

**GENETIC DIVERSITY AND GENE FLOW IN THE  
MORPHOLOGICALLY VARIABLE, RARE ENDEMICS  
*BEGONIA DREGEI* AND *BEGONIA HOMONYMA*  
(BEGONIACEAE)<sup>1</sup>**

LUZUKO ORLYN MATOLWENI,<sup>2,4</sup> KEVIN BALKWILL,<sup>3</sup> AND  
TRACY MCLELLAN<sup>2</sup>

<sup>2</sup>Department of Molecular and Cell Biology, University of the Witwatersrand, Johannesburg, Private Bag 3, Wits 2050, South Africa; and

<sup>3</sup>Department of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg, Private Bag 3, Wits 2050, South Africa

*Begonia dregei* and *B. homonyma* (Begoniaceae), rare plants endemic to coastal forests of eastern South Africa, are two closely related species with high levels of variation among populations in the shape of leaves. Distribution of genetic variation and genetic relatedness were investigated in 12 populations of *B. dregei* and seven of *B. homonyma* using polyacrylamide gel electrophoresis. Twelve of the 15 enzyme loci examined were polymorphic, but only seven loci were polymorphic within at least one population. Genetic diversity measures indicated that the among-population gene differentiation represents >90% of the total genetic component in both species considered individually or combined. This indicated restricted gene flow, consistent with the limited dispersal abilities of *Begonia* generally and the ancient separation of isolated forest patches. Genetic distances among populations are much higher than usually found within species. Allozyme data provide no support for the recognition of *B. dregei* and *B. homonyma* as distinct species.

**Key words:** African coastal forests; *Begonia*; Begoniaceae; genetic drift; genetic population structure; habitat fragmentation.

Excellent models for the study of evolutionary processes are often provided by taxa that pose the greatest problems in systematics (Wolf, Soltis, and Soltis, 1991). Endemic plants provide a superb tool for studying the dynamic processes of speciation and evolution, particularly island endemic plants (Ito and Ono, 1990; Aradya, Mueller-Dombois, and Ranker, 1991; Barrett, 1996). Evidence of most evolutionary events that formed continental biota has been lost because such biota are so ancient (Carson, 1987). Complex patterns of variation may blur species boundaries and lead to taxonomic complexity.

Fragmentation of populations may result from recent disruption of a habitat by human activities and historical biogeographic factors, or may be due to the naturally patchy distribution of appropriate environmental conditions. Several recent studies have investigated the relationship between genetic variation and the degree of isolation in plant species with fragmented distributions (e.g., Ouborg, van Treuren, and van Damme, 1991; van Treuren et al., 1991; Godt and Hamrick, 1993). In species with disjunct distributions, the effects of various evolutionary processes such as genetic drift, mating systems, and natural selection are expected to be reflected in the

genetic population structure (Lönn and Prentice, 1995). Based on the assumption that genetic markers such as allozymes are selectively neutral (Hamrick, Linhart, and Mitton, 1979), genetic drift should influence all loci equally, whereas natural selection might affect loci differentially. Many studies of population structure have focussed on spatial distribution of allelic variation for allozymes (Podolsky and Holtsford, 1995). Hamrick and Godt (1990) determined associations between variation of allozymes and life history attributes, such as mating system and life span. Nonetheless, the degree to which population structure revealed by allozymes is representative of the entire genome and reflects that of quantitative traits is still unresolved (Podolsky and Holtsford, 1995).

The present study investigates allozyme variation in *Begonia dregei* Otto and Dietr. and of *B. homonyma* Steudl. These two species form a highly variable and complex group and are distinguished from each other solely on the size and shape of leaves according to the most recent taxonomic revision (Hilliard, 1976). They share similar floral morphology and a caudex, or swollen stem base, which is not known from any other species of *Begonia* (Smith et al., 1986). There is considerable variation among populations in the size, margin, and lobing of leaves, particularly in the smaller leaved *B. dregei* (Hilliard, 1976). *Begonia dregei* and *B. homonyma* are herbaceous perennials endemic to subtropical forests along the east coast of South Africa, where they occur in small, discrete populations on deeply shaded, south-facing slopes. They are monoecious, self compatible, and pollinated by insects. Their tiny seeds are dispersed by gravity a short distance from the parent plant. The dis-

<sup>1</sup> Manuscript received 15 October 1998; revision accepted 25 June 1999.

The authors thank David Wopula and Jo Arkell for assistance with the collection of plants; and M. Slatkin, P. Lewis, and J. Goudet for programs. This study was supported by grants from the South African Foundation for Research Development (FRD) and Mellon Foundation.

<sup>4</sup> Author for correspondence: Department of Molecular and Cell Biology, University of the Witwatersrand, Johannesburg, Private Bag 3, Wits 2050, South Africa. Phone: + 27 11 716 2178; FAX: + 27 11 403 1733; e-mail: 108mal@cosmos.wits.ac.za.

tinative hemispherical caudex forms only in plants that have grown from seed and was present on over 1000 plants observed in the wild, suggesting that asexual reproduction does not occur. In cultivation, the plants are easy to cross both within and between the two species, but there is no evidence for hybridization from the distribution of variation in the wild. *Begonia homonyma* is collected from the wild for use in traditional medicine (Cunningham, 1988), but such plant use does not involve either cultivation or moving of plants.

The coastal forests in South Africa consist of patches, ranging in size from less than 1 to 1787 ha (Cooper and Swart, 1992), separated by grassland, much like an archipelago of islands. There are several hypotheses about the processes that have led to this pattern. The present distribution of forests might have resulted from human land use, involving grazing and burning over the past 300 yr, particularly the last century (Granger, 1984; Acocks, 1988). An alternate hypothesis is that the pattern is ancient and due to environmental factors including increasing aridity (Feeley, 1986; Rutherford and Westfall, 1986; Geldenhuys, 1994). The flora of coastal forests of eastern Africa, to the north of the range of *B. dregei* and *B. homonyma*, includes many endemic species with disjunct distributions that may be relicts from the breakup of the Pan African rainforests that started 17–18 million yr ago (Burgess, Clarke, and Rodgers, 1998).

