BMC Evolutionary Biology

BioMed Central

Research article

Section-level relationships of North American *Agalinis* (Orobanchaceae) based on DNA sequence analysis of three chloroplast gene regions

Maile C Neel*[1,2] and Michael P Cummings[3]

Address: [1]Department of Natural Resource Sciences and Landscape Architecture, University of Maryland, College Park, MD 20742, USA, [2]Department of Entomology, University of Maryland, College Park, MD 20742, USA and [3]Center for Bioinformatics and Computational Biology, University of Maryland, College Park, MD 20742-3360, USA

Email: Maile C Neel* - mneel@umd.edu; Michael P Cummings - mike@umiacs.umd.edu

* Corresponding author

Published: 08 June 2004

BMC Evolutionary Biology 2004, **4**:15 doi:10.1186/1471-2148-4-15

Received: 06 March 2004

Accepted: 08 June 2004

This article is available from: http://www.biomedcentral.com/1471-2148/4/15

Abstract

Background: The North American *Agalinis* are representatives of a taxonomically difficult group that has been subject to extensive taxonomic revision from species level through higher sub-generic designations (e.g., subsections and sections). Previous presentations of relationships have been ambiguous and have not conformed to modern phylogenetic standards (e.g., were not presented as phylogenetic trees). *Agalinis* contains a large number of putatively rare taxa that have some degree of taxonomic uncertainty. We used DNA sequence data from three chloroplast genes to examine phylogenetic relationships among sections within the genus *Agalinis* Raf. (=*Gerardia*), and between *Agalinis* and closely related genera within Orobanchaceae.

Results: Maximum likelihood analysis of sequences data from *rbc*L, *ndh*F, and *mat*K gene regions (total aligned length 7323 bp) yielded a phylogenetic tree with high bootstrap values for most branches. Likelihood ratio tests showed that all but a few branch lengths were significantly greater than zero, and an additional likelihood ratio test rejected the molecular clock hypothesis. Comparisons of substitution rates between gene regions based on linear models of pairwise distance estimates between taxa show both *ndh*F and *mat*K evolve more rapidly than *rbc*L, although the there is substantial rate heterogeneity within gene regions due in part to rate differences among codon positions.

Conclusions: Phylogenetic analysis supports the monophyly of *Agalinis*, including species formerly in *Tomanthera*, and this group is sister to a group formed by the genera *Aureolaria*, *Brachystigma*, *Dasistoma*, and *Seymeria*. Many of the previously described sections within *Agalinis* are polyphyletic, although many of the subsections appear to form natural groups. The analysis reveals a single evolutionary event leading to a reduction in chromosome number from $n = 14$ to $n = 13$ based on the sister group relationship of section Erectae and section Purpureae subsection Pedunculares. Our results establish the evolutionary distinctiveness of *A. tenella* from the more widespread and common *A. obtusifolia*. However, further data are required to clearly resolve the relationship between *A. acuta* and *A. tenella*.

Background

Agalinis (including *Tomanthera* Raf.) is a genus of from 40 to 70 species (depending on the taxonomy used) distributed in the eastern part of the United States, Mexico, Central America and South America. The name *Agalinis* is preferred to the older name *Gerardia* because the latter name was first applied to another taxon (now known as *Stenandrium rupestre* [Swartz] Nees) that is now a member of the family Acanthaceae [1,2]. Members of the genus *Agalinis* have zygomorphic, membranaceous, ephemeral corollas and wingless seeds with variously reticulate seed coats [3-6]. Beyond the above characteristics, life form, morphology, anatomy and floral form and color are variable in the genus, particularly in South America [6-9]. Unfortunately, South American taxa are relatively poorly known, are not included in any published classification schemes for the genus, and are not included in this study.

The North American *Agalinis* species are less variable and all have pink-purple, membranaceous, ephemeral corollas, typically with red spots on the anterior lobes. Most species also have two yellow guide lines on the anterior lobes. Except for one perennial species (*A. linifolia*), North American *Agalinis* are all annual herbs and all except three species (*A. auriculata*, *A. densiflora*, and *A. heterophylla*) have linear to filiform or scale-like leaves. Many species are hemiparasitic; in fact, *Agalinis* represents the largest genus of hemiparasitic plants in the eastern United States. Mating systems within the genus range from self incompatible in *A. strictifolia* [10], to mixed mating in *A. acuta* [11] and *A. skinneriana* [12], to highly selfing in *A. neoscotica* [13]. Most North American species are restricted to the coastal plain of the southern and eastern United States where they occupy a range of habitats including dry, sandy pine barrens, grasslands, and edges of wetlands including bogs, ponds, and salt marshes [3,4,14]. Off the coastal plain *Agalinis* species are also found in prairie habitats, other grasslands, and open habitats within shrublands or woodlands.

The North American *Agalinis* are representatives of a taxonomically difficult group that has been subject to extensive taxonomic revision from species level through higher sub-generic (e.g., subsections and sections) and generic designations. Two species currently in *Agalinis* (*A. auriculata*, and *A. densiflora*) have been considered to be in the genus *Tomanthera* [3,4]. Early species circumscriptions of, and relationships among, North American *Agalinis* taxa were originally postulated based on morphological and anatomical characteristics (Table 1) [2-4,14,15]; however, presentations of relationships were ambiguous and did not conform to modern phylogenetic standards. More recently, extensive studies of seed morphology [5], seedling morphology [16], floral development [17-20], chromosome cytology [21,22], as well as stem and leaf anatomy [23] have been used to clarify taxonomy (Table 1). While this body of work has served to revise classifications, explicit phylogenetic presentations are still lacking and only general notions of relationships (primarily section and subsection membership) within the genus have been postulated. Our research provides the first examination of DNA sequence characters within this genus and provides the first explicit hypotheses regarding phylogenetic relationships. We used DNA sequence data from three chloroplast genes to examine phylogenetic relationships among species within the genus *Agalinis* Raf. (=*Gerardia* Benth.), and between *Agalinis* and closely related genera within Orobanchaceae.

We also examined relationships of *Agalinis* to other genera. *Agalinis* has long been considered to be closely related to five other North American genera (*Aureolaria*, *Brachystigma*, *Dasistoma*, *Macranthera* and *Seymeria*). Close relationships among these taxa are reflected in the fact that they were at one time included in the genus *Gerardia* [1,2] and have been referred to as the "gerardioid genera". Relationships among the North American gerardioid genera were proposed based on morphological features [14]; again, presentations of these relationships were vague and do not conform to modern phylogenetic standards (e.g., were not presented as a phylogenetic tree). The gerardioid genera were traditionally included as subtribe Agalininae in tribe Buchnereae [1,3]. Recent molecular studies that have included *Seymeria* or *Agalinis* indicated these taxa are part of a monophyletic clade representing the tribe Rhinantheae [24-27]. These studies also provided evidence that has resulted in moving *Agalinis* and other gerardioid genera from the family Scrophulariaceae to the family Orobanchaceae. While it has yielded important insights that have influenced taxonomy, this previous molecular work has focused on broadly sampling across Scrophulariaceae *sensu lato* [27,28] or sampling intensively within putatively non-photosynthetic parasitic lineages [24-26,29]. Taxon coverage within the hemi-parasitic lineages has been relatively sparse and evolutionary relationships have not been thoroughly examined. Moreover, relationships among the gerardioid genera and between these genera and other Rhinantheae remain unclear. We used GenBank [30] sequences of *Castilleja linariifolia* and *Pedicularis foliosa* to examine placement of the gerardioid genera in Rhinantheae and *Lindenbergia philippensis* as an outgroup within Orobanchaceae [25,26,29].

