THE POPULATION GENETICS OF *ARABIDOPSIS THALIANA*

I. THE BREEDING SYSTEM

M. E. JONES
Department of Genetics, University of Birmingham, England

Received 10.vii.70

1. Introduction

Arabidopsis thaliana is widely regarded as a completely inbreeding species (Müller, 1961; Rédei, 1962; Napp-Zinn, 1964; Laibach, 1965; Pederson, 1968), a conclusion which is largely based on the morphology of the flowers which is thought to ensure self-pollination. Müller (1961) reports, furthermore, that inbreeding does not result in any obvious inbreeding depression. If *Arabidopsis* is, as these observations suggest, autogamous, then the genetic structure of natural populations of this species should take one of two simple forms. Thus, either the population is, mutation apart, genetically homogeneous, *i.e.* all individuals are of the same homozygous genotype, or it comprises a heterogeneous set of families each of which is homozygous.

There is little evidence from work on natural populations of other species of structures of either of these types. Allard, Jain and Workman (1968), in a review of the genetics of predominantly inbreeding populations, mention no example of a species with a population structure of the first type, and there appears to be only one authenticated case of a species whose natural populations have a structure of the second kind (*Festuca microstachys*; Kannenberg and Allard, 1967; Allard and Kannenberg, 1968). Otherwise, detailed studies on species previously believed to be obligate inbreeders have shown that some cross-fertilisation does occur (Vasek, 1964, 1967; Allard, 1965; Imam and Allard, 1965; Allard, Jain and Workman, 1968). Populations of this kind, unlike the two previously discussed, characteristically show genetic differences between individuals within a family as well as genetic differences between families.

The results of Napp-Zinn (1964) suggest that *Arabidopsis* populations are in fact of this type. He found that the variation shown by families from natural populations is greater than that of inbred lines. This is an outcome which is most simply explained on the assumption that the former are partially heterozygous because outcrossing takes place in natural populations of this species.

Thus the breeding behaviour of natural populations of *Arabidopsis* is open to question and it was chiefly for this reason that the present work on the structure of natural populations of this species was begun.

2. Material and methods

The evidence to be considered concerns three separate experiments.

(a) *Experiment 1*

Samples of seed were collected from five mature plants selected at random

from each of six populations (table 1). These were chosen so as to include two distinct types of habitat in which the species is commonly found, namely, gardens and disused railway tracks (table 1). Five seeds from each of the 30 plants sampled were sown in each of two independently and completely randomised blocks making a total of 300 plants in all. In order to break any residual dormancy requirement the seed was allowed to imbibe water at room temperature followed by four days' incubation at $1 \pm 1°$ C. It was then sown on a bench of soil in the glasshouse on a grid spacing of 5 cm. each way following the technique reported by Lawrence (1968). As the experiment was sown in the autumn, supplementary light was provided. Mercury-vapour lamps were used to provide a daylength of 18 hours and a light intensity at the plant level of approximately 4465 lm/m².

TABLE 1

Locations and habitats of the populations used in experiments 1 *and* 2

Population	Location	Habitat
1	Alcester, Warwickshire	Disused railway tracks
2	Broom, Warwickshire	
3	Luddington, Warwickshire	
4	Henley in Arden, Warwickshire	
5	Cannon Hill Park, Birmingham	Flower beds
6	Parks department nurseries, Ruislip, Middlesex	Frames

Twenty-five seeds of the inbred line Laibach were also sown in each block to serve as a control with which the variance of the natural progenies could be compared. The total number of plants in the experiment was thus 350.

(b) *Experiment 2*

The seed used in this experiment was obtained by the self-pollination of five randomly selected plants from each of the five families in the two garden populations (*i.e.* Cannon Hill and Ruislip) of the previous experiment. The plants were enclosed in cellophane bags and allowed to self-pollinate. The seed was pretreated as before and used to raise five plants per family in each of two independently and completely randomised blocks, giving a total of 500 plants in all. The other details of this experiment are similar to the previous one, except that no supplementary light was provided as the experiment was sown in the seond week in May. Also no inbred material was included in this experiment.

