

## Genetic Variation and Species Boundaries in *Calopogon* (Orchidaceae)

DORSET W. TRAPNELL,<sup>1,3</sup> J.L. HAMRICK,<sup>2</sup> and DAVID E. GIANNASI<sup>1</sup>

<sup>1</sup>Department of Plant Biology, University of Georgia, Athens, Georgia 30602

<sup>2</sup>Departments of Plant Biology and Genetics, University of Georgia, Athens, Georgia 30602

<sup>3</sup>Author for correspondence (dorset@plantbio.uga.edu)

Communicating Editor: John V. Freudenstein

**ABSTRACT.** Morphological and habitat similarities among the five species of the terrestrial orchid genus *Calopogon* have led to nomenclatural and taxonomic confusion. The taxa are marked by subtle character differences and little apparent reproductive isolation. Here we investigate allozyme diversity at the species level and the partitioning of genetic variation within and among species and their populations. Genetic identities are used to define species boundaries and suggest phylogenetic relationships. All five species of *Calopogon* maintain high levels of allozyme variation within their populations ( $P = 50.0\% - 94.4\%$ ,  $AP = 2.67 - 3.32$ ,  $H_e = 0.11 - 0.43$ ). *Calopogon oklahomensis*, an autotetraploid that appears to have undergone gene silencing at 13 of its 19 polymorphic loci, consistently had the highest genetic diversity values. *Calopogon multiflorus*, which has the most restricted range and rarest occurrence, had the lowest mean genetic diversity values. In *C. oklahomensis*, *C. pallidus* and *C. tuberosus* most of the genetic variation exists within rather than among populations ( $G_{ST} = 0.037 - 0.085$ ). The UPGMA phenogram generated using genetic identity data has three phenetic groups and supports designation of the taxa as separate species. The data suggest that *C. oklahomensis* most closely resembles the basal extant taxon within *Calopogon*.

The genus *Calopogon* R. Brown (Orchidaceae) includes five showy, terrestrial species of short-lived, herbaceous perennials known as Grass Pinks (Dressler 1981; 1993). *Calopogon* spp. display considerable morphological similarity, often grow in similar habitats, have overlapping ranges, and overlapping periods of anthesis. Morphological similarities between taxa (e.g. plant height, leaf length and width, number and color of flowers, petal length, width and shape, length and broadness of the lip, hair characters) are so great that taxonomic keys are unclear and characters distinguishing the taxa are often subtle, leading to nomenclatural and taxonomic confusion.

Where the *Calopogon* taxa occur sympatrically there appears to be little reproductive isolation. All *Calopogon* species share a similar pollination system with cross-pollination by bees believed to be the primary breeding mechanism (Dressler 1981; Firmage and Cole 1988; Thien and Marcks 1972). There is evidence of self-compatibility (Thien 1973) although an insect is required for pollen transfer. *Calopogon* is also capable of vegetative reproduction via root sprouting (Thien and Marcks 1972). Their typically deep pink flowers produce no nectar (with the possible exception of *C. oklahomensis* Goldman) and offer no pollen reward. Consequently, flowers must attract pollinators through mimicry of other sympatric species that offer a pollen reward and by a form of deception categorized as pseudopollen (Dafni 1984) which refers to the resemblance of the showy yellow and orange lip hairs to a mass of pollen. When a pollinator is attracted to this display and lands on the labellum, if the insect is heavy enough, the labellum swings down and the posterior of the insect comes into contact with the sticky pollinia located on the end of the column (Thien and Marcks 1972; Firmage and Cole 1988). Pollinators are

therefore varied and non-specific. Like most orchid taxa the tiny, dust-like seeds are wind dispersed (Dressler 1981).

Pace (1909) reported that haploid daughter cells of *C. tuberosus* (L.) B.S.P. had 13 chromosomes, while separate counts of diploid sporophytic tissue revealed approximately 26 chromosomes. These findings were based on only a few counts each from haploid and diploid states. In 1972, Thien and Marcks found that *C. tuberosus* from northern Wisconsin had  $2n = 40$  while a later report by Thien (1973) indicated that *C. tuberosus*, *C. pallidus* Chapman, *C. barbatus* (Walter) Ames, and *C. multiflorus* Lindley root tip cells possess  $2n = 42$  chromosomes. *Calopogon oklahomensis* is a recently described species (Goldman 1995) for which a chromosome number has not been reported.

Isolating mechanisms among the five congeners are habitat, seasonal separation, and to some extent, flower size, which necessitates pollinators of different size and weight (Thien 1973). The most common and widespread species, *C. tuberosus*, ranges from Canada to Cuba and the Bahamas and across the eastern United States. This species is comprised of two varieties, *Calopogon tuberosus* var. *tuberosus* and *C. tuberosus* var. *simpsonii* (Small) Magrath. *Calopogon tuberosus* var. *simpsonii* is found on calcareous soils in swamps and prairies in southern Florida, Bahamas and Cuba. This variety tends to have longer inflorescences that represent a larger proportion of the total plant height. *Calopogon pallidus* and *C. barbatus* are restricted to the coastal plain of the southeastern United States ranging from North Carolina to Mississippi. *Calopogon multiflorus*, which grows in open, moist pine flatwoods and meadows, has a more restricted range, being confined to Florida and its immediate borders. *Calopogon oklahomensis* occurs in mesic, acidic, sandy-loam prairies in

Arkansas, southeastern Kansas, Missouri, and eastern Oklahoma (Goldman 1995).

Reliance on morphological characters for delineation of *Calopogon* species has led to taxonomic confusion. Examination of the underlying genetic diversity of the species and the degree of genetic divergence among taxa is a useful way to elucidate the taxonomic delineation. Knowledge of the genetic variation housed within a species is also valuable since it can provide insights into the species' long-term survival potential. This information is essential for intelligent decision-making regarding conservation recommendations and policies. Genetic diversity increases the likelihood that a species can survive disease, environmental fluctuations, and natural catastrophes that invariably occur over time. The objective of this study was to determine the extent of genetic diversity within each *Calopogon* species and to describe the partitioning of genetic variation within and among species and their populations. Levels of genetic identity between these taxa are quantified and used to assess species boundaries and suggest phylogenetic relationships.

