

# COMPUTER SIMULATION OF A SPOROPHYTIC SELF-INCOMPATIBILITY BREEDING SYSTEM

By B. C. IMRIE,\* COLLEEN J. KIRKMAN,† and D. R. ROSS‡

[Manuscript received 1 November 1971]

## Abstract

A sporophytic self-incompatibility system, such as occurs in *Carthamus flavescens* Spreng., was studied by computer simulation. Equilibrium gene frequencies in an infinite population were estimated for three-allele and six-allele systems and found to be independent of initial frequencies. In a species existing as a series of more or less isolated small populations genetic drift caused rapid loss of alleles. Maintenance of the *S* allele system was enhanced by increased population size, and particularly by migration, by which alleles lost from one local population could be re-introduced from adjacent populations. Hard seed carryover had a lesser effect.

## I. INTRODUCTION

The breeding system of *Carthamus flavescens* Spreng., an annual herbaceous weed indigenous to Turkey, Syria, Lebanon, and Iraq, was found to be a single-locus, sporophytic, self-incompatibility system with at least six and probably eight or more *S* alleles (Imrie and Knowles 1971). It was further found that there was little variation among *S* alleles in plants drawn from geographically isolated populations. Since the species exists as a continuum of more or less isolated local populations, small enough for genetic drift to be an important factor in determining population gene frequencies, it was surprising to find that whereas morphologic differentiation had occurred as expected, there was little variation among *S* alleles (Imrie 1969).

The lack of variation among *S* alleles suggested that a high mutation rate was not the force maintaining allele numbers as was proposed by Wright (1964) and Mayo (1966) in studies of gametophytic self-incompatibility in finite populations. Wright calculated that mutation rates of  $2.08 \times 10^{-4}$  and  $2.20 \times 10^{-3}$  were required to maintain seven and 10 alleles respectively in a population of 50 plants. Physiological differences between gametophytic and sporophytic self-incompatibility systems, such as the site of the incompatibility reaction, have been documented. Similarly, genetic differences such as the number of alleles and their dominance relationships have been studied. It is implied above that the high mutation rates estimated to be required for the maintenance of a gametophytic system were not operative in the sporophytic system studied.

\* Division of Tropical Pastures, CSIRO, St. Lucia, Qld. 4067.

† Division of Mathematical Statistics, CSIRO; present address: Nudgee College, Sandgate Rd., Nudgee, Qld. 4014.

‡ Division of Computing Research, CSIRO, St. Lucia, Qld. 4067.

Where multiple alleles are maintained at a locus there must be a balance between the forces of selection, migration, mutation, and drift. In this paper the effects of migration, increasing population size, and carryover of hard seed are examined singly, and in combination, as potential stabilizing forces in the maintenance of a sporophytic self-incompatibility system. Simulation was chosen as the method of study because of difficulties associated with both field programs and algebraic solutions.

## II. MATERIALS AND METHODS

A program, written in FORTRAN for the CDC 3600 computer, simulated the passage of alleles in a self-incompatibility system through the sequence: mating  $\rightarrow$  zygote formation  $\rightarrow$  gamete production  $\rightarrow$  mating, with choice of parental combinations at mating and selection to maintain a fixed population size being random events. The model developed for this study deviated from the real situation only where such deviations promoted simplicity and were considered to be unlikely to materially affect the results and their interpretation. Features of the model were as follows:

