

THE EFFECT OF GENE DISPERSAL ON THE DYNAMICS AND STATICS OF GENE SUBSTITUTION IN PLANTS

DONALD A. LEVIN

Department of Botany, University of Texas at Austin, Austin, Texas
and

HAROLD W. KERSTER

Environmental Studies Center, California State University, Sacramento, California

Received 19.iii.75

SUMMARY

A computer model has been developed to determine the effect of pollen and seed dispersal patterns on the probability of fixation and time to near-fixation of a rare advantageous gene in a plant population composed of 225 individuals. In the model, pollen and seed dispersal are according to four schedules (i) zero distribution, (ii) stepping-stone distribution, (iii) leptokurtic distribution, (iv) or random distribution. All combinations of pollen and seed dispersal schedules are treated. Relative fitnesses are assigned to three genotypes, *AA*, *Aa*, and *aa*, the latter always being the superior type. Relative fitness takes the form of competitive ability among the several seedlings germinating at each of the 225 safe sites in the population.

The probability of fixation of a rare advantageous recessive gene ($a = 0.013$) is inversely correlated with neighbourhood size. Restricted pollen and seed movements also retard the tempo of gene substitution relative to that expected with panmixia. Decreasing the intensity of competition has the same effect.

Gene frequency heterogeneity among subpopulations arises as a by-product of gene flow restriction, with graded patchworks of gene frequencies being most prominent when pollen and seed dispersal is narrow.

We considered the consequences of three receipt patterns (edge, leptokurtic, and random) of extraneous genes via pollen from a population located in a prescribed direction from the recipient. If the gene is advantageous, the more restricted the distribution to one edge of the population, the slower the tempo of gene substitution. If the gene is detrimental, the equilibrium gene frequency is depressed by restricting the deposition of pollen to the edge. In both cases non-random receipt yields properties very different from that when random immigration is assumed.

The literature indicates that gene flow within and between populations is non-random in plants. Our simulation demonstrates that abandoning the notion of panmixia will be a major step in understanding the dynamics and statistics of gene frequency change in plants.

1. INTRODUCTION

THE stochastic theory of gene frequency change as first elaborated by Fisher (1930) and Wright (1931) has played a prominent role in population genetics theory. However, most theory has dealt with populations experiencing random mating, which certainly is not representative of natural populations. The misapplicability of this assumption is especially striking in plants, where assortative mating in space is almost a certainty (Levin and Kerster, 1974). Assortative mating may take the form of self-pollination, pollination among neighbouring plants or leptokurtic pollination. The progeny of these matings may become established near the seed parent. To gain insight into the

dynamics of gene frequency change in plant populations, we may turn to the small body of literature on geographically structured populations, wherein random mating is precluded. The treatments of structured populations most pertinent to plants are the neighbourhood or isolation-by-distance model of Wright (1931, 1943, 1951), and the stepping-stone model of Kimura (1953, Kimura and Weiss, 1964), both of which have served as the basis for analyses of geographically structured populations by Maruyama (*e.g.* 1970*a, b*; 1971; 1972*a, b*), Kimura and Maruyama (1971), Crow and Maruyama (1971), and Rohlf and Schnell (1971). Unfortunately, this literature deals almost exclusively with neutral genes. This bias is due in part to the difficulty of including selection in analytical treatments of geographically structured populations and to the investigators' interest in neutral gene behaviour. A theory on the genetic properties of non-panmictic populations undergoing selective gene substitution is needed to understand the evolutionary process in plants.

In this investigation we address the following questions: (1) What is the probability of fixation and loss of an advantageous recessive gene under several combinations of pollen and seed dispersal schedules? (2) What is the mean time for gene substitution under different pollen and seed dispersal schedules? (3) What is the extent and pattern of gene frequency heterogeneity among population subdivisions during the substitution process? (4) What is the effect of altering competition among genotypes on the rate of gene substitution?

There is an abundance of evidence, especially from the crop literature, that all the individuals in a population are not equally likely to receive pollen from other populations (summarised in Levin and Kerster, 1974). In both wind- and insect-pollinated plants, most extraneous pollen is deposited on the border row facing the pollen source. With this in mind, we chose to find answers to the following questions: (1) What is the effect of the receipt pattern of a favourable recessive gene immigrating each generation on the rate of gene substitution; (2) What is the effect of the receipt pattern of a detrimental recessive gene immigrating each generation on the equilibrium frequency of that gene?

A formal mathematical investigation of these questions is most formidable; therefore, our approach is primarily one of simulation by a digital computer. We have chosen a stochastic model whose elements and variables are biologically realistic. The model clearly illustrates several consequences of non-random pollination and non-random dispersal of seed. Although constructed with plants in mind, the general conclusions are applicable to other types of sessile organisms.

2. THE MODEL

Simulations were conducted under the following conditions which are depicted in part in terms of a flow chart (fig. 2).

1. The model involved a population of 225 diploid annual plants.
2. Only a single locus with two alleles (A and a) was considered possible. The A allele is dominant or semi-dominant as prescribed.
3. The plants were assumed to occupy a 15×15 grid of uniformly spaced safe sites. Every plant was one map unit from its nearest neighbours. The site at the northeast corner of the population was designated row 1, column 1.

4. The initial population was composed of 219 *AA* genotypes and 6 *Aa* genotypes. The map locations of the latter with the row given before the column are as follows: 1, 10; 6, 7; 14, 15; 3, 8; 12, 7; 13, 12. This pattern was obtained from a random number generator with equal access to all sites.

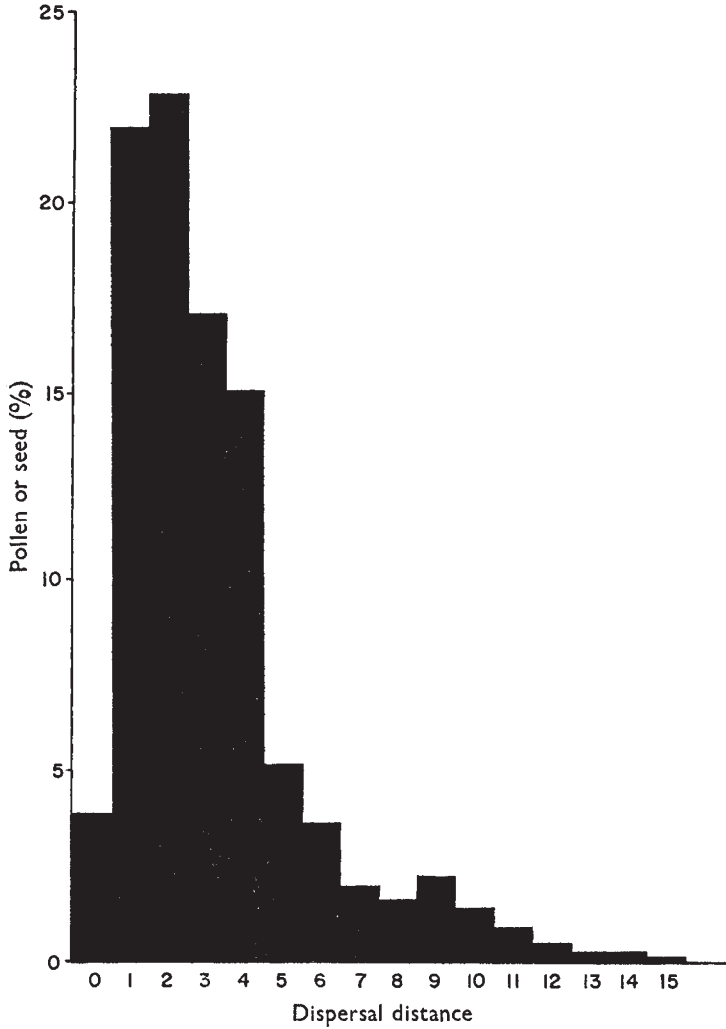


FIG. 1.—The distribution of pollen and seed dispersal distances under the leptokurtic dispersal schedule.

