Circumscription and phylogeny of Apiaceae subfamily Saniculoideae based on chloroplast DNA sequences

Carolina I. Calviño a,b,*, Stephen R. Downie a

a Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801-3707, USA
b Instituto de Botánica Darwinion, Buenos Aires, Argentina

Received 14 July 2006; revised 3 January 2007; accepted 4 January 2007

Abstract

An estimate of phylogenetic relationships within Apiaceae subfamily Saniculoideae was inferred using data from the chloroplast DNA trnQ-trnK 5′-exon region to clarify the circumscription of the subfamily and to assess the monophyly of its constituent genera. Ninety-one accessions representing 14 genera and 82 species of Apiaceae were examined, including the genera Steganoaenia, Polemanniopsis, and Lichtensteinia which have been traditionally treated in subfamily Apioideae but determined in recent studies to be more closely related to or included within subfamily Saniculoideae. The trnQ-trnK 5′-exon region includes two intergenic spacers heretofore underutilized in molecular systematic studies and the rps16 intron. Analyses of these loci permitted an assessment of the relative utility of these noncoding regions (including the use of indel characters) for phylogenetic study at different hierarchical levels. The use of indels in phylogenetic analyses of both combined and partitioned data sets improves resolution of relationships, increases bootstrap support values, and decreases levels of overall homoplasy. Intergeneric relationships derived from maximum parsimony, Bayesian, and maximum likelihood analyses, as well as from maximum parsimony analysis of indel data alone, are fully resolved and consistent with one another and generally very well supported. We confirm the expansion of subfamily Saniculoideae to include Steganoaenia and Polemanniopsis (as the new tribe Steganotaenieae C.I. Calviño and S.R. Downie) but not Lichtensteinia. Sister group to tribe Steganotaenieae is tribe Saniculeae, redefined to include the genera Actinolema, Alepidea, Arctopus, Astrantia, Eryngium, Petagnaea, and Sanicula. With the synonymization of Hacquetia into Sanicula, all genera are monophyletic. Eryngium is divided into “Old World” and “New World” subclades and within Astrantia sections Astrantia and Astrantiella are monophyletic.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Apiaceae; Saniculoideae; cpDNA trnQ-trnK 5′-exon; Phylogeny; Indels

1. Introduction

Apiaceae subfamily Saniculoideae, as treated by Drude (1898) and Wolff (1913), comprises two tribes (Saniculeae and Lagoecieae), nine genera (Actinolema Fenzl, Alepidea F. Delarocche, Arctopus L., Astrantia L., Eryngium L., Hacquetia Neck. ex DC., Lagoecia L., Petagnaea Caruel, and Sanicula L.), and approximately 330 species. The subfamily has a bipolar distribution, but is better represented in the southern hemisphere than its sister group, subfamily Apioideae (Mathias, 1971; Downie et al., 2001). Sanicula and Eryngium are each cosmopolitan and are the only genera of Saniculoideae represented in the western hemisphere. They also account for the majority of the species of the subfamily, with 39 and about 250 species, respectively (Pimenov and Leonov, 1993). Hacquetia, Lagoecia, and Petagnaea are each monotypic and occur in Europe and/or Asia, as do Actinolema (two species) and Astrantia (nine species; Pimenov and Leonov, 1993). Arctopus and Alepidea, with three and 20 species each, occur in Africa. The subfamily is of some economic and ecological importance. Several species

* Corresponding author. Address: Department of Plant Biology, 265 Morrill Hall, 505 South Goodwin Avenue, University of Illinois at Urbana-Champaign, Urbana, IL 61801-3707, USA. Fax: +1 217 244 7246.
E-mail address: ccalvino@life.uiuc.edu (C.I. Calviño).

1055-7903/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2007.01.002

are endangered, some are used for culinary or medicinal purposes, many are ornamentals, and others are noxious weeds. The plants are mostly herbaceous, with often spiny or bristly simple leaves. Their flowers are arranged primarily in simple (rarely compound) umbels or heads that are surrounded by showy bracts. Their fruits comprise an exocarp covered in scales, bristles, or prickles (or rarely are glabrous or tuberculare), a mesocarp with calcium-oxide crystals scattered throughout, and a parenchymatous endocarp. The most common base chromosome numbers of the subfamily are \( x = 7 \) and \( x = 8 \).

Since the treatments of Saniculoideae by Drude (1898) and Wolff (1913), the composition of the subfamily has changed only slightly. Lagoezia has been transferred to subfamily Apiioideae (Plunkett et al., 1996; Downie et al., 2000a; Valiejo-Roman et al., 2002), its affinity to the apioid umbellifers having been suggested previously (Koso-Poljansky, 1916; Cerceau-Larrival, 1962; Tseng, 1967; Guyot, 1971; Magin, 1980). Arctopus was transferred to and maintained within the traditionally circumscribed Apiaceae subfamily Hydrocotylinoideae (Froebe, 1964; Magin, 1980; Constance and Chuaug, 1982; Pimenov and Leonov, 1993), but later returned to Saniculoideae upon analyses of molecular data (Plunkett and Lowry, 2001; Chandler and Plunkett, 2004). Oligoocladius Chodat and Wilczek, provisionally included in subfamily Saniculoideae by Pimenov and Leonov (1993), finds affinity among the higher apioid umbellifers (C.I. Calviño and S.R. Downie, unpublished data). The most dramatic change in circumscription of Saniculoideae, however, was the recent addition of the African apioid genera Steganotaenia Hochst., Polemanniopsis B. L. Burtt, and Lichtensteinia Cham. and Schltdl., primarily on the basis of cladistic analysis of fruit anatomical characters (Liu et al., 2003). The sister group relationship between Steganotaenia and Polemanniopsis and subfamily Saniculoideae was first revealed by Downie and Katz-Downie (1999) using chloroplast DNA (cpDNA) \( rps16 \) intron sequences. Their results suggested that Steganotaenia and Polemanniopsis be removed from subfamily Apiioideae, but they were uncertain whether the circumscription of Saniculoideae should be expanded to include these genera. Steganotaenia and Polemanniopsis have morphological features similar to subfamily Saniculoideae, such as intrajugal vittae (oil ducts in the ribs of the fruits that are associated with vascular bundles) and the absence of commissural and vallecular vittae (oil ducts in the commissure and furrows of the fruits, respectively), but characters reminiscent of subfamily Apiioideae are also apparent, such as large compound umbels inflorenceses. A subsequent molecular phylogenetic study of southern African Apiaceae also indicated a sister group relationship between Steganotaenia/Polemanniopsis and subfamily Saniculoideae (Calviño et al., 2006). The genus Lichtensteinia, however, comprised a monogeneric clade sister group to all other members of subfamily Apiioideae. Furthermore, Calviño et al. (2006) reported that, in the study of Liu et al. (2003), the fruit characters uniting Lichtensteinia and Steganotaenia/Polemanniopsis with subfamily Saniculoideae were plesiomorphic in the family, thus the inclusion of these three genera within an expanded Saniculoideae was influenced by these symplesiomorphies. The placement of Steganotaenia and Polemanniopsis into an expanded Saniculoideae was not implemented by Calviño et al. (2006), given that the only evidence clearly justifying the sister group relationship between these genera and Saniculoideae is that of the \( rps16 \) intron and these sequence data supported the relationship only weakly.

