for their valuable assistance, making collecting in Bolivia a very pleasant and productive experience.

The following institutions generously provided herbarium and/or fresh garden material for study: The Royal Botanic Garden Edinburgh, the New York Botanical Garden, the Missouri Botanical Garden, the Gray Herbarium (Harvard University), and the National Herbarium at the Smithsonian. My dissertation research was funded by grants from the National Science Foundation (Doctoral Dissertation Improvement Grant), the American Society of Plant Taxonomists, and the Department of Cell Biology and Molecular Evolution at the University of Maryland.

Finally I would like to acknowledge my Swedish family who cheered me on despite the geographic distance, and most of all I want to thank my husband Dr. Peter H. Kerr who shared with me this experience from start to finish as we entered graduate school on the same day and defended our dissertations in the same week. Without Peter I would have never started on this path in the first place and undoubtedly would not have had the endurance to see it through.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I General Introduction

Introduction

The high Andes of South America support a unique flora of high endemism and diversity. The vegetation zone above the continuous treeline in the Andes is referred to as paramo (or puna in the south). Compared with the lowland Neotropical forest, the history and diversity of the paramo biota has received relatively little attention among systematists. The biogeographic and evolutionary origin of the unusual paramo flora is particularly interesting considering the relatively recent geological origin (Plio-Pleistocene) of the high Andean habitats (Simpson, 1975; Van der Hammen, 1974). Certain elements of the paramo flora appear to have evolved in situ from lowland ancestors, while other elements dispersed to the paramo from cool regions elsewhere, e.g., the Nearctic (Cleef, 1979; Gentry, 1982; Simpson and Neff, 1985; Smith and Cleef, 1988; Van der Hammen, 1979). It has been debated whether the principle mode of evolution of tropical alpine floras is vertical, i.e., gradual evolution of alpine species from lowland ancestors, or horizontal, i.e., via "hopping" or long distance dispersal from similar environments elsewhere (Chardon, 1938; Cleef, 1979; Simpson, 1975; Van der Hammen, 1979; Van der Hammen and Cleef, 1986; Vuilleumier, 1970).

The arborescent high Andean genus *Polylepis* Ruíz & Pavón (Rosaceae Juss.) is a vital element in the paramo and puna vegetation. Some species of *Polylepis* can grow at elevations above 5000 m, which is probably higher than any other arborescent angiosperm in the world (Simpson, 1979). The phylogenetic and biogeographic history of *Polylepis* and its close relatives in the tribe Sanguisorbeae

Bercht. & J.Presl has never been analyzed comprehensively. Among its close relatives are genera with markedly different distributions, some being northern hemispheric (e.g., *Sanguisorba* L.) while others have an austral subantarctic distribution (e.g., *Acaena* Mutis ex L.). Several putative morphological synapomorphies suggest that *Polylepis* constitutes a monophyletic group. However, the relationship of *Polylepis* to other genera in the Sanguisorbeae is largely unknown and it is therefore unclear whether *Polylepis* evolved via vertical evolution (possibly from a South American ancestral *Acaena*, or *Margyricarpus* Ruíz & Pavón) or via horizontal migration either from the southern temperate parts of South America or from the Nearctic via the Mesoamerican land bridge. Furthermore, it has been suggested that the evolution of *Polylepis*, and the paramo biota as a whole, was influenced, and the speciation accelerated, by the Pleistocene climatic fluctuations that repeatedly fragmented and reunited patches of paramo vegetation (Hooghiemstra, 1984; Simpson, 1983; Van der Hammen, 1974; Van der Hammen and Cleef, 1986; Vuilleumier, 1971). This hypothesis, however, has never been investigated within a phylogenetic framework.

In this dissertation, the evolutionary and biogeographic history of *Polylepis* and its close relatives was investigated through a phylogenetic analysis of extant members of petalous (Agrimoniinae J. Presl) and apetalous (Sanguisorbinae Torr. & A. Gray) Sanguisorbeae. Phylogenetically informative data were obtained from nucleotide sequences of the chloroplast and the nuclear genomes. The molecular phylogenies together with extant distribution patterns were used to estimate a biogeographic history of Sanguisorbeae with emphasis on the origin and radiation of

the southern hemispheric members of the apetalous Sanguisorbinae. The main aim of the phylogenetic analyses was to resolve the question of the geographic origin of the southern Sanguisorbinae as a whole and of *Polylepis* in particular.

The results reveal an evolutionary history driven by allopolyploid hybridization, both ancient and recent. The origin of the tribe most likely occurred in the northern hemisphere followed by two separate migrations to the southern hemisphere, within Agrimoniinae and Sanguisorbinae respectively. The southern hemispheric distribution of Sanguisorbinae in turn is explained by several transoceanic dispersal events. The molecular phylogenetic analyses suggest that the origin of *Polylepis* was "vertical" from lower elevation South American ancestors, almost certainly as a result of a hybridization event. The results are also consistent with the hypothesis that Pleistocene vicariance events influenced evolution in paramo plants, although estimates of speciation rates are ambiguous due to the lack of biological or phylogenetic species in *Polylepis*.

Morphology, ecology and distribution

The tribe Sanguisorbeae (Rosaceae: Rosoideae) is found in cool temperate or alpine regions on all continents and includes herbaceous as well as shrubby and arborescent species. Members of the Sanguisorbeae are distinguished from the rest of subfamily Rosoideae by their cup-shaped hypanthium that entirely encloses the carpel(s), resulting in a perigynous position of the flower. They are also characterized by a reduction in the number of stamens and carpels (in many cases to a single carpel). The shape of the hypanthium varies greatly, from rounded and smoothed to

winged, fleshy, or with different kinds of spines. The leaves of Sanguisorbeae are stipulate and mostly pinnate, although exceptions are found within the genus *Cliffortia*.

Agrimoniinae

The petalous members of Sanguisorbeae make up five relatively small genera, *Agrimonia* L. being the largest with approximately 15 herbaceous northern temperate species. Similar in habit to *Agrimonia* are *Aremonia* Necker ex Nestler, a monotypic genus of southeast Europe, and *Spenceria* Trimen from western China. All three genera have erect racemes producing yellow or cream flowers. The magnificent umbrella shaped trees of *Hagenia abyssinica* (Bruce) J. Gmel. can reach up to 30 m in height in the subalpine zone of East African mountains (at 2400-3600 masl). The flowers of this dioceous species are borne on large pendulous racemes. The dried female flowers are used medicinally and plant extracts purportedly contain anti-cancer compounds. The shrubby or arborescent genus *Leucosidea* Eckl. et Zeyh. grows in eastern South Africa and Zimbabwe.

Sanguisorbinae

Morphological features such as inconspicuous apetalous inflorescences and large fimbrillate stigmas have been pointed to as evidence of wind-pollination in members of the subtribe Sanguisorbinae. Although most species of *Sanguisorba* display features associated with anemophily, a few species have brightly colored calyx and compact stigmas (e.g., *S. officinalis* and allies) and have been suggested to

be insect-pollinated (Nordborg, 1966). Pollination mechanisms have not been studied carefully in Sanguisorbeae and it is possible that additional taxa at least in part are entemophilous. Large quantities of *Acaena* pollen have been found in honey (Forcone and Tellería, 2000) suggesting that at least domestic bees are visiting the flowers. Whether this occurs in the wild is unclear.

Sanguisorbinae are represented in the northern hemisphere by the widespread herbaceous northern temperate *Sanguisorba* (including *Poterium* and *Poteridium*), the monotypic Mediterranean shrub *Sarcopoterium* Spach, and the woody Canarian endemics *Bencomia* Webb & Berth., *Marcetella* Svent., and *Dendriopoterium* Svent. The remaining Sanguisorbinae are restricted to the southern hemisphere, including the Andean genera *Polylepis, Margyricarpus*, and *Tetraglochin*, the South African *Cliffortia* L., and the diverse *Acaena,* which has a widespread transcontinental distribution. *Margyricarpus* and *Tetraglochin* are dwarf sclerophyllous shrubs, while the species-rich (>100 spp.) and morphologically diverse *Cliffortia* includes a range of subfrutescent to arborescent growth forms, most of which are native to the highly endemic fynbos flora of the Cape region.

The genus *Acaena* (c.100 spp.) is herbaceous, but some species have a suffrutescent stem base or produce tough underground stolons giving rise to mat-like vegetative growth. *Acaena* displays a disjunct austral subantarctic biogeography reminiscent of a Gondwanan relict distribution. The majority of *Acaena* species are found in South and Central America; mostly in the southern temperate region and along the Andes, but one species reaches as far north as California. There are six species in Australia, 17 in New Zealand, and a single species is native to South Africa

(*A. latebrosa*) and Hawaii (*A. exigua*) respectively. One species (*A. masafuerana*) is endemic to the montane flora of the Juan Fernandez Islands, and several species are found on the Falkland Islands. A few subantarctic species reach as far south as the islands of Tristan da Cunha and South Georgia. In addition, several Australasian species of *Acaena* were introduced into Great Britain in the early 1900's and have been naturalized.

Polylepis

Polylepis consists of 15-25 spp. of small trees and shrubs with imparipinnate leaves and bisexual flowers arranged in dense racemes. The crouching crooked stems covered with exfoliating layers of bark and the sclerophyllous, often hirsute, leaves reflect adaptations to the harsh conditions of the high Andean environment. Several morphological features suggest that *Polylepis* is a monophyletic genus, e.g., arborescent habit, pubescent anthers, fruits with spines or wings (but lacking barbs). In the northern Andes (Colombia, Venezuela, Ecuador, and northern Peru), most *Polylepis* species occur in the alpine vegetation zone paramo, but a few species also grow at lower elevations in the upper montane forest. In the southern Andes (southern Peru, Bolivia and northern Chile) the corresponding alpine vegetation type is called puna and tends to be dryer and more seasonal than the paramo. Some *Polylepis* species form patches of forest reaching elevations over 5000 m, well above the upper continuous forest limit (Simpson, 1979). The success of *Polylepis* trees at such high elevations appears to be due to a combination of several physiological adaptations and the microclimate of the environment where *Polylepis* is found (Goldstein *et al.*,

1994; Rada *et al.*, 2001; Rada *et al.*, 1985; Rauh, 1956a; Rauh, 1956b; Simpson, 1979; Troll, 1959; Walter and Medina, 1969).

Previous taxonomic treatments of the Sanguisorbeae and *Polylepis*

The Rosaceae include many commercially important fruit crops and ornamentals and have therefore received considerable attention since the early history of systematic botany. However, the mostly southern hemispheric, commercially insignificant Sanguisorbeae has been largely overlooked, with the exception of a handful of monographs and revisions on particular genera. Among the few influential works on the Sanguisorbeae are Bitter's revisions of the genera *Acaena* (1911a) and *Polylepis* (1911b), Weimarck's (1934) monograph on *Cliffortia* L., Nordborg's (1966; 1967) taxonomic studies on *Sanguisorba* and closely allied genera, and Simpson's (1979) and Kessler's (1995a; 1995b) treatments of *Polylepis.* According to our current understanding of relationships within the Rosaceae, the tribe Sanguisorbeae comprises the following genera: *Acaena* (including *Ancistrum*), *Agrimonia, Aremonia, Bencomia, Cliffortia, Dendriopoterium, Hagenia, Leucosidea, Marcetella, Margyricarpus*, *Polylepis*, *Sanguisorba* (including *Poterium* and *Poteridium*) *Sarcopoterium, Spenceria* and *Tetraglochin*. Sanguisorbeae was first recognized as a group by Jussieux (1789) who included eleven genera: *Poterium, Sanguisorba, Acaena, Ancistrum, Cliffortia, Agrimonia, Neurada, Alchemilla, Aphanes* and *Sibbaldia*. Although the latter four genera have subsequently been removed from the group, Jussieux astutely included every genus that we now recognize as part of Sanguisorbeae that had been described by 1789. *Neurada* has

subsequently been removed from Rosaceae altogether and, based on morphological (e.g., lateral insertion of the style) as well as molecular characters, there is now sufficient evidence that *Alchemilla, Aphanes* and *Sibbaldia* are more closely related to *Potentilla* and *Fragaria* in the Potentilleae than to the Sanguisorbeae (Eriksson *et al.*, 1998). Since Jussieux, additional genera have been included in Sanguisorbeae as they have been discovered and described: *Polylepis* (Ruiz and Pavon, 1794), *Bencomia* and *Marcetella* (as *Poterium*) (Webb and Berthelot, 1836-50), *Hagenia* (Gmelin, 1791), *Leucosidea* (Ecklon and Zeyher, 1836), *Margyricarpus* (Ruiz and Pavon, 1794), and *Spenceria* (Trimen, 1879).

Sanguisorba

Linnaeus (1753) described two separate genera, *Sanguisorba* and *Poterium*. Among the characters he cited to distinguish the two were hermaphroditic versus unisexual flowers and one versus two styles in *Sanguisorba* and *Poterium,* respectively. Scopoli (1772) joined the two genera under *Sanguisorba.* Since then authors have alternated between recognizing one or two genera. Occasionally the species have been joined under the name *Poterium* rather than *Sanguisorba*. Most of the recent literature, however, treats them in a single genus as *Sanguisorba* and this convention was followed (a priori) to species included in this study. The Mediterranean sclerophyllous shrub *Sarcopoterium* (*Poterium spinosum* L.) and the North American herb *Poteridium* (*Poterium annuum* Nutt. Ex Hook.) were elevated to genus status by Spach in 1846. The Canarian tree *Bencomia* (*Poterium caudatum* Aiton) was likewise removed to its own genus by Webb and Berthelot (1836-50).

Cliffortia and Acaena

Cliffortia was studied by early botanists such as Linnaeus, but a comprehensive treatment of the genus was not available until Weimarck (1934) published his impressively detailed monograph of the genus. Little has been added to the taxonomy of *Cliffortia* since then, but a phylogenetic study is currently under way at the University of Cape Town (Christopher Whitehouse pers. com.).

Based on material from Mexico, Linnaeus (1771) described a single species of *Acaena* (*A. elongata*). The first description of an *Acaena* (as *Ancistrum*) from Australia was by J.R. and G. Forster as part of James Cook's second expedition (Forster and Forster, 1776). Vahl (1805) synonymized *Acaena* and *Ancistrum*. De Candolle (1825) introduced infrageneric ranks by separating the genus into the sections *Euacaena* (species with fruits entirely covered with spines), and *Ancistrum* (species with spines only at the fruit apex). Reiche used the same groupings but changed the rank of *Euacaena* and *Ancistrum* to subgenus and stressed the difference in inflorescence type (spikes versus heads) in addition to the distribution of spines on the fruit. Citerne (1897) introduced seven sections that were adopted and added to by Bitter (1911a) in his massive monograph of the genus. Before Bitter published his work, approximately 25 species of *Acaena* were recognized. Bitter, who was a proponent of splitting taxa into its smallest recognizable units, divided *Acaena* into 110 species and a myriad of subspecies and varieties. Even though Bitter's monograph is an impressive 336 pages long, it contains very few illustrations and it is difficult from his descriptions to distinguish closely related species. His species keys

often rely on characters that have the potential to vary within species as much as between them (e.g., the number of leaflets, the size of leaves and inflorescences, and the density and positions of hairs). Taxonomically more useful are Bitter's infrageneric entities. Based on inflorescence type Bitter divided *Acaena* into two "series": *Axillares* and *Terminales*, below which a total of ten sections were recognized. Grodona (1964) published a useful revision of *Acaena* that included detailed illustrations, but he only covered the 20 species found in Argentina. Even though Bitter's typological species concept can be justifiably criticized and dismissed, his complex classification scheme indicates that there is significant inter-population diversity present in some species of *Acaena*. Whether this morphological variation is the result of phenotypic plasticity, ecological differentiation, hybrid zones or apomixis, remains to be explored.

Polylepis

The genus *Polylepis* (type species: *P. racemosa* R. et P.) was first described by Ruiz and Pavon (1794) and was subsequently expanded by Hieronymus (1895; 1896) and Pilger (1906) among others. Bitter (1911b), who published the first detailed revision of *Polylepis*, recognized 33 species and numerous varieties. He divided the genus into two sections, *Dendracaena* and *Gymnopodae*, along with eleven unranked subgeneric groups. Simpson (1979) adopted a broader species concept in her revision of the genus. She described 15 species (later adding one more) forming what she considered three natural groups. Kessler (1995a) revised the

circumscription of Bolivian species of *Polylepis* adding *P. neglecta* along with several subspecies.

Previous phylogenetic studies of Sanguisorbeae

Rosoideae sensu stricto constitutes a monophyletic group encompassing the tribes Ulmarieae, Rubeae, Roseae, Coluriae, Potentilleae and Sanguisorbeae (Eriksson *et al.*, 1998; Eriksson *et al.*, 2003; Morgan *et al.*, 1994). Roseae, Potentilleae, and Sanguisorbeae share operculate pollen (Eide, 1981; Reitsma, 1966) and molecular analyses agree that these tribes are closely related. However, despite the availability of sequence data from *rbcL*, *trnL/F*, *matK* and ITS, the exact sister group of Sanguisorbeae remains ambiguous. In a phylogenetic analysis of Rosaceae based on *rbcL*, *Agrimonia* (the only Sanguisorbeae sampled) appeared as sister to a clade of Roseae and Potentilleae (Morgan *et al.*, 1994). The node joining Roseae and Potentilleae, however, had very little support (decay=0). Eriksson *et al.* (2003) conducted a phylogenetic analysis of the subfamily Rosoideae based on nuclear ITS and chloroplast *trnL/F* sequences. The analysis focused on Potentilleae but included nine Sanguisorbeae: the apetalous *Polylepis tarapacana, P. hieronymi, Acaena laevigata, A. cylindristachya, Tetraglochin strictum, Sanguisorba officinalis* and three petalous taxa (*Aremonia, Leucosidea*, and *Agrimonia*). Contrary to the *rbcL* phylogeny, ITS data as well as the combined *trnL/F*+ITS phylogeny supported a sister relationship between Sanguisorbeae and Potentilleae at the exclusion of Roseae. The phylogeny based on *trnL/F* data alone was ambiguous as to the relationship between the three groups. Mishima (2002) investigated the relationships within the

genus *Sanguisorba* using chloroplast sequence data from *matK*. The most parsimonious *matK* phylogeny showed Roseae to be sister to Sanguisorbeae at the exclusion of Potentilleae (bootstrap support 72%). There are at least two plausible explanations for the difficulty in determining the relationships among these three clades. Either the divergence of the three groups occurred in quick succession leaving little information on the internal branch, or ancient hybridization among the lineages makes chloroplast and nuclear gene trees incongruent. The analyses presented here suggest that a combination of both is most likely responsible (see chapter III).

The analysis of Eriksson *et al.* (2003) supported the monophyly of Sanguisorbeae as well as of Sanguisorbinae and Agrimoniinae. In an earlier analysis based on ITS data alone Sanguisorbeae appeared polyphyletic with Agrimoniinae nested within *Potentilla* (Eriksson *et al.*, 1998). However, forcing Sanguisorbeae to be monophyletic was shown to be only one step longer and this relationship was argued to be more plausible. Again, an ancient hybridization event involving the Agrimoniinae lineage may be responsible for the ambiguity. The relationship of *Polylepis* to other genera of the Sanguisorbeae is largely unknown. The analysis of Eriksson *et al.* (2003) supported a close relationship between *Polylepis* and *Acaena* and also hinted at the possibility that *Acaena* is paraphyletic with respect to *Polylepis* and *Tetraglochin*. Helfgott *et al.* (2000) sequenced the nuclear ITS region of 20 species of Sanguisorbeae with the aim of clarifying the origin of the Canarian endemic *Bencomia*-complex. Their results also indicated that *Polylepis* is closely related to *Acaena*. However, since the focus of this study was other relationships within the tribe, only one species each of *Polylepis* and *Acaena* were sampled, and no

representatives of *Margyricarpus* and *Tetraglochin.* A much more extensive sampling of taxa, together with more informative phylogenetic markers, is necessary to shed light on the origin of *Polylepis*. The only existing infrageneric phylogenetic study of *Polylepis* was conducted by Kessler (1995b) based on morphological data.

Species recognition in *Polylepis*

Many species of *Polylepis* have a fragmented distribution pattern with little contact between populations, which has led to considerable intraspecific variation. The problematic taxonomy of *Polylepis* is reflected in the discrepancy between the generic classifications proposed by Bitter (1911b), Simpson (1979) and Kessler (1995a; 1995b). However, its high ploidy level and complex infrageneric "phylogenetic" relationships (see chapter II and VI) suggest another level of genetic complexity. A combination of random sorting of ancestral alleles, hybridization, and/or facultative apomixis need to be invoked to explain the current data. The application of species concepts in *Polylepis* will be discussed in chapter VI.

Objectives and significance

Objectives

The objectives of this study were to address the following questions:

(1) Where did Sanguisorbinae originate and how is this group related to the petalous Agrimoniinae? What biogeographic mechanisms (vicariance, dispersal etc.) best explain the current disjunct distribution of Sanguisorbinae?

(2) How is *Polylepis* related to other members of Sanguisorbinae and what does this tell us about the geographic origin of this taxon and the paramo flora as a whole? Is *Polylepis* derived from lower elevation ancestors on the South American continent or from temperate (e.g., Nearctic or subantarctic) ancestors via migration?

(3) Was the rate or mode of speciation in *Polylepis* influenced by the Pleistocene glacial cycles; i.e., does the species-pump model apply to *Polylepis*?

Significance

Ecological studies of biotas typically emphasize local environmental conditions and interspecific interactions in the development of plant and animal communities. However, historical processes often have a strong influence on the composition of biotas. Geological and biogeographic vicariance events can have a significant effect on which lineages have the opportunity to inhabit a particular area at a particular time. In this context, phylogenies provide a means for distinguishing historical from nonhistorical contributions to the composition of communities (Brooks and McLennan, 1993). Despite the growing recognition of the importance of history in diversity and composition of biotas, phylogenetic data necessary for evaluating the historical effects are often lacking.

Polylepis is a key floral element in the high Andean environment, and determining the biogeographic origin of this genus could shed light on the history of the paramo flora as a whole. Surveys of the distribution and diversity of the genus *Polylepis* are also of high priority from a conservation standpoint. The present distribution of *Polylepis* woodland is very fragmented and it is likely that it represents

only relictual patches of what was once a much more widespread genus. It is thought that its distribution has been greatly reduced by human activities (Brandbyge, 1992; Ellenberg, 1979; Fjeldså, 1992; Van der Hammen, 1979) that have continued and even increased in recent times under the pressure of a growing rural Andean population. In addition to the threat to *Polylepis* itself, many other species are dependent on the habitat that *Polylepis* provides. The densely laminated bark and the "stacks" of dead leaves on mature plants provide diverse feeding niches for birds. Over a hundred species of bird occur regularly in *Polylepis* woodland, with many being local endemics completely dependent upon *Polylepis* (Fjeldså, 1992). Despite the strong interest in and importance of *Polylepis*, few data exist either for the distribution of the species or for genetic diversity among them.

The Sanguisorbeae provide opportunities to explore Tertiary biogeography by vicariance or dispersal. Nuclear molecular phylogenetic data of Sanguisorbeae can also shed light on the evolution of polyploid lineages. *Polylepis* is an ideal model organism not only to study Quaternary biogeography but also for studies of tropical alpine ecology, adaptations to environmental stresses, distribution and evolution of the unique avifauna, human influences on the paramo habitat, and needs and prospects for conservation of the paramo habitats. Systematic studies (in particular phylogenetic trees) provide fundamental historical data that are used in virtually all other fields of biology. The phylogenetic data from this study will provide the basis for further research on molecular and morphological evolution of Sanguisorbeae and *Polylepis*.

Chapter II Phylogeny and character evolution in the Sanguisorbeae inferred from chloroplast sequence data

Introduction

The tribe Sanguisorbeae is a morphologically diverse lineage that has received little notice among systematists, especially when compared to other taxonomic groups in the Rosaceae. Despite the lack of interest, or maybe because of it, the Sanguisorbeae clearly merit closer attention. Several members of the Sanguisorbeae are components of rare, highly endemic floras such as the tropical alpine floras of the Andes and East Africa, and the Cape fynbos vegetation of South Africa. The phylogenetic relationships of these taxa have never been studied comprehensively. Multiple origins of unisexual flowers and secondary growth in Sanguisorbeae also provide excellent opportunities to study directionality in these traits and to test correlations of morphological adaptations.

Sanguisorbeae's sister lineages, Potentilleae and Roseae, are typically herbaceous or frutescent while the Sanguisorbeae include a mix of herbaceous and arborescent taxa. Previous molecular studies have indicated that the different growth forms do not represent monophyletic lineages but that either woodiness or non-woodiness has evolved many times (Eriksson *et al.*, 2003; Helfgott *et al.*, 2000). The Sanguisorbeae are characterized by reductions in numbers of floral parts. In addition to the loss of petals and epicalyx in Sanguisorbinae, there has been a gradual reduction in the number of stamens and pistils from presumably numerous in the common ancestor (as in the sister lineages) to as few as two stamens and one carpel.

Whether this reduction has occurred sequentially or represents repeated independent losses can only be tested within a phylogenetic framework.

The phylogenetic relationships within Sanguisorbeae were investigated by analyzing DNA sequences from the non-coding chloroplast region *trnL/F*. Chloroplast and mitochondrial gene trees have often been used with great success as estimators of organismal relationships, and organellar genes are convenient phylogenetic markers as they tend to occur as single loci. The uniparental non-Mendelian mode of inheritance removes the confounding effects of recombination and crossing-over apparent in the nuclear genome. By the same token, the uniparental inheritance of organelles can result in misleading genealogies (i.e., gene trees which conflict with the organismal tree) when hybridization and introgression has occurred.

The evolution of morphological characters was optimized onto the molecular phylogeny and the ancestral conditions inferred. The analysis also sought to evaluate our current understanding of the relationships within the group. The monophyly of the Sanguisorbeae as a whole was confirmed, while revised circumscriptions were proposed for some of the traditionally recognized genera. I have refrained from applying ranks to these entities, but have instead given phylogenetic definitions to the same, in accordance with the current Phylocode (see Appendix).

Exploratory morphological analysis

A limited morphological phylogenetic analysis of the Sanguisorbeae was conducted with the intent of optimizing the taxon sampling strategy for the molecular analysis. Twenty-six species of Sanguisorbeae *sensu* Weimarck (1934) were

investigated including all genera traditionally placed in the tribe except *Hagenia* and *Dendriopoterium*. In order to test the monophyly of the Sanguisorbeae, six genera (including *Alchemilla*) were sampled from outside of Sanguisorbeae but within the subfamily Rosoideae. The tree was rooted with *Prunus* (Amygdaloideae). A total of 31 binary or multistate characters were extracted mainly from reproductive structures (i.e., stamens and carpels). The most parsimonious morphological phylogeny (not shown) suggested a paraphyletic origin of Sanguisorbeae with *Alchemilla* nested within the tribe. This is most likely due to convergent morphological similarities associated with a reduction of the floral parts. Molecular phylogenetic analyses of the Rosoideae conducted subsequent to this analysis strongly support a monophyletic Sanguisorbeae excluding *Alchemilla* which is firmly nested in Potentilleae (Eriksson *et al.*, 1998; Eriksson *et al.*, 2003). There was strong morphological support for monophyly of the apetalous genera of Sanguisorbeae. The analysis further indicated that the southern hemispheric genera *Margyricarpus*, *Cliffortia* and *Acaena* and *Polylepis* form a monophyletic clade to the exclusion of the northern taxa.

Materials and methods

Taxon sampling

Based on the morphological analysis and the subsequently published molecular analyses by Helfgott *et al.* (2000) and Eriksson *et al.* (2003) it was clear that a thorough sampling of the genus *Acaena* was necessary. Nine species and subspecies of *Polylepis*, along with fourteen species of *Acaena* and *Margyricarpus,* were collected by the author in Bolivia, Chile and Australia. Samples of an additional

seven species of *Acaena* were obtained from plants in cultivation at the Edinburough Royal Botanic Garden. John Clark (George Washington University) and Michael Kessler (Goettingen University, Germany) generously shared material of several key species of *Polylepis* and *Acaena* and Christopher Whitehouse (University of Cape Town, South Africa) kindly sent material of *Cliffortia* as well as of the enigmatic South African *Acaena latebrosa*. *Sanguisorba minor* was collected in Sweden, Chile, and on the Island of Crete. *Sarcopoterium* was collected on Crete, *Marcetella* and *Bencomia* on the Canary Islands and *Agrimonia parviflora, Potentilla indica* and *Rosa multiflora* were gathered in Maryland. A sample of *Sanguisorba alpina was* obtained from cultivation at the Missouri Botanical Garden. Herbarium specimens were acquired for species considered important but for which fresh or silica-dried material could not be obtained.

Altogether 33 accessions of *Polylepis* and 33 of *Acaena* were sampled in the molecular analysis (Table II.I). Some of these belonged to the same species but represented different geographical locations. Seven species of the South African *Cliffortia* were sequenced, along with four species of *Margyricarpus* (incl. *Tetraglochin*), nine species of *Sanguisorba*, two species of *Agrimonia*, and one species each of *Sarcopoterium, Marcetella, Bencomia, Hagenia, Aremonia, Leucosidea*. Extraction from herbarium material of *Spenceria* was attempted but failed. *Agrimonia eupatoria* and a wide selection of outgroup species had already been sequenced for *trnL/F* by Torsten Eriksson for a wider analysis of the Rosoideae (Eriksson *et al.*, 2003).

Table II.1. Species used in the *trnL/F* analysis listed alphabetically. Letters in parenthesis indicate herbaria where vouchers are deposited. [1]Sequenced by Torsten Eriksson.

Species	Source/voucher information	Geographical origin	GenBank accession
Acaena anserinifolia	P. & M. Hibbs 293 (MARY)	Australia	AY634689
Acaena argentea R. & P.	M. Hibbs 173 (MARY)	Chile	AY634690
Acaena argentea R. & P.	J. Clark 5819 (MARY)	Ecuador	AY634691
Acaena caesiglauca Bergmans	van Balgooy 4329 (MO)	New Zealand	AY634692
Acaena cylindristachya R. & P.	M. Hibbs 167 (MARY, LPB)	Bolivia	AJ512775
Acaena digitata Phil.	M. Hibbs 181 (MARY)	Chile	AY634693
Acaena echinata Nees.	D. E. Symon 13386	Australia	AY634694
Acaena elongata L.	J. Clark 5500 (MARY)	Ecuador	AY634695
Acaena elongata L.	N. Wikström 298 (MARY)	Ecuador	AY634696
Acaena elongata L.	J. Clark 5818 (MARY)	Ecuador	AY634697
Acaena eupatoria Cham. & Schltdl.	R. C. Molon *et al.* 12175 (US)	Brazil	AY634698
Acaena eupatoria Cham. & Schltdl.	R. Wasum *et al.* 10307 (US)	Brazil	AY634699
Acaena fissistipula Bitt.	M. Hibbs 65 (MARY)	New Zeeland (Cult. E)	AY634700
Acaena inermis Hook.f.	M. Hibbs 57 (MARY)	New Zeeland (Cult. E)	AY634701
Acaena latebrosa Aiton.	C. M. Whitehouse 122	South Africa	AY634702
Acaena lucida Vahl.	D. M. Moore 679 (US)	Falkland Islands	AY634703
Acaena lucida Vahl.	G. T. Prance 28561 (NY)	Falkland Islands	AY634704
Acaena macrocephala Poepp.	M. Hibbs 63 (MARY)	Chile (Cult. E)	AY634705
Acaena magellanica (Lam.) Vahl	M. Hibbs 61 (MARY)	Falkland Islands (Cult. E as *A. laevigata*)	AJ512776
Acaena masafuerana Bitt.	G. Kuschel 215 (US)	Juan Fernandez Islands	AY634706
Acaena montana Hook.	M. Hibbs 58 (MARY)	Tasmania (Cult. E)	AY634707
Acaena multifida Hook f.	M. Hibbs 183 (MARY)	Chile	AY634708
Acaena multifida Hook.f.	M. Hibbs 60 (MARY)	Chile, Argentina (Cult. E)	AY634709
Acaena novae-zelandiae Kirk	P. & M. Hibbs 292 (MARY)	Tasmania	AY634710
Acaena novae-zelandiae Kirk	D. E. Symon 15299 (MO)	Australia	AY634711
Acaena ovalifolia R. & P.	M. Hibbs 175 (MARY)	Chile	AY634712

Species	Source/voucher information	Geographical origin	GenBank accession
Acaena pinnatifida R. & P.	M. Hibbs 182 (MARY)	Chile	AY634713
Acaena pinnatifida R. & P. *ssp. grandiflora* Bitter	M. Hibbs 176 (MARY)	Chile	AY634714
Acaena splendens Gilles ex H. & A.	S. Teillier *et al.* 2324 (MO)	Chile	AY634715
Acaena subincisa Wedd.	M. Hibbs 174 (MARY)	Chile	AY634716
Acaena x anserovina Orch.	D. E. Symon 15310 (MO)	Australia	AY634717
Acaena x splendens	L. R. Landrum et al. 8227 (MO)	Chile	AY634718
Agrimonia eupatoria L. [1]	T. Eriksson *et al.* 41 (SBT)	Germany	AJ512216
Agrimonia parviflora Sol.	M. Hibbs 122 (MARY)	North America	
Aremonia agrimonioides[1] Necker ex Nestler	Karlsson 94076 (LD)	unknown	AJ512230/31
Bencomia caudata (Ait.) Webb & Berth	M. Vretblad 46 (MARY)	Canary Islands	AY634719
Cliffortia burmeana Burtt Davy	O. M. Hilliard *et al.* 14681 (E)	South Africa	AY634720
Cliffortia dentata Willd.	C. M. Whitehouse 34	South Africa	AY634721
Cliffortia graminea L.	Woodwine 66 (US)	South Africa	AY634722
Cliffortia heterophylla Weim.	J. Vlok *et al.* 81 (MO)	South Africa	AY634723
Cliffortia odorata L.f.	C. M. Whitehouse 71	South Africa	AY634724
Cliffortia ruscifolia L.	C. M. Whitehouse 72	South Africa	AY634725
Cliffortia sericea E. & Z.	unknown (US)	South Africa	AY634726
Filipendula vulgaris L. [1]	Eriksson 821	Sweden	AJ416463
Hagenia abyssinica J.Gmelin	Knox 2532 (GH)	Kenya	AY634727
Leucosidea sericea E. & Z.	F. B. Wright 1522 (E)	South Africa	AY634728
Marcetella moquiniana (Webb. & Berth.) Svent.	M. Vretblad 49 (MARY)	Canary Islands	AY634729
Margyricarpus pinnatus (Lam.) Kuntze	M. Hibbs 184 (MARY)	Chile	AY634730
Margyricarpus setosus R. & P.	M. Hibbs 136 (MARY, LPB)	Bolivia	AY634731
Polylepis australis Bitter	D. Renison s.n. 33	Argentina	AY634732
Polylepis besseri Hier. *ssp. besseri*	M. Hibbs 158 (MARY, LPB)	Bolivia	AY634733
Polylepis besseri Hier. *ssp. incarum* (Bitter) M.Kessler	M. Hibbs 172 (MARY, LPB)	Bolivia	AY634734
Polylepis besseri Hier. *ssp. subtusalbida* (Bitter) M.Kessler	M. Hibbs 164 (MARY, LPB)	Bolivia	AY634735
Polylepis crista-galli Bitter	M. Kessler 3661 (MO)	Bolivia	AY634736
Polylepis hieronymi Pilger	M. Hibbs 133 (MARY, LPB)	Bolivia	AJ512774
Polylepis incana H.B.K.	J. Clark 6228 (MARY)	Ecuador	AY634737

Species	Source/voucher information	Geographical origin	GenBank accession
Polylepis incana H.B.K.	J. Clark 4991 (MARY)	Ecuador	AY634738
Polylepis lanuginosa H.B.K.	J. Clark 6227 (MARY)	Ecuador	AY634739
Polylepis lanuginosa H.B.K.	G. Harling *et al.* 22858 (MO)	Ecuador	AY634740
Polylepis multijuga Pilger	S. Llatas Quiroz 2749 (MO)	Peru	AY634741
Polylepis neglecta M.Kessler	M. Hibbs 135 (MARY, LPB)	Bolivia	AY634742
Polylepis pauta Hieron.	B. Eriksen 59086 (NY)	Ecuador	AY634743
Polylepis pauta Hieron.	Schmidt-Lebuhn 378	Ecuador	AY634744
Polylepis pepei Simpson	St. G. Beck *et al.* 21532	Bolivia	AY634745
Polylepis pepei Simpson	I. Jimenez 1360	Bolivia	AY634746
Polylepis quadrijuga Bitter	E.G.B. Kieft *et al.* 143 (NY)	Columbia	AY634747
Polylepis quadrijuga Bitter	Cleef 4213 (US)	Colombia	AY634748
Polylepis quadrijuga Bitter	S. R. Gradstein	Colombia	AY634749
Polylepis racemosa R. & P. *ssp. lanata* (O. Kuntze) M. Kessler	M. Hibbs 169 (MARY, LPB)	Bolivia	AY634750
Polylepis racemosa R. & P. *ssp. triacontandra* (Bitter) M. Kessler	J.C. Solomon 15492 (MO)	Bolivia	AY634751
Polylepis reticulata Hier.	J. Clark 4992 (MARY)	Ecuador	AY634752
Polylepis rugulosa Bitter	M. Kumar 34	Chile	AY634753
Polylepis rugulosa Bitter	M. Kumar 35	Chile	AY634754
Polylepis sericea Wedd.	J. Clark 5820 (MARY)	Ecuador	AY634755
Polylepis sericea Wedd.	K. Romoleroux *et al.* 1495B (NY)	Ecuador	AY634756
Polylepis sericea Wedd.	P. M. Jørgensen *et al.* 1240 (MO)	Ecuador	AY634757
Polylepis subsericans Macbr.	A. Tupayachi *et al.* 1192 (MO)	Peru	AY634758
Polylepis tarapacana Philippi	M. Hibbs 163 (MARY, LPB)	Bolivia	AJ512773
Polylepis tomentella Wedd. *ssp. tomentella*	M. Hibbs 162 (MARY, LPB)	Bolivia	AY634759
Polylepis tomentella Weddell *ssp. incanoides* M. Kessler	M. Hibbs 137 (MARY, LPB)	Bolivia	AY634760
Polylepis weberbaueri Pilger	J. Clark 6229 (MARY)	Ecuador	AY634761
Polylepis weberbaueri Pilger	S. Laegaard 53795 (MO)	Ecuador	AY634762
Potentilla indica Wolf.	M. Hibbs 69 (MARY)	USA (introduced)	AY634763
Rosa multiflora Thunb. ex Murray	M. Hibbs 78 (MARY)	USA	AY634764

Species	Source/voucher information	Geographical origin	GenBank accession
Sanguisorba alpina Bunge in Ledeb.	M. Hibbs 294 (MARY)	Mongolia (Cult. MO)	AY634765
Sanguisorba annua (Nutt. ex Hook.) Torr. & Gray	E. Earle 4412 (US)	USA	AY634766
Sanguisorba annua (Nutt. ex Hook.) Torr. & Gray	B. Maguire *et al.* (PH)	USA	AY634767
Sanguisorba canadensis L.	S. Woodbury 152 (MO)	North America	AY634768
Sanguisorba filiformis (Hook. F.) Hand.-Mazz.	J. F. Rock 17963	China: SW Szechwan	AY634769
Sanguisorba hakusanensis Mak.	M. Hibbs 59 (MARY)	Japan or Korea (Cult. E)	AY634770
Sanguisorba minor Scop.	M. Vretblad 17 (MARY)	Crete	AY634771
Sanguisorba minor Scop.	M. Hibbs 89 (MARY)	Sweden	AY634772
Sanguisorba minor Scop.	M. Hibbs 178 (MARY)	Chile	AY634773
Sanguisorba officinalis L.	T. Eriksson 144 (SBT)	Sweden	AY634774
Sanguisorba stipulata Rafin.	K. Dillman 17	USA: Alaska	AY634775
Sanguisorba tenuifolia Fisch. ex Link	X. Ling 81906 (MO)	China	AY634776
Sarcopoterium spinosum (L.) Spach	M. Vretblad 16 (MARY)	Crete	AY634777
Tetraglochin cristatum (Britt.) Rothm.	M. Hibbs 150 (MARY, LPB)	Bolivia	AJ512777

Phylogenetic marker

The ostensibly recent origin of Sanguisorbeae presented a challenge to finding an organellar marker variable at sub- and intergeneric levels. Previous studies have indicated that *rbcL* is not variable at this level (Morgan *et al.*, 1994) hence amplification of this gene was not attempted in this study. Initial trial amplification and sequencing of the chloroplast genes *matK, ndhF* and *rps16* and the mitochondrial markers *nad1* and *cox-III* were aborted because of lack of sequence variation.

The non-coding chloroplast region encompassing the *trnL/F* intron and the intergenic spacer between *trnL* and *trnF* (*trnL/F* hereafter) has been informative at the species to generic level in other phylogenetic analyses of angiosperms including some studies within the Rosaceae (Bortiri *et al.*, 2001; Eriksson *et al.*, 2003; Potter *et al.*, 2000). Initial amplification and sequencing of this region suggested significant intergeneric variation in Sanguisorbeae and *trnL/F* was therefore chosen for further investigation.

DNA extraction

Plant samples collected in the field by the author were stored in silica gel, as were samples provided by collaborators. Key species for which field samples could not be obtained were extracted from herbarium specimens with the consent of the herbaria from which the specimens were borrowed (E, MO, US, NY, MARY and PH). Fresh leaf material was used in the extraction of *Agrimonia parviflora, Potentilla indica* and *Rosa multiflora*. Total DNA was isolated using the Quiagen Plant DNEasy Mini Kit. 0.005-0.025 g of dried leaf tissue was ground in liquid

nitrogen. The extraction proceeded according to the kit protocol with one exception: prior to applying the lysate to the spin column, mucilage and persistent tissue fragments were removed by centrifugation at maximum speed for 1 minute. This step was necessary because of the tough hirsute leaves of *Polylepis,* which tend to resist tissue break-up and cell lysis. A small number of extractions were performed using either a CTAB extraction protocol (Doyle and Doyle, 1987) or the resin based Nucleon Phytopure extraction kit (Amersham Pharmacia, Uppsala, Sweden). In most cases the quality of the extracted DNA was assessed using agarose gel electrophoresis and ethidium bromide staining. Because the quantity and quality of DNA extracted from herbarium specimens tended to be very low, extractions were not always visible on a gel. In these cases PCR was nevertheless attempted and in most cases the *trnL/F* amplicons were generated despite the low DNA concentration.

PCR Amplification and sequencing

The chloroplast region *trnL/F* was amplified by polymerase chain reaction (PCR) using Taberlet's (1991) primers C (forward) and F (reverse). The melting temperature of the primers was estimated by the formula $T_m(C)\ 2(N_A+N_T) + 4(N_G+N_C)$ and the annealing temperature was set to correspond to the lowest primer melting temperature. Weak DNA extractions, for which long amplicons could not be generated, were amplified in two segments using the primer combinations C+D and E+F. PCR amplification was conducted in a total reaction volume of 50µl with final concentrations of 3mM $MgCl_2$, 0.2µM each primer, 0.25µM each dNTP, 10-50ng genomic DNA, and 2.5 units Promega *Taq* polymerase.

PCR products were purified using polyethylene glycol (PEG) precipitation. 20% PEG with 2.5M NaCl was added to the PRC product in the same amount as the product and the mixture vortexed. The precipitate was centrifuged at 16,000g for 15 minutes and the supernatant removed. The pellet was washed with 80% ethanol, followed by a second centrifugation at 16,000g for 10 minutes. After removing the ethanol, pellets were air- or vacuum-dried and eluted in 10-30µl de-ionized water depending on the concentration of the PCR product.

Twenty-eight of the *trnL/F* amplicons were submitted to the University of Maryland Center for Agricultural Biotechnology for sequencing (on a ABI 377 slab gel sequencer). The remaining products were sequenced by the author on a ABI Prism 3100 sequencer. These sequencing reactions were performed in a final volume of 7 µl consisting of 1 µl PEG-purified PCR product, 3 µl de-ionized water, 1 µl 3.2 µM primer, 1.5 µl 5X buffer (400 mM Tris pH 9.0, 10 mM $MgCl_2$), and 0.5 µl BigDye Terminator Ready Reaction Mix v2 (Perkin Elmer Biosystems, Foster City, CA). Taberlet's (1991) primers C, D, E and F were used as sequencing primers.

Chromatograms were manually proof read and sequence contigs were assembled using the computer program Sequencher 3.1.1. Sequence identities were verified (i.e., as belonging to Rosaceae) by BlastN searches against published sequences (NCBI). Attempts to use alignment algorithms (e.g., Clustal, PileUp) were ineffective and resulted in clearly unsatisfactory alignments. Sequences were instead aligned manually by eye using MacClade (Maddison and Maddison, 2002). A binary (0,1) matrix of 32 indel characters was added to the end of the nucleotide matrix.

Only parsimony-informative and unambiguous (i.e., perfectly overlapping) indels were included.

Phylogenetic methods

Phylogenetic analyses were conducted under the parsimony and maximum likelihood optimality criteria. A Bayesian analysis was also conducted in order to allow for the use of mixed models (for nucleotide and indel characters) within a likelihood framework. Analyses were performed using PAUP* 4.0b10 (Swofford, 2002) and MrBayes (Huelsenbeck and Ronquist, 2001). Consistency of clade support was estimated by nonparametric bootstrap analyses in parsimony and likelihood analyses, and by posterior probabilities in the Bayesian analyses.

An initial parsimony analysis was conducted to test the monophyly of Sanguisorbeae. This analysis included three members of Potentilleae and was rooted with *Filipendula*, which has been shown to be basal in the Rosoideae (Eriksson *et al.*, 1998; Morgan *et al.*, 1994). Subsequent analyses limited the outgroup selection to *Rosa*. Two maximum parsimony analyses were conducted based on nucleotide data alone and in combination with the indel matrix. All character transformations were equally weighted. The analyses were conducted using a heuristic search strategy of tree bisection-reconnection branch swapping on 1000 random addition starting trees. Steepest decent was in effect and branches with a minimum length of 0 were collapsed. Strict consensus trees were derived from the topologies of all the most parsimonious trees. Clade support was estimated for both data sets by nonparametric bootstrap analysis based on 1000 pseudo-replicates.