This paper reports the results of an electrophoretic survey to examine genetic variation within and divergence among populations. The objectives were to: (1) describe the distribution of genetic variation, using allozymes, within and among populations of *B. dregei* and *B. homonyma*; (2) investigate genetic relatedness among populations in both species; (3) determine the relationship, if any, between genetic distance and geographic proximity of the populations; and (4) investigate whether patterns of variation in allozymes are consistent with the current taxonomic treatment. The study also explores the extent to which the structure of genetic variation of forest floor taxa can elucidate the history of the habitat.

## MATERIALS AND METHODS

**Study sites and collection of samples**—Plant material was collected throughout the known geographical ranges of *B. dregei* and *B. homonyma* (Hilliard, 1976) which occur in only two provinces along the east coast of South Africa. In KwaZulu/Natal, populations occur further inland than in the Eastern Cape, where they are restricted to northern and central regions of the province (Fig. 1). A total of 19 populations, including 12 of *B. dregei* and seven of *B. homonyma*, covering most of the variation in leaf shape in the group, was sampled (Table 1). We attempted to collect 20 individuals from each population, but several places had fewer than 20 plants, notably at the mouth of the Mtentu River, where an intensive search produced four plants. Populations occurred in limited areas, and population size was estimated by counting all plants observed. The names of forests or rivers with which populations were associated were used to refer to them; the populations are numbered from north to south for each species (Table 1).

**Isozyme electrophoresis**—Vertical polyacrylamide gel electrophoresis was performed on leaf extracts using a Hoefer Protean 16 electrophoresis unit. PAGE was chosen because of its high resolving power of electrophoretic variants (Wendel and Weeden, 1989). Fresh leaves were homogenized in a concentrated alkaline buffer made of 1.0 mol/

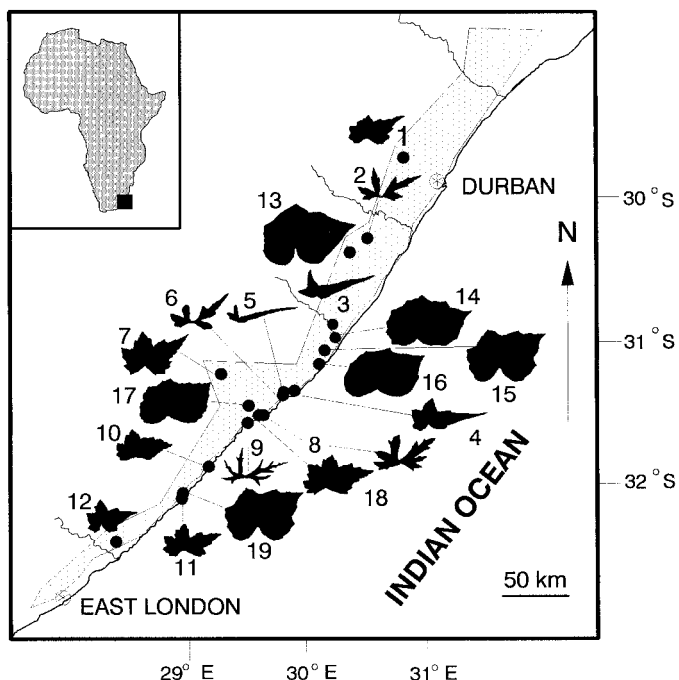


Fig. 1. Map indicating geographic distribution of collection sites and leaf shapes of *B. dregei* and *B. homonyma*. 1 = Hillcrest, 2 = Glenrosa, 3 = Longkloof, 4 = Lupatana, 5 = Ntsubane, 6 = Mbotyi, 7 = Horseshoe, Bulembu, 8 = Umzimvubu river mouth, 9 = Bulolwe river, 10 = Ku-Bomvu river mouth, 11 = Mbashe river mouth, 12 = Gxara falls, 13 = Dumisa, 14 = Bulolo river, Umtamvuna gorge, 15 = Mzamba river mouth, 16 = Mtentu river mouth, 17 = Isinuka, 18 = Mt. Sullivan, 19 = Mbanyana river mouth. Populations of each species are numbered from north to south. The shaded region represents the overall distribution of both species.

L tris-HCl pH 8.9, 20% w/v sucrose, 50 mg/mL PVP 40 (polyvinylpyrrolidone), 2.5 mg/mL DIECA (diethyldithiocarbamic acid), 2.5 mg/mL DTT (dithiothreitol), and 5  $\mu$ L/mL Triton X100. Bromophenol blue was added as tracking dye. Gels were composed of 5% acrylamide (5% was bisacrylamide) and were 140 mm wide, 143 mm long, and 1.5 mm thick. Several buffer systems were tested to determine which buffer gave the best resolution for each enzyme locus. Tris-borate pH 8.7 (0.5 mol/L tris and 0.2 mol/L boric acid) and histidine-MOPS pH 6.6 (0.25 mol/L histidine and 0.3 mol/L MOPS) continuous buffer systems (McLellan, 1982) were selected. Buffers were diluted to a ratio of 1:9 for both electrode and gel buffers. The histidine-MOPS buffer system was used for phosphoglucoisomerase (*Pgi*), phosphoglucomutase (*Pgm*), fructosebiphosphate aldolase (*Fba*), glucose-6-phosphate dehydrogenase (*G6pd*), 6-phosphogluconate dehydrogenase (*6Pgd*), and malate dehydrogenase (*Mdh*). Peroxidase (*Per*), superoxide dismutase (*Sod*), glutamic dehydrogenase (*Gldh*), malic enzyme (*Me*), formic dehydrogenase (*Fdh*), and esterase (*Est*) were best resolved in the tris-borate buffer system. Enzymes were stained using assay protocols modified from Wendel and Weeden (1989). Loci and alleles were numbered sequentially in order of decreasing anodal mobility. Allelism was tested with crosses between plants from eight of the populations; in all cases, inheritance of allozyme bands proved to be Mendelian (Ntloko, 1997).

