Variation in the mating system of *Bidens menziesii* (Asteraceae) in relation to population substructure

Kermit Ritland* and
Fred R. Ganders

Department of Botany, University of British Columbia, Vancouver, B.C., Canada V6T2B1.

The mating system of the Hawaiian plant *Bidens menziesii* subsp. *filiformis* was examined in four populations, using progeny arrays assayed at four polymorphic allozyme loci. Average selfing rates ranged near 50 per cent in all populations. Evidence that consanguineous (non-self) matings contributed to selfing was given by (1) gene frequency variation among subsites in two populations and (2) significant among-plant regressions of outcrossing-pollen gene frequency upon ovule genotype, based upon mating system estimates for individual plants. These estimates for individual plants, whose properties are discussed, showed significant among-plant variation for (1) single-locus selfing rates in three populations, (2) multilocus selfing rates in two populations, and (3) pollen gene frequency in three populations.

Expected effects of population substructure upon among-plant variation of selfing and pollen gene frequency are given in terms of Wright's *F* statistics and parameters of the "effective selfing model". In three populations, much of the single-locus selfing variation was manifested in a tendency for inbred plants to effectively self more than outbred plants. This assortment of selfing, termed "inbreeding assortative selfing", maintains the spatial variation of Wright's *F* and is a covariation of gene fixation in parents with genetic similarities between mates. This variation of inbreeding might allow the localised generation and establishment of homozygous genotypes of evolutionary significance.

INTRODUCTION

The mating system determines how genetic variation is distributed to the progeny generation, both within and among individuals, population subdivisions, and populations (Hamrick, 1982). It has been classically described by the mixed mating model (Fyfe and Bailey, 1951; Clegg, 1980), which portrays a breeding population as individuals, each practising the same proportion of self-fertilisation to random outcrossing. For substructured populations, where outcrossing is non-random, the mating system might be more generally described in terms of inbreeding coefficients that describe how genotypes gave rise to zygotes (Ritland, 1985a). However, studies of the nature, extent, and significance of among-plant variation of the mating system, particularly in relation to population substructure, are lacking (Hamrick, 1982).

* Present address: Department of Botany, University of Toronto, Toronto, Ontario M5S 1A1 Canada.

In substructured populations, outcrossing may occur to relatives (Shaw and Allard, 1982) or to individuals made genetically similar by selection (Ellstrand and Foster, 1983). Ennos and Clegg (1982) showed that localised variation of gene frequency, as measured by Wright's F_{st}, raises estimates of selfing above its actual value, and multilocus models can estimate some of this "effective" selfing (Shaw and Allard, 1981; Ritland and Jain, 1981). When random mating occurs within population subdivisions, effective selfing occurs because of a genetic similarity between mates, relative to the entire population. The exact amounts of effective selfing caused by consanguineous or near-neighbour matings has been derived in terms of gene identity coefficients (Ritland, 1984), or alternatively, in terms of coefficients based upon gene covariance (Ritland, 1985a). These derivations are based upon the "effective selfing model", which can describe genetic transmission in substructured populations. In this model, a potential source of selfing variation, the difference between

the effective selfing rates of inbred vs. outbred parents, arises.

Secondly, in a substructured population, the possibility exists that many of the pollen parents encountered by a locally outcrossing maternal plant are mutually related to one another. In the context of the mixed mating model, Schoen and Clegg (1984) proposed a model of single paternal parentage of progeny arrays descended from a common mother, applicable to insect pollinated plants where there is little pollen carryover. Such correlation of paternal parentage might also be caused by population substructure, and could be detected as variation of pollen gene frequency among maternal plants.

The extent of such mating system variation needs to be ascertained through appropriate estimation and statistical procedures (Smyth and Hamrick, 1984). Although estimation procedures commonly assume both uniform selfing rates and pollen gene frequency among maternal plants (Brown and Allard, 1970; Clegg, Kahler and Allard, 1978), studies have demonstrated among-plant variation in selfing rates (Sanders and Hamrick, 1979; Smyth and Hamrick, 1984). However, this variation in selfing rates may actually be due to correlated paternal parentage of maternal progeny arrays (Smyth and Hamrick, 1984), which in causing pollen gene frequency to vary among maternal families, causes spurious estimates of selfing variation if pollen gene frequency is assumed constant among maternal plants. Thus, estimates of among-plant variation of the mating system should jointly consider variation of selfing and pollen gene frequency.

Finally, any potential significance of the effective mating system needs identification. The effective mating system is a function of pollen and seed dispersal distance in relation to localised genetic divergence. Its allowance of the formation of neighbourhoods of inbred plants should influence the rate of appearance and maintenance of any novel variation hidden in combinations of recessive genes. Populations of the genus *Bidens* in the Hawaiian Islands characteristically occur in small, geographically isolated locales, subject to a diverse selective regime and limited dispersability (Carlquist, 1966). Infrequent interisland dispersal events have often resulted in taxonomic differentiation (Ganders and Nagata, 1984). Variation of the effective mating system among plants or population subdivisions might have influenced the adaptive radiation observed in this genus.