(c) *Experiment 3*

The material used in this experiment was obtained from 12 plants sampled at random from a disused railway track at Rubery, Worcestershire. After they had been allowed to set some seed in their natural habitat the plants were transported back to the laboratory where their later flowers were artificially self-pollinated. Thus each parent plant provided two samples of seed; one, the seed from open-pollinated flowers, and the other, the seed from self-pollinated flowers. The experiment is therefore concerned with 12 pairs of families, giving a total of 24 families in all.

Before sowing, the seed was sterilised for 10 minutes in a 1 : 1 mixture of

hydrogen peroxide (20 vols.) and absolute alcohol, and allowed to imbibe water at room tempreature for 12 hours. The material was then vernalised by incubating it in petri dishes on moist filter paper for 5 weeks in the dark at $1 \pm 1°$ C. It was then sown as before on a bench of soil in the glasshouse. Ten seeds per family were sown in each of two independently and completely randomised blocks, giving a total of 480 plants in the experiment. Supplementary light was provided as in the first experiment, for this experiment was also sown in the autumn.

Although the plants in all three experiments were scored with respect to other quantitative characters, the results for flowering time only are considered here. The time of flowering was measured in days from the day on which the first flower opened.

2. Results

(a) *Experiment 1*

The first plant to flower, 25 days after the experiment was sown, belonged to the Laibach inbred line. The first plant from a natural source did so 5 days later, and flowering continued virtually continuously for a further 104 days, when the experiment was terminated. At this time, however, not all

Table 2

Experiment 1: Analysis of variance on data pooled over blocks

Item	d.f.	M.S.
Populations	5	21,028·497***
Families within populations	24	2,350·743***
Within families	214	86·865

*** $P < 0.001$.

plants had been scored with respect to this character, either because their flowering was so abnormal as to render it unscorable, or because they showed no signs of coming into flower at all. These abnormal individuals and non-flowering individuals were however restricted to three populations, namely, Alcester, Broom and Luddington (table 3), which raises the possibility that some individuals from these populations require a period of low temperature before they will flower

From a preliminary analysis of variance it was clear that families behave consistently over blocks. Thus all further analyses are presented on the data pooled over blocks.

The overall analysis of variance is shown in table 2. As expected from the diverse origins of the populations, they differ considerably. Furthermore, these differences show good correlation with their original habitats (table 3). Thus populations 5 and 6, which are essentially weed populations, flower much earlier than the other four populations, which come from disused railway tracks. The variation between families within populations follows a similar pattern, for although families within populations are highly heterogeneous (table 2), this can be attributed mainly to differences between the families in populations 1-4 (table 3). In contrast, the families in populations 5 and 6 are relatively homogeneous.

So far the discussion of the results has been limited to the overall distribution of variation between and within populations which provides no evidence

on the nature of the breeding system. The variation which is of particular interest in relation to the breeding system is that shown between individuals within families. The within-family variances are in fact highly heterogeneous ($P < 0{\cdot}001$) although populations differ considerably in this respect.

TABLE 3

Experiment 1: Family means and varainces and populations means. The mean and variance of the inbred line Laibach is also shown at the end of the table

Pop.	Fam.	N	Mean	Variance	Population Mean
1	1	10	6·70	2·456	39·029
	2	8	47·13	54·125	
	3	10	42·20	39·511	
	4	4(5)	75·25	20·917	
	5	2(8)	80·00	2·000	
2	1	8	47·25	18·500	61·324
	2	10	48·30	414·233	
	3	8(2)	78·37	145·982	
	4	6(4)	73·00	864·000	
	5	2(6)	79·50	24·500	
3	1	3(6)	91·67	165·334	54·688
	2	10	25·20	57·733	
	3	10	39·90	63·433	
	4	7(2)	89·14	143·810	
	5	2(1)	99·50	40·500	
4	1	10	17·80	7·733	21·574
	2	10	34·40	528·711	
	3	8	21·37	130·554	
	4	9	16·33	8·750	
	5	10	17·40	9·600	
5	1	10	7·10	2·322	7·580
	2	10	7·70	3·344	
	3	10	8·00	6·667	
	4	10	7·30	4·233	
	5	10	7·80	5·511	
6	1	10	9·20	3·067	8·574
	2	9	9·56	2·778	
	3	8	8·00	1·429	
	4	10	8·90	4·100	
	5	10	7·20	5·067	
1 Laibach		49	2·33	3·570	

The number of non-flowering individuals are shown in brackets.