While molecular data have been useful to corroborate the monophyly of subfamily Saniculoideae and reveal its sister group relationship to subfamily Apiioideae, no molecular systematic study to date has focused explicitly on infrasemifamilial relationships of all of its genera and the phylogenetic placements of Steganotaenia, Polemanniopsis, and Lichtensteinia remain uncertain. Furthermore, a recent study of Apiaceae using nuclear ribosomal DNA (rDNA) internal transcribed spacer (ITS) sequences has suggested that Eryngium is paraphyletic and that Hacquetia should be treated as part of Sanicula (Valiejo-Roman et al., 2002). These phylogenetic hypotheses, however, were based exclusively on ITS sequence comparisons, and among distant members of Apiaceae alignment of these sequences is highly problematic. Moreover, several molecular genetic processes impact ITS sequences in ways that may confound phylogenetic inference (Álvarez and Wendel, 2003), such as the divergent paralogous ITS sequences detected in a few members of the early branching Annesorhiza clade within subfamily Apiioideae (Calviño et al., 2006).

The major objective of this study is to estimate phylogenetic relationships within Apiaceae subfamily Saniculoideae using molecular data. We examine the cpDNA \( trnQ-trnK \) 5’-exon region (hereafter, called \( trnQ-trnK \)), a region encompassing primarily three large noncoding loci (i.e., the \( trnQ-rps16 \) intergenic spacer, \( rps16 \) intron, and \( rps16-trnK \) intergenic spacer; Fig. 1). Over the past decade, the \( rps16 \) intron has been used increasingly in phylogenetic studies of both Apiaceae and other angiosperms (Lidén et al., 1997; Oxelman et al., 1997; Downie and Katz-Downie, 1999; Kelchner, 2002; Shaw et al., 2005), but the spacer regions flanking the \( rps16 \) gene have been rarely considered for such a purpose (Hahn, 2002). A recent study of two apioid genera using these intergenic spacers established their higher rate of molecular evolution (in both nucleotide substitutions and length mutations) over other cpDNA loci used in molecular systematic investigations of Apiaceae (Lee and Downie, 2006). Herein, we examine the efficacy of this region in resolving phylogeny at different taxonomic levels within the family, including the use of indels for phylogeny estimation. Ancillary objectives include clarification of the circumscription of subfamily Saniculoideae (with emphasis on the phylogenetic placements of Steganotaenia, Polemanniopsis, and Lichtensteinia) and an assessment of the monophyly of its constituent genera. We present the first explicit phylogenetic hypothesis for the subfamily and a revised classification that reflects this phylogeny.
The gene rps16 is interrupted by an intron, and only the 5′-exon of gene trnK is shown. The sizes of the two intergenic spacer regions and intron are presented in base pairs (bp). Scale bar is 100 bp unit. The arrows represent the directions and approximate positions of the primers used in PCR amplification and/or DNA sequencing. Forward primers are designated 1–10; reverse primers are designated A–I. These primers correspond DNA accession number and voucher information. Intron data for 13 of these accessions were obtained previously (Downie and Katz-Downie, 1996, 1999; Calviño et al., 2006). For those accessions, the modified hexadecyltrimethylammonium bromide (CTAB) protocol of Doyle and Doyle (1987) was used instead, as detailed in Downie and Katz-Downie (1996, 1999).

The region bounded by and including chloroplast genes trnQ and trnK 5′-exon and containing the rps16 intron is 3117 bp in size in tobacco (Shinozaki et al., 1986). Flanking the gene rps16 are the trnQ-rps16 and rps16-trnK intergenic spacer regions and in tobacco cpDNA these spacers are 1204 and 686 bp in size, respectively (Fig. 1). The strategies employed to obtain these sequence data are presented elsewhere (Downie and Katz-Downie, 1996, 1999; Calviño et al., 2006), with only slight modifications herein. We first performed a long-PCR on a few accessions of Eryngium, using primers anchored in genes trnQ and trnK 5′-exon that were constructed by comparisons of published gene sequences from tobacco and rice and choosing regions highly conserved between them (Shinozaki et al., 1986; Hiratsuka et al., 1989). All other primers were subsequently designed based on these Eryngium data. In total, 19 primers were used for PCR and/or DNA sequencing of the entire trnQ-trnK region (Fig. 1). Four of these primers were constructed specifically for Alepidea because primers trnQ-3 (primer 5) and trnQ-1R (primer D) could not be used due to a 238-bp deletion unique to this genus, and primer trnQ-2 (primer 3) did not anneal to Alepidea (and other African members of Saniculoideae, and Apioidae) because of a point mutation at the extreme 3′-end of the primer binding site. For PCR
amplifications of the trnQ-rps16 intergenic spacer using any of the primers internal to this region, the annealing temperature was decreased from 53 to 48°C and the MgCl₂ concentration was decreased from 2.75 to 1.5 mM. For sequencing of the rps16-trnK spacer for those DNAs where homopolymer regions or secondary structure formation obstructed the reaction, 1 μl of dGTP BigDye terminator (ABI Prism® dGTP BigDye™ Terminator v3.0) was added to the standard sequencing cocktails. Simultaneous consideration of both DNA strands across the entire region for most taxa permitted unambiguous base determination. GenBank reference numbers for all sequences are presented in the online Supplementary Appendix A.

2.3. Sequence comparisons and phylogenetic analyses

Sequence chromatograms were edited manually using Se-Al (Rambaut, 2002). DNA sequences were aligned initially using the default pairwise and multiple alignment parameters in the computer program CLUSTAL X (gap opening cost = 15.00, gap extension cost = 6.66, DNA transition weight = 0.50; Jeanmougin et al., 1998) then rechecked and adjusted manually as necessary. Gaps were positioned to minimize nucleotide mismatches. A matrix of binary-coded indels was constructed to incorporate length-mutational information into the phylogenetic analysis. Gaps of equal length in more than one sequence were coded as the same presence or absence character state if they could not be interpreted as different duplication or insertion events. Indels of similar location but with different lengths were coded as multiple binary characters. In several regions, gap coding was problematic because of homopolymers or indirect duplications of adjacent elements in two or more taxa. These gaps were not scored and these ambiguous regions were excluded from subsequent analysis.

Some regions of the alignment were scored as missing. Only data for the rps16 intron were available for Polemoniopsis marlothii (DNA Accession No. 1333; online Supplementary Appendix A) and Steganaenia araliaacea (Accession Nos. 1373 and 1385) because these accessions were used in a prior study and their DNAs were no longer available (Downie and Katz-Downie, 1999). Data for the trnQ-rps16 spacer and about half of the rps16-trnK spacer could not be obtained for Arctopus echinatus 2559 despite our repeated but unsuccessful attempts to PCR-amplify these regions. Similarly, parts of both spacers in Sanicula chinensis could not be PCR-amplified. Portions of the rps16 3′-exon were missing data (between primers G and 9; Fig 1), attributable to the positions of the primers anchored in this exon used to amplify the regions flanking it. However, this exon had little to no variation among all other accessions, hence the absence of these data did not affect the phylogenetic results. Overall, missing data represented 5.5% of the entire matrix.