The maximum likelihood analysis was conducted on the nucleotide data set alone since there is currently no software available for combining models in maximum likelihood analyses. The model of sequence evolution used in a likelihood analysis can greatly affect the resulting phylogeny and branch lengths. It is therefore desirable to use a model that is as realistic as possible. However, adding parameters to the model increases the risk of random error, thus the most parameter-rich model is not necessarily the one that best fits the data. The likelihood ratio test (Whelan and Goldman, 1999) is a widely accepted statistic for testing the goodness of fit of different models. The optimal model of sequence evolution for the *trnL/F* data set was chosen using the likelihood ratio test as implemented in the computer program ModelTest (Posada and Crandall, 1998).

Results

The alignment of *trnL/F* nucleotide sequences was straightforward with the exception of two regions of single nucleotide repeats of varying length (poly-A (char. 461-468) and poly-T (char. 1079-1094)). Sequence length ranged from 901 bp (*Aremonia*) to 1055 bp (*Sanguisorba hakusanenesis*). Aligned sequence length was 1248 bp when *Filipendula* was included, 1221 when excluded. 923 of the nucleotide characters were constant, 174 were autapomorphic and 151 were parsimony-informative. Mean GC content was 32.8%. All indel characters were parsimony-informative by design. Although the *trnL/F* region is rather conserved within genera of Sanguisorbeae, there is a phylogenetically constructive level of variation among genera. Pairwise sequence divergences within Sanguisorbeae ranged from zero (some

sequences between closely related species were identical) to 0.07270 (between *Leucosidea* and *Acaena multifida*).

The initial parsimony analysis, which included a broader outgroup sampling, resulted in 816 most parsimonious trees (length 516) with a rescaled consistency index of 0.739. The Bayesian mixed model analysis yielded a nearly identical 50% consensus tree, depicted in Figure II.1. The result of the initial broader analysis strongly supports the monophyly of the Sanguisorbeae as a whole, which is consistent with previous studies (Eriksson *et al.*, 2003). Subsequent analyses focused on Sanguisorbeae exclusively and were rooted with *Rosa multiflora* (Roseae) to avoid the long branch to *Filipendula*.

The parsimony analysis of combined data (nucleotide+indel) based on the reduced data set yielded 510 most parsimonious trees of length 337, with a rescaled consistency index of 0.775 (Figure II.2A). The likelihood ratio test suggested that the GTR+I+Γ best fit the *trnL/F* data. This model was thus applied in the maximum likelihood analysis as well as for the nucleotide characters in the Bayesian analysis.

The results of the reduced data set analyses were very similar regardless of the optimality criterion applied. The trees from the combined parsimony analysis and the mixed-model Bayesian analysis were more resolved than the maximum likelihood tree because of the added information from the indel characters. The topologies from the parsimony analyses of the nucleotide data alone and the combined data sets were congruent and only differed in the level of resolution. This tree is compared to the Bayesian mixed model phylogeny in Figure II.2B. The likelihood tree is presented in Figure II.3.

Figure II.1. *trnL/F* phylogeny of Sanguisorbeae and outgroups based on Bayesian inference using a combined model of GTR+I+Γ for nuleotide characters and Mk for indels. Posterior probabilities (as percentages) are indicated above branches. Clades with a posterior probability of less than 0.5 are collapsed. Branches marked (•) were not found in the strict consensus tree from the parsimony analysis.

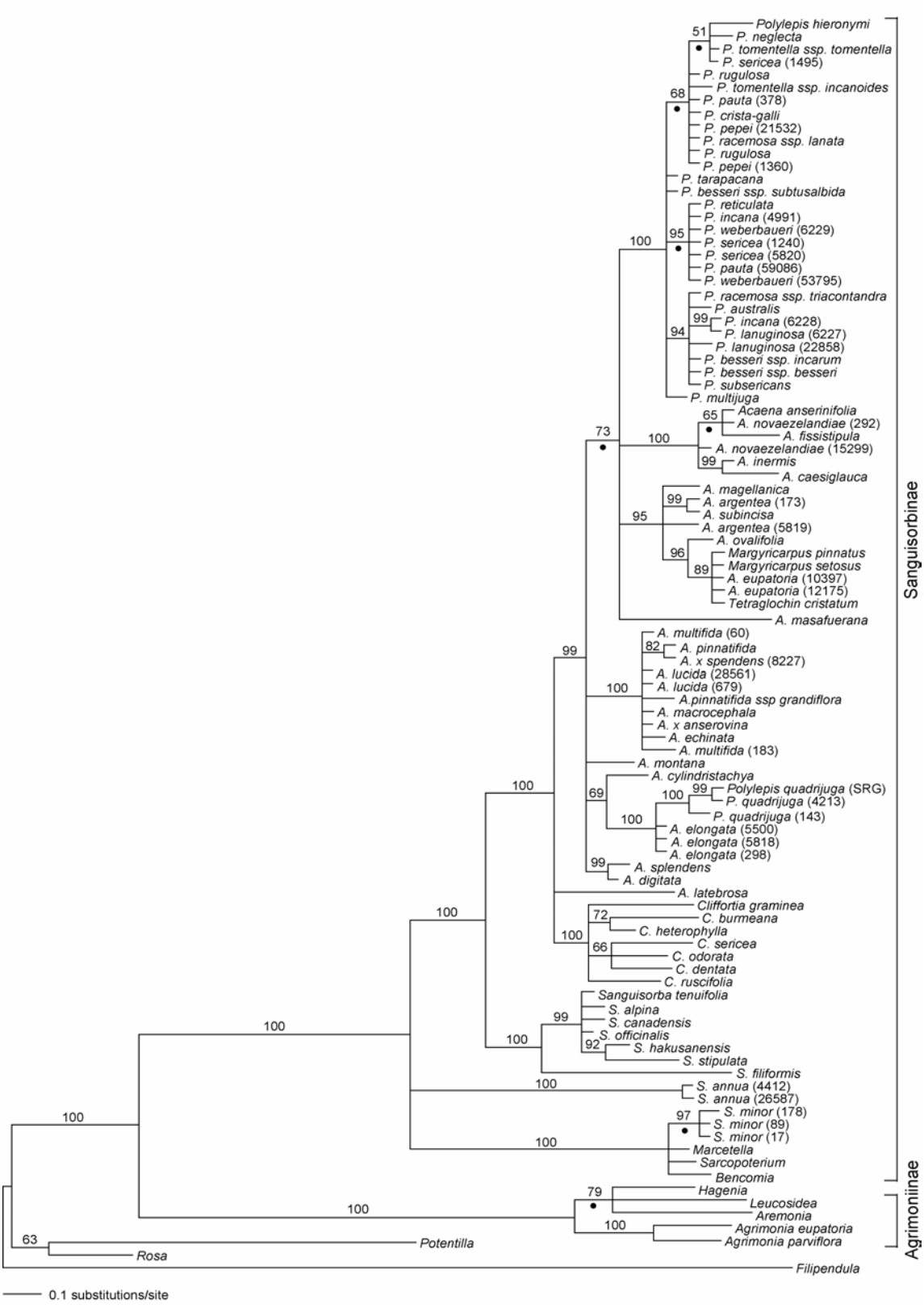
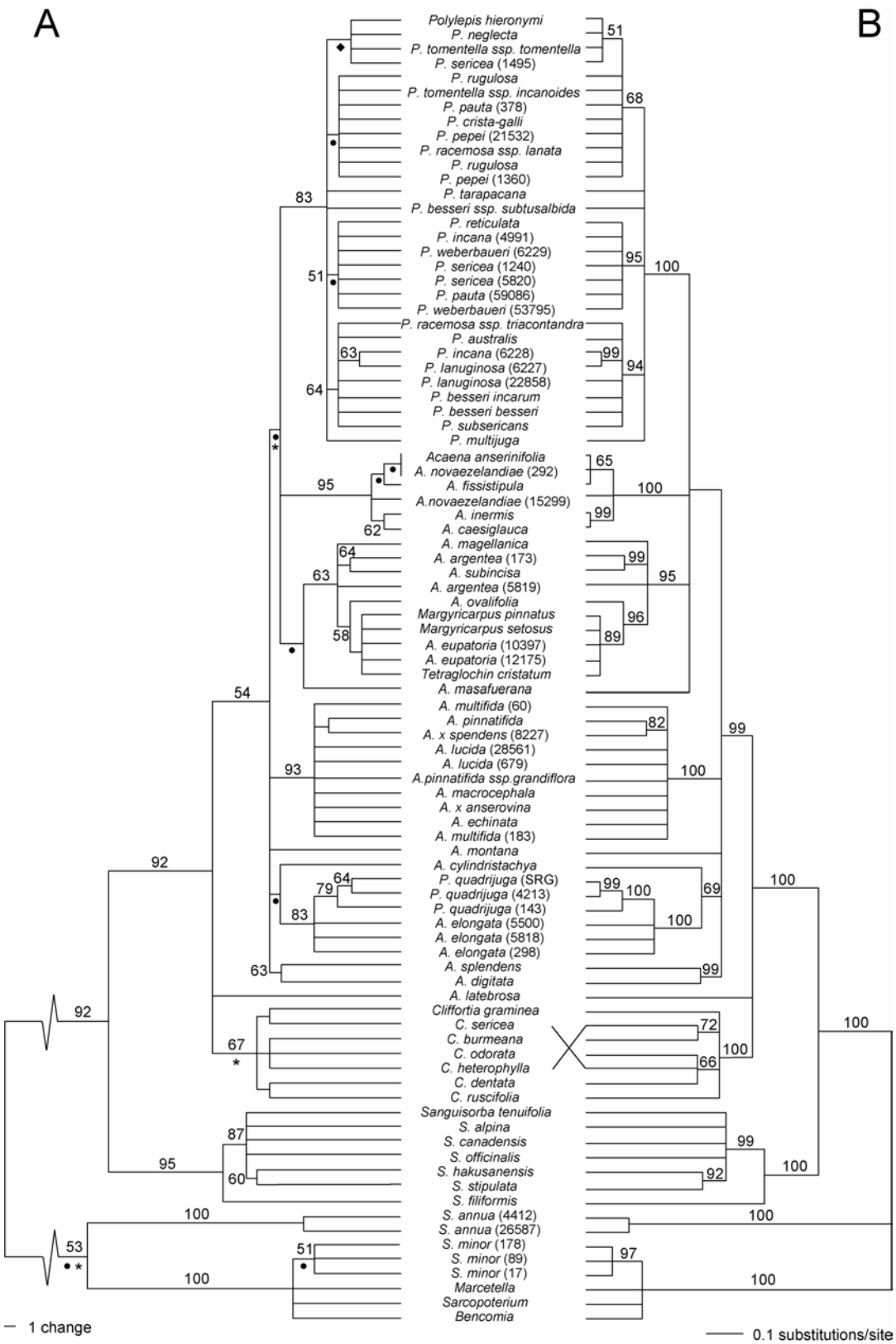

Polylepis hieronymi
P. neglecta
P. tomentella ssp. tomentella
P. sericea (1495)
P. rugulosa
P. tomentella ssp. incanoides
P. pauta (378)
P. crista-galli
P. pepei (21532)
P. racemosa ssp. lanata
P. rugulosa
P. pepei (1360)
P. tarapacana
P. besseri ssp. subtusalbida
P. reticulata
P. incana (4991)
P. weberbaueri (6229)
P. sericea (1240)
P. sericea (5820)
P. pauta (59086)
P. weberbaueri (53795)
P. racemosa ssp. triacontandra
P. australis
P. incana (6228)
P. lanuginosa (6227)
P. lanuginosa (22858)
P. besseri ssp. incarum
P. besseri ssp. besseri
P. subsericans
P. multijuga
Acaena anserinifolia
A. novaezelandiae (292)
A. fissistipula
A. novaezelandiae (15299)
A. inermis
A. caesiglauca
A. magellanica
A. argentea (173)
A. subincisa
A. argentea (5819)
A. ovalifolia
Margyricarpus pinnatus
Margyricarpus setosus
A. eupatoria (10397)
A. eupatoria (12175)
Tetraglochin cristatum
A. masafuerana
A. multifida (60)
A. pinnatifida
A. x spendens (8227)
A. lucida (28561)
A. lucida (679)
A.pinnatifida ssp grandiflora
A. macrocephala
A. x anserovina
A. echinata
A. multifida (183)
A. montana
A. cylindristachya
Polylepis quadrijuga (SRG)
P. quadrijuga (4213)
P. quadrijuga (143)
A. elongata (5500)
A. elongata (5818)
A. elongata (298)
A. splendens
A. digitata
A. latebrosa
Cliffortia graminea
C. burmeana
C. heterophylla
C. sericea
C. odorata
C. dentata
C. ruscifolia
Sanguisorba tenuifolia
S. alpina
S. canadensis
S. officinalis
S. hakusanensis
S. stipulata
S. filiformis
S. annua (4412)
S. annua (26587)
S. minor (178)
S. minor (89)
S. minor (17)
Marcetella
Sarcopoterium
Bencomia
Hagenia
Leucosidea
Aremonia
Agrimonia eupatoria
Agrimonia parviflora
Potentilla
Rosa
Filipendula
Sanguisorbinae
Agrimoniinae
0.1 substitutions/site

Figure II.2 *trnL/F* phylogeny of Sanguisorbeae based on combined nucleotide+indel characters based on (A) parsimony and (B) Bayesian inference. Internal branch lengths are drawn proportional to change. Terminal branches are not drawn proportional to allow for easier comparison between the two phylogenies. Bootstrap values and posterior probabilities are indicated above branches. The parsimony tree represents one of 510 most parsimonious hypotheses. Clades that were not found in the strict consensus tree are marked (•). Clades supported only by indel characters are marked (*).


A
B
Polylepis hieronymi
P. neglecta
P. tomentella ssp. tomentella
P. sericea (1495)
P. rugulosa
P. tomentella ssp. incanoides
P. pauta (378)
P. crista-galli
P. pepei (21532)
P. racemosa ssp. lanata
P. rugulosa
P. pepei (1360)
P. tarapacana
P. besseri ssp. subtusalbida
P. reticulata
P. incana (4991)
P. weberbaueri (6229)
P. sericea (1240)
P. sericea (5820)
P. pauta (59086)
P. weberbaueri (53795)
P. racemosa ssp. triacontandra
P. australis
P. incana (6228)
P. lanuginosa (6227)
P. lanuginosa (22858)
P. besseri incarum
P. besseri besseri
P. subsericans
P. multijuga
Acaena anserinifolia
A. novaezelandiae (292)
A. fissistipula
A.novaezelandiae (15299)
A. inermis
A. caesiglauca
A. magellanica
A. argentea (173)
A. subincisa
A. argentea (5819)
A. ovalifolia
Margyricarpus pinnatus
Margyricarpus setosus
A. eupatoria (10397)
A. eupatoria (12175)
Tetraglochin cristatum
A. masafuerana
A. multifida (60)
A. pinnatifida
A. x spendens (8227)
A. lucida (28561)
A. lucida (679)
A.pinnatifida ssp.grandiflora
A. macrocephala
A. x anserovina
A. echinata
A. multifida (183)
A. montana
A. cylindristachya
P. quadrijuga (SRG)
P. quadrijuga (4213)
P. quadrijuga (143)
A. elongata (5500)
A. elongata (5818)
A. elongata (298)
A. splendens
A. digitata
A. latebrosa
Cliffortia graminea
C. sericea
C. burmeana
C. odorata
C. heterophylla
C. dentata
C. ruscifolia
Sanguisorba tenuifolia
S. alpina
S. canadensis
S. officinalis
S. hakusanensis
S. stipulata
S. filiformis
S. annua (4412)
S. annua (26587)
S. minor (178)
S. minor (89)
S. minor (17)
Marcetella
Sarcopoterium
Bencomia
1 change
0.1 substitutions/site

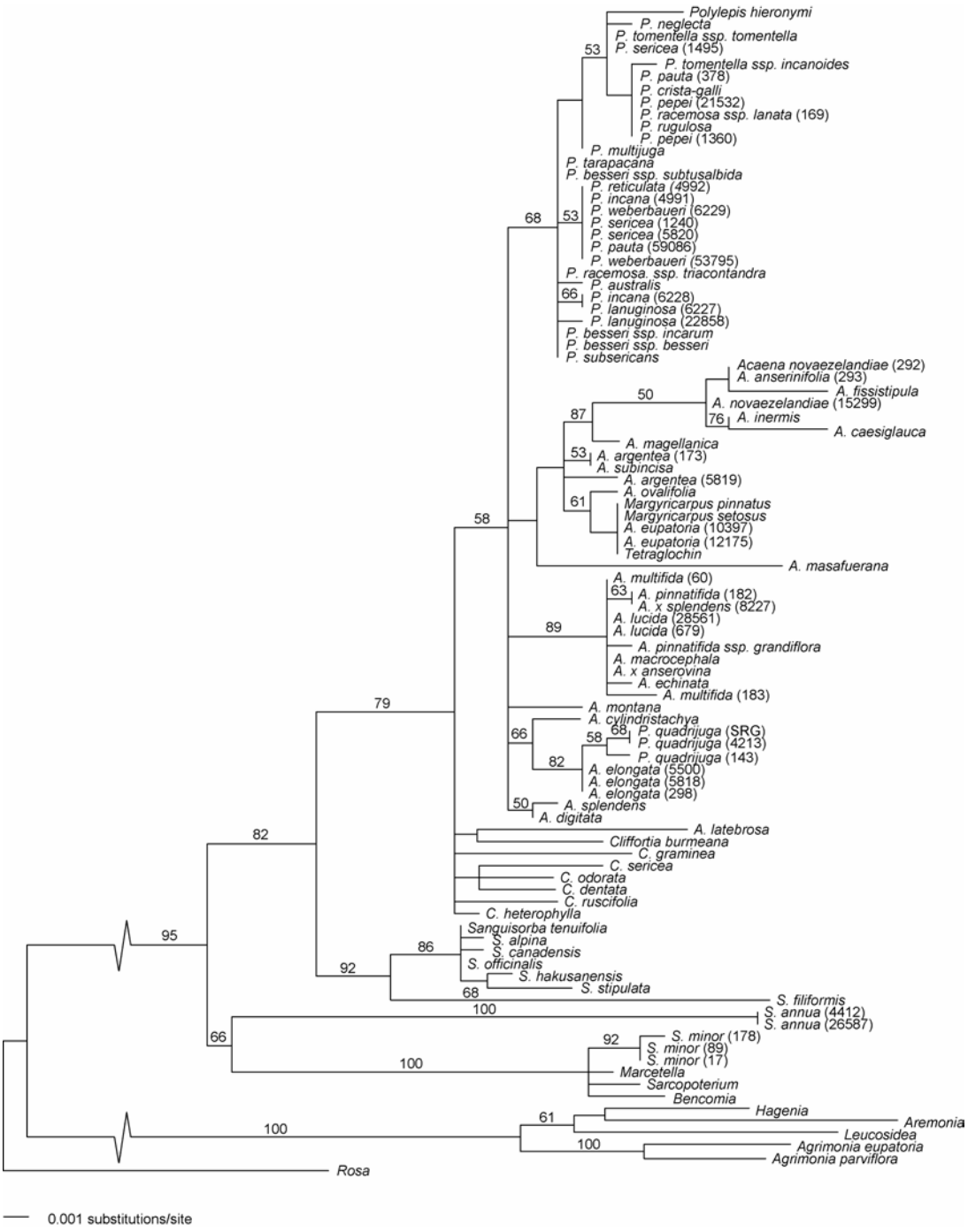


Figure II.3 *trnL/F* phylogeny of Sanguisorbeae based on maximum likelihood under a GTR+I+Γ model of sequence evolution. Bootstrap values are indicated above branches.

The *trnL/F* phylogeny (Figures II.1-3) reveals an early divergence in the ancestral Sanguisorbeae lineage, giving rise to two distinct clades corresponding to the petalous (=Agrimoniinae) and the apetalous Sanguisorbeae (=Sanguisorbinae). There is strong support for the monophyly of Agrimoniinae as well as of the Sanguisorbinae. *Sanguisorba sensu lato* is paraphyletic and split into three groups. A clade corresponding to *Sanguisorba sensu stricto* (Linnaeus), and which includes the type species *S. officinalis*, is monophyletic and sister group to a clade of *Cliffortia,* Acaena*,* and *Polylepis. Sanguisorba minor* (=*Poterium* L.) forms a well-supported unit together with the woody genera *Sarcopoterium, Marcetella*, and *Bencomia*. The placement of the third group, *Sanguisorba annua* (=*Poteridium* Spach) is unresolved. In the combined parsimony analysis *S. annua* is sister to the *S. minor* clade with low bootstrap support (53%). In the Bayesian phylogeny this relationship is unresolved in the 50% consensus (halfcompat) tree, but a position of *S. annua* basal to all of Sanguisorbinae has a posterior probability of 47% (not shown). *Cliffortia* is monophyletic with 67% bootstrap support in the combined parsimony analysis and with 100% posterior probability in the Bayesian analysis. This branch is only supported by indel characters and is therefore collapsed in the nucleotide parsimony phylogeny as well as in the likelihood tree. The South African *Acaena latebrosa* forms a polytomy with *Cliffortia* and the remaining *Acaena*.

Phylogenetic nomenclature and the application of taxon names

The current Linnaean nomenclatural system has been instrumental in communicating taxonomic information for over two hundred years. However, it has

been argued that this pre-evolutionary system is ineffective in conveying phylogenetic relationships (e.g., De Queiroz and Cantino, 2001; De Queiroz and Gauthier, 1990; De Queiroz and Gauthier, 1994; Ereshefsky, 1994). Although the phylogeny of life is inherently hierarchical, the assignment of ranks is evolutionarily meaningless and can be highly subjective. In addition, the names of clades as well as species are not stable but can change because of rank shifts or new generic assignments. Efforts are under way to develop a phylogenetic nomenclatural code that would eliminate or mitigate the flaws of the current system (Cantino and de Queiroz, 2000). Although this new code will not officially take effect until July 2004, systematists have been encouraged to publish phylogenetic names during a trial period. In this dissertation a phylogenetic classification is adopted for Sanguisorbeae, and new rank-free clade names are used in parallel with the traditional Linnaean taxonomy.

To facilitate discussion I follow the phylogenetic taxon names defined in Eriksson *et al.* (2003). New names were also proposed for previously nameless but strongly supported clades (see Appendix). I used "old" taxon names in Figures II.1-3 to reflect the predominant view of generic circumscriptions prior to this analysis, but added recircumscribed and new names in Figure II.4. For example, the name *Sanguisorba minor* has been changed to *Poterium sanguisorba* and *S. annua* to *Poteridium annuum* to reflect new evidence from the molecular analysis. Informal rank-free names were given to some infrageneric groups in *Acaena*, mostly based on names applied by previous authors (Bitter, 1911a). Because the sampling of *Acaena* species in the current analysis was incomplete, a formal revision of nomenclature in

this paraphyletic group is premature. The traditional classification of Sanguisorbeae (a composite from several sources since no complete classification of this group has been published) is contrasted with a preliminary rank-free phylogenetic classification based on new evidence (Table II.2).

Discussion

The *trnL/F* analysis was not specifically aimed at determining the position of Sanguisorbeae within subfamily Rosoideae, but some additional insights might nevertheless be gleaned from the study (for further discussion see chapter III). Phylogenetic analyses of Rosaceae based on *rbcL*, ITS (Eriksson *et al.*, 1998; Eriksson *et al.*, 2003), *matK* (Mishima *et al.*, 2002) and ITS+*trnL/F* (Eriksson *et al.*, 2003) disagree on the exact position of Sanguisorbeae within Rosoideae. According to Morgan *et al.*, *rbcL* data suggest that Sanguisorbeae is sister to a clade of Roseae+Potentilleae, although with low support. In contrast, a combined *trnL/F*+ITS phylogeny supported a sister relationship between Sanguisorbeae and Potentilleae at the exclusion of Roseae (Eriksson *et al.*, 2003). *matK* data favors a third alternative placing Sanguisorbeae with Roseae at the exclusion of Potentilleae (Mishima *et al.*, 2002). The current analysis of *trnL/F* with limited sampling outside of Sanguisorbeae supports the *rbcL* findings, i.e., Roseae and Potentilleae are more closely related to each other than either is to Sanguisorbeae. Additional molecular and/or morphological data are clearly needed to resolve this polytomy (see chapter III).

Table II.2 Traditional classification of Sanguisorbeae (composite from several sources) contrasted with a preliminary phylogenetic classification based on new molecular evidence. The para- or polyphyletic species complex "*Acaena*" requires a complete revision.

Traditional classification

Tribe Sanguisorbeae
- Subtribe Sanguisorbinae
  - Genus *Sanguisorba*
  - *Poteridium* (or incl. in *Sanguisorba*)
  - *Poterium* (or incl. in *Sanguisorba*)
  - *Sarcopoterium* (or incl. in *Sanguisorba*)
  - *Bencomia*
  - *Marcetella*
  - *Dendriopoterium*
  - *Cliffortia*
  - *Acaena*
  - *Margyricarpus*
  - *Tetraglochin*
  - *Polylepis*
- Agrimoniinae
  - *Agrimonia*
  - *Aremonia*
  - *Hagenia*
  - *Spenceria*

Phylogenetic classification (rank-free)

Sanguisorbeae
- *Sanguisorbinae*
  - *Sanguisorba*
  - *Poteridium*
  - *Poterium*
  - *Sarcopoterium*
  - *Bencomia*
  - *Marcetella*
  - *Dendriopoterium*
  - *Verruchaena*
    - *Cliffortia*
    - *Amentomorpha*
      - "*Acaena*"
      - *Margyricarpus*
      - *Tetraglochin*
      - *Polylepis*
- *Agrimoniinae*
  - *Agrimonia*
  - *Aremonia*
  - *Hagenia*
  - *Spenceria*

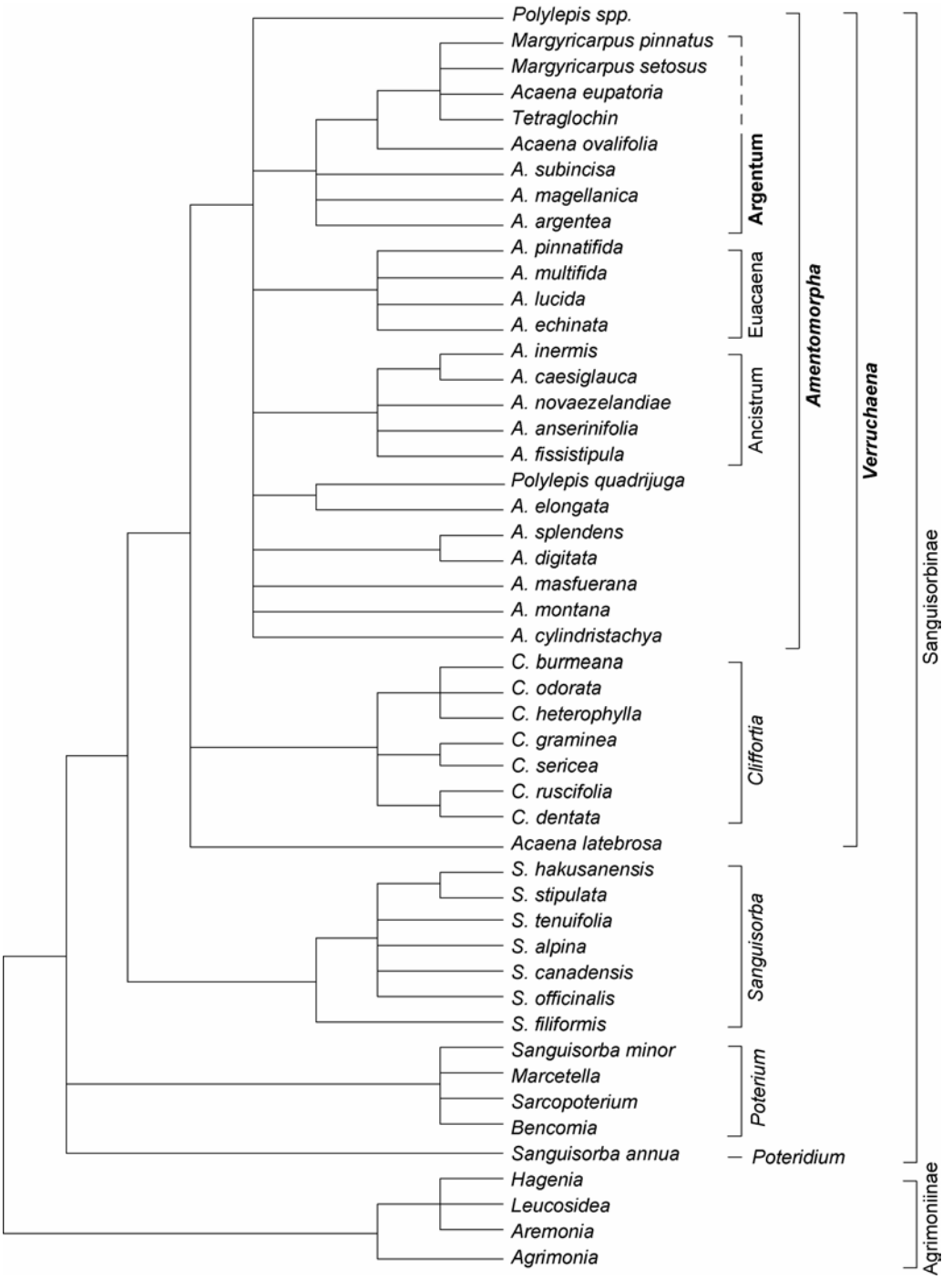


Figure II.4. Old and new taxonomic names supported by phylogenetic data. New names are in bold. For phylogenetic definitions see Appendix. Infrageneric names in *"Acaena"/Amentomorpha* are informal pending a forthcoming complete revision of this group.

The monophyly of Sanguisorbeae as a whole is consistent with the results of Eriksson *et al.* (2003), as is the monophyly of the two main subgroups Sanguisorbinae (apetalous) and Agrimoniinae (petalous). Phylogenetic definitions of these three clades were provided in Eriksson *et al.* (2003). The phylogenetic position of the Chinese genus *Spenceria* could greatly add to the understanding of the evolution of the Sanguisorbeae. *Spenceria* is traditionally considered part of the Agrimoniinae but unlike *Agrimonia* and *Aremonia, Spenceria* is diploid rather than tetraploid (chromosome counts of *Hagenia* and *Leucosidea* are unavailable).

Sanguisorba and Poterium

The monophyly of the apetalous taxa (Sanguisorbinae) is not surprising and is supported by a multitude of morphological characters associated with the reduction of the flower. The *trnL/F* phylogeny further suggests that Linnaeus was correct in initially establishing two separate genera: *Sanguisorba* for *S. officinalis* and *S. canadensis* and *Poterium* for what we now call *Sanguisorba minor, Sanguisorba hybridum* and *Sarcopoterium spinosum* (additional species were later added to both genera). Linnaues based this in part on the observation that while *Sanguisorba* has a single carpel and one style, *Poterium* has two carpels and two styles. In the following discussion *Sanguisorba minor* will be identified as *Poterium sanguisorba* to reflect the new evidence and to clarify the distinction between the two clades. The close relationship between *Poterium* and the Canarian endemics *Bencomia* and *Marcetella* is in agreement with a previous study based on ITS (Helfgott *et al.*, 2000). The *trnL/F* data do not resolve the exact position of *Sanguisorba annua* (synonyms *Poterium*

annuum, Poteridium annuum) that has traditionally been considered closely allied with *Poterium sanguisorba*. This relationship was also suggested in a *matK* analysis of Mishima *et al.* (2002). Spach (1846), however, pointed out that *Poterium annuum*, as it was called then, displayed a combination of characters from both *Poterium* and *Sanguisorba*, as well as some unique characteristics of its own (e.g., pectinately pinnatifid leaves). On this basis Spach described a new monotypic genus *Poteridium*. In addition to the characters cited by Spach, *S. annua* is the only known diploid (2n=14) *Sanguisorba*, all others (including *Poterium*) being tetraploid (2n=28) or higher. It appears that the distinct morphology and cytology as well as the divergent *trnL/F* sequence justify the separation of *S. annua* into *Poteridium*. The possibility that an ancestral *Poteridium* was involved in ancient allopolyploid speciation is discussed in chapter III.

The high alpine Himalayan species *Sanguisorba filiformis* was placed by Nordborg (1967) in *Poterium*, which she defined as a section under *Sanguisorba*. The characters cited to support this placement were *S. filiformis'* tricolporate pollen and its overall habit, which resembles *Poterium sanguisorba*. Nordborg acknowledged that the species displays a mix of characters of both *Sanguisorba* and *Poterium*. It has a single carpel like *Sanguisorba* and typically six stamens, which is more than *Sanguisorba* but less than most *Poterium*. Unlike *Sanguisorba* and *Poterium* it has strictly unisexual flowers (like *Marcetella, Bencomia* and *Sarcopoterium*). The *trnL/F* data suggest that *S. filiformis* is the sister group of *Sanguisorba sensu stricto.*

The 6-colporate pollen type of the *Sanguisorba* clade (minus *S. filiformis*) is unique within the Rosaceae. Although mere speculation at this point, it is possible

that the doubling of colpi was associated with the doubling of the genome that occurred in the ancestor of *Sanguisorba*. Several studies have suggested that pollen size and/or morphology in some angiosperm taxa vary in accordance with the ploidy level (Fukuhara, 2000; Nordborg, 1967; Pinar *et al.*, 2001; Talent and Dickinson, 2002). Nordborg (1967) reported that the pollen apertures of experimentally produced hybrids of *Sanguisorba* were often missing, irregular or anomalous in shape and number. According to this scenario, the tricolporate sister species *S. filiformis* should be diploid (like *S. annua*), which is contradicted by a recent chromosome count of 2n=42 (Mishima *et al.*, 2002). However, *S. filiformis* could have undergone genome doubling (possibly autopolyploidy) independently of the 6-colporate *Sanguisorba*. Another argument against this hypothesis is that genome doubling has occurred in many other lineages including *Poterium* without a change in the structure of the pollen grain.

A noteworthy finding is the strong support for a monophyletic clade composed of of the southern hemispheric genera *Cliffortia, Acaena* (itself para- or polyphyletic) and *Polylepis*. A detailed discussion of this clade will follow below, but, for simplicity of discussion, I will introduce here the (unranked) name *Verruchaena* to denote this clade. The *trnL/F* data strongly support a sister relationship between the northern hemispheric clade *Sanguisorba* and the southern hemispheric *Verruchaena*. This is somewhat surprising, given that, at least on a superficial level, the morphology of the southern genera is more reminiscent of *Poterium* than of *Sanguisorba*. A closer inspection of morphological characters does not shed any more light on this relationship. While *Poterium* is characterized by two

carpels and *Sanguisorba* by one, this character is polymorphic in *Cliffortia* as well as in *Acaena* (*Polylepis* has one carpel). Likewise, the evolution of monoecy and dioecy is not easily harmonized with the molecular data.

In his monograph of *Cliffortia*, Weimarck (1934) argued that the origin of the Sanguisorbinae must have occurred in the southern hemisphere, either in Africa or South America. He considered *Acaena* to represent a basal lineage from which the rest of Sanguisorbinae had sprung. According to this scenario, *Poterium* (in which he included *Sanguisorba*) had originated within *Acaena* "subgenus" *Terminales* (Bitter, 1911a) (*Acaena* with terminal inflorescences), while *Polylepis, Margyricarpus, Bencomia* and *Cliffortia* had evolved from *Acaena* "subgenus" *Axillares* (*Acaena* with axial inflorescences). These relationships are clearly not supported by the molecular data. Instead it appears most parsimonious to hypothesize that Sanguisorbinae originated in the northern hemisphere and subsequently spread southward.

Verruchaena

Although Weimarck was mistaken on the origin of *Poterium* and *Bencomia*, he was correct in recognizing a close relationship between *Cliffortia, Acaena, Margyricarpus* and *Polylepis*. Although these genera are all found in the southern hemisphere, the distribution is quite fragmented and disjunct. *Acaena* has the broadest distribution with species in South America, Australasia, South Africa, and Hawaii, as well as a scattered presence on several subantarctic islands. *Polylepis* and *Margyricarpus* are only found in South America while *Cliffortia* is restricted to South

Africa. The *trnL/F* phylogeny suggest that the South African species of *Acaena (A. latebrosa*) is quite distinct and may be more closely related to *Cliffortia* than to the rest of *Acaena*. *Polylepis, Margyricarpus* and *Tetraglochin* are nested within *Acaena* and together form a monophyletic clade separate from *Cliffortia* and *Acaena latebrosa*. Based on these data it is impossible to say whether *Verruchaena* originated in South America or South Africa (Australasia is probably less likely). A thorough discussion of different biogeographic scenarios in light of molecular data from nuclear as well as chloroplast data will follow in chapter IV.

Acaena and Margyricarpus/Tetraglochin

What Bitter (1911a) referred to as *Acaena* section *Ancistrum* is split geographically into an Australasian clade, which will henceforth informally be referred to as *Ancistrum* (no rank), and a South American clade that will be called *Argentum*. *Ancistrum* and *Argentum* may be sister taxa but because the evidence is inconclusive (parsimony bootstrap support of 29% and Bayesian PP=48) they will be considered separately here. The morphologically distinct but closely related dwarf shrubs *Margyricarpus* and *Tetraglochin* are nested within *Argentum.* Sister to these genera, and in fact identical in *trnL/F* sequence, is *Acaena eupatoria,* which, like *Margyricarpus* and *Tetraglochin*, appears out of place in this clade, based on morphology. The fruits of *Tetraglochin* are winged and lack barbs and are more similar to those of *Polylepis* than those of *Acaena*. In *Margyricarpus* the outer portion of the fruit (actually the hypanthium) is soft and fleshy. Likewise, *A. eupatoria* is morphologically very unlike the other *Acaena* species in this clade and is traditionally

placed in section *Euacaena* (Bitter, 1911a) rather than in *Argentum.* I was sufficiently suspicious of this placement that I extracted DNA from a different specimen of *A. eupatoria* and two additional accessions of *Margyricarpus* to verify the sequences. The new sequences were identical to the old, with the exception of a two-base pair deletion near the 5' end of *A. eupatoria* #10307. It appears likely that the chloroplasts of *Margyricarpus, Tetraglochin* and *A. eupatoria* have been acquired through hybridization and introgression, possibly "through" one of the lineages to the others. Identifying the male parental lineage is not possible based on this maternal phylogeny, but it appears that the female parent may have been an ancestral member of section *Argentum*. It is not inconceivable that the acquisition of an *Acaena* chloroplast into the ancestor of *Margyricarpus* and *Tetraglochin* was associated with an allopolyploid hybridization event that gave rise to this odd plant group. A hybrid origin is further supported by a chromosome number of 2n=84 (12x) reported in *Tetraglochin strictum* (Nakata and Oginuma, 1989), which is double that of most *Acaena* (2n=42). The fate of *Acaena eupatoria* may be similar but a chromosome count is unavailable to support this.

Disregarding the placement of *Margyricarpus-Tetraglochin* and *A. eupatoria*, several subclades within *Acaena* correspond rather well to subgeneric groupings established by Bitter (1911a). He divided *Acaena* into two "series": *Axillares* and *Terminales*. Of the species sampled in this analysis Bitter included the following in *Axillares*: *Acaena elongata, A. latebrosa, A. masafuerana, A. cylindristachya*, and *A. montana* (syn. *A. tasmanica*). *Acaena latebrosa* is, as already discussed, sister to all other *Acaena*, but there is not enough resolution in *trnL/F* to infer the relationships

among the rest of *Axillares* (the odd position of *Polylepis quadrijuga* will be discussed below). However, they all share the aspect of being excluded from the other three main *Acaena* clades, a fact that may suggest that they either represent several basal lineages or that they belong to their own clade separate from the rest of *Acaena*. The position of the South Pacific endemic *A. masafuerana* remains uncertain. The *trnL/F* sequence of this species is highly divergent and appears to be either basal to all of *Ancistrum, Argentum* and *Polylepis* or nested within this clade and sister to *Argentum*.

The closely linked species in the *A. pinnatifida*-clade correspond well to Bitter's section *Euacaena*, with the exception of *A. splendens* and *A. digitata*, two species included in *Euacaena* by Bitter but which here have uncertain status. The name *Euacaena* (no rank) will here be applied to the *A. pinnatifida* complex excluding *A. splendens* and *A. digitata* until further evidence suggests otherwise. It is interesting to note that although this clade is predominantly South American it also includes the Australian species *A. echinata*. Hybridization in *Verruchaena* is not only an ancient phenomenon (e.g *Margyricarpis*, *Polylepis*) but has been observed between several extant species as well. The hybrid origin of *Margyracaena skottsbergii* (from *Margyricarpus digynus* and *Acaena argentea)* on Juan Fernandez Islands was demonstrated by RAPD analysis (Crawford *et al.*, 1993). Frequent hybridization has been suggested between the Australian *A. anserinifolia* and the New Zealandic *A. novae-zelandiae* (Dawson, 1960). Hybridization has also been observed between *A. anserinifolia* and *A. echinata* of the "*A. ovina* complex". The species *A.* x *anserovina* was described specifically for this hybrid (Orchard, 1969). The hybrid

origin of plants identified as *A.* x *anserovina* is strongly supported by a phylogenetic analysis of ITS sequences (see pXX, Chapter III). In the *trnL/F* analysis, *A.* x *anserovina* is placed in the *Euacaena* clade, which, given the hybridization hypothesis, implies that a member of the *A. ovina* complex (probably *A. echinata*) was the maternal parent involved in the cross. This is in general agreement with Orchard's hybridization hypothesis, although he stated that *A. echinata* is typically the pollen donor. It is possible that hybridization between these lineages is a recurring phenomenon giving rise to a polyphyletic species *A.* x *anserovina*.

Polylepis

The Bayesian phylogeny suggests that *Polylepis* may be more closely related to the two "Ancistrum" clades than to the rest of *Acaena*, although this is only supported by one single-base pair indel. There is very little sequence variation in *trnL/F* among species of *Polylepis*, nevertheless three weakly supported clades can be distinguished. It is highly likely that hybridization among species of *Polylepis* is frequent (see Chapter VI) and the relationships among chloroplasts thus will only tell part of the story. The three weak clades in the *trnL/F* tree do not correspond to the subgeneric groups recognized by Simpson (1979) nor do they agree very well with the tree hypothesized by Kessler (1995b) based on morphology. In addition, *P. sericea* and *P. incana,* for which more than one accession was sequenced, appear to be polyphyletic. This is not surprising given that *P. sericea* has a very broad but interrupted distribution from Venezuela to Bolivia while *P. incana* shows significant morphological variation and is known to hybridize with other species (Simpson,

1979). The most aberrant result from these data is the firm placement of *Polylepis quadrijuga* as sister to *Acaena elongata*. To verify this result, three different accessions of *P. quadrijuga* were sequenced, all of which agreed on this placement. *Polylepis quadrijuga*, which is confined to the Colombian Andes, clearly belongs to *Polylepis* based on morphology (e.g., woody habit, stipules forming sheaths, hirsute stamens). Like *Margyricarpus-Tetraglochin* and *Acaena eupatoria,* it appears that either *P. quadrijuga* or all of the remaining *Polylepis* has acquired a foreign chloroplast through introgression. Nuclear data suggest, in fact, that the chlorplast of *P. quadrijuga* is "original" and that of other *Polylepis* species the result of introgression. This hypothesis and the origin of *Polylepis* as a whole will be discussed in detail in chapter IV.

Phylogenetic definitions of new or recircumscribed taxa are given in an appendix. In accordance with the Phylocode, stem based definitions were made for five clades: *Poterium, Poteridium, Sanguisorba, Verruchaena* (verrucatus=warty, achaenium= achene) and *Amentomorpha* (amentum=catkin).

Character evolution

Floral evolution

As previously mentioned there is a trend in Sanguisorbinae towards a reduction in the number of reproductive organs. In addition to lacking petals and epicalyx, the number of carpels and stamens has been severely reduced in some genera. While *Poterium* has numerous stamens, there are typically only four in *Sanguisorba*, 2-4 in *Poteridium*, 2-7 in *Acaena*, and the number in *Cliffortia* ranges

from two to numerous. *Poterium* has two carpels while *Sanguisorba* and *Poteridium* have only one. Although the majority of *Cliffortia* species have a single carpel, members of the section *Complanatae* Weimarck (represented here by *C. dentata*) have two carpels and two styles, similar to *Poterium*. *Acaena* has 1-5 carpels. The most parsimonious reconstructions of evolutionary changes in the numbers of stamens and carpels are illustrated in Figure II.5. When more than one equally parsimonious optimization was found, the one that maximized loss rather than addition of organs was preferred, based on the assumption (which may or may not be justified) that independent loss is more prevalent than independent gain.

The molecular phylogeny suggests that gender dimorphism is an example of convergent evolution in separate lineages of Sanguisorbeae. The ancestor of Sanguisorbeae most likely bore perfect flowers, a condition found in the petalous Agrimoniinae (except *Hagenia*) and in the outgroups Potentilleae and Rosaeae. The ancestral condition within the apetalous Sanguisorbinae is, however, less clear. Besides hermaphrodism, three additional sexual conditions are found in Sanguisorbeae: (1) polygamy, in which a combination of hermaphroditic and female flowers make up the inflorescence, (2) monoecy, and (3) dioecy. The evolution of this character is illustrated in figure II.6.A. Because monoecy and dioecy appear to occur interchangebly among closely related species in Sanguisorbeae, these character states were combined in one. The polygamous condition, which is found only in *Poterium*, may be considered a transitional state between perfect and strictly unisexual flowers.