As before, the populations fall into two distinct groups, namely, populations 1-4 in which the within-family variances are large and heterogeneous (figs. 1-4), and populations 5 and 6 in which the within-family variances are small and relatively homogeneous (table 3 and fig. 5). The heterogeneity of the variances in populations 1-4 could be due to either or both of the following:

1. A scalar effect such that family variance increases with the mean.
2. The result of segregation, such that small within family variances would suggest the parent plant was homozygous and inbreeding.

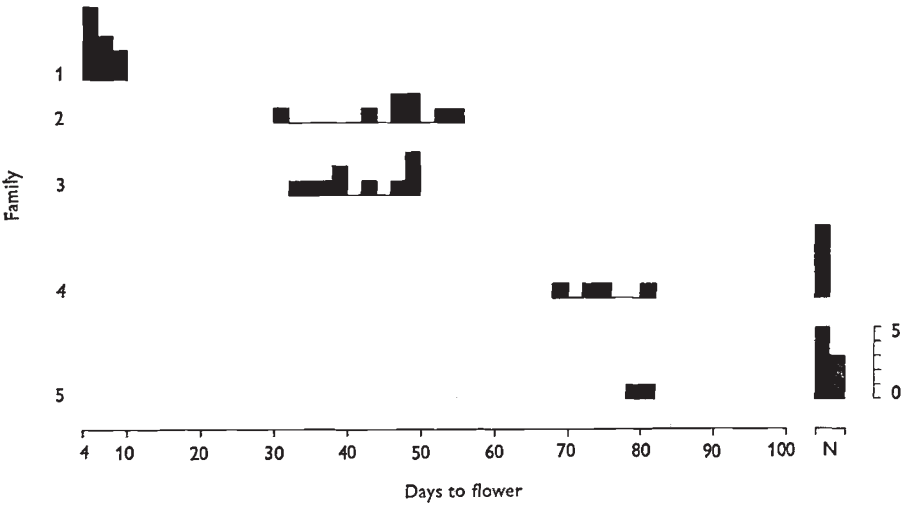

FIG. 1.—Population 1 (Alcester).

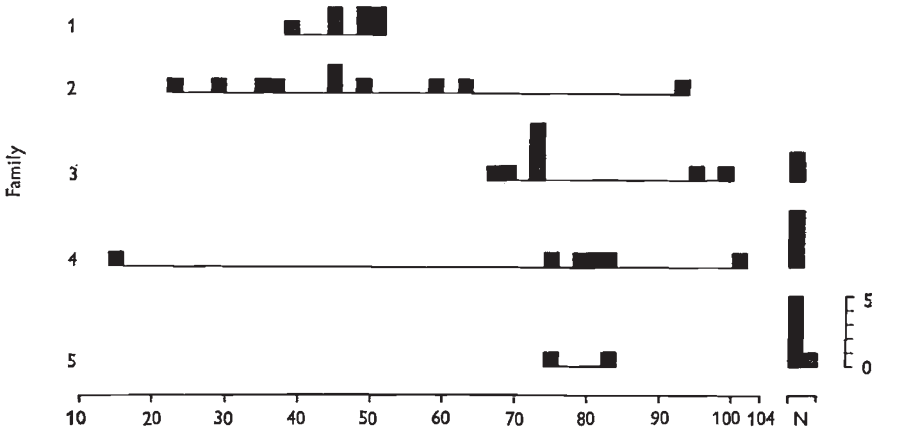

Fig. 2.—Population 2 (Broom).

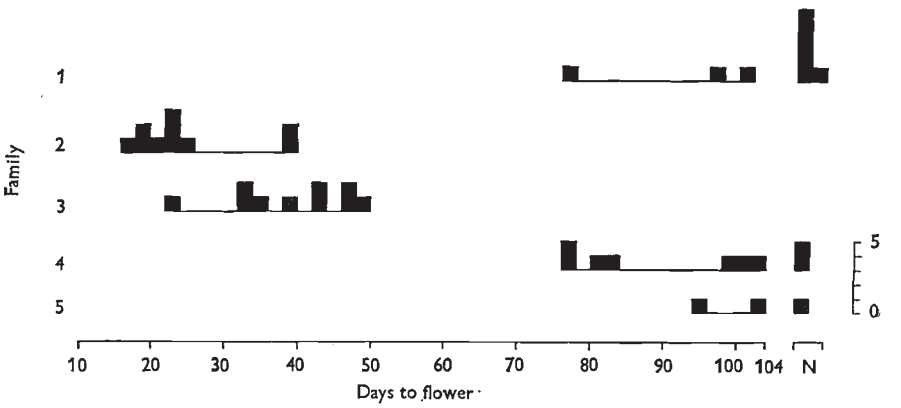

FIG. 3.—Population 3 (Luddington)