In this paper, the mating system and its variation are described for four populations of *Bidens*

menziesii (Gray) Sherff subsp. *filiformis* (Sherff) Ganders and Nagata (Asteraceae), using both single-locus and multilocus estimation procedures. This subspecies is endemic to the island of Hawaii, and it has large, dense inflorescences of many small flower heads. The flowers are self-compatible but protandrous, and populations are gynodioecious. Parameters of the effective selfing model, together with Wright's *F* statistics (Wright, 1969), are used to characterise the among-plant variation of mating system components observed in *Bidens menziesii* subsp. *filiformis*.

MATERIALS AND METHODS

(i) Collection and electrophoresis

Seeds were collected from individual plants in four populations of *B. menziesii* subsp. *filiformis* occurring on four isolated cinder cones on the south slope of Mauna Kea on the island of Hawaii. Samples from populations on the Ahumoa (elev. 2150 m) and Puu Koko (elev. 1700 m) cinder cones were collected during November 1981 and samples from populations on the Nohona o Hae (elev. 990 m) and Puu Koohe (elev. 2000 m) cinder cones were collected during November 1982. The 1982 samples were collected from 4 and 3 "subpopulations" within each respective population, these subpopulations separated by 100–200 m of intervening plants.

Progenies of individual plants (families) were kept separate, grown as seedlings, and electrophoretically assayed for four enzyme loci: phosphoglucoisomerase 1 and 2 (*Pgi*-1 and *Pgi*-2) and phosphoglucomutase 3 and 4 (*Pgm*-3 and *Pgm*-4). Sample sizes are given in table 1; family size was emphasised over number of families to maximise the power of individual plant estimates and the certainty of maternal genotype inference. Electrode buffer was 0·06 M lithium hydroxide/0·03 M boric acid, gels were 12 per cent starch in 0·03 M tris/0·05 mM citric acid monohydrate/1 per cent electrode buffer. Gels were cut in two slices, and stained for *Pgm* and *Pgi* using recipes of Roose and Gottlieb (1976). The genetic control of and assignment of alleles to the duplicated loci given in Helenurm (1983) and Helenurm and Ganders (1985).

(ii) Estimation of effective selfing model parameters

The effective selfing model has three independent parameters, and there are actually many triplets

Table 1 Gene frequencies, samples sizes, and number of effectively polymorphic loci for populations surveyed in this study. N is total sample size

Population	Pgi-1			Pgi-2			Pgm-3			Pgm-4		Number of families	Polym. loci
	a	b	c	a	b	c	a	b	c	a	b		
Nohona o Hae	0·60	0·18	0·22	0·93	0·05	0·02	0·91	0·09		0·94	0·06	42	3
N		337			337			138			138		
Ahumoa	0·67	0·29	0·04	0·94	0·03	0·03	0·55	0·35	0·10	0·89	0·11	33	3
N		342			342			336			336		
Puu Koko	0·44	0·52	0·04	1·00			0·58	0·42		1·00		29	2
N		332			332			314			55		
Puu Koohe	0·07	0·93		1·00			0·96	0·04		0·53	0·47	33	1–2
N		332			311			203			203		

of parameters that sufficiently describe this single-locus model of the mating system (Ritland, 1985a). In this paper, E, D, and F were estimated. The effective selfing rate E is the amount of genetically equivalent selfing caused by self-fertilisation and/or mating to relatives. It equals the self-fertilisation rate when there are no consanguineous matings, and population substructure causes it to be greater than the self-fertilisation rate. If there is no self-fertilisation, it equals $2F_{st} - G_{st}$, where F_{st} is given by Wright (1969) and G_{st} is a normalised third moment of gene frequency variation among randomly mating subdivisions. Positive E maintains the average level of Wright's F. The second parameter is the "inbreeding assortative selfing rate" D, which is proportional to the extent that inbred parents effectively self more (or less) than outbred parents, *i.e.*, the degree that effective selfing assorts with respect to fixation of parental alleles. Positive D maintains the variation in the level of Wright's F among individuals or population subdivisions. The third parameter is Wright's F of the parental generation.

Progeny genotypes descended from common mothers were used to estimate E, D and F (F being of maternal parents). Maximum likelihood estimates of these single-locus parameters were obtained as weighted averages over loci, using procedures in Ritland (1985b). Variance of estimates were found by inversion of the information matrix.

(iii) Estimation of the mating system for individual plants

As mentioned in the introduction, when single-locus estimates of selfing are performed for individual plants (using their progeny) and the pollen gene frequency is assumed constant among plants, the estimated selfing rate is biased if deviations of pollen gene frequency are encountered during mating.

By contrast, if selfing rate and the outcrossing-pollen pool could be separately estimated (cf. Brown and Allard, 1970) for individual plants, the selfing estimate would be unbiased. However, such estimation is not possible for a single diallelic locus, because there is only one degree of freedom available within an array of progeny descended from a common mother. But consider two loci, with alleles A and a at one locus and alleles B and b at a second locus, and where in family k, $p_k = 1 - q_k$ is the frequency of allele A and $u_k = 1 - v_k$ is the frequency of B in paternally donated gametes, and s_k is the actual self-fertilisation rate. Given maternal parent $AABB$, there are four progeny genotypes possible, with mixed mating model probabilities, for independent loci, of

$$P(AABB) = s_k + (1 - s_k)p_k u_k$$
$$P(AaBB) = (1 - s_k)q_k u_k$$
$$P(AABb) = (1 - s_k)p_k v_k$$
$$P(AaBb) = (1 - s_k)q_k v_k. \qquad (1)$$