**Analysis of allozyme variation within populations**—Allele frequencies for each population were calculated from isozyme phenotypes. A locus was considered polymorphic if it had more than one allele. Population variation within and over all populations was quantified by calculating the following metrics from allele frequencies: percentage of loci polymorphic (P), number of alleles per locus (A), and expected ( $H_e$ )

TABLE 1. Estimated population size and numbers of individuals sampled in *Begonia dregei* and *B. homonyma* in this study. Population names are accompanied by numbers that will be used to refer to the populations elsewhere.

Number	Locality	Nearest town	Latitude	Longitude	Estimated population size	No. of individuals sampled
<i>Begonia dregei</i>						
1	Hillcrest	Durban	29°46' S	30°46' E	20	18
2	Glenrosa	Port Shepstone	30°17' S	30°30' E	25	20
3	Long Kloof	Port Edward	30°54' S	30°08' E	50	20
4	Lupatana	Lusikisiki	31°25' S	29°51' E	75	20
5	Ntsubane	Lusikisiki	31°26' S	29°49' E	18	18
6	Mbotyi	Lusikisiki	31°27' S	29°49' E	100	20
7	Horseshoe, Bulembu	Flagstaff	30°06' S	29°51' E	20	15
8	Umzimvubu River Mouth	Port St Johns	31°37' S	29°33' E	50	20
9	Bulolwe River	Port St Johns	31°38' S	29°31' E	50	13
10	Ku-Bomvu River Mouth	Coffee Bay	31°59' S	29°09' E	35	20
11	Mbashe River Mouth	Elliotdale	32°14' S	28°53' E	75	20
12	Gxara Falls	Centane	32°40' S	28°33' E	100	20
<i>Begonia homonyma</i>						
13	Dumisa	Port Shepstone	30°20' S	30°25' E	15	6
14	Bulolo River, Umtamvuna Gorge	Port Edward	31°02' S	30°09' E	25	19
15	Mzamba River Mouth	Port Edward	31°06' S	30°10' E	35	20
16	Mtentu River Mouth	Bizana	31°14' S	30°02' E	4	3
17	Isinuka	Port St Johns	31°36' S	29°29' E	30	20
18	Mt. Sullivan	Port St Johns	31°36' S	29°33' E	50	16
19	Mbanyana River Mouth	Elliotdale	32°13' S	28°55' E	35	20

and observed ( $H_o$ ) frequencies of heterozygotes. Polymorphism and allelic diversity have been shown to be reliable indicators of population structure and can show the effects of bottlenecks and genetic drift (Nei, 1978; Leberg, 1992; Thomas and Bond, 1997).

**Analysis of allozyme variation among populations**—The distribution

of genetic variation within and among populations was determined by calculating Wright's (1951)  $F$  statistics ( $F_{st}$ ,  $F_{is}$  and  $F_{it}$ ) from variance components based on Weir and Cockerham's (1984) estimators  $\theta$ ,  $f$ , and  $F$ , using the program FSTAT (Goudet, 1995), which calculates unbiased estimates, standard deviations, and levels of significance. Values equal to 0 occur in a randomly mating population and those equal to 1 in a

TABLE 2. Levels of intrapopulation allozyme variation in *Begonia dregei* and *B. homonyma*. The values given are means across 15 loci.

Population	Locality	Percentage of loci polymorphic (P)	Mean number of alleles/locus (A)	Mean observed heterozygosity ( $H_o$ )	Mean expected heterozygosity ( $H_e$ )	Fixation index (F)
<i>Begonia dregei</i>						
1	Hillcrest	7	1.07	0.037	0.032	-0.164
2	Glenrosa	20	1.20	0.073	0.079	0.073
3	Long Kloof	13	1.20	0.020	0.065	0.698
4	Lupatana	13	1.13	0.020	0.043	0.542
5	Ntsubane	20	1.20	0.029	0.087	0.665
6	Mbotyi	20	1.20	0.053	0.070	0.246
7	Horseshoe, Bulembu	13	1.13	0.044	0.058	0.235
8	Umzimvubu River Mouth	7	1.07	0.043	0.061	0.296
9	Bulolwe River	7	1.07	0.015	0.027	0.446
10	Ku-Bomvu River Mouth	7	1.07	0.020	0.017	-0.152
11	Mbashe River Mouth	27	1.27	0.090	0.100	0.106
12	Gxara Falls	13	1.13	0.040	0.032	-0.208
Mean		13.9	1.15	0.038	0.054	0.277
Over all populations		80	3.00			
<i>B. homonyma</i>						
13	Dumisa	0	1.00	0.000	0.000	1.000
14	Bulolo River, Umtamvuna Gorge	13	1.13	0.035	0.035	-0.154
15	Mzamba River mouth	0	1.00	0.000	0.000	1.000
16	Mtentu River Mouth	0	1.00	0.000	0.000	1.000
17	Isinuka	20	1.20	0.000	0.067	0.911
18	Mt. Sullivan	0	1.00	0.000	0.000	1.000
19	Mbanyana River Mouth	7	1.07	0.010	0.027	0.395
Mean		5.7	1.06	0.006	0.018	0.562
Over all populations		73	2.00			
Both species		11	1.11	0.027	0.041	0.345
Both species over all populations		80	3.50			