#### MATERIALS AND METHODS

*Calopogon* belongs to subfamily Epidendroideae Lindley, tribe Arethuseae Lindley, and subtribe Bletinae Benth. The two outgroup species included in this study, *Bletia purpurea* (Lamarck) A. de Candolle and *Arethusa bulbosa* L., were selected for their moderate taxonomic distance from *Calopogon* (Goldman et al. 2001). *Bletia purpurea* belongs to the same subtribe as *Calopogon* (Dressler 1981; 1993) and both outgroup genera belong to the same tribe as *Calopogon*. Both outgroup species also have overlapping geographic ranges with some of the *Calopogon* species. Mature leaf or stem tissue was randomly collected from 15 populations representing the five *Calopogon* species and the two outgroup species. Voucher specimens from populations from which plant material was collected are listed in Appendix 1. Forty-eight individuals from each population were collected wherever possible (Appendix 1). Leaves were wrapped in moist paper towels, placed in a sealed plastic bag, and kept chilled to prevent protein denaturation. Care was taken to avoid further trauma to the tissue from either freezing or creasing.

Within 24–72 hrs of collection, the tissue was clipped into small pieces, placed in a chilled mortar, and crushed with a pestle and a pinch of sea sand to disrupt cellular compartmentalization. Stem material was crushed in liquid nitrogen with sea sand. Enzymes were extracted from the tissue with a polyvinylpyrrolidone-phosphate extraction buffer (Mitton et al. 1979). The resulting slurry containing crude protein extract was absorbed onto 4 × 6 mm wicks punched from Whatman 3 mm chromatography paper. Wicks were stored in microtest plates at –70°C until used for electrophoresis.

Wicks were placed in horizontal gels composed of 10% potato starch and electrophoresis was performed. Twelve enzyme stains in four buffer systems resolved 21 putative allozyme loci. Enzymes stained and loci identified (in parentheses) for each of the four buffer systems were: 1) system 6; alcohol dehydrogenase (ADH1, ADH2) and fluorescent esterase (FE4), 2) system 7; aspartate aminotransferase (AAT2, AAT3, AAT4), diaphorase (DIA1, DIA2), menadiene reductase (MNR2), 3) system 8; fluorescent esterase (FE1, FE2), menadiene reductase (MNR1), and triosephosphate isomerase (TPI1, TPI2) and 4) system 11; adenylate kinase (AK1), isocitrate dehydrogenase (IDH), malate dehydrogenase (MDH1,

MDH2), 6-phosphogluconate dehydrogenase (6-PGD), phosphoglucoisomerase (PGI1), phosphoglucomutase (PGM2). All stain recipes were adapted from Soltis et al. (1983) except for diaphorase which was taken from Cheliak and Pitel (1984). Buffer system 8 is a modification of buffer system 8 as described by Soltis et al. (1983). Two standard wicks from individuals of *C. tuberosus* and *C. pallidus* were placed on each gel. Banding patterns were consistent with those expected for each enzyme system (Weeden and Wendel 1989).

Levels of allozyme diversity were estimated within species and for individual populations using a computer program designed by M.D. Loveless and A.F. Schnabel. Measures of genetic diversity were percent polymorphic loci,  $P$  (a species was treated as polymorphic at a locus if two alleles were detected); mean number of alleles per locus,  $A$ ; mean number of alleles per polymorphic locus,  $AP$ ; effective number of alleles per locus,  $A_e = 1/\sum p_i^2$ ; and genetic diversity,  $H_e = 1 - \sum p_i^2$  (Nei 1973), where  $p_i$  is the frequency of the  $i$ th allele, which is the proportion of loci heterozygous per individual under Hardy-Weinberg expectations.

Observed heterozygosity ( $H_o$ ) was compared with Hardy-Weinberg expected heterozygosity for each polymorphic locus in each population by calculating Wright's fixation indices ( $F$ ; Wright 1922) and testing for significant deviations using  $\chi^2 = F^2N(a - 1)$ ;  $df = a(a - 1)/2$  where  $N$  is the total number of individuals analyzed and  $a$  is the number of alleles at the locus (Li and Horvitz 1953).

Variation among populations was estimated using Nei's (1973) measures of genetic diversity. Total genetic diversity ( $H_T$ ), mean genetic diversity within populations ( $H_S$ ), and mean genetic diversity among populations ( $D_{ST}$ ) such that  $H_T = H_S + D_{ST}$ , were determined for each polymorphic locus. The proportion of genetic variation that occurs among populations ( $G_{ST}$ ) was calculated for each polymorphic locus by  $G_{ST} = D_{ST}/H_T$  (Nei 1973) and averaged across loci. Heterogeneity in allele frequencies among populations was tested by  $\chi^2 = 2NG_{ST}(a - 1)$ ;  $df = (a - 1)(k - 1)$ , where  $N$  is the total number of individuals analyzed,  $a$  is the number of alleles at the locus, and  $k$  is the number of populations (Workman and Niswander 1970).

Nei's (1972) genetic identities and distances were calculated for each pair-wise combination of populations and species. A UPGMA phenogram of genetic identities was generated using NTSYS-PC (Rohlf 2000). A Mantel test of correspondence between genetic distances and geographic distances was performed (Smouse et al. 1986).