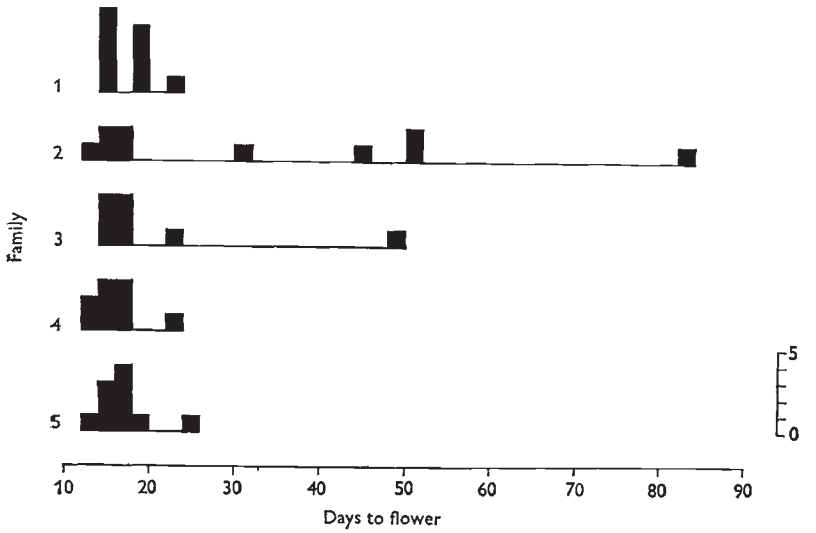

FIG. 4.—Population 4 (Henley in Arden).

FIG. 5.—Populations 5 (Cannon Hill) and 6 (Ruislip).

A scalar effect of the environment cannot fully account for this situation, for two reasons. Firstly, the family means and variances are clearly not completely correlated, for several families have similar means and widely differing variances (fig. 6). Secondly, it cannot account for the magnitude of some of the variances, for instance families 2 and 4 in population 2 contain individuals which flowered over a period of 70 days (fig. 2). There appears to be little doubt therefore that at least some individuals in populations 1, 2, 3 and 4 have arisen from outbred parents, *i.e.* that on this evidence *Arabidopsis* is at least a partially outbreeding species.

