Population genetic structure and outcrossing rate of *Arabidopsis thaliana* (L.) Heynh.

Richard J. Abbott and
Mioco F. Gomes*

Department of Biology and Preclinical Medicine, Sir Harold Mitchell Building, University of St. Andrews, St. Andrews, Fife KY16 9TH, U.K.

The population genetic structure of *Arabidopsis thaliana* was investigated in and among 16 British populations by means of an electrophoretic survey of nine enzyme systems. Seven out of 17 allozyme loci examined were polymorphic. One polymorphic locus was used as a marker locus for estimating outcrossing rates. No outcrossing was observed within any of seven populations surveyed and a joint estimate of the upper limit of outcrossing, $\hat{t}_u$, indicated that average outcrossing in the wild does not exceed 0·3 per cent. The genetic structure of *A. thaliana* as measured by Nei's gene diversity statistics, gene correlation F statistics, and estimates of multilocus genotypic diversity is concordant with that which is expected for a plant species which shows almost complete uniparental reproduction. Despite considerable genetic differentiation between populations, there was no evidence of a division between English and Scottish populations based on genetic distance.

INTRODUCTION

Arabidopsis thaliana (L.) Heynh. has long been regarded as a plant species supremely suited for genetic studies (Redei, 1975). This view has been confirmed in recent years since the species small genome size has made it particularly suitable as a tool in plant molecular genetics (Meyerowitz, 1987). It is rather surprising, therefore, that little is known about the population genetics of *A. thaliana* other than for quantitative traits (Jones, 1971a, b; Cetl, 1987). Despite the fact that allozymes are resolved fairly easily in the species and the genetics of several isozyme systems have been analysed (Cammaerts and Jacobs, 1973; Jacobs and Schwind, 1976; Cammaerts and Jacobs, 1983) there are no published accounts of the genetic structure of natural populations of *A. thaliana* based on allozyme diversity. Here we report the findings of a study of genetic variation within and between 16 British populations of *A. thaliana* based on an electrophoretic survey of nine enzyme systems. In addition, we provide estimates of outcrossing as measured at a single locus coding for peroxidase.

* Present address: Departamento Biologia Molecular, CCEN-Universidade Federal da Paraiba, 58.000 J. Pessoa PB, Brazil.

MATERIALS AND METHODS

Seed was collected separately from 30 individual plants from each of 11 populations of *A. thaliana* located in Derbyshire, England and five populations from Scotland (table 1). Plants were grown from seed in a greenhouse and three offspring per individual were assayed by electrophoresis to determine maternal genotype (Brown, 1975). In the very rare instances when a heterozygote was observed among the three offspring of an individual a further seven offspring of that individual were assayed to confirm maternal genotype. Horizontal electrophoresis was conducted in 12·5 per cent starch gels on enzyme extracts from rosette leaves of four-week-old plants. Extracts were made using an extraction buffer consisting of 0·1 M Tris-HCl pH 7·2 and 0·1 per cent (v/v) 2-mercaptoethanol. Electrophoresis was performed using standard procedures (Ferguson, 1981). Three different buffer systems were used: (*A*) Discontinuous lithium borate/Tris-citrate pH 8·3 (after Scandalios, 1969); (B) Discontinuous sodium borate/Tris-citrate pH 8·8 (Adams and Joly, 1980); (C) Continuous Tris-citrate pH 6·2 (Adams and Joly, 1980). For buffer systems A and C voltage was held constant at 250 V and 115 V respectively. For system B the current was

Table 1 Locations of British populations of *Arabidopsis thaliana*

Designation	Location	Grid ref.	Habitat
England			
E1*	Dimmins Dale, Derbyshire	SK169704	Limestone Pasture
E2	Hay Dale, Derbyshire	SK178724	Limestone Pasture
E3*	Hipley Hill, Derbyshire	SK208542	Limestone Pasture
E4	Lathkill Dale, Derbyshire	SK163661	Limestone Pasture
E5	Lathkill Dale, Derbyshire	SK160662	Limestone Pasture
E6	Lathkill Dale, Derbyshire	SK155665	Limestone Pasture
E7	Lathkill Dale, Derbyshire	SK183658	Limestone Outcrop
E8*	Monsall Head, Derbyshire	SK184719	Limestone Outcrop
E9*	Wetton Mill, Derbyshire	SK095562	Leisure Park
E10	Winnats Pass, Derbyshire	SK137827	Limestone Outcrop
E11	Matlock, Derbyshire	SK303609	Car Park
Scotland			
S1*	Glasgow, Strathclyde	NS613674	Garden
S2*	Glasgow, Strathclyde	NS552679	Waste Site
S3	Fortingall, Grampian	NN739471	Walls/Roadside
S4	Glen Lyon, Grampian	NN718479	Walls/Roadside
S5*	St. Andrews, Fife	NO510165	Garden