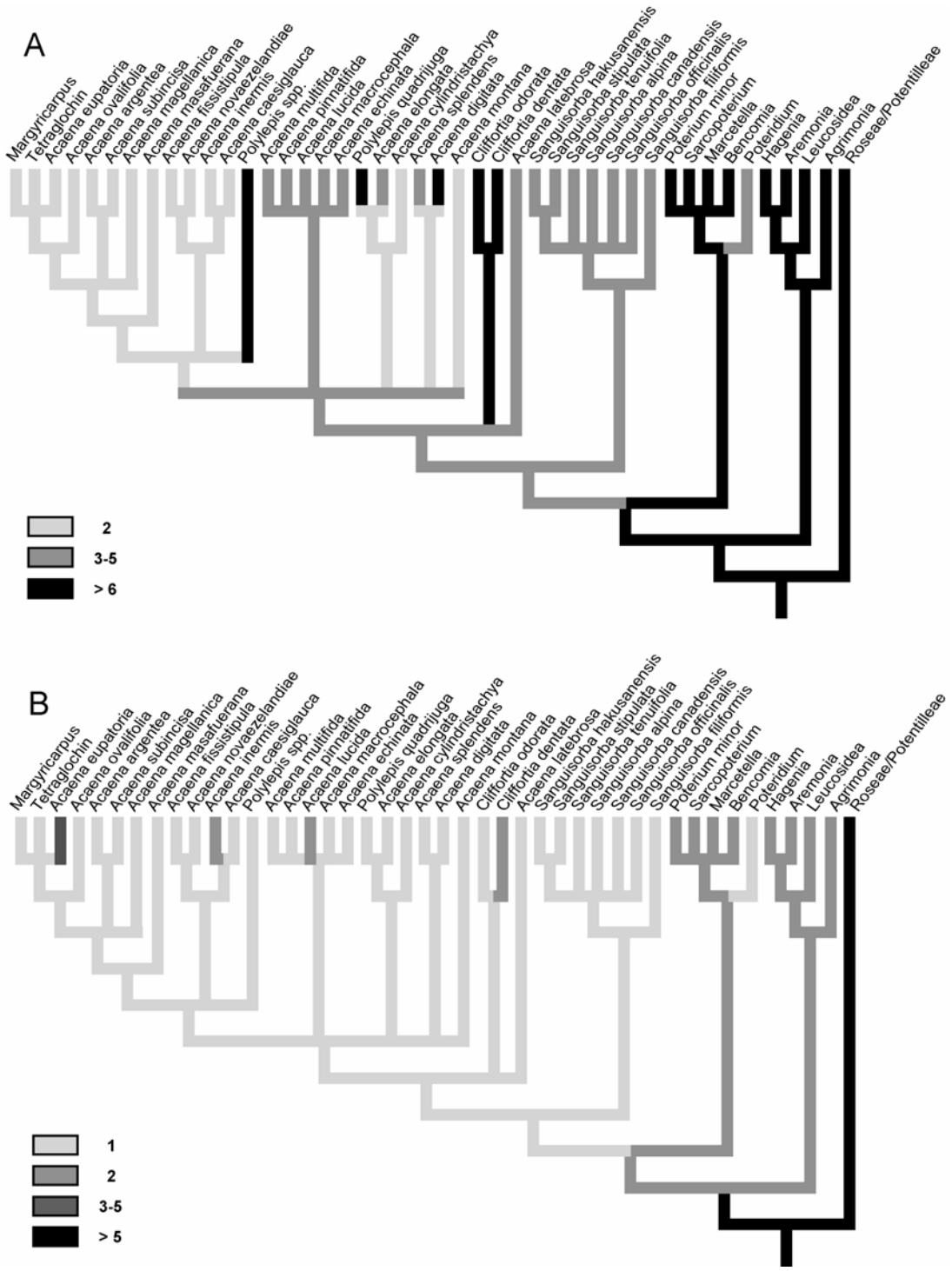


Figure II.5. Reconstruction of the evolutionary reduction in the number of (A) stamens and (B) carpels in Sanguisorbeae.

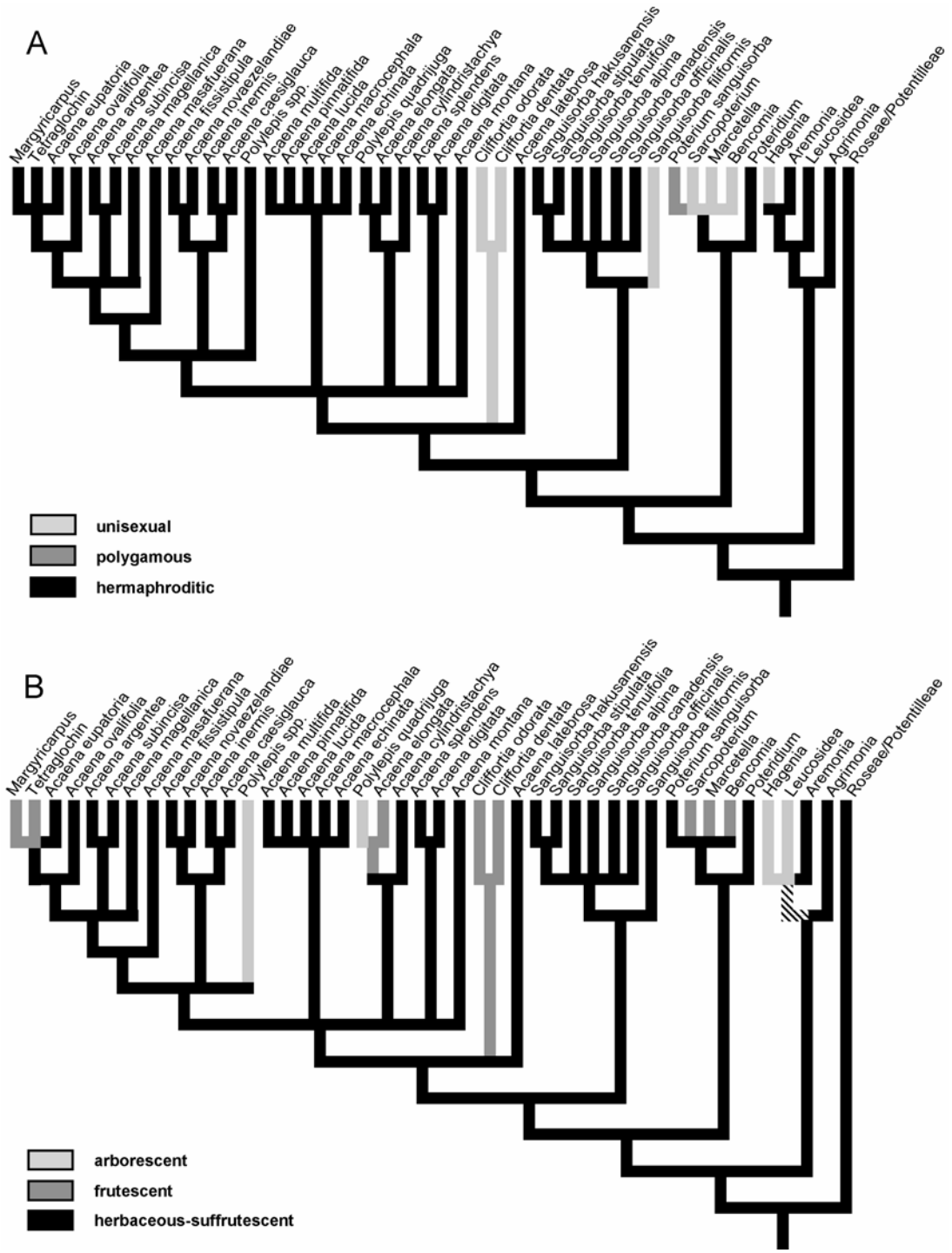


Figure II.6. Reconstruction of the evolutionary change in (A) gender and (B) habit in Sanguisorbeae.

The molecular phylogeny suggests, however, that unisexual flowers may have evolved before polygamous ones in the *Poterium* clade. All the 6-colporate *Sanguisorba* as well as *Poteridium* have hermaphrodite flowers. The anomalous *Sanguisorba filiformis* has monoecious flowers, a condition that ties it to *Poterium* rather than to *Sanguisorba*. Gender dimorphism is also found in *Hagenia* (dioecious), *Marcetella* (dioecious), *Bencomia* (mono- and dioecious) and *Sarcopoterium* (monoecious) of the *Poterium* clade, and in the South African genus *Cliffortia* (monoecious). The most parsimonious reconstruction suggests four independent origins of unisexual flowers in Sanguisorbeae. Constraining the ancestor of the apetalous Sanguisorbinae to be monoecious or dioceous adds at least one parsimony step not counting the transition to polygamy in *Poterium*.

There is clearly a tendency within the Sanguisorbeae to evolve unisexual flowers. Apart from the polygamous *Poterium* and the monoecious *Sanguisorba filiformis*, unisexuality is restricted to shrubs and trees. Interestingly, the evolution of gender dimorphism has recently been associated with polyploidization (Miller and Venable, 2000; Miller and Venable, 2002), something that may be true in Sanguisorbeae as well (see chapter III).

Evolution of habit

A woody habit, like unisexual flowers, appears to have evolved multiple times in Sanguisorbeae (Figure II.6.B). There are no known annual species in the group (*Poteridium* was initially mischaracterized as an annual); all herbaceous species are perennial. Within the Agrimoniinae secondary growth is found in *Hagenia*, which can

grow as tall as 30 meters in the subalpine tropics in East Africa, and in *Leucosidea*, a shrub or small tree from southern Africa. Within the *Poterium* clade *Sarcopoterium* is a sclerophyllous shrub, and *Marcetella* and *Bencomia* are shrubs that sometimes attain a tree-like appearance. *Cliffortia* consists of mostly shrubs but some species become arborescent. Finally, *Margyricarpus* and *Tetraglochin* are dwarf shrubs while *Polylepis* species range from shrubs to large trees. It should be noted that some of the herbaceous perennials tend to have suffrutescent basal branches, especially in *Acaena*, and in some cases the species are not easily classified under either herb or shrub. Nevertheless there is a significant morphological difference between the suffrutescent habit of *Acaena* and the shrubby growth form of *Cliffortia* and *Bencomia* and certainly with the arborescent habit of some species of *Polylepis*. It is considerably more parsimonious to assume multiple origins of frutescent/arborescent growth from herbaceous ancestors than to postulate a woody ancestor of either Sanguisorbeae, Sanguisorbinae, Agrimoniinae or *Verruchaena*.

Chapter III Phylogenetic analyses of the Sanguisorbeae based on duplicated polymorphic nuclear sequence data

Introduction

The phylogenetic analysis of Sanguisorbeae based on chloroplast sequences shed new light on the relationships among basal members of the group (chapter II). However the level of variation in *trnL/F* was not enough to sort out relationships within the southern hemispheric clade *Verruchaena* and the results left several other key questions unanswered. Furthermore, the analysis hinted at the possibility of hybrid speciation being an important component of the evolutionary history of Sanguisorbeae. Hypotheses of hybridization cannot be tested on a single gene tree, in particular if that gene tree is uniparentally inherited.

It can be argued that relying on nucleotide data from a single genome in a phylogenetic study is equivalent to producing a single-character phylogeny (Doyle, 1992). Gene trees can differ from organismal trees because of lineage sorting, introgression or because paralogous, not orthologous, genes were compared. Nuclear genes are typically part of gene families. Using members of gene families in phylogenetic analyses can be problematic due to the risk of mistaking paralogy for orthology. The possibility of recombination between different gene copies can further obscure organismal phylogeny. Nevertheless, the nuclear genome comprises an immense depository of genetic information that can be a valuable source of phylogenetic data if the markers are chosen carefully.

To complement the chloroplast data, two nuclear genes were chosen: the protein coding alcohol dehydrogenase (*Adh*) and the ribosomal internal transcribed spacers (ITS). *Adh* is a small gene family with only a few loci in most seed plants studied. Ribosomal genes like ITS, on the other hand, consist of numerous tandemly repeated units. *Adh* seemed appropriate as a phylogenetic marker based on its putative low copy number, and the ITS gene was chosen because previous studies in the Rosaceae had indicated a higher rate of sequence evolution in ITS than in chloroplast genes. ITS is one of the most commonly used phylogenetic markers and have provided invaluable phylogenetic data in countless studies on fungi, animals and plants. The multiple copies of the gene typically retain a high degree of sequence identity due to the homogenizing effects of concerted evolution. However, one has to be mindful of the risk that if speciation is faster than homogenization, some intragenomic nucleotide diversity may persist (Buckler *et al.*, 1997).

The phylogenies of *Adh* and ITS collectively reveal a complex evolutionary history of the Sanguisorbeae, including, but not limited to, ancient and recent hybridization, allopolyploidization, and inter-locus concerted evolution. Intergeneric relationships in Sanguisorbinae were clarified, leading to an improved understanding of relationships within the problematic *Verruchaena* clade.

Polyploidy and gene duplication in Sanguisorbeae

To facilitate the interpretation of the nuclear molecular phylogenies, a brief review of reproductive biology and chromosome evolution in Sanguisorbeae is necessary. For definitions of homology terms as applied here, see glossary (p. 176).

Gene duplication followed by mutation can lead to new gene function and is almost certainly a driving force in evolutionary innovation. It has also long been known that polyploidization (i.e., duplication of the entire genome) has had a major effect on the size and composition of the nuclear genome, particularly in Angiosperms. As early as 1950, Stebbins argued that most living angiosperms with a haploid chromosome number over 11 were polyploid. More recently, Masterson (1994) used guard cell size to infer chromosome numbers of extinct lineages and from that determined that approximately 70% of extant angiosperms are polyploid. It appears reasonable to assume that at least a portion of the duplicated genes observed in a particular plant taxon originally arose through genome doubling.

Rosaceae has a rich history of polyploidy, including the demonstrated ancient allopolyploid origin of Maloideae (Evans and Campbell, 2002; Morgan *et al.*, 1994), giving rise to such familiar plants as apples, pears and hawthorns. Polyploids are especially frequent in a number of agamic complexes, which occur in *Rubus*, *Alchemilla*, *Crataegus*, *Amelanchier*, and *Sorbus* among other rosaceous genera. The ancestral haploid chromosome number in Rosaceae is thought to have been either n=7 (Darlington, 1963) or n=9 (Challice, 1981; Kalkman, 1988; Raven, 1975; Robertson, 1974) and any number divergent from this (7 or 9) is thought to have arisen through polyploidy or aneuploidy. The ancestral condition in Sanguisorbeae as in all of subfamily Rosoideae is undoubtedly x=7, which is the base chromosome number found in all genera currently recognized in Rosoideae, with the exception of *Alchemilla,* which is x=8 (Morgan *et al.*, 1994). Paleodiploid members of Sanguisorbeae should thus be $2n$=14, a condition found only in the petalous

Spenceria and in the apetalous *Poteridium* (=*Sanguisorba annua*). All the remaining species that have been studied are either 2n=28 (4x), 42 (6x), 56 (8x), 70 (10x), 84 (12x), or 126 (18x) (see Table III.1 for species sampled in this analysis).

Briefly summarized, the sporophytic paleopolyploid chromosome number is 28 (4x) in *Agrimonia* and *Aremonia, Sanguisorba sensu stricto*, and *Poterium*, 42 (6x) in the paraphyletic *Acaena* and 84 (12x) in *Margyricarpus*. No chromosome counts have been published for *Cliffortia* but a tentative count of 2n=42 for *C. dregeana* was made by Christopher Whitehouse (pers. comm.). Exact chromosome counts for *Polylepis* species are unavailable and appear to be difficult to obtain. Both Simpson (1979) and Kessler (1995b) reported ineffective attempts at chromosome staining. Kessler, however, was able to observe that the chromosomes were very small and numerous and estimated the number to be approximately 2n=80 (based on *P. neglecta, P. racemosa ssp. triacontandra* and *P. tarapacana*). With a base chromosome number of x=7 (documented in *Sanguisorba* and *Agrimonia*), 2n≈80 is close to 84, which represents a ploidy level of 12x.

Materials and methods

Taxon sampling

The taxon sampling strategy was similar to that of the *trnL/F* analysis (chapter II). Multiple accessions were extracted from certain species of *Polylepis* that were considered to be morphologically or geographically diverse. A total of 110 taxa were included in the analysis. A few species of Sanguisorbeae and numerous outgroup taxa had been sequenced for previous ITS studies and some of these were included in the

Table III.1. Ploidy levels of taxa sampled in this study (or their close relatives). Number of rDNA loci listed when available (Mishima *et al.* 2002).

Taxon	Close relative in this study	Sporophytic chromosome count	Geographic origin of referenced specimens	Cytological reference	18S-5.8S-26S rDNA loci (Mishima *et al.* 2002)
Acaena argentea Ruiz & Pav.		42 (6x)	Argentina	Roulet 1981	
Acaena integerrima Gill. ex H. & A.	*A. splendens*	42 (6x)	Argentina	Roulet 1981	
Acaena leptacantha Phil.	*A. macrocephala*	c. 41-42 (6x)	Argentina	Roulet 1981	
Acaena magellanica (Lam.) Vahl.		42 (6x)	Sub-antarctic Isls, E. Tierra del Fuego	Moore 1972	
Acaena magellanica (Lam.) Vahl.		84 (12x)	Argentina, Chile	Moore 1972	
Acaena microcephala Schltdl.	*A. montana*	c. 42 (6x)	Argentina	Roulet 1981	
Acaena ovalifolia Ruiz & Pav.		42 (6x)	Peru, Argentina	Oginuma *et al.* 1988	
Acaena pinnatifida Ruiz & Pav.		42 (6x)	Argentina	Moore 1981	
Acaena caesiglauca (Bitter) Bergmans		42 (6x)	New Zealand	Beuzenberg *et al.* 1983	
Acaena fissistipula Bitter		42 (6x)	New Zealand	Beuzenberg *et al.* 1983	
Acaena inermis Hook.f.		c. 42 (6x)	New Zealand	Beuzenberg *et al.* 1983	
Acaena dp. Aff. anserinifolia (J.R. et G. Forst.) Druce	*A. anserinifolia*	42 (6x)	New Zealand	Beuzenberg *et al.* 1983	
Agrimonia eupatoria L.		28 (4x) [1]	Northern temperate	Murín 1997, Hollingsworth, et al 1992, Dobes 1997 (among others)	(6 – A. nipponica)
Aremonia agrimonoides (L.) DC		42 (6x)		Murín et al 1987, Baltisberger 1991	
Bencomia caudata (Ait.) Webb & Berth		28 (4x)	Canary Islands	Larsen 1956	
Marcetella moquiniana (Webb. & Berth.) Svent.		28 (4x)	Canary Islands	Larsen 1956	
Sanguisorba annua (Nutt. ex Hook.) Torr. & Gray		14 (2x)	North America		4
Sanguisorba alpina Bunge in Ledeb.		28 (4x)	Asia	G. Nordborg 1966	

Taxon	Close relative in this study	Sporophytic chromosome count	Geographic origin of referenced specimens	Cytological reference	18S-5.8S-26S rDNA loci (Mishima *et al.* 2002)
Sanguisorba canadensis L.		28 (4x) [1]	North America	G. Nordborg 1966	
Sanguisorba hakusanensis Mak.		28 (4x)	Japan	Sakai 1935	4
Sanguisorba minor Scop. *(=Poterium)*		28 (4x) [1]	Eurasia	Böcher & Larsen 1957, Erdtman *et al.* 1961, Dobes *et al.* 1997 (among others)	4
Sanguisorba officinalis L.		28 (4x) [1]	Eurasia	Krogulevich 1984, Mishima *et al.* 2002	6
Sanguisorba stipulata Rafin.		28 (4x)	Japan	Mishima *et al.* 2002	4
Sanguisorba filiformis (Hook. F.) Hand.-Mazz.		28 (4x)	China	Mishima *et al.* 2002	4
Sanguisorba tenuifolia Fisch. ex Link		56 (8x) [1]	Japan	G. Nordborg 1966, Mishima *et al.* 1996, Mishima *et al.* 2002	8
Sarcopoterium spinosum (L.) Spach		28 (4x)	Mediterranean	Slavizk 1993. Larsen 1955	
Spenceria ramalana Trimen		14 (2x)	China	Takahashi, A. et al 1996	
Tetraglochin cristatum (Britt.) Rothm.		84 (12x)	Andes	Nakata *et al.* 1989	
Potentilla		14 (2x) [1]	temperate cosmopolitan	Krogulevich 1978 (among others)	
Rosa		14 (2x) [1]	temperate cosmopolitan	Malecka *et al.* 1994 (among others)	2 (R. multiflora)

current analysis. No *Adh* sequences closely related to Sanguisorbeae were available prior to the study. *Potentilla, Rosa* and *Filipendula* were chosen as outgroups in this analysis (Table III.2).

Phylogenetic markers

The *Adh* gene family has been well characterized in organisms ranging from bacteria, to animals, fungi (yeast) and plants. Within Angiosperms it has been the marker of choice to trace the origins of polyploid crop plants. Among other studies, *Adh* was used to examine evolutionary relationships and nucleotide substitution rates in grasses and palms (Gaut *et al.*, 1996; Gaut *et al.*, 1999), the origin of rice species (Ge *et al.*, 1999), hybrid speciation in peonies (Ferguson and Sang, 2001; Sang *et al.*, 1997; Sang and Zhang, 1999) and duplication and allopolyploidy in cotton (Small *et al.*, 1999; Small and Wendel, 2000a; Small and Wendel, 2000b; Small and Wendel, 2002).

Most *Adh* loci consist of nine introns and ten exons, but genes with fewer introns have also been documented. Its popularity as a phylogenetic marker stems from a low copy number that typically ranges from 2-5. It appears that the gene has been duplicated independently many times and it is impossible to establish orthology above the "family level". Because of the lack of understanding of the global evolution of *Adh*, the naming of different *Adh* loci in different taxa has been inconsistent and confusing. The names given to loci used in this study, *Adh1* and *Adh2,* may not be orthologous to genes with the same name in other taxa.

Table III.2. Species used in the phylogenetic analyses of *Adh* and ITS listed alphabetically. Letters in parentheses indicate herbaria where vouchers are deposited. [1]Sequenced by Torsten Eriksson. [2]Sequences obtained from GenBank.

Species	Source/voucher information	Geographical origin	ITS GenBank accession
Acaena anserinifolia (JR Forst. & G. Forst.) Druce.	P. & M. Hibbs 293 (MARY)	Australia	AY634778
Acaena argentea R. & P.	M. Hibbs 173 (MARY)	Chile	AY634781-87
Acaena argentea R. & P.	J. Clark 5819 (MARY)	Ecuador	AY634849-54
Acaena caesiglauca Bergmans	van Balgooy 4329 (MO)	New Zealand	AY634788
Acaena cylindristachya R. & P.	M. Hibbs 167 (MARY, LPB)	Bolivia	AY634789-93
Acaena digitata Phil.	M. Hibbs 181 (MARY)	Chile	AY634794-99
Acaena echinata Nees.	D.E. Symon 13386	Australia	
Acaena elongata L.	J. Clark 5500 (MARY)	Ecuador	AY634801-04
Acaena elongata L.	N. Wikström 298 (MARY)	Ecuador	
Acaena elongata L.	J. Clark 5818 (MARY)	Ecuador	AY634800
Acaena eupatoria Cham. &Schltdl.	R. C. Molon *et al.* 12175 (US)	Brazil	AY634805
Acaena eupatoria Cham. &Schltdl.	R. Wasum *et al.* 10307 (US)	Brazil	
Acaena fissistipula Bitt.	M. Hibbs 65 (MARY)	New Zeeland (RBGE)	AY634806-08
Acaena inermis Hook.f.	M. Hibbs 57 (MARY)	New Zeeland (RBGE)	AY634809-12
Acaena latebrosa Aiton.	C. M. Whitehouse 122	South Africa: Komsberg	AY634813-15
Acaena lucida Vahl.	D. M. Moore 679 (US)	Falkland Islands	
Acaena lucida Vahl.	G. T. Prance 28561 (NY)	Falkland Islands	AY634816
Acaena macrocephala Poepp.	M. Hibbs 63 (MARY)	Chile (RBGE)	AY634817-18
Acaena magellanica (Lam.) Vahl	M. Hibbs 61 (MARY)	Falkland Islands (RBGE)	AY634819-21
Acaena masafuerana Bitt.	G. Kuschel 215 (US)	Juan Fernandez Islands	
Acaena montana Hook.	M. Hibbs 58 (MARY)	Tasmania (RBGE)	AY634822
Acaena multifida Hook.f.	M. Hibbs 60 (MARY)	Chile, Argentina (RBGE)	AY634823-24
Acaena multifida Hook f.	M. Hibbs 183 (MARY)	Chile	AY634846-48
Acaena novae-zelandiae Kirk	P. & M. Hibbs 292 (MARY)	Tasmania	AY634826-29
Acaena novae-zelandiae Kirk	D. E. Symon 15299 (MO)	Australia	AY634825
Acaena ovalifolia R. & P.	M. Hibbs 175 (MARY)	Chile	AY634830-33
Acaena pinnatifida R. & P.	M. Hibbs 182 (MARY)	Chile	AY634840-43

Species	Source/voucher information	Geographical origin	ITS GenBank accession
Acaena pinnatifida R. & P. *ssp. grandiflora* Bitter	M. Hibbs 180 (MARY)	Chile	AY634838-39
Acaena pinnatifida R. & P. *ssp. grandiflora* Bitter	M. Hibbs 176 (MARY)	Chile	AY634834-37
Acaena splendens Gilles ex H. & A.	S. Teillier *et al.* 2324 (MO)	Chile	AY634855-59
Acaena subincisa Wedd.	M. Hibbs 174 (MARY)	Chile	AY634860-61
Acaena x anserovina Orch.	D. E. Symon 15310 (MO)	Australia	AY634779-80
Acaena x splendens	L. R. Landrum *et al.* 8227 (MO)	Chile	
Agrimonia eupatoria[1] L.	Eriksson *et al.* 41 (SBT)	Germany	U90798
Agrimonia parviflora Sol.	M. Hibbs 122 (MARY)	North America	AY634862-63
Aremonia agrimonioides (L.) DC. ssp. pouzarii Skalicky *(Adh)*	E. M. Rix *et al* 612 (E)	Greece	
Aremonia agrimonioides[1] Necker ex Nestler (ITS)	Eriksson *et al.* 1998	unknown	U90799
Bencomia caudata (Ait.) Webb & Berth	M. Vretblad 46 (MARY)	Canary Islands	AY634864-66
Cliffortia burmeana Burtt Davy	O.M. Hilliard *et al.* 14681 (E)	South Africa	
Cliffortia cuneata Dryand.	Helfgott *et al.* 2000	unknown	AF183520/43
Cliffortia dentata Willd.	C. M. Whitehouse 34	South Africa	AY634867
Cliffortia heterophylla H. Weim.	unknown	South Africa	AY634873
Cliffortia nitidula R.E. & T.C.E. Fries	Helfgott *et al.* 2000	Kenya	AF183521/44
Cliffortia odorata L.f.	C. M. Whitehouse 71	South Africa	AY634874
Cliffortia ruscifolia L.	C. M. Whitehouse 72	South Africa	AY634868-72
Filipendula vulgaris[1] L.	Eriksson *et al.* 2003	Sweden	AJ416463
Hagenia abyssinica J.Gmelin *(Adh)*	Knox 2532 (GH)	Kenya	
Hagenia abyssinica J.Gmelin (ITS)	Eriksson *et al.* 1998	Kenya	U90800
Leucosidea sericea E. & Z.	Helfgott *et al.* 2000	South Africa	AF183524/47
Marcetella moquiniana (Webb. & Berth.) Svent.	M. Vretblad 49 (MARY)	Canary Islands	AY634875
Margyricarpus pinnatus (Lam.) Kuntze	M. Hibbs 66 (MARY)	Andes (RBGE)	AY634877-78
Margyricarpus pinnatus (Lam.) Kuntze	M. Hibbs 184 (MARY)	Chile	AY634876
Margyricarpus setosus R. & P.	M. Hibbs 136 (MARY, LPB)	Bolivia	AY634879-80
Polylepis australis Bitter	D. Renison s.n. 33	Argentina	AY634881-84
Polylepis australis Bitter	D. Renison s.n. 39	Argentina	AY634885-89
Polylepis besseri Hier. *ssp. besseri*	M. Hibbs 158 (MARY, LPB)	Bolivia	AY634890-98

Species	Source/voucher information	Geographical origin	ITS GenBank accession
Polylepis besseri Hier. *ssp. incarum* (Bitter) M.Kessler	M. Hibbs 172 (MARY, LPB)	Bolivia	AY634899-AY634910
Polylepis besseri Hier. *ssp. subtusalbida* (Bitter) M.Kessler	M. Hibbs 164 (MARY, LPB)	Bolivia	AY634911-14
Polylepis crista-galli Bitter	M. Kessler 3661 (MO)	Bolivia	AY634915-20
Polylepis hieronymi Pilger	M. Hibbs 133 (MARY, LPB)	Bolivia	AY634921
Polylepis incana H.B.K.	J. Clark 6228 (MARY)	Ecuador	AY634927-28
Polylepis incana H.B.K.	J. Clark 4991 (MARY)	Ecuador	AY634922-26
Polylepis lanuginosa H.B.K.	J. Clark 6227 (MARY)	Ecuador	AY634933-36
Polylepis lanuginosa H.B.K.	G. Harling *et al.* 22858 (MO)	Ecuador	AY63492932
Polylepis multijuga Pilger	S. Llatas Quiroz 2749 (MO)	Peru	AY634937-40
Polylepis neglecta M.Kessler	M. Hibbs 135 (MARY, LPB)	Bolivia	AY634941-45
Polylepis pauta Hieron.	B. Eriksen 59086 (NY)	Ecuador	AY634951-52
Polylepis pauta Hieron.	Schmidt-Lebuhn 378	Ecuador	AY634946-50
Polylepis pepei Simpson	St.G. Beck *et al.* 21532	Bolivia	
Polylepis pepei Simpson	I. Jimenez 1360	Bolivia	AY634953
Polylepis quadrijuga Bitter	E.G.B. Kieft *et al.* 143 (NY)	Columbia	AY634957
Polylepis quadrijuga Bitter	P. Franco *et al.* 5632 (MO)	Colombia	AY634958-60
Polylepis quadrijuga Bitter	S. R. Gradstein	Colombia	AY634954-56
Polylepis racemosa R. & P. *ssp. lanata* (O. Kuntze) M. *Kessler*	M. Hibbs 169 (MARY, LPB)	Bolivia	AY634961-64
Polylepis racemosa R. & P. *ssp. triacontandra* (Bitter) M. *Kessler*	J.C. Solomon 15492 (MO)	Bolivia	AY634965-67
Polylepis reticulata Hier.	J. Clark 4992 (MARY)	Ecuador	AY634968-78
Polylepis rugulosa Bitter	M. Kumar 34	Chile	AY634985-90
Polylepis rugulosa Bitter	M. Kumar 35	Chile	AY634991-98
Polylepis sericea Wedd.	J. Clark 5820 (MARY)	Ecuador	AY634999-AY635000
Polylepis sericea Wedd.	K. Romoleroux *et al.* 1495B (NY)	Ecuador	
Polylepis sericea Wedd.	P.M. Jørgensen *et al.* 1240 (MO)	Ecuador	
Polylepis subsericans Macbr.	A. Tupayachi *et al.* 1192 (MO)	Peru	AY635005-11
Polylepis tarapacana Philippi	M. Hibbs 163 (MARY, LPB)	Bolivia	AY635012-15
Polylepis tomentella Wedd. *ssp. tomentella*	M. Hibbs 162 (MARY, LPB)	Bolivia	AY635016-23

Species	Source/voucher information	Geographical origin	ITS GenBank accession
Polylepis tomentella Wedd. *ssp. incanoides M. Kessler*	M. Hibbs 137 (MARY, LPB)	Bolivia	
Polylepis weberbaueri Pilger	J. Clark 6229 (MARY)	Ecuador	AY634979-84
Polylepis weberbaueri Pilger	S. Laegaard 53795 (MO)	Ecuador	AY635024
Potentilla erecta[1] *(L.)Räusch.*	Eriksson *et al.* 1998	Sweden	AH006918
Potentilla indica Wolf. (= *Duchesnea indica* (Andrews) Focke)	M. Hibbs 69 (MARY)	N. America (introduced)	AY635025
Prunus sp.	M. Hibbs 67 (MARY)	N. America	
Rosa majalis[1] Herrm.	Eriksson *et al.* 1998	Sweden	U90801
Rosa multiflora Thunb. ex Murray	M. Hibbs 78 (MARY)	N. America	AY635026
Rosa persica[1] Michx.	Eriksson *et al.* 2003	Iran, Afghanistan (Uppsala Botanic Garden)	AJ416468
Rubus idaeus[2] *L.*	Alice *et al.* 2001	Sweden	AF055757
Sanguisorba alpina Bunge in Ledeb.	M. Hibbs 294 (MARY)	Mongolia (MOBOT)	AY635027-29/44
Sanguisorba ancistroides (Desf.) Cout. (=*Poterium ancistroides*)	Helfgott *et al.* 2000	Morocco	AF183530/53
Sanguisorba annua (Nutt. ex Hook.) Torr. & Gray	E. Earle 4412 (NH)	N. America	AY635031-32
Sanguisorba annua (Nutt. ex Hook.) Torr. & Gray	B. Maguire *et al.* (PH)	N. America	AY635030
Sanguisorba canadensis L.	S. Woodbury 152 (MO)	N. America	AY635033-34
Sanguisorba filiformis (Hook. F.) Hand.-Mazz.	J. F. Rock 17963 (NH)	China: SW Szechwan	
Sanguisorba hakusanensis Mak.	M. Hibbs 59 (MARY)	Japan or Korea (RBGE)	AY635035
Sanguisorba minor Scop. (=*Poterium sanguisorba*)	M. Hibbs 62 (MARY)	Mediterranean (RBGE)	AY635036-39
Sanguisorba minor Scop. (=*Poterium sanguisorba*)	M. Hibbs 89 (MARY)	Sweden	
Sanguisorba officinalis L.	T. Eriksson 144 (SBT)	Sweden	AY635040-41
Sanguisorba parviflora[1] (Maxim.) Tak.	Eriksson *et al.* 1998	Cult. (ex Siberia)	U90797
Sanguisorba stipulata Rafin.	K. Dillman 17 (NH)	Alaska	AY635042-43
Sanguisorba tenuifolia Fisch. ex Link	X. Ling 81906 (MO)	China	
Sarcopoterium spinosum[2] (L.) Spach	Helfgott *et al.* 2000	unknown	AF 183534/57
Tetraglochin cristatum (Britt.) Rothm.	M. Hibbs 150 (MARY, LPB)	Bolivia	AY635045-47

The internal transcribed spacers ITS1 and ITS2 are located between the 18S and 5.8S rRNA genes and between the 5.8S and 26S rRNA genes, respectively. Nuclear ribosomal DNA (rDNA) consists of thousands of copies of tandemly repeated paralogs. Due to the homogenizing effect of concerted evolution, these paralogs tend to be identical. Assuming that concerted evolution proceeds faster than speciation, the number of paralogs should not be an obstacle to phylogenetic reconstruction. The number of phylogenetic studies in which ITS has been used is impressive and continuous to accumulate. Part of the appeal is that despite the high rate of sequence evolution, amplification is generally unproblematic using universal primers that match the highly conserved flanking ribosomal genes. In this study ITS1, ITS2 and the intervening 5.8S were used.

DNA extraction, PCR amplification and Sequencing

The DNA extraction protocol was outlined in detail in chapter II. The locus I call *Adh1* was initially amplified from a few species using primers AdhF1 and AdhR1 of Sang *et al.* (1997). Based on initial sequences and GenBank data two new primers Adh2F and Adh9R were designed to more perfectly match Sanguisorbeae (Table III.3). These primers, as it turns out, selectively amplified a different *Adh* locus (*Adh2*). Another set of primers (AdhF1b and AdhR2b) was designed to target *Adh1* at the exclusion of *Adh2*. A complete list of primer sequences is found in table III.3. ITS1 and ITS2 were amplified in one segment using universal primers ITS4 (White *et al.*, 1990) and ITS.LEU.I (Urbatsch *et al.*, 2000).

Table III.3. Primers used in the PCR amplification and sequencing of *Adh*.

Primer name[1]	Sequence
ADHF1 (Sang *et al.* 1997)	5'-TACTTCTGGGAAGCYAAGG-3'
ADHF2 (Sang *et al.* 1997)	5'-CCTCGCATATTTGGTCACGAAG-3'
ADHF1b	5'-TCTGGGAAGCTAAGGTAAA-3'
ADHF1.Rosaceae1	5'-TACTTCTGGGAAGCCAAGG-3'
ADHF1.Rosaceae2	5'-TACTTCTGGGAAGCTAAGG-3'
ADHF1.Sanguisorba	5'-TGGGAAGCTAAGGTAAATGT-3'
ADH1.4F	5'-AGTCAGAGGAAAGCAACAT-3'
ADH1.4Fb	5'-ACATGTGTGATCTCCTAAG-3'
ADH1.7F	5'-TGAATTTGTGAATCCAAAAGA-3'
ADH1.7Fb	5'-ATTTGTGAATCCAAAAGA-3'
ADH1.7Fc	5'-GTTTGTGAATCCAAAAGA-3'
ADH1S.A.cyl.i4F	5'-TATATTACTGCAACAATATACA-3'
ADH1S.Polylepis.i2F	5'-AGGTAAATATTTAACTTACT-3'
ADHR1 (Sang *et al.* 1997)	5'-CCCTTRAGMGTCCTCTCATTC-3'
ADHR2 (Sang *et al.* 1997)	5'-GGGCACACCAACAAGTACTG-3'
ADHR2b	5'-GGCACACCAACAAGTACTGC-3'
ADHR2c	5'-GGCACACCAACAAGTACAGC-3'
ADH1R.Bencomia	5'-GTGTCTTTAAAATCAGGGTT-3'
ADH1.4R	5'-ATGTTGCTTTCCTCTGACT-3'
ADH1.7R	5'-TCTTTTGGATTCACAAATTCA-3'
ADH1.7Rb	5'-TCTTTTGGATTCACAAAT-3'
ADH1.7Rc	5'-TCTTTTGGATTCACAAAC-3'
ADH1L.Ra	5'-ACTGAGTAGTGCTAAGAGC-3'
ADH1L.Rb	5'-AGTGTAGCAGACTTGAAACTCT-3'
ADH1L.Polylepis.i8R	5'-AGAAAATGAGTTACAGTGTT-3'
ADH1L.Polylepis.i8-e9R	5'-AACCCCCCAACCCTGATTTA-3'
ADH1L.Acaena.i8-e9R	5'-AACCCCCCAACCCTAATTTA-3'
ADH1S.Polylepis.i8-e9R	5'-AACCCCCCAACCCTAATTTT-3'
ADH1S.B.Ra	5'-ACTGCAACCCCCCAACCCTG-3'
ADH1S.B.Rb	5'-AACCCCCCAACCCTGATTTT-3'

[1]S=small locus, L=large locus. 4 through 9 refer to Adh exon numbers unless specified as either i=intron or e=exon.

PCR amplification of both *Adh* and ITS was conducted in 25µl reaction volume with final concentrations of 3mM $MgCl_2$, 0.2µM each primer, 0.25µM each dNTP, 10-50ng of genomic DNA and 2.5 units of Promega *Taq* polymerase or a combination of 2 units of *Taq* and 0.05µl of Stratagene Pfu polymerase. Each template was amplified with and without 2.5µl DMSO. PCR products were purified using polyethylene glycol (PEG) precipitation (see chapter II for details). The precipitate was eluted in 5-10µl de-ionized water. Visualization on agarose gels revealed that *Adh* products included two bands of approximately 1500 and 1900bp each. Although ITS products were visible as a single band, direct sequencing revealed that these products were polymorphic as well. Because of the complete conservation of the flanking regions of rDNA combined with too much variation in the ITS sequence itself, it was not possible to design primers to selectively target the different loci. To ensure that all copies of the ITS gene were found, a range of PCR conditions were applied to each template including different annealing temperatures, $MgCl_2$ concentrations and denaturing agents.

Polymorphic *Adh* and ITS products were ligated into Promega pGEM-T Easy vectors following the manufacturer's protocol. Ligations were transformed in either Promega or Stratagene competent cells. After heat-shocking the cells for 45 seconds at 42°C, 1ml SOC medium was added to the transformation and cultures were allowed to stabilize in a shaking water bath at 37°C for 1.5 hours. The cultures were then plated on LB-agar plates (containing ampicillin, X-gal and IPTG for blue/white screening) and incubated at 37°C over night. Between 5 and 25 white colonies from each transformation (species) were PCR amplified directly using vector primers T7

and M13R. The reaction volume was scaled down to 20µl and only 0,25 units of *Taq* were needed per reaction. Otherwise the reaction mix was the same as the initial PCR (without DMSO).

Sequencing reactions were performed on an ABI Prism 3100 automated sequencer, as outlined in chapter II. Sequencing primers for both *Adh* and ITS were the vector primers T7 and M13R, and for *Adh* four additional internal sequencing primers were designed: Adh4F, Adh7F, Adh4R and Adh7R. A total of 592 *Adh* clones and 857 ITS clones were sequenced. Chromatograms were manually proof read and sequence contigs were assembled using the computer program Sequencher 3.1.1 (Gene Codes Corporation, Ann Arbor, MI). Sequence identities were verified as belonging to Rosaceae by BlastN searches against published sequences (NCBI). Sequences were aligned by eye using MacClade (Maddison and Maddison, 2002). Portions that could not be aligned unambigously were excluded from the analysis.

Optimal and suboptimal secondary structures of ITS2 were explored based on thermodynamic properties as implemented in MFold (Zucker). Published secondary structures of the ITS2 (Goertzen in press) were used to constrain the sequence adjascent to 5.8S/26S when the four-domain structure was not found automatically. Minimum free energy values were compared between different ITS loci to determine functionality.

Phylogenetic methods

Phylogenetic analyses were conducted under the maximum parsimony criterion and Bayesian inference. Because of the size of the data sets maximum

likelihood analyses were not practical. Analyses were performed using PAUP* 4.0b10 (Swofford, 2002) and MrBayes (Huelsenbeck and Ronquist, 2001). Clade support was estimated by nonparametric bootstrap analyses in parsimony and posterior probabilities in the Bayesian analyses. Likelihood ratio tests were conducted using ModelTest to determine the model of sequence evolution with the best fit for *Adh* and ITS data respectively. Optimal models were then applied to the Bayesian analyses.

Results

Adh

In the first round of sequencing, two very divergent loci of *Adh* were found in the Sanguisorbeae. These will be referred to as *Adh1* and *Adh2*. Orthology of sequences from different species was inferred from shared gene structure and nucleotide similarity. After a small data set of *Adh1* and *Adh2* had been gathered, the exon sequences were compiled and compared with published sequences in GenBank (www.ncbi.nclm.nih.gov). A simple phylogenetic analysis was conducted using a neighbor-joining algorithm on corrected (GTR) distances. The unrooted "global" *Adh* phylogeny is presented in figure III.I. The purpose of this analysis was simply to investigate whether sequences from orthologous loci in other taxa were available and to estimate how deep the split between *Adh1* and *Adh2* is. An attempt was made to include a broad spectrum of angiosperm taxa but with emphasis on sequences with high nucleotide similarity to Sanguisorbeae *Adh1* and *Adh2*. There are four previously published sequences of *Adh* for Rosaceae, three of which

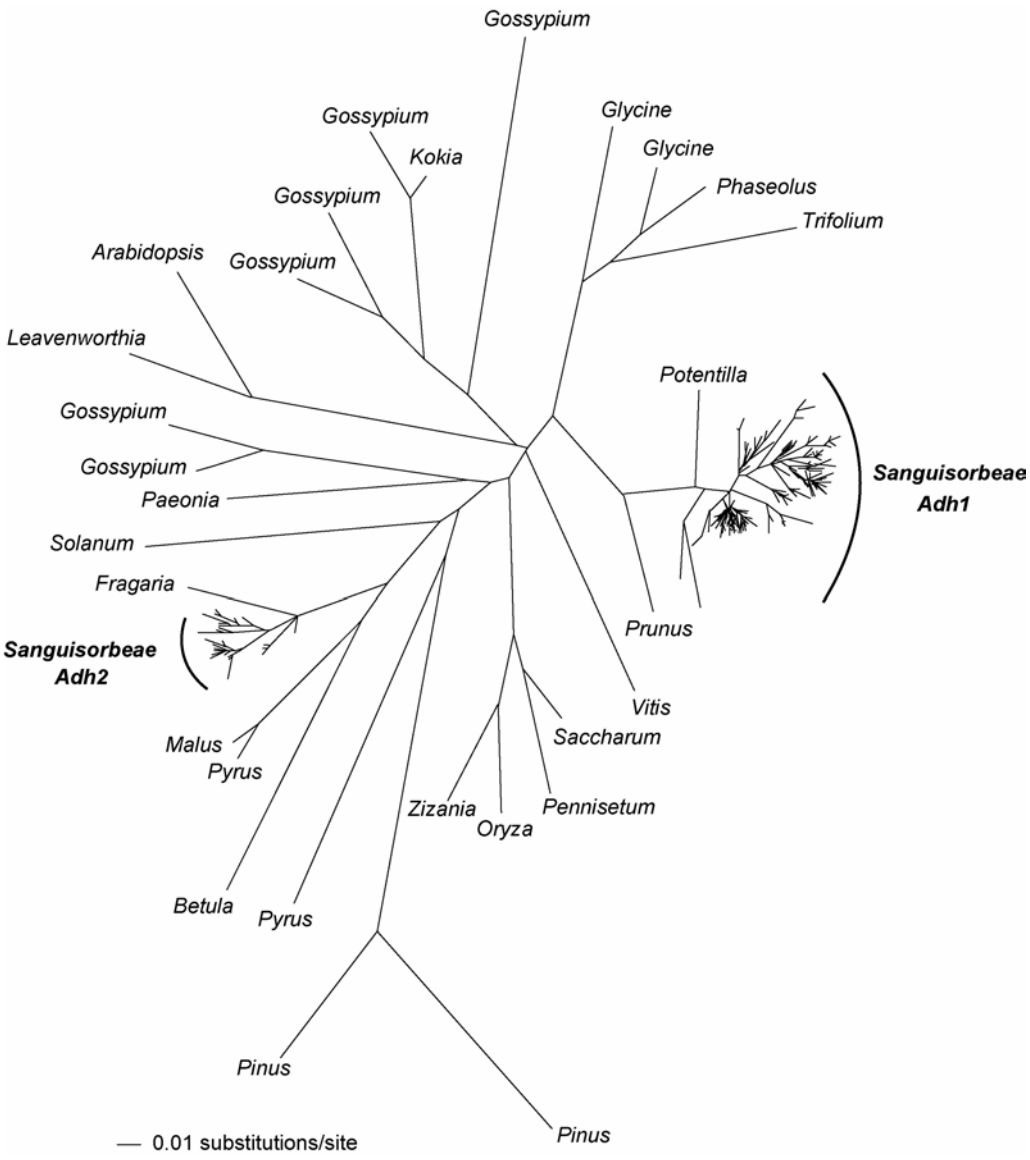


Figure III.1. *Adh* loci "1" and "2" of Sanguisorbeae in relation to published spermatophyte sequences. Unrooted phylogram based on a neighbor joining analysis of correceted (GTR+I+Γ) distances.

