Genetic Structures in Finite, Open-pollinated Plant Populations: A Model and its Application to Seed Orchards

Hans-Rolf Gregorius and Gerhard Müller
Lehrstuhl für Forstgenetik und Forstpflanzenzüchtung der Universität Göttingen (BRD)

Summary. A model has been constructed to investigate the consequences of the rate of self-fertilization, pollen-dispersal, population-size, and number of clones on the genetic structure of finite seed plant populations. Derivations have been performed for two different cases:

A) Parental genetic structure explicitly given: inferences for the expected genetic structure of the resulting seed population;

B) Extension of case A) to several non-overlapping generations.

If random cross-fertilization is assumed for case A) the genetic composition does not change and the genetic distance between the corresponding Hardy-Weinberg-structure and the expected offspring-structure is 0 if the rate of self-fertilization is equal to \( \frac{1}{N} \) (N = population-size); any deviation from \( \frac{1}{N} \) causes an increase in genetic distance.

In case B) the expected genetic structures have been derived for all generations and it was possible to establish a comparatively simple dependence on the coefficient of inbreeding. In addition the variance of the allele-frequency has been presented. All the above influential components can be summarized by a single quantity, called M. After proving that \( \frac{1}{N} \) can be conceived as the effective population-size, all the results obtained could be presented depending on this effective size and the average rate of self-fertilization only.

Applying the findings of the model to the situation realized approximately in a seed-orchard, the following statements can be made:

Case A) Again assuming random cross-fertilization, a deviation of the parental population from the corresponding Hardy-Weinberg-proportions can, with increasing rate of self-fertilization, be exceeded by the respective deviation of the seed population. Case B) The influence of pollen dispersal on the effective population size has been investigated, assuming no variation of the individual rates of self-fertilization, pollen and seed production within the population. Only extremely small differences between effective and actual population size were obtained, which indicates that the influence of pollen dispersal is of minor importance in this case. For different rates of self-fertilization, significant differences in the increments per generation for the coefficients of inbreeding, as well as the frequency of homozygotes, were obtained for the first generation only. Decreasing number of clones influences the rate of self-fertilization and the effective population size simultaneously by increasing the first and decreasing the latter. This is transferred to the coefficient of inbreeding, frequency of the homozygotes and the variance of the allele frequency by an increase of increments for all generations.

Introduction

Any kind of breeding system may be regarded as acting upon the genetic variability contained in a real or hypothetical base population, where the pattern of variability is completely determined by the set of corresponding genotypic frequencies - denoted as the genetic structure - which in turn fix the set of gene frequencies - denoted as the genetic composition. This variability may be reduced where the predominating components of the breeding system are, for example special kinds of selecting or drift in small populations, which at the extreme can cause a high amount of inbreeding and a loss of genes originally present in the base population. A phenomenon of this kind constitutes a risk, the importance of which depends heavily on the kind of organism considered. If, for instance, plant species are considered which occupy very heterogeneous environments or have long generation cycles, reduced genetic variability will imply low adaptability and restrains separate generations from genetic exchange. Situations like these are commonly met in forest tree breeding when establishing seed orchards.

The present paper treats especially the consequences implied for their genetic structures by characteristic components of the mating system found in seed orchards. The emphasis will be put on the role played by the rate of self-fertilization, mode and range of pollen dispersal, population size and number of clones. This set of parameters determines to a great extent the amount of inbreeding and kinship and deviations from random-mating.

Of course there are many other factors which influence the breeding system displayed in seed orchards, such as variation in flowering times, seed and pollen production. But too little is known about the genetic control of components like these, so that reasonable assumptions concerning implications for the genetic structure can not be made.
An investigation of genetic structures in seed orchards can be performed from two different points of view:

A) Representations refer exclusively to the expected genetic structure within the seed production of a given seed orchard. The aim is to state how far the parental structure is reflected by the seed population.

B) Assuming that populations derived from the seed production of a seed orchard are used to establish future generations, considerations concerning the risk caused by reduction of genetic variability obtain increasing importance. To demonstrate this expected risk, investigations are extended to an arbitrary number of generations.

1. Theoretical derivations

1.1. Preliminary remark

As already mentioned, an investigation of genotype-frequencies (genetic structure) and gene-frequencies (genetic composition) in finite populations can be based on two different prerequisites: firstly, one may be interested in finding the expected genetic structure of the progeny of a parental population, whose members are explicitly described by their genotypes, thus considering merely the realized genetic structures of the parental generation and not their probabilities of occurrence; secondly, no exact knowledge about the actual genetic structure of the parental population is available, and therefore some assumptions concerning the probability distribution of this structure - which then has to be regarded as a random variable - have to be made.

Obviously, the expected genetic structure of the progeny in the latter case is obtained from the first case by taking the expectation with respect to the probability-distribution of the parental structure.

A clear distinction between these two situations is necessary when applying theoretically obtained results to actually existing populations, especially seed-plant populations, which are able to produce a comparatively large number of offspring in the form of seeds. Such a seed population may be accepted as representing sufficiently well the expected genetic structure of the progeny of a specified set of parents. It will be very interesting to find the degree of concordance between the genetic structure and the composition of the parental and offspring generations.

A comparison of this kind can not be performed if the information about the genetic structure of the parental population is not more explicitly given, but rather refers to the expected structure. In this case all the genotype- and gene-frequencies of the two generations can not be regarded as actually existing genetic structures in the sense mentioned above; they merely denote the probability of a single genotype or gene occurring in the population. The same concept underlies the definition of the coefficient of inbreeding and kinship, both of which therefore have to be treated within the scope of this concept. One of the aims of this paper shall be to demonstrate the relationship between the coefficients of inbreeding and kinship and the genotypic probability structure for the model applied.

1.2. The model

Populations are assumed to have finite size $N$ and consist of monoecious, diploid seed-plants. The individuals are distributed over their habitat according to a specified pattern, so that each plant can be identified by its location. The following conditions shall be realized approximately: no immigration, mutation, gametic and zygotic selection; no genetically caused variation of the amounts of pollen- and seed-production, of the types of pollen-dispersal and of the rates of self-fertilization among the plants. The overall pollen production is sufficiently large to pollinate all eggs present in the population. All plants flower at the same time.

Where several generations are considered, these are assumed to be discrete and non-overlapping, and each offspring generation is obtained by taking a random sample from the overall seed-production of the parental generation; the assignment of the individuals of this sample to their locations is at random.

Because we intend to describe the influence of the type of pollen-dispersal and the rate of self-fertilization on the genetic structure, respectively on the average coefficient of inbreeding and kinship of a population, all the following derivations assume that the probabilities of mating for all pairs of individuals are given as functions of their locations, and that the seed-production of each (mother-) plant can be split into one part resulting from self-fertilization and the remaining part resulting from cross-fertilization.

One autosomal gene locus with an arbitrary number of alleles $A_i$ is considered.