* Populations for which outcrossing rates were determined at the *PER* locus.

maintained at 35 mA. Electrophoresis was terminated after the front had moved 8 cm from the origin towards the anode. Gels were sliced twice, and each of the three slices were stained for a different enzyme system. Details of the enzyme systems stained together with the number of loci scored, and buffer systems used are given in table 2. After incubation, gels were rinsed several times with tap water before scoring. If required, gels were fixed in 50 per cent glycerol except for those stained for GOT which demanded immediate scoring.

For a given enzyme system the locus coding for the most anodally migrating protein was designated 1, the next 2 etc. Likewise at each locus the allele coding for the most anodally migrating allozyme (cathodally in the case of peroxidase) was designated a, the next b and so on. Resolution of ACP activity was equivalent to that reported by Jacobs and Schwind (1976). Variation occurred at

three clearly distinguished zones due to allelic segregation at three distinct ACP loci. Gels stained for EST were multi-banded; however, variation was confined to only two zones of activity. Grover (1975) has reported that variation within each of these zones is due to allelic variation at the *EST*-1 and *EST*-2 loci respectively. A genetic analysis of the fast and slow isozymes of *IDH*-1 and *PER*-1 revealed single locus control with codominant alleles at each locus. *IDH*-1 produced a three banded phenotype in the heterozygote and is therefore dimeric. Null alleles were present at the *ACP*-2, *EST*-1 and *EST*-2 loci. The frequencies of these alleles in populations were calculated directly from the frequencies of the respective null homozygotes. This method was considered accurate since heterozygotes for other "normal" alleles were absent in all bar one population (see below).

Table 2 Enzymes, number of loci, buffer systems and strain references

Enzyme		No. of loci scored	Buffer System	Stain reference
Acid phosphatase	(ACP)	3*	A	Gottlieb (1973)
Esterase	(EST)	2*	A	Scandalios (1969)
Glutamate dehydrogenase	(GDH)	2	C	Brewer and Singh (1970)
Glutamate-oxaloacetate transaminase	(GOT)	2	B	Adams and Joly (1980)
Isocitric dehydrogenase	(IDH)	1	C	Adams and Joly (1980)
Malate dehydrogenase	(MDH)	2	C	Adams and Joly (1980)
Peroxidase	(PER)	1*	A	Shaw and Prasad (1970)
Phosphoglucose isomerase	(PGI)	2	A	Gottlieb (1973)
Phosphoglucomutase	(PGM)	2	B	Adams and Joly (1980)

* Additional loci coding for these enzyme are likely to be present as extra bands showing variable expression were resolved on gels.

Table 3 Mean number of alleles per locus ($\bar{K}$), per cent of loci polymorphic (*PLP*), and mean observed heterozygosity ($\bar{H}_0$) in British populations of *A. thaliana*

Populations	No. of Populations	No. of loci	$\bar{K}$(SD)	PLP	$\bar{H}_0$
English	11	17	1·182 (0·474)	14·97	0·00000
Scottish	5	17	1·282 (0·648)	20·00	0·00078
Total	16	17	1·213 (0·535)	16·50	0·00024

N.B. A locus is treated as polymorphic when the frequency of the most common allele is $\leqq$99 per cent.

The electrophoretic survey showed that the *PER*-1 locus coding for two allozymes of peroxidase was polymorphic in most populations surveyed (see below). This locus was subsequently used as a marker locus for the estimation of outcrossing rate in seven populations (see table 1) in which the alleles of *PER*-1 were at relative intermediate frequencies. To this end seven more progeny from each of the 30 plants per populations were classified according to genotype at the *PER*-1 locus.

RESULTS

Of the 17 loci surveyed seven were polymorphic in at least one population. Allele frequencies at polymorphic loci are given in the Appendix. The other ten loci (*GDH*-1, *GDH*-2, *MDH*-1, *MDH*-2, *GOT*-1, *GOT*-2, *PGM*-1, *PGM*-2, *PGI*-1, *PGI*-2) were monomorphic in all populations surveyed.