(*Fragaria* X15588, *Malus* Z48234 (mRNA) and *Pyrus* AF031900 (mRNA)) are closely related to Sanguisorbeae *Adh2* and one (*Pyrus* AH031899 (mRNA)) that is highly divergent and very distantly related to any of the Sanguisorbeae sequences. There are no Rosaceae sequences of *Adh1* in GenBank (NCBI). The unrooted phylogeny supports a monophyletic Rosaceae *Adh1* distantly sister to a clade of legume sequences. However, without genetic mapping data, an assessment of orthology beyond Rosaceae is tentative at best. The duplication giving rise to *Adh1* and *Adh2* appears to be very ancient, occuring well before the origin of Rosaceae.

A phylogenetic analysis of the limited initial *Adh2* data set revealed an early duplication event within *Adh2* followed by apparent additional duplication events giving rise to multiple sequences for each species. Intra-genomic sequences appeared to have undergone extensive post-duplication recombination (incomplete concerted evolution) and, based on prior knowledge of relationships, there were no discernable phylogenetic structure among species. Because *Adh2* appeared to poorly reflect organismal phylogeny subsequent sequencing focused exclusively on *Adh1*.

Like *Adh2*, the *Adh1* lineage has undergone a duplication event early in Sanguisorbeae evolution (possibly associated with the same genomic rearrangement as the early duplication in *Adh2*). The duplication gave rise to two loci of different size. *Adh1L* (L=large) appears to have the standard *Adh* gene order of nine exons and eight introns (although the first exon and intron was not sequenced), while *Adh1S* (S=small) is missing intron seven. The gene structure of *Adh1S, Adh1L* and *Adh2* is outlined and compared in Figure III.2.

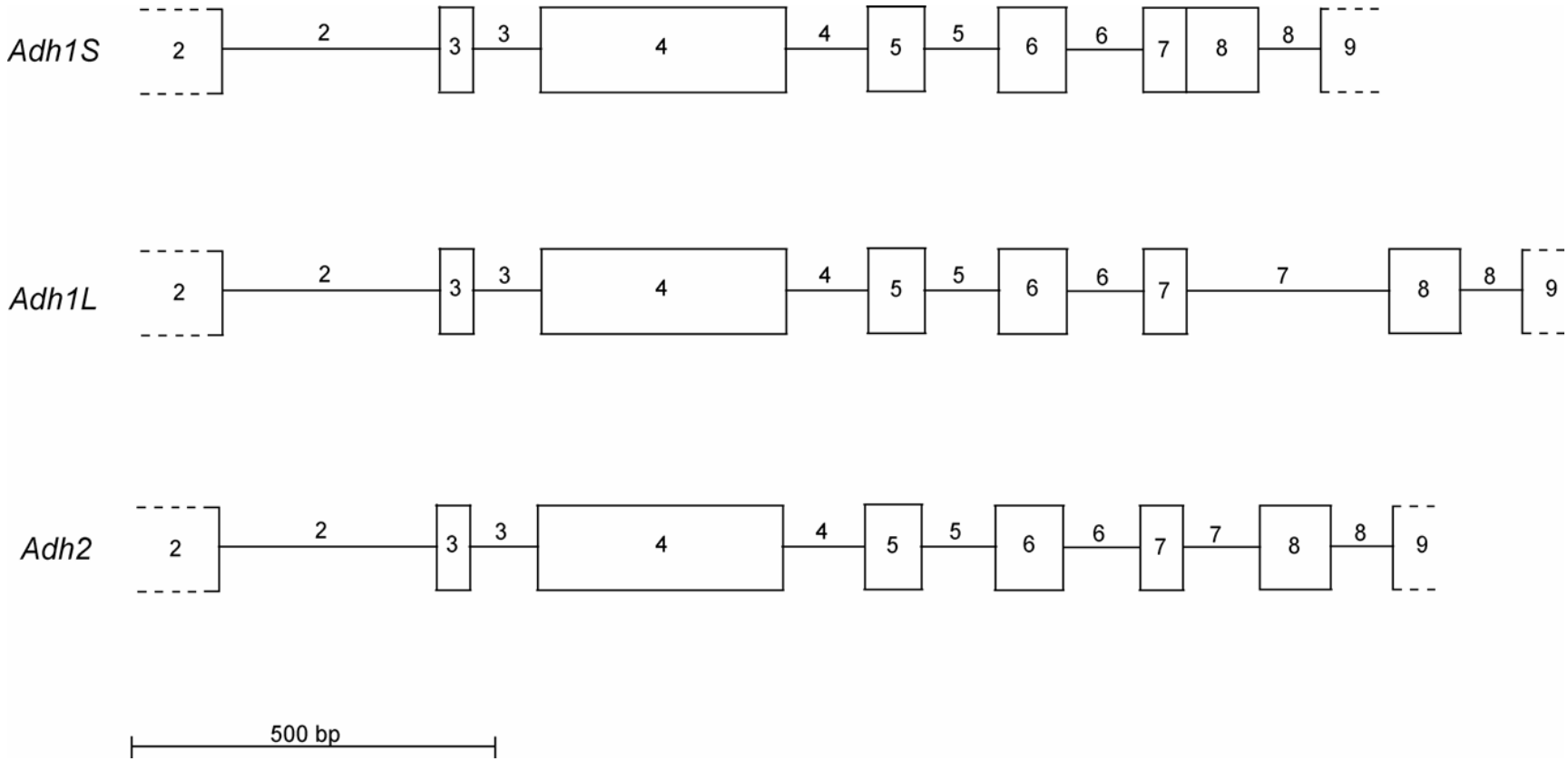


Figure III.2 *Adh* gene structure in Sanguisorbeae. *Adh1L* and *Adh1S* were used as phylogenetic markers.

In an effort to find a complete set of *Adh1* copies from each species, several PCR amplifications were conducted for each species, using different combinations of six 5' primers and ten 3' primers, designed to amplify different copies. Copies of *Adh1L* and *Adh1S* could be distinguished on an agarose gel because of their difference in size, but copies within *Adh1S* could not be distinguished based on size. In order to ensure detection of polymorphism within the same PCR product, a large number of clones were sequenced from each cloned amplicon. It is possible that additional *Adh1* copies exist but were not detected because of PCR selection and drift. In addition, DNA extracted from herbarium material could in some cases not be amplified. *Adh* was more difficult to amplify than ITS because of the length of the sequence. Because of the occurrence of multiple loci, amplifying the gene in smaller segments with internal primers was not an option.

In addition to the presence of four major intergeneric *Adh1* copies all species displayed low-level polymorphism that was manifested as closely related species "combs" (polytomies). It was suspected that in cases where intraspecific variation was strictly autapomorphic, this might be caused by *Taq* polymerase error. This was corroborated by re-amplification of three species using Pfu, a polymerase with high proof-reading ability. The cloned, Pfu-amplified sequences again included representatives from the different *Adh1* loci but the short autapomorphic intraspecific branches were removed. Because polymerase error is random and autapomorphic it was judged that this variation would not bias the phylogenetic analyses other than slightly overestimating terminal branch lengths. After initial phylogenetic analyses polytomic species combs were reduced to one sequence, retaining the sequence with

the shortest terminal branch length (presumably with the lowest level of error). It has to be acknowledged that by excluding polymorphic sequences some real allelic differences may also have been erased.

The phylogeny of Sanguisorbeae *Adh1* was essentially the same regardless of the optimality criterion, with the exception of the placement of one *Sanguisorba* clade (see below). An outline of the entire *Adh1* genealogy in Sanguisorbeae based on the Bayesian analysis is depicted in figure III.3. Because of the size of the data set the phylogenies of *Adh1L* and *Adh1S* are presented separately in figures III.4A and B. Clades of *Polylepis* are reduced to one (average length) branch and the evolution of *Polylepis* will be discussed in detail in chapter V and VI.

Adh1L

Prunus, *Potentilla*, and the petalous Sanguisorbeae (*Agrimonia, Aremonia and Hagenia*), form single (or independently duplicated) lineages basal to a monophyletic apetalous Sanguisorbeae *Adh1.S+Adh1.L* suggesting that the duplication in *Adh1* occurred after the divergence of Agrimoniinae. No *Adh1S* genes were found in *Poterium, Marcetella* and *Bencomia*. Instead two closely related adh1L copies were found suggesting an independent duplication in the ancestor of this clade. Instead of being sister to the rest of Sanguisorbinae, as was shown by chloroplast data (chapter II), *Poterium, Marcetella* and *Bencomia* are here sister to the *Adh1L* copy of *Sanguisorba*. *Poteridium*, like the *Poterium* group, has no *Adh1S* but an internally duplicated *Adh1L* gene that is basal to the *Poterium* group and *Sanguisorba*.

Figure III.3. Overview of *Adh1* genealogy in Sanguisorbeae based on Bayesian inference under a GTR+I+Γ model of sequence evolution. Branch lengths are drawn proportional to the amount of change.

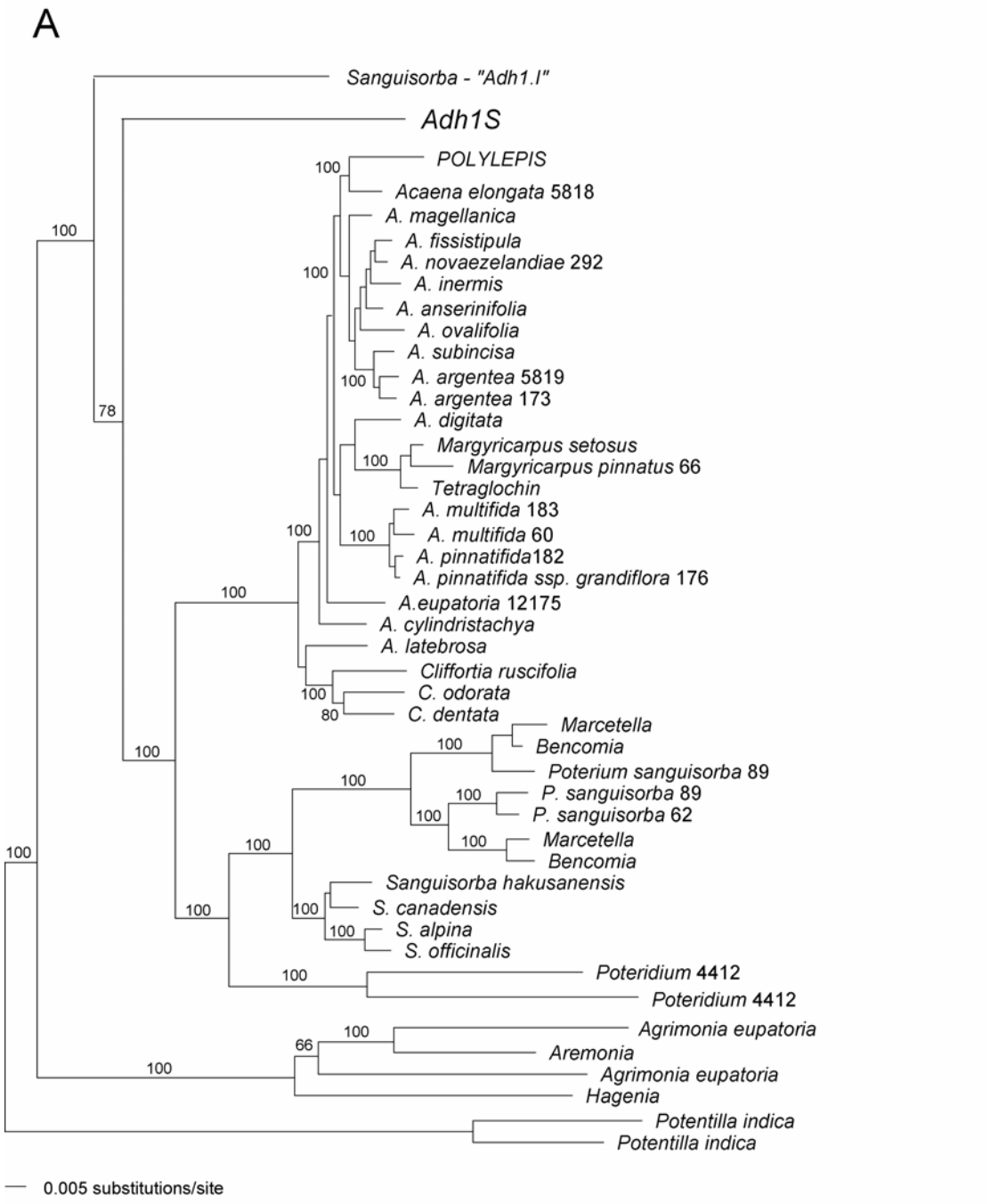
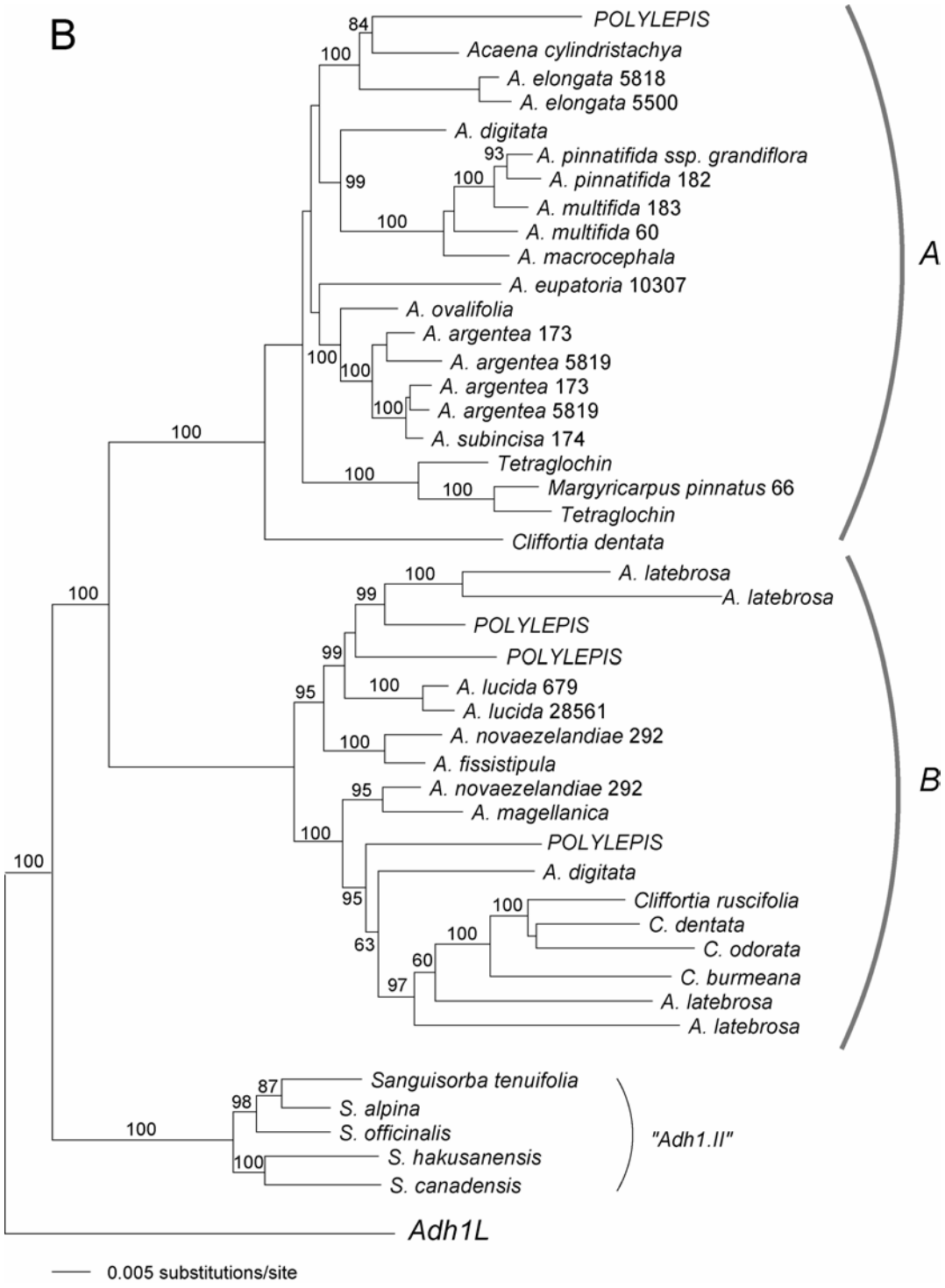


Figure III.4. Phylogeny of (A) *Adh1L* and (B) *Adh1S* in Sanguisorbeae. Subclades magnified from the overall phylogeny depicted in Figure III.3 based on Bayesian inference. Posterior probabilities indicated above branches. Clades of *Polylepis* sequences were reduced to one representative per clade. Collection numbers are indicated for those species that were represented by more than once accession (see Table III.2).


B
POLYLEPIS
Acaena cylindristachya
A. elongata 5818
A. elongata 5500
A. digitata
A. pinnatifida ssp. grandiflora
A. pinnatifida 182
A. multifida 183
A. multifida 60
A. macrocephala
A. eupatoria 10307
A. ovalifolia
A. argentea 173
A. argentea 5819
A. argentea 173
A. argentea 5819
A. subincisa 174
Tetraglochin
Margyricarpus pinnatus 66
Tetraglochin
Cliffortia dentata
A
A. latebrosa
A. latebrosa
POLYLEPIS
POLYLEPIS
A. lucida 679
A. lucida 28561
A. novaezelandiae 292
A. fissistipula
A. novaezelandiae 292
A. magellanica
POLYLEPIS
A. digitata
Cliffortia ruscifolia
C. dentata
C. odorata
C. burmeana
A. latebrosa
A. latebrosa
B
Sanguisorba tenuifolia
S. alpina
S. officinalis
S. hakusanensis
S. canadensis
"Adh1.II"
Adh1L
0.005 substitutions/site

As in the *trnL/F* phylogeny, the southern hemispheric genera *Cliffortia*, *Acaena*, and *Polylepis* form a monophyletic group (Figure III.4A). Within the southern hemispheric clade, *Cliffortia* is sister to the South African species *Acaena latebrosa*, confirming the chloroplast results that *A. latebrosa* is only distantly related to the rest of *Acaena*. There is very little variation among the basal lineages of *Acaena* and *Polylepis* but several well-supported *Acaena* clades can be distinguished. These correlate well with subgeneric taxa established by Bitter (1911a) and with chloroplast data. *Acaena* section *Ancistrum* sensu Bitter (including both the Australasian *Ancistrum sensu stricto* and the South American *Argentum*, see p.44) forms a well-supported clade with the subantarctic *A. magellanica* being basal. The sequences of section *Euacaena* are very similar and form a sister relationship with *A. digitata* from Bitter's series *Axillares*. The relationship of *A. cylindristachya* (*Axillares*) and *A. eupatoria* (*Eucaena*) is unresolved, while *A. elongata* is sister to *Polylepis*. As with the chloroplast data, the internal nodes in *Polylepis* are weakly supported and do not correspond with traditional subgeneric taxa.

Adh1S

Scattered nucleotide insertions and deletions in the exons of *Adh1S* suggest that some or all of these sequences are no longer functional. The phylogenetic analysis reveals that *Adh1S* was duplicated at the base, giving rise to *Adh1S.A* and *Adh1S.B* (Figure III.4B). These two copies have the same general gene structure (lacking intron 7), and consequently orthology can only be inferred from phylogenetic analysis. The *Cliffortia-Acaena-Polylepis* clade of *Adh1SB* is internally relatively

homogeneous but highly divergent and easily distinguished from the *Adh1SA* sequences of the same taxa. However, the orthology between these two clades and the two *Adh1S* clades of *Sanguisorba* is more difficult to establish. Because of this uncertainty the *Sanguisorba* clades will be referred to as *Adh1S.I* and *Adh1S.II* until further evidence allows for a conclusive homology assessment. *Adh1S.I* appears to have retained many plesiomorphic characters and in most analyses will appear basal to *Adh1L+Adh1S*. This placement was considered dubious based on the fact that it shares the fundamental gene order of the rest of *Adh1S* and in an analysis using a reduced data sets this *Sanguisorba* clade is in fact sister to the rest of *Adh1S* (not shown). The phylogenetic reconstruction of *Adh1S* is probably impeded by the highly divergent *Adh1S.B*.

Assuming that the two copies of *Sanguisorba* are each basal to *Adh1S.A* and *Adh1S.B* the relationships of *Adh1S* are similar to *Adh1L* but slightly more complex with apparent gene loss in several taxa. Despite numerous amplification attempts with *Adh1S.A*-specific primers a *Cliffortia* copy of *Adh1S.A* could only be detected in *C. dentata*. Again, due to drift and selection in the polymerase chain reaction, this cannot be taken as proof that more copies do not exist. Similarly, *Adh1S.A* copies were not found from any of the Australasian *Acaena* section *Ancistrum* and *A. magellanica*. The fact that this represents an entire missing clade supports the notion that these copies have in fact been lost. Within *Polylepis* the same pattern of "missing" species is evident (see chapter VI). Disregarding the missing taxa, the phylogenetic relationships agree with *Adh1L*, with the exception that *Acaena cylindristachya* rather than *A. elongata* is sister to *Polylepis*.

The evolution of *Adh1S.B* is extraordinarily complex and difficult to tease apart. A close inspection of the sequences points to an early duplication followed by recombination between the two copies. Many of the sequences are clearly chimeric between two *Adh1S.B* "types". An alternative explanation is that the sequences of this clade are especially susceptible to PCR recombination and that the chimeric sequences are artifacts. This hypothesis is challenged by the fact that identical chimeric sequences were amplified multiple times (with and without the denaturing agent dimethyl sulfate). A detailed discussion of *Adh1S.B* in *Polylepis* follows in chapter VI.

Maybe the most interesting result of the *Adh1S* data is that species tend to have a copy of either *Adh1S.A* or *Adh1S.B* but not both. The only exception to this is *Cliffortia dentata*. Although it is difficult to explain this pattern, the fact that primer combinations specific to *Adh1S.B* failed to amplify anything among species of the *Adh1S.A*-clade suggest that this was not simply due to random preferential amplification by PCR. If this pattern is real it suggests a strong propensity to retain a single copy of *Adh1S.* It does not explain, however, why the two copies were kept for so long in the ancestral lineages of *Acaena* and *Polylepis*.

ITS

Alignment of ITS sequences was relatively straight-forward within Sanguisorbeae but not always so between the ingroup and the outgroup. An ambiguous region of approximately 26 bp was excluded from the analyses. Similar to the cloned *Adh* sequences, some polymerase error was detected in the ITS data but to

a lesser extent due to the much shorter sequences. Species "combs" of very similar sequences were reduced to one sequence, retaining the one with shortest terminal branch length (which was typically a zero-length branch).

Even before any analyses had been conducted, an inspection of the sequence alignment revealed multiples highly divergent copies of the ITS region in most Sanguisorbeae species. It was immediately obvious that some of these copies had lost its function and become pseudogenes. This assessment was based on (1) the presence of insertions and deletions in the 5.8S region, (2) highly divergent 5.8S region (i.e., similar amount of sequence change as in the spacers), and (3) the inability to find optimal or suboptimal thermostable secondary structures with the characteristic four domain configuration (see further discussion on this below). A series of analyses were conducted including and excluding pseudogene clades. These will be discussed in detail below.

Initial heuristic parsimony searches on ITS1, ITS2 and 5.8S showed signs of overwhelming homoplasy and a vast number of most parsimonious trees were gathered. After removing the 5.8S region the rescaled consistency index value surprisingly increased substantially, even though the 5.8S region typically is highly conserved. After close inspection of the normally highly conserved 5.8S region, a variable four-base region was found, corresponding to the small hairpin of the third domain of 5.8S. Among the putatively functional ITS sequences, five different sequence motifs were found (GUUA, AUUA, GUCA, UUCA, and UUUA), but their distribution appeared to make no phylogenetic sense. If there are no structural constraints on the hairpin loop one would expect these characters to behave in a

similar way as the ITS sequences, which appeared not to be the case. The only apparent explanation for the distribution of these characters is convergent evolution of this motif among unrelated taxa. As expected, a partition-homogeneity test suggested that the ribosomal gene was not combinable with ITS1 and ITS2. Subsequent analyses excluded the 5.8S region.

Pseudogenization

Optimal and suboptimal secondary structures of ITS2 were explored based on minimum free-energy configurations as implemented in the computer program MFold (Walter *et al.*, 1994). Suboptimal foldings within 5% of the minimum free energy value were also examined. For sequences of single-copied outgroups and of *Verruchaena* locus A-A.I , optimal configurations corresponding to the characteristic four-domain model of ITS2 (Goertzen in press) were found immediately without constraining any stem regions. For the remaining sequences (copies B, C and D), however, this structure was not found, even when the stem region connecting 5.8S to ITS2 was constrained to pair. In addition, the minimum free-energy values for these alternative configurations were significantly higher than for sequences of locus A (Figure III.5). Several additional highly divergent species or clades of ambiguous orthology appeared to lack functional secondary structure including a divergent *Poterium* clade, one clone of *Sanguisorba officinalis* and *S. alpina* respectively. As would be expected, base frequencies were very different in functional genes and pseudogenes.

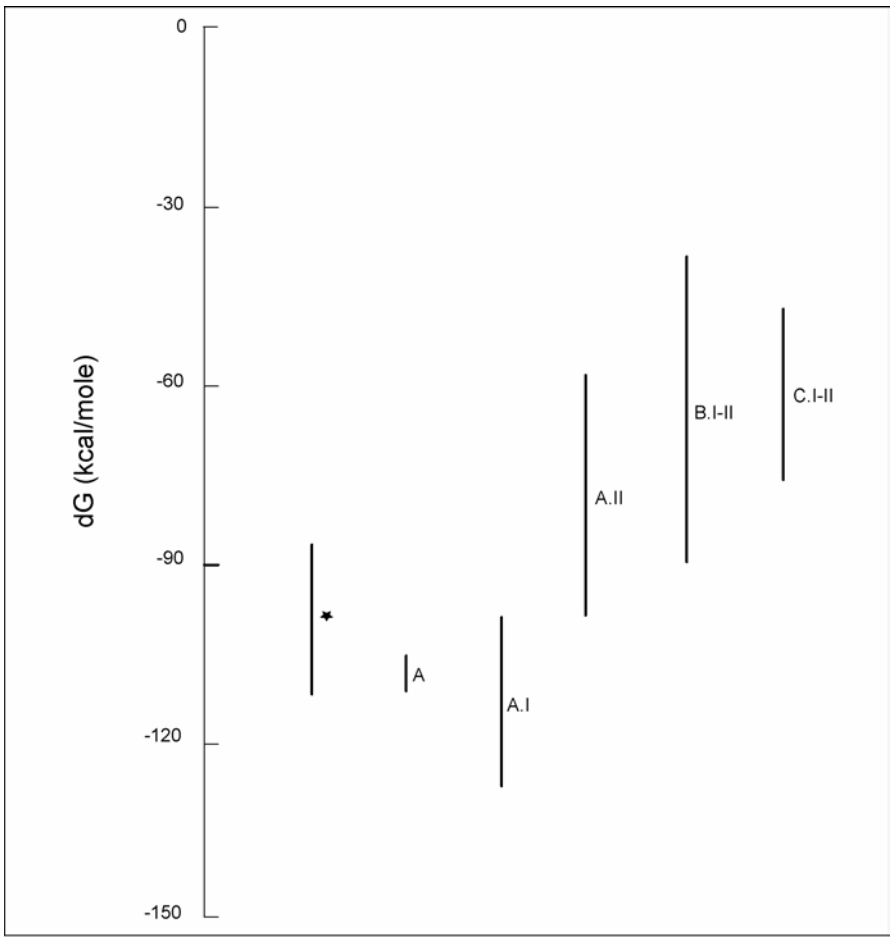


Figure III.5. Free energy values of estimated optimal seccondary structure configurations for ITS2 sequences representing six different clades (loci) in Sanguisorbeae. (*) functional locus *Sanguisorba, Poterium,* Agrimoniinae, (A) functional locus Cliffortia, (A.I) functional locus *Amentomorpha*, (A.II) pseudogene *Amentomorpha*, (B.I-II) pseudogene *Verruchaena,* (C.I-II) pseudogene *Polylepis*. The bars represent the range from the lowest to highest free energy value found within the clade. Sequences from species outside of *Verruchaena* were assumed to be orthologous and placed in the same (unnamed) "locus group" with the exception of one sequence each of *Poteridium, Sanguisorba officinalis* and *S. alpina* for which orthology could not be established. All sequences of *Polylepis, Acaena,* and *Cliffortia* were "assigned" to a locus based on phylogenetic affinities with the exception of two sequences of *Acaena splendens* (sister to locus A.II) for which homology could not be established.

While GC content was 61.1% in single-locus Sanguisorbeae, and 62.6% in the putatively functional sequences of *Amentomorpha* (locus A.I), it was close to 50% in all of the putative pseudogenes suggesting a release from selection for a stable secondary structure (Table III.4). The assessment of functionality based on sequence evolution in 5.8S was thus corroborated by the analysis of secondary structure as well as base frequencies.

Predicted and observed numbers of ITS loci

The phylogenetic signal in the ITS dataset is obscured by the great variation in evolutionary rates caused by the pseudogene formation. In addition, there is evidence of loss of copies and/or missing data due to preferential PCR amplification. An attempt was made to estimate the completeness of the data set while establishing orthology among sequences. According to fluorescence in situ hybridization (FISH) experiments by Mishima *et al.* (2002) the diploid outgroups *Rubus* and *Rosa* have one 18S-5.8S-26S rDNA locus per haploid genome while most of the paleotetraploid *Sanguisorba* and *Poterium* as well as the diploid *Poteridium* have two 18S-5.8S-26S rDNA sites (Table III.1). The paleotetraploid *Agrimonia, Sanguisorba officinalis*, and *S. japonensis* were found to have three loci per haploid genome, one more than would be expected by the ploidy levels. *S. tenuifolia* and *S. albiflora* had four sites consistent with their octaploid genomes. Mishima's findings suggest that although most duplications in the Sanguisorbeae ITS data occurred as a result of polyploidization, additional duplications have occurred in the absence of genome doubling.

Table III.4. Base frequences of putative functional versus pseudogenized ITS loci.

ITS locus	% A	% C	% G	% T	% GC
single-copied species (functional)	19.0	31.2	29.9	19.9	61.1
A.I (functional)	18.7	32.0	30.7	18.6	62.6
A.II (pseudogene)	22.9	27.9	26.3	22.9	51.0
B.I-II (pseudogene)	22.7	28.3	26.7	22.4	54.9
C.I-II (pseudogene)	24.5	26.6	24.0	24.9	50.5

In light of the FISH data, and assuming no random lineage sorting of ancestral alleles, a comparison was made between expected and observed number of ITS loci. It appears as though a complete set of ITS copies were found in the following northern hemispheric taxa: *Agrimonia* (2, albeit from different species), *Poteridium* (2), and *Poterium* (>2). Autotetraploids are common in *Poterium sanguisorba* (Nordborg, 1966; Nordborg, 1967) thus the four copies found here do not necessarily contradict the FISH data but suggest that the particular specimen sampled may have been a tetraploid. Another explanation could be that the gene has undergone another duplication in addition to the genomic doubling. Research on genomic rearrangements following polyploidization suggests that additional duplications beyond those resulting directly from the genome doubling may be common in recent polyploids (Paterson *et al.*, 2000). Two ITS loci were found in *Sanguisorba officinalis*, which is one fewer than predicted by FISH. However, because the third copy must have arisen through simple gene duplication it is possible that two of the copies have retained identical or near identical sequences through concerted evolution (near identical autapomorphic sequences were removed because of suspected polymerase error). Except for *S. officinalis* and *S. alpina,* the remaining *Sanguisorba* lack a second locus. Judging from the FISH data and based on the monophyly of *Sanguisorba sensu stricto* (see chapter II) it is unlikely that this represents a real loss.

Mishima (2002) did not investigate the southern hemispheric genera in the Sanguisorbeae but qualified predictions of copy numbers were inferred from ploidy levels. The different loci have been named A through C.II based on genealogical

affinities. Despite phylogenetic data and information from structural gene organization, orthology cannot be established unambiguously for all loci. Nevertheless it was considered informative to label the clades according to the current best estimate of homology. As predicted, *Acaena* was split into three clades likely representing the three different loci expected from a ploidy level of 2n=42 (6X). However, many *Acaena* species are only found in one or two of the clades and a complete set has so far only been found in four species. An additional duplication event appears to have occurred within locus B not associated with an increase in ploidy. This duplication appears to have affected the subgeneric groups *Ancistrum* and *Argentum*, and was likely followed by recombination between the two copies.

Assuming that the tentative ploidy estimate of 2n=42 (6x) is correct, *Cliffortia* appears to be missing locus B. In fact, in the majority of *Cliffortia* species, only a single ITS copy was found. Once again, future studies may show that this locus has not been lost but were missed despite repeated amplification attempts, or that concerted evolution has homogenized the sequences. Based on an approximate ploidy of 2n=84 (12x) (Kessler, 1995b), six ITS loci are predicted in *Polylepis*. Four monophyletic clades of *Polylepis* were found, two of which appear to be assemblages of two copies each (B and C). Because of the lack of phylogenetic structure in the *Polylepis* clades, along with indications of hybridization and/or random assortment of ancestral alleles, orthology among ITS copies from different species was difficult to establish. In addition, there is evidence of inter-locus recombination, especially within what are putatively loci B.I-II and C.I-II, respectively. Judging from the

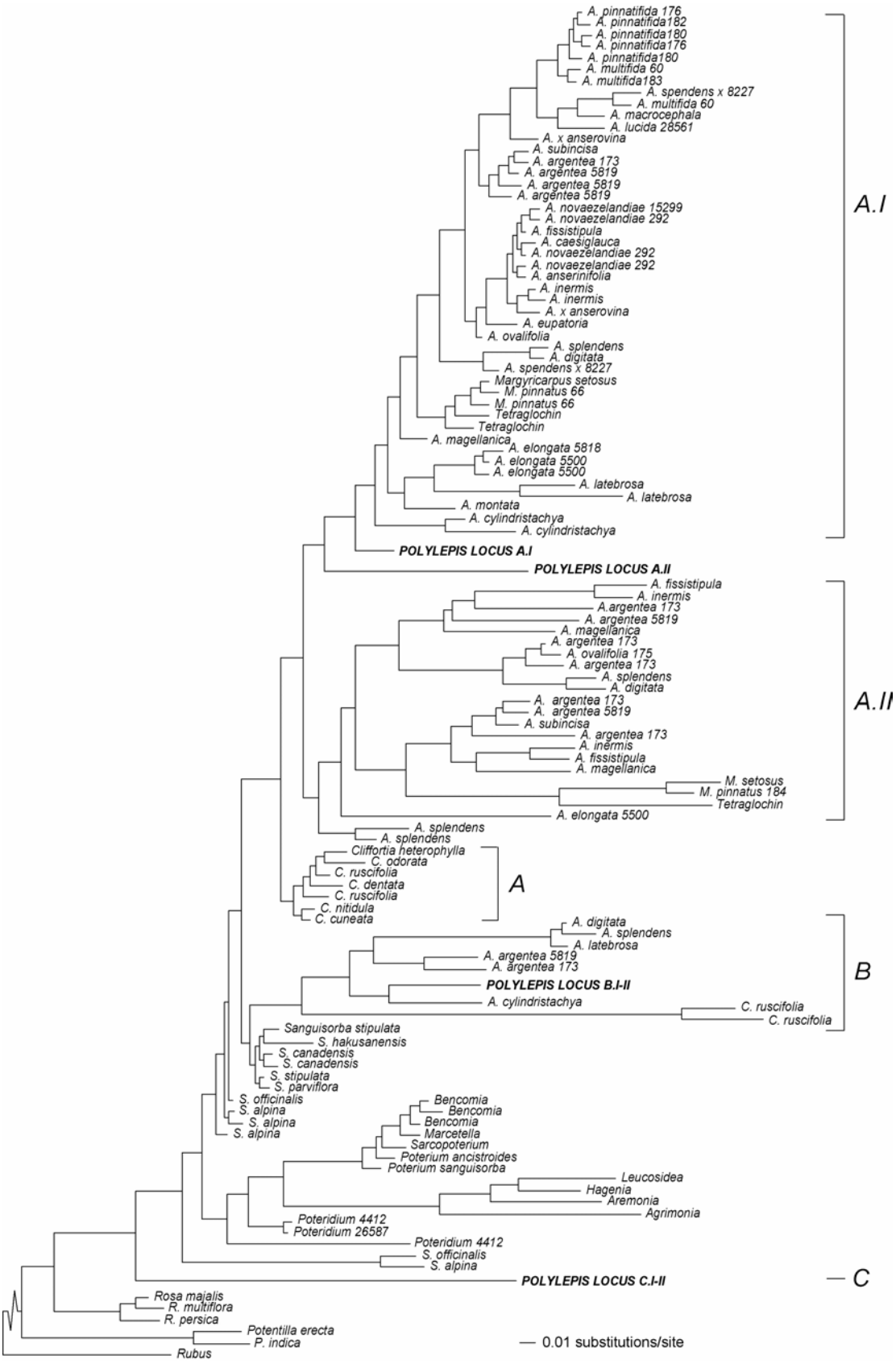
numbers of copies found in some *Polylepis* species, it appears that additional duplication may have taken place.

ITS Phylogeny

The results of the phylogenetic analyses of ITS are at first sight perplexing, but a careful inspection of the data suggests that they are in fact informative on several levels. Several instances of hybridization can be detected, and some phylogenetic relationships within the functional clade A have strong support (as measured by bootstrap and posterior probabilities). The ITS sequences also provide considerable insight into the evolutionary fates of duplicated nuclear genes in polyploids.

Oddly enough, *Polylepis* loci C.I and C.II were reconstructed as sister group to the rest of Sanguisorbeae in all analyses (Figure III.6). This is probably unlikely and could be due to the severe degree of divergence found in these sequences. The position of the petalous Agrimoniinae is ambiguous. In most analyses it was placed sister to the two *Poterium* clades, and not basal within Sanguisorbeae as expected from *trnL/F* (chapter II) and *matK* data (Mishima *et al.*, 2002). The sequences of Agrimoniinae seemed, on close inspection, to be chimeric, with some sections matching well against *Poterium* and other sections more similar to *Potentilla*. One could argue that such a combination of sequence motifs simply reflects the expected appearance of a lineage that it is sister to Sanguisorbeae but has retained a certain degree of ancestral characteristics. However, if that were true one would expect

Figure III.6. Overall genealogy of ITS1 and ITS2 in Sanguisorbeae based on Bayesian inference under a GTR+I+Γ model. *Verruchaena* clades A through C represent different loci. Branch lengths drawn proportional to change. Collection numbers are indicated for those species that were represented by more than once accession. Clades of *Polylepis* sequences were reduced to one.


A. pinnatifida 176
A. pinnatifida182
A. pinnatifida180
A. pinnatifida176
A. pinnatifida180
A. multifida 60
A. multifida183
A. spendens x 8227
A. multifida 60
A. macrocephala
A. lucida 28561
A. x anserovina
A. subincisa
A. argentea 173
A. argentea 5819
A. argentea 5819
A. argentea 5819
A. novaezelandiae 15299
A. novaezelandiae 292
A. fissistipula
A. caesiglauca
A. novaezelandiae 292
A. novaezelandiae 292
A. anserinifolia
A. inermis
A. inermis
A. x anserovina
A. eupatoria
A. ovalifolia
A. splendens
A. digitata
A. spendens x 8227
Margyricarpus setosus
M. pinnatus 66
M. pinnatus 66
Tetraglochin
Tetraglochin
A. magellanica
A. elongata 5818
A. elongata 5500
A. elongata 5500
A. latebrosa
A. latebrosa
A. montata
A. cylindristachya
A. cylindristachya
A.I
POLYLEPIS LOCUS A.I
POLYLEPIS LOCUS A.II
A. fissistipula
A. inermis
A.argentea 173
A. argentea 5819
A. magellanica
A. argentea 173
A. ovalifolia 175
A. argentea 173
A. splendens
A. digitata
A. argentea 173
A. argentea 5819
A. subincisa
A. argentea 173
A. inermis
A. fissistipula
A. magellanica
M. setosus
M. pinnatus 184
Tetraglochin
A. elongata 5500
A.II
A. splendens
A. splendens
Cliffortia heterophylla
C. odorata
C. ruscifolia
C. dentata
C. ruscifolia
C. nitidula
C. cuneata
A
A. digitata
A. splendens
A. latebrosa
A. argentea 5819
A. argentea 173
POLYLEPIS LOCUS B.I-II
A. cylindristachya
C. ruscifolia
C. ruscifolia
B
Sanguisorba stipulata
S. hakusanensis
S. canadensis
S. canadensis
S. stipulata
S. parviflora
S. officinalis
S. alpina
S. alpina
S. alpina
Bencomia
Bencomia
Bencomia
Marcetella
Sarcopoterium
Poterium ancistroides
Poterium sanguisorba
Leucosidea
Hagenia
Aremonia
Agrimonia
Poteridium 4412
Poteridium 26587
Poteridium 4412
S. officinalis
S. alpina
POLYLEPIS LOCUS C.I-II
C
Rosa majalis
R. multiflora
R. persica
Potentilla erecta
P. indica
Rubus
0.01 substitutions/site

Agrimoniinae to be basal to other Sanguisorbeae in the phylogenetic analysis. This was not seen in any of the analyses. In fact when *Poterium* and *Poteridium* were removed from the analysis (not shown) the Agrimoniinae clade was suddenly embedded within Potentilleae. A hybrid origin of tetraploid Agrimoniinae is thus a possibility.

Chimeric sequences - genomic and/or PCR recombination

Because of the putative chimeric nature of Agrimoniinae ITS, this clade was removed from further analyses. Removing Agrimoniinae also increased the consistency index suggesting a decrease in the amount of homoplasy in the phylogeny. However, even after removing Agrimoniinae, certain terminal branches in the tree were suspiciously divergent, even within the putative functional locus A. A comparison of these sequences to the rest of the alignment revealed that they were chimeric either among separate loci or among separate species. Similarly to the *Adh* data a few sequences where found that were easily recognized as recombinants of two loci from the same species. This pattern was found in one clone of *Acaena cylindristachya*, which was a recombinant between locus A and B and one clone of *Acaena* splendens #2324 that is a recombinant of clone A and C. These sequences were experimentally cut in two based on sequence identity/similarity to the two parental loci. In most cases this was straightforward, but, the exact location of the recombination between the two strands could not be determined when a portion of the sequence perfectly matched both “parents”. In that case the sequence was cut arbitrarily at the midpoint between the closest characters suggesting contradictory

affinities. Phylogenetic analysis based on this experimental data set placed the two segments of each recombinant unambiguously with their respective "parental" sequence. Whether these sequences represent true genomic recombinants or PCR artifacts is unclear. Interestingly, *A. cylindristachya* was also found to have a recombinant *Adh* sequence, but to postulate that ITS and *Adh* were involved in the same recombination event is highly speculative. In addition, non-recombinant copies of *A. cylindristachya* were found for both ITS loci suggesting that if the recombination is real, locus one and three must be heterozygous.

The interspecific recombinants were less difficult to interpret and their chimeric sequence patterns clearly suggested hybridization. The sequences were cleaved the same way as the inter-locus recombinants. A phylogenetic analysis based on the experimental data set again placed the two segments of each putative hybrid unambiguously with their respective putative parental lineage (or closest living relatives thereof). The following interspecific ITS recombinants were found: *A.* x *anserovina* (parental lineages= *Acaena* section *Ancistrum* x section *Euacaena*), *A.* x *splendens* No. 8227 (parental lineages= *Acaena splendens* x *Acaena* section *Euacaena*), and *Acaena latebrosa* (closest living relatives of parental lineages= *Acaena* section *Axillares* x *Cliffortia*). Further evidence of hybridization in these taxa is discussed below.

In subsequent analyses recombinants were excluded from the data set. In analyses that focused on the overall gene phylogeny the cut recombinant sequences were also excluded because the large amounts of missing data was making heuristic search strategies intractable. Phylogenetic analyses in which Agrimoniinae and

recombinants were excluded showed significantly less homoplasy as measured by consistency indeces. Within the putatively functional clade, relationships are very similar to those found in *trnL/F*. *Poterium* is sister to *Poteridium* and the two together form a clade sister to *Verruchaena* (all other Sanguisorbinae). *Cliffortia* is monophyletic and sister to *Amentomorpha*. The relationship among different clades of *Acaena* and their connection to *Polylepis* is ambiguous. In some analyses *Polylepis* is sister to all of *Acaena*, in others it is sister to *Acaena cylindristachya,* similar to the *Adh1S.A* phylogeny. Two of the *Acaena* clades found in *Adh* and *trnL/F* are again recovered, namely Ancistrum and Euacaena (Figure III.7). Argentum, on the other hand is paraphyletic with respect to Ancistrum and *A. montana* joins *A. splendens* and *A. digitata.* The relationship between the basal clades of *A. magellanica, A.elongata* and the *Margyricarpus-Tetraglochin* is unresolved. As with previous data, there was no resolution within the functional *Polylepis* clade. The higher rate of sequence evolution within locus B.I-II and C.I-II is of little help due to what appears to be interlocus recombination, possibly accompanied by random lineage sorting and interspecific hybridization (see Chapter VI).