#### RESULTS

**Population Level Variation.** All species of *Calopogon* displayed high levels of genetic diversity (Table 1). Of the five species, *C. multiflorus* had the least genetic variability ( $P = 50.0\%$ ,  $AP = 2.67$ ,  $H_e = 0.11$ ) while *C. oklahomensis* had the highest genetic diversity values ( $P = 94.4\%$ ,  $AP = 3.32$ ,  $H_e = 0.43$ ). Outgroup species had lower levels of genetic variability with values of  $P = 30.8\%$ ,  $AP = 2.00$ , and  $H_e = 0.02$  for *B. purpurea* and  $P = 41.2\%$ ,  $AP = 2.14$ , and  $H_e = 0.13$  for *A. bulbosa*.. The only two species for which the observed mean heterozygosity was higher than Hardy-Weinberg expectations were *C. barbatus* and *A. bulbosa*. *Calopogon oklahomensis* appears to be an autotetraploid displaying banding patterns consistent with tetrasomic inheritance at six of its 19 polymorphic loci (IDH1, MDH1, MDH2, PGM2, 6-PGD, and TPI1). These six loci had an unusually high proportion of heterozygotes, and both balanced and unbalanced heterozygotes (corre-

TABLE 1. Summary of allozyme variation at 21 loci for the five *Calopogon* species. Variation is described by proportion of all loci that are polymorphic (P; if a second allele is observed at a locus it is considered polymorphic), mean number of alleles per polymorphic locus (AP), mean number of alleles per locus (A), mean effective number of alleles per locus ( $A_e$ ), mean observed heterozygosity ( $H_o$ ), and mean expected heterozygosity ( $H_e$ ). Standard deviations are shown in parentheses. Gene frequency data are available from DWT upon request.

Population	P (%)	AP	A	$A_e$	$H_o$ (sd)	$H_e$ (sd)
<i>Calopogon tuberosus</i>						
Ct-FL1	71.4	2.93	2.38	1.44	0.153 (0.012)	0.194 (0.051)
Ct-FL2	82.4	2.79	2.47	1.39	0.180 (0.014)	0.187 (0.051)
Ct-SC	88.9	3.13	2.89	1.54	0.209 (0.014)	0.250 (0.056)
Ct-TN	66.7	2.50	2.00	1.38	0.176 (0.013)	0.190 (0.051)
Ctsimp-FL	31.6	2.17	1.37	1.12	0.078 (0.009)	0.075 (0.032)
Population level mean	68.2	2.70	2.22	1.37	0.159 (0.006)	0.179 (0.022)
Species level	90.5	3.42	3.19	1.46	—	0.215
<i>Calopogon pallidus</i>						
Cp-FL1	68.4	3.46	2.68	1.63	0.215 (0.014)	0.252 (0.060)
Cp-FL2	70.0	2.79	2.25	1.48	0.168 (0.012)	0.204 (0.053)
Cp-FL3	61.9	3.23	2.38	1.50	0.206 (0.013)	0.217 (0.056)
Cp-SC	55.6	3.20	2.22	1.63	0.251 (0.024)	0.258 (0.067)
Population level mean	64.0	3.17	2.38	1.56	0.210 (0.008)	0.233 (0.030)
Species level	81.0	3.59	3.10	1.59	—	0.252
<i>Calopogon oklahomensis</i>						
Co-AR	95.0	3.37	3.25	2.08	0.396 (0.054)	0.442 (0.050)
Co-OK	93.8	3.27	3.13	1.99	0.348 (0.053)	0.415 (0.056)
Population level mean	94.4	3.32	3.19	2.04	0.372 (0.038)	0.429 (0.038)
Species level	95.0	3.68	3.55	2.04	—	0.414
<i>Calopogon barbatus</i>	66.7	3.07	2.38	1.64	0.237 (0.014)	0.227 (0.058)
<i>Calopogon multiflorus</i>	50.0	2.67	1.83	1.25	0.076 (0.012)	0.107 (0.056)
<i>Bletia purpurea</i>	30.8	2.00	1.31	1.02	0.017 (0.006)	0.020 (0.009)
<i>Arethusa bulbosa</i>	41.2	2.14	1.47	1.24	0.144 (0.021)	0.130 (0.050)

sponding to equal and unequal dosages of different alleles) were evident suggesting autotetraploidy. At two of these loci as many as four alleles were detected within individuals while up to three alleles per individual were observed at another two loci. Diploid gene expression was observed at the remaining 13 polymorphic loci.

In the five *Calopogon* species, fixation indices (F) were significantly different from Hardy-Weinberg expected values in 25% (43 of 170 tests) of the chi-square tests. Based on chance alone, one would expect 5% (9 of 170) to be significantly different from the expected Hardy-Weinberg value. The percentage of significant

deviations ranged from 21% (*C. barbatus*) to 33% (*C. multiflorus*). Mean F values across all polymorphic loci were  $-0.024$  in *C. barbatus*,  $-0.064$  in *C. oklahomensis*,  $0.090$  in *C. tuberosus*,  $0.117$  in *C. pallidus*, and  $0.180$  in *C. multiflorus* suggesting that there is a slight excess of homozygotes in most populations. In *B. purpurea* 25% (one of four) of the F values were significantly different from zero, while none differed significantly in *A. bulbosa*.

**Species Level Variation.** Estimates of genetic diversity within species were only calculated for the three species (*C. tuberosus*, *C. pallidus*, and *C. oklahomensis*) with more than one population available. All three species displayed high levels of genetic diversity (Table 2). In *C. oklahomensis* 95% of the loci were polymorphic, each averaging 3.68 alleles (AP) and genetic diversity ( $H_e$ ) calculated over all loci was 0.414. In *C. tuberosus* and *C. pallidus* the percentages of polymorphic loci were also high (91% and 81% respectively). In *C. tuberosus*, AP = 3.42 and  $H_e$  = 0.215. In *C. pallidus*, AP = 3.59 and  $H_e$  = 0.252.

Summed across all 21 monomorphic and polymorphic loci, *C. tuberosus* and *C. pallidus* each had 67 alleles while *C. barbatus* had 51 alleles. In the other four species fewer loci were resolved. Seventy-two alleles were detected in *C. oklahomensis* (20 loci), 33 alleles in *C.*

TABLE 2. Nei's (1973) mean genetic diversity statistics for *C. tuberosus*, *C. pallidus* and *C. oklahomensis*: mean total genetic diversity ( $H_T$ ), mean genetic diversity within populations ( $H_S$ ), the proportion of genetic variation that occurs among populations ( $G_{ST}$ ) and mean fixation index ( $F_{IS}$ ) which indicates the degree to which populations differ from Hardy-Weinberg equilibrium. Values are based on 19 polymorphic loci.  $G_{ST}$  data for individual loci are available from DWT upon request.