Much of our motivation for studying *Agalinis* is the large number of putatively rare taxa in the genus that have some degree of taxonomic uncertainty. The genus includes six globally vulnerable (G3), imperiled (G2) or critically imperiled (G1) taxa, and several additional taxa are of uncertain global conservation status (Table 2) [31]. Further, 26 of the approximately 40 North American species

Table 1: Alternative proposed classification schemes for the genus *Agalinis*, with synonomies from the USDA Plants Database [32]

Canne and co-authors [5,22,23]	Pennell (1929) [3]	Pennell (1935) [4]
	Tomanthera	*Tomanthera*
	T. auriculata	*T. auriculata*
	T. densiflora	*T. densiflora*
Agalinis	*Agalinis*	*Gerardia*
Section Linifoliae	Section Linifolieae	Section Spartorhizoma
A. linifolia	*A. linifolia*	*G. linifolia*
		Section Chytra
Section Heterophyllae	Section Heterophyllae	Subsection Heterophyllae
A. auriculata	*A. calycina*	*G. calycina*
A. calycina	*A. heterophylla*	*G. heterophylla*
A. densiflora		
A. heterophylla		
	Section Asperae	Subsection Asperae
	A. aspera	*G. aspera*
Section Purpureae	Section Purpureae	
Subsection Purpureae	Subsection Purpureae	Subsection Purpureae
A. fasciculata	*A. albida*	*G. fasciculata*
A. harperi = *A. pinetorum*	*A. borealis*	*G. georgiana* = *A. fasciculata*
A. maritima	*A. caddoensis*	*G. harperi* = *A. pinetorum*
A. neoscotica	*A. fasciculata*	*G. maritima*
A. paupercula	*A. georgiana* = *A. fasciculata*	*G. paupercula*
A. pinetorum	*A. harperi* = *A. pinetorum*	*G. pulchella*
A. purpurea	*A. maritima*	*G. purpurea*
A. tenuifolia	*A. neoscotica*	*G. racemulosa* = *A. fasciculata*
A. virgata = *A. fasciculata*	*A. paupercula*	
	A. pinetorum	
	A. purpurea	
	A. virgata = *A. fasciculata*	
Subsection Setaceae	Subsection Setaceae	Subsection Setaceae
A. filifolia	*A. filifolia*	*G. aphylla*
A. laxa	*A. holmiana* = *A. setacea*	*G. caddoensis*
A. plukenetii	*A. keyensis*	*G. filifolia*
A. setacea	*A. laxa*	*G. laxa*
A. stenophylla	*A. oligophylla*	*G. microphylla* = *A. oligophylla*
	A. pseudaphylla = *A. oligophylla*	*G. plukenetii*
	A. setacea	*G. pseudaphylla* = *A. oligophylla*
	A. stenophylla = *A. setacea*	*G. pulcherrima* = *A. pulchella*
		G. setacea
	Subsection Aphyllae	*G. stenophylla* = *A. setacea*
	A. aphylla	*G. strictifolia*
Subsection Pedunculares	Subsection Pedunculares	
A. aspera	*A. caddoensis*	
A. edwardsiana	*A. peduncularis*	
A. peduncularis	*A. pulchella*	
A. pulchella	*A. strictifolia*	
A. strictifolia		
A. homalantha		
Section Tenuifoliae	Section Tenuifoliae	Subsection Tenuifoliae
A. divaricata	*A. divaricata*	*G. divaricata*
A. filicaulis	*A. edwardsiana*	*G. edwardsiana*
A. nutallii	*A. filicaulis*	*G. filicaulis*
	A. homalantha	*G. homalantha*
	A. longifolia = *A. nuttallii*	*G. longifolia* = *A. nuttallii*
	A. polyphylla	*G. tenuifolia*
	A. tenuifolia	
Section Erectae	Section Erectae	Section Chloromone
A. acuta	*A. acuta*	*G. acuta*
A. aphylla	*A. decemloba* = *A. obtusifolia*	*G. decemloba* = *A. obtusifolia*

Table 1: Alternative proposed classification schemes for the genus *Agalinis*, with synonomies from the USDA Plants Database [32]

A. decemloba = *A. obtusifolia*	*A. erecta*	*G. gattingeri*
A. gattingeri	*A. gattingeri*	*G. obtusifolia*
A. keyensis	*A. obtusifolia*	*G. skinneriana*
A. obtusifolia	*A. skinneriana*	*G. tenella* = *A. obtusifolia*
A. oligophylla	*A. tenella* = *A. obtusifolia*	*G. viridis*
A. skinneriana	*A. viridis*	
A. tenella = *A. obtusifolia*		
A. viridis		

A. caddoensis, *A. albida*, *A. georgiana* were not included in Canne-Hilliker papers. *A. peduncularis* was left out of Pennell 1935 [4] – he mentions it in the text but does not include it in the keys or descriptions; *A. erecta* is mentioned only in Pennell 1929 [3].

Table 2: North American *Agalinis* species examined with sectional and subsectional classification following J. Canne-Hilliker

Taxon (= synonym in The Plants Database)[1]	Status[2]	Locality	Voucher	GenBank Accession Numbers		
				*rbc*L	*mat*K	*ndh*F
Section Linifoliae (*n* = 14)						
A. linifolia (Nutt.) Britt.	S	FL USA	JCH 3554	AY563949	AY563923	AY563929
Section Heterophyllae (*n* = 14)						
A. auriculata (Michx.) Blake	G3/S	Midewin, Will Co., IL USA	J. Koontz 5	AY563938	AY563917	na
A. heterophylla (Nutt.) Small ex Britt.		Boca Chica Beach, Cameron Co., TX USA	Cabrera and Dieringer 1057	AY563934	AY563918	AY563928
Section Purpureae						
Subsection Purpureae (*n* = 14)						
A. fasciculata (Ell.) Raf.	S	Long Co., GA USA	JCH 3529	AY563944	AY563919	na
A. tenuifolia (Vahl) Raf.	S	Ames, IA, Story Co., USA	ISC 424636	AY563936	AY563916	AY563927
Subsection Setaceae (*n* = 14)						
A. plukenetii (Ell.) Raf.	G3–G5	Washington Co, FL USA	JCH 3558	AY563933	AY563915	na
A. setacea (J. F. Gmel.) Raf.	S	VA USA	JCH 3499	AY563941	AY563914	na
Subsection Pedunculares (*n* = 13)						
A. pulchella Pennell		GA USA	JCH 3544	AY563935	AY563912	na
A. strictifolia (Benth.) Pennell		Brackenridge Field Lab, Travis Co., TX USA	JLN 01-10-07-03	AY563945	AY563913	na
Section Tenuifoliae (*n* = 14)						
A. divaricata (Chapman) Pennell	G3?/S	AL USA	JCH 3559	AY563946	AY563906	na
A. filicaulis (Benth.) Pennell	G3–G4/S	AL USA	JCH 3569	AY563937	AY563907	na
Section Erectae (*n* = 13)						
A. acuta Pennell	G1/S	Waquoit Bay NERR, Branstable Co., MA USA	no voucher	AY563943	AY563908	AY563930
A. aphylla (Nutt.) Raf.	G3–G4/S	FL USA	JCH 3545	AY563939	AY563911	AY563931
A. obtusifolia Raf.	S	FL USA	JCH 3598	AY563950	AY563910	AY563932
A. tenella Pennell = *A. obtusifolia*	S	Ware Co., GA USA	JCH 3537	AY563948	AY563909	na