The mean number of alleles per locus ($\bar{K}$) averaged over populations and all loci, including monomorphic loci (table 3), was very similar for both English and Scottish populations as were the mean values for the average percentage of loci that are polymorphic (*PLP*). Of the 16 populations examined, only two, E2 and E10, were invariant at all 17 allozyme loci screened. Despite the presence of allozyme variability in most populations, mean observed heterozygosity per individual ($\bar{H}_0$) was extremely low. Indeed, out of a total of 480 maternal genotypes tested, only two proved to be heterozygous at any locus. These genotypes were heterozygous at the *ACP*-3 locus and occurred in the S2 population only.

Gene diversity

To quantify the degree of gene differentiation in and among the 16 populations examined Nei's gene diversity statistics (Nei, 1973) were computed. The level of gene diversity within populations, H_S, was obtained for each locus in turn by

calculating the expected heterozygosity (assuming panmixia), H_e, in each population, given by $H_e = 1 - \sum_{i=1}^{k} x_i^2$ (where x_i is the *i*th of k alleles in a given population), and then averaging H_e over all populations. H_T, the gene diversity in the total population, was calculated as $H_T = 1 - \sum_{i=1}^{k} \bar{x}_i^2$ where $\bar{x}_i^2$ is the mean frequency of the *i*th of k alleles for the populations surveyed. Finally the gene diversity between populations, D_{ST} was calculated by subtracting H_S from H_T. From these statistics it was possible to compute two measures of relative gene diversity; these were G_{ST}—the gene diversity between populations relative to that of the total population—given by $G_{ST} = D_{ST}/H_{ST}$; and R_{ST}—the between population gene diversity relative to the within sub-population gene diversity—calculated as $R_{ST} = \bar{D}_m/H_S$. $\bar{D}_m$ is computed to exclude comparisons of populations with themselves and is given by $\bar{D}_m = sD_{ST}/(s-1)$ where s is the number of populations.

Gene diversity statistics for 17 loci in *A. thaliana* are presented in table 4. Averaging over all loci the total gene diversity, H_T, in *A. thaliana* is 0·148. Examination of the mean G_{ST} value shows that the greatest proportion of this diversity is attributable to between population diversity, D_{ST}. However, within population diversity, H_S, is never insignificant and for three loci—*IDH*-1, *PER*-1

Table 4 Measures of gene diversity in British populations of *Arabidopsis thaliana*

Locus	H_T	H_S	D_{ST}	G_{ST}	R_{ST}
ACP-1	0·4738	0·1330	0·3408	0·7193	2·7332
ACP-2	0·1729	0·0910	0·0819	0·4737	0·9600
ACP-3	0·1117	0·0481	0·0636	0·5694	1·3222
EST-1	0·6818	0·2150	0·4668	0·6847	2·3159
EST-2	0·5899	0·1737	0·4162	0·7055	2·5558
IDH-1	0·0523	0·0306	0·0217	0·4149	0·7564
PER-1	0·4387	0·2753	0·1634	0·3725	0·6331
Mean	0·1483	0·0569	0·0914	0·6163	1·7134

N.B. Means of H_T, H_S and D_{ST} are averages over all polymorphic and monomorphic loci (*i.e.*, 17 loci). Means for G_{ST} and R_{ST} are computed from the mean values for H_T, H_S and D_{ST}.

and ACP-2—H_S was greater than D_{ST}. Direct comparison of the amount of diversity between populations relative to within populations (R_{ST}) showed that there was approximately 1.7 times as much variation between populations as within them.

Using the allele frequencies computed at each locus (Appendix) it is possible to calculate a measure of genetic distance between each pair of populations in turn using Nei's genetic distance statistic, $\hat{D}$, (Nei, 1972). $\hat{D}$ can range from zero when populations show all alleles at the same frequency to ∞ when populations have no alleles in common. The average genetic distances between all populations, English populations and Scottish populations were respectively 0.110, 0.097 and 0.109. An unweighted pair group method cluster analysis of genetic distance between all pairs of populations (fig. 1) revealed no separation of English from Scottish populations and an absence, therefore, of any overall association between genetic distance and geographical distance. Further tests of association between genetic distance and geographical distance within each of the two geographical regions revealed no association in

Figure 1 Phenogram showing the relationships of 16 populations of *A. thaliana* generated by UPGMA clustering based on Nei's genetic distance $\hat{D}$.

Scotland ($r = 0.09$ ns) but a small yet significant association for English populations ($r = 0.39**$). The latter tended to reflect a very positive association between populations that are very near neighbours e.g., the Lathkill Dale populations E4–E7 ($r = 0.81$).