- (1) An eight-allele system was chosen for study as this appeared to be the most likely number of alleles in *C. flavescens*. *S* alleles were numbered  $S_1, S_2, \dots, S_8$ , with  $S_1$  being dominant to all other alleles,  $S_2$  being dominant to alleles  $S_3$ – $S_8$ , and so on. Dominance was expressed in the pollen but gene action was independent in the style. These dominance relationships were determined for *C. flavescens* and are common in the sporophytic incompatibility systems of other species (Imrie, loc. cit.).
- (2) A base population of zygotes was generated by randomly pairing *S* alleles from a finite pool in which the alleles were present with equal frequency. The genotype  $S_1S_1$  was disallowed because this genotype cannot be formed in nature.
- (3) It was observed in the field in areas where *C. flavescens* occurs naturally that populations varied considerably in both size and density. Although most populations contained less than 100 plants an occasional population in a favourable location could contain in excess of 1000 plants (P. F. Knowles, personal communication). The harmonic means, and thus the effective population sizes (Li 1955) of several estimated distributions of population size (the actual distributions were not known) were calculated to be between 30 and 40 plants. It was assumed that size fluctuations due to environmental variability over years in any one location would approximate variations between locations observed in one year. Consequently, for studies of migration and hard seed carryover effects, a constant population size of 32 plants was chosen. In a population of this size it was possible to generate base populations having equal frequencies of all *S* alleles. The selection of 32 plants to form generation  $n+1$  from the progeny of generation  $n$  was random. The effect of population size on rate of *S* allele loss was studied in populations of 8, 16, 32, and 64 plants. The study of larger populations was precluded by computer time restrictions.
- (4) Mating was achieved by taking each member of the population in turn as a female, and combining it with a randomly selected member of the population as a male. *C. flavescens* is insect-pollinated. Considering the foraging behaviour of bees (Free 1970) there is no reason to suspect that pollination is non-random. The combination was tested for compatibility, and if compatible the resultant zygotes were added to a progeny population. For example, let plant  $X_1$  of the population, chosen as a female, have the genotype  $S_2S_5$ . By random choice, plant  $X_n$ , having the genotype  $S_2S_5$ , is chosen as a male. Since  $S_2$ , the dominant allele in the male, is also in the female the mating is incompatible and no progeny are derived. The mating process was repeated with, for example, plant  $X_m$ , having the genotype  $S_4S_5$ , chosen as a male. The progeny from this mating, assumed to be one of each possible genotype ( $S_2S_4, S_2S_5, S_5S_4, S_5S_5$ ), were added to the population of progeny. Ten matings were performed for each plant in the parent population.

- (5) Migration was achieved by increasing the progeny population by a predetermined proportion (referred to as percentage migrants) of zygotes generated from a gametic pool in which the frequency of different *S* alleles was equal.
- (6) Hard seed carryover was similarly achieved by adding to the progeny of generation  $n+1$  a predetermined proportion (percentage carryover) randomly selected from the progeny of generation  $n$ . It was possible for a seed to be carried over for several generations although the probability of a particular seed being carried over reduced rapidly. With a 20% carryover factor the probability of an individual in generation  $n$  being carried over to generation  $n+i$  was  $0.2^i$ .

A second computer program was used to determine the equilibrium gene frequencies. Crow and Kimura (1970) discuss the equilibrium situation for a gametophytic system where, for any number of alleles  $n$ , the equilibrium frequencies are approximated by  $P_i = 1/n$ . The exact solution for a gametophytic system is intractable (Crow and Kimura, loc. cit.) and for a sporophytic system is even more difficult. Arithmetic solutions were therefore derived by simulation for several examples. The program had the following features:

- (1) Input was in the form of frequencies of genotypes.
- (2) All possible matings were made with zygotic output determined by (i) compatibility, (ii) frequency of mating (product of genotypic frequencies). No selection was imposed and the population was assumed to be of infinite size.
- (3) Genotypic frequencies were calculated for generation  $n+1$  from the zygotic output of generation  $n$ , following which gene frequencies were calculated.
- (4) The program was run until equilibrium was attained.

### III. RESULTS

Equilibrium gene frequencies were determined for three-allele (the simplest case) and six-allele systems. The same equilibrium frequencies were obtained regardless of the initial gene and genotypic frequencies (Table 1). For the three-allele