Since there are three degrees of freedom, s_k, p_k and u_k are now jointly estimable. Generally, multilocus data allows such joint estimation; the exception is when the maternal parent is heterozygous for at least $n - 1$ diallelic loci with $p = 0.5$ at all these loci (Ritland and El-Kassaby 1985). Diallelic loci with p near 0.5 would likewise contribute little information. For (1), the joint maximum likelihood estimates, which exactly fit the data to expectations, are, for observed frequencies f,

$$\hat{s}_k = [f(AABB)f(AbBb)$$
$$- f(AaBB)f(AABb)]/f(AaBb)$$

$$\hat{p}_k = f(AABb)/[f(AABb) + f(AaBb)]$$

$$\hat{u}_k = f(AaBB)/[f(AaBB) + f(AaBb)]. \qquad (2)$$

If consanguineous matings occur, a second problem arises. The multilocus segregation pattern *within* a progeny array is not generally expected to fit mixed mating model expectations (the single-locus pattern does fit, as proven in Ritland, 1984). This is because for a given mating, self-fertilisation causes selfing at all loci simultaneously, whereas a consanguineous mating is surmised to effectively result in selfing at some loci and random mating at other loci. To characterise this, define the two-locus joint selfing rate as s_k and the single-locus effective selfing rate given two locus outcrossing as e_k. The probabilities of offspring are then

$$P(AABB) = s_k + (1 - s_k)[e_k^2 + e_k(1 - e_k)(p_k + u_k)$$

$$+ (1 - e_k)^2 p_k u_k]$$

$$P(AaBB) = (1 - s_k)[e_k(1 - e_k)q_k + (1 - e_k)^2 q_k u_k]$$

$$P(AABb) = (1 - s_k)[e_k(1 - e_k)v_k + (1 - e_k)^2 p_k v_k]$$

$$P(AbBb) = (1 - s_k)(1 - e_k)^2 q_k v_k. \qquad (3)$$

There are too many parameters for the available degrees of freedom. But the estimates obtained by imposing model (3) and using (2) are

$$\hat{s}_k = s_k$$

$$\hat{p}_k = p_k + e_k q_k$$

$$\hat{u}_k = u_k + e_k v_k. \qquad (4)$$

This shows that if mixed-mating model estimates are performed on individual progeny arrays derived in part from consanguineous matings, we can obtain the n-locus selfing rate, whereas the pollen gene frequency estimate includes a component, $e_k q_k$, of pollen alleles related to the maternal genotype (caused by consanguineous matings). However, it should be noted that mating to closer relatives also causes some n-locus selfing. In summary, the multilocus selfing estimate would include selfing caused by mating to stronger relatives, and the multilocus pollen gene frequency estimate can include both a correlated paternal parentage component and a consanguineous-mating component.

With the *Bidens menziesii* allozyme data, first, single-locus estimates of selfing were found for each plant as a weighted mean over loci, assuming pollen gene frequency was constant among plants, and inferring maternal parentage (Brown and Allard, 1970). Negative estimates of selfing (possible with statistical error) were truncated to zero so that the statistical error could be partitioned

out (see discussion). Second, the multilocus model of Ritland and Jain (1981) estimated selfing rate and pollen gene frequencies for each plant.

Since these individual plant estimates are based upon relatively few progeny, the asymptotic variances of s and p obtained by inverting the expected family information matrix (Ritland and El-Kassaby, 1985) are too small. Instead, the significance of among-plant variances was found with Monte Carlo methods. Monte Carlo data sets were created as in Ritland and Jain (1981), with parameters equal to the population estimates of s and p and with the same distribution of family sizes as in the data; s (and p for the multilocus model) estimates were then performed for each family in the Monte Carlo data sets. The distribution of these Monte Carlo estimates of s or p were then compared to the actual distribution of individual plant estimates by (1) subtracting the Monte Carlo estimates from their mean and squaring this deviation, (2) doing likewise for the actual estimates, and (3) performing the non-parametric Mann–Whitney U test to compare the two sets of variates. A significantly larger rank for the actual estimates indicated a significant among-plant variance.

RESULTS AND DISCUSSION

(i) Effective selfing model estimates

The four loci surveyed in this study were chosen on the basis of prior knowledge of genetic polymorphism, necessary for mating system estimation. Adequate polymorphism at some loci, as well as moderate among population genetic differentiation, is evident in table 1.

Effective selfing rates E of plants, relative to the population, ranged from 0·37 to 0·65, with an average of 0·48, and most between-population differences for selfing rate were not statistically significant (table 2). The Puu Koohe population did show a significant excess of selfing; however, we are primarily interested in variation of the

Table 2 Effective selfing model estimates with 95 per cent confidence intervals

Population	Effective selfing rate $\hat{E}$	Fixation index of parents $\hat{F}$	Inbreeding assortative selfing rate $\hat{D}$
Nohona o Hae	0·41 ± 0·09	0·17 ± 0·18	−0·03 ± 0·05
Ahumoa	0·37 ± 0·08	0·23 + 0·18	0·15 ± 0·05
Puu Koko	0·49 ± 0·12	0·25 ± 0·24	0·10 ± 0·11
Puu Koohe	0·65 ± 0·10	0·57 ± 0·23	0·20 ± 0·09

mating system within populations, and fortunately, these intermediate rates of selfing provide the most statistical power for detecting among-plant variation of selfing.