Population structure

An analysis of population structure was conducted by computing the parameters F, θ and f in the manner given by Weir and Cockerham (1984). F describes the correlation of genes within individuals in the total "species" population and is related to Wright's (1951) F_{IT}. θ describes the correlation of genes *between* individuals within populations relative to the total "species" population, and is related to Wright's F_{ST}. θ is a measure of the amount of differentiation between populations. f measures the correlation of genes *within* individuals within populations. It is related to Wright's F_{IS}, and describes the deviation of observed heterozygosity in populations from that expected under panmixia.

Estimates of $\hat{F}$, $\hat{\theta}$ and $\hat{f}$ averaged over all loci were respectively 0.9979 (0.0009), 0.6354 (0.0030), and 0.9943 (0.0023). The values in brackets are the standard errors derived by the method of Miller (1974) as described in Weir and Cockerman (1984). The high values for $\hat{F}$ and $\hat{f}$ (close to 1) are expected in a species such as *A. thaliana* which shows very high levels of self-fertilisation (see below) and exhibits, therefore, very little heterozygosity. The value for $\hat{\theta}$, which is also high, indicates significant genic differentiation among populations.

Multilocus genotypic diversity

From the allelic variation resolved at the seven polymorphic loci it is possible to construct 97,200 theoretical multilocus genotypes. Among the sample of 470 plants examined, 45 different multilocus genotypes were uncovered (0·094 genotypes per individual). The distribution of the number of individuals per genotype (fig. 2) shows that 18 genotypes were represented by either one or two individuals out of the 480 surveyed. In contrast one genotype was present in 61 individuals. The number of genotypes per population ranged from 1–8 with a mean of 4·62. The distribution of genotypes per population was relatively even (fig. 3(a)) with only two populations composed of a single genotype. In contrast, the distribution of the number of populations per genotype was highly
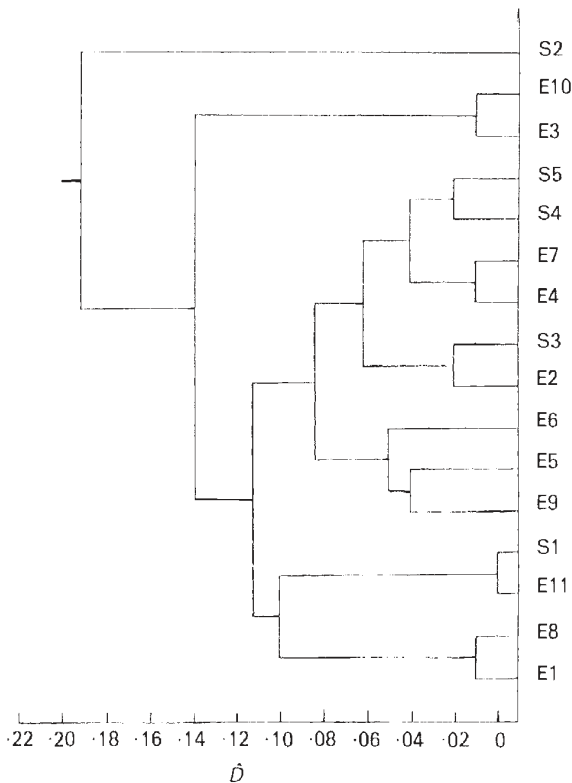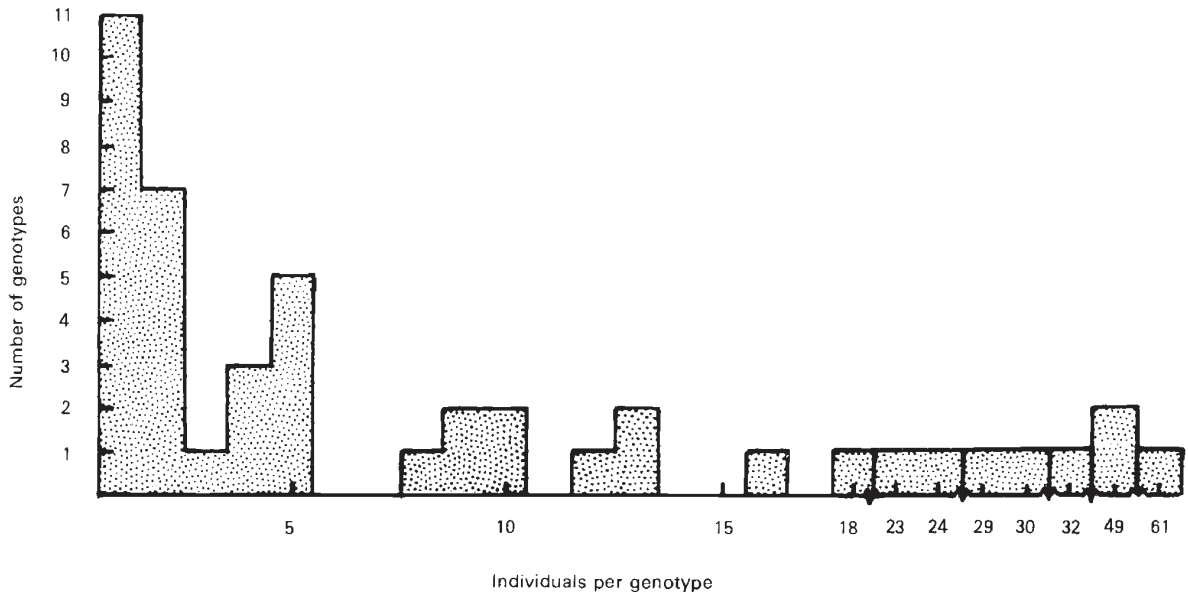
Figure 2 Distribution of number of individuals per genotype over the 45 multilocus genotypes detected.