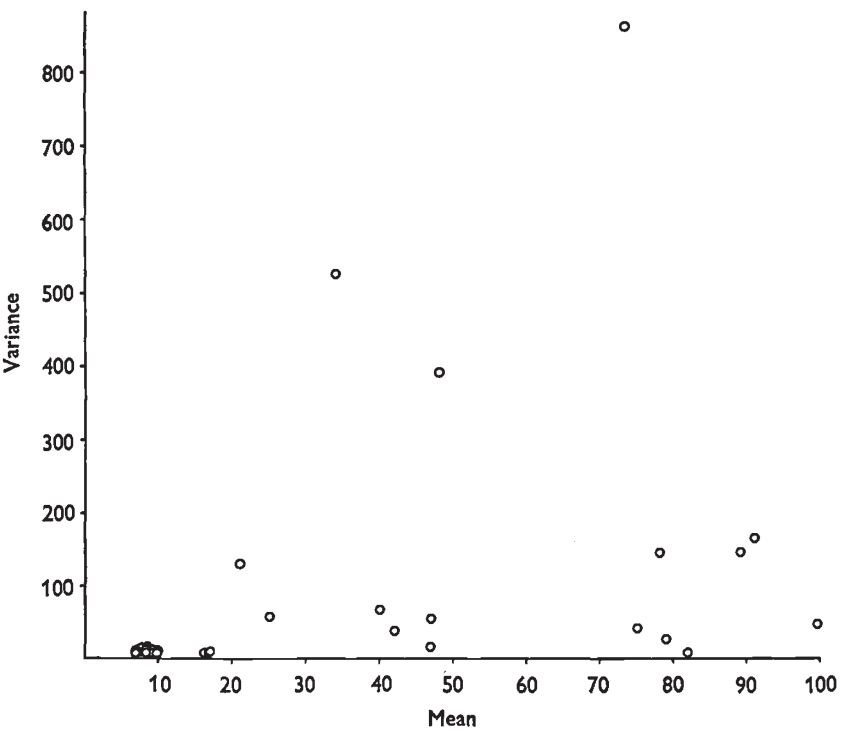

FIG. 6.—Experiment 1: Family mean plotted against family variance.

In contrast, the variation shown within families in populations 5 and 6 is small and similar to that of the inbred line (table 3). However, if cross-pollination does occur in these populations, and segregation would be difficult to detect by this method of investigation because of the narrow range of variation. With this in mind, further investigations on populations 5 and 6 were initiated using the progeny of the first selfed generation.

This extension of the experiment allows any variation between individuals within families of the first experiment to be detected, for it can now be tested for its significance against a further rank of statistics, namely the average variation between the progeny of each member of the family selfed. For clarity, the progeny resulting from the self-pollination of an individual in the previous experiment is defined as a family, and families with a common grandmother, *i.e.* the original wild parent, as a family group.

(b) *Experiment 2*

The first plant flowered 25 days after sowing and, as expected, all the individuals came into flower within the scoring period of the experiment. Again, families were found to behave consistently over blocks and thus all further analyses are presented on the data pooled over blocks.

TABLE 4

Experiment 2: Hierarchical analysis of variance on data pooled over blocks

Item	d.f.	M.S.
Populations	1	1366·731***
Groups within populations	8	9·115
Families within groups	35	7·219*
Within families	335	4·867

*** $P<0{\cdot}001$; * $P = 0{\cdot}05-0{\cdot}01$.

As expected from earlier results, the analysis of variance shows that most of the variation falls to a comparison between populations (table 4). However, the significance of the families within groups item shows that on average the individuals originating from a single wild parent differ, thereby suggesting that families are segregating in these populations. However, if each family group is examined individually (table 5) it turns out that the significance of

TABLE 5

Experiment 2: Analysis of between family mean squares for each family group separately

		Between families		Between individuals within families		
Population	Family group	M.S.	d.f.	M.S.	d.f.	P
5	1	9·162	3	8·564	27	
	2	7·874	3	15·477	21	
	3	38·383	3	6·229	29	***
	4	7·581	4	6·902	35	
	5	5·847	2	3·304	14	
6	1	1·266	4	4·664	39	
	2	1·918	4	2·868	42	
	3	1·972	4	2·129	41	
	4	3·150	4	2·395	44	
	5	4·151	4	2·578	43	

*** $P<0{\cdot}001$.

the between families within groups item is attributable solely to family 3 in population 5. Thus it is evident that one family is segregating in population 5 and it can therefore be concluded that some outbreeding has occurred in this population. In population 6, however, no outbreeding is detectable by this method of investigation.

(c) *Experiment 3*

In the previous two experiments the method by which outbreeding was

detected was by a comparison of the variation within families *inter se* or with the variation shown by a known inbred line. An alternative method is to compare the mean and variance of the open and self-pollinated progenies of the same plant. If self-pollination occurs, then this method has the advantage of a comparison of material of identical gene content. However, if cross-pollination occurs, then in most instances the mean and variance of the open-pollinated progeny will be significantly different from those of the self-pollinated progeny, thus providing a means of discrimination.

The first plant flowered 28 days after sowing and scoring was continued for a further 78 days, by which time all but a few plants had come into flower. As before, families behaved consistently over blocks and thus the results are presented on the data pooled over blocks.

In the analysis of variance (table 6) one of the most important sources of

TABLE 6

Experiment 3: Analysis of variance on data pooled over blocks

Item	d.f.	M.S.
Pollinations	1	942·261
Families	11	311·603***
F × P	11	337·439***
Within	348	10·855

*** $P < 0{\cdot}001$.

variation is that relating to differences between the means of the open and self-pollinated progeny in different families. Thus it is evident that some outbreeding has occurred in this population, although the results are clearly not consistent over families.