In addition to the total analysis of all ITS loci, an analysis was also conducted based only on putatively functional sequences (Figure III.8 and III.9). To reduce the data set this way may compromise the results if the copies that have retained functionality are not orthologous, i.e., pseudogenization has occurred independently in paralogous copies. Consequently the results of this analysis have to be viewed with caution. The reason for reducing the data set was to minimize the confounding effect

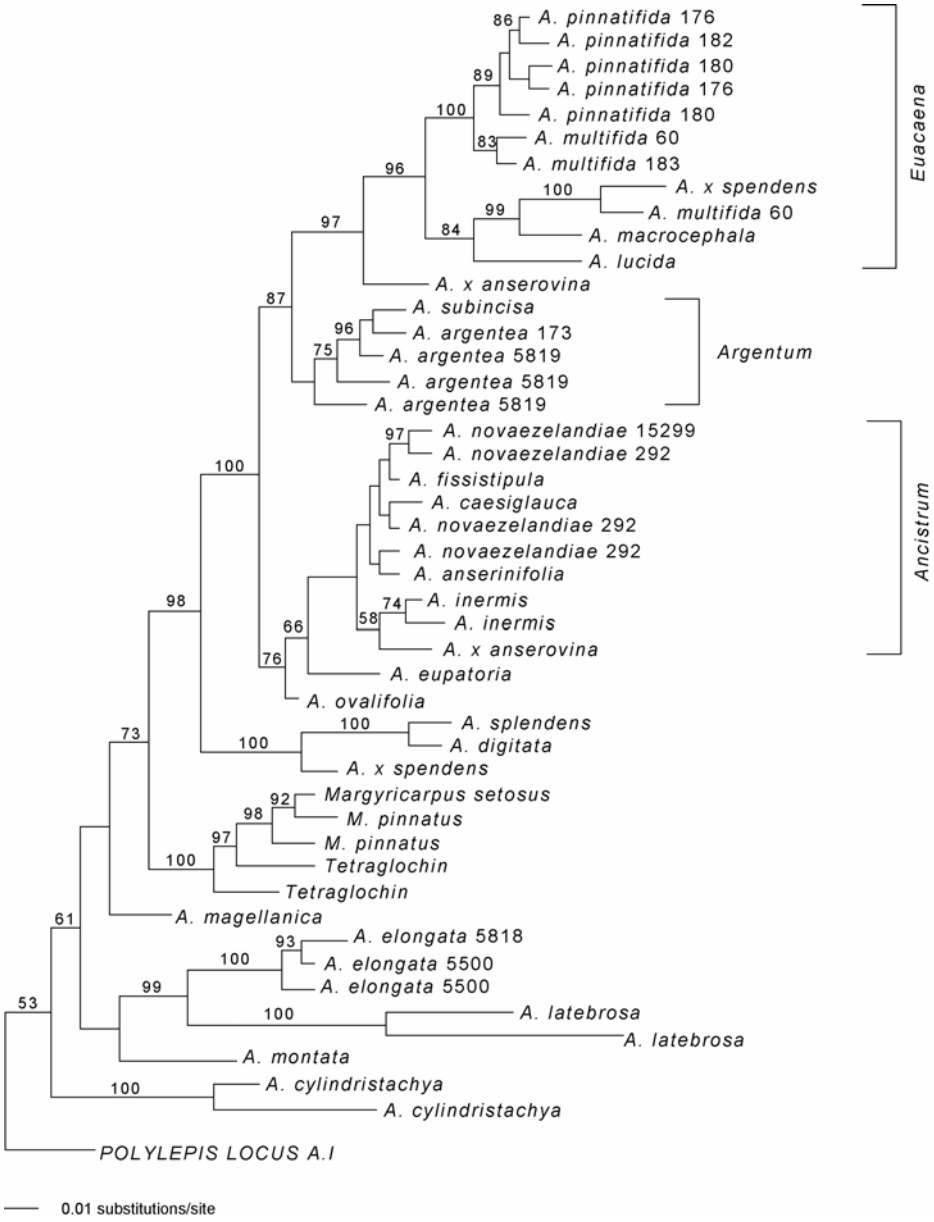


Figure III.7. Phylogenetic relationships among sub-clades of *Acaena, Margyricarpus*, *Tetraglochin* and *Polylepis*, based on ITS locus A.I.

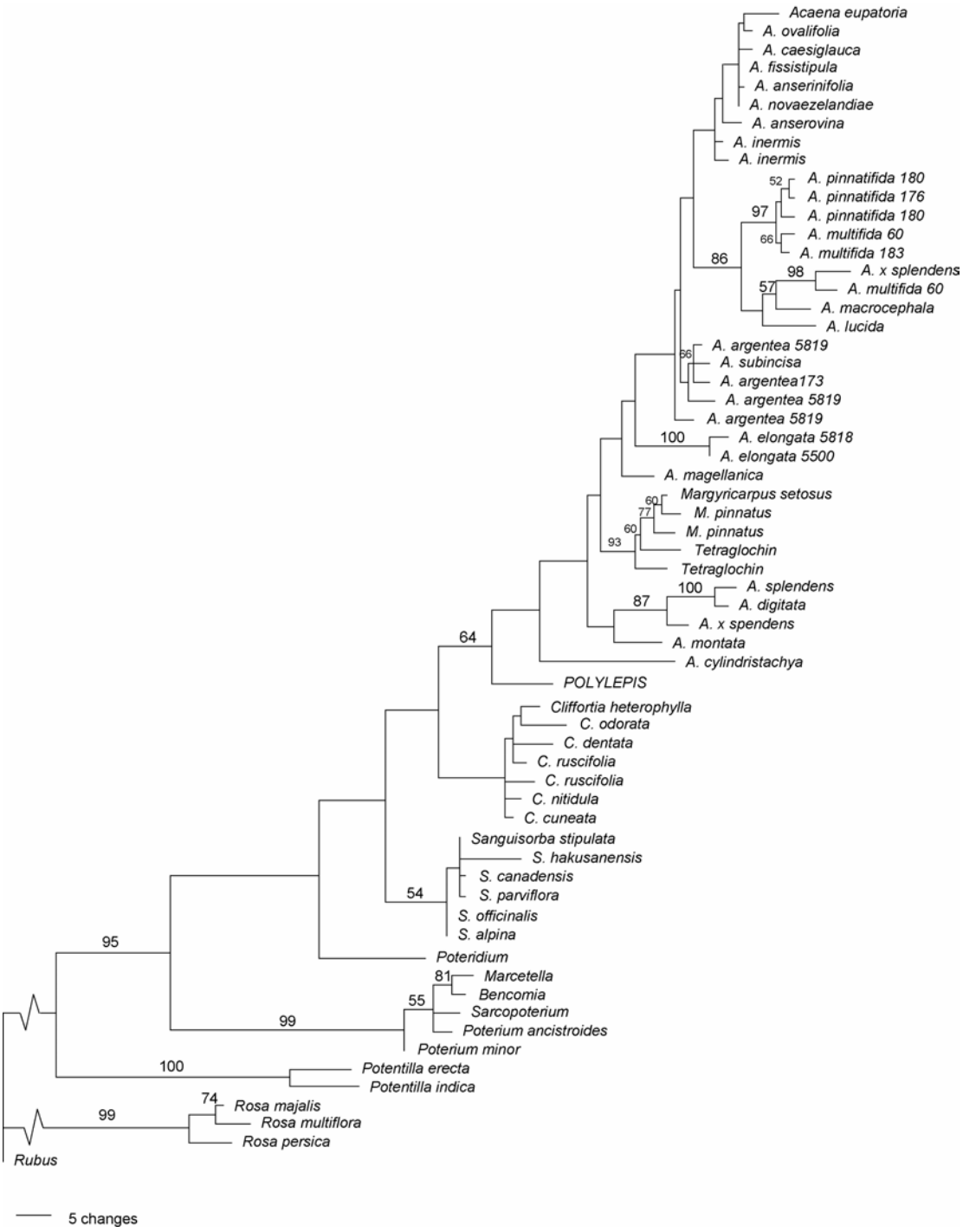


Figure III.8. Gene tree of functional ITS sequences in Sanguisorbeae based on maximum parsimony. Bootstrap values indicated above branches. Putative chimeric (hybrid) Agrimoniinae sequences excluded from the analysis. *Polylepis* clade reduced to a single representative. Orthology of sequences presumed but not confirmed.

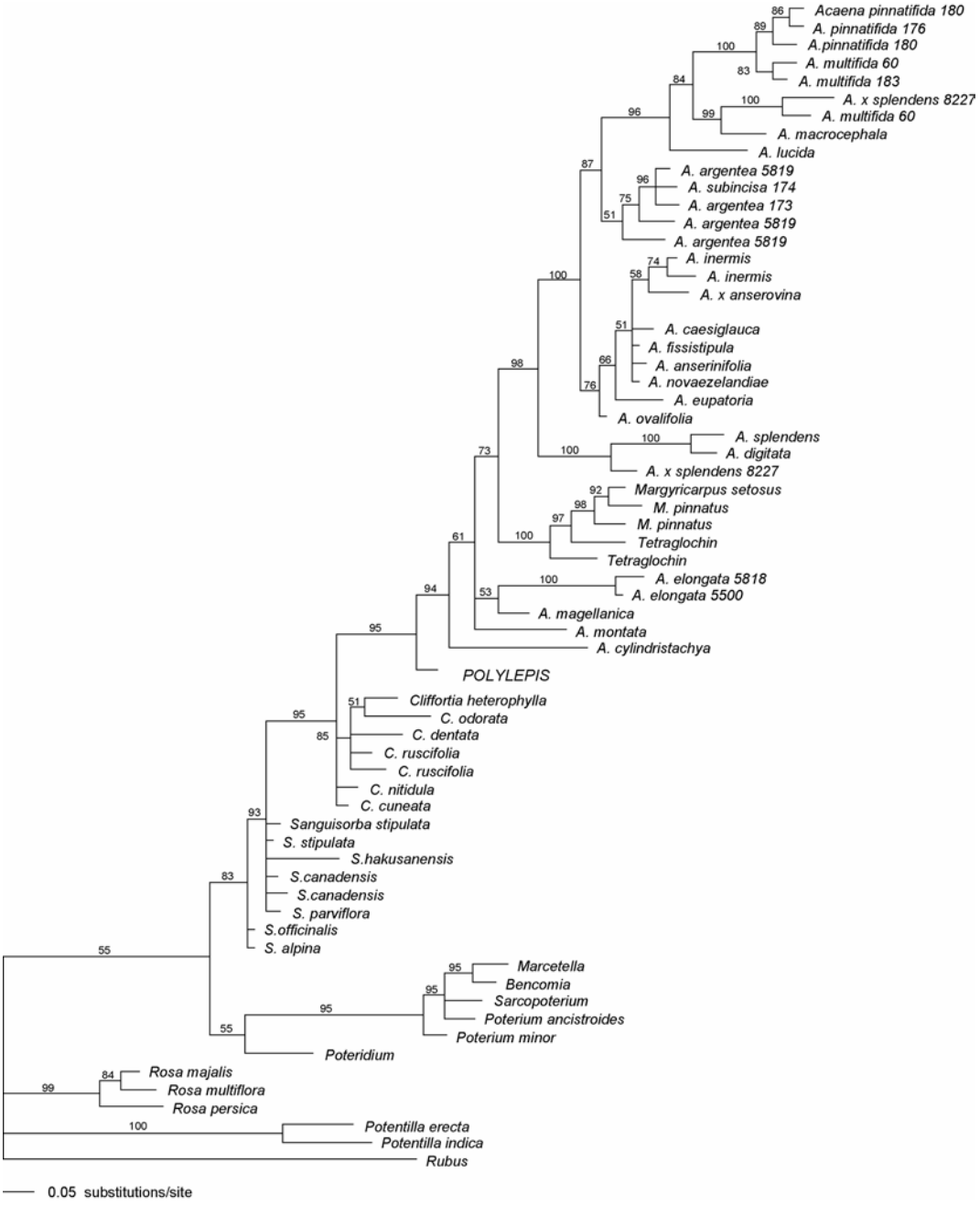


Figure III.9. Gene tree of functional ITS sequences in Sanguisorbeae based on Bayesian inference under a GTR+I+Γ model of sequence evolution. Posterior probabilities indicated above branches. Orthology of sequences presumed but not confirmed. Putative chimeric (hybrid) Agrimoniinae sequences excluded from the analysis. *Polylepis* clade reduced to a single representative.

that some highly divergent sequences were having on the phylogenetic reconstruction.

Discussion

Sanguisorbeae's sister group - Roseae or Potentilleae

The relationship among the tribes Sanguisorbeae, Roseae and Potentilleae has remained elusive despite the accumulation of molecular data (Eriksson *et al.*, 1998; Eriksson *et al.*, 2003; Mishima *et al.*, 2002; Morgan *et al.*, 1994). Chloroplast data either support Sanguisorbeae as sister group to Roseae to the exclusion of Potentilleae (*matK*) or Roseae and Potentilleae as sister taxa with Sanguisorbeae being more distantly related (*rbcL* and *trnL/F*, although weakly supported). The ITS data are ambiguous depending on phylogenetic optimality criterion and taxon sampling within Sanguisorbeae. Parsimony analyses tended to favor a relationship between Sanguisorbeae and Potentilleae while Bayesian analyses supported a closer relationship with Rosae, similar to *matK*. Amplifications of *Adh1* in *Rosa* were unsuccessful and consequently this gene is not informative in this case. The fact that three uniparentally inherited genes of varying evolutionary rates have difficulty resolving this node suggests that the internal branch, wherever it lies, must be "short" and that the two divergence events giving rise to the three lineages probably occurred within a relatively short geological time span. The ambiguities in the nuclear ITS data, however, are difficult to explain by rapid radiation and may suggest that a hybridization event involving Potentilleae and an ancient Sanguisorbeae lineage has obscured the data.

Although the evidence is far from conclusive the scenario that appears to best fit the current data is one in which the tetraploid Agrimoniinae originated from an allopolyploid hybridization event between an ancestral Potentilleae and a diploid ancestor of either *Poterium*, *Poteridium* or both (see figure III.10). The "missing link" required to solve this problem may be the Chinese species *Spenceria*, the only diploid member of Agrimoniinae. Fresh material was unavailable and several attempts to extract from herbarium specimens failed. Because the affinities of *Spenceria* are unknown, there are several possible scenarios. The morphology of *Spenceria* is in some respects intermediate between Potentilleae and Agrimoniinae and it is possible, although it has never been suggested, that *Spenceria* arose from within Potentilleae. The hypothesis presented in Figure III.10 illustrates an allopolyploid hybridization event between an ancestral Potentilleae and a diploid early member of Sanguisorbinae (possibly *Poteridium*) giving rise to the tetraploid Agrimoniinae. If the apetalous Sanguisorbinae (including *Poteridium*) are sister to Roseae, this scenario agrees with the *matK* phylogeny and suggests that the ambiguous ITS data from Agrimoniinae result from recombination between synologous ITS loci. Given that *matK* represents the maternal line, the pollen donor must have been Potentilleae. An additional line of evidence comes from floral morphology. While the Agrimoniinae have an epicalyx similar to Potentilleae, the Sanguisorbinae and Roseae lack such a structure. For this reason Weimarck (1934) suggested that Sanguisorbeae may be polyphyletic but the same arguments can be used to argue for a hybrid origin of Agrimoniinae.

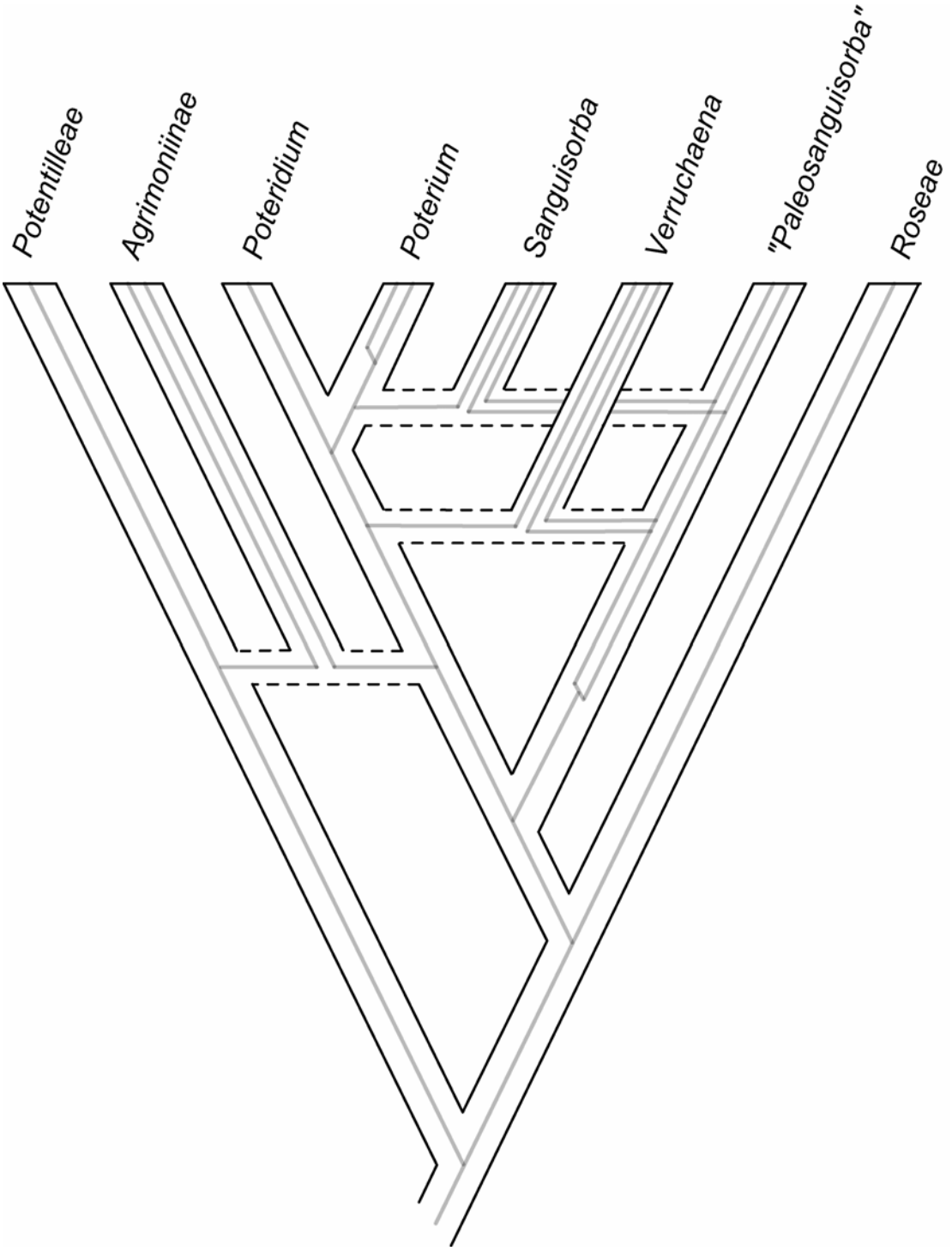


Figure III.10. Hypothesis on ancient hybridizations in Sanguisorbeae giving rise to the complex genealogical patterns found in *Adh* and ITS. "Pipes" represent organismal phylogenetic relationships and internal lines represent relationships among nuclear genes/diploid genomes. Dotted lines represent parental lineages involved in hybridization.

Allopolyploid hybrid origin of Sanguisorba and Verruchaena

While ITS provided insights into the origin of Agrimoniinae, the *Adh* phylogeny appears to be the key to resolving the origin of *Sanguisorba* and *Verruchaena.* The only extant diploid within Sanguisorbinae is *Poteridium*. Although the diploid number suggest that *Poteridium* represents a basal lineage in Sanguisorbinae, an aneuploid reduction of the chromosome number was suggested by Mishima (2002) and cannot be ruled out. The tetraploidy of the *Poterium* clade appears to have arisen through ancient autopolyploidy since the two *Poterium* clades are sister groups.

The most striking anomaly in the *Adh* tree is the occurrence of three distantly related loci in *Sanguisorba sensu stricta*, one of which is sister to *Poterium* with very strong support. It appears that the only way to explain this is to invoke a hybrid origin of *Sanguisorba*. Given the strongly supported *trnL/F* phylogeny we know that the maternal parent of this hybrid could not have been from the *Poterium* clade. Thus *Poterium sensu lato*, or rather a diploid ancestor thereof, must have been the paternal contributor. *Sanguisorba sensu stricto* departs morphologically from other Sanguisorbeae by having six-colporate pollen, all other taxa being tricolporate. An association between the doubling of colpi and hybridity (or at least polyploidy) is certainly plausible although it has never been proposed.

The *trnL/F* phylogeny reveals that the six-colporate species (here referred to as *Sanguisorba sensu stricto*) are related to the morphologically dissimilar tricolporate *Sanguisorba filiformis* (syn. *Poterium filiformis*). Hybridization between an ancestral "paleosanguisorba", possibly similar to *S. filiformis*, and an ancestral

Poterium could have given rise to the six-colporate clade (Figure III.10). The tetraploid chromosome count of *Sanguisorba* suggests that one of the *Adh1* loci originated through duplication unrelated to genome doubling. This is also supported by the fact that only two ITS loci have been found. The scenario that best fits the *Adh* phylogeny (taking into account also *Sanguisorba's* relationship with *Verruchaena*) predicts that a gene duplication without ploidy increase occurred in the ancestral *S. filiformis* (or other maternal parent). Although a recent chromosome count for *S. filiformis* was tetraploid (Mishima *et al.*, 2002), its ancestor could have been diploid or a set of chromosomes could have been lost in *Sanguisorba* through aneuploid reduction.

The relationship among the three *Adh* clades of *Verruchaena* suggests yet another allopolyploid hybridization event. Several scenarios are possible and no single hypothesis can be unambiguously established. The hypothesis that appears to best explain the current data involves an early member of *Sanguisorba* ("paleosanguisorba") and an ancestral *Poteridium* (Figure III.10). The support for basal nodes in the ITS phylogeny is too low to either support or refute the hypotheses based on the *Adh* data. However, the ITS data may solve another problematic case: *Acaena latebrosa.* This sole South African *Acaena* is morphologically quite unlike other members of the genus. The *trnL/F* phylogeny suggested that it might be basal to all of *Acaena* and *Polylepis*. The *Adh1L* copy was sister to the *Cliffortia* clade and there were several divergent, ambiguously placed copies of *Adh1S.B.* The two ITS copies of *A. latebrosa* appeared to belong to the functional clade A but there was little support for any exact placement. A closer inspection of the sequences revealed that

some sections matched closely with basal members of *Acaena* (series *Axillares* Bitter) while other segments were nearly identical to sequences of *Cliffortia*. The boundaries were less clear than in the previous cases of intra- and interspecific chimeras but an attempt was made to experimentally cleave the sequences. In an analysis using the cut sequences one part was closely related to *Cliffortia* and the other part was sister to *Acaena elongata* (Figure III.11). It could be argued that this pattern of mixed sequence motifs indicates that *A. latebrosa* is a basal *Verruchaena* lineage sharing plesiomorphic characters with both *Cliffortia* and *Acaena*. However, the fact that phylogenetic analysis fails to recognize this pattern perhaps suggests that a hybrid origin between an ancestral *Cliffortia* and an ancestral *Acaena* "*Axillares"* better explains the data. Another possibility is that hybridization was not involved in the origin of *A. latebrosa* but that later gene exchange with *A. elongata* led to introgression of certain parts of the foreign sequence. The biogeographic consequences of a hybrid origin of *A. latebrosa* are discussed in chapter IV. The origin of *Polylepis* is discussed in chapter V.

The genealogies of *Adh* and ITS suggest a fixation of heterozygocity in allopolyploid taxa followed by limited incidence of recombination among homeologous loci (or quite frequent among some copies). A genomic region that exhibits perfect concerted evolution in a diploid organism may not do so in a polyploid. The recombination responsible for concerted evolution typically occurs during meiosis when the chromosomes line up in homologous pairs. If the polyploid genome has become "diploidized", i.e., has disomic inheritance, homeologous chromosomes will act independently and have limited opportunity to exchange

genetic material. For this reason it is generally thought that crossing-over is rare between non-homologous chromosomes and among dispersed (i.e., not tandemly repeated) gene families. However, this may be less so for genes that are located near the telomeres. FISH imaging has clearly demonstrated the telomeric position of the 18S-5.8S-26S repeats in *Sanguisorbeae* (Mishima *et al.*, 2002).

Hybrid species origins

Unlike the ancient hybridizations among the main clades of Sanguisorbeae, a few cases of interspecific hybridization could be established with relatively high confidence. The species *Acaena* x *anserovina* was originally described as a putative hybrid between *A. anserinifolia* and *A. echinata,* based on an apparent intermediate morphology (Orchard, 1969). Two ITS copies of locus A were amplified from *A.* x *anserovina*. One was embedded within *Ancistrum* (close to *A. anserinifolia*) and the other appeared to be a recombinant between the putative *Ancistrum* homeolog and a *Euacaena* homoelog. When the chimeric sequence was cleaved, analysis placed the *Ancistrum* segment with the unrecombined copy and the *Euacaena* segment next to *A. multifida* in *Euacaena* (Figure III.11). Although ITS was not amplified from *A. echinata, trnL/F* data suggest that *A. multifida* is a close relative. Thus a morphological hypothesis of hybridization was confirmed by molecular data. Based on morphological observations (Orchard, 1969) hybridization between these lineages may be a recurring phenomenon giving rise to a "polyphyletic" species *A.* x *anserovina*.

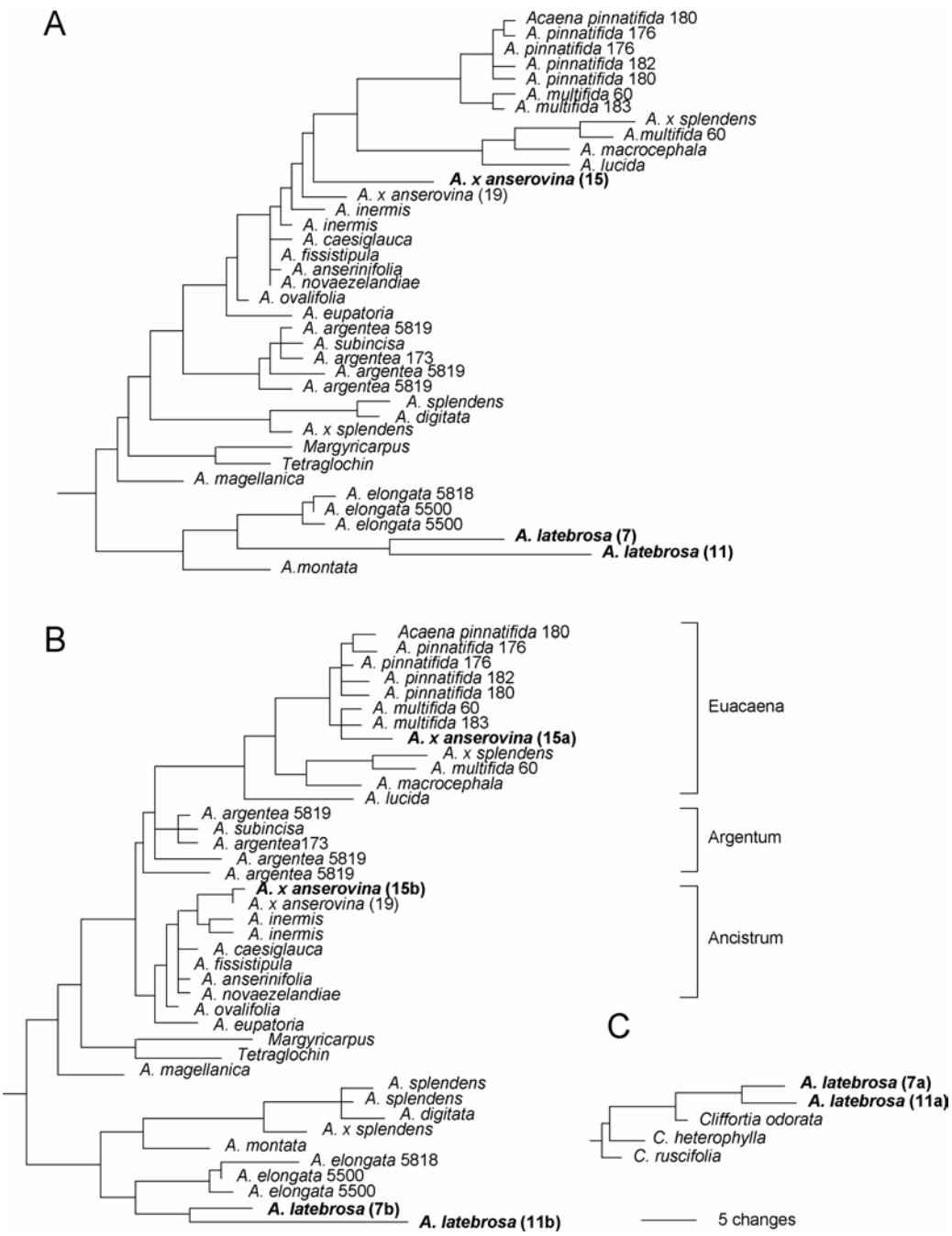


Figure III.11. Phylogenetic position of putative chimeric (hybrid) ITS sequences from *Acaena x anserovina* (clone 15) and *A. latebrosa* (clones 7 and 11). (A) Maximum parsimony analysis with all sequences intact. (B-C) Maximum parsimony analysis after putative hybrid sequences had been cleaved into two segments.

A second case of interspecific hybridization was found in *Acaena* x *splendens* (accession #8227). This species was amplified from herbarium material that had been originally identified as *A. splendens* by the collector. Because of its close resemblance to *A. pinnatifida*, the specimen was conditionally re-labelled as *Acaena* x *splendens.* The ITS data suggested that locus A was heterozygous and phylogenetic analysis confirmed one copy (or allele) of *Euacaena* type (most closely related to *A. multifida*) and a second copy that appeared as sister to *A. splendens* + *A. digitata.* Because chromosome counts are not available for either *A.* x *anserovina* or the specimen of *Acaena* x *splendens*, it is unclear whether these hybridization events were associated with polyploidy.

Chapter IV The biogeographic history of Sanguisorbinae

Introduction

The apetalous members of the Sanguisorbeae, Sanguisorbinae, are disjunctly distributed between the temperate zones of the northern and southern hemispheres. The southern members of the group are in turn split between South America, Africa, and Australasia. The low molecular diversity within the group (see chapters II, III), together with limited fossil evidence, suggests that the Sanguisorbinae represent a relatively young lineage. Consequently, the current disjunct distribution pattern can probably not be explained by tectonic vicariance events involved in the break-up of Gondwana, but must be explained by long-distance dispersal, continental migrations or a combination of both. Molecular sequence data provide opportunities to assess biogeographical hypotheses, especially when combined with age estimates of clades. Phylogenetic trees can be used to optimize the ancestral distribution of a lineage before its diversification, and to determine the direction and chronology of the subsequent radiation.

Gene phylogenies of *trnL/F*, *Adh1* and ITS were used as backbones to explore the biogeographic history of the Sanguisorbeae. Three questions in particular were addressed: (1) where did Sanguisorbinae originate? (2) where did *Verruchaena* originate? and (3) how many trans-oceanic disjunctions have to be invoked between Australasia and South America to explain the phylogenetic and geographic relationships within *Acaena*? The origin of the Andean genus *Polylepis* will be discussed separately in chapter V.

There are now several analytical methods available for estimating divergence times in the absence of a molecular clock, e.g., nonparametric rate smoothing (NPRS) (Sanderson, 1997), penalized likelihood (PL) (Sanderson, 2002), and parametric Bayesian approaches (Thorne *et al.*, 1998) among others. All methods rely on some form of a priori knowledge of the age of one or more clades in the phylogeny of interest. This prior knowledge, which usually comes from the fossil record, is necessary to calibrate the analysis. Fortunately, well-preserved flower and fruit fragments of a rosaceous species (*Paleorosa similkameenensis*) have been documented from Eocene. Based on a combination of morphological characteristics this fossil species appears to represent a rather basal lineage in the family (Basinger, 1976; Cevalloz-Ferriz *et al.*, 1993). In addition to *Paleorosa*, a limited number of more recent Rosaceae fossils have been found. Within Sanguisorbeae approximate fossil dates are available for *Sanguisorba* (Menke, 1976), *Polylepis* (Van der Hammen 1974, 1997) and *Cliffortia* (Scott *et al.*, 1995).

Berry (1919, 1922, 1939) documented what he thought were leaves of *Polylepis* from sediments deposited on the Bolivian altiplano. These formations were later determined by radiometric dating to be almost 11 million years old (Gregory-Wodzicki *et al.*, 1998). The elevation of this area was considerably lower at that time (1160 ± 600 m) and it has been estimated that the climate was 10-13C warmer than today (Gregory-Wodzicki, 2002; Gregory-Wodzicki *et al.*, 1998). It is difficult to imagine *Polylepis* growing in this kind of environment. What is thought to be one of the more ancient species, *P. multijuga*, grows in the montane forest rather than the paramo, but it still does not grow below 2500m. Judging from published photos, the

small leaflets are similar to those of extant *Polylepis* species (although not to *P. multijuga*) but could also be the leaves of *Acaena*, other rosaceous lineages, or even a non-rosaceous plant.

Palynological data seem to suggest a much younger age of *Polylepis*. The identification is complicated by the fact that the pollen grains of *Acaena* and *Polylepis* are identical. Only relative dates are available, but it appears that pollen of *Polylepis*/*Acaena* type started appearing in South America at least 2.5 million years ago. Van der Hammen *et al.* (1997; 1973) found *Polylepis/Acaena* pollen in what they called "biozone IV" of the Colombian Eastern Cordillera. They estimated this zone to be 3.2-2.4 million years old. Judging from the large quantity of pollen recorded, it appears more likely that it was deposited by an ancestral *Polylepis* rather than *Acaena*. (Extant *Polylepis* species can form rather dense forest stands while herbaceous or suffrutescent *Acaena* rarely become dominant elements of the vegetation.)

Estimates of divergence times are more accurate the more calibration points that can be included, and if more than one molecular tree is used. Using several data sets, of course, requires that the topologies are congruent among gene trees. Because of the hybrid history of Sanguisorbeae, molecular trees from different genomes or even within duplicated genes show different topologies and cannot be combined. In this analysis dates were estimated using the PL method based on the *trnL/F* and *Adh* trees independently and the results compared.

Materials and Methods

Biogeographic analyses were based on a previously generated *trnL/F* phylogeny (chapter II). For details on extraction, amplification, and sequencing see chapter II. Possible ancestral areas for different clades were explored with dispersal/vicariance analysis as implemented in DIVA (Ronquist, 1996). This free software reconstructs ancestral distributions without taking area relationships into account. Optimal solutions are those that explain a given distribution pattern by minimizing the number of dispersal events. Speciation by vicariance or by sympatric speciation costs nothing while dispersal and extinction each have a cost of one. DIVA tends to estimate widespread ancestral areas, especially at deeper nodes. The optimizations were primarily used to pinpoint the occurrence and number of geographic disjunctions between the South American and Australasian *Acaena*.

Two new data sets of *trnL/F* and ITS sequences of Rosaceae were compiled for the dating analysis. In order to use the fossil species *Paleorosa* as a calibration point, the phylogenetic analysis had to be expanded to include a wide sampling outside of subfamily Rosoideae (table IV.1). Cevalloz-Ferriz *et al.* (1993) demonstrated that *Paleorosa* possessed a combination of characteristics now associated with members of Spiroideae and Maloideae respectively. They argued that the fossil species therefore must have belonged to a lineage older than the split between Spiroideae and Maloideae. Several previously published sequences of Spiroideae, Maloideae and Amygdaloideae representing divergent clades (based on previous analyses of Rosaceae (Morgan *et al.*, 1994; Potter *et al.*, 2002) data) were

Table IV.1. Species used in analyses of divergence times in Sanguisorbeae based on *trnL/F* and ITS. Letters in parentheses indicate herbaria where vouchers are deposited.

Species	Gene	Source/voucher information	Geographic origin
Acaena x anserovina Orch.	ITS	D. E. Symon 15310 (MO)	Australia
Acaena argentea R. & P.	both	M. Hibbs 173 (MARY)	Chile
Acaena argentea R. & P.	*trnL/F*	J. Clark 5819 (MARY)	Ecuador
Acaena caesiglauca Bergmans	*trnL/F*	van Balgooy 4329 (MO)	New Zealand
Acaena cylindristachya R. & P.	both	M. Hibbs 167 (MARY)	Bolivia
Acaena digitata Phil.	both	M. Hibbs 181 (MARY)	Chile
Acaena echinata Nees.	*trnL/F*	D. E. Symon 13386	Australia
Acaena elongata L.	both	J. Clark 5500 (MARY)	Ecuador
Acaena eupatoria Cham. &Schltdl.	ITS	R. C. Molon *et al.* 12175 (US)	Brazil
Acaena fissistipula Bitt.	*trnL/F*	M. Hibbs 65 (MARY)	New Zealand (Cult. in RBGE)
Acaena inermis Hook.f.	both	M. Hibbs 57 (MARY)	New Zealand (Cult. in RBGE)
Acaena latebrosa Aiton.	*trnL/F*	C. M. Whitehouse 122	South Africa
Acaena macrocephala Poepp.	ITS	M. Hibbs 63 (MARY)	Chile, Argentina (Cult. in RBGE)
Acaena magellanica (Lam.) Vahl	both	M. Hibbs 61 (MARY)	Falkland Islands (Cult. in RBGE)
Acaena masafuerana Bitt.	*trnL/F*	G. Kuschel 215 (US)	Juan Fernandez Islands
Acaena montana Hook.	both	M. Hibbs 58 (MARY)	Tasmania (Cult. in RBGE)
Acaena multifida Hook.f.	ITS	M. Hibbs 60 (MARY)	Chile, Argentina (Cult. in RBGE)
Acaena multifida Hook.f.	both	M. Hibbs 183 (MARY)	Chile
Acaena novae-zelandiae Kirk	ITS	P. & M. Hibbs 292 (MARY)	Tasmania
Acaena ovalifolia R. & P.	both	M. Hibbs 175 (MARY)	Chile
Acaena pinnatifida R. & P. *ssp. grandiflora* Bitter	*trnL/F*	M. Hibbs 180 (MARY)	Chile
Acaena pinnatifida R. & P.	*trnL/F*	M. Hibbs 182 (MARY)	Chile
Acaena pinnatifida R. & P. *ssp. grandiflora* Bitter	ITS	M. Hibbs 176 (MARY)	Chile
Acaena splendens Gilles ex H. & A.	both	S. Teillier *et al.* 2324 (MO)	
Acaena subincisa Wedd.	ITS	M. Hibbs 174 (MARY)	Chile
Acaena x splendens	ITS	L. R. Landrum *et al.* 8227 (MO)	Chile
Adenostoma fasciculatum[2] Hook. & Arn.	*trnL/F*	Potter *et al.* 2002	USA

Species	Gene	Source/voucher information	Geographic origin
Agrimonia eupatoria[1] L.	*trnL/F*	Eriksson *et al.* 2003	Sweden
Agrimonia parviflora Sol.	*trnL/F*	M. Hibbs	USA
Aremonia agrimonioides Necker ex Nestler	*trnL/F*	T. Eriksson 147	unknown
Bencomia caudata (Ait.) Webb & Berth	both	M. Vretblad 12 (MARY)	Canary Islands
Ceanothus americanus L.	ITS	Hardig *et al.* 2000	USA
Cliffortia burmeana Burtt Davy	*trnL/F*	O. M. Hilliard *et al.* 14681 (E)	South Africa
Cliffortia dentata Willd.	both	C. M. Whitehouse 34	South Africa
Cliffortia graminea L.	*trnL/F*	Woodwine 66 (US)	South Africa
Cliffortia heterophylla Weim.	both	J. Vlok, B-E. *et al.* 81 (MO)	South Africa
Cliffortia odorata L.f.	both	C. M. Whitehouse 71	South Africa
Cliffortia ruscifolia L.	both	C. M. Whitehouse 72	South Africa
Cliffortia sericea E. & Z.	*trnL/F*	unknown	South Africa
Exochorda racemosa[2] (Lindl.) Rehd.	*trnL/F*	Potter *et al.* 2002	China
Fallugia paradoxa[2] (D. Don) Endl. ex Torr.	ITS	Smedmark *et al.* 2003	USA
Fallugia paradoxa[2] (D. Don) Endl. ex Torr.	*trnL/F*	Potter *et al.* 2002	USA
Filipendula vulgaris[1] L.	*trnL/F*	Eriksson *et al.* 2003	Sweden
Fragaria nilgerrensis[2] Schltdl.	ITS	Potter *et al.* 2000	
Fragaria vesca[2] L.	ITS	Potter *et al.* 2000	USA
Gillenia stipulata[2] (Muhl. ex Willd.) Baillon	ITS	Bortiri *et al.* (unpublished)	USA
Gillenia trifoliata[2] (L.) Moench	*trnL/F*	Potter *et al.* 2002	USA
Hagenia abyssinica J.Gmelin	*trnL/F*	Knox 2532 (GH)	Kenya
Leucosidea sericea E & Z.	*trnL/F*	F. B. Wright 1522 (E)	South Africa
Marcetella moquiniana (Webb. & Berth.) Svent.	both	M. Vretblad 11 (MARY)	Canary Islands
Margyricarpus pinnatus (Lam.) Kuntze	ITS	M. Hibbs 66 (MARY)	Andes (Cult. in RBGE)
Margyricarpus pinnatus (Lam.) Kuntze	*trnL/F*	M. Hibbs 184 (MARY)	Chile
Margyricarpus setosus R. & P.	ITS	M. Hibbs 136 (MARY, LPB)	Bolivia
Polylepis hieronymi Pilger	*trnL/F*	M. Hibbs 133 (MARY)	Bolivia
Polylepis incana H.B.K.	ITS	J. Clark 4991 (MARY)	Ecuador
Polylepis lanuginosa H.B.K.	*trnL/F*	J. Clark 6227 (MARY)	Ecuador

Species	Gene	Source/voucher information	Geographic origin
Polylepis multijuga Pilger	*trnL/F*	D. Llatas Quiroz 27495 (MO)	Peru
Polylepis quadrijuga Bitter	*trnL/F*	E.G.B. Kieft *et al.* 143 (NY)	Colombia
Polylepis quadrijuga Bitter	both	S. R. Gradstein	Colombia
Polylepis reticulata Hier.	*trnL/F*	J. Clark 4992 (MARY)	Ecuador
Polylepis sericea Wedd.	ITS	J. Clark 5820 (MARY)	Ecuador
Polylepis tarapacana Philippi	ITS	M. Hibbs 163 (MARY)	Bolivia
Potentilla anserina[1] L.	*trnL/F*	Eriksson *et al.* 2003	Sweden
Potentilla anserina[1] L.	ITS	Eriksson *et al.* 2003	Sweden
Potentilla arguta[1] Pursh.	ITS	Eriksson *et al.* 1998	USA
Potentilla bifurca[1] L.	ITS	Eriksson *et al.* 1998	Sweden
Potentilla erecta[1] L.	ITS	Eriksson *et al.* 2003	Sweden
Potentilla indica Wolf.	both	M. Hibbs 69 (MARY)	USA (introduced)
Potentilla stenophylla[1] Diels	ITS	Eriksson *et al.* 2003	China
Prunus laurocerasus[2] L.	*trnL/F*	Potter *et al.* 2002	USA (UC Davis Arb.)
Prunus padus[2] L.	ITS	Bortiri *et al.* 2001	
Rhamnus californica[2] Eschsch.	*trnL/F*	Potter *et al.* 2002	USA (UC Davis Arb.)
Rhodotypos scandens[2] (Thunb.) Mak.	*trnL/F*	Potter *et al.* 2002	(Berkeley Bot. Gard.)
Rosa majalis[1] Herrm.	ITS	Eriksson *et al.* 1998	Sweden
Rosa majalis[1] Herrm.	*trnL/F*	Eriksson *et al.* 2003	Sweden
Rosa multiflora Thunb. ex Murray	both	M. Hibbs 78	USA
Rosa persica[1] Michx.	ITS	Eriksson *et al.* 2003	Iran or Afghanistan (Uppsala Bot. Gard.)
Rubus chamaemorus[1] L.	*trnL/F*	Eriksson *et al.* 2003	Sweden
Rubus chamaemorus[2] L.	ITS	Alice *et al.* 1999	Sweden
Rubus idaeus[2] L.	ITS	Alice *et al.* 2001	Sweden
Sanguisorba alpina Bunge in Ledeb.	*trnL/F*	M. Hibbs (MARY)	Mongolia (MOBOT)
Sanguisorba annua (Nutt. ex Hook.) Torr. & Gray	*trnL/F*	E. Earle 4412 (NH)	USA
Sanguisorba annua (Nutt. ex Hook.) Torr. & Gray	ITS	B. Maguire *et al.* 26587 (PH)	USA
Sanguisorba canadensis L.	both	S. Woodbury 152 (MO)	USA
Sanguisorba filiformis (Hook. F.) Hand.-Mazz.	*trnL/F*	J. F. Rock 17963	China

Species	Gene	Source/voucher information	Geographic origin
Sanguisorba hakusanensis Mak.	both	M. Hibbs 59 (MARY)	Japan or Korea (RBGE)
Sanguisorba minor Scop.	ITS	M. Hibbs 62 (MARY)	Mediterranean (RBGE)
Sanguisorba minor Scop.	*trnL/F*	M. Hibbs 178 (MARY)	Chile (introduced)
Sanguisorba officinalis L.	both	T. Eriksson 144 (SBT)	Sweden
Sanguisorba stipulata Rafin.	*trnL/F*	K. Dillman 17 (NH)	Alaska
Sanguisorba tenuifolia Fisch. ex Link	*trnL/F*	X. Ling 81906 (MO)	China
Sarcopoterium spinosum (L.) Spach	ITS	M. Hibbs 16 (MARY)	Mediterranean
Sarcopoterium spinosum[2] (L.) Spach	*trnL/F*	Helfgott *et al.* 2000	Mediterranean
Sorbus aucuparia[2] L.	ITS	Campbell *et al.* 1995	
Sorbus californica[2] E. Greene	*trnL/F*	Potter *et al.* 2002	USA
Sorbus torminalis[2] (L.) Crantz	ITS	Robinson *et al.* 2001	
Spiraea cantonensis[2] Lour.	ITS	Bortiri *et al.* 2001	
Spiraea densiflora[2] Torr. & A. Gray	*trnL/F*	Potter *et al.* 2002	USA
Tetraglochin cristatum (Britt.) Rothm.	ITS	M. Hibbs 150 (MARY)	Bolivia
Waldsteinia geoides[2] Willd.	ITS	Smedmark *et al.* 2003	Unknown (Cult. Stockh. Univ.)

included in the phylogenetic and PL analyses. *Paleorosa* was assumed to represent the most recent common ancestor of Spiroideae and Maloideae. The deposits where *Paleorosa* was found have been dated by potassium-argon ratios, which probably makes this fossil a more reliable calibration point than most other Rosaceae fossils that have been dated by relative dating techniques. Because of the great age discrepancy between purported *Polylepis* macrofossils and palynological data, a wide range of 2.5-11 mya was set as the calibration age of *Polylepis.* Fossils and dates used in the dating analyses are listed in table IV.2. The ITS data set of Sanguisorbeae were reduced to functional gene copies with the assumption that these sequences are orthologous to the monomorphic ITS copies of the outgroups. The Agrimoniinae were excluded from the ITS analysis because of their potential hybrid origin and the chimeric nature of their sequences. *Acaena latebrosa* was likewise excluded for the same reason. *Rhamnus californica* Eschsch. and *Ceanothus americanus* L. (Rhamnaceae) were used to root the *trnL/F* and ITS phylogenies respectively.