skewed (fig. 3(b)) with the majority of genotypes (28) limited in distribution to a single population. One genotype, however, occurred in six of the 16 populations surveyed.

To obtain a measure of multilocus genotypic diversity in each population, Pielou's corrected version of the Gini Index, $\tilde{D}$, was calculated (Pielou, 1969). This index of heterogeneity takes into account both the number of genotypes present and the evenness with which they are distributed in a sample (Peet, 1974). $\tilde{D}$ may range from zero

(a uniform sample) to 1 (a completely heterogeneous sample). The values for the 16 populations surveyed are illustrated in fig. 3(c). They ranged from 0·00 to 0·83 with a mean of 0·55.

Outcrossing rates

Out of a total of 2100 progeny (seven populations × 30 maternal plants × 10 progeny) tested for genotype at the *PER*-1 locus, not one was found to be heterozygous. This means that the estimated
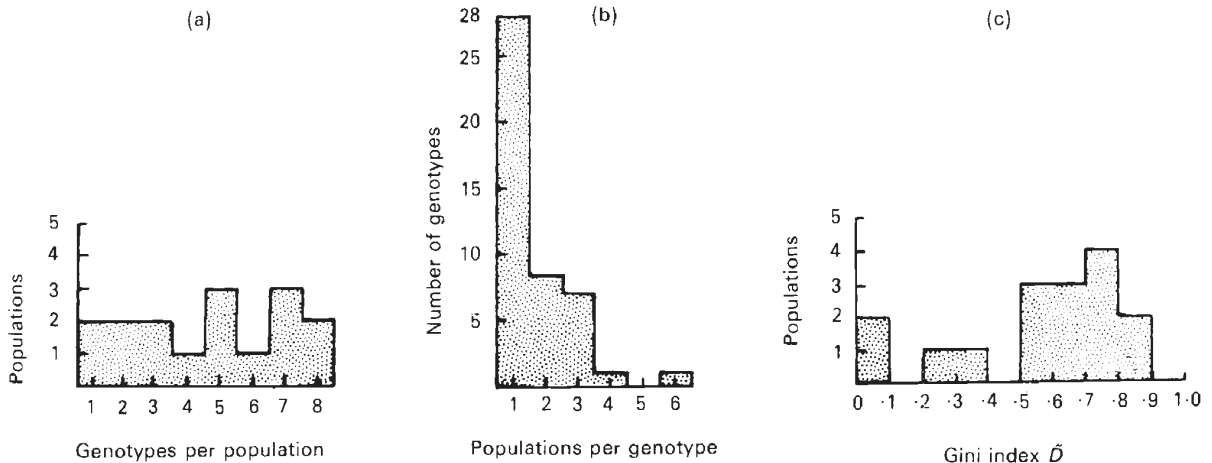
Figure 3 Distribution of (a) genotypes per population; (b) populations per genotype; and (c) corrected Gini Index $\tilde{D}$, over populations.

rate of outcrossing ($\hat{t}$) in each population equals zero. Although it is not possible to calculate interval estimates of $\hat{t}$, as the variance of t is also zero, a 95 per cent upper limit of $\hat{t}$, $\hat{t}_u$, can be calculated for each population following a procedure described by Adams and Allard (1982). Using this procedure the upper limit of outcrossing rate, $\hat{t}_u$, over the seven populations was found to range from 2·00–4·32 per cent. However, because not a single outcross was detected in any of the seven populations examined, actual outcrossing (t) in natural populations may be considerably lower than is indicated by $\hat{t}_u$. Following Adams and Allard (1982), a joint $\hat{t}_u$ was calculated from the combined data based on the assumption that t is at a uniformly low value in each of the seven populations surveyed. The joint value of $\hat{t}_u$ equals 0·285 per cent.