Boundaries of the genes trnQ, rps16, and trnK 5′-exon were determined by comparison of the DNA sequences to corresponding boundaries in tobacco cpDNA (Shinozaki et al., 1986). The determination of boundary sequences for the six major structural domains of the rps16 group II intron was based on similar boundary sequences inferred for tobacco, mustard, and other Apiaceae (Michel et al., 1989; Neuhaus et al., 1989; Downie and Katz-Downie, 1999). Characterization of the three cpDNA regions and six rps16 intron structural domains was facilitated using MacClade version 4.07 (Maddison and Maddison, 2005), BioEdit version 6.0.7 (Hall, 1999), and PAUP version 4b10 (Swofford, 2002). Uncorrected pairwise nucleotide distances of unambiguously aligned positions were determined using the distance matrix option of PAUP*. Average sequence divergence estimates were calculated both within and between genera. Relative evolutionary rates were estimated by plotting the pairwise corrected genetic distance (according to the model selected by Modeltest version 3.7; Posada and Crandall, 1998) of one region versus the other. The slope of the linear regression was taken as the relative rate value (Pochon et al., 2006). Because both variables are independent or subject to natural variability, we performed a Model II (geometric mean) linear regression and not the standard Model I. The Model II method (for equations see Barker et al., 1988) effectively minimizes the sum of the squares of the deviations of the observations from the line in both axis by measuring the offsets along a line perpendicular to the regression line.

The data matrices were each analyzed using maximum parsimony (MP) as implemented by PAUP*. For matrices representing either the entire trnQ-trnK region, the entire region plus scored indels, or only scored indels, heuristic searches were performed for 100,000 replicates with random addition of taxa and tree-bisection-reconnection (TBR) branch swapping. For matrices of partitioned data (i.e., intron or spacer regions), with and without indels, the heuristic search strategies employed by Calviño et al. (2006) were followed. Bootstrap values were calculated from 100,000 replicate analyses using “fast” stepwise-addition of taxa and only those values compatible with the 50% majority-rule consensus tree were recorded. The number of additional steps required to force particular taxa into a monophyletic group was examined using the constraint option of PAUP*.

The relative utility of the three cpDNA regions in resolving phylogenetic relationships and the effect of incorporating length-mutational events into the phylogenetic analyses were assessed by comparing the results of MP analysis of each data partition, with or without indels included as additional characters, against those major clades inferred from MP analysis of the entire trnQ-trnK region plus binary-scored alignment gaps (because analysis of the latter yielded trees of greatest resolution and highest bootstrap support overall). Comparisons were made of the number of major clades recovered in each of these analyses and their corresponding bootstrap support values (Felsenstein, 1985). Additional comparative data included the number of parsimony informative characters, the number and length
of maximally parsimonious trees, and measures of relative homoplasy. In comparing the consistency and retention indices of each data partition, with or without indels, each group of characters was optimized onto the most parsimonious trees inferred from analysis of the entire \( trnQ-trnK \) region plus binary-scored alignment gaps.

Bayesian inference of the entire \( trnQ-trnK \) region (indel characters not included) was conducted using MrBayes version 3.1.1 (Huelsenbeck and Ronquist, 2001). The program was run in parallel on an IBM pSeries 690 system at the National Center for Supercomputing Applications at UIUC. Prior to analysis, Modeltest was used to select an evolutionary model of nucleotide substitution that best fits these data, as selected by the Akaike Information Criterion estimator. The settings appropriate for the best-fit TVM+I+G model were put into a MrBayes block in PAUP* (nst = 6, rates = invgamma). The priors on state frequencies and rates and variation across sites (shape of the gamma distribution) were estimated automatically from the data assuming no prior knowledge about their values (Dirichlet default option). From different random starting trees, four independent Bayesian analyses (nruns = 4) were run for 10 million generations and the trees saved to a file every 100 generations (i.e., a total of 400,000 trees was sampled). Four simultaneous Markov chain Monte Carlo (MCMC) chains were used and branch lengths of the trees were saved. Variation in likelihood scores to determine apparent stationarity was examined graphically for each independent run using the program Tracer version 1.2.1 (A. Rambaut and A. Drummond, University of Oxford, unpublished data). The states of the chain that were sampled before stationarity (i.e., the “burn in” of the chain) were discarded and the posterior probability values for each bipartition of the phylogeny were determined from the remaining trees. To summarize and compare the samples from each analysis, the sump and sumt commands of MrBayes were used. MCMC convergence was also explored by examining the potential scale reduction factor (PSRF) convergence diagnostics for all parameters in the model (provided by the sump and sumt commands) and graphically using the cumulative, compare, and absolute difference options of the program AWTY online (Wagenbusch et al., 2004).

The data matrix of the entire \( trnQ-trnK \) region (indel characters not included) was also analyzed using the maximum likelihood method as implemented by PAUP*. The results obtained were congruent to those inferred by the Bayesian analysis; hence they will not be discussed further.

3. Results

3.1. Sequence comparisons

Sequence characteristics of each of the three cpDNA data partitions are presented in Table 1. On average, the size of the \( trnQ-rps16 \) intergenic spacer in Saniculoideae is smaller than that of outgroups Apioideae and \( Hermas \) as a result of several large deletions. Sizes of the \( rps16 \) intron and \( rps16-trnK \) intergenic spacer are approximately the same between ingroup and outgroup taxa. Alignment of all partitioned regions for 91 accessions of Apiaceae resulted in a matrix of 4846 positions. Of these, 445 were excluded from the analysis because of alignment ambiguities. The remaining 4401 aligned positions yielded 871 parsimony informative characters. In addition, 322 unambiguous alignment gaps were inferred, of which 189 were parsimony informative. The latter ranged in size from 1 to 1009 bp and their frequency in relation to size for each data partition is shown in Fig. 2. Most indels were 10 bp or shorter in size. The average size of insertion across all three regions was 6 bp, whereas the average size of deletion across all regions ranged from 5 to 56 bp as a result of several large deletions in the \( trnQ-rps16 \) inter-
genic spacer (black bars in Fig. 2). Relative to the outgroup *Hermas*, these 189 informative gaps represent a minimum of 77 deletion and 112 insertion events. For each data partition, the ratio of informative substitutions to informative indels is similar (4.4–4.9:1); when the three regions were considered collectively, this ratio is 4.6:1.

Of the three data partitions, the *trnQ-rps16* region displays the most parsimony informative characters (i.e., nucleotide substitutions plus indels; Table 1). The two intergenic spacers had high levels of sequence divergence at all taxonomic levels considered, with maximum divergence values of 3.1% among saniculoid congeners (*rps16-trnK*), 10.8% between saniculoid genera (*trnQ-rps16*), and 14.0% across subfamilies. The slopes of the linear regressions in pairwise comparisons indicate that the intergenic spacers are evolving about twice as fast as that of the *rps16* intron (*trnQ-rps16* vs. *rps16*, $m=1.75$, $R^2=0.9189$; *trnQ-rps16* vs. *rps16-trnK*, $m=1.01$, $R^2=0.8402$; *rps16-trnK* vs. *rps16*, $m=1.73$, $R^2=0.8191$). No significant differences were observed in the relative evolutionary rates between the two intergenic spacer regions.

