OUTCROSSING IN NATURAL POPULATIONS.
II. *CLARKIA UNGUICULATA*

F. C. VASEK

Department of Life Sciences, University of California, Riverside, California

Flower structure suggests that self-pollination would be expected in *Clarkia exilis* and cross-pollination would be expected in *C. unguiculata*. In the former species, pollen shedding anthers are clustered around the receptive stigma at anthesis. Nevertheless, the surprisingly high estimate of 44.5 per cent outcrossing was obtained from one natural population (Vasek, 1964). In contrast, *C. unguiculata* has long, exserted styles and is markedly protandrous. One would, therefore, expect this species to be cross-fertilizing.

The present paper describes an experiment to estimate the frequency of outcrossing in *C. unguiculata*. It thereby provides outcrossing data with which to compare the breeding systems in these two closely related species.

METHODS

A population near Atascadero in San Luis Obispo County, California (1.8 miles east of Cerro Alto public campground on the Atascadero–Morro Bay road), was selected for this experiment because it was known to be polymorphic for flower color (H. Lewis, personal communication). The population was paced off into approximate five-meter squares. The phenotypes were counted in each square on June 10, 1962. Some of the plants in each square were selected at random and marked with tags indicating the phenotypes. Seeds were collected from the tagged wild plants on July 28, 1962.

The following winter, progenies were grown in a University of California greenhouse in Riverside, California, and scored for flower color. To estimate outcrossing, flower color was scored to determine: 1) the proportion of dominants in progenies of wild recessives; 2) the proportion of re-

cessives in progenies of wild heterozygotes; and 3) the proportion of homozygous to heterozygous dominant wild plants.

Determination of flower color can be made at the seedling stage because anthocyanous seedlings develop into plants with pink flowers and seedlings which lack anthocyanin grow into plants which produce white flowers. Consequently, large numbers of plants can be scored for flower color at the seedling stage in a relatively small space and short time. Several dozen seedlings were transplanted to pots and grown to maturity to verify the correlation between anthocyanous seedlings and pink flowers, and to have plants available for making crosses to determine the mode of inheritance of flower color.

INHERITANCE OF FLOWER COLOR

White flower color in *C. unguiculata* is reported due to homozygosity for two recessive genes (Rasmuson, 1920). Although Rasmuson provides no data, and his statements are not translated without some ambiguity, I interpret his conclusions as indicating that flower color can be produced by either of two dominant genes, one producing "salmon red" and the other producing "light red." Since a salmon color was not observed in the population under study, a single gene difference between pink ("light red"?) and white flowers would be expected.

The seedlings transplanted to pots for verification of flower color included both "pink" flowered offspring of wild white-flowered plants and white-flowered plants. The former were white × pink or F_1 heterozygotes. The latter were homozygous recessives. The pink heterozygotes were self-pollinated, crossed to each other, and test-crossed to white recessives. To test for pos-

TABLE 1. *Inheritance of flower color.*

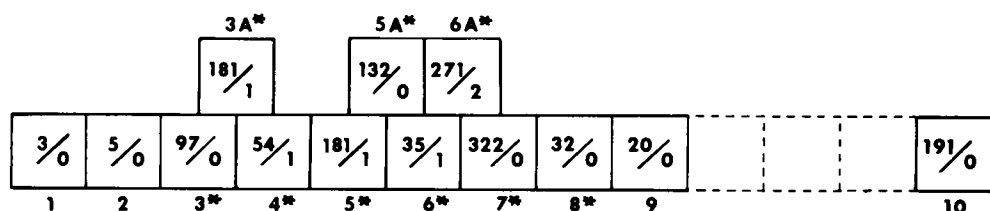| Type of cross	No. of crosses	Progeny		Ratio	Goodness of fit		Heterogeneity	
		Pink	White		χ^2	P	χ^2	P
$F_1 \times$ self	28	1,253	416	3 : 1	0.0049	0.95	30.9137	0.32
$F_1 \times F_1$	16	730	225	3 : 1	1.0991	0.30	13.3245	0.65
White $\times F_1$	8	236	215	1 : 1	0.9778	0.33	17.1546	0.03
$F_1 \times$ white	14	522	469	1 : 1	2.8344	0.09	12.6024	0.56
Sum of test crosses	22	758	684	1 : 1	3.7976	0.06	29.7570	0.10
White $\times$ white	12	0	624	—	—	—	—	—

FIG. 1. Plant density and phenotype distribution in the Atascadero population of *Clarkia unguiculata*. An asterisk denotes those five-meter squares from which seed samples were taken. The numbers in each square indicate the number of pink-flowered and white-flowered plants observed. Total: 1,524 pink, 6 white.

sible genic complementarity, white recessives were crossed to other white recessives.

The results of these crosses, which are found in Table 1, strongly indicate that, in the population under investigation, the difference between pink and white is due to a single gene. These results are consistent with the interpretation of Rasmuson's conclusions if one assumes that the gene for pink is the same as the gene Rasmuson called "light red." Furthermore, since salmon flower color was not observed in this population, the population is evidently homozygous for the recessive allele of "salmon red" or the "salmon red" gene is in very low frequency.