TABLE 1  
INITIAL GENE FREQUENCIES AND EQUILIBRIUM FREQUENCIES OF THREE-ALLELE AND SIX-ALLELE SPOROPHYTIC SELF-INCOMPATIBILITY SYSTEMS

| Allele number       | Initial frequency |      |      |      |      |      | Equilibrium frequency |       |
|---------------------|-------------------|------|------|------|------|------|-----------------------|-------|
| Three-allele system |                   |      |      |      |      |      |                       |       |
| 1                   | 0.45              | 0.35 | 0.25 | 0.10 | 0.20 | 0.30 | 0.40                  | 0.198 |
| 2                   | 0.45              | 0.35 | 0.25 | 0.45 | 0.40 | 0.35 | 0.30                  | 0.258 |
| 3                   | 0.10              | 0.30 | 0.50 | 0.45 | 0.40 | 0.35 | 0.30                  | 0.544 |
| Six-allele system   |                   |      |      |      |      |      |                       |       |
| 1                   | 0.10              | 0.30 | 0.18 | 0.14 | 0.10 |      |                       | 0.107 |
| 2                   | 0.18              | 0.14 | 0.18 | 0.14 | 0.10 |      |                       | 0.115 |
| 3                   | 0.18              | 0.14 | 0.18 | 0.14 | 0.10 |      |                       | 0.127 |
| 4                   | 0.18              | 0.14 | 0.18 | 0.14 | 0.10 |      |                       | 0.147 |
| 5                   | 0.18              | 0.14 | 0.18 | 0.14 | 0.10 |      |                       | 0.181 |
| 6                   | 0.18              | 0.14 | 0.10 | 0.30 | 0.50 |      |                       | 0.323 |

system 8–19 generations were required to reach equilibrium but in the six-allele system, equilibrium was reached in 27–46 generations.

Equilibrium frequencies were dependent on the level of dominance, with alleles increasing in frequency as their degree of dominance decreased. This is due to the compatibility reaction of pollen being determined by the genotype of the plant on which it is produced.  $S_1$ , being dominant to all other alleles, is never masked, and plants carrying the  $S_1$  allele cannot be fertilized by  $S_1$  pollen. At the other end of the scale  $S_6$  is recessive to all other alleles in a six-allele system. Plants carrying  $S_6$  can be fertilized by  $S_6$  pollen provided it has the compatibility reaction of one of the five more dominant alleles with which it is paired in the sporophyte. Similarly  $S_5$  can be masked by four more dominant alleles, and so on. The number of successful matings involving any particular allele is therefore proportional to the level of dominance of that allele.

For the study of the effects of population size, migration, and hard seed carry-over in small populations results are expressed in terms of generations to loss of an  $S$  allele. Forty replications were run to determine population size effects, and migration and hard seed carryover effects were estimated from 100 replications. The distributions of numbers of generations to loss of an allele were strongly skewed and deviated significantly ( $P < 0.01$ ) in all cases from both normal and lognormal distributions. Because of the skewness it was considered most appropriate to express results as harmonic means rather than as arithmetic means.

#### (a) Population Size

Most local breeding populations of *C. flavescens* in nature are rather small. The effect of population size on the number of generations to loss of an  $S$  allele (Table 2)

TABLE 2  
EFFECT OF POPULATION SIZE ON LOSS OF  $S$  ALLELES

| Population size | No. of generations to loss of an allele |       |
|-----------------|---|-------|
|                 | Harmonic mean                           | Range |
| 8               | 2.32                                    | 2-5   |
| 16              | 3.20                                    | 2-14  |
| 32              | 4.90                                    | 2-18  |
| 64              | 6.40                                    | 2-68  |

was significant ( $P < 0.01$ ). However, in all cases, as would be expected with such small populations, allele loss due to random genetic drift occurred within a few generations. In the absence of migration or a strong frequency-dependent selection mechanism small populations would rapidly become extinct.

#### (b) Migration and Hard Seed Carryover

The numbers of generations to loss of an  $S$  allele when migration and hard seed carryover are included in the model are presented in Table 3. The effect of migration was relatively large. Mean numbers of generations to loss of an allele were 5.60, 6.37,

9.93, and 9.67 for 0, 1, 10, and 20% migration respectively. The effect of hard seed carryover was comparatively small, and in most cases non-significant.

TABLE 3  
EFFECTS OF MIGRATION AND HARD SEED CARRYOVER ON LOSS OF S ALLELES FROM A POPULATION OF 32 PLANTS

| Percentage hard seed carryover | No. of generations to loss of an allele* |               | Mean |
|--------------------------------|--|---------------|------|
|                                | No migration                             | 10% migration |      |
| 0                              | 5.60                                     | 9.93          | 7.76 |
| 10                             | 4.89                                     | 8.07          | 6.48 |
| 20                             | 5.95                                     | 11.11         | 8.53 |
| Mean                           | 5.48                                     | 9.71          |      |

\* Harmonic mean of 100 replicates.