[1]Chromosome counts represent those known for the section or subsection. [2]Conservation Status: G1, G2, and G3 specify globally vulnerable or imperiled; S specifies imperiled (S1 or S2) in at least one state (USA); when a range or question mark (?) is given the precise conservation status is uncertain.

in the genus are considered imperiled (S2) or critically imperiled (S1) in at least one state (USA) in which they occur [31]. A number of these state-rare taxa are considered possibly synonymous with more widely ranging taxa (e.g., *A. paupercula* with *A. purpurea*, *A. tenella* with *A. obtusifolia*, *A. decemloba* with *A. obtusifolia*, and *A. keyensis* with *A. oligophylla*) [32]. Additionally, a number of recognized taxa are taxonomically challenging to delineate in the field, making it unclear which populations warrant conservation attention and protection (e.g., *A. skinneriana* and *A. fasciculata*). Confirming or clarifying the evolutionary distinctiveness of putatively rare taxa provides information that can inform conservation actions including determining legal status and setting priorities.

Table 3: Species in the gerardioid genera and outgroup taxa in the Rhinantheae examined in this study

Taxon	GenBank Accession Numbers		
	*rbc*L	*mat*K	*ndh*F
Representatives of Gerardioid genera			
Aureolaria pedicularia (L.) Raf.; VA USA; JCH 3497	AY563940	AY563920	AY563926
Brachystigma wrightii (Gray) Pennell; Huachuca Mtns., Cochise Co., AZ USA; JCH 3569	AY563942	AY563922	AY563924
Dasistoma macrophylla (Nutt.) Raf.; Ames, Story Co., IA USA; no voucher	AY563947	AY563921	AY563925
Seymeria pectinata Pursh	AF026837	AF051999	AF123691
Representatives of other Rhinantheae genera			
Castilleja linariifolia Benth.	AF026823	AF051981	na
Pedicularis foliosa L.	AF026836	AF489959	AF123689
Other species within Orobanchaceae			
Lindenbergia philippensis (Cham.) Benth.	AF123664	AF051990	AF123686

We have included 12 taxa of conservation concern in this study (Table 2). We are particularly interested in clarifying the distinctiveness of *A. tenella* from *A. obtusifolia*, and *A. acuta* from *A. tenella*. *Agalinis tenella* occurs on the coastal plain from North Carolina to Florida and Alabama [3,4,33]. It is considered to be "significantly rare" in North Carolina [33] and a species of concern in South Carolina [34]. While this species continues to be recognized by some authors [33], it is considered by others to be synonymous with the widespread, common *A. obtusifolia* [35]. *Agalinis acuta* is a federally-listed endangered species that occurs in sandplain grasslands on the coastal plain in Connecticut, Rhode Island, Massachusetts, New York (Long Island) and at one location on the piedmont in Maryland [31,36]. *Agalinis acuta* and *A. tenella* have been distinguished morphologically by shorter corollas, smaller seeds and shorter pedicels in *A. acuta* but their evolutionary distinctiveness from one another has recently been called into question. Clarifying whether these taxa are distinct is essential to understanding their rarity status. It is especially important in the case of *A. acuta* because the assumption of evolutionary (*i.e.*, taxonomic, phylogenetic, or genetic) distinctiveness is a fundamental requirement for listing as an endangered species [37].

The specific objectives of this research were to test phylogenetic hypotheses concerning relationships within North American *Agalinis* including following.

1. Monophyly of *Agalinis* as currently defined including two species previously included in the separate genus *Tomanthera*.

2. Congruence between the molecular-based phylogeny and the sectional and subsectional classifications based on anatomy, morphology and cytogenetics. Specific alternative hypotheses we examined correspond to the monophyly of sections and subsections recognized by Pennell [3,4] and Canne-Hilliker [5,6,16,21,23].

3. Phylogenetic distinctiveness of putatively rare taxa including *A. tenella* from *A. obtusifolia*, and *A. acuta* from *A. tenella*.

4. Relationships among the gerardioid genera.

Results and Discussion

Basic data description

We determined 18 partial *rbc*L sequences and these were 929–1322 bp ($\bar{x}$ = 1293 bp) in length. Four additional gerardioid, Rhinantheae, and outgroup *rbc*L sequences were obtained from GenBank [30] (Table 3) yielding a 22 taxon data set with an aligned length of 1322 bp. We determined 9 partial *ndh*F sequences and these were 2096–2122 bp ($\bar{x}$ = 2115 bp) in length. We were unable to amplify *ndh*F from nine *Agalinis* species (*A. auriculata*, *A. divaricata*, *A. fasciculata*, *A. filicaulis*, *A. plukenetii*, *A. pulchella*, *A. setacea*, *A. strictifolia*, and *A. tenella*). We were unable to amplify any part of *ndh*F using numerous internal and external primer combinations under a variety of PCR conditions, indicating that the lack of amplification was not due simply to modification of one priming site. Difficulty in amplifying *ndh*F sequences in some Scrophulariaceae genera has been attributed to absence or divergence [27]. This phenomenon has not been previously reported among species within a single genus. Three additional sequences for *ndh*F were obtained from GenBank (Table 3) resulting in sequences for 12 taxa with an aligned length of 2132 bp. We determined 18 *mat*K sequences and these were 2715–3740 bp ($\bar{x}$ = 2953 bp)

in length. Four additional sequences were obtained from GenBank (Table 3), yielding 22 taxa with an aligned length of 3869 bp. The combined data set included 4042–7155 bp ($\bar{x}$ = 5186 bp) of newly determined sequences for each of 18 taxa. Together with sequences of *Seymeria macrophylla* and three additional taxa, the combined data set of 22 taxa had an aligned length of 7323 bp.

Sampling from multiple regions of a genome and sampling moderately long total sequence length (approximately 5000 bp or more) have both been shown to significantly improve resolution and support phylogenetic analyses [38-42]. The sequence length used in this study is substantially greater than that included in most single studies in molecular systematics. This amount of sequence data was necessary to meet our challenging objectives of elucidating relationships at multiple levels including potentially relatively recently diverged taxa. In general, longer sequences are more likely to provide more robust estimation of phylogenetic relationships [38,39,41], and improve computational efficiency by increasing differentiation among alternative topologies [41,42]. Because differences in rate variation among different regions [43,44], and differences in rate variation among different codon positions within genes is higher than overall rate variation among genes, all regions sequenced provided some phylogenetic information across all levels of investigation.