In order to examine these differences more closely the means and variances of the two types of progeny were calculated for each family separately and tested for heterogeneity (table 7). The families clearly fall into three main groups namely:

1. Families 1, 2, 3, 8 and 9 in which the means and variances are relatively homogeneous.
2. Families 6 and 11 in which the means and variances of self-pollinated progeny are significantly greater than those of the open-pollinated progeny.
3. Families 4, 5, 7, 10 and 12 in which the means and variances of the open-pollinated progeny are significantly greater than those of the self-pollinated progeny.

Thus it can be concluded with some certainty that in 7 out of the 12 families under investigation some outcrossing has occurred. It is clearly not possible with the present information to determine whether the open-pollinated progeny of the remaining 5 families have originated from self or cross-pollination. However, it is worth noting that, although the means and variances of families 1 and 3 are homogeneous, the open-pollinated progeny differ from the self-pollinated progeny in that they contain one non-flowering individual, thereby suggesting that some cross-pollination has also occurred in these families.

4. Discussion

The distribution of variation within a population of a sexually reproducing plant such as *Arabidopsis* is largely dependent on its breeding system. On the balance of evidence before this investigation was begun, it was expected that *Arabidopsis* populations would show a distribution of variation typical

Table 7

Experiment 3: Results of t *tests and variance ratio tests for the heterogeneity of the means and variances of the open (1) versus the self-pollinated (2) progeny of individual families*

Family	Poll.	Mean	d.f.	P	Variance	d.f.	P
1	1	14·35	25(1)	NS	241·961	18	NS
	2	13·00			116·500	(1)7	
2	1	31·00	22	NS	205·444	18	NS
	2	32·00			45·335	4	
3	1	11·84	23(1)	NS	46·499	17	NS
	2	14·25			32·645	(1)6	
4	1	25·60	35	***	109·655	18	**
	2	12·16			22·460	17	
5	1	18·65	36	***	64·250	18	***
	2	7·75			5·850	18	
6	1	9·58	33	*	14·421	17	**
	2	14·94			77·778	16	
7	1	25·31	31(1)	***	95·741	(1)14	***
	2	4·00			1·919	17	
8	1	16·74	26	NS	53·465	17	NS
	2	15·09			86·948	9	
9	1	25·63	33	NS	59·127	17	*
	2	21·78			148·548	16	
10	1	23·45	34	***	69·806	18	**
	2	11·44			15·111	16	
11	1	5·68	33	***	2·994	17	***
	2	13·78			48·498	16	
12	1	67·00	17(11)	***	150·667	(11) 3	***
	2	9·69			14·991	14	

The number of non-flowering individuals are shown in brackets.
*** $P < 0{\cdot}001$; ** $P = 0{\cdot}01$-$0{\cdot}001$; * $P = 0{\cdot}05$-$0{\cdot}01$.

of an inbreeding species. The results obtained, however, are obviously not compatible with those expected of inbreeding populations, and thus it is clear that some outbreeding occurs in natural populations of this species.

The amount of outcrossing which occurs in a population will depend on several factors, the most important of which are the morphology of the flowers and pollination agents. Part of the evidence put forward in favour of obligate inbreeding in this species is based on the morphology of the flowers which is thought to ensure self-pollination. However, observations on individuals growing both in the glasshouse and under natural conditions

Plate I

Flower head showing stages in flower development.

1. Immature flower bud.
2. Protogynous stage.
3. Autogamous stage.

FIGS. 1-5.—Experiment 1: Histograms showing the distribution of flowering time in each of the six populations. The data has been pooled over 2-day intervals. The non-flowering individuals (N) are shown separately.

have revealed a series of developmental stages which show a capacity for both self and cross-pollination (plate I). These stages are, in order of maturity:

1. Protogynous stage: the anthers are immature and the stigma protrudes from the flower.
2. Autogamous stage: the anthers grow up and dehisce at level of the stigma.

The stigma is known to be receptive at the protogynous stage as an earlier stage than this is used during artificial fertilisation. It seems clear, therefore, that there is some opportunity for cross-fertilisation during flower development.