For each of the six major structural domains of the cpDNA group II *rps16* intron, characteristics of the aligned sequences are presented in Table 2. Domain I is the largest, ranging between 475 and 503 bp among Saniculoideae, whereas domains V (34 bp) and VI (24–38 bp) are the smallest. Domains V and VI are also the most conserved evolutionarily, with few informative positions, low nucleotide sequence divergence, and very few or no alignment gaps. The small sizes of domains III and VI in Saniculoideae relative to the outgroup taxa are due to two deletions of 21 and 13 bp, respectively. These deletions occur in all non-African saniculoid taxa and in “Old World” *Eryngium*, respectively.

### 3.2. Phylogenetic analyses

MP analysis of 4401 unambiguously aligned positions representing the entire *trnQ-trnK* region and 189 binary-
scored informative indels resulted in 4320 minimal length trees, each of 2249 steps (consistency indices, CIs = 0.7781 and 0.7215, with and without uninformative characters, respectively; retention index, RI = 0.9487). The strict consensus of these trees is presented in Fig. 3. No relationship was apparent between size of indel and its level of homoplasy, with the exception that all indels 16 bp or greater in size were not homoplastic when considered across all MP trees (Fig. 4). Indeed, 82% of all indels were not homoplastic when optimized on these trees. The percentage of homoplastic indels from each of the three data partitions was similar, ranging from 11 to 16%. Twenty-four major clades were identified on the strict consensus tree and are described in Table 3. These clades represent a variety of taxonomic levels, such as sections and other infrageneric groupings, genera, subtribes, tribes, and subfamilies. Sup-

port for each of these clades is generally quite strong, with bootstrap values ranging between 72 and 100% (averaging 96%). The same 24 major clades were recovered when indels were excluded from the analysis; bootstrap values on the resultant strict consensus tree (not shown) ranged between 68 and 100% (averaging 93%; Table 3). For some clades (i.e., Nos. 3, 4, and 18; Table 3), bootstrap support levels decreased considerably upon the exclusion of indel characters. The results of MP analyses of partitioned data, with or without indels as additional characters, and their comparisons to the results of the aforementioned analyses are presented in Table 3. The trnQ-trnK data partition plus indels was the only matrix that recovered all major clades inferred by analysis of the entire trnQ-trnK region plus indels.

Table 3
Comparison of bootstrap support values calculated from MP analysis of combined or partitioned data, with and without their corresponding binary-coded indel matrices, for the 24 major clades of Apiaceae identified in Fig. 3 and described here

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Clade</th>
<th>trnQ-trnK+indels</th>
<th>trnQ-trnK</th>
<th>trnQ-trnK+indels</th>
<th>trnQ-trnK</th>
<th>trnQ-trnK+indels</th>
<th>trnQ-trnK</th>
<th>trnQ-trnK+indels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subfamilial 1</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>87</td>
<td>79</td>
<td>84</td>
<td>78</td>
<td>77</td>
<td>63</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>97</td>
<td>77</td>
<td>95</td>
<td>74</td>
<td>70</td>
<td>59</td>
<td>32</td>
</tr>
<tr>
<td>Tribal</td>
<td>4</td>
<td>94</td>
<td>82</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>78</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>96</td>
<td>94</td>
<td>96</td>
<td>89</td>
<td>94</td>
<td>86</td>
<td>83</td>
</tr>
<tr>
<td>Subtribal</td>
<td>6</td>
<td>91</td>
<td>83</td>
<td>57</td>
<td>53</td>
<td>95</td>
<td>90</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>100</td>
<td>98</td>
<td>92</td>
<td>92</td>
<td>94</td>
<td>94</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>72</td>
<td>68</td>
<td>51</td>
<td>47</td>
<td>38</td>
<td>33</td>
<td>55</td>
</tr>
<tr>
<td>Generic</td>
<td>11</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>100</td>
<td>99</td>
<td>n/a</td>
<td>99</td>
<td>97</td>
<td>94</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>100</td>
<td>99</td>
<td>99</td>
<td>88</td>
<td>79</td>
<td>79</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>100</td>
<td>99</td>
<td>99</td>
<td>84</td>
<td>80</td>
<td>83</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>85</td>
<td>72</td>
<td>55</td>
<td>54</td>
<td>84</td>
<td>72</td>
<td>46</td>
</tr>
<tr>
<td>Infrageneric</td>
<td>19</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>99</td>
<td>96</td>
<td>88</td>
<td>91</td>
<td>78</td>
<td>78</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>96</td>
<td>93</td>
<td>63</td>
<td>63</td>
<td>65</td>
<td>64</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>91</td>
<td>91</td>
<td>63</td>
<td>63</td>
<td>60</td>
<td>55</td>
<td>26</td>
</tr>
<tr>
<td>Average</td>
<td>96</td>
<td>93</td>
<td>89</td>
<td>86</td>
<td>85</td>
<td>83</td>
<td>81</td>
<td>74</td>
</tr>
</tbody>
</table>
$rps16$ plus indels matrix recovered all major clades but one (a subtribe, clade No. 10), whereas the $rps16$-$trnK$ data partition plus indels performed poorest, with several major clades not recovered in the strict consensus tree derived from these data. Considering the three data partitions, bootstrap support values are generally highest for the $trnQ$-$rps16$ plus indels matrix (averaging 89%), but for several major clades these values are higher for one of the other two regions. In general, the incorporation of indels into an analysis resulted in higher bootstrap support values than when they were not included, and the simultaneous analysis of data from all three regions plus binary-scored alignment gaps resulted in trees of greatest resolution and highest bootstrap support values (Table 3). This was also true when compared to results of MP analyses of any combination of two regions (data not shown).

MP analysis of the matrix representing only the 189 binary-scored informative indels from the entire $trnQ$-$trnK$...
region resulted in 2016 minimal length trees, each of 217 steps (CI = 0.8710, RI = 0.9842). The strict consensus of these trees (Fig. 5) is highly consistent with the strict consensus tree inferred using both nucleotide substitutions and indels (Fig. 3), with most major clades recovered. For the former, bootstrap support values ranged from 51 to 100%, averaging 81%. To reveal the distribution of indels throughout the phylogeny, the pattern of indel distribution along one arbitrarily selected minimal length tree was mapped onto the strict consensus tree inferred from indel data. Patterns of indel distribution are important in supporting the sister group relationship between Saniculoideae and subfamily Saniculoideae, as well as the sister group relationship between Lichtensteina and other members of subfamily Apioidae. Two unique indels (a 6-bp insertion in trnQ-rps16 and a 12-bp insertion in rps16 intron domain I) are shared by Steganotaeina, Polemanniospis, and Polemanniospis. Four to five synapomorphic indels support the branch leading to Steganotaeina, Polemanniospis, and Saniculoideae, depending upon the reconstruction. Four of these (representing 4- and 148-bp deletions, and 4- and 5-bp insertions) occur within the trnQ-rps16 intergenic spacer region; a single 1-bp deletion occurs within rps16 intron domain IV. Three of these deletions, however, are nested within larger deletions in other saniculoid taxa. As an example, the 148-bp deletion in Steganotaeina, Polemanniospis, and Polemanniospis, Alepidea, Arctopus, Actinolema, and Astrantia cpDNAs is nested within a 272-bp deletion in Sanicula, Hacquetia, and Eryngium and a 1009-bp deletion in Petagna. Four synapomorphic indels support the clade of Lichtensteina, Anginon, and Annesorhiza (representing a single 2-bp deletion and 3-, 4-, and 5-bp insertions). Not a single indel supports the union of Lichtensteina with Steganotaeina and Polemanniospis or with any member of subfamily Saniculoideae. Each of the genera Eryngium, Petagna, Astrantia, Actinolema, Arctopus, Alepidea, and Steganotaeina are supported by synapomorphic indels; the genera Sanicula and Hacquetia unite as a clade, supported by nine synapomorphic indels. No autapomorphic indel was found for the monotypic genus Polemanniospis.