5. The breeding structure of the population was defined by the dispersal of pollen and seeds. Dispersal followed four schedules which are designated (a) zero, (b) stepping-stone, (c) leptokurtic, and (d) random. The pollen and seed schedules are independent. Zero dispersal refers to no movement of pollen or seeds. The stepping-stone schedule involves the movement of pollen or seeds to one of the four nearest sites at the cardinal compass points. The particular site was determined by a random number generator. For

the leptokurtic schedule, pollen and seed dispersal distances assumed the distribution shown in fig. 1, but seeds produced from self-fertilisation were discarded. The direction of broadcast was designated by a random number generator. This distribution is similar to that described for pollen and seeds in natural populations (Levin and Kerster, 1974). With random dispersal, pollen and seed had equal likelihood of reaching any site in the population, but no self-fertilisation was permitted.

6. Reproduction is modelled as a sequential series of events repeated several times each generation. Specifically, each plant produces 18 pollen grains. The pollen grains are distributed in 18 episodes, each episode involving the dispersal of one pollen grain per plant according to the dispersal schedule. Each grain landing within the population is considered to be deposited on a stigma leading immediately to production and dispersal of a seed. The number of ovules at each site (or per plant) is large so that seed production is limited by pollen receipt. Subsequent episodes of pollination and seed dispersal follow until all pollen has been dispersed. Although the mechanics are not realistic, the results are. Each plant produces about 18 pollen grains (some are carried beyond the boundaries of the population) which are effective in the reproductive process. On the average, each plant produces nearly 18 seeds which reach sites suitable for germination and the establishment of an adult plant. The choice of 18 pollen grains was dictated by the desire to fill all of the safe sites with at least one seed. The location of each seed and the genotype thereof was stored on arrival for use in the selection procedure.

7. Competitive selection occurs if a seed arrives at an occupied site. The prior occupant and the new arrival divide between them the domain of a uniform random number generator, each assigned a portion of the domain proportional to its relative fitness. A random number is drawn and the site allotted to the competitor in whose portion of the domain the number has fallen. For example, an occupant has fitness 1.0 and a new arrival fitness 0.2. The occupant is assigned the region 0.0 and 1.0 and the challenger the region 1.0 to 1.2 of a random number generator domain of 0.0 to 1.2 of a random number generator domain of 0.0 to 1.2. If the random number was 0.5 (within the occupant's region), the occupant would retain possession of the site, but might have to defend it against yet another challenger, should another seed arrive at the site during that generation's events. This procedure yields results like that obtained analytically by Nei (1971) for competition among more than two individuals.

8. We have explored the condition where the relative fitnesses (as defined by competitive ability) of the dominant homozygote and heterozygote were the same: $W_{AA} = 0.20$, $W_{Aa} = 0.20$, and $W_{aa} = 1.0$; departures from the basic model (uniform selection against the dominant phenotype) will be discussed as they are performed. Genotype fitness is constant throughout the population.

9. Simulations commenced with the *a* gene frequency being 0.013, and were run until this gene reached or surpassed a frequency of 0.98, herein referred to as the point of near-fixation. The cost of simulating to fixation was prohibitive. Twenty replicates of each pollen-seed dispersal combination were run to near-fixation.

Our model involves competition among plants in a saturated environment. Although competitive selection of this type undoubtedly occurs in