RESULTS

Plants in the population sampled were scattered along and above a road cut for a distance of approximately 65 meters. Thickets of shrubs divide the population into subunits of varying size. The non-uniformity of distribution is indicated by the variation in plant density from 0 to 322 flowering plants per five-meter square, as shown in Fig. 1. Only six, or 0.39 per cent of the wild plants observed were white flowered.

The number of wild plants from which seed samples were taken is shown in Table 2 for each of the eight squares sampled. The frequency of pink- and of white-flowered phenotypes in the progenies grown from the wild seed samples is also shown in Table 2. Of the 80 pink-flowered wild plants sampled, 68, or 85 per cent, produced only pink-flowered offspring and were presumably homozygous. The other 12, or 15 per cent, pink-flowered wild plants were doubtless heterozygous inasmuch as 6.42 per cent of their offspring were white flowered. Five white-flowered wild plants produced progenies in which 88.6 per cent of the offspring were pink flowered.

ANALYSIS AND DISCUSSION

The proportion of pink-flowered plants in the progenies of white-flowered plants, 88.6 per cent in this experiment, is a direct measure of natural crossing, or the frequency of cross-fertilization between dif-

TABLE 2. *Progenies of wild plants.*

Square no.	Pink female parents, Non-segregating progenies			Pink female parents, Segregating progenies			Rh	White female parents			Dr
	No. wild plants	Offspring		No. wild plants	Offspring			No. wild plants	Offspring		
		Pink	White		Pink	White			Pink	White	
3	7	976	–	2	244	16	0.0615				
3A	4	656	–	2	164	16	0.0889	1	34	2	0.9444
4	11	1,916	–	1	218	22	0.0917				
5	13	1,910	–	1	202	18	0.0818	1	15	0	1.0000
6	9	1,798	–	1	112	4	0.0345	1	326	51	0.8647
6A	9	1,518	–	5	663	34	0.0488	2	208	22	0.9043
7	11	2,511	–	–	–	–	–				
8	4	792	–	–	–	–	–				
Total	68	12,077*	–	12	1,603	110	0.0642	5	583	75	0.8860

* Progeny size ranges from 34 to 459 with a mean of 177.6.

ferent genotypes. Since the crosses white × white and pink × pink are not involved in the measure of natural crossing, total outcrossing should be greater.

Two methods may be employed to estimate the total amount of outcrossing. A direct method, which is based on gene frequencies in the outcrossing fraction of the pollen pool (Fyfe and Bailey, 1951; Vasek, 1964), utilizes two sets of data: Dr, the proportion of dominants observed in the progenies of recessives; and Rh, the proportion of recessives observed in the progenies of heterozygotes. The theoretical expectations are respectively λp and $(1 + \lambda q - \lambda p)/4$ (Vasek, 1964), where λ is the frequency of outcrossing, p is the frequency of the dominant gene, and $1 - p = q$ is the frequency of the recessive gene. The observed frequencies are set equal to the expected frequencies,

$$Dr = \lambda p \quad \text{(equation I)},$$
$$Rh = \frac{1 + \lambda q - \lambda p}{4} \quad \text{(equation II)},$$

and the two equations are simultaneously solved for λ and p.

An indirect method is based on the zygotic frequencies expected according to Wright's equilibrium: $(p^2 + pqF) + (2pq - 2pqF) + (q^2 - pqF) = 1$. This method uti-

lizes two other sets of data: R, the frequency of recessives directly observed in the population, and $H/(1-R)$, the proportion of heterozygotes, determined by progeny tests, observed among the dominant phenotypes in the population. The theoretical expectations are respectively

$$q^2 + pqF \quad \text{and} \quad \frac{2pq - 2pqF}{1 - (q^2 + pqF)},$$

where F is the coefficient of inbreeding and relates to λ in that, at equilibrium,

$$F = \frac{1 - \lambda}{1 + \lambda} \quad \text{and} \quad \lambda = \frac{1 - F}{1 + F}$$

(Nei and Syakudo, 1958). Observed frequencies are set equal to expectations,

$$R = q^2 + pqF \quad \text{(equation III)},$$
$$\frac{H}{1 - R} = \frac{2pq - 2pqF}{1 - (q^2 + pqF)} \quad \text{(equation IV)},$$

and the two equations are simultaneously solved for λ and p.

The four equations utilize four independent sets of data. Heterogeneity χ^2 tests

TABLE 3. χ^2 *test for heterogeneity.*

Data set	χ^2	d.f.	Probability
Dr	5.60	3	0.10–0.20
Rh	10.47	5	0.05–0.10
R	12.21	13	about 0.40
H/(1 − R)	10.60	7	0.10–0.20

(Table 3) indicate that each of the four sets of data is homogeneous. Therefore, the population will be considered as a unit, without subdivision of the data sets. However, it may be worth noting that equations 1, 2, and 4 derive data only from subpopulation areas 3 to 8.

In applying the direct method to the present experiment,

$$0.8860 = \lambda p \quad \text{(equation I)},$$

$$0.0642 = \frac{1 + \lambda q - \lambda p}{4} \quad \text{(equation II)},$$

and the simultaneous solutions are $\lambda = 1.0288$, and $p = 0.8612$.