DISCUSSION

It is clear from this study that outcrossing in natural populations of *A. thaliana* is a very rare phenomenon. Previous estimates of outcrossing rate in the species have been based on studies of artificial populations constructed from mutant lines. Studies by Snape and Lawrence (1971), and Röbbelen (1971) showed outcrossing to range from 1–2 per cent. However, under conditions of high pollinator activity and/or in mutant lines showing varying degrees of male sterility, rates of outcrossing have been known to rise to 6–26 per cent (Cetl and Plchova, 1975; Trněná *et al.*, 1987; Drescher and Kranz, 1987). From the results presented here, outcrossing in natural populations is, in fact, much rarer than indicated previously. According to the estimated joint upper limit of outcrossing, $\hat{t}_u$,

average outcrossing is normally below 0·3 per cent. Indeed, the actual outcrossing values recorded in populations suggest that often there is complete selfing. However, direct evidence that some outcrossing occurs occasionally in British populations comes from the finding of two heterozygous individuals (at the *ACP*-3 locus) among 30 individuals sampled from the Scottish population S2. Taken overall, the pattern and level of outcrossing found in *A. thaliana* is very similar to that reported previously for the annual grass *Festuca microstachys* Nutt. (Adams and Allard, 1982). In *F. microstachys*, Adams and Allard found that although recorded outcrossing was normally zero in most populations, outcrossing could rise in a population to approximately 7 per cent in a particular year only to return to zero the following year.

Given the exceptionally low rate of outcrossing exhibited by *A. thaliana*, it is of value to compare the levels of genetic variation found in *A. thaliana* with that found in other predominantly selfing species. Comparisons with other selfers for measures of gene diversity (table 5) show that in most respects the level and partitioning of diversity in and among populations of *A. thaliana* is representative of a typical selfing species. That said, some differences are apparent. Most notably, *A. thaliana* contains more alleles per polymorphic locus than is normal and, in addition, the observed level of heterozygosity ($\bar{H}_0$) is only a quarter of that commonly found in other selfers, presumably because of the exceptionally low outcrossing rate of the species.

Estimates for the F statistics, $\hat{F}$, $\hat{\theta}$ and $\hat{f}$, are again as expected for a species that exhibits virtually no outcrossing. In a recent study of the genetic structure of *Impatiens capensis* (Knight and Waller, 1987) values of F, θ and f based

Table 5 Measures of gene diversity in *Arabidopsis thaliana* as compared with means for predominantly selfing species. Standard errors of the selfer's means are given in brackets

	No. of populations*	No. of enzymes*	No. of allozyme loci*	$\bar{K}$*	PLP*	$\bar{H}_0$*	I*	H_T†	H_S†	G_{ST}†
Selfer's means	14·04 (3·62)	8·96 (0·66)	16·61 (1·05)	2·26 (0·10)	0·183 (0·032)	0·001 (0·0004)	0·975 (0·011)	0·291 (0·086)	0·128 (0·030)	0·523 (0·211)
A. thaliana	16	9	17	2·86	0·165	0·00024	0·897	0·360	0·138	0·563

I = Mean genetic identity of all pairs of populations.

* Values from Gottlieb (1981).

† Values from Loveless and Hamrick (1984) based on different average number of populations, enzymes and loci from those given by Gottlieb.

N.B. For the purpose of comparison, estimates of $\bar{K}$, H_T, H_S and G_{ST} are averages over polymorphic loci only.

on an electrophoretic survey of the progenies of cleistogamous flowers (enforced selfing), were 0.927, 0·482 and 0·858 respectively which are slightly lower than the values recorded for *A. thaliana* here. From the estimates recorded in the present study it is clear that in *A. thaliana* there is considerable local inbreeding ($\hat{f}$) due to selfing, considerable genetic differentiation between populations ($\hat{\theta}$) and, as a result of these two factors, fixation in the total population ($\hat{F}$) is almost complete.

Finally, in regard to measures of multilocus genotypic diversity in *A. thaliana*, the species is found to contain a very similar level of variation to that found by Lyman and Ellstrand (1984) in North American populations of the apomict Dandelion, *Taraxacum officinale*. In *T. officinale*, the mean number of multilocus genotypes per population and the average corrected Gini Index, $\tilde{D}$, (calculated for multiclonal populations only) were respectively 5 and 0·57 (see Ellstrand and Roose, 1987), whereas in *A. thaliana* these values were 4·6 and 0·62 respectively. A comparison is

drawn between these two species as in both the present study and that of Lyman and Ellstrand a similar number of characters, individuals and populations were examined.