Sequences were aligned by eye. Two regions of ITS could not be unambiguously aligned across Rosaceae and were therefore excluded from the analysis. Phylogenetic analyses were performed using Bayesian inference. PL analysis requires a phylogram with non-zero branch lengths. Because of the low interspecific variation within *Polylepis* and parts of *Acaena*, some of these taxa had to be excluded from the *trnL/F* analysis. Bayesian analyses were performed based on two million generations of Markov chain Monte Carlo, sampling every 100 trees. Trees sampled before the chain reached stationarity were excluded from the "allcompat" Bayesian consensus tree.

Table IV.2. Fossils used as calibration points for estimating divergence times in Sanguisorbeae.

Clade constrained	Fossil name (kind)	Published fossil date	Constrained date	Reference
Cliffortia	*Cliffortia sp.* (pollen)	late Miocene-Pliocene	max. 11.2 - min. 1.8 mya	Scott, Steenkamp *et al.* 1995
Maloideae+Spiroideae	*Paleorosa similkameenensis* (macro)	44.25 mya	fixed 44.25 mya	Basinger 1976; Cevalloz-Ferriz *et al.* 1993
Polylepis	*Polylepis sp.* (pollen)	min. 3.2-2.2 mya	min. 2.2 mya	Van der Hammen 1974, 1997
	Polylepis sp. (macro)	Max. 10.7 mya	max. 11 mya	Berry 1919, 1922, 1939; Gregory-Wodzicki *et al.* 1998
Sanguisorba sensu stricto	*Sanguisorba officinalis* (pollen?)	Pliocene	max. 5.3 - min. 1.8 mya	Menke 1976

The likelihood ratio test was used to investigate rate heterogeneity in the *trnL/F* and ITS phylogenies. Log likelihoods were compared under a molecular clock model and an unconstrained model. A significant difference of log likelihoods was detected between the unconstrained and constrained analyses and consequently a molecular clock was rejected. Assuming non-clocklike molecular evolution, divergence dates were estimated using the semiparametric PL method as implemented in the free unix-based software r8s (Sanderson, 2003a). The PL method allows for different substitution rates on each branch of the tree but combines this with a roughness penalty that discourages abrupt rate changes. The optimality criterion is the log likelihood score minus the roughness penalty (Sanderson, 2003b). The contribution of the roughness penalty is determined by a smoothing parameter.

A cross-validation procedure implemented in r8s was used to optimize the smoothing parameter for the *trnL/F* and ITS phylogeny, respectively, and the dating analyses were run using the TN algorithm. *Paleorosa* was used to fix the age of the clade uniting Maloideae and Spiroideae. Because analyses on the origin of *Polylepis* suggest that *P. quadrijuga* has retained the "original" *Polylepis* chloroplast (see chapter V) the divergence of this species was assigned the fossil calibration date for *Polylepis* in the *trnL/F* phylogeny. Two non-identical accessions of *P. quadrijuga* were included allowing for a non-zero branch uniting this "clade". However, these accessions may not represent the entire diversity of this clade and may therefore underestimate the age of *Polylepis*. Penalized likelihood output files were inspected and manipulated in PAUP* (Swofford, 2002).

Results

The DIVA analysis predictably estimated wide ancestral distributions for all deeper nodes. The ancestor of *Verruchaena*, for example, was estimated to have been distributed across South America, South Africa and Australasia. This is probably unlikely. Nevertheless, the analysis was useful in pinpointing the occurrence of disjunctions between South America and Australasia. The minimum number of dispersal events required to explain this disjunction is two or three depending on the phylogenetic position of *Acaena montana.*

The Bayesian *trnL/F* phylogeny based on the extended data set agreed in large part with previous analyses (chapter II). The main difference was the placement of *Acaena latebrosa* within *Cliffortia*, as opposed to basal to *Verruchaena* suggested by the previous analyses. The ITS phylogeny was congruent with previous analyses with the exception that there was no synapomorphic support for *Cliffortia*, and there was an overall reduction in branch lengths in Sanguisorbeae, which was expected given the exclusion of hypervariable regions. The relationships among the outgroups differed markedly between the *trnL/F* and ITS phylogenies. *trnL/F* data suggested that the taxa sampled from outside of subfamily Rosoideae (i.e., *Gillenia, Sorbus, Spiraea, Adenostoma, Prunus, Exochorda* and *Rhodotypos*) formed a single clade sister to Rosoideae. ITS on the other hand placed Amygdaloideae (*Prunus*) and most Spiroideae (*Stephanandra, Neillia, Physocarpus, Lyonothamnus, Rhodotypos* and *Sorbaria*) sister to Rosoideae to the exclusion of Maloideae (*Sorbus*) and other Spiroideae (*Spiraea, Gillenia*). This discrepancy had a significant effect on the placement of the calibration date of *Paleorosa*. According to the *trnL/F* analysis this

fossil represent the most recent common ancestor of the sister clade of Rosoideae. The ITS data, on the other hand, suggest that *Paleorosa* was the ancestor of all of Rosaceae. This incongruity had a significant effect on the age estimates of older nodes but appeared to have less severe consequences for age estimates of more recent taxa. The results of a cross-validation test as implemented in r8s suggested that the optimal smoothing parameter was approximately *S*=10 for *trnL/F* as well as for ITS.

The divergence estimates resulting from the dating analyses are presented in Figures IV.1 and 2. Taxa outside of Rosoideae that were used for calibration purposes were pruned from the trees. Branches with posterior probabilities of less than 0.5 were collapsed. The PL analysis of *trnL/F*, using a constraint of 2.2-11 mya for *Polylepis quadrijuga*, suggested an origin of Sanguisorbeae in early Miocene, or about 22 mya. The extant lineages of Agrimoniinae and Sanguisorbinae diverged soon thereafter. The ancestor of *Poteridium* and *Poterium sensu lato* appears to have diverged approximately 16 mya, while the most recent common ancestor of *Verruchaena* existed at about 10 mya (late Miocene). The origin of the maternal lineage of the putative hybrid *Sanguisorba sensu stricto* is estimated to be 3.6 million years old (mid Pliocene) and *Amentomorpha* is estimated to have originated around 8.8 mya.

Because the Agrimoniinae were excluded from the ITS data set, an overall age of Sanguisorbeae is missing from this analysis. The age estimates of the apetalous Sanguisorbinae as a whole (14.4 mya) and of *Poterium* (3.4 mya) are significantly younger than those estimated by the chloroplast data. *Sanguisorba sensu stricto*, on the other hand, appears to be much older (11.1 mya). The ITS estimated age of

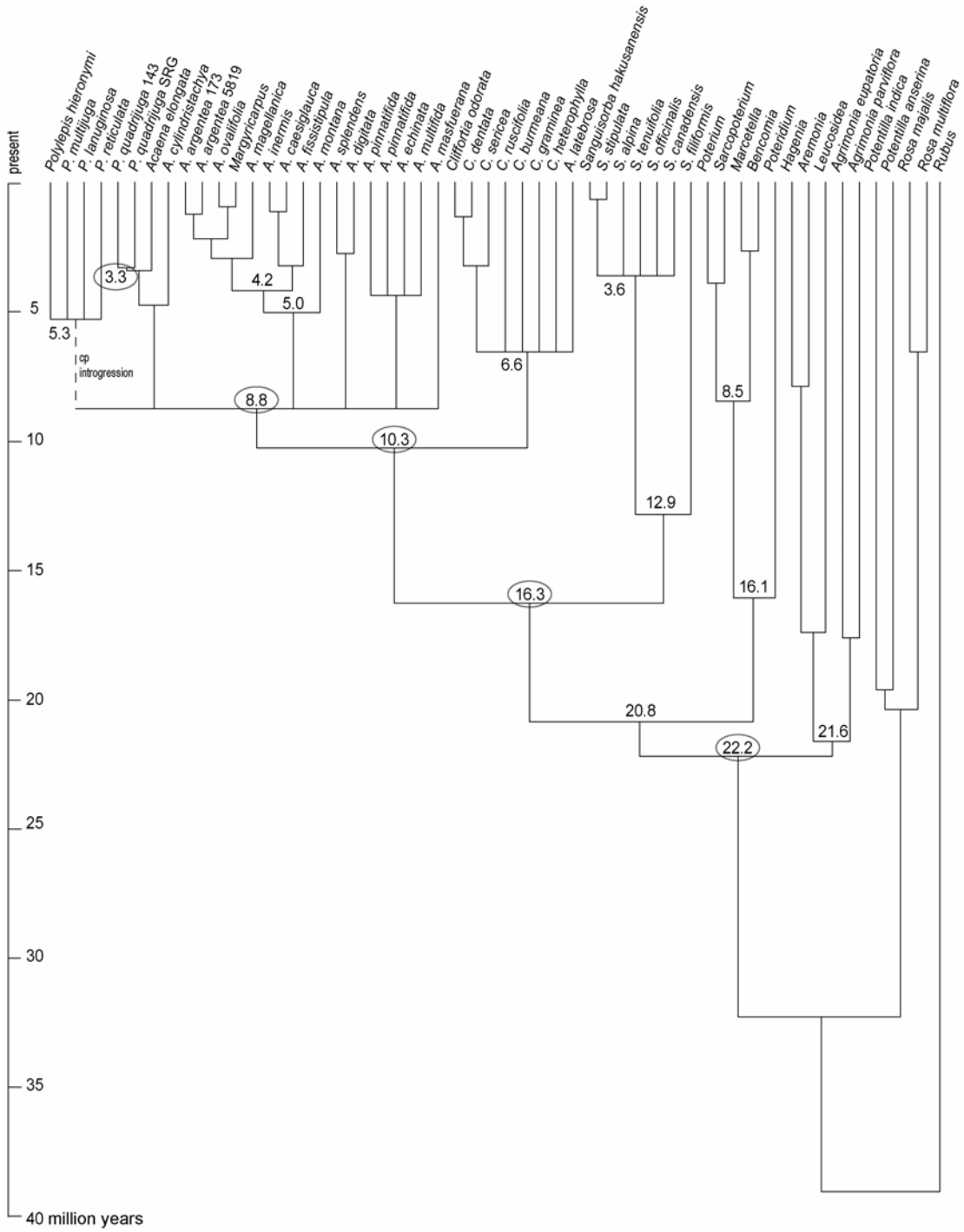


Figure IV.1. Chronogram of Sanguisorbeae based on *trnL/F* sequences using penalized likelihood (PL) analysis.

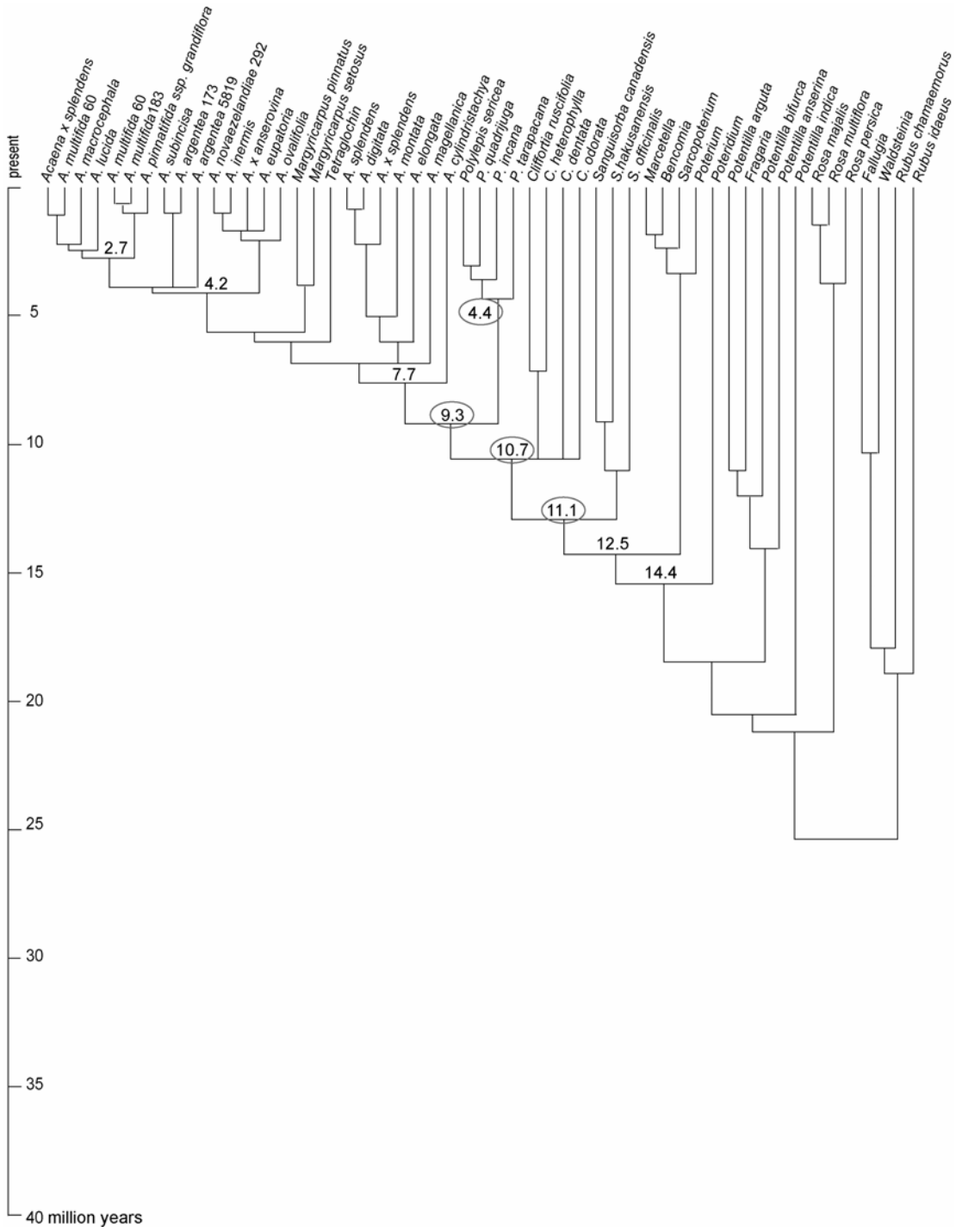


Figure IV.2. Chronogram of Sanguisorbeae based on ITS sequences using penalized likelihood (PL) analysis.

Verruchaena (10.7 mya) is remarkably similar to that estimated by *trnL/F* (10.3 mya). The origin of *Amentomorpha* and *Polylepis* was estimated at 7.7. and 4.4 mya, respectively.

Discussion

Bitter (1911a) recognized what he argued were two fundamentally different groups within the genus *Acaena.* To *Acaena* Axillares he counted those species with axillary inflorescences and fruits typically entirely covered with spines. The second group, *Acaena* Terminalis, was characterized by terminal inflorescences and apical spines. In his monograph on *Cliffortia*, Weimarck (1934) argued that *Acaena* was the basal stock from which other Sanguisorbinae had evolved. In his view, *Poterium sensu lato* (incl. *Poteridium* and *Sanguisorba*) had originated from within *Acaena* Terminales, while *Polylepis*, *Margyricarpus*, *Tetraglochin*, *Bencomia* and *Cliffortia* had sprung from *Acaena* Axillares. Weimarck also argued that the origin of the group occurred at a time when the two continents of Africa and South America were by some means connected. The continental drift theory was highly controversial at the time and was competing with hypotheses of ancient land bridges, so the exact nature of the connection was unclear to Weimarck.

In contrast with Weimarck's hypothesis, the current data show that northern hemispheric Sanguisorbinae represent a basal paraphyletic grade with respect to *Acaena* and not the other way around. It is also clear from the dating analyses that *Verruchaena* is too young to have been around before the break-up of Gondwana and the disjunct distribution must be explained by migration or long-distance dispersal of

taxa. However, the transoceanic distances may not have been as great as they are now. In addition to a larger and warmer Antarctic landmass during Tertiary, a series of islands in the southern Atlantic Ocean may have "linked" southwest Africa with southern South America long after the split up of Gondwana.

Because of the propensity for hybridization in Sanguisorbeae, divergence times estimates based on molecular data could be misleading. The parental lineages involved in a hybridization event must obviously coexist in time and space, and molecular analyses of uniparentally inherited genes could be powerful tools in tracing and dating ancient hybridization events. However, nuclear data such as ITS may have undergone recombination between parental copies which could have a confounding effect on phylogenetic as well as dating analyses.

Origin and dispersal of Verruchaena

Based on molecular phylogenetic support for a monophyletic origin of this group, the southern hemispheric Sanguisorbinae were previously given taxonomic status as *Verruchaena* (see chapter II). One of the most intriguing questions addressed in this study is the biogeographic origin of this clade. In other words, which came first, the South African *Cliffortia* or the South American-Australasian *Amentomorpha*? Based on the phylogenetic data alone, the ancestral distribution of *Verruchaena* cannot be established. Phylogenetic data from chloroplast as well as nuclear data suggest an early split of this lineage, giving rise to *Amentomorpha* and *Cliffortia* respectively. Based on both dating analyses this occurred in the late Miocene. At this time Africa was already well separated from South America but

islands may have persisted in the Atlantic possibly aiding as stepping stones in transoceanic dispersals.

However, neither the phylogeny nor the timing of the origin of *Verruchaena* tells us unambiguously where it took place. The current distributions of *Verruchaena*'s sister lineages *Poteridium*, *Poterium* and *Sanguisorba* seem to favor an origin of *Verruchaena* in America. While *Cliffortia* is currently completely isolated from other members of Sanguisorbinae, *Acaena* reaches all the way up to California where it shares habitat with the diploid *Poteridium*. Based on a hybridization hypothesis presented in chapter III an ancestral *Poteridium* could have been the paternal lineage involved in the allopolyploid origin of *Verruchaena*. Although the center of diversity of *Amentomorpha* is currently in South America an origin of this lineage in North America is not impossible. Perhaps changing climatic conditions forced the ancestral lineage southward until it reached southern South America where it a) colonized a wide range of South American habitats and b) dispersed across the Atlantic to South Africa. Assuming an origin of *Verruchaena* in South America requires that the "jump" to South Africa happened very early on in the evolution of this lineage since there are no extant *Acaena*-like lineages older than the divergence of *Cliffortia* (although some species of *Acaena* remain to be sampled in the phylogenetic analyses) (see Figure IV.3).

Another consideration is that the current distribution of Sanguisorbinae in Africa may represent a significant narrowing of a former more widespread element. The African climate has changed significantly the last few million years and paleoclimatological evidence suggests that the highly endemic fynbos biota, now

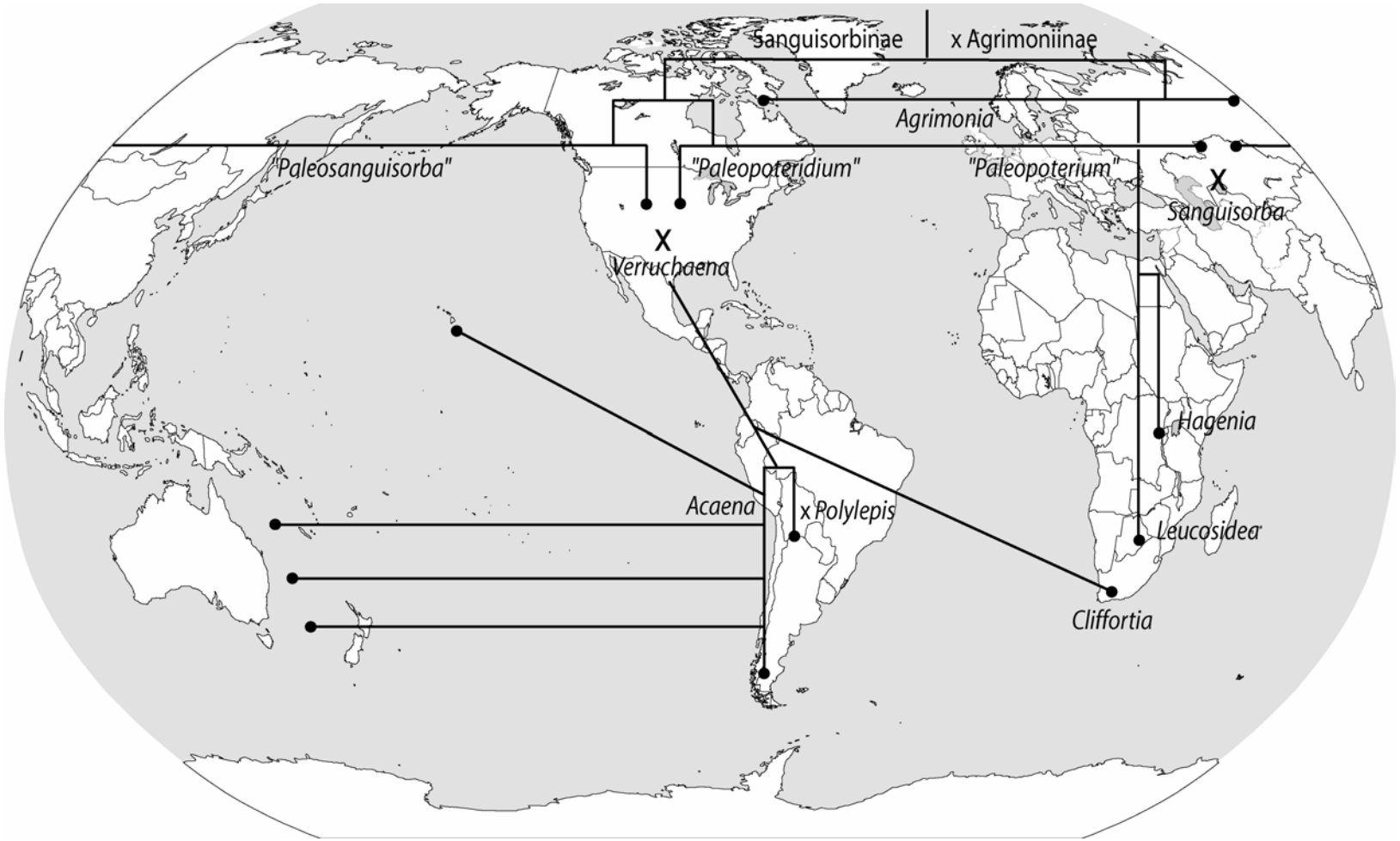


Figure IV.3. Hypothesized biogeographic history of Sanguisorbeae including ancient hybridization giving rise to *Sanguisorba sensu stricto* and *Verruchaena*.

restricted to the Cape peninsula, was much more widespread earlier in Quaternary (Davis *et al.*, 2002). One could imagine a scenario in which other basal *Verruchaena* lineages were present in South Africa during late Tertiary but were subsequently forced out by climatological change or were outcompeted by *Cliffortia*. Even so, given the current data, an origin of *Verruchaena* from early members of *Sanguisorba* and *Poteridium*, whether by hybridization or cladogenesis, appears unlikely to have happened in Africa. Unless other evidence emerges, the most plausible scenario is an origin in the New World followed by an early dispersal to South Africa, and subsequent spread to the subantarctic, Hawaii and Australasia.

Acaena *disjunctions- one or many colonization events to Australasia?*

The austral connection is not unique to *Acaena*. Patterns of plant distributions across the southern temperate parts of the circum-Pacific continents were noted already by Hooker (1847) and have since been documented in numerous plant and animal taxa. Thorne listed 48 genera that exhibit austral distribution patterns. Some of these disjunctions are probably old enough to represent remnants of Gondwanan distributions, e.g., *Auracaria, Nothofagus,* and *Proteaceae,* while others appear to be of more recent origin, e.g., *Aristotelia* and *Acaena*. Simpson suggested that the present fragmented distribution of *Acaena* might represent remnants of a more continuous range in the southern hemisphere. The possibility of a widespread southern temperate ancestor of *Acaena* appears unlikely given the young age of the clade. It is, however, possible that *Acaena* formerly had a greater presence in the Antarctic region than it has today.

As discussed above, the Australasian *Acaena* are probably the result of later dispersal events from one or several lineages that originated in the New World. The greatest diversity of extant *Acaena* is clearly in South America where the majority of species are found. Approximately 20 species occur in Australia and New Zealand, one (sometimes split into two) in South Africa and one (possibly extinct) in Hawaii. A few species are also known from the subantarctic islands of Georgia and Tristan de Cunha, as well as from the Juan Fernandez Islands off the coast of Chile. The ITS phylogeny suggests that the oldest lineages of the paraphyletic *Acaena* are among the South American *Axillares* (not counting the putative hybrid *A. latebrosa*).

Chloroplast as well as nuclear data suggest that the Australasian species do not form a monophyletic group but represent at least two and possibly three independent lineages, each of which is more closely related to South American allies than to each other. The *trnL/F* and ITS phylogenies show that the *A. ovina* complex (represented by *A. echinata*) is allied with the South American *Euacaena* while the *A. anserinifolia* complex form a distinct and separate clade possibly sister to one or all of the South American *Argentum*. The *Acaena ovina* complex is morphologically uniform and their origin appears to represent a single introduction from South America. Orchard (1969) argued for a recent origin of no more than 10,000 years ago. European colonization probably resulted in further spread of this group within Australia as well as introduction to New Zealand (Orchard, 1969). Current evidence is conflicting regarding the placement of the New Zealandic and Tasmanian *A. montana.* The *trnL/F* data suggest that this species is sister to *A. cylindristachya* or a larger clade of *Acaena* "*Axillares"*, while ITS supports a connection to *A. digitata*

and *A. splendens*. In either case *A. montana* appears to have evolved independently from other Australasian lineages. *A. montana* may represent the most ancient of three separate dispersal events to New Zealand and Australia. Judging from the ITS phylogeny, reversed dispersal events from Australasia back to South America did not take place.

Exceptional dispersal ability is required to traverse the vast distances separating the continents. Dispersal via high-altitude westerly jet streams can probably be ruled out based on of the relatively large size and weight of the *Acaena* achenes. The achene's viability after immersion in seawater has not been tested but a waterborne trip of that distance and duration is unlikely. The most plausible scenario, which was discussed by Carlquist (1965), is the dispersal by migrating sea birds. The barbed spines of many *Acaena* achenes have a remarkable ability to latch on to any rough surface, as anyone knows who has had the pleasure of hiking through *Acaena* vegetation. The spines have in fact been found in the down of young petrels on Juan Fernandez Islands as well as in New Zealand. Petrels are known to migrate large distances across the subantarctic region.

The epizoic dispersal mechanism of *Amentomorpha* could also have contributed to rapid colonization within terrestrial areas. *Acaena* species that bear barbed spines (as opposed to naked spines), in particular, appear to be exceptional colonizers. Lee (2001) demonstrated a correlation between fruit features and geographic distribution among New Zealandic species, with the barb-spined species having much wider geographic ranges than species with barbless spines. It is likely that the fruits that were carried with birds across the Pacific Ocean belonged to a

barbed species. Barbed spines are present in most species of *Ancistrum* and *Argentum* and in some members of *Euacaena* and the paraphyletic *Axillares* including *A. montana.* The Australasian species that lack barbs (*A. inermis* and closely related species) are phylogenetically embedded within a clade of barbed taxa suggesting a possible secondary loss.

Whether the Hawaiian endemic *Acaena exigua* is more closely related to South American or Australasian clades remains to be discovered. This species, which is native to remote alpine bogs of Maui, is greatly endangered, and may be near extinction. It was long considered to be extinct until a single plant individual was located on West Maui in 1997 (Meidell *et al.*, 1998). Bitter (1911a) placed *A. exigua* in the same section as *A. pumila,* native to the Magellan region and the sub-Antarctic islands, and *A. masafuerana*, an endemic of the alpine flora of the Juan Fernandez Islands. The *trnL/F* data suggest an early divergence of *A. masafuerana,* but without complete sampling it is impossible to establish when this species diverged from its sister lineage. The oldest of the volcanic islands in the Juan Fernandez archipelago is thought to be four million years old. It is possible that *A. masafuerana* constitutes a link between the South American-subantarctic *Acaena* and the Hawaiian *A. exigua,* although the distance between the islands is admittedly vast. An eastern origin of *A. exigua* also contradicts the prevailing view that the majority of Hawaiian plants are derived from the west and south (Carlquist, 1965). Transoceanic dispersal from New Zealand to Hawaii via Polynesian islands have been inferred in Myrtaceae (Wright *et al.*, 2000; Wright *et al.*, 2001). However, *Acaena* requires a temperate or alpine environment and is not known from any other tropical or subtropical Pacific islands.

The origin of *Acaena latebrosa* remains a mystery. Axelrod and Raven (1978) argued that *Acaena* probably arrived via long-distance dispersal to South Africa fairly recently, but they did not discuss the connection with *Cliffortia*. The current study suggests a possible hybrid origin between an ancestral *Cliffortia* and a later *Acaena* colonizer, possibly *A. elongata.*

Chapter V The biogeographic origin of *Polylepis* and the paramo flora

Introduction

The biogeographic and evolutionary history of the paramo flora is of particular relevance owing to its high degree of diversity and endemism, especially in light of the relatively young geological age of the high Andean habitats. The basis for determining the origin and composition of any biota is an understanding of the relative contributions of different source areas in providing ancestral (or extant) colonizers. In the case of tropical alpine biotas this can be restated as the relative importance of vertical versus horizontal evolution, i.e., the importance of regional lower elevation zones versus distant temperate/alpine regions as source areas. Several thorough taxonomic studies have been published related to the composition of the paramo flora, but taxonomy-based analyses are unable to unambiguously differentiate between vertical and horizontal elements. Previous surveys have, however, clearly demonstrated that the paramo flora developed through a combination of in situ evolution and immigration from cool regions elsewhere. In addition, these studies point to the combined effects of evolutionary history and ecology by highlighting the morphological convergence between taxonomically disparate groups among tropical alpine floras on different continents (e.g., Smith and Cleef, 1988). The logical next step is to apply phylogenetic methods to these problems.

At present, phylogenetic studies of paramo taxa are very few and initiating such research should be of high priority. Phylogenetic analyses of paramo clades and their closest relatives can provide detailed information of the extent (and possibly timing) of interchange between the South and North American continent. Studies of

the great American interchange have so far focused mainly on faunal exchanges and the understanding of phytogeographical connections between the two continents is still very limited. Because *Polylepis* constitutes a frequent and sometimes dominant component of paramo habitats, determining the biogeographic origin of this genus is especially relevant.

A clear understanding of the phylogenetic relationships of the Sanguisorbeae is a prerequisite for evaluating different biogeographical hypotheses with respect to the origin of *Polylepis*. Several authors have suggested that *Polylepis* is derived from within the southern hemispheric *Acaena* (Bitter, 1911a; Bitter, 1911b; Weimarck, 1934). Bramwell (1976), on the other hand, suggested a connection between *Polylepis*, the Canarian *Bencomia*-alliance and the East African *Hagenia*. The molecular phylogenetic analyses presented in this dissertation support a close association between *Polylepis* and certain members of the paraphyletic *Acaena*, but the exact relationship is ambiguous. By comparing the phylogenies from different genes and different genomes, a scenario is proposed to reconcile incongruent phylogenetic signal.

Tropical alpine environments

Tropical alpine environments are those regions within the tropics occurring between the continuous treeline and the upper limit of plant life (approx. 3500-5000m). These environments are unique and fundamentally different from temperate alpine and arctic environments. The main environmental difference lies in the periodicity of important climatic variables such as temperature and sunlight. Tropical

alpine regions go through daily fluctuations in temperature of the same magnitude as the seasonal changes in temperate or arctic environments. Frost can occur on any night of the year and is often followed by rapid surface heating during the day. Troll (1968) referred to this climate as the diurnal temperature climate. As opposed to temperate plants, tropical alpine plants cannot retreat to dormancy to escape the cold temperatures, but have to have adaptations to cope with daily environmental stresses. Furthermore tropical alpine habitats are like ecological islands in a sea of warm lowland tropical climate and vegetation, and this insularity has a profound effect on the development of these biotas.

Tropical alpine vegetation is found in Central and South America, East Africa, Malaysia and Hawaii (Monasterio and Vuilleumier, 1986). The physiognomy of tropical alpine plants varies to some extent between regions, but certain distinct growth forms are characteristic of nearly all tropical alpine biotas and appear to have evolved independently in different regions and different taxa. Such dominant growth forms are tussock grasses, tall woody erect rosette plants, acaulescent rosette plants, cushion plants and sclerophyllous shrubs (Hedberg and Hedberg, 1979; Hedberg, 1992). Although the dominant growth forms are similar between regions, the taxonomic composition and diversity of the floras are markedly different (Smith and Cleef, 1988).

The biogeographic and evolutionary origins of the unusual tropical alpine floras are particularly interesting considering the relatively recent geological origin of today's tropical mountains. Most tropical alpine environments are of late Tertiary or Quaternary age. The recent development of the tropical alpine habitats also makes it

possible to, with reasonable accuracy, track the biogeographic origin of individual elements of their biotas.

The tropical alpine zone of the northern Andes is called paramo and has been studied in some detail by Simpson (1974; 1983), Van der Hammen (1979), Cleef (1981), Hooghiemstra (1984), Van der Hammen & Cleef (Van der Hammen and Cleef, 1986), Simpson & Todzia (Simpson and Todzia, 1990), among others. The diversity and composition of the paramo biota has been analyzed and compared with other tropical alpine habitats in East Africa, Malaysia and Hawaii (Dorst and Vuilleumier, 1986; Hedberg, 1992; Mena and Balslev, 1986; Smith and Cleef, 1988). The paramo vegetation appears to be a mix of both locally adapted elements and immigrants from the Nearctic as well as the Austral-Antarctic regions (Cleef, 1979; Simpson and Todzia, 1990; Smith and Cleef, 1988; Van der Hammen, 1979). Exactly how much various source areas have contributed to today's paramo flora is difficult to estimate and remains unresolved. The origin of the paramo flora is also closely tied to the climatic and geological changes in the region during Pleistocene, a topic that will be discussed in more detain in chapter VI.

Materials and methods

Nucleotide sequence data from *trnL/F*, *Adh* loci *1S.A* and *1L,* and one ITS locus (putatively functional, see chapter III) were used to clarify the position of *Polylepis* in Sanguisorbinae and infer the ancestral distribution of the most recent common ancestor (or ancestors) of *Polylepis* and its sister group/groups. The extraction, amplification and sequencing of *Polylepis* species and putative sister

groups were described in detail in chapter II and III. Based on the previous broad analyses of Sanguisorbeae, reduced data sets were generated that focused on the origin of *Polylepis*. The ITS data set was reduced to include only functional sequences (see chapter III) and excluded all chimeric sequences (e.g., *A. latebrosa*).

Phylogenetic analyses were performed under maximum parsimony for all data sets and under maximum likelihood for *Adh1L, Adh1S.A* and ITS. Bayesian analysis was preferred over maximum likelihood for the *trnL/F* data in order to use a mixed likelihood model for nucleotides plus indels. 1000-replicate heuristic searches were performed under parsimony and likelihood and the Bayesian analysis was run for two million generations (mcmc) and "burnin" discarded. Branch support was estimated using non-parametric bootstrap analysis under parsimony and likelihood, and with posterior probabilities in the Bayesian analysis. All analyses were performed using PAUP* (Swofford, 2002) and MrBayes (Huelsenbeck and Ronquist, 2001).

Results

Previous broader analyses (chapter II and III) had already indicated that *Polylepis* is closely associated with the genus *Acaena*. The results from the reduced analyses were largely congruent with the previous findings. Because the results were very similar regardless of optimality criterion only the likelihood trees for *Adh1S.A, Adh1.L* and ITS and the Bayesian tree of *trnL/F* are presented here (Figure V.I). Chloroplast *trnL/F* data suggests that *Polylepis* is sister group to either *Acaena*

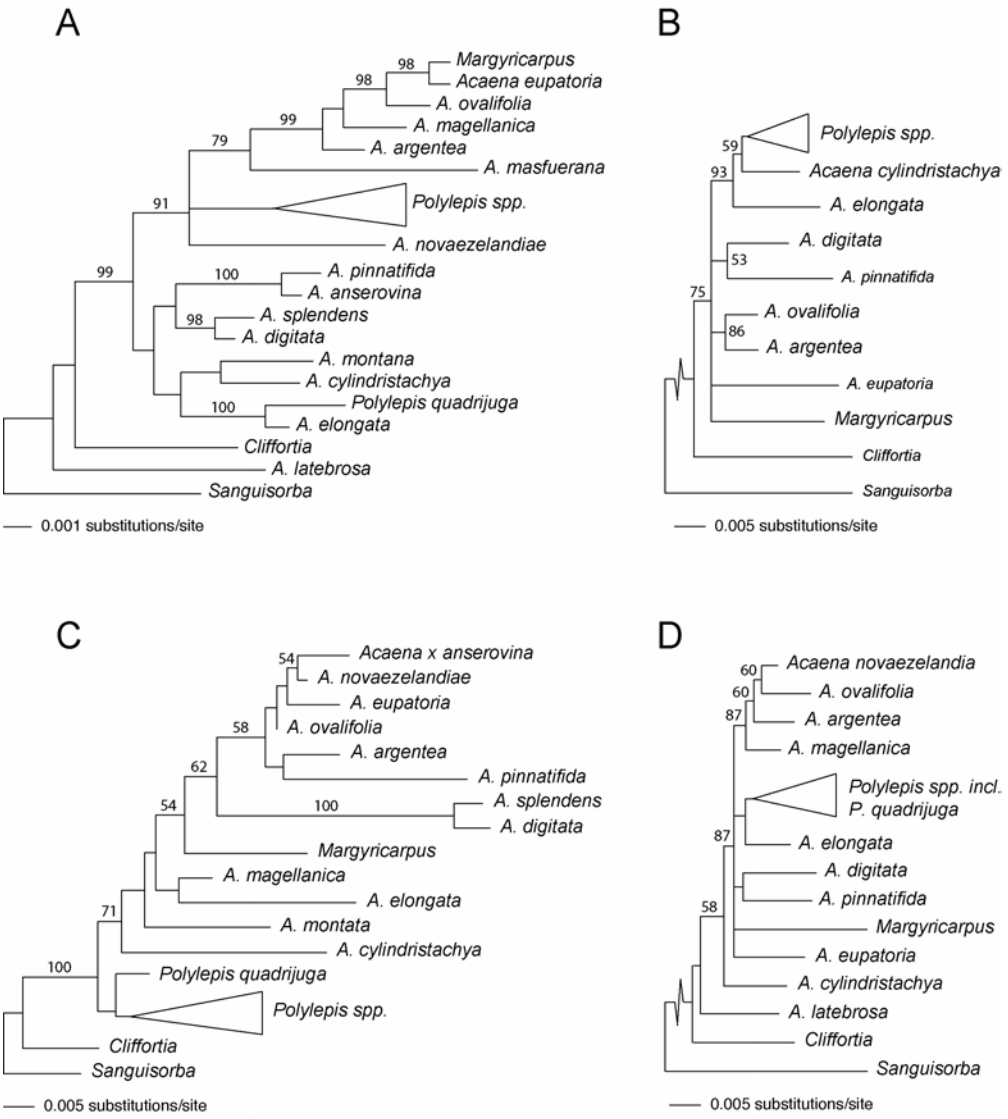


Figure V.1. Gene trees of (A) *trnL/F*, (B) *Adh1S.A*, (C) ITS, and (D) *Adh1L* representing phylogenetic hypotheses on the relationship between *Polylepis* and sister taxa. The phylogenetic analyses were performed under Maximum likelihood in all cases except in the *trnL/F* analysis, which was conducted using Bayesian inference under a mixed model (Mk for indels and GTR+I+Γ for nucleotides). Likelihood bootstrap values or posterior probabilities exceeding 50% are indicated above the branches.

novaezelandiae (representing section *Ancistrum sensu stricto*), *Acaena* section *Argentum* (included in *Ancistrum* by Bitter (1911a)), or a clade of both. One species, *P. quadrijuga*, is separated from other *Polylepis* and is sister to *A. elongata* with 100% posterior probability. The ITS data suggest an early split between *Polylepis* and *Acaena*, and consequently a monophyletic origin of *Acaena* and *Margyricarpus* (excluding *A. latebrosa*) (71% ML bootstrap). However, when all the ITS sequences are analyzed together under parsimony, *Polylepis* is sister to *A. cylindristachya* and the two together are sister to the rest of *Acaena* and *Margyicarpus* (not shown). *Polylepis* and *A. cylindristachya* as sister taxa is also supported by ITS pseudogene B (see Figure III.6). Unlike the *trnL/F* data, the ITS phylogeny indicates that *P. quadrijuga* is closely allied with the rest of *Polylepis*, but in a basal position. The two *Adh* loci present two different alternatives. *Adh1L* supports a relationship between *Polylepis* and *Acaena elongata* at the exclusion of the rest of *Acaena*. *Adh1S.A* on the other hand supports a relationship with *A. cylindristachya* although *A. elongata* is only one node removed. *P. quadrijuga* is nested within the *Polylepis* clade in *Adh1L*, while no *Adh1.SA* locus was recovered in *P. quadrijuga*.

In chapter II it was argued that the *trnL/F* sequence of *P. quadrijuga* might represent introgression from an ancestral *Acaena elongata.* With the added information from three nuclear loci this hypothesis is put on its head and instead it appears that all the rest of *Polylepis* obtained their chloroplast through an early introgression event. This is consistent with the ITS data regarding infrageneric *Polylepis* relationships showing *P. quadrijuga* as basal to a (tokogenic) group

representing the remaining *Polylepis*. Disregarding the introgression in *trnL/F* there is still disagreement as to the sister group of *Polylepis*. *Adh1L* and *trnL/F* (represented now by *P. quadrijuga*) suggest a sister relationship with *A. elongata*, *Adh1S.A* indicates a closer relationship with *A. cylindristachya* (with 59% ML bootstrap support and 79% Bayesian posterior probability) and ITS supports either a basal position of *Polylepis* (current analysis) or, like *Adh1S.A,* a sister relationship with *A. cylindristachya* (see large scale analysis, chapter III). A hybrid origin of *Polylepis* between ancestors of *A. elongata* and *A. cylindristachya* could account for the incongruence among gene trees.

Discussion

Evolutionary origin of Polylepis

Bitter (1911b) argued for a derivation of *Polylepis* from what he called *Acaena* series *Axillares*, and more specifically he discussed a link with *A. elongata,* which have multipinnate leaves and long dense racemes similar to *Polylepis*. Kessler (1995b) agreed with Bitter that *Polylepis*' closest relative was probably to be found in *Acaena* section *Elongatae*. The results of this study support this view but suggest a slightly more complicated picture. The hypothesis advocated here is illustrated in figure V.2. According to this scenario *Polylepis* as a whole originated from a hybridization event between ancestral members of the *A. elongata* and the *A. cylindristachya* lineages. The hybridization was probably allopolyploid as is suggested by a ploidy of 2n=42 (6x) for most *Acaena* (although *A. elongata* and *A.*

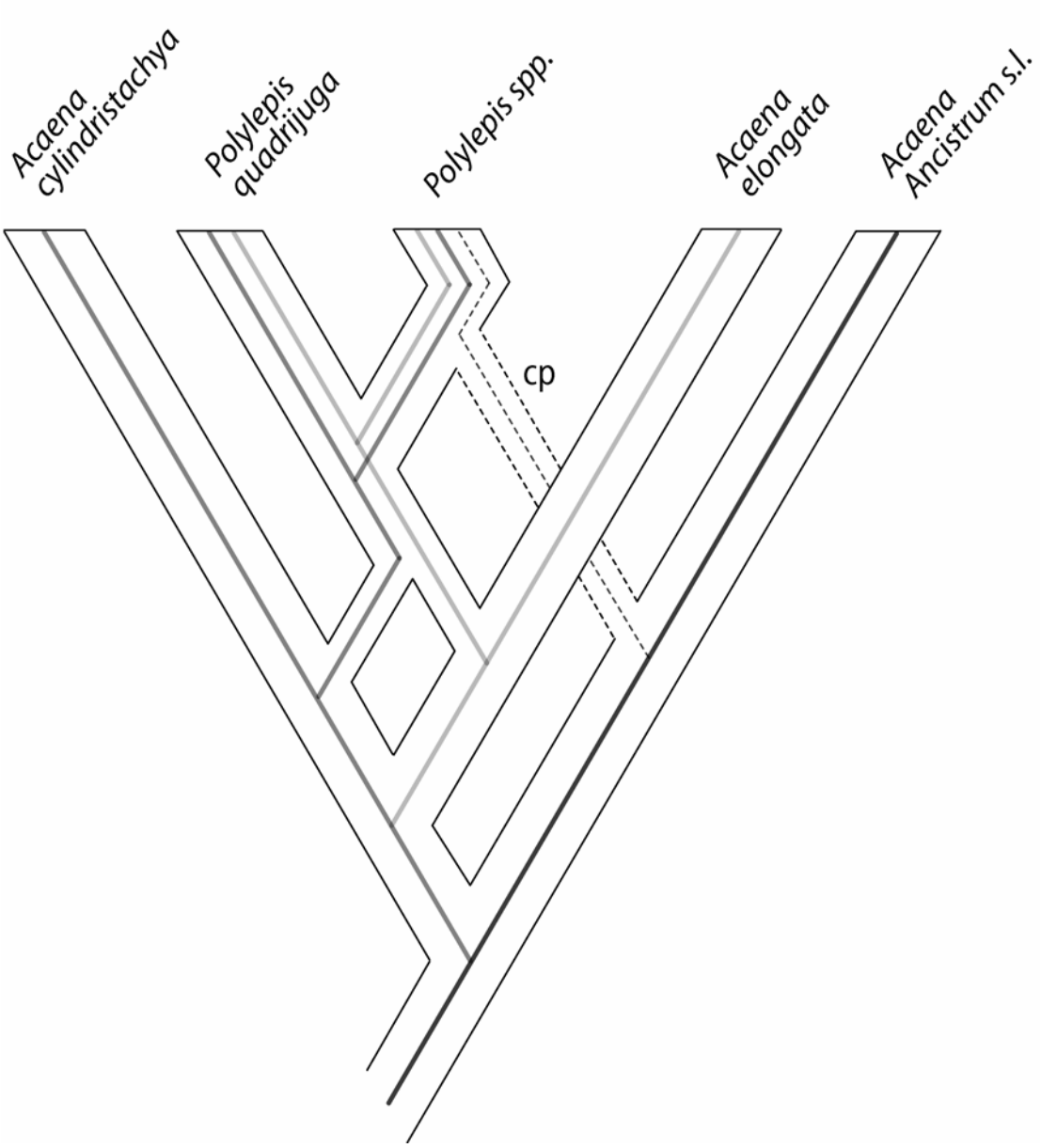


Figure V.2. Hypothesis on the origin of *Polylepis* based on reconciled molecular phylogenetic data. An initial hybridization event between ancestral lineages similar to *Acaena elongata* and *A. cylindristachya* gave rise to an ancestral *Polylepis quadrijuga*. Subsequent chloroplast introgression (dotted line) from *Acaena* sect. *Ancistrum s.l.* account for the anomalous chloroplast sequences in *Polylepis et al. exl. P. quadrijuga*.

cylindristachya have not been cytologically investigated) and a putative ploidy of 2n=84 (12x) for *Polylepis* (see chapter III). Subsequent to this event the ancestral *Polylepis quadrijuga* hybridized again with an ancestral member of *Acaena* section *Ancistrum* (*sensu lato*) giving rise to the ancestor of all remaining species of *Polylepis*. This hybridization appears to have left the original *Polylepis* nuclear genome intact but resulted in chloroplast capture (introgression) from *Ancistrum*.