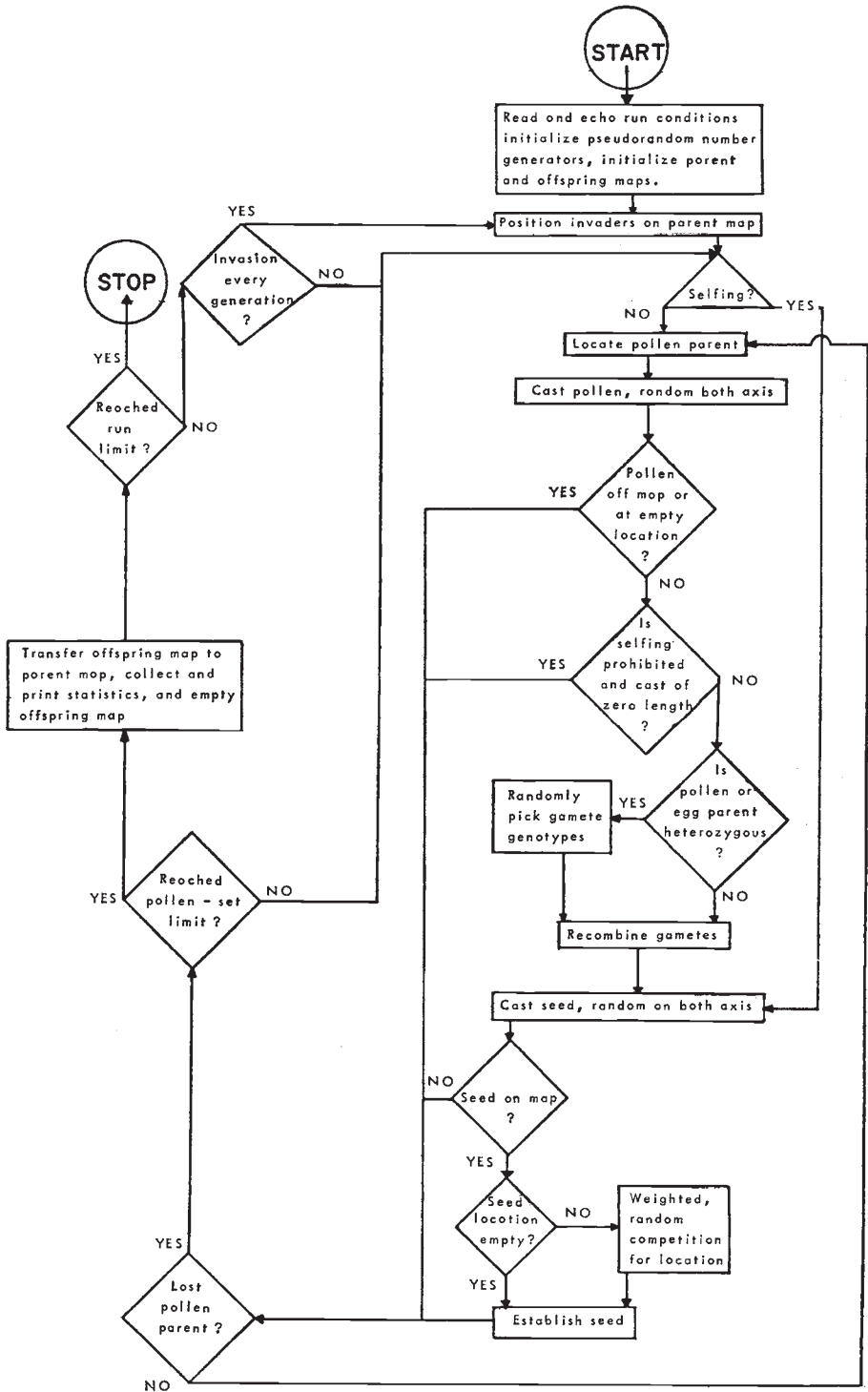


FIG. 2.—Flow chart of the simulation programme. See text for additional information.

natural populations, it has received little attention. Mather (1969) and Nei (1971) developed mathematical models for competitive selection in regulated populations and showed that many classical formulae for the change in gene frequency hold true under this form of selection given random mating.

3. PROBABILITY OF FIXATION OF A MUTANT GENE

The results of the Monte Carlo experiments demonstrating the effect of pollen and seed dispersal distances on the probability of gene fixation are summarised in table 1. The schedule of pollen and seed dispersal has a pronounced effect on the probability of fixation. Where pollen and seed dispersal is random, this probability is 0.49, which approximates 0.51 expected from the equations formulated for continuous stochastic processes (*e.g.* Maruyama and Kimura, 1971). As dispersal becomes more restricted

TABLE 1

Fixation probability of a mutant with initial frequency of 0.013 as a function of gene dispersal

| Pollen dispersal | Seed dispersal | | | |
|------------------|----------------|----------------|-------------|--------|
| | Zero | Stepping-stone | Leptokurtic | Random |
| Zero | — | 0.85 | 0.70 | 0.55 |
| Stepping-stone | 0.82 | 0.78 | 0.69 | 0.55 |
| Leptokurtic | 0.61 | 0.64 | 0.57 | 0.50 |
| Random | 0.50 | 0.52 | 0.53 | 0.49 |

the probability of fixation increases, progressing to 0.57 with leptokurtic pollen and seed dispersal and to 0.78 with stepping-stone pollen and seed dispersal. Fixation probabilities of *c.* 0.80 are forthcoming when self-pollination is obligatory and seed dispersal follows the stepping-stone schedule.

Since the fixation probability of an advantageous gene is dependent upon the gene dispersal schedule, it is of value to express this schedule in the form of neighbourhood size (N_e) based upon Wright's (1943, 1951) isolation by distance model. Neighbourhood sizes for each pollen and seed dispersal combination have been calculated according to the following formulae using the variance of pollen and seed dispersal distances (in arbitrary distance units) which accrued in our simulation. The expression for neighbourhood size which may be applied to plants is $N_e = 6.3 rd (\Sigma p^2/Np + \Sigma s^2/Ns)$ where p = pollen dispersal distance, s = seed dispersal distance, N_p = number of pollen grains, d = effective density and r = the proportion of self-fertilised seed (Levin and Kerster, 1971). Neighbourhood sizes are presented in table 2. On a semi-log plot, there is linear relationship between the probability of fixation and N_e . The coefficient of correlation between the fixation probability and the log transformed N_e is a surprising -0.98 .

We have explored the probability of fixation of an advantageous recessive mutant gene as a function of the relative fitness of the dominant phenotype. In addition to the fitnesses assigned in our basic model ($W_{AA} = 0.20$, $W_{Aa} = 0.20$, and $W_{aa} = 1.0$), we treated the cases where the fitnesses of both dominant homozygote and heterozygote both were 0.40, 0.60, and

0.80. The initial gene frequency is 0.013 in the form of 6 heterozygotes. The results of our Monte Carlo experiments are summarised in table 3, which also includes the theoretical values for panmixia at $N_e = 225$ obtained from the formulae of Maruyama and Kimura (1971). Two features are conspicuous. First, as the relative fitness of the dominant phenotype increases, the probability of fixation of the advantageous gene decreases. The values obtained from our random dispersal simulations are in good accord with theoretical expectations. Second, the dispersal schedule notably affects fixation probabilities, relative fitness notwithstanding. Again we see that restricted gene dispersal enhances the probability of fixation.

TABLE 2
Neighbourhood size as a function of pollen and seed dispersal

| Pollen dispersal | Seed dispersal | | | |
|------------------|----------------|----------------|-------------|--------|
| | Zero | Stepping-stone | Leptokurtic | Random |
| Zero | — | 9 | 113 | 907 |
| Stepping-stone | 13 | 23 | 128 | 929 |
| Leptokurtic | 452 | 475 | 615 | 1809 |
| Random | 900 | 940 | 1063 | 2715 |

TABLE 3
Fixation probability of an advantageous gene with initial frequency of 0.013 as a function of relative fitness

| Relative fitness | Pollen and seed dispersal schedule | | | Theoretical expectation† |
|-------------------------------------|------------------------------------|-------------|--------|--------------------------|
| | Stepping-stone | Leptokurtic | Random | |
| $W_{AA} = W_{Aa} = 0.2, W_{aa} = 1$ | 0.76 | 0.57 | 0.49 | 0.51 |
| $W_{AA} = W_{Aa} = 0.4, W_{aa} = 1$ | 0.64 | 0.36 | 0.37 | 0.34 |
| $W_{AA} = W_{Aa} = 0.6, W_{aa} = 1$ | 0.40 | 0.32 | 0.24 | 0.23 |
| $W_{AA} = W_{Aa} = 0.8, W_{aa} = 1$ | 0.32 | 0.24 | 0.18 | 0.15 |

† Based on formulae of Maruyama and Kimura (1971).