Levels of significance of differences between distributions of number of generations to loss of an S allele are presented in Table 4. Because of the nature of the distributions the Kolmogorov-Smirnov non-parametric statistical test (Siegel 1956) was used to test the null hypothesis that the distributions were the same.

TABLE 4  
LEVEL OF SIGNIFICANCE OF DIFFERENCES BETWEEN DISTRIBUTIONS OF NUMBER OF GENERATIONS TO LOSS OF AN S ALLELE IN POPULATIONS OF 32 PLANTS

Significance levels: \* = 10%, \*\* = 5%, \*\*\* = 1%

| Migration (%) → |                           | 1  | 10  | 20  | 0    | 0   | 10   | 10   |
|-----------------|---------------------------|----|-----|-----|------|-----|------|------|
| ↓               | Hard seed carryover (%) → | 0  | 0   | 0   | 10   | 20  | 10   | 20   |
|                 | ↓                         |    |     |     |      |     |      |      |
| 0               | 0                         | ** | *** | *** | n.s. | *   | ***  | ***  |
| 1               | 0                         |    | *** | *** | **   | *   | ***  | ***  |
| 10              | 0                         |    |     | **  | ***  | *** | n.s. | ***  |
| 20              | 0                         |    |     |     | ***  | *** | *    | n.s. |
| 0               | 10                        |    |     |     |      | *   | ***  | ***  |
| 0               | 20                        |    |     |     |      |     | **   | ***  |
| 10              | 10                        |    |     |     |      |     |      | **   |

In addition to their effects on the number of generations to loss of an S allele, migration and hard seed carryover can cause the reintroduction of an allele to a population one or more generations after it has been lost due to genetic drift. In one computer run of 100 generations of a 32-plant population subjected to 10% migration a total of 24 alleles were lost, but in all cases the lost alleles were reintroduced in subsequent generations. At generation 100 the parental gene frequencies of alleles one to eight respectively were 0.09, 0.16, 0.17, 0.05, 0.08, 0.09, 0.03, and 0.33, and the progeny gene frequencies were 0.10, 0.14, 0.16, 0.05, 0.09, 0.10, 0.05, and 0.31.

(c) *Number of Alleles*

The effect of number of alleles was measured with three population sizes (Fig. 1). As the number of alleles decreased the number of generations to loss of a further allele increased. The point at which a curve becomes asymptotic to the *Y*-axis provides an estimate of the maximum number of alleles which can be maintained at that population size. This was found to be four alleles in a population of 32 plants. Populations of size 8 and 16 plants went to extinction, being unable to maintain the minimum three alleles necessary for the operation of an incompatibility system.

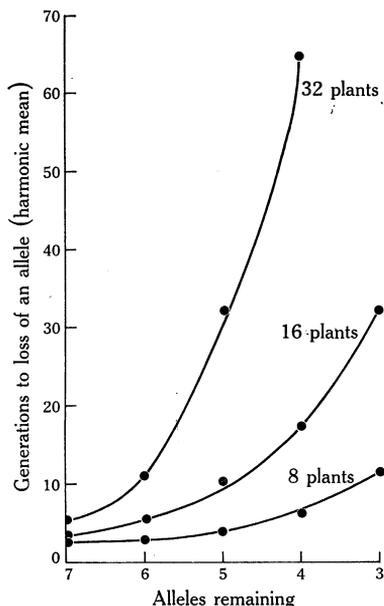


Fig. 1.—Rate of loss of *S* alleles due to genetic drift.