Species	$H_T$	$H_S$	$G_{ST}$	$F_{IS}$
<i>C. tuberosus</i>	0.237	0.197	0.085	0.090
<i>C. pallidus</i>	0.311	0.283	0.078	0.117
<i>C. oklahomensis</i>	0.436	0.415	0.037	$-0.064$

*multiflorus* (18 loci), 25 alleles in *A. bulbosa* (17 loci), and 17 alleles in *B. purpurea* (13 loci). The number of alleles unique to each species was 14 in *C. oklahomensis* (two populations), five in *C. tuberosus* (five populations), four in *C. pallidus* (four populations), three in *C. barbatus* (one population), and one in *C. multiflorus* (one population). At the TPI1 locus 11 alleles were detected in the five *Calopogon* species. Due to limitations of the software used, the three rarest alleles were pooled to make one synthetic allele. All nine were scored for *C. pallidus*, while eight occurred in *C. tuberosus* and *C. barbatus*, seven in *C. oklahomensis* and six in *C. multiflorus*.

Mean total genetic diversity at the polymorphic loci of *C. tuberosus* was relatively low ( $H_T = 0.237$  or  $0.253$  if *C. tuberosus* var. *simpsonii* was excluded; Table 2), indicating that allele frequencies at many loci were skewed. Most of this variation occurs within populations ( $H_S = 0.197$  or  $0.233$  if *C. tuberosus* var. *simpsonii* was excluded). Consequently the proportion of genetic variation among populations was low ( $G_{ST} = 0.085$  or  $0.045$  if *C. tuberosus* var. *simpsonii* was excluded). However, tests for heterogeneity in allele frequencies among populations of *C. tuberosus* indicate that 84% (16 of 19 polymorphic loci) are significant ( $P > 0.05$ ).

*Calopogon pallidus* had a higher mean total genetic diversity at the polymorphic loci ( $H_T = 0.311$ ), most of which was attributable to within-population variation. The proportion of genetic variation among populations ( $G_{ST} = 0.078$ ) was somewhat lower than in *C. tuberosus*. Tests for allele frequency differences among populations revealed that 77% (13 of 17 polymorphic loci) were significant ( $P > 0.05$ ).

*Calopogon oklahomensis* had the highest mean total genetic diversity at its polymorphic loci ( $H_T = 0.436$ ; Table 2). Most of this variation was found within populations ( $G_{ST} = 0.037$ ). Chi-square tests for allele frequency differences among populations show that 47% (9 of the 19 polymorphic loci) are significant ( $P > 0.05$ ). Based on chance alone, only one significantly different locus was expected.

**Genetic Relationships Among Taxa.** Mean genetic identities among the five *Calopogon* species were calculated from pairwise comparisons between populations and ranged from 0.634 to 0.888 (Table 3), with an overall mean value of 0.764. *Calopogon pallidus* and *C. multiflorus* had the highest genetic identity (0.888), followed closely by *C. barbatus* and *C. multiflorus* (0.876), and *C. barbatus* and *C. pallidus* (0.875). *Calopogon oklahomensis* had the lowest mean genetic identities with its congeners (0.634–0.696).

Mean genetic identity between *B. purpurea* and the five *Calopogon* species was 0.684, ranging from 0.509 (with *C. oklahomensis*) to 0.819 (with *C. multiflorus*; Table 3). Mean genetic identity between *A. bulbosa* and the *Calopogon* species was 0.557, ranging from 0.481

TABLE 3. Mean genetic identities for pairwise comparisons of species. Mean intraspecific identities are given on the diagonal. *Cb* = *C. barbatus*, *Cm* = *C. multiflorus*, *Co* = *C. oklahomensis*, *Cp* = *C. pallidus*, *Ct* = *C. tuberosus*, *Bp* = *Bletia purpurea*, *Ab* = *Arethusa bulbosa*.

	<i>Cm</i>	<i>Co</i>	<i>Cp</i>	<i>Ct</i>	<i>Bp</i>	<i>Ab</i>
<i>Cb</i>	0.876	0.684	0.875	0.783	0.716	0.585
<i>Cm</i>	—	0.696	0.888	0.768	0.819	0.481
<i>Co</i>		0.900	0.662	0.634	0.509	0.584
<i>Cp</i>			0.945	0.775	0.728	0.528
<i>Ct</i>				0.932	0.649	0.605
<i>Bp</i>					—	0.385

(with *C. multiflorus*) to 0.605 (with *C. tuberosus*). The lowest genetic identity value (0.385) was found between *A. bulbosa* and *B. purpurea*.

Mean genetic identity among populations of *C. tuberosus* was 0.932, ranging from 0.853 (Ctsimp-FL and Ct-TN) to 0.969 (Ct-FL1 and Ct-FL2). When *C. tuberosus* var. *simpsonii* (Ctsimp-FL) was removed from the calculation, average genetic identity rose to 0.950 and ranged from 0.904 (between Ct-FL1 and Ct-TN) to 0.969. The Mantel test showed that there was no significant correlation between decreased genetic identity with increasing geographic distance ( $r = -0.66$ ;  $P = 0.955$ ). *Calopogon tuberosus* var. *simpsonii* had a mean genetic identity of 0.904 with populations of *C. tuberosus* var. *tuberosus*, ranging from 0.853 (Ctsimp-FL and Ct-TN) to 0.953 (Ctsimp-FL and Ct-FL2). The genetic distance of *C. tuberosus* var. *simpsonii* from *C. tuberosus* var. *tuberosus* populations therefore warrants its exclusion from some calculations that follow. Genetic identities among the *C. pallidus* populations ranged from 0.898 (Cp-FL2 and Cp-SC) to 0.980 (Cp-FL1 and Cp-SC) with a mean of 0.945 (Table 3). Genetic identity decreased with increasing geographic distance but was not significant ( $r = -0.22$ ;  $P = 0.683$ ). Sampling of only one population of *C. barbatus* and *C. multiflorus* precluded calculation of intraspecific identity values for these two species.