4. THE TEMPO OF GENE SUBSTITUTION

The effects of different pollen and seed dispersal on the tempo of gene substitution are summarised in fig. 3 and table 4. The figure illustrates mean gene frequency as a function of time, and the table provides mean near-fixation ($q \geq 0.98$) times and their standard errors. Consider the case where the relative fitness of the genotypes are $W_{AA} = 0.20$, $W_{Aa} = 0.20$, and $W_{aa} = 1.0$. Gene frequencies over time form curves varying from convex to concave and sigmoid, depending on the dispersal parameters (fig. 3). The convex curves, characterised by an initial rapid increase in gene frequency followed by an asymptotic approach to fixation, are diagnostic of populations undergoing complete self-pollination. This pollination schedule accompanied by leptokurtic or random seed dispersal afford the most favourable circumstances for the saltational increase in the frequency of a novel advantageous gene (fig. 3, A). The concave curves, characterised by

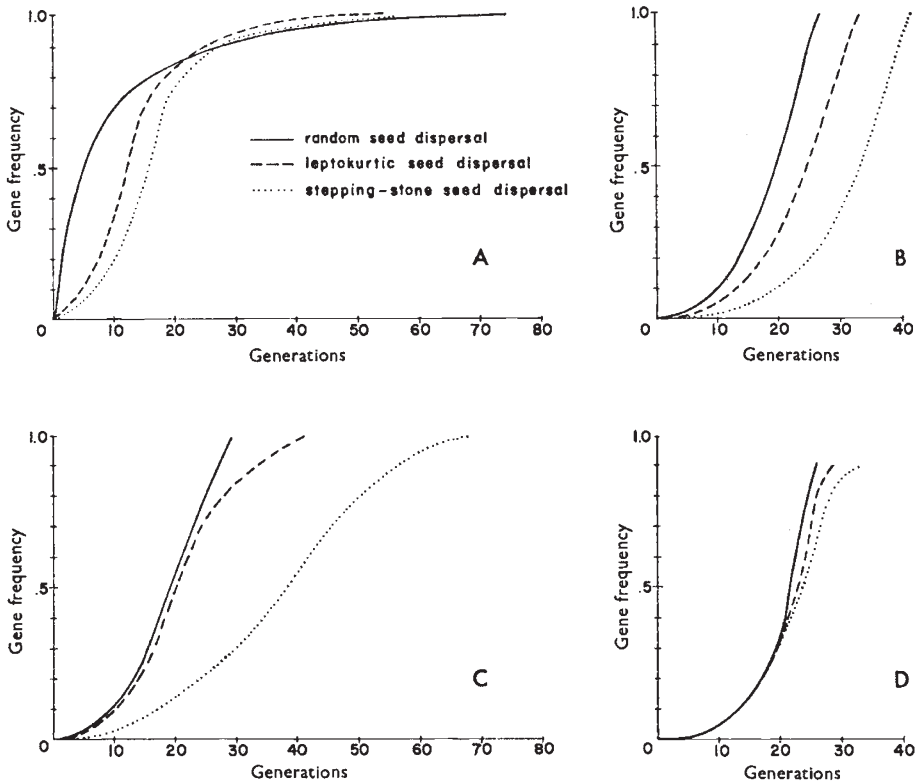


FIG. 3.—The tempo of gene substitution under different seed dispersal schedules with pollen dispersal a constant. A. Zero pollen dispersal. B. Leptokurtic pollen dispersal. C. Stepping-stone pollen dispersal. D. Random pollen dispersal. Genotype fitnesses: $W_{AA} = W_{Aa} = 0.20$, and $W_{aa} = 1.0$.

TABLE 4

Average number of generations until near-fixation of the *a* gene. Genotype fitnesses: $W_{AA} = 0.20$, $W_{Aa} = 0.20$, $W_{aa} = 1.0$

| Pollen dispersal | Seed dispersal | | | |
|------------------|------------------|------------------|------------------|------------------|
| | Zero | Stepping-stone | Leptokurtic | Random |
| Zero | — | 54.8 ± 12.29 | 70.8 ± 16.57 | 76.4 ± 15.28 |
| Stepping-stone | 68.7 ± 14.60 | 41.2 ± 8.75 | 28.0 ± 7.28 | 27.2 ± 5.80 |
| Leptokurtic | 42.8 ± 9.72 | 33.0 ± 7.38 | 26.1 ± 7.54 | 25.6 ± 6.13 |
| Random | 34.1 ± 8.12 | 28.9 ± 6.94 | 26.0 ± 6.28 | 25.1 ± 5.04 |

a sluggish increase in gene frequency early in the substitution process and a rapid increase thereafter, are diagnostic of random and leptokurtic pollen dispersal. The seed dispersal schedule has little effect upon the gene frequency curves when pollen dispersal is random (fig. 3, D). However, when pollen dispersal is leptokurtic, the more restricted is seed dispersal, the longer is the period before gene frequency changes rapidly (Fig. 3, B). Stepping-stone pollen flow with random or leptokurtic seed flow yields a curve like

that obtained with panmixia. Seed broadcasting via the stepping-stone schedule, or no broadcasting result in sigmoid curves, the slope being greater in the case of the former (Fig. 3, C).

The near-fixation time of a favourable recessive gene is greatly influenced by the breeding structure of the population (table 4). Mean near-fixation times vary from 25 generations to 76 generations. The longest times are in obligately self-pollinating populations, regardless of the seed dispersal regime. This is especially interesting since these populations experience the most rapid increase in gene frequency during the initial phase of substitution. As the breeding structure of the population become more open, the tempo of substitution is accelerated. Just the shift from self-pollination to the stepping-stone schedule results in a three-fold decrease in the time to fixation, with leptokurtic or random seed movement.

We should have an agreement between theoretical predications for panmictic systems and the results from our Monte Carlo experiments with random pollen and seed dispersal if our model is working satisfactorily. Employing the following recurrence equation:

$$p' = \frac{(p^2 + \Delta F p(1-p))W_{AA} + (2p(1-p) - \Delta F 2p(1-p))W_{Aa}}{1 - (p^2 + \Delta F p(1-p) + 2p(1-p) - \Delta F 2p(1-p))} \quad (1)$$

where p' is the frequency of A in the n th generation, p the frequency of A in the previous generation, ΔF is the change in the inbreeding coefficient per generation, and W_{AA} and W_{Aa} are the relative fitnesses of the dominant phenotypes. The theoretical fixation time is 28 generations versus 26 generations which we obtained. In view of our standard error of 5.04, the agreement is quite good.

5. MEAN EXTINCTION TIME

All pollen and seed combinations were simulated until 30 extinction episodes were obtained. The mean extinction time for each combination is presented in table 5. Extinction time is a positive function of the distance genes are dispersed. The mean number of generations to extinction varies from 4.0 when there is severe inbreeding to *c.* 14.0 when either pollen or seed

TABLE 5

Mean time to extinction of advantageous gene as a function of pollen and seed dispersal. Initial gene frequency, a = 0.013. Genotype fitnesses: $W_{AA} = W_{Aa} = 0.20$; $W_{aa} = 1.0$

| Pollen dispersal | Seed dispersal | | | |
|------------------|----------------|----------------|-------------|-------------|
| | Zero | Stepping-stone | Leptokurtic | Random |
| Zero | — | 4.0 ± 1.15 | 8.4 ± 1.72 | 13.4 ± 3.10 |
| Stepping-stone | 4.2 ± 1.02 | 5.1 ± 1.22 | 8.8 ± 1.45 | 13.8 ± 2.65 |
| Leptokurtic | 9.6 ± 2.14 | 10.5 ± 2.40 | 11.2 ± 2.63 | 12.9 ± 2.91 |
| Random | 13.8 ± 2.85 | 14.2 ± 3.60 | 13.6 ± 2.90 | 13.1 ± 2.75 |

dispersal is at random. The mean extinction time is inversely correlated with the probability of fixation, $r = -0.65$ ($P < 0.01$). This relationship emerges because dispersal schedules characterised by high fixation probabilities show rapid increases in the frequency of the advantageous genes soon after the substitution process commences. Thus, if the gene is to be lost, it is most likely to happen in the first few generations.