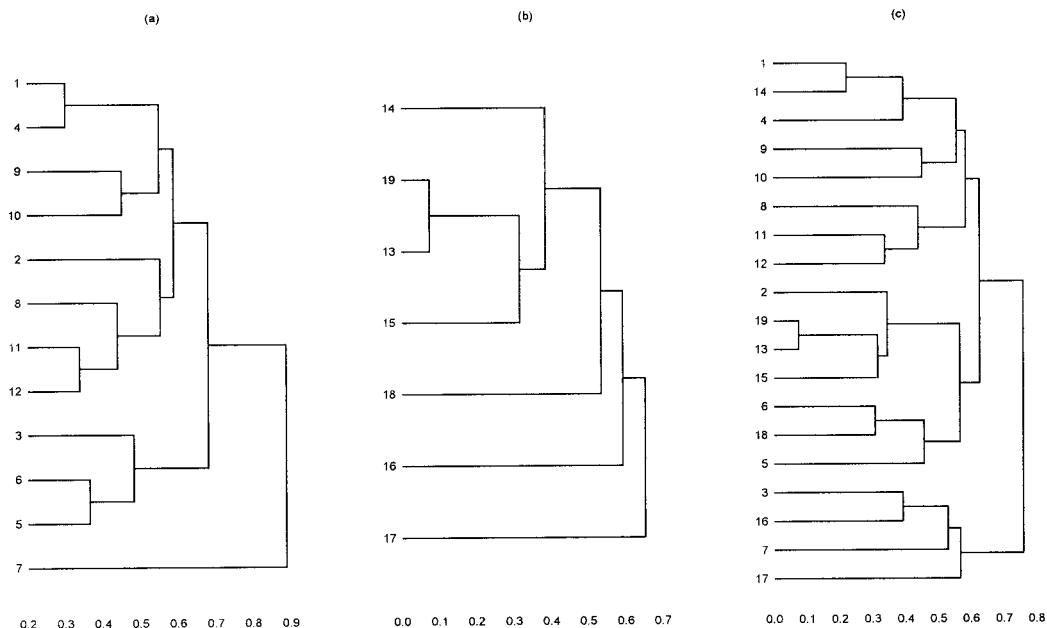


Fig. 2. Dendrograms based on UPGMA of allozyme data for: (a) *B. dregei*, (b) *B. homonyma*, and (c) *B. dregei* and *B. homonyma*. 1 = Hillcrest, 2 = Glenrosa, 3 = Longkloof, 4 = Lupatana, 5 = Ntsubane, 6 = Mbotyi, 7 = Horseshoe, Bulembu, 8 = Umzimvubu river mouth, 9 = Bulolwe river, 10 = Ku-Bomvu river mouth, 11 = Mbashe river mouth, 12 = Gxara falls, 13 = Dumisa, 14 = Bulolo river, Umtamvuna gorge, 15 = Mzamba river mouth, 16 = Mtentu river mouth, 17 = Isinuka, 18 = Mt. Sullivan, 19 = Mbanzana river mouth. Populations are numbered from north to south.

selfing population, as these *F* statistics represent correlations between uniting gametes (Wright, 1951). Variance components of  $F_{it}$  and  $F_{is}$  measure heterozygosity of an individual relative to the total population and its subpopulations, respectively;  $F_{st}$  is a measure of genetic differentiation among populations (Wright, 1951; Weir and Cockerham, 1984; Weir, 1996).

**Cluster analysis and genetic divergence among populations**—The genetic divergence between populations was estimated from their genetic distances (*D*) obtained from Nei's (1978) index of genetic similarity calculated using Genetic Data Analysis (GDA), a computer program written by Lewis and Zaykin (1997). Dendrograms for both species separately and together were constructed using the UPGMA (un-

weighted pair group method using arithmetic means) algorithm using NTSYS-PC (Rohlf, 1993). The relationship between distances among populations and levels of genetic differentiation were determined with Mantel tests to compare matrices and by using Slatkin's (1993) isolation by distance program. Plotting of  $\hat{M}$  (an estimate of gene flow) instead of genetic distance against geographic distances and then performing a regression analysis between the two variables has many advantages, ranging from increased resolution power to nonbiased estimates (Slatkin, 1993).

RESULTS

**Allelic variation**—Twelve enzymes, presumably coded by 15 loci, were scored: *Fba*, *Fdh*, *Gldh*, *G6pdh*, *Mdh*,

TABLE 3. Estimates of genetic distances (*D*) among populations for the two species, *B. dregei* (numbers 1–12) and *B. homonyma* (numbers 13–19). For population names refer to Table 1.

1	0.000																			
2	0.565	0.000																		
3	0.628	0.591	0.000																	
4	0.302	0.661	0.618	0.000																
5	0.884	0.585	0.437	0.912	0.000															
6	0.551	0.517	0.528	0.529	0.365	0.000														
7	0.997	1.000	0.539	0.769	0.954	0.810	0.000													
8	0.449	0.456	0.579	0.656	0.732	0.489	0.797	0.000												
9	0.544	0.658	0.675	0.636	0.584	0.464	0.781	0.575	0.000											
10	0.457	0.591	0.943	0.566	0.523	0.595	1.000	0.513	0.451	0.000										
11	0.671	0.551	0.854	0.502	0.858	0.695	0.926	0.394	0.641	0.424	0.000									
12	0.590	0.653	0.781	0.460	0.914	0.827	0.692	0.486	0.664	0.805	0.339	0.000								
13	0.578	0.347	0.466	0.830	0.472	0.634	1.000	0.444	0.582	0.412	0.780	0.929	0.000							
14	0.223	0.362	0.544	0.489	0.755	0.509	0.782	0.408	0.538	0.602	0.800	0.607	0.401	0.000						
15	0.762	0.400	0.352	0.685	0.678	0.684	0.833	0.639	0.859	0.775	0.644	0.734	0.310	0.487	0.000					
16	0.578	0.791	0.392	0.726	0.844	0.868	0.519	0.600	0.948	0.883	0.831	0.646	0.511	0.484	0.629	0.000				
17	0.779	0.980	0.447	0.914	1.000	0.623	0.593	0.631	0.738	1.000	0.967	0.720	0.648	0.578	0.635	0.662	0.000			
18	0.516	0.364	0.529	0.424	0.549	0.308	1.000	0.444	0.649	0.601	0.579	0.614	0.629	0.417	0.511	0.762	0.790	0.000		
19	0.532	0.288	0.423	0.786	0.489	0.655	1.000	0.473	0.679	0.418	0.809	1.000	0.075	0.269	0.322	0.581	0.602	0.581	0.000	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19		

TABLE 4. *F* statistics for levels of gene flow within and among populations of *B. dregei*. Standard deviations are in parentheses.<sup>a</sup>

Locus	<i>F</i> <sub>is</sub>	<i>F</i> <sub>it</sub>	<i>F</i> <sub>st</sub>
<i>Est-1</i>	0.326 (0.259) n.s.	0.878 (0.053) ***	0.819 (0.064) ***
<i>Est-2</i>	-0.054 (0.954) n.s.	0.985 (0.015) ***	0.986 (0.014) ***
<i>Gldh</i>	0.161 (0.276) n.s.	0.737 (0.167) ***	0.687 (0.197) ***
<i>Per</i>	0.051 (0.269) ***	0.936 (0.041) ***	0.869 (0.061) ***
<i>Pgi</i>	0.321 (0.289) **	0.828 (0.092) ***	0.747 (0.089) ***
<i>6Pgd</i>	0.214 (0.179) n.s.	0.604 (0.111) ***	0.496 (0.125) ***
<i>Sod-1</i>	0.002 (0.880) n.s.	0.897 (0.122) ***	0.897 (0.122) ***
Over all	0.273 (0.052) ***	0.914 (0.033) ***	0.882 (0.043) ***

<sup>a</sup> n.s. = not statistically significantly different from 0; \*\* *P* < 0.005; \*\*\* *P* < 0.001.