Although the greatest rate difference is among the three codon positions within each gene region, the pairwise distances among taxa based on the three gene regions differed (Figure 1), as might be expected. Using the slope of the relationships of the pairwise distances between each pair of taxa as estimated for each pair of gene regions based on linear models we have estimated that the *mat*K gene region evolves 1.615 × faster than *rbc*L, and *ndh*F evolves 2.184 × faster than *rbc*L in these taxa. A simple generalization was not possible for comparing *mat*K and *ndh*F because the estimated linear model fit to the pairwise distance values for these two gene regions crossed null model expectation of rate equality (Figure 1).

Further analysis suggests that there is rate heterogeneity across lineages among the taxa we examined when *Lindenbergia* is used as the outgroup for rooting. A test of the molecular clock hypothesis based on the likelihood ratio of the topology with the highest log likelihood (Figure 2) with (-18764.414) and without (-18429.128) assuming a molecular clock significantly rejected the molecular clock hypothesis ($P < 0.001$).

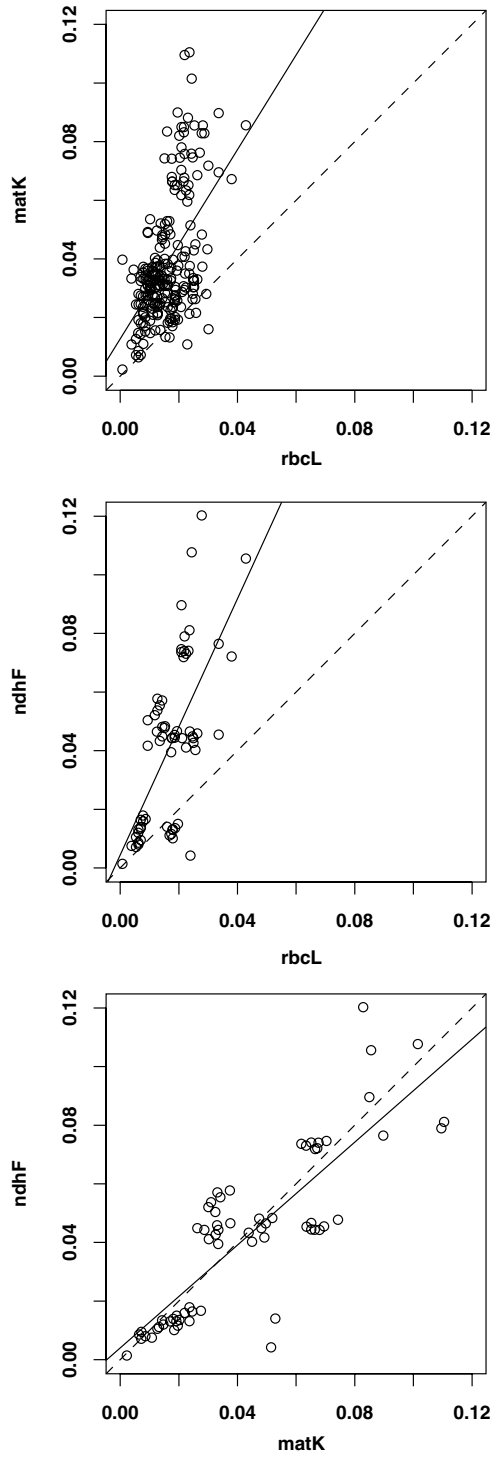

Figure 1
Plots depicting the relationship of pairwise maximum likelihood distances between taxa estimated from different gene regions Dashed lines represent the null hypothesis of equal distance for the gene regions being compared. Solid line represent the simple linear model estimated from the data: *mat*K = 0.013 + 1.615 × *rbc*L; *ndh*F = 0.004 + 2.184 × *rbc*L; *ndh*F = 0.004 + 0.878 × *mat*K.

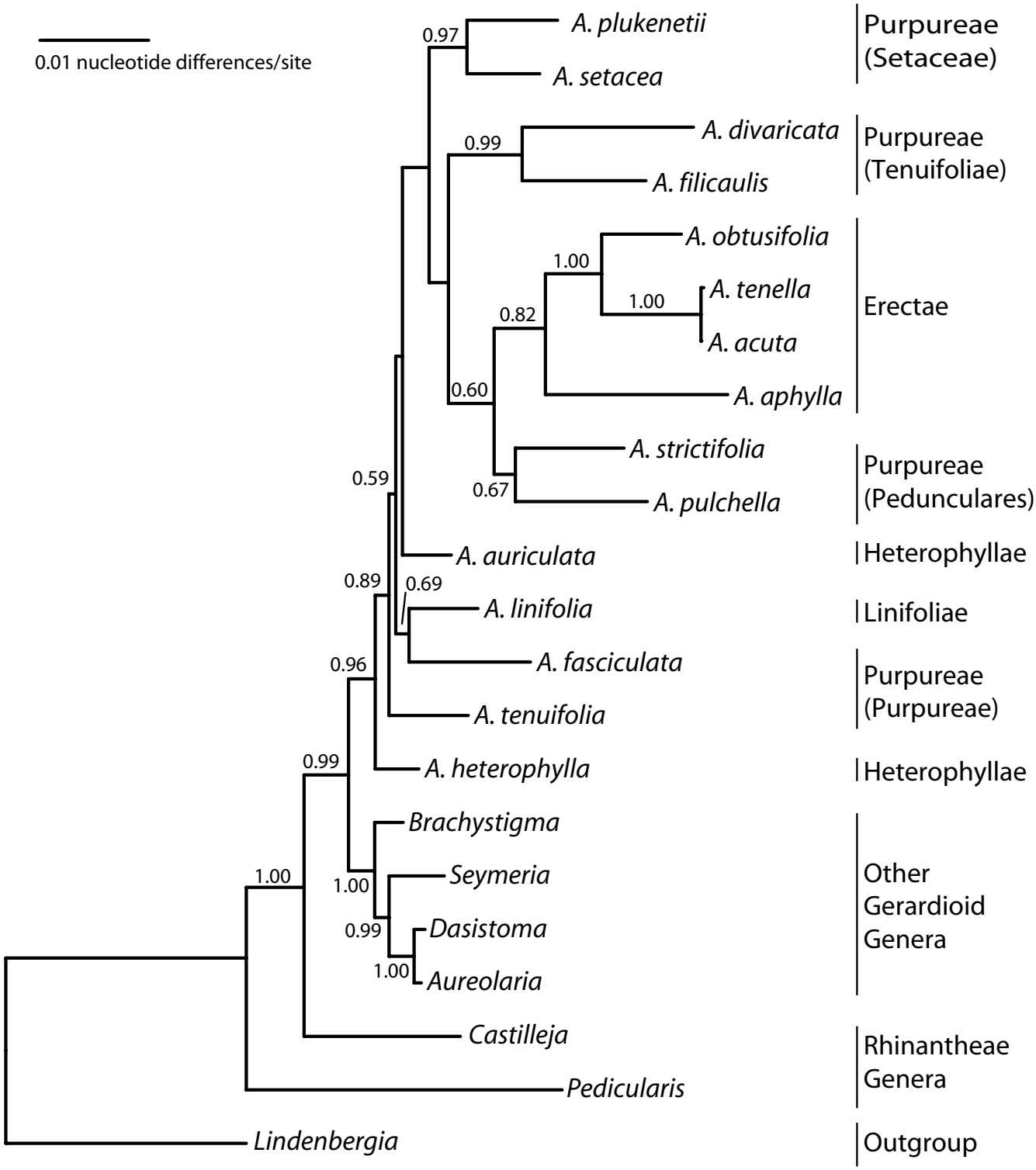

Figure 2
Phylogenetic tree depicting the relationships among species of *Agalinis*, other gerardioid genera, and other genera within Rhinantheae inferred by maximum likelihood with branch lengths proportional to the amount of inferred nucleotide differences. Numerals adjacent to branches denote the proportion of 2000 bootstrap replicates supporting the clade. The ln likelihood value is -18429.128.