The relative proportions of self-fertilisation vs. consanguineous matings that contribute to this observed E cannot be estimated with single-locus family data. Consanguineous matings are broadly termed here as matings between genotypes similar with respect to population substructure; this includes the spatial arrangement of lineal relatives, as well as the spatial variation of gene frequency as caused by either genetic drift or selective differentiation.

Wright's fixation index F, which measures excess of homozygosity above Hardy–Weinberg expectations, ranged from 0·17 to 0·57 for parent plants, being significantly positive in all but the Nohona O Hae population (table 2). Thus, moderately to strongly inbred plants exist in these populations.

Inbreeding assortative selfing D was significantly positive in three of four populations; the Nohona o Hae population showed a near zero D (table 2). Positive inbreeding assortative selfing maintains the variation of F among population subdivisions through the tendency of plants in inbred neighbourhoods to practise more effective selfing than plants in outbred neighbourhoods. This can occur even if inbred plants receive pollen from the same distance as outbred plants, since the strength of relatedness between adjacent plants is likely to be larger in the more inbred neighbourhoods of the population.

Thus, the population substructure in these three populations with positive D includes an association between (1) areas of closer relationship among mates and (2) levels of gene fixation F of parents. Positive D also implies that the level of gene fixation varies in the population. Such variation of gene fixation would arise through variation among randomly breeding neighbourhoods for either (1) the rate of genetic drift (due to differing neighbourhood sizes) or (2) the strength of natural selection.

(ii) One measure of consanguineous matings

The effective selfing rate E of a maternal parent is twice the normalised covariance F' between two alleles, one chosen randomly from each mate, minus the normalised third central moment G of the two maternal parent alleles with a randomly chosen paternal parent allele (Ritland, 1985a). This suggests that an "E_{st}", caused by random

mating within a population subdivision, can be found using gene frequency data. The Nohona o Hae and Puu Koohe populations were sampled as 4 and 3 subsites respectively; however, when few subdivisions are censused, the sampling properties of the third moment are poor. At least, E_{st} should lie between F_{st} and $2F_{st}$, since G_{st} lies between 0 and 1. Nei's (1973) D_{st} (not the D of this paper), which is a weighted average of F_{st} over loci (Nei, 1973), was found to be 0·011 in the Nohona o Hae population and 0·094 in the Puu Koohe population. Thus, effective selfing due to localised mating lies between 0·01 and 0·02 in the Nohona o Hae population, and between 0·09 and 0·18 in the Puu Koohe population.

This indicates that no more than 25 per cent of the total selfing in the Puu Koohe population was caused by consanguineous matings. However, since the effective selfing rate of inbred plants is $(EF + D)/F$, the estimates in table 2 indicate near 100 per cent effective selfing was practised by inbred plants of the Puu Koohe population. This predominance of consanguineous matings in presumably inbred neighbourhoods magnifies the role of mating to relatives in maintaining patches of excessively homozygous plants.

(iii) Among-plant variation of selfing: preliminary theory

Estimates of selfing rates for individual plants have a large component of "null" statistical variance. If these estimates are restricted to the interval (0, 1), the among-plant "normalised" variance of selfing (variance divided by $\bar{s}(1 - \bar{s})$, where $\bar{s}$ is the average selfing rate), also ranges in the interval (0, 1). The normalised, among-plant variance of estimates, denoted σ_e^2, can then be expressed, like the partition of Wright's F (Wright, 1969), as

$$\sigma_e^2 = \sigma_n^2 + (1 - \sigma_n^2)\sigma_S^s,$$

where σ_n^2 is the normalised, null statistical variation expected of estimates if there is no true variation of selfing among plants, and σ_S^2 is the normalised, true variance of selfing among plants. This true variance is then estimated as

$$\hat{\sigma}_S^2 = (\hat{\sigma}_e^2 - \hat{\sigma}_n^2)/(1 - \hat{\sigma}_n^2),$$

given that the null statistical variance has been estimated by robust methods, such as Monte-Carlo simulation. Like arguments hold for pollen gene frequency.

Secondly, when inbreeding assortative selfing D is non-zero, there is variation of effective selfing correlated with F. The among-plant variance of

this selfing is denoted $\text{Corr}(\sigma_E^2, F)$, and is the variation about E of the selfing rate of inbred plants, $s^i = (EF + D)/F$, and the selfing rate of outbred plants, $s^o = (E(1-F) - D)/(1-F)$,

$$\text{Corr}(\sigma_E^2, F) = F(s^i - E)^2 + (1-F)(s^o - E)^2,$$

$$= \frac{D^2}{F(1-F)}. \qquad (5)$$

This variation of selfing, correlated with F, maintains the spatial variation of F; variation of selfing uncorrelated with F does not (if D is positive). Its proportion of the estimated true variation of selfing would measure the importance of the observed among-plant variation of selfing as a factor maintaining this aspect of population substructure.