To sum up, therefore, the amount and distribution of genetic variation in *A. thaliana* estimated in terms of gene diversity, levels of fixation (as measured by F statistics) and multilocus genotypic diversity is concordant with that which is expected for a species which shows almost complete uniparental reproduction. Whether this variation for allozyme characters in *A. thaliana* reflects equivalent diversity for quantitative traits remains to be established.

APPENDIX

Allele frequencies for seven polymorphic loci surveyed in sixteen British populations of *Arabidopsis thaliana*

Locus	Allele	Population															
		E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	S1	S2	S3	S4	S5
ACP-1	a	1·00	1·00	0·07	0·53	0·03	0·27	0·43	1·00	—	—	—	0·97	1·00	0·53	1·00	1·00
	b	—	—	0·93	0·47	0·97	0·73	0·57	—	1·00	1·00	1·00	0·03	—	0·47	—	—
ACP-2	a	—	—	0·03	—	—	—	—	—	—	—	—	0·03	0·43	—	—	—
	b	1·00	1·00	0·97	1·00	1·00	1·00	0·97	1·00	0·97	1·00	1·00	0·97	0·14	0·50	1·00	1·00
	c	—	—	—	—	—	—	—	—	0·03	—	—	—	0·43	0·33	—	—
	nl	—	—	—	—	—	—	0·03	—	—	—	—	—	—	0·17	—	—
ACP-3	a	—	—	0·07	—	—	—	—	—	—	—	—	—	0·44	—	—	—
	b	1·00	1·00	0·93	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	0·13	1·00	1·00	1·00
	c	—	—	—	—	—	—	—	—	—	—	—	—	0·43	—	—	—
EST-1	a	—	—	—	0·47	0·03	—	0·93	—	—	—	0·77	1·00	0·90	0·03	0·40	1·00
	b	0·66	1·00	—	0·47	0·67	0·73	0·07	1·00	—	—	0·07	—	—	0·97	0·34	—
	c	0·17	—	—	0·06	0·03	0·27	—	—	—	—	—	—	—	—	0·13	—
	nl	0·17	—	1·00	—	0·27	—	—	—	1·00	1·00	0·16	—	0·10	—	0·13	—
EST-2	a	0·77	—	1·00	0·47	0·03	—	0·57	1·00	—	1·00	0·17	—	0·43	0·03	0·07	—
	b	0·23	1·00	—	0·53	0·90	1·00	0·43	—	1·00	—	—	—	0·57	0·97	0·73	1·00
	nl	—	—	—	—	0·07	—	—	—	—	—	0·83	1·00	—	—	0·20	—
IDH-1	a	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	0·57	1·00	1·00	1·00
	b	—	—	—	—	—	—	—	—	—	—	—	—	0·43	—	—	—
PER-1	a	0·53	1·00	0·67	0·80	1·00	0·30	0·77	0·20	0·63	1·00	0·43	0·43	0·20	0·97	1·00	0·87
	b	0·47	—	0·33	0·20	—	0·70	0·23	0·80	0·37	—	0·57	0·57	0·80	0·03	—	0·13

nl = null allele.

REFERENCES

ADAMS, W. T. AND ALLARD, R. W. 1982. Mating system variation in *Festuca microstachys*. *Evolution*, *36*, 591-595.

ADAMS, W. T. AND JOLY, R. J. 1980. Genetics of allozyme variants in loblolly pine. *J. Hered.*, *71*, 33-40.

BREWER, G. J. AND SINGH, C. F. 1970. *An Introduction to Isozyme Techniques*. Academic Press, New York.

BROWN, A. H. D. 1975. Optimum experimental designs for the estimation of genetic parameters in plant populations. *Biometrics*, *31*, 145-160.

CAMMAERTS, D AND JACOBS, M. 1975. Study of the intracellular location and the genetic control of malate dehydrogenase isozymes in *Arabidopsis thaliana*. *Plant Sci. Lett.*, *4*, 249-256.

CAMMAERTS, D. AND JACOBS, M. 1983. A study of the polymorphism and the genetic control of the glutamate dehydrogenase isozymes in *Arabidopsis thaliana*. *Plant Sci. Lett.*, *31*, 65-73.

CETL, I. 1987. Genetic analysis of *Arabidopsis thaliana* (L.) populations from Czechoslovakia. *Arabid. Inf. Serv.*, *25*, 67-84.