The UPGMA phenogram generated using the genetic identity data groups the *C. tuberosus* and *C. pallidus* populations into two distinct clusters (Fig. 1). The phenogram indicates that *C. pallidus* is more closely related to *C. multiflorus* than to *C. barbatus* and that these three species differed less from one another than they did from *C. tuberosus*. The data also suggest that *A. bulbosa* is the more distantly related of the two outgroups. A rather perplexing result is that *C. oklahomensis* is shown as the basal taxon to its four congeners as well as *B. purpurea*.

## DISCUSSION

**Variation Within Populations.** All five species of *Calopogon* maintain high levels of allozyme variation

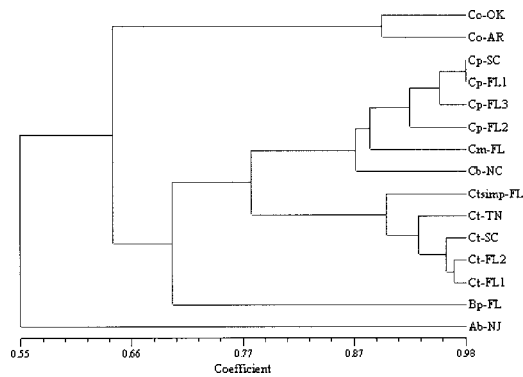


FIG. 1. UPGMA phenogram displaying levels of genetic identity between populations. The coefficient of similarity is indicated at the bottom of the phenogram. Cb = *C. barbatus*; Cm = *C. multiflorus*; Co = *C. oklahomensis*; Cp = *C. pallidus*; Ct = *C. tuberosus*; Ctsimp = *C. tuberosus* var. *simpsonii*; Ab = *Arethusa bulbosa*; Bp = *Bletia purpurea*. Population codes are indicated next to the species codes and are listed in Appendix 1.

within their populations. It has generally been observed that more widespread species have more genetic diversity within their populations than among them. This is only partially true for the five *Calopogon* species. Although *C. oklahomensis* and *C. multiflorus* have the most restricted geographic distributions, *C. oklahomensis* has the highest genetic diversity measures of the five species, perhaps because it is an autotetraploid. *Calopogon tuberosus*, the most widespread of the five, has the second highest mean percentage of polymorphic loci (77% without *C. tuberosus* var. *simpsonii*). However, the mean AP value in *C. tuberosus* is less than that of the other three more geographically restricted species. Even so, *C. tuberosus* has mean population measures of genetic variation that exceed the mean values reported by Hamrick and Godt (1989) for species with regional distributions ( $P = 36.4$ ,  $A = 1.55$ ,  $A_e = 1.16$ , and  $H_e = 0.118$ ). *Calopogon multiflorus* has the lowest mean values for all measures of genetic variation. However, compared with mean values of other plant species that share a suite of eight taxonomic and life history traits, *C. multiflorus* has higher within-population values for three of the four genetic parameters used (Hamrick and Godt 1989). Only its genetic diversity ( $H_e = 0.107$ ) was lower than the mean found for species with similar traits. Such a discrepancy between these genetic diversity parameters is probably due to the large number of low frequency alleles in the single *C. multiflorus* population examined. The lower genetic diversity of *C. multiflorus* may be explained by the possibility that it is a recently derived species from *C. barbatus* with a limited portion of its progenitor's genome. This is supported by the finding that *C. multiflorus* has only one unique allele while its congeners have three (*C. barbatus*), four (*C. pallidus*), five (*C. tuberosus*), and 14 (*C. oklahomensis*) unique alleles. In ad-

dition, *C. multiflorus* may have declined in abundance within the last few decades as a result of increased development, habitat degradation, and fire suppression. Reductions in the size and number of populations, as well as the increased isolation of the remaining populations, may have resulted in the loss of genetic diversity within individual populations. This species is also the most fire dependent, which has rendered it the most susceptible to decline by the recent history of fire suppression in the Southeast.

**Species Level Variation.** Variation at this level can only be compared for *C. tuberosus*, *C. pallidus*, and *C. oklahomensis*, species for which multiple populations were examined. For all three species the proportion of polymorphic loci is much greater than mean values for plants sharing the same eight life history characteristics (Hamrick and Godt 1989). On average, plant species are polymorphic at only 50% of their allozyme loci (Hamrick and Godt 1989; Hamrick et al. 1991) but the highest mean value ( $P = 59.2\%$ ) was reported for monocot species (Hamrick and Godt 1989). Sixteen orchid species surveyed by Hamrick and Godt (1996) had a mean  $P = 45\%$ . However, *C. oklahomensis*, *C. tuberosus* and *C. pallidus* have  $P$  values of 95.0%, 90.5% and 81.0%, respectively. These species had a mean number of alleles per locus of 3.55, 3.19 and 3.10, respectively, compared with a mean value of 2.53 for temperate-tropical plants. *Calopogon tuberosus* had a genetic diversity value ( $H_e$ ) of 0.22 compared to 0.15 in geographically regional species while *C. oklahomensis* and *C. pallidus* had genetic diversity values of 0.41 and 0.25, respectively, compared with a mean of 0.14 found for species with narrow ranges (Hamrick and Godt 1989). Mean genetic diversity for all plant species studied is 0.15 (Hamrick and Godt 1989; Hamrick et al. 1991) and for the 16 species of orchids is 0.14 (Hamrick and Godt 1996).