6. THE GEOGRAPHICAL STRUCTURE OF POPULATIONS DURING GENE SUBSTITUTION

The process of gene substitution typically is considered without regard to the distribution of genes in space. However, since some dispersal schedules have a profound effect on the geographical structure of populations, it is of interest to ascertain how the spatial distribution of genes changes through time. We will describe geographical structure in two ways. We will present sequences of maps from two simulations in which the genotypes of plants at the 225 sites are shown at eight intervals during gene substitution. Pollen and seed dispersal follow the stepping-stone schedule in one simulation and the leptokurtic schedule in the second. Populations were divided into 25 3×3 plant quadrants, and the true standardised genetic variance among these quadrants was computed for each pair of pollen and seed dispersal schedules. This value, referred to as F_{ST} is given by $\sigma^2 p / \bar{p} \bar{q}$ minus a contribution to the variance due to sampling error which is approximately $\bar{p} \bar{q} / 2N$, where $\sigma^2 p$ is the variance of the A gene among quadrants, \bar{p} and \bar{q} are the mean frequencies of A and a genes throughout the population and N is the sample size (Wright, 1951; Workman and Niswander, 1970).

The maps portraying a population with stepping-stone dispersal are shown in fig. 4. Each map is annotated with the frequency of the a gene and the F_{ST} for that generation. We start tracking the population after the frequency of the favoured gene reaches 0.10, and follow its distribution every other generation until it reaches a frequency of 0.91. By chance, the a gene is most common in the southeast quadrant of the population (fig. 4, map I). This area is largely occupied by aa homozygotes by the time the gene frequency of a is 0.43 (fig. 4, map IV), and this area of homozygotes expands during subsequent generations. The upper left corner of the population happens to have five homozygous (aa) plants early in the substitution process (fig. 4, map I). The area expands through time and coalesces with the major area of aa homozygotes when the frequency of the a gene is 0.60 (fig. 4, map V). The most striking feature of this series of maps is the patchiness in the distribution of the A and a genes, and "infective" spread of the a gene to neighbouring portions of the population. For the leptokurtic sequence, the geographical structure is depicted for each generation between gene frequencies of 0.1 and 0.9 (fig. 5). By chance, the a gene initially is more common in the lower half of the population (fig. 5, map I). The imbalance in gene frequency is retained for the remainder of the generations we are charting, being a lasting by-product of gene flow restriction. Gene substitution proceeds at similar rates throughout much of the population in spite of the imbalance, and the F_{ST} is small. This is in contrast to the localised shifts in gene frequency per generation and the high F_{ST} in the population with stepping-stone dispersal.

The standardised genetic variance for 15 combinations of pollen and gene dispersal schedules is presented in table 6. The value given is the average over the generations wherein the frequency of a was between 0.1 and 0.9. The greatest gene frequency heterogeneity occurs when pollen dispersal is zero and seed dispersal follows the stepping-stone schedule ($F_{ST} = 0.70$) and when the converse is true ($F_{ST} = 0.65$). Stepping-stone dispersal of both pollen and seeds also results in pronounced heterogeneity ($F_{ST} = 0.52$). Combining a leptokurtic distribution with a more restricted dispersal schedule affords moderate local differentiation ($F_{ST} = 0.13-0.19$) whereas

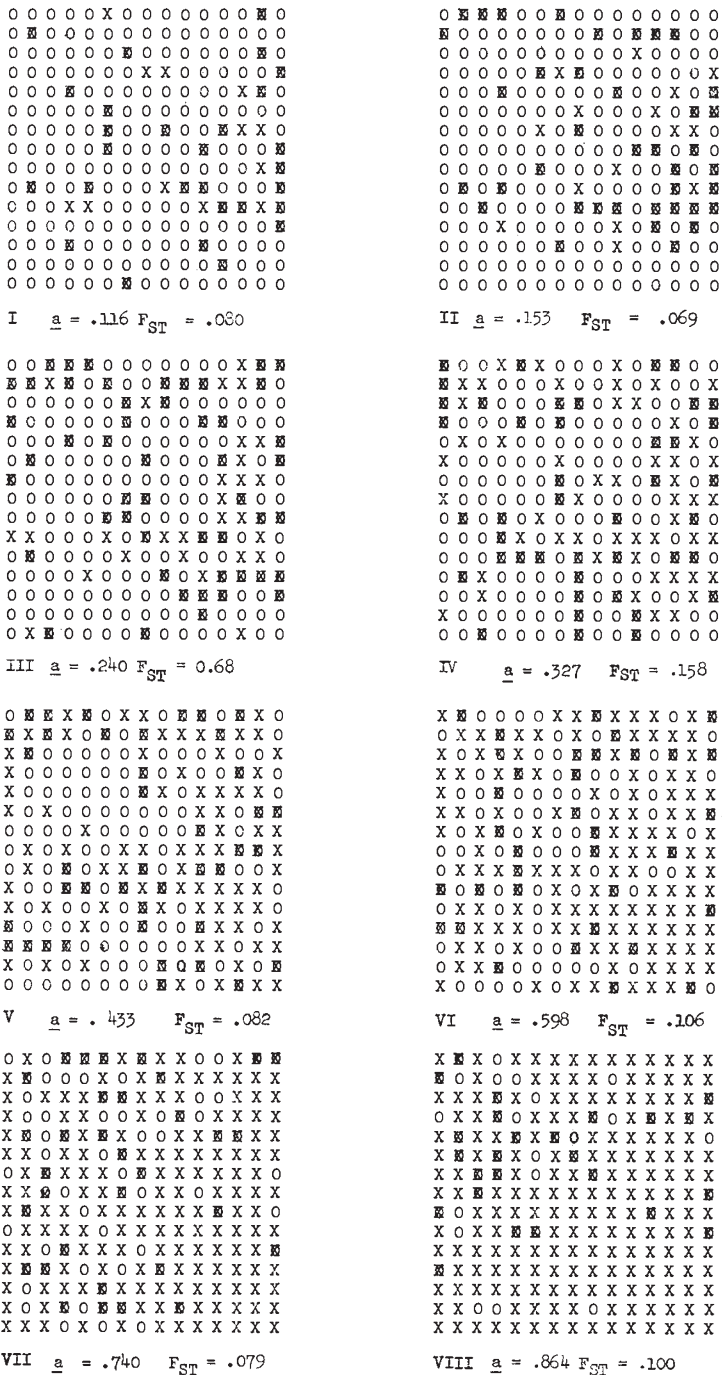


FIG. 5.—The distribution of genotypes in space during gene substitution. Pollen and seed are dispersed according to the leptokurtic schedule. Genotype designations are as follows: $\circ = AA$; $\otimes = Aa$; $\times = aa$.

leptokurtic dispersal of both pollen and seeds yields only weak differentiation ($F_{ST} = 0.06$). The geographical structure developed under restricted gene flow is striking, and it is achieved in the absence of geographically heterogeneous selection pressures.