A hybridization event between ancestral *Acaena elongata* and *A. cylindristachya* as the parental stock, giving rise to *Polylepis,* fits agreeably with morphological characters. Both lineages have multipinnate leaves similar to *Polylepis* and both species have robust suffrutescent stems, especially *A. elongata,* which can be characterized as a small shrub. While the mature fruit (hypanthium) of *Polylepis* is more like that of extant *A. elongata,* the dense racemose inflorescence is very similar to *A. cylindristachya* (even though it is pendulous in the former and upright in the latter). Although this is the scenario that fits the current data best, future evidence may change or add to the picture. It is in any case clear that hybridization and/or chloroplast introgression has played a vital role in the origin of *Polylepis*.

Polylepis multijuga has been considered the most ancestral and unspecialized species of *Polylepis*, and the one which shares most similarities with *Acaena* (Bitter, 1911b; Simpson, 1979). Among the postulated ancestral characters are large leaves with many leaflets, long racemes and fruits adapted for animal dispersal (Simpson, 1986). This study suggests that *P. quadrijuga* is perhaps the oldest extant member of *Polylepis*. This species is also characterized by long dense racemes and spine-covered fruits adapted for epizoic dispersal, but the leaflets of *P. quadrijuga* are not as many

or as sizeable as those of *P. multijuga.* While *P. multijuga* is restricted to the lower montane zone in northern Peru, the Colombian species *P. quadrijuga* is found at higher elevations in humid boggy paramos or along streams (Simpson, 1979). According to Simpson (1986) the habitat of *P. multijuga* has probably experienced a relatively stable climate since mid Tertiary, and the area is known for harboring a high concentration of endemic and relict elements. This suggests that *P. multijuga* may be the second most ancestral lineage among extant *Polylepis*. Interestingly, a previous multilocus ITS analysis shows that the sequence of *P. quadrijuga* is basal in a clade representing the pseudogenized ITS loci C.I, while *P. multijuga* appears to be basal in C.II (see chapter III). There was no C.I copy found in *P. multijuga* and no C.II copy in *P. quadrijuga*. It has to be emphasized however, that the pseudogene data are difficult to interpret because of what appears to be intragenomic recombination between C.I and C.II (although possibly excluding the two former species).

Biogeographic origin of Polylepis

Polylepis and the paraphyletic *Acaena* both occur in the Andes but while *Polylepis* is found mainly in the tropical alpine region, the greatest diversity of *Acaena* occurs in the temperate regions in Chile and Patagonia. The discovery of a putative hybrid origin of *Polylepis* meant that the ancestral distribution had to be optimized so that the distributions of the two potential ancestors overlap. *Acaena elongata* and *A. cylindristachya* are among the highest growing extant species of *Acaena* and have been found thriving at elevations above 4000m (although more

typically between 3000-4000m). Extant *Acaena elongata* has a wide distribution from Mexico and Central America to Ecuador, Colombia, and Peru. It is a frequent member of the paramo community and can grow at elevations rivaling that of *Polylepis quadrijuga*. *Acaena cylindristachya* occupies similar habitats between 3000-4100 meters but is restricted to Venezuela, Colombia, Peru and Bolivia. In Colombia we find a convergence of the distributions among the three extant species *A. elongata, A. elongata* and *Polylepis quadrijuga*. This fits remarkably well with a hypothesized hybridization between the former two lineages.

Pollen of the *Polylepis*/*Acaena* type is found in fossil formations dating back to 2.4-3.2 Ma from the Colombian Cordillera (Van der Hammen and Hooghiemstra, 1997) but because this pollen could have belonged to either *Polylepis* or *Acaena* it does not pinpoint the origin of *Polylepis*. However, between 2.4-2.2 Ma (biozone V (Van der Hammen and Hooghiemstra, 1997)) the same type of pollen becomes very abundant indicating that the source species is dominating the vegetation. Judging from the present habit and local distributions of species of *Acaena* versus *Polylepis* this pollen layer was most likely deposited by *Polylepis*. It is unlikely that either *Acaena* elongata or *A. cylindristachya* could have deposited this amount of pollen. The large quantitative increase in *Acaena*/*Polylepis* pollen towards the end of Pliocene in any case suggests a significant change in vegetation, which would be consistent with the origin of a new species with a more dominant presence in the protoparamo vegetation. However, changes in vegetation could also be explained by the well documented climatic changes that greatly affected the region starting around the transition from Pliocene to Pleistocene. A hypothesized origin of *Polylepis*

between 2.2-2.4 corresponds to the beginning of the first in a series of global cold spells which has characterized temperate and alpine areas during Quaternary (Van der Hammen and Hooghiemstra, 1997).

The origin of the paramo flora

The word "origin" as a phytogeographical concept is rather ambiguous and has been used in different contexts by different authors. Generally, origin refers simultaneously to both phylogenetic ancestry (or taxonomic affiliation) and biogeographic origin (i.e., source area). Naturally, when speaking of the origin of entire biotas there can be no single phylogenetic origin (at least it is difficult to imagine a community with a single common ancestor). Likewise, when referring to the biogeographic origin of biotas there is seldom a single contributing source area. However, certain source areas have typically contributed more than others to a particular community and the biogeographic origin of biotas can be expressed as the relative contribution of different source areas to the present species composition. Alternatively, the origin of biotas can be expressed as the relative importance of local evolution versus long-distance dispersal. Any given biota is likely to be a combination of species that evolved where they now occur (adaptive radiation) and those that evolved elsewhere and dispersed into these areas (adaptive shifts) (Brooks and McLennan, 1993; Ricklefs and Schluter, 1993). In the case of tropical alpine environments the floras will thus represent a combination of locally adapted tropical elements, and "pre-adapted" elements that have immigrated from different cool areas in the southern or northern temperate zones. The representation of different taxa in

the tropical alpine habitats must have been influenced by differences in the intrinsic capacity for dispersal and adaptation to the new environments, and the geographic position of these habitats on each continent (Monasterio and Vuilleumier, 1986).

In Late Jurassic, the supercontinent Pangaea was breaking up and South America rifted away from North America. Beginning in Early Cretaceous, South America also started to separate from Africa, which initiated a long period of isolation of the Neotropical continent (Storey, 1995). Fossil evidence suggests, however, that the isolation was interrupted by transient periods of limited exchange across the Proto-Antillean and the Central American archipelagoes (Estes and Baez, 1985; Vuilleumier, 1984; 1985). Towards the end of the Tertiary, tectonic events and eustatic changes formed a continuous Central American land bridge between North and South America (about 3.5 mya), and the isolation of South America was abruptly ended. The resultant waves of biotic exchange between North and South America are referred to as the Great American Interchange. The biotic exchange was also facilitated by vegetative shifts and sea level changes. During glacial maxima the lowered sea levels increased both the area and elevation of the Central American land bridge. The dry conditions (occurring from approx. two mya) caused savanna to expand toward the equator and form a continuous habitat corridor for the migration of many species adapted to these open habitats (Colinvaux, 1996; Webb, 1991).

The Plio-Pleistocene final uplift of the northern Andes made migration of northern temperate plants into the Neotropical region more direct. Many of the elements that invaded from North America have temperate requirements and remained at high elevations or high latitude southward in South America. Since

Pleistocene (or earlier), the Andes have extended along the entire length of the western continental margin. The Andes have during this time served as a corridor between northern temperate and southern temperate (Patagonian or Fuegian) elements. Many common northern temperate plant families appear to have immigrated into South America after the elevation of the Andes, such as Apiaceae, Brassicaceae, Crassulaceae, Papaveraceae, Primulaceae, and Ranunculaceae, among many others (Raven and Axelrod, 1974).

Taxonomic approaches

In phytogeographical analyses of floras, the first step generally involves defining the phytogeographical categories, often called floral elements, in which individual species (or higher taxonomic groups) are classified. Floral elements in a phytogeographical analysis of the paramo could be, for instance, the paramo element, the wide Neotropical element, the southern temperate element, etc. Because there is often a high level of endemism at the species level and because endemic taxa are not informative about the ultimate derivation of a flora, the taxonomic unit most often used is the genus. The common practice is to then compile a complete list of the genera of the flora of interest and subsequently assign each genus to one of the floral elements based on the distribution of congeneric members elsewhere (e.g., Smith & Cleef 1988), or alternatively by selecting the area that has the greatest species diversity for a particular taxon (e.g., Simpson & Todzia 1990).

The taxonomic phytogeographical method has been used in at least three separate analyses aimed at estimating the origin of the paramo flora of the Colombian

Eastern Cordillera (Cleef, 1979; Simpson and Todzia, 1990; Smith and Cleef, 1988). The discrepancy of these three studies illustrates the complexity of the problem of estimating biotic origins. Cleef (1979) divided the paramo flora into seven different geographical elements and concluded that the generic vascular paramo flora is made up of half local tropical elements and half immigrated taxa. Not unexpectedly, widely distributed genera in most cases represented the temperate component.

In a later paper, Smith and Cleef (1988) investigated the geographical affinities of the tropical alpine flora of ten different regions worldwide, of which the Colombian Eastern Cordillera was one. They concluded that tropical alpine floras in all regions are predominantly of temperate origin. Migration of cool-adapted plants from both north and south temperate regions, over long distances, appeared to account for 83% of the paramo flora. Endemic genera represented 27%, 8% were tropical, 29% northern or southern temperate, and 36% were widespread. Smith and Cleef further suggested that degrees of endemism may suggest relative ages of the floras. The high degree of endemism in the paramo flora compared to other tropical alpine areas therefore may indicate that the paramo habitat is the oldest tropical alpine environment, which also is in agreement with geological evidence (Smith and Cleef, 1988)(Smith & Cleef 1988).

Simpson and Todzia (1990) compared the diversity and phytogeographic origin of three distinct Andean floras: the paramo flora in Colombia, the puna flora in Argentina, and the austral-alpine flora in Tierra del Fuego. They recognized fewer paramo genera than the previous studies, and did not state the results in terms of local versus immigrant contribution. However, according to their table of data (p. 1430),

59% of the genera are assigned to temperate and cosmopolitan elements, while the tropical element is represented by 41% of the genera, of which 33% are Neotropical.

These studies of the paramo flora have resulted in independent contrasting conclusions. Cleef (1979) concluded that the flora is half local and half immigrated. Smith and Cleef (1988), however, state that as much as 83% of the paramo flora is of immigrated temperate origin. Simpson and Todzia's (1990) analysis is intermediate between the former studies and suggests (although not stated explicitly) that 59% of the Colombian paramo flora is of temperate immigrated origin.

The disagreement may be caused in part by a difference in methods. Cleef (1979) and Smith and Cleef (1988) designated as source area (element) the total area in which congeneric species occurred, while Simpson and Todzia (1990) selected the area that had the greatest species diversity for that particular genus. Unfortunately, both of these approaches are susceptible to error. The first approach may designate large, imprecise source areas and give only a vague indication of the biogeographic origin of a particular phylogenetic lineage. The second approach depends on the assumption that the area of highest diversity of a genus is the area of origin of that genus. There is generally no direct evidence to support this assumption. Similarly, the ultimate area of origin of a genus may not be the area of origin of a particular species/lineage within the genus. In other (phylogenetic) words, the most recent non-tropical alpine ancestor of a tropical alpine species/lineage may not have been inhabiting the same area as the common ancestor of a larger more inclusive clade (e.g., a genus). Finally, there is also a possibility that the area of origin is not within the present distribution of the genus.

The main weakness of the taxonomic approach to estimating phytogeographic origins is that it disregards phylogeny, which is essential for an accurate inference of the biogeographic history of individual clades and thus, of entire biotas. In the taxonomic analyses summarized above, it is implied that the total distribution of non-Andean congeneric species can be extrapolated to indicate the biogeographic origin of the Andean species or species-groups of the same genus. The strict use of the genus level when determining geographic affinities may seem to promote consistency, but is in fact entirely arbitrary. The rank of genus does not represent equal divergence times or equal phylogenetic affinities among congeners across taxa. Even with this aside, the use of the genus level often leads to unnecessarily imprecise estimates of biogeographic origins. The fact that a genus that is distributed in both the southern and northern hemispheres does not mean that the paramo representatives of that genus are both southern and northern in origin. The total distribution of a widespread genus provides little information regarding the origin of a single species.

Phylogenetic approaches

Biogeographic origins are inferred from recognizing the distribution of the most recent common ancestor of a particular lineage of interest. Thus, knowing the phylogenetic relationships is a fundamental requirement of biogeographic analysis. Once a phylogeny is obtained, there are theoretically two ways to infer the ancestral area of a particular clade: by using fossil evidence, or by mapping the ancestral distribution using criteria based on parsimony or maximum likelihood (Schluter *et al.*, 1997).

Pollen strata from both the Andes and potential source areas can be useful in reconstructing the origin of the tropical alpine floras by providing information of ancestral distributions, changes in the flora over time and ages of certain species or lineages (i.e., how long they have inhabited a certain area). It must be kept in mind that not only are different habitats preserved differentially in the fossil record, within a particular stratum certain taxonomic groups are preferentially preserved while others are rarely or never preserved. The palynological record is more or less restricted to anemophilous species, which tend to produce much larger quantities of pollen than entemophilous taxa. Due to the incompleteness of the fossil record, phylogenetic optimization is in many cases the only option for reconstructing ancestral distributions.

A disadvantage of using a phylogenetic approach when estimating the origin of biotas is the dearth of phylogenetic data and the massive work involved in gathering such data for a large number of clades. The origin of *Polylepis* represents a single data point in the history of the paramo flora. Nevertheless, *Polylepis* was an early and dominant member of the paramo flora and information the evolutionary history and derivation of this lineage can inform phylogenetic studies of other paramo plants or of the paramo community as a whole.

Chapter VI Radiation of *Polylepis* during Pleistocene

Introduction

Many different mechanisms have been proposed that may result in enhanced variation and speciation, including hybridization, polyploidy, disruptive selection, vicariance events and founder effect. The relative importance of the various modes of speciation is influenced by environmental factors such as climate and topography (Vrba *et al.* 1995; Graham 1997). Taxonomic groups living in regions with distinct environmental histories are therefore likely to have radiated via different speciation processes. If the Pleistocene climatic cycles caused repeated events of isolation and contact between populations of *Polylepis*, one would expect that the mode of evolution in *Polylepis* should differ from that of a taxon that evolved in a stable habitat during the same time. Nuclear sequence data from *Polylepis* and sister taxa were gathered and analyzed phylogenetically. The aim of the analysis was to use the phylogenetic tree to compare the rate of speciation (i.e., number of cladogenic events) in *Polylepis* and its phylogenetic sister lineage. Because two sister lineages have had the same amount of time to speciate, this comparison was considered to be the most appropriate. Unexpectedly, the phylogenetic analyses revealed high levels of polymorphism in ITS as well as *Adh* within all sampled individuals of *Polylepis*. The polymorphic clones were not monophyletic within species, nor were they phylogenetically structured in any other way (e.g., similar species relationships among subclades). This intra-individual genetic diversity may suggest random lineage

sorting of ancestral alleles and/or asexual reproduction in some or all "species" of *Polylepis.*

Because of the apparent absence of biological or phylogenetic species in *Polylepis*, a straightforward comparison of speciation rates could not be conducted. In addition, the discovery of a hybrid origin of *Polylepis* (chapter V) further undermined a test of correlation based on phylogenetic contrast. Nevertheless, the taxonomic and genomic complexity of *Polylepis* was compared to biogeographic data on the Pleistocene climatic fluctuations to assess whether the glacial cycles had an effect on the radiation of *Polylepis* in comparison to its sister group(s). In addition, evidence for the occurrence of apomictic reproduction in *Polylepis* is discussed.

Climatological history of the Andes during Pliocene and Pleistocene

The main uplift of the Eastern Cordillera took place between 4.5-3.5 mya and had ceased entirely before the beginning of Pleistocene (Van der Hammen and Hooghiemstra, 1997). The rich paramo flora must therefore have evolved in a time span of less than 3.5 mya. During this period the earth experienced several glacial-interglacial cycles that had a profound effect on biotas, in particular in temperate and alpine regions. Many have suggested that the cycles of drastic climate change associated with the glacial advances could have stimulated speciation processes in the Andes by repeatedly separating and reconnecting populations (Van der Hammen *et al.*, 1973). Glacial advances allowed Andean species to expand their ranges as vegetation belts were lowered, a trend that may have been especially dramatic in the southern Andes (Vuilleumier, 1971). The recent development of sophisticated

paleoclimatatological techniques has revealed that the earth underwent at least ten glacial cycles during the last two million years.

Most extant species of *Polylepis* grow in parts of the Andes that reached their present elevation in the late Tertiary or Pleistocene (Simpson, 1986). Fluctuations in distribution patterns and population sizes of *Polylepis* species during Pleistocene and Holocene have been documented by palynological work. In addition, current distributions have also been greatly influenced by human activities during the last few thousand years. It is though that the current distribution represents a mere fraction of the pre-settlement range of *Polylepis* (Brandbyge, 1992; Fjeldså, 1992; Van der Hammen, 1979).

Materials and Methods

Gene trees of *Polylepis* and sister taxa were generated from multiple loci of the nuclear ribosomal internal transcribed spacers (ITS) and three loci of *Adh1* (see chapter III). Other more distant *Adh* loci were not amplified. The sequences were collected previously for a large-scale analysis of the Sanguisorbeae and the DNA extraction, amplification, cloning and sequencing was described in detail in chapter III. Nearly all currently recognized species of *Polylepis* were sampled, and in some cases more than one accession of the same species (Table III.2). All PCR amplicons were cloned and multiple clones were sequenced from each species.

Sequence contigs were assembled using Sequencher 3.1.1 (Gene Codes Corporation, Ann Arbor, MI) and alignments were assembled manually using MacClade (Maddison and Maddison, 2002). Phylogenetic analyses were conducted

using maximum parsimony analysis and Bayesian inference. Likelihood ratio tests were performed to determine which model(s) of sequence evolution best fit the data. ITS sequences from all loci were analyzed simultaneously. The three *Adh* loci *S.A*, *S.B* and *L* were analyzed separately. Parsimony analyses were conducted using PAUP* (Swofford, 2002) performing 500 heuristic random addition replicates with TBR branch swapping. 100-replicate bootstrap analyses were performed to assess node support. Bayesian analyses were conducted using the program MrBayes (Huelsenbeck and Ronquist, 2001). Two million generations of mcmc were performed sampling trees every 100 generation. Bayesian consensus trees of all compatible clades were constructed after "burnin" trees had been discarded. Because of multiple loci and obvious incongruencies between ITS and *Adh,* no attempt was made to combine the data. Intra- and interspecific pairwise sequence distances were calculated using corrected distances under a GTR+I+Γ model of evolution. The genetic diversity of *Adh1* and ITS was assessed in *Polylepis* and compared to sister groups in the Sanguisorbeae.

Results

Adh1

The likelihood ratio test suggested a superior fit with GTR+I+Γ and this model was incorporated into the Bayesian analyses of the three loci *Adh1L, Adh1S.A* and *Adh1S.B*. The topologies from parsimony and Bayesian analyses were largely congruent for all three genes. The Bayesian trees and posterior probabilities are presented in Figure VI.1A-C. Within *Adh1* three divergent loci were found in

Polylepis and its sister group *Acaena*. In addition, the *Adh1S.B* gene appears to have been duplicated at the base, giving rise to two daughter copies (*B.I* and *B.II*, Figure VI.1A). Multiple copies were found for each locus and each species. This polymorphism occurred at two levels, one of which was shallow and strictly autapomophic and was determined to be the result of polymerase error (see chapter III). Intraspecific polytomic species combs resulting from polymerase error were reduced to a single species, keeping the sequence with the shortest terminal branch length.

The deeper level of intraspecific intra-locus polymorphism was pervasive in *Polylepis,* but absent in all but a few species of *Acaena*. No consistent interspecific phylogenetic pattern could be observed within any of the *Adh1* loci. Most species were polyphyletic/heterozygous within any one locus and no species was monophyletic/homozygous for all loci. All species of *Polylepis* except *P. hieronymi* had multiple copies of *Adh1L* (Table VI.1). Because of selection and drift in the polymerase chain reaction, it is likely that more copies exist that were not detected. Locus-specific primers were designed as far as possible to exhaustively sample the genome of *Polylepis* (Table III.3), but within-locus variation was not structured in a way that allowed for specific priming. Needless to say, only non-identical copies can be distinguished with this approach and the actual number of loci is unknown.

The sequence patterns observed in the alignment of *Adh1S.B* and the high degree of homoplasy in the phylogenetic analysis suggest that the two copies of this locus have undergone extensive inter-locus recombination in several or all the species

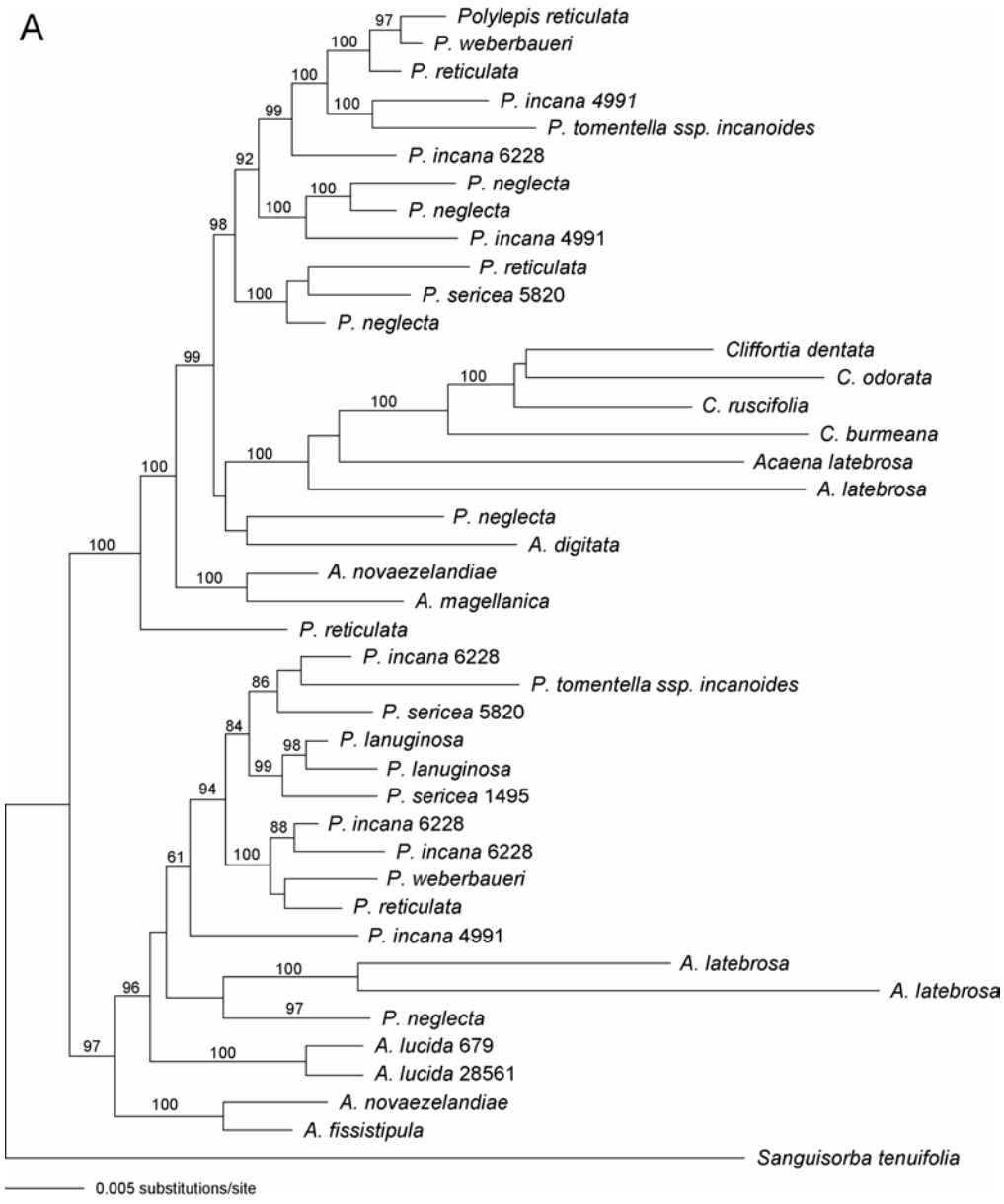
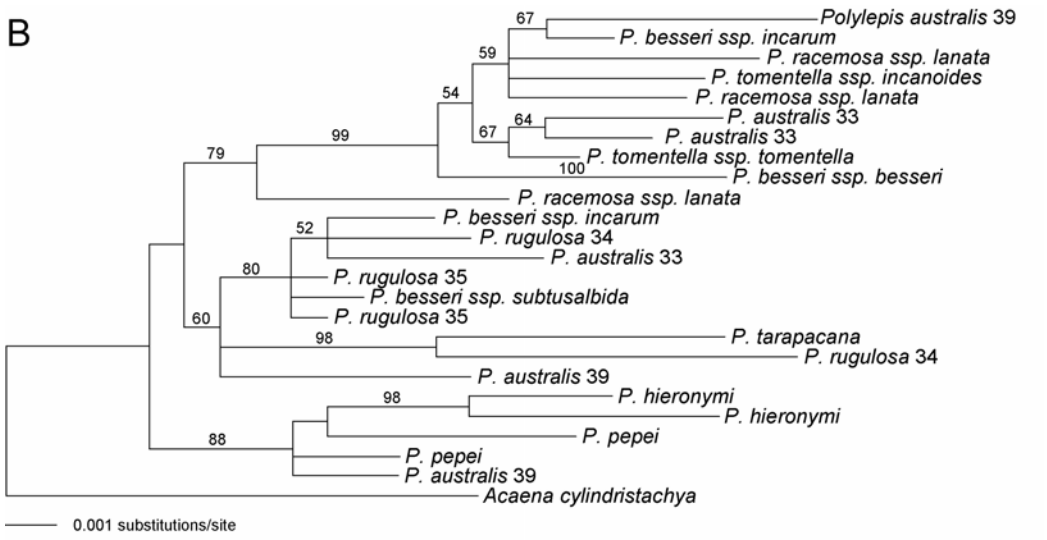


Figure VI.1. Gene trees of *Polylepis* for (A) *Adh1S* (B), *Adh1S.A,* and (C) *Adh1L* based on Bayesian inference. Posterior probabilities indicated above branches.


B
Polylepis australis 39
P. besseri ssp. incarum
P. racemosa ssp. lanata
P. tomentella ssp. incanoides
P. racemosa ssp. lanata
P. australis 33
P. australis 33
P. tomentella ssp. tomentella
P. besseri ssp. besseri
P. racemosa ssp. lanata
P. besseri ssp. incarum
P. rugulosa 34
P. australis 33
P. rugulosa 35
P. besseri ssp. subtusalbida
P. rugulosa 35
P. tarapacana
P. rugulosa 34
P. australis 39
P. hieronymi
P. hieronymi
P. pepei
P. pepei
P. australis 39
Acaena cylindristachya
0.001 substitutions/site

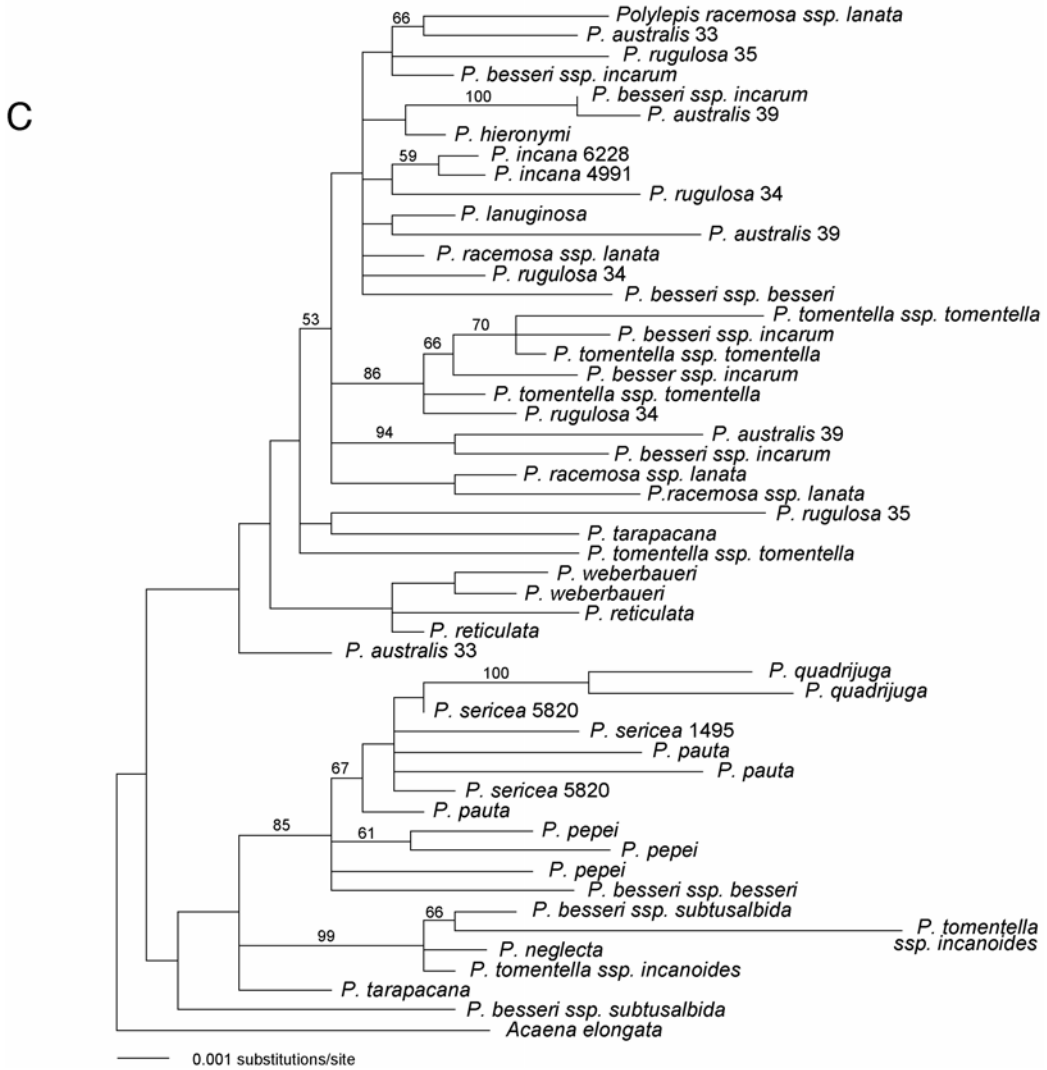

C
Polylepis racemosa ssp. lanata
P. australis 33
P. rugulosa 35
P. besseri ssp. incarum
P. besseri ssp. incarum
P. australis 39
P. hieronymi
P. incana 6228
P. incana 4991
P. rugulosa 34
P. lanuginosa
P. australis 39
P. racemosa ssp. lanata
P. rugulosa 34
P. besseri ssp. besseri
P. tomentella ssp. tomentella
P. besseri ssp. incarum
P. tomentella ssp. tomentella
P. besser ssp. incarum
P. tomentella ssp. tomentella
P. rugulosa 34
P. australis 39
P. besseri ssp. incarum
P. racemosa ssp. lanata
P.racemosa ssp. lanata
P. rugulosa 35
P. tarapacana
P. tomentella ssp. tomentella
P. weberbaueri
P. weberbaueri
P. reticulata
P. reticulata
P. australis 33
P. quadrijuga
P. quadrijuga
P. sericea 5820
P. sericea 1495
P. pauta
P. pauta
P. sericea 5820
P. pauta
P. pepei
P. pepei
P. pepei
P. besseri ssp. besseri
P. besseri ssp. subtusalbida
P. tomentella ssp. incanoides
P. neglecta
P. tomentella ssp. incanoides
P. tarapacana
P. besseri ssp. subtusalbida
Acaena elongata
0.001 substitutions/site

Table VI.1. Number of *Adh1* loci in species of *Polylepis* and corrected (GTR+I+Γ) distances between them.

Species	Number of *Adh1S.A* copies found	Number of *Adh1S.B* copies found	Number of *Adh1L* copies found	Mean corrected distance between *S* and *L* locus	Mean *AdhS* intralocus distance	Mean *AdhL* intralocus distance
Polylepis australis No. 33	2		2	0.1037	0.01529	0.00885
P. australis No. 39	3		3	0.1088	0.01649	0.01225
P. besseri ssp. besseri	1		?			
P. besseri ssp. incarum	2		5	0.1030	0.0123	0.0084
P. besseri ssp. subtusalbida	1		2	0.0991		
P. hieronymi	1		1	0.0995		
P. pepei	1		1	0.0989		
P. rugulosa No. 34	2		3	0.1037	0.01669	0.00786
P. rugulosa No. 35	1		2	0.0998		0.01482
P. tarapacana	1		2	0.1052		0.01246
P. racemosa ssp. lanata	2		3	0.1047	0.01381	0.00669
P. tomentella ssp. tomentella	1		3	0.1017		0.00746
P. incana No. 4991		3	1	0.0969	0.02847	
P. incana No. 6228		3	1	0.0944	0.01687	
P. lanuginosa		1	1	0.0907		
P. neglecta		4	1	0.0956	0.02129	
P. reticulata		4	1	0.0913	0.01644	
P. sericea No. 1495		1	1	0.1028		
P. sericea No. 5820		2	1	0.0968	0.01949	
P. weberbaueri		2	1	0.0944	0.02389	
P. pauta			2			0.00648

in which *Adh1S.B* was found. The possibility of PCR mediated recombination cannot be ruled out but the repeated amplification of identical chimeras suggest that the recombinants may be real. In addition, the chimeric sequences appear to have many break-points suggesting multiple crossing-over or gene conversion events. An attempt was made to cleave chimeric sequences into its two constituent copies to test their original affinities. Because of the multiplicity of crossing-over events, this was not uncomplicated. Nevertheless, a phylogenetic analysis based on the imperfectly split sequences indicates that *Adh1S.B* indeed consists of two subclades and that some sequences have segments "belonging" to both clades (Figure VI.2).

With the sole exception of *P. tomentella ssp. incanoides*, species of *Polylepis* and *Acaena* had either *Adh1S.A* or *Adh1S.B* but not both. As discussed in chapter III, if this pattern is real (and not a PCR artifact) it suggests a propensity to retain a single copy of *Adh1S.* Based on previous phylogenetic data (see chapter II and III) the loss must have occurred more than once in *Acaena* but could potentially represent a single event in *Polylepis.* The corrected genetic distances were substantially greater among *Adh1* loci within the same species than between species within the same locus (see table VI.1). The phylogeny further suggested little or no recombination among loci, with the exception of the two copies of *Adh1S.B.*

ITS

Four distinct ITS loci (AI, AII, B and C) were found, two of which (C and D) had been further duplicated resulting in a total of six copies (Figure VI.2). The relationship among the different loci has been discussed elsewhere (see chapter III).

Figure VI.2. Gene trees of *Polylepis* for ITS loci (A) A.I (B) A.II and (C) B.I-II and (D) C.I-II based on maximum parsimony.

Although the sequences were analyzed together, the four clades are presented separately here for two reasons: (1) relationships among loci was poorly supported and (2) the aim of this study was to elucidate intergeneric relationships of *Polylepis* and that of its most closely related sister lineage. As with the *Adh1* data, autapomorphic species combs were reduced to a single sequence, after it had been confirmed with high-fidelity *Pfu* polymerase that this variation was caused by *Taq* error. The sequence with the shortest terminal branch length was retained which in most cases was a zero-length branch. Similar to *Adh1*, multiple non-identical copies of each locus were found in many species of *Polylepis*. Separate loci were identical in sequence in conserved regions and it proved impossible to design locus-specific primers. For this reason it is possible that despite the sequencing of 20-40 clones from each species, some ITS copies may have been missed.

All of the sampled species and subspecies of *Polylepis* carry the functional A.I locus, but not all of them carry the other loci. Only four species were found to have locus A.II, thirteen were carrying either B.I, B.II or both and fourteen were carrying either C.I, C.II or both. Again, the apparent lack of certain copies in some species could be partly due to PCR selection and drift.

There was little phylogenetic resolution among *Polylepis* species at the same locus. The functional locus A.I, in particular, demonstrated a very low level of variation. Despite much higher substitution rates in the pseudogenized loci, interspecific relationships appeared to be "scrambled" and random. Levels of

homoplasy were high in loci B and C possibly suggesting recombination between daughter copies.

Discussion

Interspecific relationships in Polylepis

Based on the relatively young age of the high Andean habitats, Simpson suggested that the oldest members of *Polylepis* are those species found today at lower elevations (e.g., *P. multijuga, P. pauta* and *P. hieronymi).* Simpson further argued that differences in speciation pattern among three subgeneric groups of *Polylepis* were caused by the history of the geographic region where each group had evolved. She inferred sequential morphological changes as each lineage adapted to higher elevation environments during the gradual uplift of the Andes.

The phylogenetic analyses of *Polylepis* sequences presented here revealed practically no genealogical congruence among different loci of ITS and *Adh1*. Genomic-level characters such as the presence or absence of different loci appear to be only slightly more informative with regards to organismal relationships. If the loss of either *Adh1S.A* or *Adh1S.B* occurred early in the evolution of *Polylepis*, species that carry the same locus would be expected to be more closely related than species that carry different loci. In other words, *P. australis, P. racemosa, P. besseri, P. rugulosa, P. tomentella, P. tarapacana, P. hieronymi* and *P. pepei* should represent a monophyletic group while *P. reticulata, P. weberbaueri, P. incana, P. neglecta, P. sericea* and *P. lanuginosa* should represent another (*P. quadrijuga* and *P. pauta,* appeared to only carry locus *Adh1L).* The dichotomy of *Polylepis* into these two

groups does not correspond with the infrageneric relationships suggested by Simpson (1979) and Kessler (1995b), although some agreement can be found. Simpson suggested that *P. racemosa, P. australis, P. besseri* and *P. tomentella* were closely affiliated, a relationship that is in agreement with a putative shared loss of *Adh1S.B*. However, she also considered *P. incana* a member of this group, something that is not supported by the *Adh* data. In addition, *P. hieronymi*, and *P. pepei* appear to also have lost *Adh1S.B*. Based on a morphological cladistic analysis, Kessler proposed a close relationship among *P. besseri, P. rugulosa, P. tomentella, P. tarapacana*, *P. incana* and *P. racemosa* and suggested that these taxa, in turn, form a sister group to *P. neglecta, P. australis* and *P. crista-galli* (not sampled in *Adh* analysis). Again, with the exception of *P. incana* and *P. neglecta,* all of these species carry the *Adh1S.A* locus and share a putative loss of *Adh1S.B*.

Intragenomic nuclear DNA polymorphism and correlation with reproductive biology

Nuclear regions of ribosomal DNA (rDNA) have been applied as molecular markers in countless phylogenetic analyses of organisms across the tree of life. Because of their importance for cellular function rDNA loci occur as large clusters of serial units. In diploid sexually reproducing individuals these tandemly repeated units tend to be extremely homogeneous making them suitable for phylogenetic inference. The homogeneity across hundreds or thousands of repeats has been attributed to a remarkably efficient process of genetic exchange among repeats through unequal crossing over and/or gene conversion (Dover *et al.*, 1982). As long as this process is

faster than speciation the entire rDNA region will evolve in concert (i.e., concerted evolution).

Despite the common assertion of almost ubiquitous concerted evolution in rDNA, several studies suggest that intraspecific or even intra-individual rDNA polymorphisms may be more common than previously thought. Polymorphic rDNA appear to be especially prevalent in polyploid and apomictic plants suggesting that concerted evolution of repetetive DNA can be interrupted or reduced by polyploidy and asexual reproduction (Aguilar *et al.*, 1999; Anamthawat-Jonsson and Heslop-Harrison, 1995; Gruber, 1991; Polanco and Delavega, 1995; Waters and Schaal, 1996; Zhang *et al.*, 1992). Campbell *et al.* (1997) found intraspecific polymorphic ITS sequences in the agamic complex *Amelanchier* and argued that the maintenance of this pattern could be due to polyploidy and/or agamospermy. Genetic studies of the agamic species complexes in *Taraxacum* and *Chondrilla* suggest significant intrapopulation diversity for several nuclear markers (van Dijk, 2003). In addition, high intraspecific or intraindividual diversity of rDNA has been observed within individuals of parthenogenic animals (Aguilar *et al.*, 1999; Gandolfi *et al.*, 2001; Harris and Crandall, 2000; Hugall *et al.*, 1999; Schön *et al.*, 2000) and fungi (Lanfranco *et al.*, 1999).

High genetic diversity within asexual species appears to contradict traditional theory which asserts that the very nature of asexual reproduction must lead to low genetic diversity (Darlington, 1939; Grant, 1981). This, however, is only true if one assumes obligate asexuality. Occasional influx of new genetic material through hybridization with sexual species can maintain diversity in facultative apomictic taxa.

In addition, genetic diversity in apomicts can be due to multiple independent origins of apomictic lineages through hybridization of the same sexual progenitors (Hugall *et al.*, 1999; Schön *et al.*, 2000; Soltis and Soltis, 1991).

Naturally, allopolyploid hybrids will at least initially have a greater genetic diversity because of hybrid heterozygozity. In addition, studies of artificially generated polyploids have demonstrated that polyploid species can generate extensive "new" genetic diversity in a short period of time (Song *et al.*, 1995). Although absence of meiosis in apomictic species should limit the possibility for repeat homogenization (Campbell *et al.*, 1997), gene conversion and crossing-over has been documented in some parthenogenic organisms (Crease and Lynch, 1991).

Polylepis –*an agamic complex?*

The pattern of rDNA diversity in *Polylepis* is contradictory. On the one hand, the intraindividual heterogeneity of sequences suggests that concerted evolution is absent or inefficient. On the other hand, the sequences of ITS loci B and C (as well as *Adh1S.B*) suggest recombination in the form of crossing-over and/or partial gene conversion. This pattern has been observed in other parthenogenic organisms and could indicate that *Polylepis* is apomictic. However, the incidence of apomixis is always correlated with polyploidy and the effects of these two components are difficult to separate. A third explanation is that the lack of phylogenetic structure in *Polylepis* represents dynamics at the population level and that the entire *Polylepis* complex is essentially acting as a single species. This does not, however, explain the lack or suppression of concerted evolution.

Despite being purportedly wind-pollinated, the dispersal of pollen from extant *Polylepis* species appears to be rather limited (Graf, 1986; Salgado-Labouriau, 1979). Wind-pollinated plants tend to produce large amounts of pollen but some species of *Polylepis*, in particular at high elevations, have highly reduced inflorescences with sometimes only a few flowers. Although the reproductive biology of *Polylepis* has not been studied comprehensively, geitonogamy (selfing between different flowers on the same plant) has been observed in *Polylepis australis* (Olsen, 1976) and it is possible that selfing is common throughout *Polylepis*. Although the evidence is circumstantial, many characteristics of *Polylepis* indicate that some or all species are facultatively apomictic. The complex pattern of evolution in *Polylepis* apparent from its multiple polymorphic copies of nuclear genes and their interrelatedness, the documented high ploidy levels and the complex and continuous morphological variation, are all indications of an abnormal reproductive system. In addition, parthenogenic taxa are often found in arid environments and at high elevations, similar to the habitat of *Polylepis*. Apomixis has been documented in at least 18 genera of the Rosaceae (Czapik, 1996), e.g., *Crataegus, Amelanchier, Sorbus, Rubus*, and *Potentilla*, among others.

Asexual populations tend to have a particular tendency to colonize previously glaciated areas. The occurrence of apomixis in *Polylepis* may have coincided with the Pleistocene glacial cycles that provided repeated opportunities for hybridization between previously isolated populations. Similar scenarios have been suggested in the apomictic rosaceous genera *Rubus* (Haskell, 1966) and *Aronia* (Hardin, 1973). Sinnott and Phipps (1983) suggested that a combination of Pleistocene vicariance and

settler disturbance was responsible for the complex evolutionary patterns seen in North American *Crataegus*.

The species-pump effect in Polylepis

The species-pump model is related to, and an extension of, Haffer's (1969) refuge hypothesis for the lowland tropics. He argued that changes in vegetation patterns caused by the Pleistocene climatic cycles were responsible for the animal diversity and species composition of today's tropics. Similarly, the evolution of alpine taxa must have been influenced by the repeated shifts in climate and distribution caused by the glacial cycles.