Mean total genetic diversity at polymorphic loci for *C. tuberosus* ( $H_T = 0.237$  or 0.253 when *C. tuberosus* var. *simpsonii* is excluded) was lower than all mean values found for species that share a suite of eight taxonomic and life history traits (Hamrick and Godt 1989; Table 2). For *C. pallidus*, however,  $H_T$  (0.311) exceeded the mean of all species with comparable taxonomic and life history characteristics except monocots (0.320) and species with sexual reproduction (0.311; Table 2). The higher mean  $H_T$  value for *C. pallidus* is due to the presence of more alleles per polymorphic locus and alleles that occur at more even frequencies than those of *C. tuberosus*. In *C. oklahomensis*,  $H_T$  (0.436) exceeded the mean of all species examined by Hamrick and Godt (1989; Table 2). The higher levels of genetic diversity observed in *C. oklahomensis* ( $H_T = 0.436$ ) compared to *C. tuberosus* ( $H_T = 0.237$ ) or *C. pallidus* ( $H_T = 0.311$ ) are probably a reflection of the autotetraploidy exhibited

at six of the 19 polymorphic loci examined ( $H_T$  for the six autotetraploid loci = 0.512).

The best explanation for diploid expression in 13 of the 19 polymorphic loci and autotetraploid expression in the remaining six loci is the duplication of genes through autopolyploidy followed by the fixation of null alleles at one of the duplicated loci (i.e. gene silencing; Li 1980). Ultimately this can lead to the complete loss of duplicate gene expression at the silenced loci (Li 1980). Eventually a polyploid can become genetically diploidized such that gene expression will be identical to that of the diploid ancestor (Ohno 1970; Grant 1981; Leipoldt and Schmidtke 1982; Haufler 1987; Soltis and Soltis 1993). For example, in tetraploid fish (*Salmo* spp. and *Cyprinus carpio*), in every enzyme system examined, gene loss was observed in at least one species, confirming that normal function is maintained by single locus expression. Some species have lost 70% or more of their duplicate genes through gene silencing (Li 1980). It is interesting that in natural populations of tetraploid fish, gene silencing occurs to a greater extent in isozyme loci with high levels of polymorphism (Allendorf 1978). Ferns and fern allies, which generally have exceptionally high chromosome numbers but diploid gene expression, are also believed to represent ancient polyploids that have undergone genetic diploidization (Haufler 1987; Haufler 1989; Soltis and Soltis 1989; Soltis et al 1992).

**Variation Among Populations.** The  $G_{ST}$  values found for *C. oklahomensis* (0.037), *C. tuberosus* (0.085 or 0.045 without *C. tuberosus* var. *simpsonii*) and *C. pallidus* (0.078) are lower than that found in plants with comparable taxonomic and life history traits (Hamrick and Godt 1989). However,  $G_{ST}$  values for these three *Calopogon* species are only slightly lower than that found for 16 species of terrestrial orchids (mean  $G_{ST}$  = 0.087) surveyed by Hamrick and Godt (1996). Animal pollination and seed dispersal by wind may both contribute significantly to the low levels of among population variation seen in *Calopogon*. Seed dispersal probably plays a major role since the dust-like orchid seeds may potentially spread long distances. The dispersal of diploid seeds has twice the effective gene flow rate of a haploid pollen grain. The  $G_{ST}$  values found for *Calopogon* species are consistent with moderate to high rates of gene flow and suggest that genetic drift within populations has not played a major role in the genetic structuring of these species.

**Variation Among Taxa.** The UPGMA phenogram clusters intraspecific populations together. This clustering of species is consistent with reproductive isolation among the species. Since *C. tuberosus*, for example, requires pollinators of sufficient size and weight to effect pollination (Thien and Marcks 1972; Firmage and Cole 1988) it is reproductively isolated from its congeners by its floral morphology. Three

phenetic groups comprise the phenogram: 1) *C. pallidus*, *C. multiflorus* and *C. barbatus*; 2) *C. tuberosus*; 3) *C. oklahomensis*. *Calopogon oklahomensis* is placed basally to its congeners suggesting that it may be similar to an ancestral species. The degree of gene silencing observed in this species would also suggest that it is a rather old polyploid. The most logical explanation for these results is that *C. pallidus*, *C. multiflorus*, *C. barbatus* and *C. tuberosus* may have arisen from the diploid progenitor of *C. oklahomensis*.

Gottlieb (1981) found that congeneric species have mean genetic identities of  $0.67 \pm 0.04$ , while Crawford (1983) reported values of 0.75. Populations of the five *Calopogon* species have a mean genetic identity of 0.764 between taxa. Mean genetic identities among the three genera examined were closer to Gottlieb's average for congeneric species with values of 0.684 (between *B. purpurea* and the five *Calopogon* species) and 0.557 (between *A. bulbosa* and the five *Calopogon* species). The data suggest that *A. bulbosa* is the more distant outgroup. However, Goldman et al. (2001) provide convincing evidence that, of these two outgroups, *Arethusa* is closer to *Calopogon* phylogenetically. A perplexing finding is the exceptionally high genetic identity between *C. multiflorus* and *B. purpurea* (0.819). Of the 13 loci shared by these species, seven were monomorphic for the same allele. Of the remaining six loci, two were monomorphic in one species for the common allele at polymorphic loci of the other species. According to Gottlieb (1981) the genetic identity statistic is not sensitive to differences among populations resulting from low frequency alleles. Although the *C. multiflorus* population possessed seven alleles that were not detected in *B. purpurea* while the latter had four alleles absent in *C. multiflorus*, most were low frequency alleles. In systematic studies, electrophoretic data are best applied to comparisons of conspecific populations and congeneric species (Crawford 1983; Giannasi and Crawford 1986). Furthermore, to treat electromorphs that have identical migration patterns but are from widely divergent taxa as if they represented the same degree of genetic similarity as identical electromorphs from conspecific populations could distort interpretations (Crawford 1983).

Conspecific populations tend to be similar genetically and have much higher genetic identities in *Calopogon* ( $0.956 \pm 0.011$ ) than the primarily outcrossing congeneric species surveyed by Gottlieb (1981). Such findings indicate that one or a few populations may constitute an adequate sample for ascertaining the genetic diversity of the species (Gottlieb 1977; 1981; Hamrick and Godt 1989), at least for alleles that occur in moderate to high frequencies ( $\geq 0.20$ ; Gottlieb 1981).