*Me*, *Per*, *Pgi*, and *6Pgd* each exhibited one locus while *Est*, *Pgm* and *Sod* exhibited two loci. Twelve of these loci were polymorphic over all populations of both species, and seven had more than one allele within at least one population (Appendix). In all the enzyme loci, consistent banding patterns with the same subunit structure for each enzyme were observed throughout all the populations. The number of isozymes detected for all of the enzymes was the same as those commonly found in diploid plants (Gottlieb and Weeden, 1979), and there was no evidence for duplication of genes in any enzyme system (Ntloko, 1997). These observations together with the known small chromosome number of these plants (Legro and Doorenbos, 1969, 1971) are consistent with diploidy.

**Genetic diversity**—Four of the seven populations of *B. homonyma* had no variation at any locus; the others averaged 1.07–1.2 alleles per locus, with an average over all populations of 1.07 (Table 2). These values were lower than those found in *B. dregei*, which ranged between 1.07 and 1.27, with an average of 1.16 alleles per locus. Percentage of polymorphic loci ranged from 0 to 20%, with an average of 5.7% among the seven populations of *B. homonyma*. *Begonia dregei* had an average of 13.9% loci that were polymorphic, ranging from 7 to 27%. The values of these measures of genetic diversity were low relative to averages found within populations of other species of plants (Hamrick and Godt, 1996). Across all populations of both species, 80% of loci were polymorphic, with an average of 3.5 alleles per locus; these values were higher than the averages reported for variation within species (Hamrick and Godt, 1996).

Twelve of the 52 alleles detected occurred only in one of each of the 19 populations. In *B. dregei*, populations 4 (Lupatana) and 7 (Horseshoe) had three alleles that were not found in any other population, and population 2 (Glenrosa) had two unique alleles, while four other populations each had one allele not found in other populations. In *B. homonyma*, only population 17 (Isinuka) had an allele found in no other population. Ten of these unique alleles occurred at a frequency of 1.0, an indication that there is a minimum of exchange of genes among these populations.

Observed and expected heterozygosities, *H*<sub>o</sub> and *H*<sub>e</sub>, varied among populations, with no value >0.10 for either species (Table 2). Intrapopulation variation was higher in *B. dregei* than in *B. homonyma*. On average, the expected heterozygosity (*H*<sub>e</sub> = 0.043) was lower than that reported

TABLE 5. *F* statistics for levels of gene flow within and among populations of *B. homonyma*. Standard deviations are in parentheses.<sup>a</sup>

Locus	<i>F</i> <sub>is</sub>	<i>F</i> <sub>it</sub>	<i>F</i> <sub>st</sub>
<i>Gldh</i>	1.000 (0.000) ***	1.000 (0.000) ***	0.831 (0.203) ***
<i>Pgi</i>	1.000 (0.000) ***	1.000 (0.000) ***	0.860 (0.143) ***
<i>6Pgd</i>	0.242 (0.266) n.s.	0.828 (0.107) ***	0.774 (0.159) ***
<i>Sod-1</i>	-0.193 (0.968) n.s.	0.896 (0.112) ***	0.912 (0.094) ***
Over all	0.576 (0.289) ***	0.973 (0.019) ***	0.937 (0.027) ***

<sup>a</sup> n.s. = not statistically significantly different from 0; \*\*\* *P* < 0.001.

for endemic species (*H*<sub>e</sub> = 0.063) by Hamrick and Godt (1996), which in general are expected to have much less genetic variation than more widespread species. Significant departure from Hardy Weinberg expectations was found for most loci (Table 2), with a deficiency of heterozygotes. The negative fixation index values may indicate an outbred population, although this may be a function of sampling error since those populations were polymorphic at only one locus (Table 2). The four populations of *B. homonyma*, 13, 15, 16, and 18, that had no polymorphism at all, had a fixation index of 1.00; one other population of *B. homonyma* (17) and three of *B. dregei* (3, 4, and 5) had fixation indexes above 0.60. The mean fixation index for *B. dregei* was 0.277 and for *B. homonyma* was 0.562, indicating that *B. homonyma* is more highly inbred than *B. dregei*.

**Relatedness between populations**—Nei's genetic distances in both *B. dregei* and *B. homonyma* (Table 3) were high compared to those usually found for populations within a species. The mean genetic distances were 0.640 and 0.520, respectively. Clustering of populations based on *D* shows that only one pair of populations, 13 and 19, in *B. homonyma* cluster closely together (Fig. 2a, b), but others do not show an overall pattern of clear groups (Fig. 2a). When both species are considered together, *B. dregei* and *B. homonyma* do not form separate groups (Fig. 2c). The highest genetic distance values of 1.00 were found within *B. dregei* and *B. homonyma*, as well as in the comparisons between the two species (Table 3).