There is little information on the way in which pollen is transferred from plant to plant under natural conditions. The *Flora of the British Isles* (Clapham, Tutin and Warburg, 1962) states that the plant is visited by several small insects. Furthermore, hoverflies have been seen visiting plants in the glasshouse. It is also possible that some cross-pollination may occur by contact as the plants are often found growing at close proximity in the wild.

Clearly more observations are required on wild populations *in situ* recording insects which visit the flowers. An experimental approach which could provide useful information is to surround a plant heterozygous for a recessive mutant with normal individuals and assess both the amount of selfing in the heterozygote and spread of mutant genes to the surrounding individuals. Varying the density of the surrounding plants would give information on the effect of proximity and the use of insect proof cages, on the importance of insects for cross-pollination.

5. Summary

1. Three experiments are described which were designed to investigate the nature of the breeding system in seven natural populations of *Arabidopsis thaliana.*

2. The wide range of variation shown by the progeny of individual plants sampled in the wild suggests that some outbreeding occurred in all but two of the populations under investigation.

3. An experiment which involved selfing the progeny of a random sample of plants from the latter further suggested that one of these populations also practised some outbreeding.

4. In a more detailed analysis of a seventh population, in at least 7 out of the 12 families under investigation, the naturally pollinated seed had resulted from cross-pollination.

5. Contrary to expectation, it appears that natural populations of this species are not uniformly inbreeding. This is discussed with reference to the factors which affect the frequency of outcrossing, namely the morphology of the flowers and pollination agents.

Acknowledgments.—I would like to thank both Dr M. J. Lawrence and Professor J. L. Jinks for their help and guidance throughout this work. A Science Research Council Studentship is gratefully acknowledged.

6. References

ALLARD, R. W. 1965. Genetic systems associated with colonizing ability in predominantly self-pollinated species. In *The Genetics of Colonizing Species*, Ed. Baker, H. G., and Stebbins, G. L., pp. 49-75. Academic Press, New York and London.

ALLARD, R. W., JAIN, S. K., AND WORKMAN, P. L. 1968. The Genetics of inbreeding populations. *Advances in Genetics, 14*, 55-131.

ALLARD, R. W., AND KANNENBERG, L. W. 1968. Population studies in predominantly self-pollinated species. XI. Genetic divergence among the members of the *Festuca microstachys* complex. *Evolution, 22*, 517-528.

CLAPHAM, A. R., TUTIN, T. G., AND WARBURG, E. F. 1962. *Flora of the British Isles*. Cambridge University Press.

IMAM, A. C., AND ALLARD, R. W. 1965. Population studies in predominantly self-pollinated species. VI. Genetic variability between and within natural populations of wild oats from differing habitats in California. *Genetics, 51*, 49-62.

KANNENBERG, L. W., AND ALLARD, R. W. 1967. Population studies in predominantly self-pollinated species. VIII. Genetic variability in the *Festuca microstachys* complex. *Evolution, 21*, 227-240.

LAIBACH, F. 1965. 60 Jahre *Arabidopsis*—Forschung, 1905-1965. " Arabidopsis Research ", *Rep. Int. Symp. Göttingen*, 14-18.

LAWRENCE, C. W. 1968. Radiation-induced polygenic mutation in *Arabidopsis thaliana*. 1. Selection for flowering time. *Heredity, 23*, 321-337.

MÜLLER, A. 1961. Zur Charakterisierung der Bluten und infloreszenzen von *Arabidopsis thaliana* (L.) *Heynh. Kulturpflanze, 9*, 364-393.

NAPP-ZINN, K. 1964. Uber genetische und entwicklungsphysiologische Grundlagen jahreszeitlicher Aspekte von Pflanzengesellschaften. In *Beitrage zur Phytologie* Festschrift Walter, H. Edited by Kreeb, K. 33-49.

PEDERSON, D. G. 1968. Environmental stress, heterozygote advantage and genotype-environment interaction in *Arabidopsis*. *Heredity, 23*, 127-138.

RÉDEI, G. P. 1962. Supervital mutants of *Arabidopsis*. *Genetics, 47*, 443-460.

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

VASEK, F. C. 1967. Outcrossing in natural populations. III. The deer creek population of *Clarkia exilis*. *Evolution, 21*, 241-248.