TABLE 6

The true standardised genetic variance as a function of gene dispersal

| Pollen dispersal | Seed dispersal | | | |
|------------------|----------------|----------------|-------------|--------|
| | Zero | Stepping-stone | Leptokurtic | Random |
| Zero | — | 0.65 | 0.17 | -0.02 |
| Stepping-stone | 0.70 | 0.52 | 0.13 | 0.02 |
| Leptokurtic | 0.19 | 0.15 | 0.06 | 0.03 |
| Random | -0.01 | 0.03 | 0.01 | -0.02 |

7. THE EFFECT OF COMPETITION ON THE TEMPO OF GENE SUBSTITUTION

We have assumed that fitness is a function of the genotype's ability to exploit a limiting resource, and thus selection in our model is through competition. Thus the extent to which the fitness differential is translated into gene frequency change is dependent upon the incidence of contact between different genotypes. If populations are subdivided, contact is primarily among the same genotype and the intergenotype interaction is relatively weak. The incidence of contact between two genotypes is also a function of population density. Since contact between two different genotypes is less likely in sparse populations than in dense ones, the selection against a deleterious gene or genotype will be less in sparse populations. Mather (1969) has analysed selection through competition, emphasising the difficulty of describing the phenomena by relatively simple fitness sets.

In order to modify seedling densities, we varied pollen production per plant. As earlier, we assume that all pollen deposited in the population leads to seed production and that all seeds germinate. Four levels of pollen production were considered: 6, 12, 18 and 24 grains. Pollen and seed dispersal is leptokurtic. Genotype fitnesses are $W_{AA} = 0.20$, $W_{Aa} = 0.20$, and $W_{aa} = 1.0$.

The results of our simulations are depicted in fig. 6. It is clear that the rate of substitution is dependent upon seedling density. The mean near-fixation times for the four pollen production levels are as follows: 19.2 ± 5.62 generations with 24 grains; 26.1 ± 7.54 generations with 18 grains, 32.4 ± 7.23 generations with 12 grains; 40.7 ± 9.44 generations with 6 grains. It is also apparent that the rates of gene substitution between the frequencies of 0.1 and 0.9 are similar, pollen production notwithstanding. With 24 grains it takes about 8 generations for a recessive gene to reach a frequency of 0.9, whereas with 6 grains it takes 12 generations. These data indicate that altering seed production via pollen production and seedling competition primarily affects the time to the "log" phase of gene frequency increase. Once 0.1 is reached, the rate at which the recessive gene increases and the time to near-fixation are not greatly affected by seedling density or its correlate, competition.

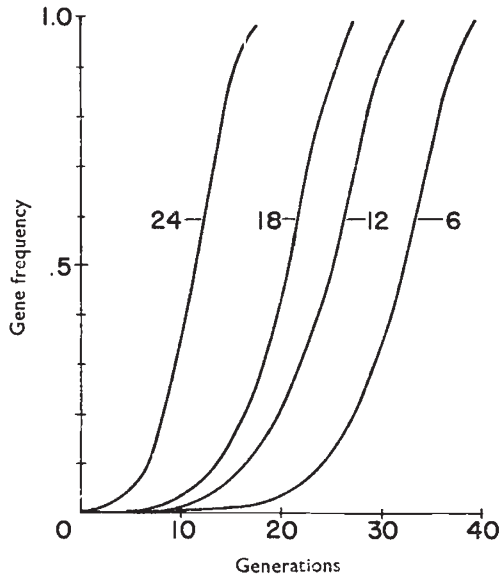


FIG. 6.—The tempo of gene substitution as a function of competition expressed in terms of pollen production per plant. See text for additional explanation

8. GENE SUBSTITUTION OR EQUILIBRIUM WITH IMMIGRATION

Migration represents one of the standard variables in the repertoire of theoretical population genetics. When migration is included in analytical or stimulation studies, the incoming genes are assumed to be randomly distributed throughout the population. This assumption is not in accord with the long-standing observations of plant breeders. Indeed most plants producing "hybrid" seed are located along or near the border of the population proximal to the pollen source, be the alien pollen wind- or animal-borne. In order to determine the effect of the receipt patterns on the rate of gene substitution (be the immigrant advantageous) or on the equilibrium gene frequency (be the immigrant detrimental), we began with our basic model using leptokurtic pollen and seed dispersal as constants. The extraneous pollen receipt pattern assumes three forms: (1) all pollen is received only by the proximal row or edge of the population; (2) pollen receipt follows a leptokurtic distribution like that used for intra-population pollen flow, with most pollen being deposited near the proximal edge of the population; (3) pollen receipt is random. Six hybrid seed are produced each generation; the immigration rate is 0.013; the extraneous gene is recessive.

With the extraneous gene being advantageous, we explored the rate of gene substitution given the following relative fitnesses: $W_{AA} = 0.20$, $W_{Aa} = 0.20$, and $W_{aa} = 1.0$; a second set of simulations were run with fitness of $W_{AA} = 0.60$; $W_{Aa} = 0.60$ and $W_{aa} = 1.0$. The change of gene frequency through time and the mean time to near-fixation are depicted in fig. 7. Both the gene frequency curve and the near-fixation time are functions of the gene receipt pattern. The recessive gene increases most rapidly with random receipt and most slowly with edge receipt, relative fitness of the heterozygote notwithstanding. Similarly, the near-fixation time is the

shortest with random receipt and the longest with edge receipt. When the fitness of the heterozygote is 0.20, the mean times to near-fixation are as follows: random receipt, 32 generations; leptokurtic receipt, 37 generations; edge receipt, 42 generations. The simulations show that the edge receipt pattern, which is the most realistic of those treated, clearly results in a slower tempo of gene substitution than does random receipt. The effect of edge receipt would be enhanced had we used a more restricted pair of pollen and seed dispersal schedules within the population.

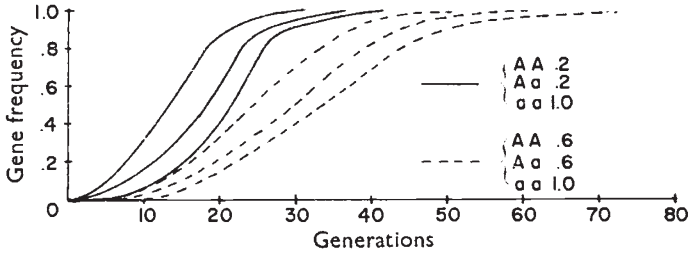


Fig. 7.—The tempo of gene substitution under different immigration patterns and relative fitnesses. For each family of curves the one on the left involves random pollen receipt, the center one leptokurtic receipt, and the right one edge receipt.