A comparison of the number and length of minimal length trees resulting from MP analysis of combined or separate data partitions, with or without indels, or with just binary-scored indels, is presented in Table 4. Analyses of partitioned data results in the preset maximum tree limit of 20,000 trees, whereas MP analyses of the entire trnQ-trnK region, with and without indels, or of just the indels matrix alone, results in a lower and definite number of trees. Comparisons of overall homoplasy, calculated by optimization of each data partition (with or without indels), the entire trnQ-trnK region (without indels), or the matrix of indel characters onto the MP trees inferred by analysis of the trnQ-trnK plus indels matrix, reveal that the indels matrix had the least level of homoplasy. In general, when indels were incorporated into the phylogenetic analysis, CI and RI values increased. The trnQ-rps16 plus indels matrix had the lowest level of homoplasy (CI = 0.7514, RI = 0.9547), whereas the rps16-trnK plus indels matrix was the most homoplastic (CI = 0.6794, RI = 0.9387).

The four independent Bayesian analyses showed MCMC convergence for all parameters in the best-fit model (PSRF reached 1 for all parameters). Moreover, the absolute difference graphic produced by AWY online showed no significant variability among independent runs. Pairwise comparisons between tree files of each run showed no difference in the posterior probabilities of all splits for paired MCMC analyses. In all independent runs, the likelihood values reached stationarity by generation 200,000; however, the cumulative graphics produced by the program AWY online showed that the posterior probabilities of the splits stabilize after 5 million generations, showing that tree topologies are finally being sampled in proportion to their posterior distribution and that the chains actually reached stationarity after 5 million generations. Given these results, the first 50,000 trees of each run were discarded as “burn in” and a 50% majority rule consensus tree that summarizes topology and branch length information was calculated based upon the remaining 200,000 trees (Fig. 6).

The phylogenies estimated using MP, Bayesian, and maximum likelihood analyses of the entire trnQ-trnK region are each highly resolved and consistent with one another. The five included accessions of Lichtensteina comprise a clade sister group to the clade of Anginon + Annesorhiza (87% bootstrap, 100% posterior probability). The genera Eryngium, Sanicula, Hacquetia, Petagna, Astrantia, Actinolema, Arctopus, and Alepidea comprise a strongly supported monophyletic group (99% bootstrap, 100% posterior probability). This clade comprises subfamily Saniculoideae, as traditionally circumscribed (but excluding Lagoecia). Steganotaeina and Polemanniospis are well supported sister taxa (94% bootstrap, 100% posterior probability) and this clade is sister group to Saniculoideae (97% bootstrap, 97% posterior probability). Constraining Steganotaeina + Polemanniospis and the clade comprised of Anginon, Annesorhiza, and Lichtensteina to monophyly in a subsequent MP analysis

<table>
<thead>
<tr>
<th>Data partition</th>
<th>Number of MP trees</th>
<th>Length</th>
<th>CI</th>
<th>RI</th>
</tr>
</thead>
<tbody>
<tr>
<td>trnQ-trnK</td>
<td>7455</td>
<td>2026</td>
<td>0.7036</td>
<td>0.9415</td>
</tr>
<tr>
<td>All indels</td>
<td>2016</td>
<td>217</td>
<td>0.8475</td>
<td>0.9089</td>
</tr>
<tr>
<td>trnQ-rps16 + indels</td>
<td>&gt;20,000</td>
<td>1141</td>
<td>0.7514</td>
<td>0.9547</td>
</tr>
<tr>
<td>rps16 + indels</td>
<td>&gt;20,000</td>
<td>1016</td>
<td>0.7408</td>
<td>0.9494</td>
</tr>
<tr>
<td>rps16</td>
<td>&gt;20,000</td>
<td>352</td>
<td>0.6920</td>
<td>0.9430</td>
</tr>
<tr>
<td>rps16-trnK + indels</td>
<td>&gt;20,000</td>
<td>699</td>
<td>0.6794</td>
<td>0.9387</td>
</tr>
<tr>
<td>rps16-trnK</td>
<td>&gt;20,000</td>
<td>642</td>
<td>0.6558</td>
<td>0.9282</td>
</tr>
</tbody>
</table>

Homoplasy indices were calculated by optimizing each data partition over the MP trees obtained from the analysis of the trnQ-trnK + indels matrix (No. of MP trees = 4320; length = 2249 steps; CI = 0.7215; RI = 0.9487). CI, consistency index, excluding uninformative characters; RI, retention index.
resulted in trees eight steps longer than those without the constraint invoked (length = 2257 steps), indicating that Stegnotaenia and Polemanniopsis are indeed more closely related to subfamily Saniculoideae than they are to subfamily Apioideae (the latter including Lichtensteinia). With the exception of Sanicula, all genera are monophyletic. The monotypic genus Hacquetia arises within a paraphyletic Sanicula and comprises a clade along with Sanicula euro-

Fig. 6. Fifty-percent majority-rule consensus of 200,000 trees derived from Bayesian analysis of 91 cpDNA trnQ-trnK sequences. Numbers at nodes are posterior probabilities values; these values are indicated only for Stegnotaenia, Polemanniopsis, Anginon + Anginosorhiza, and the 24 recognized clades indicated in Table 3 and discussed in the text.
World.” Within Astrantia, sections Astrantia and Astrantia-ella are each resolved as monophyletic.

4. Discussion

4.1. Phylogenetic utility of the cpDNA \textit{trnQ-trnK} region

The cpDNA \textit{trnQ-trnK} region encompasses three large noncoding loci, of which only the \textit{rps16} intron has been extensively characterized and used in phylogenetic study (summarized in Kelchner, 2002). To obtain greater phylogenetic resolution at the intergeneric and infrageneric levels than that which could be obtained using only the \textit{rps16} intron (Downie and Katz-Downie, 1999; Calviño et al., 2006), additional data from the \textit{trnQ-rps16} and \textit{rps16-trnK} intergeneric spacer regions were considered. Such noncoding loci are under less functional constraints than coding or intron regions and, therefore, should provide greater levels of variation for phylogenetic analyses (Learn et al., 1992; Gielly and Taberlet, 1994; Downie et al., 1996; Kelchner, 2000, 2002; Shaw et al., 2005).