(iv) Among-plant variation of selfing rates: results

The normalised, among-plant variances for both single- and multilocus selfing were quite high for the raw estimates ($\hat{\sigma}_e^2$, table 3), and although statistical variance contributed greatly to this variation ($\hat{\sigma}_n^2$, table 3), the significance test and partitioning formula showed significant single-locus variances of selfing for the Ahumoa, Puu Koko, and Puu Koohe populations ($\hat{\sigma}_E^2$, table 3) and significant multilocus variance of selfing for two of these three populations ($\hat{\sigma}_S^2$, table 3). The distribution of two or three locus selfing rates among families is plotted for each population in fig. 1 (the number of loci depended upon available polymorphism). The "null" distribution of estimate variation, superimposed on fig. 1, graphically shows the size of the random statistical component of among-plant variation of selfing estimates.

Non-statistical variation of single-locus selfing rates can be caused by variation of self-fertilisation among plants (through inherited outcrossing mechanisms such as male sterility or through environmental factors) or by variation in the strength of relatedness among mates. Variation of multilocus selfing rates would be less influenced by mating among relatives, particularly weaker relatives. Variation of multilocus selfing was apparently a large component of the total selfing variation in the Ahumoa population (the ratio $\hat{\sigma}_S^2/\hat{\sigma}_F^2$, table 3) but a small component in the other two populations.

The Nohona o Hae population showed no selfing variance (table 3), yet male sterility was observed in this population (21 per cent male steriles in Nov. 1983) and subsequently in the other populations. However, the development of male-sterility appears to be environmentally sensitive in this species (FRG, personal observation), so that levels of male sterility in the years of seed collection cannot now be accurately determined. The mechanism of inheritance for male sterility (selfed progeny tend to be male steriles) would tend to favour negative D values on a local level. Since Nohona o Hae has little variation of gene frequency among subsites and its value of D is near zero, in contrast to the other three populations, population substructure may be largely responsible for the significant estimates of selfing variation in these other three populations.

The predicted variance of single-locus selfing caused by differences between the effective selfing rates of inbred vs. outbred parents was found by applying formula (5) to the estimates of table 2 and normalising by $E(1-E)$. Half (Puu Koko) or all (Ahumoa and Puu Koohe) of the non-statistical variation of single-locus selfing was accounted for by positive D (table 3). This is additional evidence

Table 3 Estimates of normalised, true among-plant variances of (1) effective selfing E (σ_E^2) and (2) multilocus selfing S (σ_S^2); their ratio, and the predicted normalised variance of effective selfing correlated with F, $\text{Corr}(\sigma_E^2, F)/(E(1-E))$. The normalised variance of actual estimates (σ_e^2) and the "null", normalised statistical variance of estimates (σ_n^2, found by simulation) determine the true variance with the partition formula given in the text

Population	1. Effective selfing			2. Multilocus selfing			Ratio†	Predicted
	$\hat{\sigma}_e^2$	$\hat{\sigma}_n^2$	$\hat{\sigma}_E^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_n^2$	$\hat{\sigma}_S^2$	$\hat{\sigma}_S^2/\hat{\sigma}_F^2$	$\text{Corr}(\sigma_E^2, F)/E(1-E)$
Nohona o Hae	0·48	0·43	0·09	0·19	0·19	0·00	0·00	0·00
Ahumoa	0·53	0·25	0·37*	0·43	0·19	0·30*	0·81	0·54
Puu Koko	0·63	0·27	0·49*	0·23	0·13	0·11*	0·22	0·21
Puu Koohe	0·86	0·35	0·78*	0·16	0·15	0·01	0·01	0·71

† significance not determined.
* $p < 0.01$.
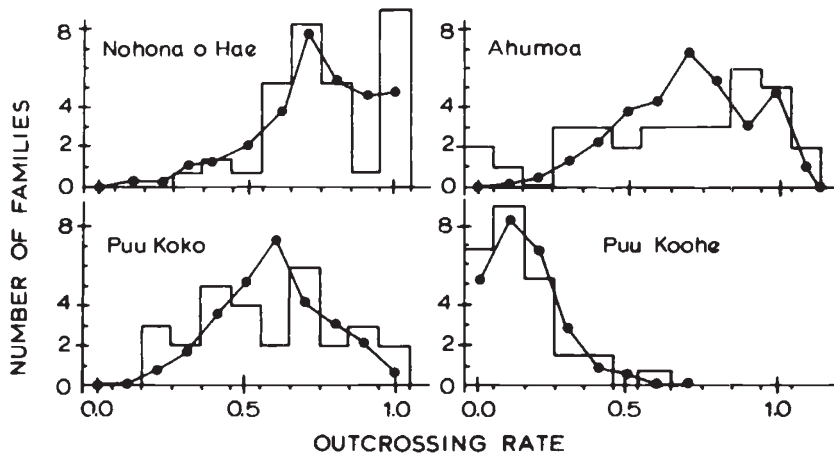
Figure 1 Among-plant (family) variation of multilocus selfing rates, for each of the four populations surveyed in this study. Histograms are actual estimates, and lines form "null" distribution found by Monte-Carlo simulation (outcrossing rate = 1 − selfing rate).

for significant effects of population substructure upon selfing variation.