Geographic distances between populations varied from 1 to 387 km. A Mantel test to compare the matrices of genetic and geographic distances showed that there was no relationship between them (*r* = -0.060, *P* = 0.348). In addition, regression analysis of  $\hat{M}$  against the log of

TABLE 6. *F* statistics for levels of gene flow within and among populations of *B. dregei* and *B. homonyma*. Standard deviations are in parentheses.<sup>a</sup>

Locus	<i>F</i> <sub>is</sub>	<i>F</i> <sub>it</sub>	<i>F</i> <sub>st</sub>
<i>Est-1</i>	0.328 (0.263) **	0.921 (0.037) ****	0.882 (0.049) ****
<i>Est-2</i>	-0.051 (0.988) n.s.	0.989 (0.011) ****	0.990 (0.010) ****
<i>Gldh</i>	0.383 (0.275) ****	0.835 (0.090) ****	0.733 (0.124) ****
<i>Per</i>	0.051 (0.273) ****	0.951 (0.031) ****	0.900 (0.048) ****
<i>Pgi</i>	0.461 (0.266) ****	0.890 (0.060) ****	0.796 (0.071) ****
<i>6Pgd</i>	0.220 (0.089) **	0.666 (0.096) ****	0.571 (0.146) ****
<i>Sod-1</i>	-0.075 (0.141) n.s.	0.895 (0.078) ****	0.902 (0.073) ****
Over all	0.325 (0.68) ****	0.933 (0.036) ****	0.901 (0.025) ****

<sup>a</sup> n.s. = not statistically significantly different from 0; \*\* *P* < 0.005; \*\*\*\* *P* < 0.0002.

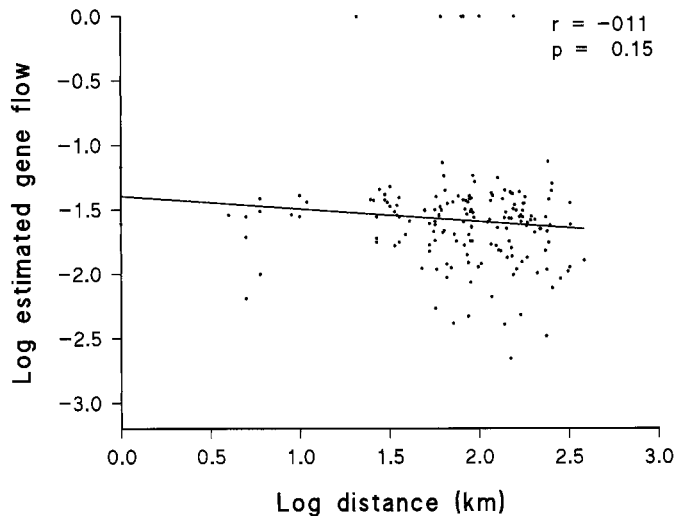


Fig. 3. Plot of the log of estimated gene flow ( $\hat{M}$ ) against the log of geographic distance in kilometres for all pairs of populations.

geographic distance showed that these populations did not fit a model of isolation by distance ( $P = 0.15$ ) (Fig. 3). Negative values of  $\log \hat{M}$  mean that the true values of  $F_{st}$  for those pairs of locations were extremely high. Presence of unique alleles and high rates of fixation at some loci contributed to the low values of  $\hat{M}$ . The linear distance between populations is unlikely to be the route of dispersal; it is more likely that gene flow would occur along watercourses, where forests are more or less continuous. However, the four populations from the Umzimvubu River (7, 8, 17, and 18) were not similar to each other, nor were the two from the Umtamvuna River valley (3 and 14).

**F statistics**—In *B. dregei* both  $F_{it}$  and  $F_{st}$  had high values that were significantly different from zero (Tables 4, 6). *Begonia homonyma* shows similar results, where  $F_{it}$  and  $F_{st}$  are significantly different from zero for all loci (Tables 5, 6). Values of  $F_{is}$  were significantly different from zero for three of the seven loci in *B. dregei* and two of the four in *B. homonyma*, an indication of deficiency of heterozygotes within populations. In both species the high values of  $F_{it}$  and  $F_{st}$  imply deficiencies of heterozygotes globally and among populations, indicating that there have been low levels of gene flow among populations (Hamrick and Godt, 1990). These estimates for all three metrics were higher than those reported for other groups of plants (Linhart et al., 1981; Guries and Ledig, 1982; Alvarez-Buylla and Garay, 1994).

## DISCUSSION

The pattern of distribution of allozyme variation in both species includes low genetic variation within populations and pronounced divergence among populations, consistent with the prediction for the effects of limited gene flow among small populations that have been separated for a long time. Heterozygote deficiency might be due to self-fertilization in these self-compatible plants or inevitable mating between relatives in small populations. *Begonia dregei* and *B. homonyma* occur in populations

with total numbers ranging from four to a few hundred individuals; *B. homonyma* has both less genetic variation and smaller populations than *B. dregei*. Over the summer flowering season, from 10 to 50% of the individuals in a population may produce flowers, and not all of those will produce seed (data not shown).

There is evidence for little, if any, gene flow among populations of either *B. dregei* or *B. homonyma* or between the species. Unique alleles and the high values of  $F_{st}$  are indicators that there is little exchange of genes among populations, even those that are a few kilometres apart within the same forest. The structuring of variation, with very high average values of  $F_{st}$ , is above average even for selfing plants. Populations 8 and 18 in the same forest patch within a kilometre of each other exhibit very little gene flow between them.

The high values of  $F_{st}$  and genetic distance compared to those found in other surveys of plant population structure might possibly be due to greater resolution of variants by polyacrylamide gel electrophoresis rather than the more routinely used starch gels. The first estimates of population heterozygosities were based on polyacrylamide gels (Hubby and Lewontin, 1966), and estimates based on starch gels do not differ substantially. Only a few studies have compared the resolution of the different types of gels with the same samples. A few more electrophoretic variants of human blood proteins were detected on starch rather than on polyacrylamide gels (McLellan, Jorde, and Skolnick, 1984). Comparison of an enzyme from *Drosophila* showed that each type of gel revealed unique variants, and that there were differences between laboratories in the resolution of starch gels (Coyne et al., 1979). Therefore, the resolution of allelic variants on polyacrylamide gels performed in this study is likely to be within the range of results obtained from starch gels.

Limited potential for dispersal is likely to contribute to the low rates of gene flow between populations. Ågren and Schemske (1993) also found very low outcrossing rates and substantial inbreeding in *Begonia hirsuta* Aubl. and *Begonia semiovata* Liebm. due to lack of pollen transfer and short-distance seed dispersal. The distribution of variation among populations of *B. dregei* and *B. homonyma* may in part be attributable to a lack of seed dispersal among populations. The lack of correlation between geographic and genetic distances might be a result of genetic drift, over a long period of time, obscuring any geographic pattern in the distribution of genetic variation. The lack of similarity between populations along the same river is consistent with a lack of gene flow as well.