The issue of the relative frequencies of nucleotide substitutions vs. indels in noncoding cpDNA sequences has been much discussed. Depending upon the region and taxonomic rank considered, indels may occur more, equally, or less often than nucleotide substitutions (Golenberg et al., 1993; Clegg et al., 1994; Gielly and Taberlet, 1994; Small et al., 1998; Shaw et al., 2005). In our study and across the entire \textit{trnQ-trnK} region, the ratio of informative substitutions to indels was 4.6:1; this value was similar to those ratios calculated for each of the separate regions. Despite their lower frequency than nucleotide substitutions, the inclusion of indels in phylogenetic analyses improves levels of branch support and resolution of relationships. In analyses of combined or partitioned data, the inclusion of indel characters yielded more resolved phylogenies with lower levels of overall homoplasy and with higher bootstrap support values. These results add to the growing body of knowledge that indels provide reliable characters for phylogenetic analysis (Graham et al., 2000; Simmons et al., 2001; Geiger, 2002; Vogt, 2002; Hamilton et al., 2003; Müller and Borsch, 2005). Our observation that indels of 10 bp or shorter are the most common size class is in agreement with previous observations of other noncoding markers (Graham et al., 2000). It has been suggested that there is no apparent trend of decreased homoplasy with increased indel size (Simmons et al., 2001; Müller and Borsch, 2005) and our data support these findings. However, all indels of 16 bp or longer were not homoplastic, and a close examination of the results of Müller and Borsch (2005) reveals the same, suggesting that there might be a size threshold for decreased probability of length mutations.

Sequence comparisons of the three cpDNA loci revealed that the two intergenic spacers are evolving about twice as fast as that of the \textit{rps16} intron and pairwise nucleotide divergence values for the spacers are, on average, two times higher than that for the \textit{rps16} intron at all taxonomic levels considered. Group II introns in the chloroplast genomes of land plants, such as the \textit{rps16} intron, show a strong relationship between the functional importance of its secondary structure and the probability of mutational change, with those domains essential for intron-associated functions most conserved evolutionarily (Michel et al., 1989; Lenz et al., 1992; Downie et al., 1996, 1998, 2000b; Kelchner, 2002). While the numerous conserved regions within the \textit{rps16} intron may explain its lower rate of mutational change relative to the intergenic spacers, the functional (and, consequently, structural) constraints within the latter regions and their effects on mutational processes are less understood.

No significant differences were detected in relative evolutionary rates between the two intergenic spacers. However, the \textit{trnQ-rps16} spacer contributes more informative characters to the phylogenetic analysis than do both the \textit{rps16} intron and \textit{rps16-trnK} spacer region combined. MP analysis of separate data sets reveals that the \textit{trnQ-rps16} region yields phylogenies that are most resolved and with greater branch support than the other regions, and resolution of relationships is obtained at diverse hierarchical levels. Also, relative to the \textit{rps16} and \textit{rps16-trnK} regions, the \textit{trnQ-rps16} intergeneric spacer region results in trees with higher CI and RI values. A major problem with the \textit{trnQ-rps16} locus, however, is the size of deletions which can occur in this region (such as the 1009 bp deletion in \textit{Petagnaea}). Such large deletions may be disastrous for phylogenetic analysis if they result in a reduction of available informative nucleotide substitutions.

Overall, the greatest resolution of relationships and the highest levels of branch support were achieved by simultaneous analysis of all nucleotide and indel data. Had we restricted this study to a continued sampling of only \textit{rps16} intron sequences, all major clades inferred by MP analysis of combined data plus indels would have been recovered save one, but many of these would have been supported only weakly to moderately. Similarly, resolution of infrageneric relationships would have been minimal. Continued investigations of Apiaceae phylogeny, at a variety of hierarchical levels, would benefit by consideration of the entire \textit{trnQ-trnK} locus.

4.2. Subfamily Saniculoideae

Wolff’s (1913) revision of subfamily Saniculoideae is comprehensive and predominant and, until recently, changes to his system have been minimal. Like the classification of Drude (1898), Wolff recognized two tribes (Sanieuleae and Lagoecieae) and nine genera (\textit{Actinolema}, \textit{Alepidia}, \textit{Arctopus}, \textit{Astrantia}, \textit{Eryngium}, \textit{Hacquetia}, \textit{Lagoecia}, \textit{Petagnaea}, and \textit{Sanicula}) within the subfamily. Molecular studies have supported the transfer of \textit{Lagoecia} to subfamily Apiiodeae (Plunkett et al., 1996; Downie et al., 2000a). The Argentinean genus \textit{Oligocladus} has been described as \textit{incertae sedis} within Saniculoideae (Pimenov and Leonov, 1993); our unpublished work, however, places
this genus in tribe Selineae of subfamily Apioidae (C.I. Calviño and S.R. Downie, unpublished data). Liu et al. (2003), primarily on the basis of cladistic analysis of fruit anatomical characters but also considering the available molecular evidence (Downie and Katz-Downie, 1999), expanded the circumscription of subfamily Saniculoideae to include the African genera Steganothaenia, Polemanniopsis, and Lichtensteinia. Based on our results and those obtained in a concurrent study of southern African Apiaceae (Calviño et al., 2006), we present a revised classification of Saniculoideae that more accurately reflects its evolutionary history. We reject the transfer of the genus Lichtensteinia into Saniculoideae and recognize Steganothaenia and Polemanniopsis as comprising a new, distinct tribe within the subfamily that is a well-supported sister group to tribe Saniculeae. The latter is redefined to include the genera Actinolema, Alepidea, Arctopus, Astraantia, Eryngium, Petagnaea, and Sanicula. Hacquetia epipactis is included within Sanicula and, as such, all genera are monophyletic. Both tribes are clearly delineated morphologically, but only molecular data support their sister group relationship. In addition to nucleotide substitutions, the distribution of several synapomorphic indels supports the inclusion of Steganothaenia and Polemanniopsis in Saniculoideae. As yet, however, we are unaware of any unique anatomical or morphological character that would support the union of these taxa.

The Namibian genus Phlyctidocarpa Cannon and W.L. Theob., also treated in subfamily Apioidae (Cannon and Theobald, 1967; Pimenov and Leonov, 1993), may constitute yet another member of subfamily Saniculoideae based on our preliminary molecular study. Like Steganothaenia and Polemanniopsis, Phlyctidocarpa resembles most apioid umbellifers in habit, but has fruit features similar to those found in Saniculoideae (Theobald and Cannon, 1973). Fruit anatomical and further molecular studies of this unusual monotypic genus are in progress and the results should clarify its phylogenetic placement (B.E. van Wyk et al., University of Johannesburg, unpublished data).

4.3. Lichtensteinia

Liu et al. (2003) included the African apioid genus Lichtensteinia in their broadened circumscription of Saniculoideae based primarily on the shared presence of large intrajugal secretory ducts and the absence of both commissural and vallecular vittae. These characters, however, occur among other early branching members of Apiaceae, thus the inclusion of Lichtensteinia alongside Steganothaenia, Polemanniopsis, and other Saniculoideae is based on shared plesiomorphies rather than synapomorphies (Calviño et al., 2006). We are not aware of any morphological or anatomical character supporting the union of Lichtensteinia with Saniculoideae that does not also include members of Apiaceae subfamilies Azorelloideae and Mackinlayioideae. The results of a concurrent molecular phylogenetic study of southern African Apiaceae with expanded sampling of early branching lineages of subfamily Apioidae revealed that Lichtensteinia comprises a monogenic clade that is sister group to all other members of subfamily Apioidae (Calviño et al., 2006). The same results are supported herein upon expanded sampling of Saniculoideae and consideration of additional molecular data. Lichtensteinia, therefore, should be maintained within subfamily Apioidae.