(v) Among-plant variation of pollen gene frequency: preliminary theory

Among-plant variation of the outcrossing-pollen gene pool is the second and more complex aspect of variation in the mating system, complex because its variation is confounded by any variable components of effective selfing. For clarity, the outcrossing-pollen gene pool is defined as consisting of two components, (1) the "effective selfing-pollen gene pool": pollen alleles among effectively selfed progeny, and (2) the "effective outcrossing-pollen gene pool": pollen alleles among effectively random outcrossed progeny.

For the k-th plant whose self-fertilisation rate is s_k and whose effective selfing rate is E_k ($E_k \geqq s_k$), let $e_k = (E_k - s_k)/(1 - s_k)$ be the effective selfing rate among outcrossed progeny, let m_k be the gene frequency in the effective selfing-pollen gene pool (m_k depends directly upon maternal genotype), and let p_k be the gene frequency in the effective outcrossing-pollen gene pool. For equal family sizes, the normalised variance of outcrossing-pollen gene frequency is the expectation

$$\sigma_P^2 = E[(e_k m_k + (1 - e_k)p_k - p)^2]$$
$$= [e^2(1 - \sigma_e^2) + e\sigma_e^2]\sigma_m^2$$
$$+ (1 - e)[1 - e(1 - \sigma_e^2)]\sigma_p^2 \qquad (6)$$

where $p = E[m_k] = E[p_k]$, $e = E[e_k]$, and σ_e^2, σ_m^2 and σ_p^2 are the *normalised* variances of e_k, m_k and p_k respectively.

Equation (6) shows there are several components of pollen gene frequency variance. If male and female parents have the same levels of gene fixation, then for pairs of pollen alleles,

$$F_{st} \leqq \sigma_p^2 \leqq \sigma_m^2 = \frac{(1 + F_{it})}{2}$$

where F_{it} and F_{st} are Wright's F statistics (Wright, 1969). The lower range for σ_p^2 occurs with open-pollination, the upper range with single-male parentage; in most natural populations, some intermediate value would occur. Examination of (6) shows three factors that increase variance of pollen gene frequency:
(a) More frequent consanguineous matings (higher e), since such matings increase the contribution of σ_m^2 over σ_p^2;
(b) Greater fixation of alleles within individuals (higher F), since greater F increases σ_m^2 and/or σ_p^2;
(c) Greater variance of effective selfing and/or greater inbreeding assortative selfing D.

(vi) Among-plant variation of pollen gene frequency: results

The estimates of normalised variance of pollen gene frequency were significantly positive for the Ahumoa population at the *Pgi*-1 and *Pgm*-3 loci, for the Puu Koko population at the *Pgi*-1 locus, and for the Puu Koohe population at the *Pgm*-4 locus (table 4). Results of only the most frequent alleles are shown because variation of other alleles

Table 4 Estimates of normalised, true among-plant variances of pollen gene frequency, σ_P^2. This variance is related to the normalised variance of actual estimates, σ_e^2, and its null statistical variance, σ_n^2 (determined by simulation; not given), as $\sigma_P^2 = \sigma_n^2 + (1 - \sigma_n^2)\sigma_e^2$

Population	$Pgi - 1$		$Pgi - 2$		Pgm 3		$Pgm - 4$	
	$\hat{\sigma}_e^2$	$\hat{\sigma}_P^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_P^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_P^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_P^2$
Nohona o Hae	0·29	0·05	0·28	0·00	1·00	1·00	0·92	0·74
Ahumoa	0·46	0·27*	0·23	0·00	0·48	0·31*	0·44	0·26
Puu Koko	0·41	0·17*	—	—	0·34	0·03	—	—
Puu Koohe	0·74	0·00	—	—	—	—	1·00	1·00*

* $p < 0·05$.
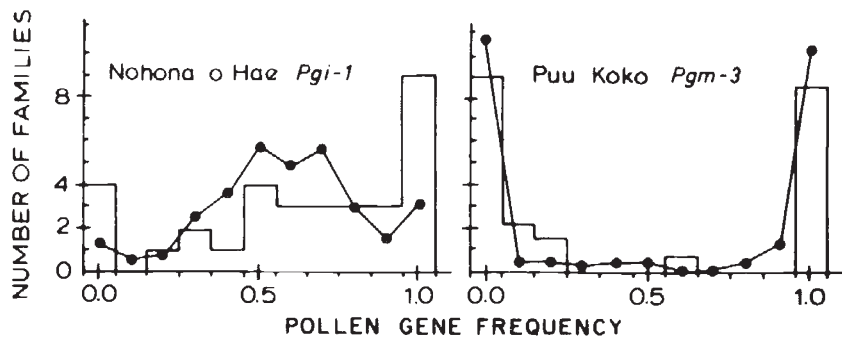
Figure 2 Among-plant (family) variation of outcrossing-pollen gene frequency, for the Nohona o Hae $Pgi - 1$ locus and the Puu Koko $Pgm - 3$ locus. Histograms are actual estimates, and lines form "null" distribution found by Monte-Carlo simulation.

at the same locus are highly correlated with the most frequent allele.

Two distributions of pollen gene frequency, Nohona o Hae Pgi-1 and Puu Koko Pgm-3, are plotted in fig. 2 with their superimposed null distribution, found by Monte Carlo simulation. Pollen gene frequency tended to a (0, 1) distribution in the more highly selfing Puu Koohe population. This near (0, 1) distribution would indicate a tendency for effective parentage by one inbred male. However, the null Pgm-3 distribution, superimposed on fig. 2, assumed a uniform 0·5 gene frequency and also gave a near (0, 1) distribution.