Lower genetic identities between *C. oklahomensis* and *C. tuberosus* and their congeners suggest that *C. barbatus*, *C. multiflorus*, and *C. pallidus* are more closely

related to one another than to *C. tuberosus* or *C. oklahomensis*. This suggests that these three species diverged from one another more recently than they did from *C. tuberosus*. Gottlieb (1977) found that when species pairs possessed very high genetic identities such as seen in *Calopogon*, they were usually closely related progenitor and derivative species of recent origin with little genetic divergence (Gottlieb 1981; Giannasi and Crawford 1986).

The genetic divergence of *C. tuberosus* var. *simpsonii* probably results from its ecological separation from *C. tuberosus* var. *tuberosus*. *Calopogon tuberosus* var. *simpsonii* only occurs in approximately neutral marl soil in the Everglades of southern Florida (Correll 1950), compared with the acidic soils in which the parental variety typically grows. The lower genetic identity suggests that this taxon has been reproductively isolated for a substantial period of time.

In summary, these data verify high genetic similarity within *Calopogon*, which corresponds well with the high degree of morphological resemblance between the species. It is possible that introgression has contributed to the observed genetic and morphological similarities among the taxa. However, there is sufficient genetic distance to warrant recognition as distinct taxa. The five species of *Calopogon* all have high measures of genetic diversity. In the species where multiple populations were sampled, low  $G_{ST}$  values were found indicating that most of the species' genetic variation is housed within rather than among populations. These results further suggest that there are moderate to high rates of gene flow and that genetic drift within these populations has not been significant in the genetic structuring of these species. Although *C. tuberosus* has the broadest geographic distribution, it is not the most genetically diverse taxon. Inclusion of the Tennessee population lowers the mean genetic diversity values for this species. This population had four fewer polymorphic loci and fewer alleles per polymorphic locus. These four loci are monomorphic for the allele that is most common in the other populations. These characteristics suggest that this population may have been founded by relatively few individuals. Lastly, *C. oklahomensis* is the most genetically distinct species. The genetic distance of this apparently ancient polyploid may be the result of its occupation of a separate glacial refugium west of the Mississippi river.

**ACKNOWLEDGMENTS.** The authors wish to thank Drs. Jim Afolter and Mike Moore for their valuable advice and field assistance and Dr. John Nason for his generous help in the lab. We thank Doug Goldman, Dennis Horn, R.G. Reimus, and Bob Yannetti for kindly sending us plant specimens. We are also most grateful to Dr. John Beckner, Dave Darrow, Kris Delaney, Don Ford, Bob Hale, Scott Hedges, Julie Herrlinger, David Jones, Cathy King, Richard J. LeBlond, Dr. John Nelson, Bob Progulskie Jr., Dena Thompson, Walter M. Thomson, Bruce Sorrie and Chuck Wilson for their assistance and companionship in the field. Special thanks

go to Angus Gholson, Jr. for his extensive field advice, assistance and patience. Financial support was provided by a Palfrey Grant from the Department of Botany at the University of Georgia.

#### LITERATURE CITED

- ALLENDORF, F. W. 1978. Protein polymorphism and the rate of loss of duplicate gene expression. *Nature* 272: 76–78.
- CHELIAK, W. M. and J. A. PTEL. 1984. Techniques for starch gel electrophoresis of enzymes from forest tree species. Pp. 49 in *Petawawa National Forestry Institute, Information Report P1-X-42*. Chalk River, Ontario: Canadian Forestry Service, Agriculture.
- CORRELL, D. S. 1950. *Native orchids of North America north of Mexico*. Waltham, Massachusetts: Chronica Botanica Company.
- CRAWFORD, D. J. 1983. Phylogenetic and systematic inferences from electrophoretic studies. Pp. 257–287 in *Isozymes in plant genetics and breeding*, part A, eds. S. D. Tanksley and T. J. Orton. Amsterdam: Elsevier Science Publishers, Inc..
- DAFNI, A. 1984. Mimicry and deception in pollination. *Annual Review of Ecology and Systematics* 15: 259–278.
- DRESSLER, R. L. 1981. *The orchids: natural history and classification*. Cambridge, Massachusetts: Harvard University Press.
- . 1993. *Phylogeny and classification of the orchid family*. Portland: Dioscorides Press.
- FIRMAGE, D. H. and F. R. COLE. 1988. Reproductive success and inflorescence size of *Calopogon tuberosus* (Orchidaceae). *American Journal of Botany* 75: 1371–1377.
- GIANNASI, D. E. and D. J. CRAWFORD. 1986. Biochemical systematics II. A reprise. Pp. 25–248 in *Evolutionary biology*, Vol. 20, eds. M. K. Hecht, B. Wallace, and G. T. Prance. New York: Plenum Publishing Corp.
- GOLDMAN, D. H. 1995. A new species of *Calopogon* from the mid-western United States. *Lindleyana* 10: 37–42.
- , J. V. FREUDENSTEIN, P. J. KORES, M. MOLVRAY, D. C. JARRELL, W. M. WHITTEN, K. M. CAMERON, R. K. JANSEN and M. W. CHASE. 2001. Phylogenetics of Arethuseae (Orchidaceae) based on plastid *matK* and *rbcl* sequences. *Systematic Botany* 26: 670–695.
- GOTTLIEB, L. D. 1977. Electrophoretic evidence and plant systematics. *Annals of the Missouri Botanical Garden* 64: 161–180.
- . 1981. Electrophoretic evidence and plant populations. *Progress in Phytochemistry* 7: 1–46.
- GRANT, V. 1981. *Plant speciation*, 2nd edition. New York: Columbia University Press.
- HAMRICK, J. L. and M. J. W. GODT. 1989. Allozyme diversity in plant species. Pp. 46–63 in *Plant population genetics, breeding and genetic resources*, eds. A. H. D. Brown, M. T. Clegg, A. L. Kahler, and B. S. Weir. Sunderland, Massachusetts: Sinauer Associates, Inc.
- and ———. 1996. Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 351: 1291–1298.
- , ———, D. A. MURAWSKI, and M. D. LOVELESS. 1991. Correlation between species traits and allozyme diversity: implications for conservation biology. Pp. 75–86 in *Genetics and conservation in rare plants*, eds. D. A. Falk and K. E. Holsinger. New York: Oxford University Press.
- HAUFLER, C. H. 1987. Electrophoresis is modifying our concepts of evolution in homosporous pteridophytes. *American Journal of Botany* 74: 953–966.
- . 1989. Towards a synthesis of evolutionary modes and mechanisms in homosporous pteridophytes. *Biochemical Systematics and Ecology* 17: 109–115.
- LEIPOLDT, M. and J. SCHMIDTKE. 1982. Gene expression in phylogenetically polyploid organisms. Pp. 219–236 in *Genome evolution*, eds. G. A. Dover and R. B. Flavell. London: Academic Press.
- LI, C. C. and D. G. HORVITZ. 1953. Some methods of estimating