In applying the indirect method, but using only the data from subpopulation areas 3 to 8,

$$0.0045 = q^2 + pqF \quad \text{(equation III)},$$

$$0.15 = \frac{2pq - 2pqF}{1 - (q^2 + pqF)} \quad \text{(equation IV)},$$

and the simultaneous solutions are $\lambda = 1.0506$, and $p = 0.9208$. (For the entire population, $R = 0.0039$, $p = 0.9214$, and $\lambda = 1.0650$.)

The two methods used in this study are based upon different sets of data, and are also based upon different assumptions. The direct method assumes the same outcrossing frequency for both heterozygotes and recessives. Thus some error would be introduced if a pollinator preferentially visited pink rather than white flowers. The indirect method, on the other hand, assumes that there is no selection and that the population is in equilibrium. Some error, therefore, is likely if, for example, selection occurs against recessives. The very low number of recessives in the population suggests that selection against recessives probably does occur. A statistical error is another consequence of the low number of recessives in the population, leading to very wide confidence limits for the data set R. Because of these wide confidence limits, and because no estimate of the severity of the probable selection against recessives is available, the estimates of outcrossing from equation III are unsatisfactory.

Other statistical errors are possible from the properties of equation II, and especially equations III and IV, where small changes in gene frequency result in very large changes in outcrossing frequency. Thus, the most reliable single set of data is Dr, used in equation I.

Despite the possible errors, the two estimates of λ and p are in close agreement. In addition, four more estimates are possible involving simultaneous solution of equations I or II with equations III or IV. All six possible simultaneous solutions are listed in Table 4 for convenient comparison. Considering the errors, the solutions give remarkably similar values for λ and p, the estimates of p ranging from 0.86 to 0.96, and the estimates of λ ranging from 0.80 to 1.05.

Perhaps the best estimates of λ and p are the averages of the individual estimates. On this basis, $\lambda = 0.9449$ and $p = 0.9209$. If the estimates based on equation III are eliminated, and justifiably so in view of the above discussion, the averages of the three remaining solutions are $\lambda = 0.9639$ and $p = 0.8976$. These values for λ and p are entirely reasonable since the most reliable data, $Dr = 0.886 = \lambda p$, establish that both λ and p are between 0.886 and 1.00. Furthermore, the presence of approximately 15 per cent heterozygotes in the population establishes that q probably is at least 0.07 and, conversely, that p is likely not to exceed 0.93. Therefore, it may reasonably be concluded that p is about 0.90 and that λ is about 0.96.

TABLE 4. *Outcrossing estimates.*

Simultaneous equations	p	λ
1–2	0.8612	1.0288
3–4	0.9208	1.0506
1–4	0.9174	0.9658
2–4	0.9143	0.8970
1–3	0.9471	0.9355
2–3	0.9645	0.8000
Average, all six solutions	0.9209	0.9449
Average, excluding equation three	0.8976	0.9639

The estimated frequency of outcrossing is not significantly different from 1.00, indicating that essentially all the ovules are cross-fertilized and this high frequency of outcrossing correlates with expectations based on the long exserted style and marked protandry.

By comparison, in a population of *C. exilis* (Vasek, 1964), outcrossing occurred with a frequency of 0.445, which is much lower than the frequency in *C. unguiculata*, but is more than would be expected on the basis of flower structure. Therefore, the frequency of cross-fertilization differs markedly between these two species and correlates satisfactorily with structural features of the breeding systems.

SUMMARY

A direct method of estimating outcrossing is based on the proportion of dominant offspring in progenies of wild recessives and the proportion of recessive offspring in progenies of wild heterozygotes. Observations were set equal to expectations and the two equations were solved simultaneously for $\lambda = 1.029$, the outcrossing frequency, and $p = 0.861$, the dominant gene frequency.

An indirect method of estimating outcrossing is based on the proportion of heterozygotes among dominant phenotypes, and on the proportion of recessives, expected in the population according to Wright's equilibrium. Observations were set equal to expectations and the two equations were solved simultaneously for $\lambda = 1.051$ and $p = 0.921$.

The four equations can be solved in four

other simultaneous combinations. However, the data for the frequency of recessives in the population was considered unsatisfactory, because the very small numbers of plants gave very wide confidence limits, and estimates based on this data were excluded. An average of the three remaining simultaneous solutions gave estimates of $\lambda = 0.96$ and $p = 0.90$.

The high frequency of outcrossing correlates with expectations based on the long exserted style and marked protandry.

LITERATURE CITED

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

NEI, M., AND K. SYAKUDO. 1958. The estimation of outcrossing in natural populations. Japanese J. Genetics, **33**: 46–51.

RASMUSON, H. 1920. Die Hauptergebnisse von einigen genetischen Versuchen mit verschiedenen Formen von *Tropaeolum, Clarkia* und *Impatiens*. Hereditas, **1**: 270–276.

VASEK, F. C. 1964. Outcrossing in natural populations. I. The Breckinridge Mountain population of *Clarkia exilis*. Evolution, **18**: 213–218.