The distinction between *B. dregei* and *B. homonyma*, based on the size and shape of leaves (Hilliard, 1976), is not supported by the allozyme data. Most populations appear too divergent for allozymes to be informative about relatedness and evolutionary history, although allozymes are effective indicators of gene flow in these plants. The high differentiation among populations, with mean genetic distances of over 0.50 in both species, suggests that perhaps additional taxa should be recognized. There is heterogeneity within both *B. dregei* and *B. homonyma* in leaf shape and size (Hilliard, 1976; McLellan, 2000). The two populations with the lowest value of genetic distance, populations 13 and 19, despite their geo-

graphic separation, are similar in terms of leaf shape and other aspects of morphology. For the remainder of the group, perhaps additional information from morphology, as well as other molecular data, will clarify an appropriate taxonomic framework. Some populations with similar leaf shapes and near each other geographically differ greatly in allozymes (populations 8 and 9 in *B. dregei*; populations 14, 15, and 16 in *B. homonyma*). The lack of concordance between leaf shape and allozymes suggests a lack of linkage between genes for the two sets of traits, so that they could have evolved independently of each other. It is possible that leaf shape and size are influenced by selection, and allozymes, which are usually assumed to be selectively neutral, are more affected by genetic drift, but we do not have data to distinguish selection from drift for either source of data.

Although many plant species have declined rapidly during the past few hundred years due to human activities, many others owe their population disjunctions to climatic changes during the quaternary (Sage and Wolff, 1986). The small patches of coastal forest in which *B. dregei* and *B. homonyma* occur are apparently remnants of a once continuous forest, which began to break up in the Miocene (Burgess, Clarke, and Rodgers, 1998). The distribution of genetic variation in *B. dregei* and *B. homonyma* is consistent with the ancient separation of forest patches, limiting the dispersal of plants that are specific to forest habitats.

By current categories for threatened taxa (IUCN, 1994), neither *B. dregei* nor *B. homonyma* would be considered threatened. We have found 50 populations; there are likely to be at least twice that many. Although population sizes are small, with an average of <20 mature individuals, the total number of individuals is reasonably high. Most of the forests are well protected, and we have seen only a few cases of recent habitat destruction. *Begonia homonyma* is collected for traditional medicine in KwaZulu/Natal (Cunningham, 1988; N. Crouch, personal communication, 1998). The probability of extinction of any single population is high, since the populations are so small (Hughes, Daily, and Ehrlich, 1997). Almost every one of the populations that we have considered here is unique in its combination of molecular and morphological traits. It is highly likely that additional species from African coastal forests and other habitats with this sort of island distribution and limited dispersal abilities will show similar patterns of variation. This type of variation demonstrates that conservation efforts should be oriented below the level of species, especially in highly variable taxa.

#### LITERATURE CITED

- ACOCKS, J. P. H. 1988. Veld types of South Africa. *Memoirs of the Botanical Survey of South Africa* 28: 1–146.
- ÅGREN, J., AND D. W. SCHEMSKE. 1993. Outcrossing rate and inbreeding depression in two annual monoecious herbs, *Begonia hirsuta* and *B. semiovata*. *Evolution* 47: 125–135.
- ALVAREZ-BUYLLA, E. R., AND A. A. GARAY. 1994. Population genetic structure of *Crecropia obtusifolia*, a tropical tree species. *Evolution* 48: 437–453.
- ARADYA, K. M., D. MUELLER-DOMBOIS, AND T. A. RANKER. 1991. Genetic evidence for recent and incipient speciation in the evolution of Hawaiian *Metrosideros* (Myrtaceae). *Heredity* 67:129–138.
- BARRETT, S. H. C. 1996. The reproductive biology and genetics of island plants. *Philosophical Transactions of the Royal Society of London* 351: 725–733.
- BURGESS, N. D., G. P. CLARKE, AND W. A. RODGERS. 1998. Coastal forests of eastern Africa: status, endemism patterns and their potential causes. *Biological Journal of the Linnean Society* 64: 337–367.
- CARSON, H. L. 1987. The processes whereby species originate. *BioScience* 37: 715–720.
- COOPER, K. H., AND W. SWART. 1992. Transkei forest survey. Wildlife Society of Southern Africa, Durban, South Africa.
- COYNE, J. A., W. F. EANES, J. A. M. RAMSHAW, AND R. K. KOEHN. 1979. Electrophoretic heterogeneity of  $\alpha$ -glycerophosphate dehydrogenase among many species of *Drosophila*. *Systematic Zoology* 28: 164–175.
- CUNNINGHAM, A. B. 1988. An investigation into the herbal medicine trade in Natal/KwaZulu. Institute of Natural Resources, University of Natal, Pietermaritzburg, South Africa.
- FEELEY, J. M. 1986. The distribution of iron age farming settlement in Transkei: 470–1870. Master's thesis, University of Natal, Pietermaritzburg, South Africa.
- GELDENHUYS, C. J. 1994. Bergwind fires and the location pattern of forest patches in the southern Cape landscape, South Africa. *Journal of Biogeography* 21: 49–62.
- GODT, M. J., AND J. L. HAMRICK. 1993. Genetic diversity and population structure in *Tradescantia hirsuticaulis* (Commelinaceae). *American Journal of Botany* 80: 959–966.
- GOTTLIEB, L. D., AND N. F. WEEDEN. 1979. Gene duplication and phylogeny in *Clarkia*. *Evolution* 33: 1024–1039.
- GOUDET, J. 1995. FSTAT (Version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* 86: 485–486.
- GRANGER, J. E. 1984. Fire in forests. In P. de V. Booysen and N. M. Tainton [eds.], *Ecological effects of fire in South African ecosystems*, 177–197. Springer-Verlag, Berlin, Germany.
- GURIES, R. P., AND F. T. LEDIG. 1982. Genetic diversity and population structure in pitch pine (*Pinus rigida* Mill.). *Evolution* 36: 387–402.
- HAMRICK, J. L., AND M. J. W. GODT. 1990. Allozyme diversity in plant species. In H. D. Brown, M. T. Clegg, M. A. Kahler, and B. S. Weir [eds.], *Plant population genetics, breeding and genetic resources*, 43–63. Sinauer, Sunderland, Massachusetts, USA.
- , AND ———. 1996. Conservation genetics of endemic plant species. In J. C. Avis and J. L. Hamrick [eds.], *Conservation genetics: case histories from nature*, 281–304. Chapman and Hall, New York, New York, USA.
- HAMRICK, J. L., Y. B. LINHART, AND J. B. MITTON. 1979. Relationship between life history characteristics and electrophoretically detectable genetic variation in plants. *Annual Review of Ecology and Systematics* 10: 173–200.
- HILLIARD, O. M. 1976. Begoniaceae. In J. H. Ross [ed.], *Flora of Southern Africa*, vol. 22, 136–144. Botanical Research Institute, Department of Agriculture Technical Services, Pretoria, South Africa.
- HUBBY, J., AND R. C. LEWONTIN. 1966. A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. *Genetics* 54: 577–594.
- HUGHES, J. B., G. C. DAILY, AND P. R. EHRLICH. 1997. Population diversity: its extent and extinction. *Science* 278: 689–692.
- ITO, M., AND M. ONO. 1990. Allozyme diversity and the evolution of *Crepidiastrum* (Compositae) on the Bonin (Ogasawara) Islands. *Tokyo Botanical Magazine* 103: 449–459.
- IUCN. 1994. IUCN Red list categories. IUCN, Gland, Switzerland.
- LEBERG, P. L. 1992. Effect of population bottlenecks on genetic diversity as measured by allozyme electrophoresis. *Evolution* 46: 477–494.
- LEGRO, R. A. H., AND J. DOORENBOS. 1969. Chromosome numbers in *Begonia*. *Netherlands Journal of Agricultural Science* 17: 189–202.
- , AND ———. 1971. Chromosome numbers in *Begonia*, 2. *Netherlands Journal of Agricultural Science* 19: 176–183.
- LEWIS, P., AND D. ZAYKIN. 1997. Genetic data analysis, version 1 d 10, University of New Mexico, Albuquerque, New Mexico, USA.
- LINHART, Y. B., J. B. MITTON, K. B. STURGEON, AND M. L. DAVIS. 1981. Genetic variation in space and time in a population of ponderosa pine. *Heredity* 46: 407–426.
- LÖNN, M., AND H. C. PRENTICE. 1995. The allozyme structure and leaf