The time to fixation of an advantageous gene where $m = 0.013$ and genotype fitnesses are $W_{AA} = 0.20$, $W_{Aa} = 0.20$, and $W_{aa} = 1.0$, may be computed from the following recurrence equation:

$$p' = \frac{(p^2 + \Delta F p(1-p))W_{AA} + (2p(1-p) - \Delta F 2p(1-p))W_{Aa} + m}{1 - (p^2 + \Delta F p(1-p)) + 2p(1-p) - \Delta F 2p(1-p) + m} \quad (2)$$

where m is the immigration rate; the other symbols are the same as in equation (1).

The population size is assumed to be 225 and intrapopulation mating and the distribution of immigrants is assumed to be random. The recessive gene approaches fixation in 12 generations. This is in contrast to 32 generations in our stochastic model where the pollen receipt is random but intrapopulation gene flow is leptokurtic. The near-fixation time obtained from the formula is less than 30 per cent of that obtained from our model when the gene receipt pattern is edge or leptokurtic. When selection is less rigorous ($W_{AA} = 0.60$, $W_{Aa} = 0.60$, and $W_{aa} = 1.0$), the near-fixation time dictated by the formula is 24 generations. Again we see that disregard of receipt pattern and the breeding structure of the recipient population leads to a gross underestimate of the near-fixation time of the immigrant gene.

Consider next the gene frequency equilibria reached when the immigrating gene, a , is deleterious, and the receipt pattern is the variable. The results of our simulations are depicted in fig. 8. Given a recessive phenotype fitness of 0.20, the highest mean equilibrium frequency is achieved with a random receipt pattern, $a = 0.112$. Equilibrium values for the leptokurtic and edge pattern are considerably less, $a = 0.064$ for the former and 0.044 for the latter. A similar disparity is evident with relaxed selection against the dominant phenotype ($W = 0.60$). The equilibria are as follows: $a = 0.199$ with random receipt; $a = 0.126$ with leptokurtic receipt; $a = 0.083$ with edge receipt.

The introduction pattern is a prime determinant of gene frequency equilibria, because in conjunction with the breeding structure of the recipient population it determines the percentage of heterozygosity and its corollary the effectiveness of selection. With leptokurtic pollen and seed dispersal schedules within the recipient population, random gene receipt, and recessive phenotype fitness of 0.20, heterozygosity averages 21.6 per cent. This compares with mean heterozygosities of 13.7 per cent with leptokurtic receipt and 8.8 per cent with edge receipt. A large fraction of pollinations are among neighbouring plants. Thus the more restricted the area inhabited by heterozygotes, the greater will be the proportion of recessive homozygotes at a

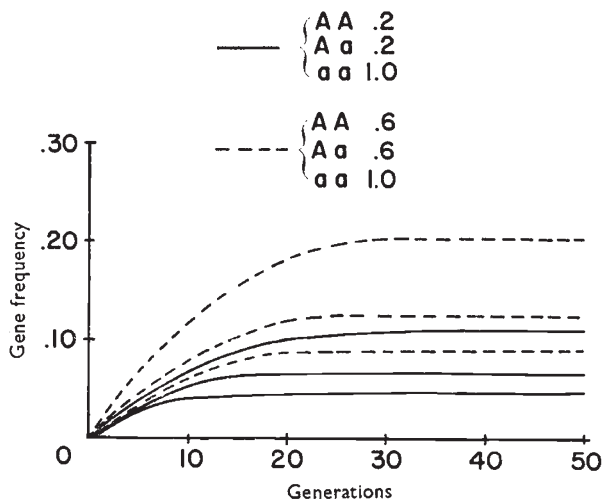


FIG. 8.—The tempo and level of gene frequency equilibria under different immigration curves and fitnesses. For each family of curves, the upper one involves random pollen receipt, the center one leptokurtic receipt, and the lower one edge receipt.

given gene frequency. Should a more restricted breeding structure have been used, the disparity in equilibria caused by different receipt patterns would have been even greater than described here.

When migration is counterbalanced by selection, the equilibrium level may be found by iteration of the expression given earlier. With the immigration rate at 0.013, a population size of 225 and a phenotype fitness of 0.20, the equilibrium frequency of the alien gene is 0.14. This value is twice that obtained in our simulations assuming leptokurtic receipt, and is three times that obtained assuming edge receipt. A similar disparity between the deterministic equilibrium level and that obtained in our simulations is present when the fitness of the phenotype is 0.60. The theoretical equilibrium frequency is 0.20, *vs.* 0.20 with random receipt, 0.13 with leptokurtic receipt and 0.08 with edge receipt. Should the breeding structure of the population have been more restrictive, the disparity between the deterministic equilibrium values and that from the simulations would have been even greater. Thus the formula used for calculating gene frequency equilibria when immigration is opposed by selection grossly overestimates \hat{q} , by ignoring realistic patterns of gene receipt and the breeding structure of the

recipient population. When this is done, the rule of thumb that one immigrant per generation precludes drift (Kimura and Ohta, 1971) is inoperative.

Each pattern of extraneous gene input leaves a characteristic spatial distribution of genotypes in the recipient population, even though pollen and seed dispersal therein is leptokurtic. To illustrate this point, we chose a

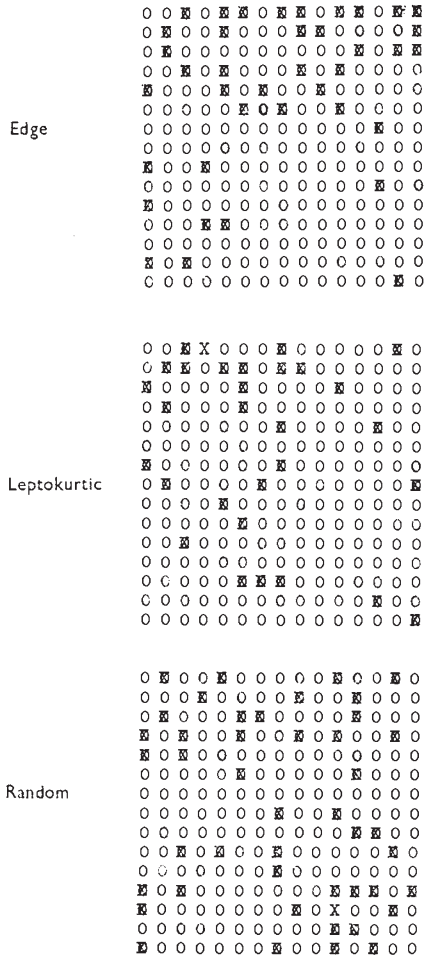


FIG. 9.—Representative distributions of genotypes in space after gene frequency equilibria have been reached under random, leptokurtic and edge pollen receipt patterns. Genotype designations are as follows: ○ = AA; ⊙ = Aa; × = aa. The north edge of the population faces the extraneous pollen source.

representative map of genotype locations for each input pattern after the population began to fluctuate about the equilibrium frequency. When extraneous gene input only is along the edge of the population facing the source, the gene tends to remain near the edge although there is some penetration of the interior of the population (fig. 9). The fact that geographical structure is maintained in spite of leptokurtic gene dispersal

bespeaks of the weakness of this schedule in homogenising gene frequencies when structure is promoted by a systematic force. There is also a tendency for the extraneous gene to reside near the proximal edge of the population when the input pattern is leptokurtic, the result being a gene frequency cline from the proximal to the distal edge of the population (fig. 9). When the gene input is random, no geographical structure develops.