CETL, I. AND PLCHOVÁ, Z. 1975. *Arabidopsis thaliana* (L.) Heyn. as an object for studying population structure of predominantly self-pollinating plant species. *Arabid. Inf. Serv.*, *12*, 21-24.

DRESCHER, W AND KRANZ, A. R. 1987. First evidence for entomogamous hybridization in *Arabidopsis thaliana* (L.) Heyn. *Arabid. Inf. Serv.*, *23*, 41-45.

ELLSTRAND, N. C. AND ROOSE, M. L. 1987. Patterns of genotypic diversity in clonal plant species. *Amer. J. Bot.*, *74*, 123-131.

FERGUSON, A. 1980. *Biochemical Systematics and Evolution*. Blackie, Glasgow.

GROVER, N. S. 1975. Characterisation of *Arabidopsis thaliana* ecotypes on the basis of genetic variation at ten isozyme loci. *Arabid. Inf. Serv.*, *12*, 19-21.

GOTTLIEB, L. 1973. Genetic differentiation, sympatric speciation, and the origin of a diploid species of *Stephanomeria*. *Amer. J. Bot.*, *60*, 545-553.

GOTTLIEB, L. 1981. Electrophoretic evidence and plant populations. *Prog. Phytochem.*, *7*, 1-46.

JACOBS, M. AND SCHWIND, F. 1976. Biochemical genetics of acid phosphatase isozymes in *Arabidopsis thaliana* (L.) Heyn. *Arabid. Inf. Serv.*, *13*, 56-73.

JONES, M. E. 1971a. The population genetics of *Arabidopsis thaliana*. I. The breeding system. *Heredity*, *27*, 39-50.

JONES, M. E. 1971b. The population genetics of *Arabidopsis thaliana*. II. Population structure. *Heredity*, *27*, 51-58.

KNIGHT, S. E. AND WALLER, D. M. 1987. Genetic consequences of outcrossing in the cleistogamous annual, *Impatiens capensis*. I. Population genetic structure. *Evolution*, *41*, 969-978.

LOVELESS, M. D. AND HAMRICK, J. L. 1984. Ecological determinants of genetic structure in plant populations. *Ann. Rev. Ecol. Syst.*, *15*, 65-95.

LYMAN, J. C. AND ELLSTRAND, N. C. 1984. Clonal diversity in *Taraxacum officinale* (Compositae), an apomict. *Heredity*, *53*, 1-10.

MEYEROWITZ, E. M. 1987. *Arabidopsis thaliana*. *Ann. Rev. Genet.*, *21*, 93-111.

MILLER, R. G. 1974. The jacknife—a review. *Biometrika*, *61*, 1-15.

NEI, M. 1972. Genetic distance between populations, *Amer. Natur.*, *106*, 283-292.

NEI, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl Acad. Sci. USA*, *70*, 3321-3323.

PEET, R. 1974. The measurement of species diversity. *Ann. Rev. Ecol. Syst.*, *5*, 285-307.

PIELOU, E. C. 1969. *An Introduction to Mathematical Ecology*. Wily-Interscience, New York.

REDEI, G. P. 1975. *Arabidopsis* as a genetic tool. *Ann. Rev. Genet.*, *9*, 111-127.

RÖBBELEN, G. 1971. Further data on the extent of outcrossing in *Arabidopsis*. *Arabid. Inf. Serv.*, *8*, 7.

SCANDALIOS, J. G. 1969. Genetic control of molecular forms of isozymes in plants—a review. *Biochem. Genet.*, *3*, 37-79.

SHAW, C. R. AND PRASAD, R. 1970. Starch gel electrophoresis of enzymes—A compilation of recipes. *Biochem. Genet.*, *4*, 297-320.

SNAPE, J. W. AND LAWRENCE, M. J. 1971. The breeding system of *Arabidopsis thaliana*. *Heredity*, *27*, 299-302.

TRNĚNA, L., ŠKÁROVÁ, M., RELICHOVÁ, J. AND CETL, I. 1987. High outcrossing rates in a partially fertile line and their significance in the genetic structure of populations in *Arabidopsis thaliana* (L.) Heynh. *Arabid. Inf. Serv.*, *23*, 31-40.

WEIR, B. S. AND COCKERHAM, C. C. 1984. Estimating F- statistics for the analysis of population structure. *Evolution*, *38*, 1358-1370.

WRIGHT, S. 1951. The genetical structure of populations. *Ann. Eugen.*, *15*, 323-354.