4.4. Tribe Steganothaenieae

The African genera Steganothaenia and Polemanniopsis, arborescent and shrubby plants that were treated previously in subfamily Apioidae, are monophyletic sister taxa and this clade is a well-supported sister group to subfamily Saniculoideae, as traditionally circumscribed. Such relationships were proposed initially using rps16 intron sequences (Downie and Katz-Downie, 1999) and later corroborated through a greater sampling of early branching lineages of subfamily Apioidae (Calviño et al., 2006), but in each of these studies bootstrap support for the sister group relationship between Steganothaenia/Polemanniopsis and Saniculoideae was weak. As a result of examination of fruit anatomical characters and in light of the molecular evidence, Liu et al. (2003) transferred Steganothaenia and Polemanniopsis to subfamily Saniculoideae. Upon consideration of additional cpDNA data from previous studies, and comprehensive sampling of all genera and most infrageneric groups traditionally considered in Saniculoideae, our results corroborate the expansion of the subfamily to include Steganothaenia and Polemanniopsis, as the entire group is strongly supported as monophyletic. The alternative would be to recognize the clade of Steganothaenia/Polemanniopsis as a new subfamily but, by doing so, its sister group relationship to Saniculoideae would not be reflected in the classification and it would comprise only two genera and four species.

Steganothaenia and Polemanniopsis are woody, occur in southern Africa (Sleganothaenia is widespread throughout tropical Africa), and have fruits with 2–3 wings and 5 prominent intrajugal secretory ducts per mericarp. At maturity, those intrajugal ducts associated with the wings expand forming enormous cavities (Liu et al., 2003). This combination of characters sets them apart from all other genera of the subfamily. We formally recognize this clade at the tribal level and propose the new name, Steganothaenieae C. I. Calviño and S. R. Downie.

4.5. Tribe Saniculoideae

With the transfer of Lagoecia to subfamily Apioidae and taxonomic realignments suggested by the molecular
phylogenies, further changes to the infrasubfamilial classification of Saniculoideae are necessary. Tribe Saniculeae, as traditionally circumscribed (Drude, 1898; Wolff, 1913), consists of Actinolema, Alepidea, Astrantia, Hacquetia, Eryngium, and Sanicula. The inclusion of both Petagnaea and Actopous (genera traditionally treated in tribe Lagoeceae) among these taxa suggests that tribe Saniculeae, to be monophyletic, must be redefined to include these genera. Tribe Saniculeae was characterized by Drude and Wolff as having mature fruits with both mericarps fertile (as is typical of most Apiaceae), whereas members of tribe Lagoeceae have fruits with only one fertile (or developed) mericarp. Our expanded circumscription of Saniculoideae includes plants characterized by fruits with an exocarp often covered in scales, bristles, or prickles, a mesocarp with calcium-oxalate crystals scattered throughout, and umbels or heads that are surrounded by showy bracts. This revised definition of tribe Saniculeae corresponds precisely to Drude’s conception of subfamily Saniculoideae (upon the exclusion of Lagoeceae).

Alepidea is well supported as monophyletic (100% bootstrap, 100% posterior probability) and is sister group to all other members of tribe Saniculeae. The genus comprises perennial herbs with roslulate, simple leaves and capitulate inflorescences. It is mainly distributed in southern Africa, with a few species extending northward to Zimbabwe, Kenya, and Ethiopia (van Wyk, 2000). Historically, Alepidea and Eryngium have been considered closely related (Bentham, 1867; Dümmer, 1913; Turmel, 1950). Both genera share similar inflorescence types and leaves, with the main difference between them being the absence of floral bracts in Alepidea. It was further thought that Eryngium and Alepidea had a common Asian origin, whereupon the lineages that diverged into these genera dispersed along two different routes: one northwestward to colonize Europe and then the New World, and the other southward towards eastern Africa (Turmel, 1950). Our results, however, reject the hypothesis of a sister group relationship between Alepidea and Eryngium. The morphological characters once thought to unite these genera, while similar in appearance, have evolved independently for there are clear structural differences in the capitate inflorescences and leaves between the genera (C. I. Calviño et al., unpublished data). Moreover, the origin of Alepidea is likely southern African rather than Asian. The infrageneric taxonomy of Alepidea has long been problematic despite several revisionary studies (Sonder, 1862; Dümmer, 1913; Wolff, 1913; Weimarck, 1949; Hilliard and Burtt, 1982; Burtt, 1991). Weimarck (1949) recognized six sections based on fruit surface features and types of marginal cilia on the basal leaves. The sampling of Alepidea herein represents all sections save one, and the phylogenetic results obtained indicate that these sections are artificial. Features of the marginal cilia display intergradation of character states among the sections (Hilliard and Burtt, 1982; Burtt, 1991) and we are of agreement with Burtt (1991) in suggesting that these sections be abandoned.

Actopous consists of three dioecious species endemic to the Cape Floristic Region of South Africa. Its distinctive characters include an acaulcous habit, deciduous leaves, spiny leaf margins, a single developed mericarp, and lack of differentiation between sepals and petals in female flowers. Monophyly of the genus was confirmed using ITS and cpDNA trnL-F sequences (A.R. Magee et al., University of Johannesburg, unpublished data). Once treated in the traditionally recognized Apiaceae subfamily Hydrocotyloideae (Froebe, 1979; Magin, 1980; Constance and Chuang, 1982; Pimenov and Leonov, 1993), molecular data clearly support the original placement of this genus within subfamily Saniculoideae. Actopous is sister group to the large clade of Actinolema through Eryngium (Fig. 3).

Astrantia and Actinolema are each monophyletic and collectively comprise a strongly supported clade sister group to the clade of Petagnaea, Sanicula (including Hacquetia), and Eryngium. The sister group relationship between Actinolema and Astrantia is also supported by ITS sequences (Valiejo-Roman et al., 2002) and several shared morphological characters (Grinztesco, 1910; Wolff, 1913). Actinolema comprises two annual species distributed in Anatolia and the Caucasus. They have entire basal leaves and sessile umbels with leaf-like involucral bracts. Astrantia includes nine perennial species characterized by divided basal leaves and conspicuous, colored involucral bracts. Its distribution is sympatric to Actinolema in SW Asia, but its range extends further west to the Balkan Peninsula, Carpathian Mountains, Alps, Apennines, and Pyrenees (Wörz, 1999). Two sections are recognized in Astrantia: sect. Astrantia (= sect. Macraster Grinztesco), including A. major, A. maxima, and A. colchica; and sect. Astrantiella, including A. bavarica and A. minor (Grinztesco, 1910; Wolff, 1913). Section Astrantia displays rigid involucral bracts and acute calyx teeth longer than the petals, whereas section Astrantiella has membranous involucral leaves and obtuse calyx teeth longer or as long as the petals (Wollf, 1913; Wörz, 1999). Each of these sections is monophyletic in the molecular phylogenies presented herein.