The Nohona o Hae population showed no significant variation of pollen gene frequency, while the other populations showed significant variation at half of their loci (table 4). Nohona o Hae displayed similar E values and lower F and D values relative to the other three populations (table 2), and displayed a notable lack of selfing variation (table 3). This association of higher F and D values with higher variance of the pollen gene pool accords with the above predictions. Thus, population substructure appears to be partly responsible for pollen gene frequency variation in the Ahumoa, Puu Koko and Puu Koohe populations of *Bidens menziesii* ssp. *filiformis*.

(vii) A second measure of consanguineous mating

The effective selfing rate of maternal plants can also be measured by the regression of outcrossing pollen gene frequency on ovule genotype (Ritland, 1985a), where ovule genotypes are assigned the additive values $AA = 1$, $Aa = 0·5$, and $aa = 0$. Although this regression should ideally be conditioned upon maternal fixation of alleles, the pollen allele frequency of each progeny array (inferred with the mutilocus model) regressed significantly with its maternal genotype in the Ahumoa, Puu Koko and Puu Koohe populations (table 5), further evidence for consanguineous matings in these populations. However, the significant regressions were not associated with significant pollen gene variances in table 4, indicating that consanguinity between mates (e) is not a major determinant of pollen gene variation; rather the components of $\sigma_p^2(F_{it}, F_{st})$ likely determine pollen gene variation.

CONCLUSIONS

Both theory and data from populations of *Bidens menziesii* subsp. *filiformis* have illustrated that

Table 5 Regressions of outcrossing pollen gene frequency on additive ovule genotypic value. Regressions for alleles with $p > 0.95$ or $p < 0.05$ are not given because of large errors

Population	Locus			
	$Pgi - 1$	$Pgi - 2$	$Pgm - 3$	$Pgm - 4$
Nohona o Hae	0·06	0·28	−0·57	—
Ahumoa	0·50†	0·22†	0·09	0·20
Puu Koko	0·22	—	0·28*	—
Puu Koohe	−0·05	—	—	0·49†

* $p < 0.05$, † $p < 0.01$.

among-plant variation of selfing can be symptomatic of a certain aspect of population substructure, namely the joint localised covariation of gene fixation with the strength of relatedness between mates. This aspect of population substructure has been rarely, if at all, studied, and requires that the level of gene fixation vary among neighbourhoods within the population.

The mating system variation found in this study is partially that of the "effective" mating system, and it is partly "passive" variation, since the physical processes of self-fertilisation and dispersal distance of pollen and seed may be the same for all neighbourhoods, yet the effective mating system may vary because of the abovementioned aspect of population substructure. However, note that variation of that self-fertilisation, seed dispersal, and pollen dispersal likewise contribute to the formation of this population substructure. Thus the effective mating system is the outcome of an interaction between the factors that form population substructure, and population substructure itself.

Overall, *Bidens menziesii* subsp. *filiformis* shows an intermediate level of selfing with a relatively small component of among-population variation. Variance of gene frequency among subsites and significant regressions of outcrossing-pollen gene frequency upon ovule genotype were separate lines of evidence for consanguineous matings in three of the four populations surveyed. Although only approximately $\frac{1}{4}$ of the selfing was apparently caused by consanguineous matings in the Puu Koohe population, the impact of such relatively low levels of consanguineous matings is magnified by their tendency to occur at higher frequency in more inbred neighbourhoods.

Within each population is mating system variation distinguishable from the high level of statistical "noise" inherent in the relatively small family data sets. Significant variation of single-locus selfing and pollen gene frequencies occurred in the three populations that showed significant regressions of pollen gene frequency on ovule genotype. Although significant variation of multilocus selfing occurred in two of these three populations (multilocus estimates are less affected by consanguineous matings), only in one of these populations did the multilocus variation approach the levels of single-locus variation. Overall, these lines of evidence suggest that variation of effective selfing rates among plants occurs when consanguineous matings are frequently practised in natural populations.

This variation of effective selfing is relatable to parameters of the effective selfing model. Significant variation of selfing occurred in populations with greater inbreeding assortative selfing D, and significant variation of pollen gene frequency occurred in populations with greater F and D values. Theoretical expectations of these variances, in terms of effective selfing model parameters, were given by equations (5) and (6). Much or all of the non-statistical variation of single-locus selfing was correlated with the fixation index of parent plants. This indicates that much of the observed variation of selfing serves to maintain variation in Wright's fixation index F among population subdivisions.

The significant D values detected in three of the four *B. menziesii* populations show that among-plant variation in the level of gene fixation F exists and is being maintained through generations. If the neutral processes of genetic drift and/or migration are causing this variation of F, then all loci in plants within a neighbourhood would attain equal expected levels of gene fixation. Thus in these populations, there should exist inbred neighbourhoods of plants that are excessively homozygous at all loci, relative to some "average" neighbourhood. Each inbred neighbourhood would consist of a subset of alleles that have been sampled from some imaginary, outbred base population.