Phylogenetic relationships

The best fit likelihood model based on both likelihood ratio tests and Akaike Information Criterion included six nucleotide substitution rate parameters, gamma distributed substitution rates and some invariant sites (*i.e.*, GTR + G + I model). Successive heuristic searching and parameter value estimation yielded a tree of highest log likelihood of -18429.128. The principal implications of this tree are discussed below.

Agalinis, as represented by the North American taxa we sampled and including *Tomanthera*, is monophyletic with a bootstrap value of 0.96 (Figure 2). We also show that *Agalinis* is sister to a group composed of the gerardioid genera with yellow or red corollas: *Aureolaria*, *Brachystigma*, *Dasistoma*, and *Seymeria* with a bootstrap value of 0.99 (Figure 2). These taxa as a whole form a monophyletic group within Rhinantheae as previously suggested based on sequences from *Agalinis* and *Seymeria* [24].

Our results suggest that the current section-level classification within the genus needs revision. While a number of sectional and subsectional classifications appear to be natural (*i.e.*, taxa within some sections and subsections form monophyletic groups), there are numerous cases of polyphyly. Sections Erectae and Tenuifoliae as defined by Canne-Hilliker (Table 1) appear to be the only proposed sections comprising more than one taxon that are monophyletic (Figure 2). In contrast, sections Purpureae and Heterophyllae are clearly polyphyletic. Section Chytra proposed by Pennell [4] roughly in place of section Purpureae also is not monophyletic. Section Purpureae is the largest proposed section in *Agalinis* and is usually treated as being composed of multiple subsections (Table 1). While the section is not monophyletic, each of the subsections as they are represented in our sample except subsection Purpureae appears to be monophyletic (Figure 2). One member of subsection Purpureae, *A. fasciculata*, appears to be most closely related to *A. linifolia* although the bootstrap support is low. *Agalinis linifolia* has always been considered sufficiently distinct from the rest of the members of the genus due to its perennial life history as well as anatomical characters of the stems, roots, and leaves, to be placed in its own section (Table 1 and citations therein). While nodes in the vicinity of *A. linifolia* do not have high bootstrap values, it does not appear that this species is basal or that it forms a single-species sister group to all other species in the genus (Figure 2). The placement of the other member of subsection Purpureae, *A. tenuifolia*, is also not well supported as indicated by low bootstrap values in our analysis. However, our data do support removing *A. tenuifolia* from section Tenuifoliae, as has been suggested based on stem anatomy [23].

Relationships of the other subsections within Purpureae are also weakly supported. The maximum likelihood tree indicates that Subsection Pedunculares is a sister group to section Erectae and together they appear to form a monophyletic clade, although the bootstrap value was low (0.60). This clade unites the two groups within the genus that have a chromosome number of $n = 13$. This result is surprising because subsection Pedunculares and section Erectae have been considered only distantly related due to differences in floral and vegetative characters [5,21,22]. This finding indicates that Pennell's submersion of subsection Pedunculares into subsection Setaceae (Table 1) [4] was in error. While the results support the resurrection of subsection Pedunculares proposed by Canne-Hilliker (Table 1), they do not support inclusion of subsection Pedunculares in section Purpureae. It appears that the inferred ancestral chromosome number for *Agalinis* is $n = 14$, and that there was a single chromosome number reduction within the genus.

The two species in section Tenuifoliae in our sample form a strongly supported monophyletic group (bootstrap value = 0.99); a result supporting the naturalness of this section. Section Tenuifoliae appears as a sister group to the $n = 13$ taxa (section Erectae and subsection Pedunculares of section Purpureae); however, this relationship has low bootstrap support (< 0.50). Section Tenuifoliae was originally considered to be more closely related to section Purpureae [3,4]. Section Erectae was originally considered only distantly related to other sections based on its lighter, yellow-green foliage, lack of tannins and numerous inflorescence and floral characteristics [4]. More recently detailed analysis has found many similarities between Tenuifoliae and Erectae [22] that are corroborated by our data.

Although the three groups Tenuifoliae, Pedunculares, and Erectae form a clade; the bootstrap values uniting these groups and uniting the Pedunculares and Erectae to the exclusion of the Tenuifolieae are low to moderate. This clade is interesting because as discussed above it unites two groups that have been considered closely related based on a number of morphological characters (Erectae and Tenuifoliae) and two groups that are morphologically distinct but that share chromosome number (Erectae and Pedunculares). While bootstrap support uniting Erectae and Pedunculares as sister groups is only 0.60, the alternative grouping of Erectae and Tenuifoliae as sister groups never appeared in any of the 2000 bootstrap replicates. Additional molecular data are necessary to provide statistically significant support for relationships among these taxa.

Our results also clarify placement of certain taxa whose relationships have been debated such as placement of *A.*

aphylla in section Erectae [22]. Pennell originally considered this taxon to be in its own subsection (Aphyllae) within section Purpureae based on its minute, scale-like, appressed leaves [3]. Later he placed *A. aphylla* in subsection Setaceae of section Chytra [4]. Canne-Hilliker moved *A. aphylla* to section Erectae based on chromosome number (n = 13) [22], seed characteristics [5], and stem anatomy [23]. That move is strongly supported by our data.

In another case, *A. auriculata*, which was placed in the genus *Tomanthera* [3,4], clearly falls within the genus *Agalinis* as has been proposed [23]. The genus *Tomanthera* was distinguished from *Agalinis* by lack of the yellow guide lines on the lower corolla lip and by having large, lobed leaves; foliaceous calyx lobes; retrorsely hispid stems; raised seed reticulations; and reduced anther cells on the posterior stamens [3]. Whereas we demonstrate that *A. auriculata* is part of *Agalinis* it is not sister taxon to the other species in section Heterophyllae we sampled (*A. heterophylla*; Figure 2), making the section Heterophyllae polyphyletic. Although *A. auriculata* falls well within *Agalinis* (i.e., it shares a number of inferred common ancestors with other *Agalinis* species), *A. heterophylla*, another species placed within section Heterophyllae, is basal to the rest of the species in the genus.