- the inbreeding coefficient. *American Journal of Human Genetics* 5: 107–117.
- LI, W. H. 1980. Rate of gene silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes. *Genetics* 95: 237–258.
- MITTON, J. B., Y. B. LINHART, K. B. STURGEON and J. L. HAMRICK. 1979. Allozyme polymorphisms detected in mature needle tissue of ponderosa pine. *Heredity* 70: 86–89.
- NEI, M. 1972. Genetic distance between populations. *American Naturalist* 106: 283–292.
- . 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Science, USA*. 70: 3321–3323.
- OHNO, S. 1970. *Evolution by gene duplication*. New York: Springer.
- PACE, L. 1909. The gametophytes of *Calopogon*. *Botanical Gazette* 48: 126–137.
- ROHLF, F. J. 2000. Numerical taxonomy and multivariate analysis system version 2.1. Setauket, New York: Exeter Publishing.
- SMOUSE, P. E., J. C. LONG and R. R. SOKAL. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology* 35: 627–632.
- SOLTIS, D. E. and P. S. SOLTIS. 1989. Polyploidy, breeding systems and genetic differentiation in homosporous pteridophytes. Pp. 241–258 in *Isozymes in plant biology*, eds. D. E. Soltis and P. S. Soltis. Portland: Dioscorides Press.
- and ———. 1993. Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences* 12: 243–273.
- , C. H. HAUFLE, D. C. DARROW and G. J. GASTONY. 1983. Starch gel electrophoresis of ferns: a compilation of grinding buffers, gel and electrode buffers, and staining schedules. *American Fern Journal* 73: 9–27.
- SOLTIS, P. S., J. J. DOYLE and D. E. SOLTIS. 1992. Molecular data and polyploid evolution in plants. Pp. 177–201 in *Molecular systematics of plants*, eds. P. S. Soltis, D. E. Soltis and J. J. Doyle. New York: Chapman and Hall.
- THIEN, L. B. 1973. Isolating mechanisms in the genus *Calopogon*. *American Orchid Society Bulletin* 42: 794–797.
- and B. G. MARCKS. 1972. The floral biology of *Arethusa bulbosa*, *Calopogon tuberosus*, and *Pogonia ophioglossoides* (Orchidaceae). *Canadian Journal of Botany* 50: 2319–2325.
- WEEDEN N. F. and J. F. WENDELL. 1989. Genetics of plant isozymes. Pp. 46–72 in *Isozymes in Plant Biology*, eds. D. E. Soltis and P. S. Soltis. Portland: Dioscorides Press.
- WORKMAN, P. L. and J. D. NISWANDER. 1970. Population studies on southwestern Indian tribes. II. Local genetic differentiation in the Papago. *American Journal of Human Genetics* 22: 24–49.
- WRIGHT, S. 1922. Coefficients of inbreeding and relationship. *American Naturalist* 56: 330–338.

## APPENDIX 1

List of populations sampled for allozyme analysis, with the number of individuals collected and the voucher specimen for each population. The herbarium in which each specimen is deposited is indicated by the acronym in parentheses.

- Calopogon barbatus* Cb-NC (48 individuals) = R.J. LeBlond 3797, Carteret Co., NC (GA). *Calopogon multiflorus* Cm-FL (34 individuals) = D.W. Trapnell 371, Osceola Co., FL (GA). *Calopogon oklahomensis* Co-AR (48 individuals) = D.H. Goldman 553, Prairie Co., AR (TEX); Co-OK (48 individuals) = D.H. Goldman 514, Muskogee Co., OK (TEX). *Calopogon pallidus* Cp-FL1 (48 individuals) = D.W. Trapnell 158, Liberty Co., FL (GA); Cp-FL2 (48 individuals) = D.W. Trapnell 374A, Highlands Co., FL (GA); Cp-FL3 (48 individuals) = D.W. Trapnell 375A, Liberty Co., FL (GA); Cp-SC (19 individuals) = D.W. Trapnell 378, Williamsburg Co., SC (GA). *Calopogon tuberosus* var. *tuberosus* Ct-FL1 (48 individuals) = D.W. Trapnell 157, Levy Co., FL (GA); Ct-FL2 (48 individuals) = D.W. Trapnell 376A, Franklin Co., FL (GA); Ct-SC (48 individuals) = D.W. Trapnell 377, Richland Co., SC (GA); Ct-TN (48 individuals) = D.D. Horn 687, Coffee Co., TN (GA). *Calopogon tuberosus* var. *simpsonii* Ctsimp-FL (48 individuals) = R.G. Reimus 711, Dade Co., FL (GA).
- Bletia purpurea* Bp-FL (41 individuals) = D.W. Trapnell 373A, Collier Co., FL (GA).
- Arethusa bulbosa* Ab-NJ (19 individuals) = B. Yannetti s.n. September 1994, Burlington Co., NJ (GA).