Without recognizable species units, an absolute rate of speciation in *Polylepis* could not be measured. In addition, testing for a species-pump effect in *Polylepis* against is sister lineage proved to be an invalid contrast, not only because of the hybrid origin of *Polylepis* (see chapter V), but because the putative parental lineages, *A. elongata* and *A. cylindristachya* are both capable of occupying high elevation habitats and may have been influenced by the same "species-pump" effect as *Polylepis*.

Disregarding causal correlations between vicariance and speciation, it is clear that *Polylepis,* as a clade, exhibits considerable morphological and molecular diversity. This becomes even more evident when *Polylepis* is compared to its phylogenetic sister taxa *A. elongata* and *A. cylindristachya*. Although the exact relationship between *Polylepis* and these two lineages remains vague, the contrast in morphological and molecular variation is conspicuous. *A. cylindristachya* is

morphologically distinct from other *Acaena* and was placed in its own section by Bitter (1911a). According to nuclear and chloroplast genealogies (chapter II, III and V) the lineage is monospecific. Bitter split *A. elongata* into two species (*A. elongata* L. and *A. torilicarpa* Bitter) and placed them in the same section as the Bolivian-Argentine species *A. stricta* Griseb. Plant material of *A. torilicarpa* and *A. stricta* was not obtained for this analysis. Nevertheless, the molecular phylogenies suggest that the lineage to which *A. elongata* belongs is species poor and morphologically relatively uniform. In contrast, the taxonomic splitting of *Polylepis* into as many as 34 species (Bitter, 1911a) is indicative of the range of morphological diversity in this group. The rugged xerophytic shrubs of high alpine *P. tarapacana* are strikingly different from the tall, large- and thin-leaved trees of *P. multijuga*. The inflorescences of different *Polylepis* "species" range from very small (a couple of flowers) to long dense racemes. Fruits are variously adorned with spines, wings and ridges. Whether the Pleistocene "species pump" contributed to the morphological diversity seen in *Polylepis* cannot be confirmed. Perhaps the best working hypothesis is that a combined effect of vicariance events and facultative apomixis is responsible for the current patterns of morphological and molecular variation of *Polylepis*.

Species recognition in Polylepis

An agamic complex is a population where frequent hybridization between sexuals and/or apomicts produce a large amount of reproductively isolated morphological subtypes (Darlington, 1939; Grant, 1981). The advantages of apomixis in crop production (predictable uniform progeny) has recently caused a boom in

apomixis research (van Dijk and van Damme, 2000). Consequently we now know much more about the genetic basis of apomixis than was possible only a few years ago. This does not, however, eliminate the problems systematists face when trying to apply a phylogenetic, evolutionary or biological species concept to apomictic entities. It is frequently assumed in many taxonomic treatments that morphologically distinct entities coincide with phylogenetic units of most closely related individuals. This assumption may be violated in biparentally reproducing organisms and is an even less useful criterion for classifying apomictic or hybrid complexes. Morphology in apomictic groups does not necessarily correspond well with phylogenetic relationships (Campbell and Dickinson, 1990). Likewise, the biological species concept is difficult to apply to a group that reproduces (sexually) infrequently or never. Hybridization complicates classification further by introducing polyphyletic taxa (Campbell and Dickinson, 1990). If apomictic species are defined as morphospecies they may be polyphyletic, arising repeatedly from hybridization between two mixed but separate populations.

Phylogenetic species concepts, and in particular the explicit genealogical concordance concept (GCC) (Macklin and Phipps, 2002; Mayden, 1999), are appealing and provide practical criteria for the classifications of many taxonomic groups. Difficulties arise in groups where there is evidence of occasional interspecific or even intergeneric hybridization. Following the GCC, all units that are hybridizing, even at very rare occasions, belong to the same species. The propensity of hybridization in *Polylepis* and *Verruchaena* as a whole makes the application of GCC impractical. Despite evidence of a hybrid origin of *Polylepis* and subsequent

chloroplast introgression, lumping all of *Polylepis* and most of *Acaena* into a single species seems absurd considering the morphological diversity within the group.

Approaches to classifying apomictic groups span from extreme lumping to extreme splitting. The most radical solution is to treat the entire agamic complex as one species, including the hybridizing parental sexual species (Carman, 1997; Cruise, 1964; Kellogg, 1990). The other extreme is to recognize each morphological variant as a microspecies, which may not correspond to any meaningful evolutionary or even ecological unit. The currently recognized species delineations in *Polylepis* were based on careful morphological studies (Kessler, 1995a; Kessler, 1995b; Simpson, 1979). In lieu of a unifying criterion for the application of species concepts in agamic or reticulate taxa, a new classification based on molecular evidence is not warranted.

Chapter VII Conclusions and suggestions for further study

Conclusions

This study was designed with the aim of addressing a number of straight-forward hypotheses regarding phylogeny and biogeography in the tribe Sanguisorbeae. As is often the case in science, the results triggered larger questions, and led the researcher onto paths not initially considered. The molecular analyses of Sanguisorbeae and *Polylepis* revealed an exceptionally complex phylogenetic and genomic history and suggested that evolutionary innovation in this group is driven in large part by allopolyploidy.

It is now considered uncontroversial that evolution in plants as well as other organisms has been strongly influenced by whole genome duplication (e.g., Levy and Feldman, 2002; Ozkan *et al.*, 2001; Ozkan *et al.*, 2002; Paterson *et al.*, 2000). Sophisticated sequence analysis have lead to the discovery of ancient polyploidization even in lineages with very small genomes, e.g., yeast (Wong *et al.*, 2002). Numerous studies have suggested that polyploidy and asexual reproduction are strongly correlated with certain geographic and environmental variables (e.g., geographic parthenogenesis). In addition, polyploid and/or asexual plants and animals appear to be better at colonizing extreme and disturbed habitats than their diploid and sexual relatives and are especially common at high latitudes and high elevations (Bell, 1982; Glesener and Tilman, 1978; Lynch, 1984; Stebbins, 1950).

Nuclear genomic evolution in Sanguisorbeae

Phylogenies of nuclear gene families may include orthologous as well as paralogous copies and are potentially powerful resources for resolving hypotheses of allopolyploidy. Even so, evidence of ancient polyploidization can be difficult to detect because of genome rearrangements that may erase similarities between homeologous chromosomes. In addition, gene loss, divergence or recombination can obscure the relationships among gene homeologs. When using multi-copy markers in phylogenetic analysis the genes generally preferred are those that exhibit either a very high degree of inter-locus recombination (complete concerted evolution) or those that do not recombine at all. In reality, there are probably a number of nuclear genes that fall somewhere in between. This could be especially true in polyploid organisms where some recombination between synologous gene copies may be expected.

Despite, or maybe owing to, multiple duplications both before and after the divergence of Sanguisorbeae, *Adh* proved to be highly informative for resolving cladogenic as well as tokogenic relationships within Sanguisorbeae. The exceedingly complex ITS genealogy appeared, at first sight, not to be very informative with regard to organismal phylogeny. Even so, careful analysis of sequences revealed tokogenic relationships not apparent from phylogenetic analysis of either *trnL/F* or *Adh* data. Fixed heterozygosity in ITS and/or *Adh1* in *Sanguisorba sensu stricto, Verruchaena, Amentomorpha* and *Polylepis* supported hypotheses of allopolyploid origins of these taxa. Circumstantial evidence based on chimeric ITS sequences likewise suggested an allopolyploid origin of the tetraploid petalous Agrimoniinae through hybridization of diploid ancestral members of Potentilleae and Sanguisorbeae. In addition, insights

were gleaned into the genomic evolutionary processes of rDNA in polyploid Sanguisorbeae, which may be applicable to other polyploid plant taxa. Because ITS is one of the most popular phylogenetic markers in plant systematics, it is important to understand its underlying evolutionary mechanisms. When ITS genealogies are used to infer organismal phylogenetic relationships it is particularly important to appreciate that multiple loci may be present even in a diploid species (e.g., *Poteridium*).

Biogeography of Sanguisorbinae

An ancestral *Verruchaena* lineage, morphologically similar to *Acaena,* originated, by hybridization or cladogenesis, from *Poterium, Poteridium* and/or *Sanguisorba* somewhere in the New World in middle to late Miocene. An early member of this lineage appears to have spread to South Africa in late Miocene where it radiated into *Cliffortia*. Another lineage of *Verruchaena* remained in South America where it gave rise to *Amentomorpha (Acaena, Polylepis* and *Margyricarpus-Tetraglochin).* Basal members of *Amentomorpha* colonized Juan Fernandez Island and Hawaii and three independent dispersal events to Australasia gave rise to the *Acaena* species there. The first of these transoceanic dispersals (giving rise to *A. montana*) could have occurred as early as the Miocene-Pliocene transition, while the latest one (the *A. echinata* complex) probably occurred much more recently.

An alternative scenario is possible if the hybrid origin of *Acaena latebrosa* is a misinterpretation. In this case *A. latebrosa* could be the oldest representative of *Verruchaena*, giving rise to *Acaena* in South America (where it subsequently went extinct) and radiating into *Cliffortia* after dispersal to South Africa. A complete

understanding of the biogeographic history of *Verruchaena* may not be possible, but a more comprehensive sampling of extant species of *Acaena* could improve on the assessments presented here.

Evolution and biogeography in Polylepis

In line with the propensity for hybridization in Sanguisorbeae as a whole, *Polylepis* appears to be yet another example of allopolyploidy as a driving force of evolutionary innovation in this group. A hybrid origin of *Polylepis* from two ancestral *Acaenas* appears to reconcile the incongruent gene phylogenies exhibited by different nuclear loci. This scenario is also in agreement with extant distribution patterns of putative parental lineages.

Although a direct phylogenetic test of a "species pump" effect in *Polylepis* was not applicable, the high molecular and morphological diversity in *Polylepis* compared to its sister taxa suggests a fundamental difference in the evolutionary processes (past and present) governing these lineages. Based on documented apomixis in other rosaceous taxa and parallels to patterns of sequence evolution in a wide variety of parthenogenic organisms, it appears reasonable to extrapolate that *Polylepis* at least partly comprise of apomictic assemblages.

Suggestions for further study

Fluorescent in situ hybridization of rDNA in *Polylepis, Acaena* and *Cliffortia*, along with detailed cytological investigations should complement the current sequence analysis and allow for a more complete picture of ITS evolution in Sanguisorbeae. Studies of reproductive biology of Sanguisorbeae in general and of *Polylepis* in particular would greatly add to our understanding of the evolution of these polyploid taxa. Recently developed technologies in flow cytometry of seeds could be useful not only in place of conventional chromosome counts, but also as a means to detect polyploidy and apomixis in *Polylepis*.

A complete taxonomic revision of Sanguisorbeae is forthcoming. In addition to new clade names proposed here, a thoroughly sampled analysis of *Acaena* will result in a revised nomenclature of this paraphyletic taxon.

Chapter VIII . Appendix

Nomenclature

Phylogenetic definitions of taxa supported by molecular phylogenetic data are provided. The definitions follow the guidelines of the current Phylocode (www.phylocode.org) and are stem-based. Because the Phylocode requires that the dissemination of new taxonomic names be peer-reviewed before considered validly published, this nomenclature will also be published elsewhere.

Poteridium

Stem-based definition. *Poteridium* is the most inclusive clade containing *Poteridium annuum* (syn. *Poterium annuum, Sanguisorba annua*) but not *Poterium sanguisorba* (syn. *Sanguisorba minor*), *Sanguisorba filiformis* or *Sanguisorba officinalis.*

ICBN at genus rank. *Poteridium* in Spach 1846 p.43. Typus *Poteridium annuum* Nutt ex Hook.

Poterium

Stem-based definition. *Poterium* is the most inclusive clade containing *Poterium sanguisorba* (syn. *Sanguisorba minor*) but not *Sanguisorba filiformis*, *Sanguisorba officinalis* or *Poteridium annuum*.

ICBN at genus rank. *Poterium* in Linnaeus 1753 p. 994. Typus *Poterium sanguisorba* L.

Verruchaena

Stem-based definition. *Verruchaena* is the most inclusive clade containing *Acaena elongata,* but not *Sanguisorba filiformis, Sanguisorba officinalis, Poteridium annuum* or *Poterium sanguisorba.*

No rank.

The clade *Verruchaena* comprises the genera *Acaena, Ancistrum, Polylepis, Cliffortia*, *Margyricarpus* and *Tetraglochin*.

Amentomorpha

Stem-based definition. *Amentomorpha* is the most inclusive clade containing *Acaena elongata,* but not *Cliffortia dentata, Sanguisorba filiformis*, and *Poteridium annuum*.

No rank.

The clade *Amentomorpha* comprises the genera *Acaena, Ancistrum, Polylepis, Margyricarpus* and *Tetraglochin*.

Chapter IX Glossary

homeologous	genes/chromosomes arising through hybridization.
orthology	homology of gene lineages attributable to speciation.
paleopolyploid	ancient polyploids that behave like diploids in that they have disomic inheritance (becoming "diploidized") (definitions vary)
paralogy	homology of gene lineages attributable to gene duplication. Paralogy can be further separated into (a) strict paralogy referring to duplicated genes within the same organism and (b) *metalogy* referring to paralogous genes from two different organisms.
plerology	homologous genes that have undergone partial or complete concerted evolution.
synology	homology of gene lineages attributable to genomic fusion. This term is usually used in the context of endosymbionts but is here applied to the chimeric genomes of hybrids.
tokogeny	reticulate relationships

Chapter X References

Aguilar, J. F., J. A. Rossello, and G. N. Feliner. 1999. Nuclear ribosomal DNA (nrDNA) concerted evolution in natural and artificial hybrids of Armeria (Plumbaginaceae). Molecular Ecology 8:1341-1346.

Anamthawat-Jonsson, K., and J. S. Heslop-Harrison. 1995. Molecular cytogenetics of Icelandic birch species - physical mapping by in-situ hybridization and rDNA polymorphism. Canadian Journal of Forest Research 25:101-108.

Basinger, J. F. 1976. *Paleorosa similkameenensis*, gen. et sp. nov., permineralized flowers (Rosaceae) from Eocene of British-Columbia. Canadian Journal of Botany 54:2293-2305.

Bell, G. 1982. The masterpiece of nature: the evolution and genetics of sexuality. University of California Press, Berkeley.

Bitter, G. 1911a. Die Gattung *Acaena*. Bibiotecha Botanica 74:1-80.

Bitter, G. 1911b. Revision der Gattung *Polylepis*. Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie 45:564-656.

Bortiri, E., S. H. Oh, J. Jiang, S. Baggett, A. Granger, C. Weeks, M. Buckingham, D. Potter, and D. E. Parfitt. 2001. Phylogeny and systematics of *Prunus* (Rosaceae) as determined by sequence analysis of ITS and the chloroplast *trnL-trnF* spacer DNA. Systematic Botany 26:797-807.

Bramwell, D. 1976. VI. The endemic flora of the Canary Islands; distribution, relationships and phytogeography. Pages 207-240 *in* Biogeography and Ecology in the Canary Islands (G. Kunkel, ed.) Dr. W. Junk, The Hague.

Brandbyge, J. 1992. Planting of local woody species in the páramo. Pages 265-274 *in* Paramo - An Andean ecosystem under human influence (H. Balslev, and J. L. Luteyn, eds.). Academic Press, London.

Brooks, D. R., and D. A. McLennan. 1993. Historical ecology: Examining phylogenetic components of community evolution. Pages 267-280 *in* Species diversity in ecological communities. Historical and geographical perspectives The University of Chicago Press, Chicago.

Buckler, E. S. I., A. Ippolito, and T. P. Holtsford. 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. Genetics 145:821-832.

Campbell, C. S., and T. A. Dickinson. 1990. Apomixis, patterns of morphological variation, and species concepts in subfam. Maloideae (Rosaceae). Systematic Botany 15:124-135.

Campbell, C. S., M. F. Wojciechowski, B. G. Baldwin, L. A. Alice, and M. J. Donoghue. 1997. Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae). Molecular Biology and Evolution 14:81-90.

Cantino, P. D., and K. de Queiroz. 2000. The PhyloCode: a phylogenetic code of biological nomenclature. http://www.ohiou.edu/phylocode/.

Carlquist, S. C. 1965. Island Life: A Natural History of the Islands of the World. Natural History Press, Garden City, N.Y.

Carman, J. G. 1997. Asynchronous expression of duplicated genes in angiosperms may cause apomixis, bispory, tetraspory and polyembryony. Biological Journal of the Linnean Society 61:51-94.

Cevalloz-Ferriz, S. R. S., D. M. Erwin, and R. A. Stockey. 1993. Further observations on *Paleorosa similkameenensis* (Rosaceae) from the Middle Eocene Princeton chert of British Columbia, Canada. Review of Palaeobotany and Palynology 78:277-292.

Challice, J. S. 1981. Chemotaxonomic studies in the family Rosaceae and the evolutionary origins of the subfamily Maloideae. Preslia 53:289-304.

Chardon, C. E. 1938. Apuntaciones sobre el origen de la vida en los Andes. Boletin Sociedad Ciencias Naturales (Caracas, Venezuela) 5:1-47.

Citerne, P. 1897. Du genre *Acaena*. Revue des Sciences Naturelles de l'Ouest 7.

Cleef, A. M. 1979. The phytogeographical position of the neotropical vascular páramo flora with special reference to the Colombian Cordillera Oriental. Pages 175-184 *in* Tropical Botany. Proceedings of a symposium held at the University of Aarhus on 10-12 August 1978 (K. Larsen, and L. B. Holm-Nielsen, eds.). Academic Press, London.

Cleef, A. M. 1981. The vegetation of the páramos of the Colombian Cordillera Oriental. Cramer, Vaduz.

Colinvaux, P. A. 1996. Quaternary environmental history and forest diversity in the Neotropics. Pages 359-405 *in* Evolution and environment in tropical America (J. B. C. Jackson, A. F. Budd, and A. G. Coates, eds.). The University of Chicago Press, Chicago.

Crawford, D. J., S. Brauner, M. B. Cosner, and T. F. Stuessy. 1993. Use of RAPD markers to document the origin of the intergeneric hybrid *Margyracaena skottsbergii* (Rosaceae) on the Juan Fernandez Islands. American Journal of Botany 80:89-92.

Crease, T. J., and M. Lynch. 1991. Ribosomal DNA variation in *Daphnia pulex* reproducing by cyclic and obligate parthenogenesis. Molecular Biology and Evolution 8:620-640.

Cruise, J. E. 1964. Studies of natural hybrids in *Amelanchier*. Canadian Journal of Botany 42:651-663.

Czapik, R. 1996. Frequency of agamospermy. Apomixis Newsletter 9.

Darlington, C. D. 1939. The evolution of genetic systems. Cambridge University Press, Cambridge, UK.

Darlington, C. D. 1963. Chromosome botany and the origins of cultivated plants, 2nd edition. George Allen and Unwin, London.

Davis, C. C., C. D. Bell, S. Mathews, and M. J. Donoghue. 2002. Laurasian migration explains Gondwanan disjunctions: Evidence from Malpighiaceae. Proceedings of the National Academy of Sciences of the United States of America 99:6833-6837.

Dawson, J. W. 1960. Natural *Acaena* hybrids in the vicinity of Wellington. Transactions of the Royal Society of New Zealand 88:13-27.

De Candolle, A. P. 1825. Prodromus systematis naturalis regni vegetabilis. Treuttel et Würtz, Paris.

De Queiroz, K., and P. D. Cantino. 2001. Phylogenetic nomenclature and the PhyloCode. Bulletin of Zoological Nomenclature 58:254-271.

De Queiroz, K., and J. Gauthier. 1990. Phylogeny as the central principle in taxonomy: phylogenetic definitions of taxon names. Systematic Zoology 39:307-322.

De Queiroz, K., and J. Gauthier. 1994. Toward a phylogenetic system of biological nomenclature. Trends in Ecology and Evolution 9:27-31.

Dorst, J., and F. Vuilleumier. 1986. Convergences in bird communities at high altitudes in the tropics (especially the Andes and Africa) and at temperate latitudes (Tibet). Pages 120-150 *in* High altitude tropical biogeography (F. Vuillemier, and M. Monasterio, eds.). Oxford University Press, Oxford.

Dover, G., E. Coen, and T. Strachan. 1982. The dynamics of sequence turnover and concerted evolution of families of rDNA, histone genes and non-coding DNA of *Drosophila*. Heredity 49:141-141.

Doyle, J. A., and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19:11-15.

Doyle, J. J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. Systematic Botany 17:144-163.

Ecklon, C. F., and C. Zeyher. 1836. Enumeratio plantarum Africae australis extratropicae, Hamburg.

Eide, F. 1981. Key for northwest European Rosaceae pollen. Grana 20:101-118.

Ellenberg, H. 1979. Man's influence on tropical mountain ecosystems in South America. Journal of Ecology 67:401-416.

Ereshefsky, M. 1994. Some problems with the Linnean hierarchy. Philosophy of Science 61:186-205.

Eriksson, T., M. J. Donoghue, and M. S. Hibbs. 1998. Phylogenetic analysis of *Potentilla* using DNA sequences of nuclear ribosomal internal transcribed spacers (ITS), and implications for the classification of Rosoideae (Rosaceae). Plant Systematics and Evolution 211:155-179.

Eriksson, T., M. S. Hibbs, A. D. Yoder, C. F. Delwiche, and M. J. Donoghue. 2003. The phylogeny of Rosoideae (Rosaceae) based on sequences of the internal transcribed spacers (ITS) of nuclear ribosomal DNA and the *trnL/F* region of chloroplast DNA. International Journal of Plant Sciences 164:197-211.

Estes, R., and A. Baez. 1985. Herpetofaunas of North and South America during the Late Cretaceous and Cenozoic: Evidence for interchange? Pages 140-200 *in* The great American biotic interchange (F. G. Stehli, and S. D. Webb, eds.). Plenum, New York.

Evans, R. C., and C. S. Campbell. 2002. The origin of the apple subfamily (Maloideae; Rosaceae) is clarified by DNA sequence data from duplicated GBSSI genes. American Journal of Botany 89:1478-1484.

Ferguson, D., and T. Sang. 2001. Speciation through homoploid hybridization between allotetraploids in peonies (*Paeonia*). Proceedings of the National Academy of Sciences of the United States of America 98:3915-3919.

Fjeldså, J. 1992. Biogeography of the birds of the *Polylepis* woodlands of the Andes. Pages 31-44 *in* Paramo - An Andean ecosystem under human influence (H. Balslev, and J. L. Luteyn, eds.). Academic Press, London.

Forcone, A., and M. C. Tellería. 2000. Palynological characterization of honeys from the plain of the Senguerr river (Chubut-Argentina). Darwiniana 38:267-271.

Forster, J. R., and G. Forster. 1776. Characteres Generum Plantarum. B. White,T. Cadell, P. Elmsly, London.

Fukuhara, T. 2000. Variations in pollen and ovule parameters among different ploidy levels of *Corydalis*. Plant Systematics and Evolution 224:1-12.

Gandolfi, A., P. Bonilauri, V. Rossi, and P. Menozzi. 2001. Intraindividual and intraspecies variability of ITS1 sequences in the ancient asexual *Darwinula stevensoni* (Crustacea : Ostracoda). Heredity 87:449-455.

Gaut, B. S., B. R. Morton, B. C. McCaig, and M. T. Clegg. 1996. Substitution rate comparisons between grasses and palms: Synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcL*. Proceedings of the National Academy of Sciences of the United States of America 93:10274-10279.

Gaut, B. S., A. S. Peek, B. R. Morton, and M. T. Clegg. 1999. Patterns of genetic diversification within the *Adh* gene family in the grasses (Poaceae). Molecular Biology and Evolution 16:1086-1097.

Ge, S., T. Sang, B. R. Lu, and D. Y. Hong. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. Proceedings of the National Academy of Sciences of the United States of America 96:14400-14405.

Gentry, A. H. 1982. Neotropical floristic diversity - phytogeographical connections between Central and South-America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny. Annals of the Missouri Botanical Garden 69:557-593.

Glesener, R. R., and D. Tilman. 1978. Sexuality and the components of environmental uncertainty: clues from geographic parthenogenesis in terrestrial animals. American Naturalist 112:659-673.

Gmelin, J. F. 1791. *in* Systema Naturae (C. Linnaeus, ed.), Leipzig.

Goldstein, G., F. C. Meinzer, and F. Rada. 1994. Environmental biology of a tropical treeline species, *Polylepis sericea*. Pages 129-149 *in* Tropical alpine environments: plant form and function (P. W. Rundel, A. P. Smith, and F. C. Meinzer, eds.). Cambridge University Press, Cambridge.

Graf, K. 1986. Klima und Vegetationsgeographie der Anden - Grundzüge Südamerikas und pollenanalytische Spezialuntersuchungen Boliviens, Zürish.

Grant, V. 1981. Plant speciation, 2nd edition. Columbia Univ. Press, New York.

Gregory-Wodzicki, K. M. 2002. A late Miocene subtropical-dry flora from the northern Altiplano, Bolivia. Palaeogeography Palaeoclimatology Palaeoecology 180:331-348.

Gregory-Wodzicki, K. M., W. C. McIntosh, and K. Velasquez. 1998. Climatic and tectonic implications of the late Miocene Jakokkota flora, Bolivian Altiplano. Journal of South American Earth Sciences 11:533-560.

Grodona, E. 1964. Las especies Argentinas del género "*Acaena*". Darwiniana 13.

Gruber, V. 1991. Heterozygosity in tetraploid potatoes revealed by rDNA polymorphism analysis of their dihaploid progenies - a contribution to chromosome assignment. Plant Molecular Biology 17:1045-1054.

Hardin, J. W. 1973. The enigmatic chokeberries (*Aronia*, Rosaceae). Bulletin of the Torrey Botanical Club 100:178-184.

Harris, D. J., and K. A. Crandall. 2000. Intragenomic variation within ITS1 and ITS2 of freshwater crayfishes (Decapoda : Cambaridae): Implications for phylogenetic and microsatellite studies. Molecular Biology and Evolution 17:284-291.

Haskell, G. 1966. The history, taxonomy and breeding system of apomictic British Rubi. Pages 141-151 *in* Reproductive biology and taxonomy of vascular plants (J. G. Hawkes, ed.) Clarendon, Oxford.

Hedberg, I., and O. Hedberg. 1979. Tropical-alpine life-forms of vascular plants. Oikos 33:297-307.

Hedberg, O. 1992. Afroalpine vegetation compared to páramo: Convergent adaptations and divergent differentiation. Pages 31-44 *in* Paramo - An Andean ecosystem under human influence (H. Balslev, and J. L. Luteyn, eds.). Academic Press, London.

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

Hieronymus, G. 1895. Plantae Lehmannianae. Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie 20 (III, Supplement 49):1-72.

Hieronymus, G. 1896. Plantae Stuebelianae novae. Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie 21:306-378.

Hooghiemstra, H. 1984. Vegetational and climatic history of the high plain of Bogotá, Colombia: a continuous record of the last 3.5 million years. Dissertationes Botanicae 79:1-138.

Hooker, J. D. 1847. Introductory essay. Pages 3 pts *in* The botany of the Antartic voyage, 1839-1843. II. Flora Novae-Zelandiae. Reeve Brothers, London.

Huelsenbeck, J. P., and F. Ronquist. 2001. MrBayes: bayesian inference of phylogenetic trees. Biometrics 17:754-755.

Hugall, A., J. Stanton, and C. Moritz. 1999. Reticulate evolution and the origins of ribosomal internal transcribed spacer diversity in apomictic *Meloidogyne*. Molecular Biology and Evolution 16:157-164.

Jussieux, A. L. d. 1789. Genera plantarum secundum ordines naturales dispositae, Parisiis.

Kalkman, C. 1988. The phylogeny of the Rosaceae. Botanical Journal of the Linnean Society 98:37-59.

Kellogg, E. A. 1990. Variation and species limits in agamospermous grasses. Systematic Botany 15:112-123.

Kessler, M. 1995a. The genus *Polylepis* (Rosaceae) in Bolivia. Candollea 50:131-171.

Kessler, M. 1995b. *Polylepis*-Wälder Boliviens: Taxa, Ökologie, Verbreitung und Geschichte. J. Cramer, Berlin.

Lanfranco, L., M. Delpero, and P. Bonfante. 1999. Intrasporal variability of ribosomal sequences in the endomycorrhizal fungus *Gigaspora margarita*. Molecular Ecology 8:37-45.

Lee, W. G., B. H. Macmillan, T. R. Partridge, R. Lister, and K. M. Lloyd. 2001. Fruit features in relation to the ecology and distribution of *Acaena* (Rosaceae) species in New Zealand. New Zealand Journal of Ecology 25:17-27.

Levy, A. A., and M. Feldman. 2002. The impact of polyploidy on grass genome evolution. Plant Physiology 130:1587-1593.

Linnaeus, C. 1753. Species plantarum. Salvius, Stockholm.

Linnaeus, C. 1771. Mantissa altera plantarum. Salvius, Stockholm.

Lynch, M. 1984. Destabilizing hybridization, general-purpose genotypes and geographic parthenogenesis. Quarterly Review of Biology 59:257-290.

Macklin, J. A., and J. B. Phipps. 2002. Sargent's names in *Crataegus* series *Coccineae* (Rosaceae). Harvard Papers in Botany 7:25-35.

Maddison, D. R., and W. P. Maddison. 2002. MacClade, version 4.05 for OSX. Sinauer Associates, Inc.

Masterson, J. 1994. Stomatal Size in Fossil Plants - Evidence for Polyploidy in Majority of Angiosperms. Science 264:421-424.

Mayden, R. L. 1999. Consilience and a hierarchy of species concepts: Advances toward closure on the species puzzle. Journal of Nematology 31:95-116.

Meidell, J. S., H. L. Oppenheimer, and R. T. Bartlett. 1998. New Plant Records from West Maui. Pages 6-8 *in* Records of the Hawaii Biological Survey for 1997 (G. W. Staples, ed.).

Mena, P., and H. Balslev. 1986. Comparación entre la vegetación de los páramos y el cinturón afroalpino. Reports from the Botanical Institute, University of Aarhus 12:1-54.

Menke, B. 1976. Pliozäne und ältestquartäre Sporen- und Pollenflora von Schleswig-Holstein. Geologisches Jahrbuch Reihe A A 32:3-197.

Miller, J. S., and D. L. Venable. 2000. Polyploidy and the evolution of gender dimorphism in plants. Science 289:2335-2338.

Miller, J. S., and D. L. Venable. 2002. The transition to gender dimorphism on an evolutionary background of self-incompatibilty: An example from *Lycium* (Solanaceae). American Journal of Botany 89:1907-1915.

Mishima, M., N. Ohmido, K. Fukui, and T. Yahara. 2002. Trends in site-number change of rDNA loci during polyploid evolution in *Sanguisorba* (Rosaceae). Chromosoma 110:550-558.

Monasterio, M., and F. Vuilleumier. 1986. Introduction: High tropical mountain biota of the world. Pages 3-7 *in* High altitude tropical biogeography (F. Vuilleumier, and M. Monasterio, eds.). Oxford University Press, Oxford.

Morgan, D. R., D. E. Soltis, and K. R. Robertson. 1994. Systematic and evolutionary implications of *rbcL* sequence variation in Rosaceae. American Journal of Botany 81:890-903.

Nakata, M., and K. Oginuma. 1989. Cytological studies on phanerogams in southern Peru. II. Chromosomes of *Tetraglochin strictum* (Rosaceae) and *Ephedra americana* (Ephedraceae). Bulletin of the National Science Museum: Series B: Botany 15:63-66.

Nordborg, G. 1966. *Sanguisorba* L., *Sarcopoterium* Spach, and *Bencomia* Webb et Berth. Opera Botanica 11:1-103.

Nordborg, G. 1967. The genus *Sanguisorba* section *Poterium*. Opera Botanica 16:1-166.

Olsen, O. 1976. *Polylepis australis* Bitter - en ny hårdfor vedplante for Denmark. Dansk Dendrologisk Årsskrift 5:21-31.

Orchard, A. E. 1969. Revision of the *Acaena ovina* A. Cunn. (Rosaceae) complex in Australia. Transactions of the Royal Society of South Australia 93:91-109.

Ozkan, H., A. A. Levy, and M. Feldman. 2001. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops-Triticum*) group. Plant Cell 13:1735-1747.

Ozkan, H., A. A. Levy, and M. Feldman. 2002. Rapid differentiation of homeologous chromosomes in newly-formed allopolyploid wheat. Israel Journal of Plant Sciences 50:S65-S76.

Paterson, A. H., J. E. Bowers, M. D. Burow, X. Draye, C. G. Elsik, C. X. Jiang, C. S. Katsar, T. H. Lan, Y. R. Lin, R. G. Ming, and R. J. Wright. 2000. Comparative genomics of plant chromosomes. Plant Cell 12:1523-1539.

Pilger, R. 1906. Rosaceae andinae. Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie 37:534-541.

Pinar, N. M., H. N. Büyükkartal, and H. Çölgeçen. 2001. Pollen and seed morphology of diploids and natural tetraploids of *Trifolium pratense* L. (Leguminosae). Acta Biologica Cracoviensia Series Botanica 43:27-32.

Polanco, C., and M. P. Delavega. 1995. Length polymorphism in the ribosomal DNA intergenic spacer of rye and slender wild oats. Journal of Heredity 86:402-407.

Posada, D., and K. Crandall. 1998. Modeltest: Testing the model of DNA substitution. Bioinformatics 14:817-818.

Potter, D., F. Gao, P. E. Bortiri, S. H. Oh, and S. Baggett. 2002. Phylogenetic relationships in Rosaceae inferred from chloroplast *matK* and *trnL-trnF* nucleotide sequence data. Plant Systematics and Evolution 231:77-89.

Potter, D., J. J. Luby, and R. E. Harrison. 2000. Phylogenetic relationships among species of *Fragaria* (Rosaceae) inferred from non-coding nuclear and chloroplast DNA sequences. Systematic Botany 25:337-348.

Rada, F., C. Garcia-Nunez, C. Boero, M. Gallardo, M. Hilal, J. Gonzalez, F. Prado, M. Liberman-Cruz, and A. Azocar. 2001. Low-temperature resistance in *Polylepis tarapacana*, a tree growing at the highest altitudes in the world. Plant, Cell and Environment 24:377-381.

Rada, F., G. Goldstein, A. Azocar, and F. Meinzer. 1985. Daily and seasonal osmotic changes in a tropical treeline species. Journal of Experimental Botany 36:898-1000.

Rauh, W. 1956a. Peruanische Vegetationsbilder, I: Die grossen Gegensatze: Die Küstenwüste und die Walder des Osten. Die Umschau 56:140-143.

Rauh, W. 1956b. Peruanische Vegetationsbilder, II: Die Andenwesthange. Die Umschau 56:171-174.

Raven, P. H. 1975. The bases of angiosperm phylogeny: cytology. Annals of the Missouri Botanical Garden 62:724-764.

Raven, P. H., and D. I. Axelrod. 1974. Angiosperm biogeography and past continental movements. Annals of the Missouri Botanical Garden 61:539-673.

Reitsma, T. 1966. Pollen morphology of some European Rosaceae. Acta Botanica Neerlandica 15:290-379.

Ricklefs, R. E., and D. Schluter. 1993. Introduction to part 4. Historical and phylogenetic perspectives. Pages 241-242 *in* Species diversity in ecological communities. Historical and geographical perspectives (R. E. Ricklefs, and D. Schluter, eds.). The University of Chicago Press, Chicago.

Robertson, K. R. 1974. The genera of Rosaceae in the southeastern United States. Journal of the Arnold Arboretum 55:611-662.

Ronquist, F. 1996. DIVA version 1.1. Computer program and manual available by anonymous FTP from Uppsala University (ftp.uu.se or ftp.systbot.uu.se).

Ruiz, L. H., and J. A. Pavon. 1794. Florae Peruvianae et Chilensis Prodromus. Sancha, Madrid, Spain.

Salgado-Labouriau, M. 1979. Modern pollen deposition in the Venezuelan Andes. Grana 18:53-59.

Sanderson, M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Molecular Biology and Evolution 14:1218-1231.

Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. Molecular Biology and Evolution 19:101-109.

Sanderson, M. J. 2003a. r8s - Analysis of rates ("r8s") of evolution, version 1.60.

Sanderson, M. J. 2003b. r8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19:301-302.

Sang, T., M. J. Donoghue, and D. M. Zhang. 1997. Evolution of alcohol dehydrogenase genes in peonies (*Paeonia*): Phylogenetic relationships of putative nonhybrid species. Molecular Biology and Evolution 14:994-1007.

Sang, T., and D. M. Zhang. 1999. Reconstructing hybrid speciation using sequences of low copy nuclear genes: Hybrid origins of five *Paeonia* species based on *Adh* gene phylogenies. Systematic Botany 24:148-163.

Schluter, D., T. Price, A. Ø. Mooers, and D. Ludwig. 1997. Likelihood of ancestor states in adaptive radiation. Evolution 51:1699-1711.

Schön, I., A. Gandolfi, E. Di Masso, V. Rossi, H. I. Griffiths, K. Martens, and R. K. Butlin. 2000. Persistence of asexuality through mixed reproduction in *Eucypris virens* (Crustacea, Ostracoda). Heredity 84:161-169.

Scopoli, I. A. 1772. Flora Carniolica. I, Vienna.

Scott, L., M. Steenkamp, and P. B. Beaumont. 1995. Palaeoenvironmental conditions in South Africa at the Pleistocene-Holocene transition. Quaternary Science Reviews 14:937-947.

Simpson, B. B. 1974. Glacial migrations of plants: island biogeographical evidence. Science 185:698-700.

Simpson, B. B. 1975. Pleistocene changes in the flora of the high tropical Andes. Paleobiology 1:273-294.

Simpson, B. B. 1979. A revision of the genus *Polylepis* (Rosaceae: Sanguisorbeae). Smithsonian Contributions to Botany 43:1-62.

Simpson, B. B. 1983. An historical phytogeography of the high Andean flora. Revista Chilena de Historia Natural 56:109-122.

Simpson, B. B. 1986. Speciation and specialization of *Polylepis* in the Andes. Pages 304-315 *in* High altitude tropical biogeography (F. Vuillemeir, and M. Monasterio, eds.). New York, Oxford Univ. Press.

Simpson, B. B., and J. L. Neff. 1985. Plants, their pollinating bees, and the great American interchange. Pages 427-452 *in* The great American biotic interchange (F. G. Stehli, and S. D. Webb, eds.). Plenum Press, New York.

Simpson, B. B., and C. A. Todzia. 1990. Patterns and processes in the development of the high Andean flora. American Journal of Botany 77:1419-1432.

Sinnott, Q. P., and J. B. Phipps. 1983. Studies in *Crataegus* L. (Rosaceae: Maloideae). IV. Variation patterns in *Crataegus* series *Pruinosae* (Rosaceae) in southern Ontario. Systematic Botany 8:59-70.

Small, R. L., J. A. Ryburn, and J. F. Wendel. 1999. Low levels of nucleotide diversity at homoeologous *Adh* loci in allotetraploid cotton (*Gossypium* L.). Molecular Biology and Evolution 16:491-501.

Small, R. L., and J. F. Wendel. 2000a. Copy number lability and evolutionary dynamics of the *Adh* gene family in diploid and tetraploid cotton (*Gossypium*). Genetics 155:1913-1926.

Small, R. L., and J. F. Wendel. 2000b. Phylogeny, duplication, and intraspecific variation of *Adh* sequences in new world diploid cottons (*Gossypium* L., Malvaceae). Molecular Phylogenetics and Evolution 16:73-84.

Small, R. L., and J. F. Wendel. 2002. Differential evolutionary dynamics of duplicated paralogous *Adh* loci in allotetraploid cotton (*Gossypium*). Molecular Biology and Evolution 19:597-607.

Smith, J. M. B., and A. M. Cleef. 1988. Composition and origins of the world's tropicalpine flora. Journal of Biogeography 15:631-645.

Soltis, P. S., and D. E. Soltis. 1991. Multiple origins of the allotetraploid *Tragopogon mirus* (Compositae): rDNA evidence. Systematic Botany 16:407-413.

Song, K. M., P. Lu, K. L. Tang, and T. C. Osborn. 1995. Rapid Genome Change in Synthetic Polyploids of *Brassica* and Its Implications for Polyploid Evolution. Proceedings of the National Academy of Sciences of the United States of America 92:7719-7723.

Spach, E. 1846. Revisio generis *Poterium*. Annales des Sciences Naturelles 3. Ser. Botanique 5:31-43.

Stebbins, G. L. 1950. Variation and evolution in plants. Columbia University Press, New York.

Storey, B. C. 1995. The role of mantle plumes in continental breakup: Case histories from Gondwanaland. Nature 377:301-308.

Swofford, D. L. 2002. PAUP* (Phylogenetic Analysis Using Parsimony (*and Other Methods)), version 4.0b10. Sinauer Associates, Inc.

Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three noncoding regions of chloroplast DNA. Plant Molecular Biology 17:1105-1109.

Talent, N., and T. A. Dickinson. 2002. Fruit and seed formation from inter-ploidy and intersectional pollination in *Crataegus* (Rosaceae). Pages 170 *in* Abstracts Proceedings of the Botany 2002, Madison, WI, August 2-7, 2002.

Thorne, J. L., H. Kishino, and I. S. Painter. 1998. Estimating the rate of evolution of the rate of molecular evolution. Molecular Biology and Evolution 15:1647-1657.

Trimen, J. 1879. Journal of Botany 17:97.

Troll, C. 1959. Die tropische Gebirge. Bonner Geographische Abhandlungen 25:1-93.

Troll, C. 1968. The cordilleras of the tropical Americas. Pages 15-56 *in* Geo-ecology of the mountainous regions of the tropical Americas (C. Troll, ed.) Dummlers, Bonn.

Urbatsch, L. E., B. G. Baldwin, and M. J. Donoghue. 2000. Phylogeny of the coneflowers and relatives (Heliantheae : Asteraceae) based on nuclear rDNA internal transcribed spacer (ITS) sequences and chlorplast DNA restriction site data. Systematic Botany 25:539-565.

Vahl, M. 1805. Enumeratio Plantarum. N. Mölleri et Filii, Hauniae.

Van der Hammen, T. 1974. The Pleistocene changes of vegetation and climate in tropical South America. Journal of Biogeography 1:3-26.

Van der Hammen, T. 1979. History of flora, vegetation and climate in the Colombian Cordillera Oriental during the last five million years. Pages 25-32 *in* Tropical botany. Proceedings of a symposium held at the University of Aarhus on 10-12 August 1978 (K. Larsen, and L. B. Holm-Nielsen, eds.). Academic Press, London.

Van der Hammen, T., and A. M. Cleef. 1986. Development of the high páramos flora and vegetation. Pages 153-201 *in* High altitude tropical biogeography (F. Vuilleumier, and M. Monasterio, eds.). Oxford University Press, New York.

Van der Hammen, T., and H. Hooghiemstra. 1997. Chronostratigraphy and correlation of the Pliocene and Quaternary of Colombia. Quaternary International 40:81-91.

Van der Hammen, T., J. H. Werner, and H. van Dommelen. 1973. Palynological record of the upheaval of the northern Andes: A study of the Pliocene and lower Quaternary of the Colombian Eastern Cordillera and the early evolution of its high Andean biota. Review of Palaeobotany and Palynology 16:1-122.

van Dijk, P., and J. van Damme. 2000. Apomixis technology and the paradox of sex. Trends in Plant Science 5:81-4.

van Dijk, P. J. 2003. Ecological and evolutionary opportunities of apomixis: insights from *Taraxacum* and *Chondrilla*. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 358:1113-1120.

Vuilleumier, B. S. 1971. Pleistocene changes in the fauna and flora of South America. Science 173:771-779.

Vuilleumier, F. 1970. Insular biogeography in continental regions .1. Northern Andes of South-America. American Naturalist 104:373-&.

Vuilleumier, F. 1984. Faunal turnover and development of fossil avifaunas in South America. Evolution 38:1384-1396.

Vuilleumier, F. 1985. Fossil and recent avifaunas and the Interamerican interchange. Pages 387-424 *in* The great American biotic interchange (F. G. Stehli, and S. D. Webb, eds.). Plenum, New York.

Walter, A. E., D. H. Turner, J. Kim, M. H. Lyttle, P. Muller, D. H. Mathews, and M. Zuker. 1994. Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. Proceedings of the National Academy of Sciences of the United States of America 91:9218-9222.

Walter, H., and E. Medina. 1969. La temperatura del suelo como factor determinante para la caracterización de los pisos subalpino y alpino en los Andes de Venezuela. Boletin Venezolana de Ciencias Naturales 115/116:201-210.

Waters, E. R., and B. A. Schaal. 1996. Biased gene conversion is not occurring among rDNA repeats in the *Brassica* triangle. Genome 39:150-154.

Webb, D., and C. Berthelot. 1836-50. Historie Naturelle des Isles Canaries, Paris.

Webb, S. D. 1991. Ecogeography and the Great American Interchange. Paleobiology 17:266-280.

Weimarck, H. 1934. Monograph of the genus *Cliffortia*. Gleerupska Univ. Bokhandeln, Lund.

Whelan, S., and N. Goldman. 1999. Distributions of statistics used for the comparison of models of sequence evolution in phylogenetics. Molecular Biology and Evolution 16:1292-1299.

White, T. J., T. Bruns, S. Lee, and J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pages 315-322 *in* PCR protocols: a guide to methods and applications (M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White, eds.). Academic Press, London.

Wong, S., G. Butler, and K. H. Wolfe. 2002. Gene order evolution and paleopolyploidy in hemiascomycete yeasts. Proceedings of the National Academy of Sciences of the United States of America 99:9272-9277.

Wright, S. D., C. G. Yong, J. W. Dawson, D. J. Whittaker, and R. C. Gardner. 2000. Riding the ice age El Nino? Pacific biogeography and evolution of *Metrosideros* subg. *Metrosideros* (Myrtaceae) inferred from nuclear ribosomal DNA. Proceedings of the National Academy of Sciences of the United States of America 97:4118-4123.

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

Zhang, Q. F., M. A. S. Maroof, and P. G. Yang. 1992. Ribosomal DNA polymorphisms and the oriental-occidental genetic differentiation in cultivated barley. Theoretical and Applied Genetics 84:682-687.