One of our objectives was to evaluate the evolutionary distinctiveness of *A. tenella* from *A. obtusifolia*, and *A. acuta* from *A. tenella*. We were able to determine that *A. tenella* and *A. obtusifolia* are not synonymous as has been suggested [35], and thus submerging *A. tenella* is not warranted. These taxa are closely related, but are evolutionarily distinct (Figure 2). The branch length from the inferred most recent common ancestor to *A. obtusifolia* is significant (0.00738, S.E. = 0.00112, $P < 0.001$), as is the branch length to *A. tenella* (0.00026, S.E. = 0.00026, $P = 0.003$). The number of pairwise differences between *A. obtusifolia* and *A. tenella* over 3657 aligned nucleotides positions includes 66 substitutions (0.018) and 5 indels involving 7 positions. The differentiation of these taxa is important because *A. tenella* has been considered to be imperiled in at least two states but merging these taxa would eliminate *A. tenella* for consideration for protection.

In contrast, we found very little divergence between *A. acuta* and *A. tenella*, indeed the smallest amount of divergence among any of the taxa we examined (Figure 2); one substitution over 4048 aligned nucleotide positions (0.0002). Unfortunately, we were unable to obtain *ndh*F sequence from *A. tenella*, thus limiting our ability to detect divergence. It is critical to more thoroughly examine this issue using molecular markers that will provide sufficient resolution to determine if these two taxa are really on independent evolutionary trajectories. Given the apparent lack of differentiation in chloroplast genes, yet recognized morphological differentiation, it appears be that nuclear genome markers (e.g., expressed sequence tags [ESTs] and/or microsatellites) sampled from multiple populations combined with coalescent-based analysis and detailed morphological measurement would be useful for studying the relationships of these taxa.

Beyond the genus *Agalinis* we were able to provide some insights into relationships among the gerardioid genera. Previous thoughts regarding evolution of the these genera that *Aureolaria* is the most primitive genus and closely resembles the common ancestor of the group [14] are clearly incorrect. *Aureolaria* is among the most derived genera in our sample (Figure 2). Further, there is a close relationship between *Aureolaria* and *Dasistoma*. Divergence between *D. macrophylla* and *A. pedicularia* is less than divergence between all pairs of *Agalinis* species except *A. acuta* and *A. tenella*. Similarities in vegetative parts in *Aureolaria* and *Dasistoma* have been noted, but the two genera have been considered distinct based on floral morphology [14]. On the whole the differentiation among the gerardioid genera (as indicated by relatively short branch lengths) is modest compared to that among species of *Agalinis* and among other Rhinantheae genera (Figure 2).

Conclusions

As the first molecular systematic study and phylogenetic analysis of *Agalinis*, this research contributes toward understanding of relationships among taxa in the genus. It provides support for some, and refutes other, previous suggestions regarding classification of a number of species, subsections and sections. Furthermore this work contributes to understanding relationships among members of the Orobanchaceae in general.

Phylogenetic analysis supports the monophyly of *Agalinis*, including species formerly in *Tomanthera*, and this group as sister to the gerardioid genera *Aureolaria*, *Brachystigma*, *Dasistoma*, and *Seymeria*. Many of the previously described sections within *Agalinis* are polyphyletic, although many of the subsections appear to form natural groups. The analysis reveals a single evolutionary event leading to a reduction in chromosome number from n = 14 to n = 13 based on the sister group relationship of section Erectae and section Purpureae subsection Pedunculares. Our results establish the evolutionary distinctiveness of *A. tenella*, as species of conservation concern, from the more widespread and common *A. obtusifolia*. However, further data are required to clearly resolve the relationship of *A. acuta* and *A. tenella*.

Methods

Taxa sampled

A total of 22 taxa were included in our study (Tables 2 and 3). We sampled 15 taxa representing all North American sections of the genus *Agalinis* as well as *Aureolaria pedicularia* (L) Raf., *Brachystigma wrightii* (A. Gray) Pennell, and *Dasistoma macrophylla* (Nutt.) Raf. (Tables 2 and 3). Sequences from four additional taxa (*Seymeria pectinata* Pursh, *Castilleja linariifolia* Benth., *Pedicularis foliosa* L, and *Lindenbergia philippensis* (Cham.) Benth.) were obtained from GenBank [30] (Table 3) to help elucidate phylogenetic relationships among *Agalinis* and the other gerardioid genera.

DNA preparation

DNA was isolated from fresh or frozen leaves and flower buds by grinding 50–75 mg of tissue to powder in liquid nitrogen with a mortar and pestle, and then using GenElute Plant Genomic DNA Kits (Sigma Chemical Company, St. Louis, Missouri, USA) following manufacturer's instructions.

Sequence regions and PCR amplification

We sampled sequences of the following chloroplast gene regions: *rbc*L, which encodes the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase; *mat*K, which includes partial sequence of the gene for ribosomal protein S16 (*rps*16), an intergenic spacer, the 5' exon of lysine tRNA (*trn*K), the gene for maturase K (*mat*K), an intron sequence and the 3' exon of lysine tRNA (*trn*K), and a partial gene for photosystem II D1 protein (*psb*A); and *ndh*F which encodes NADH dehydrogenase subunit F.

We selected these gene regions for sequencing because previous studies indicated that *ndh*F and *mat*K are among the most rapidly evolving protein coding genes in the chloroplast genome [44,45], and thus have been firmly established as useful regions at the levels of divergence that are the primary focus of this research. *rbc*L evolves more slowly, and collectively the three regions provide information across the range of divergence levels we examined (from among species within *Agalinis* to among genera within Orobanchaceae). Further, the majority of these regions are protein coding thus minimizing insertions and deletions and allowing unambiguous alignments. Finally, the large amount of comparative data from other studies of these gene regions in angiosperms [45,46] and particularly in the Scrophulariaceae *sensu lato* [25-28] allowed us to incorporate existing data for other taxa, provided a strong comparative framework for our results, and allowed us to take advantage of primer sequences designed and tested by others.

PCR amplification

Polymerase chain reactions (PCR) were based on Eppendorf MasterTaq PCR kits (Brinkman, Westbury, New York, USA) run on an MJ Research PTC-200 Thermal Cycler. For *rbc*L we used primer sequences Z-1 and Z-1375 from Zurawski et al. [47] to amplify the whole target region. We also used their internal primer Z-1204R and five internal primers of our own design for sequencing. For the *mat*K region we used primers *rps*16-4547F and *psb*A-R from Johnson and Soltis [45] to amplify the whole region. We used seven of their internal sequencing primers and six primers of our own design. For *ndh*F we used primers 1 and 2110R from Olmstead and Sweere [41] to amplify the entire *ndh*F region, six of their internal sequencing primers, and two primers of our own design (Additional file 1). Specific amplification conditions for each primer combination in each of gene region varied (1). In general, the PCR temperature profile was 30 cycles of 94°C for 60 s, annealing temperature set approximately 5°C below the lower of the two primer melting temperatures for 90 s, 72°C for 150 s, and a final 15 min elongation period at 72°C. PCR products were separated by agarose gel electrophoresis and DNA from fragments of the expected size was extracted and purified using the QIAQuick DNA cleanup system according to manufacturer's instructions (Qiagen Inc., Valencia, California, USA).