9. CONCLUSIONS

Under the criteria for our simulations, we found that the probability of near-fixation or fixation of an advantageous recessive gene is profoundly affected by the breeding structure of a population. When expressing the breeding structure in terms of neighbourhood size, there is a compelling inverse relationship between the fixation probability and N_e , the relationship being linear when N_e is transformed into logarithms.

Our results differ from that obtained by Maruyama (1971, 1972*a*, 1974) in an analytical and simulation study of the fixation probability of mutant genes in a subdivided population. He showed that population subdivision (or inbreeding) did *not* alter the fixation probability of a mutant gene if fitness were additive. The basis for the differences lies in the treatment of heterozygote fitness; Maruyama assumed intermediate heterozygote fitness, whereas we assumed the fitness of the heterozygote was identical to that of the dominant homozygote. To our knowledge, Maruyama's models are the only ones other than our own which explore the effect of inbreeding on either the probability of fixation or the mean fixation time.

The rate of gene substitution after a single immigration episode is profoundly affected by the pollen and seed dispersal schedules when heterozygote fitness is like that of the dominant homozygote. The rate of substitution of a dominant gene by a recessive gene relative to that obtained with panmixia is retarded by inbreeding, and decreasing the density of organisms competing for limiting resources.

Gene frequency heterogeneity among subpopulations arises as a by-product of gene flow restriction and is very pronounced in some cases. If one encountered the graded patchwork of gene frequencies seen above (fig. 4, IV) without knowledge of the breeding structure, it might seem that selection acting alone had given rise to a spatial pattern which varies in a continuous, one-to-one way with a heterogeneous environment. As emphasised by Wright (1951) and others more recently, environmental heterogeneity is not a prerequisite for pronounced gene frequency variance among subpopulations.

When immigration occurs in each generation, the pattern of gene receipt via pollen notably affects the rate of substitution if the novel gene was advantageous, and the equilibrium level of the novel gene were it detrimental. The more localised the pattern of receipt the greater the deviation from theory, which is expressed as slower rates of substitution, and lower equilibrium gene frequencies. The near-fixation time obtained in our experiments may be more than 30 per cent greater than that predicted by theory. As a corollary the equilibrium gene frequency in our experiments may be only one-third that predicted by theory. Clearly, we can no longer neglect the gene input pattern when considering the effect of immigration in plants.

The specific models presented here were not designed to correspond exactly to any natural populations. We have included four pollen and four seed dispersal schedules, three of which are realistic and one (random) which provides a standard for comparison. The general conclusions reached either were in accord with expectations derived intuitively or by analytical methods, or were novel in demonstrating the effect of variables heretofore not considered in conjunction with gene substitution. In any event, they could prove a useful guide in thinking about natural populations of plants and about the applicability of simple genetic formulae to them.

We endorse Bradshaw's (1972) admonition stated in an essay on the evolutionary consequences of being a plant. He wrote, "To understand what is *actually* happening in plant species, we need to assume very different premises from many of those currently exercising the minds of many population geneticists." Clearly, abandoning the notion of panmixia will be a major step in this regard.

Acknowledgements.—We are indebted to M. Nei and S. K. Jain for their critical reading of the manuscript and to C. Pankratz for running the simulations. The study was supported by National Science Foundation grant BMS 73-06687-A02 to D. A. L.

10. REFERENCES

- BRADSHAW, A. D. 1972. Some evolutionary consequences of being a plant. In Th. Dobzhansky, M. K. Hecht, and W. C. Steere (eds.), *Evolutionary Biology*, Vol. 5, pp. 21-44. Appleton-Century-Crofts, New York.
- CROW, J. F., AND MARUYAMA, T. 1971. The number of neutral alleles maintained in a finite geographically structured population. *Theoret. Pop. Biol.*, 2, 437-453.
- FISHER, R. A. 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- KIMURA, M. 1953. Stepping stone model of population. *Ann. Rep. Nat. Inst. Genetics*, 3, 63-65.
- KIMURA, M., AND MARUYAMA, T. 1971. Pattern of neutral polymorphism in a geographically structured population. *Genet. Res. Camb.*, 18, 125-133.
- KIMURA, M. AND OHTA, T. 1971. *Theoretical Aspects of Population Genetics*. Princeton Univ. Press, Princeton, New Jersey.
- KIMURA, M., AND WEISS, G. H. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, 49, 561-576.
- LEVIN, D. A., AND KERSTER, H. W. 1971. Neighbourhood structure in plants under diverse reproductive methods. *Amer. Nat.*, 105, 345-354.
- LEVIN, D. A., AND KERSTER, H. W. 1974. Gene flow in seed plants. In Th. Dobzhansky, M. T. Hecht, and Wm. C. Steere (eds.), *Evolutionary Biol.*, Vol. 7, pp. 139-220. Plenum Press, New York.
- MARUYAMA, T. 1970a. On the fixation probability of mutant genes in a subdivided population. *Genet. Res.*, 15, 221-225.
- MARUYAMA, T. 1970b. Effective number of alleles in a subdivided population. *Theoret. Pop. Biol.*, 1, 273-306.
- MARUYAMA, T. 1971. Speed of gene substitution in a geographically structured population. *Amer. Nat.*, 105, 253-265.
- MARUYAMA, T. 1972a. Some invariant properties of a geographically structured finite population: distribution of heterozygotes under irreversible mutation. *Genet. Res., Camb.*, 20, 141-149.
- MARUYAMA, T. 1972b. Rate of decrease of genetic variability in a two-dimensional continuous population of finite size. *Genetics*, 70, 639-651.
- MARUYAMA, T. 1974. A simple proof that certain quantities are independent of the geographical structure of a population. *Theoret. Pop. Biol.*, 5, 148-154.
- MARUYAMA, T., AND KIMURA, M. 1971. Some methods for treating continuous stochastic processes in population genetics. *Japan J. Genet.*, 46, 407-410.
- MATHER, K. 1969. Selection through competition. *Heredity*, 24, 529-540.
- NEI, M. 1971. Fertility excess necessary for gene substitution in regulated populations. *Genetics*, 68, 169-184.

- ROHLF, F. J., AND SCHNELL, G. D. 1971. An investigation of the isolation by distance model. *Amer. Nat.*, 105, 295-324.
- WORKMAN, P. L., AND NISWANDER, J. D. 1970. Population studies on south-western Indian tribes. II. Local differentiation in the Papago. *Amer. J. Human Genet.*, 22, 24-49.
- WRIGHT, S. 1931. Evolution in Mendelian populations. *Genetics*, 16, 97-159.
- WRIGHT, S. 1943. Isolation by distance. *Genetics*, 28, 114-138.
- WRIGHT, S. 1951. The genetical structure of populations. *Ann. Eugenics*, 15, 323-354.