Petagnaea is endemic to the Nebo Rocky Mountains of NE Sicily, Italy. Its only species, P. gussonei (= Petagnia saniculifolia Gussone), is categorized as endangered according to the International Union for the Conservation of Nature and Natural Resources (IUCN) Red List (Gianguzzi et al., 2004). Petagnaea comprises a monotypic lineage sister to the clade of Eryngium + Sanicula (including Hacquetia). Its isolated phylogenetic position in Saniculoideae corroborates the need for further conservation actions for this species.

The genus Sanicula, upon the inclusion of Hacquetia epipactis, comprises a well supported monophyletic group. Both genera are morphologically very similar, with Hacquetia separated from Sanicula primarily by its showy involucral bracts and the absence of prickles on the fruit exocarp. Many species of Sanicula, however, also lack fruit prickles. Indeed, the absence of such prickles characterizes three different lineages of Sanicula (Vargas et al., 1999).
Sanicula orthacantha (sect. Pseudopetagna) and S. hacquetoides (sect. Tuberculatae, not sampled here) also display prominent and leafy involucres similar to those seen in Hacquetia. Shan and Constance (1951) suggested an affinity among Hacquetia and Sanicula sects. Tuberculatae and Pseudopetagna and in our study, Hacquetia epipactis allies with S. orthacantha and S. europaea. The results of phylogenetic analyses of ITS sequences also show Hacquetia nested within Sanicula (Valiejo-Roman et al., 2002). On the bases of morphological similarities, the intergradation of characters previously used to delimit the genera, and results of molecular phylogenetic analyses, H. epipactis is best treated within the genus Sanicula. The name Sanicula epipactis (Scop.) E.H. Krause is already available for use.

Eryngium is the largest genus in the family with about 250 species (Pimenov and Leonov, 1993). It is easily distinguished by its capitate inflorescences and single bract per flower and it has long been recognized as a natural group. The genus, however, is extremely variable morphologically and interspecific relationships are obscure. In the present study, we examined 35 species of Eryngium, a 4- to 6-fold increase in sampling over any previous study (Downie and Katz-Downie, 1999; Valiejo-Roman et al., 2002). These species represent 34 of 36 sections of Eryngium recognized by Wolff (1913). On the bases of the phylogenetic results presented herein, Eryngium is confirmed as monophyletic (85% bootstrap, 100% posterior probability). Furthermore, three synapomorphic indels support the monophyly of Eryngium (Fig. 5). These results are in contrast to those inferred by Valiejo-Roman et al. (2002) using ITS sequences in which Sanicula occurred within a paraphyletic Eryngium. A re-analysis of the DNA sequences used by Valiejo-Roman et al., which contained representatives of all major subfamilies of Apiaceae, could not be aligned unambiguously because of high nucleotide sequence divergence and numerous indels. Tree topologies were not stable to alternative pairwise and multiple alignment parameters assigned by CLUSTAL X and, in contrast to the results of Valiejo-Roman et al., a monophyletic Eryngium could result in some of our analyses of their data. At such deep levels of comparison, ITS sequences are just too divergent for phylogenetic analyses (Downie et al., 2001; Hardway et al., 2004).

In our analyses, Eryngium is sister group to Sanicula; this relationship, however, is not strongly supported (72% bootstrap, 68% posterior probability).

The cpDNA phylogenies presented herein show two strongly supported major clades within Eryngium: “Old World,” representing species from Eurasia and North Africa; and “New World,” representing species primarily from the western hemisphere, but also from Australia and the western Mediterranean. Similar groupings of species were recognized by Wolff (1913): “Species gerontogaeae” and “Species americanae and australienses,” respectively. Wolff (1913) included Eryngium tenue, E. visiparum, and E. corniculatum (distributed in the western Mediterranean) in his “Species gerontogaeae” (“Old World”) group, but indicated that these species are closely related to those of North America. In the molecular phylogenies, these three species fall as successive sister lineages at the base of the “New World” clade, suggesting that Eryngium of the New World may have had their origin from western Mediterranean ancestors. A recent, yet preliminary, classification of Eryngium has been presented by Wörz (2004) based on morphology, but it does not reflect the major clades recognized herein using molecular data. Further molecular systematic studies on Eryngium are in progress and the results should illuminate the naturalness of the sections described by Wolff (1913) while contributing to a modern classification of this taxonomically complex genus (C.I. Calviño et al., unpublished data).

In summary, we revise the circumscription of Apiaceae subfamily Saniculoideae and present an estimate of phylogenetic relationships within the subfamily using data from the cpDNA trnQ-trnK 5′-exon region, which includes two intergenic spacers that have been heretofore previously underutilized in molecular systematic studies. The region presents a large number of phylogenetically informative indels that, when analyzed by themselves, yield trees of great resolution and high levels of overall branch support. With the exception of the monotypic genus Polemanniopsis, synapomorphic indels support the monophyly of all recognized genera within the subfamily. When used with nucleotide substitution data, the inclusion of indels in phylogenetic analyses of both combined and partitioned data sets improves resolution of relationships, increases branch support values, and decreases levels of overall homoplasy. The continued acquisition of sequence data from the trnQ-trnK 5′-exon region for assessing relationships at a variety of hierarchical levels within Apiaceae seems worthwhile, given the large number of parsimony informative characters obtained, as well as adequate levels of sequence divergence observed even among congeners.

We have redefined subfamily Saniculoideae to include the African genera Steganotaenia and Polemanniopsis (as the new tribe Steganoaenae), but not Lichtensteinia. The latter should be maintained within subfamily Apiioideae. Sister group to tribe Steganoaenae is an expanded tribe Saniculeae, which includes Actinolema, Alepidea, Arctopus, Astrantia, Eryngium, Petagnaea, and Sanicula. The genus Hacquetia is synonymized with Sanicula and, as such, all genera are monophyletic. Intergeneric relationships within tribe Saniculeae are fully resolved and, generally, very well supported. Studies are in progress which will use the resultant phylogeny to evaluate patterns of character state evolution within the subfamily and to formulate hypotheses of ancestral character states and biogeography using objective methods.

Acknowledgments

The authors thank the curators of herbaria BOL, E, ILL, ILS, JACA, JEPS, MO, NBG, PAL, SI, UC, and WIS for access to specimens, M. Watson, G. Domina, and P. Vargas for sending leaf material and/or DNAs, I. Hedge and R.
Pozner for helping with the Latin description, and D.S. Katz-Downie, S.G. Martinez, F.O. Zuloaga, and three anonymous reviewers for comments on the manuscript. This work was supported by grants to S.R. Downie from the National Science Foundation (DEB 0089452) and the National Center for Supercomputing Applications (DEB 030005N), utilizing the IBM pSeries 690 system at the University of Illinois at Urbana-Champaign (UIUC). Travel funds to C.I. Calviño were provided by UIUC’s School of Integrative Biology Enhancement Fund and the Department of Plant Biology John R. Laughman Award. This paper represents part of a Ph.D. dissertation (C.I.C.), for which funding from CONICET is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2007.01.002.

References


sequence of the rice (*Oryza sativa*) chloroplast genome; intermolecular recombination between distant tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. Mol. Gen. Genet. 217, 185–194.