DNA sequencing

Direct sequencing of PCR-generated templates was done using reactions based on the chemistry of BigDye Terminator v3.0 Cycle Sequencing Ready Reaction Kits (Applied Biosystems, Foster City, California, USA) with reactions set up in 96-well microtiter plates using a robotic workstation. Cycle sequencing was performed on MJ Research PTC-200 Thermal Cyclers. Sequencing reactions were cleaned by isopropanol precipitation to remove unincorporated labeled terminators prior to running the samples on an Applied Biosystems 3700 DNA Analyzer. We conducted bi-directional sequencing with ≥ fourfold coverage to ensure high accuracy of the sequence data.

Data analysis

Sequence trace curves were collected on the computer controlling the sequencer. After completion of each set of sequencing runs the trace curves were transferred as SCF-formatted files to a Linux workstation for all subsequent processing and analysis. Base calling and quality assignments were made using the program phred [48,49]. Data from individual sequencing runs were assembled into final complete sequence using the program phrap [50]. Contigs were evaluated with the help of the program consed [51].

Multiple alignments for each gene region were performed using ClustalW [52] and edited if deemed appropriate.

These multiple alignments were simple and results were unambiguous because the majority of the sequences coded for proteins and showed relatively few insertion or deletion events. Multiple alignments for each gene region were concatenated to form a single combined data set.

Phylogenetic relationships were determined by maximum likelihood analysis [53] of the aligned nucleotide sequences. Data exploration was done to determine the most appropriate model using preliminary phylogenetic trees and the programs Modeltest [54] and PAUP* [55]. Successive heuristic searching (with multiple random taxon addition and tree bisection-reconnection branch swapping) and model parameter value estimation was done to find the highest likelihood tree using PAUP*. Support for specific relationships was assessed with the bootstrap [56] using PAUP*. For the bootstrap analysis, the parameters of the likelihood model were set to those of the highest likelihood tree based on the original data. For each of 2000 replicates a single simple addition heuristic search was conducted with tree bisection-reconnection branch swapping. Estimating likelihood parameter values and applying them to bootstrap replicates is more efficient then re-estimating the values for each bootstrap replicate [57]. As a supplementary evaluation, we estimated branch lengths with standard errors, and tested their significance with likelihood ratio tests using PAUP*.

To compare the rates of evolution of the three gene regions, we estimated pairwise distances among all pairs of taxa using the same likelihood model described above using PAUP* [55]. We used linear models as implemented by the R system for statistical computing [58] to describe the relationships of distances among these taxon pairs as estimated by each of the three gene regions. We also tested the molecular clock hypothesis (i.e., that all lineages evolved at the same rate) based on the likelihood ratio of the topology with the highest log likelihood with and without assuming a molecular clock.

Authors' contributions

MCN conceived of the study and participated in the molecular genetic analyses. MPC carried out the data analyses. Both authors participated in the design of the study, interpreted results and drafted the manuscript. Both authors read and approved the final manuscript.

Additional material

Additional File 1

Primers used and their estimated melting temperatures (Tm)
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-4-15-S1.pdf]

Additional File 2

Basic PCR Recipes (using Eppendorf MasterTaq PCR kit; Brinkman, Westbury, New York, USA) and PCR conditions
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-4-15-S2.pdf]

Acknowledgements

We thank JM Canne-Hilliker, G Dieringer, J Koontz, D Lewis, JL Neff, P. Polloni, and P. Somers for assistance with collecting specimens; JM Canne-Hilliker for helpful discussions regarding the taxonomy of *Agalinis*; and AR Cohen, E Lasek-Nesselquist, and LA McInerney for help with laboratory work.

References

1. von Wettstein R: **Scrophulariaceae.** In *Die Natürlichen Pflanzenfamilien Volume 4*. Issue 3b Edited by: Engler A, Prantl K. Leipzig: Wilhelm Engelmann; 1891:39-107.
2. Pennell FW: **Studies in the Agalinanae, a subtribe of the Rhinanthaceae. I. nomenclature of the neartic genera.** *B Torrey Bot Club* 1913, **4:**119-130.
3. Pennell FW: ***Agalinis* and allies in North America – II.** *P Acad Nat Sci Phila* 1929, **81:**111-249.
4. Pennell FW: **The Scrophulariaceae of eastern temperate North America (41 *Gerardia*).** *Academy of Natural Sciences of Philadelphia Monographs* 1935, **1:**419-476.
5. Canne JM: **A light and scanning electron microscope study of seed morphology in *Agalinis* (Scrophulariaceae) and its taxonomic significance.** *Syst Bot* 1979, **4:**281-296.
6. Canne JM: **Seed surface features in *Aureolaria*, *Brachystigma*, *Tomanthera*, and certain South American *Agalinis* (Scrophulariaceae).** *Syst Bot* 1980, **5:**241-252.
7. Barringer K: **Two new species of *Agalinis* (Scrophulariaceae) from South America.** *Brittonia* 1985, **37:**352-354.
8. Canne-Hilliker JM: ***Agalinis* (Scrophulariaceae) in Peru and Bolivia.** *Brittonia* 1988, **40:**433-440.
9. Barringer K: **New and noteworthy South American species of *Agalinis*.** *Brittonia* 1987, **39:**353-357.
10. Dieringer G: **Variation in individual flowering time and reproductive success of *Agalinis strictifolia* (Scrophulariaceae).** *Am J Bot* 1991, **78:**497-503.
11. Neel MC: **Conservation implications of the reproductive ecology of *Agalinis acuta* (Scrophulariaceae).** *Am J Bot* 2002, **89:**972-980.
12. Dieringer G: **Reproductive biology of *Agalinis skinneriana* (Scrophulariaceae), a threatened species.** *J Torrey Bot Soc* 1999, **126:**289-295.
13. Stewart HM, Stewart SC, Canne-Hilliker JM: **Mixed mating system in *Agalinis neoscotica* (Scrophulariaceae) with bud pollination and delayed pollen germination.** *Int J Plant Sci* 1996, **157:**501-508.
14. Pennell FW: ***Agalinis* and allies in North America – I.** *P Acad Nat Sci Phila* 1928, **80:**339-449.
15. Pennell FW: **Studies in the Agalinanae, a subtribe of the Rhinanthaceae. II. Species of the Atlantic coastal plain.** *B Torrey Bot Club* 1913, **4:**401-439.
16. Canne JM: **The taxonomic significance of seedling morphology in *Agalinis* (Scrophulariaceae).** *Can J Bot* 1983, **61:**1868-1874.
17. Canne-Hilliker JM: **Patterns of floral development in *Agalinis* and allies (Scrophulariaceae). II. Floral development of *Agalinis densiflora*.** *Am J Bot* 1987, **74:**1419-1430.
18. Kampny CM: **Aspects of development in Scrophulariaceae, striking early differences in three tribes.** In *Aspects of Floral Development* Edited by: Leins P, Tucker S, Endress P. Berlin: J Cramer; 1988:147-157.
19. Kampny CM, Canne-Hilliker JM: **Patterns of floral development in *Agalinis* and allies (Scrophulariaceae) I. floral development of *Agalinis fasciculata* and *A. tenuifolia*.** *Can J Bot* 1987, **65:**2255-2262.