## IV. DISCUSSION

The most important factors contributing to maintenance of the self-incompatibility system in *C. flavescens* were found to be population size and migration. It was estimated that, in the absence of migration, four alleles would be maintained with an effective population size of 32. Computer time limitations prevented the generation of curves for larger population sizes but the trend indicated in Figure 1 is obvious. These data also support those of Wright (1964) who showed that with a mutation rate of  $10^{-5}$  five *S* alleles would be maintained in a population of between 20 and 30 plants and seven *S* alleles in a population of approximately 100 plants. Wright's estimations were for a gametophytic system which, although biologically different, is mathematically comparable for purposes of the above comparison. The proportion of successful matings increases at the same rate in both systems as the number of alleles increases, but the sporophytic system consistently has 3% fewer successful matings than the gametophytic system.

Migration increased the number of generations to loss of an *S* allele but in terms of maintenance of the self-incompatibility system its most important contribution was probably the reintroduction of an allele to a population following its loss due to

genetic drift. Following reintroduction there was an apparent tendency for alleles to approach their equilibrium frequencies. Equilibrium frequencies were not calculated for an eight-allele system but the pattern would be similar to that of the six-allele system (Table 1). The frequencies listed above for generation 100 of a 32-plant population subjected to 10% migration adhere to the pattern, particularly with respect to the high frequency of the allele lowest on the dominance scale. The approach to equilibrium was indicated by the changes in gene frequency between parent and progeny populations. Of the five alleles having frequencies less than 0.10 in the parent population four had higher frequencies in the progeny population and the fifth remained the same. Similarly, there was a reduction in the frequency of the more frequent alleles.

The measured effect of hard seed carryover on the retention of *S* alleles in the population was found to be small compared to the effect of migration. This could have been due to the proportion of seed carried over and may not be indicative of the real situation. King (1966) referred to the vast stores of hard seeds of weed species in the soil. No data are available on the effectiveness of hard seeds as a survival mechanism in *C. flavescens*, but Quinlivan and Peirce (1968), in a study of *C. lanatus* L., found seed germination spread over three seasons with the highest proportion germinating in the second season. If longevity is greater than 2–3 yr, or the proportion of hard seed carryover significantly in excess of 20%, then our model would tend to underestimate the effects of hard seed carryover. Epling, Lewis, and Ball (1960) gave considerable weight to hard seed carryover as a factor affecting gene frequencies in *Linanthus*. In our experiment the combined effect of migration and hard seed carryover was greater than the effect of migration alone.

From this simulation we have concluded that four *S* alleles would be maintained in a population of effective size 32 in the absence of migration and hard seed carryover. It is likely that the maintenance of the observed larger number of alleles in *C. flavescens* would be possible due to the reintroduction of lost alleles in subsequent generations as a consequence of migration and hard seed carryover.

## V. REFERENCES

- CROW, J. F., and KIMURA, M. (1970).—"An Introduction to Population Genetics Theory." (Harper and Row: New York.)
- EPLING, C., LEWIS, H., and BALL, F. M. (1960).—The breeding group and seed storage: a study in population dynamics. *Evolution* **14**, 238–55.
- FREE, J. B. (1970).—"Insect Pollination of Crops." (Academic Press: London and New York.)
- IMRIE, B. C. (1969).—Studies of variability and self-incompatibility in *Carthamus flavescens* Spreng. and its genetic relationship to *C. tinctorius* L. Ph.D. Thesis, University of California.
- IMRIE, B. C., and KNOWLES, P. F. (1971).—Genetic studies of self-incompatibility in *Carthamus flavescens* Spreng. *Crop Sci.* **11**, 6–9.
- KING, L. J. (1966).—"Weeds of the World." (Interscience: New York.)
- LI, C. C. (1955).—"Population Genetics." (University of Chicago Press.)
- MAYO, O. (1966).—On the problem of self-incompatibility alleles. *Biometrics* **22**, 111–20.
- QUINLIVAN, B. J., and PEIRCE, J. R. (1968).—The long-term field germination of saffron thistle (*Carthamus lanatus* L.) and the life span of dormant seeds in the Geraldton region, W.A. *J. Aust. Inst. agric. Sci.* **34**, 231–2.
- SIEGEL, S. (1956).—"Nonparametric Statistics for the Behavioral Sciences." (McGraw-Hill Book Company: New York.)
- WRIGHT, S. (1964).—The distribution of self-incompatibility alleles in populations. *Evolution* **18**, 609–19.