This aspect of subpopulation differentiation is distinct from the more classical notion of differentiation as consisting of changes of gene frequency; here it would involve a depletion of heterozygosity or allelic diversity at all loci in those localised neighbourhoods practising excessive inbreeding. It can perhaps be termed "inbreeding differentiation" because neighbourhoods differ in their level of gene fixation.

Such differentiation can affect the evolutionary potential of the population. In neighbourhoods with high levels of consanguinity between mates, novel multilocus genotypes may appear because recessive alleles at several loci appear simultaneously, and continued high levels of consan-

guineous matings perpetuate these types through generations. Natural selection may favour some of these types, or if they are maladaptive, outbred neighbourhoods loosely linked to inbred neighbourhoods through limited pollen flow may serve as a reservoir of genetic variability for further "genetic tinkering" by inbred neighbourhoods.

However, it would be premature to attribute any of the adaptive radiation in Hawaiian *Bidens* to this type of population substructure. It is likely that gradual genetic divergence in allopatric populations has been the most important factor in the evolution of Hawaiian *Bidens* (Ganders and Nagata, 1984). Further studies of the nature, extent, and consequences of variation in the mating system are needed.

REFERENCES

BROWN, A. H. D. AND ALLARD, R. W. 1970. Estimation of the mating system in open-pollinated maize using isozyme polymorphisms. *Genetics, 66,* 133–145.

CARLQUIST, S. 1966. The biota of long-distance dispersal. I. Principles of dispersal and evolution. *Quarterly Review of Biology, 41,* 247–270.

CLEGG, M. 1980. Measuring plant mating systems. *Bioscience, 30,* 814–818.

CLEGG, M., KAHLER, A. L. AND ALLARD, R. W. 1978. Estimation of life cycle components of selection in an experimental plant population. *Genetics, 89,* 765–792.

ELLSTRAND, N. C. AND FOSTER, K. W. 1983. Impact of population structure on the apparent outcrossing rate of grain sorghum (*Sorghum bicolor*). *Theoretical and Applied Genetics, 66,* 323–327.

ENNOS, R. A. AND CLEGG, M. T. 1982. Effect of population substructuring on estimates of outcrossing rate in plant populations. *Heredity, 48,* 283–292.

FYFE, J. L. AND BAILEY, N. T. J. 1951. Plant breeding studies in leguminous forage crops. I. Natural crossbreeding in winter beans. J. Agric. Sci., *41,* 371–378.

GANDERS, F. R. AND NAGATA, K. M. 1984. The role of hybridization in the evolution of *Bidens* on the Hawaiian Islands. In Grant, W. F., (ed.) *Plant Biosystematics,* Don Mills, Ontario, Academic Press, Canada, pp. 179–194.

HAMRICK, J. L. 1982. Plant population genetics and evolution. *American Journal of Botany, 69,* 1685–1693.

HELENURM, K. 1983. Genetic differentiation of Hawaiian *Bidens*. MSc Thesis, University of British Columbia, Vancouver, Canada.

HELENURM, K. AND GANDERS, F. R. 1985. Adaptive radiation and genetic differentiation in Hawaiian *Bidens. Evolution, in press.*

NEI, M. 1973. Analysis of gene diversity in subdivided populations. *Proceedings National Academy of Sciences (USA), 70,* 3321–3323.

RITLAND, K. 1984. The effective proportion of self-fertilization with consanguinous matings in inbred populations. *Genetics, 106,* 139–152.

RITLAND, K. 1985a. The genetic mating structure of subdivided populations. I. Open mating model. *Theoretical Population Biology, 27,* 51–74.

RITLAND, K. 1985b. Estimation and optimal progeny size for the mixed mating and effective selfing models. *Biometrics, in press.*

RITLAND, K. AND JAIN, S. K. 1981. A model for the estimation of outcrossing rate and gene frequencies using *n* independent loci. *Heredity, 47,* 35–52.

RITLAND, K. AND EL-KASSABY, Y. A. 1985. The nature of inbreeding in a seed orchard of Douglas-fir as shown by an efficient multilocus model. *Theoretical and Applied Genetics, in press.*

ROOSE, M. AND GOTTLIEB, L. 1976. Genetic and biochemical consequences of polyploidy in *Tragopogon. Evolution, 30,* 818–830.

SANDERS, T. B. AND HAMRICK, J. L. 1979. Variation in the breeding system of *Elymus canadensis. Evolution, 34,* 117–122.

SCHOEN, D. J. AND CLEGG, M. T. 1984. Estimation of mating system parameters when outcrossing events are correlated. *Proceedings National Academy of Sciences (USA), 81,* 5258–5262.

SHAW, D. V. AND ALLARD, R. W. 1981. A multilocus estimator of mating system parameters in plant populations. *Proceedings National Academy of Sciences (USA), 78,* 1298–1302.

SHAW, D. V. AND ALLARD, R. W. 1982. Estimation of outcrossing rates in Douglas-fir using isozyme markers. *Theoretical and Applied Genetics, 62,* 113–120.

SMYTH, C. A. AND HAMRICK, J. L. 1984. Variation in estimates of outcrossing in musk thistle populations. *Journal of Heredity, 75,* 303–307.

WRIGHT, S. 1969. Evolution and the Genetics of Populations. Volume II. *The Theory of Gene Frequencies.* University of Chicago Press, Chicago.