- shape variation in isolated range-margin populations of the shrub *Hippocrepis emerus* (Leguminosae). *Ecography* 18: 276–285.
- MCLELLAN, T. 1982. Electrophoresis buffers for polyacrylamide gels at various pH. *Analytical Biochemistry* 126: 94–99.
- . 2000. Geographic variation and plasticity of leaf shape and size in *Begonia dregei* and *B. homonyma* (Begoniaceae). *Botanical Journal of the Linnean Society* 132: 79–95.
- , L. B. JORDE, AND M. H. SKOLNICK. 1984. Genetic distances between the Utah Mormons and related populations. *American Journal of Human Genetics* 36: 836–857.
- NEI, M. 1978. Estimation of average heterozygosity and genetic distance from small numbers of individuals. *Genetics* 89: 330–590.
- NTLOKO, T. S. S. 1997. Molecular variation in *Begonia dregei*. M.Sc. Thesis, University of the Witwatersrand, Johannesburg, South Africa.
- OUBORG, N. J., R. VAN TREUREN, AND J. M. M. VAN DAMME. 1991. The significance of genetic erosion. II. Morphological variation and fitness components in populations of varying size of *Salvia pratensis* L. and *Scabiosa columbaria* L. *Oecologia* 86: 359–367.
- PODOLSKY, R. H., AND T. P. HOLTSFORD. 1995. Population structure of morphological traits in *Clarkia dudleyana* I. Comparison of  $F_{ST}$  between allozymes and morphological traits. *Genetics* 140: 733–744.
- ROHLF, F. J. 1993. NTSYS-pc. Numerical taxonomy and multivariate analysis system, version 1.80. Exeter Software, Setauket, New York, USA.
- RUTHERFORD, M. C., AND R. H. WESTFALL. 1986. Biomes of southern Africa—an objective categorization. *Memoirs of the Botanical Survey of South Africa* 54: 1–98.
- SAGE, R. D., AND J. O. WOLFF. 1986. Pleistocene glaciations, fluctuating ranges, and low genetic variability in a large mammal (*Ovis dalli*). *Evolution* 40: 1092–1095.
- SLATKIN, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* 47: 264–279.
- SMITH, L. B., D. C. WASSHAUSEN, J. GOLDING, AND C. E. KAREGEANNES. 1986. Begoniaceae. Part I: illustrated key, Part II: annotated species list. Smithsonian Institution Press, Washington, DC, USA.
- THOMAS, J. C., AND W. J. BOND. 1997. Genetic variation in an endangered cedar (*Widdringtonia cedarbergensis*) versus two congeneric species. *South African Journal of Botany* 63: 133–140.
- VAN TREUREN, R., R. BIJLSMA, W. VAN DELDEN, AND N. J. OUBORG. 1991. The significance of genetic erosion in the process of extinction. I. Genetic differentiation in *Salvia pratensis* and *Scabiosa columbaria* in relation to population size. *Heredity* 66: 181–189.
- WEIR, B. S. 1996. Genetic data analysis II. Methods for discrete population genetic data. Sinauer, Sunderland, Massachusetts, USA.
- , AND C. C. COCKERHAM. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.
- WENDEL, J. E., AND N. F. WEEDEN. 1989. Visualization and interpretation of plant isozymes. In D. E. Soltis and P. S. Soltis [eds.], *Isozymes in plant biology*, 5–45. Chapman and Hall, Washington, DC, USA.
- WOLF, P. S., P. S. SOLTIS, AND D. E. SOLTIS. 1991. Genetic relationships and patterns of allozymic divergence in the *Ipomopsis aggregata* complex and related species (Polemoniaceae). *American Journal of Botany* 78: 515–526.
- WRIGHT, S. 1951. The genetic structure of populations. *Annals of Eugenics* 15: 323–354.