20. Stewart HM, Canne-Hilliker JM: **Floral development of *Agalinis neoscotica*, *Agalinis paupercula* var. *borealis* and *Agalinis purpurea* (Scrophulariaceae): implications for taxonomy and mating system.** *Int J Plant Sci* 1998, **159:**418-439.
21. Canne JM: **Chromosome counts in *Agalinis* and related taxa (Scrophulariaceae).** *Can J Bot* 1981, **59:**1111-1116.
22. Canne JM: **Chromosome numbers and the taxonomy of North American *Agalinis* (Scrophulariaceae).** *Can J Bot* 1984, **62:**454-456.
23. Canne-Hilliker JM, Kampny CM: **Taxonomic significance of leaf and stem anatomy of *Agalinis* (Scrophulariaceae) from the U.S.A. and Canada.** *Can J Bot* 1991, **69:**1935-1950.
24. dePamphilis CW, Young ND, Wolfe AD: **Evolution of plastid gene *rps*2 in a lineage of hemiparasitic and holoparasitc plants: many losses of photosynthesis and complex patterns of rate variation.** *Proc Natl Acad Sci, USA* 1997, **94:**7367-7372.
25. Wolfe AD, dePamphilis CW: **The effect of relaxed functional constraints on the photosynthetic gene *rbc*L in photosynthetic and nonphotosynthetic parasitic plants.** *Mol Biol Evol* 1998, **15:**1243-1258.
26. Young ND, dePamphilis CW: **Purifying selection detected in the plastid gene *mat*K and flanking ribozyme regions within a group II intron of non-photosynthetic plants.** *Mol Biol Evol* 2000, **17:**1933-1941.
27. Olmstead RG, dePamphilis CW, Wolfe AD, Young ND, Elisons WJ, Reeves PA: **Disintegration of the Scrophulariaceae.** *Am J Bot* 2001, **88:**348-361.
28. Olmstead RG, Reeves PA: **Evidence for the polyphyly of the Scrophulariaceae based on chloroplast *rbc*L and *ndh*F sequences.** *Ann Mo Bot Gard* 1995, **82:**176-193.
29. Young ND, Steiner KE, dePamphilis CW: **The evolution of parasitism in Scrophulariaceae/Orobanchaceae: plastid gene sequences refute an evolutionary transition series.** *Ann Mo Bot Gard* 1999, **86:**876-893.
30. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL: **GenBank.** *Nucleic Acids Res* 2004, **32:**D23-26.
31. **NatureServe** [http://www.natureserve.org/explorer/]
32. **The Plants Database. Version 3.5** [http://plants.usda.gov/]
33. **Flora of the Carolinas and Virginia** [http://www.herbarium.unc.edu/flora.htm]
34. **South Carolina Rare, Threatened & Endangered Species Inventory** [http://www.dnr.state.sc.us/pls/heritage/county_species.list?pcounty=all]
35. Kartesz JT: **A synonymized checklist and atlas with biological attributes for the vascular flora of the United States, Canada, and Greenland.** In *Synthesis of the North American Flora* 1.0th edition. Edited by: Kartesz JT, Meacham CA. Chapel Hill: North Carolina Botanical Garden; 1999.
36. US Fish and Wildlife Service: **Determination of *Agalinis acuta* to be an endangered species.** *Federal Register* 1988, **53:**34701-34705.
37. Clegg MT, Brown GM Jr, Brown WY, Fink WL, Harte J, Houck OA, Lynch M, Maguire LA, Murphy DD, O'Brien PY, Pickett STA, Pulliam HR, Ralls K, Simpson BB, Sparrowe RD, Steadman DW, Sweeney JM: *Science and the Endangered Species Act* Washington DC: National Research Council; 1995.
38. Cummings MP, Otto SP, Wakeley J: **Sampling properties of DNA sequence data in phylogenetic analysis.** *Mol Biol Evol* 1995, **12:**814-822.
39. Cummings MP, Otto SP, Wakeley J: **Genes and other samples of DNA sequence data for phylogenetic inference.** *Biol Bull* 1999, **196:**345-350.
40. Otto SP, Cummings MP, Wakeley J: **Inferring phylogenies from DNA sequence data: the effects of sampling.** In *New Uses for New Phylogenies* Edited by: Harvey PH, Brown AJ, Maynard Smith J, Nee S. Oxford: Oxford University Press; 1996:103-115.
41. Olmstead RG, Sweere JS: **Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae.** *Syst Biol* 1994, **43:**467-481.
42. Soltis DE, Soltis PS, Mort ME, Chase MW, Savolainen V, Hoot SB, Morton C: **Inferring complex phylogenies using parsimony: an empirical approach using three large DNA data sets for angiosperms.** *Syst Biol* 1998, **47:**32-42.
43. Clegg MT, Gaut BS, Learn GH, Morton BR: **Rates and patterns of chloroplast DNA evolution.** *Proc Natl Acad Sci, USA* 1994, **91:**6795-6801.
44. Hilu KW, Liang H: **The *mat*K gene: sequence variation and application in plant systematics.** *Am J Bot* 1997, **84:**830-839.
45. Johnson LA, Soltis DE: **Phylogenetic inference in Saxifragaceae sensu stricto and *Gilia* (Polemoniaceae) using *mat*K sequences.** *Ann Mo Bot Gard* 1995, **82:**149-175.
46. Olmstead RG, Michaels HJ, Scott KM, Palmer JD: **Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of *rbc*L.** *Ann Mo Bot Gard* 1992, **79:**249-265.
47. Zurawski G, Perrot B, Bottomley W, Whitfeld PR: **The structure of the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase from spinach chloroplast DNA.** *Nucleic Acids Res* 1981, **9:**3251-3270.
48. Ewing B, Green P: **Base-calling of automated sequencer traces using phred. II. Error probabilities.** *Genome Res* 1998, **8:**186-194.
49. Ewing B, Hillier L, Wendl MC, Green P: **Base-calling of automated sequencer traces using phred. I. Accuracy assessment.** *Genome Res* 1998, **8:**175-185.
50. **phrap** [http://www.phrap.org/]
51. Gordon D, Abajian C, Green P: **Consed: a graphical tool for sequence finishing.** *Genome Res* 1998, **8:**195-202.
52. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22:**4673-4680.
53. Felsenstein J: **Evolutionary trees from DNA sequences: a maximum likelihood approach.** *J Mol Evol* 1981, **17:**368-376.
54. Posada D, Crandall KA: **Modeltest: testing the model of DNA substitution.** *Bioinformatics* 1998, **14:**917-918.
55. Swofford D: *PAUP*: Phylogenetic Analysis Using Parsimony* (*and other methods), Version 4* Sunderland, Massachusetts: Sinauer Associates; 2002.
56. Felsenstein J: **Confidence limits on phylogenies: an approach using the bootstrap.** *Evolution* 1985, **39:**783-791.
57. Cummings MP, Handley SA, Myers DS, Reed DL, Rokas A, Winka K: **Comparing bootstrap and posterior probability values in the four-taxon case.** *Syst Biol* 2003, **52:**477-487.
58. Ihaka R, Gentleman R: **R: a language for data analysis and graphics.** *J Comput Graph Stat* 1996, **5:**299-314.