

Stochastic Simulation of Population Management Strategies for Tree Breeding: a New Decision-Support Tool for Personal Computers

By T. J. MULLIN¹) and Y. S. PARK²)

(Received 1st February 1995)

Abstract

This paper describes the development and application of a decision-support software tool, called POPSIM, for the simulation and analysis of tree breeding population management strategies on modern personal computers. Genetic effects are simulated according to an additive-dominance-epistasis model for a base population of trees that are mated according to a user-defined breeding plan. Individual progeny from the breeding plan are generated, simulating the recombination of genetic effects and the random environmental effects during field testing. Selection is carried out among the test progeny to assemble various kinds of production populations for deployment of improved stock, and to advance the breeding population to the next cycle of breeding. Genetic variances are calculated directly from the simulated data and reported for each generation, together with the average level of inbreeding, effective population size, and the cumulative genetic gain realized by each type of production population. The theoretical basis for the simulation is given, together with a description of the options currently available to the user. An example is provided to illustrate the statistics generated by the program, and demonstrate the flexibility of the tool to handle diverse approaches to management of genetic gain and diversity.

Key words: stochastic models, breeding strategy, genetic diversity, effective population size, genetic gain, computer simulation, sublines, assortative mating, Monte Carlo methods.

FDC: 165.4—015.5; 165.5.

Introduction

The prediction of genetic gain from one round of selection in a tree improvement program is a relatively straightforward calculation, requiring only appropriate estimates of heritability, selection intensity and the phenotypic variance of the selection units. Several assumptions are usually made regarding the normal distribution of variances, and the absence of inbreeding. There are several papers in the literature demonstrating how this basic calculation can be applied to a wide variety of situations found in applied breeding programs (e.g., COTTERILL, 1986; COTTERILL and JACKSON, 1989; MATHESON and LINDGREN, 1985; NAMKOONG *et al.*, 1966; SHELBOURNE, 1969, 1991).

Although such gain calculations are useful, and indeed essential for the evaluation of the economic returns from breeding, comparisons of breeding strategies based only on gain estimates from one round of selection are generally inadequate. More often the objective of an improvement program will be to maximize gain over time after several cycles of breeding. This objective will often be further constrained by the loss of genetic variance, accumulation of inbreeding, effective population size, and ultimately cost (MULLIN, 1994). While predicting the gain from a single round of selection may be relatively straightforward, resulting changes in population structure are less

readily described, yet these have a profound impact on the outcome of future breeding and selection. Algebraic expressions of such sequences of breeding and selection are, if not impossible, extremely complex.

Using computers to simulate genetic systems is not really new; the technique has been reported in the literature for at least 35 years (FRASER, 1957). Even so, the application of simulation to investigate the outcome of breeding efforts with forest trees and other plant crops has been very limited, although recently several authors have incorporated simulation techniques in their comparisons of specific breeding strategies. In her analysis of the effects of positive assortative mating on cumulative gains and inbreeding, MAHALOVICH (1990) used a gene effects model, including additive and partial dominance effects at 50 loci. KING and JOHNSON (1991, 1993) demonstrated the flexibility of computer simulation methods when they used a parameter based stochastic model for additive genetic effects to consider gains and effective population size for 5 mating schemes for advanced-generation breeding of *Pinus radiata* D. DON in New Zealand. To date, the most significant progress in the simulation of artificial selection systems has been made by domestic animal breeders (e.g., DE ROO, 1987; DE VRIES *et al.*, 1990; MEYER and SMITH 1990; RUANE and THOMPSON, 1991).

Simulation models may be classified into 2 broad categories: (i) "stochastic" models, often referred to as Monte Carlo simulation, where random processes are mimicked by the generation of pseudo-random numbers, and (ii) "deterministic" models, where the outcome of processes is predicted through an algebraic mechanism. In genetic systems, the random recombination of alleles through sexual reproduction is a process that lends itself to simulation by stochastic methods. By repeating the simulation through several iterations, one can predict the mean outcome of stochastic events and describe the variance of outcomes about this mean.

Using computers to simulate stochastic processes, the investigator can set very broad limits on the complexity of model parameters and interactions, without resorting to highly theoretical or abstract mathematics (LEVIN, 1969). Provided that genetic effects are simulated in a way that makes biological sense, the simulation of complex breeding plans requires elementary mathematics and is primarily an exercise in logic and efficient programming. A deterministic model of the same system will, at best, require extremely complex mathematics and will likely require simplification to the point that the model is no longer a realistic analogue (KEMPTHORNE, 1988).

A typical breeding plan describes numerous decision steps and employs a variety of limits, many of them somewhat arbitrary, to maintain diversity in the population while achieving as much gain as possible in each generation. Comparisons of breeding strategies based on gain estimates from one round of selection are generally inadequate, when the objective is usually the optimization of gain over time for several breeding cycles. Criteria for evaluation of breeding strategies are rarely limited to genetic gain alone, but will most likely include limits for reduction of genetic variance, accumulation of inbreeding,

¹) Genesis Forest Science Canada Inc., P.O. Box 1321, Truro, N.S., Canada B2N 5N2. e-mail: genesis@fox.nstn.ns.ca

²) Canadian Forest Service – Maritimes Region, P.O. Box 4000, Fredericton, N.B., Canada E3B 5P7. e-mail: ypark@fcmr.forestry.ca

effective population size, and ultimately cost (MULLIN, 1994). Simulation techniques can make such complex comparisons relatively straightforward.

Previous simulation tools developed for tree improvement applications have been limited to specific problems and were designed for use on large, main-frame computers. This paper describes the design and application of a second-generation software tool, called POPSIM, for the simulation and analysis of tree breeding population management strategies on modern personal computers. Our objective was to provide breeders with a generalized decision-support tool to critically examine multi-generation breeding plans. The theoretical basis for the simulation is given, together with a description of the options currently available to the user. Examples are provided that illustrate the statistics generated by the program, and demonstrate the flexibility of the tool to handle diverse approaches to management of genetic gain and diversity.

General description of the POPSIM simulation software

The general approach used by the POPSIM simulator is to generate genetic and environmental effects for a base population of trees that are mated according to a user-defined breeding plan. Individual progeny from the breeding plan are generated, simulating the recombination of genetic effects and the random environmental effects during field testing. Selection is carried out among the test progeny to assemble various kinds of production populations for deployment of improved stock, and to advance the breeding population to the next cycle of breeding. Genetic variances are calculated directly from the simulated data and reported for each generation, together with the average level of inbreeding and effective population size. The cumulative genetic gain realized by each type of production population is presented as a percentage of the trait mean in the base breeding population.

Defining the phenotype and genotype of simulated trees

In this simulation, each tree's phenotype is considered to be the sum of independent genetic and environmental effects:

$$P = A + D + I + E$$

where

P is the phenotypic, or observed value for a given trait or an index of traits;

A is the additive genetic effect of substituting one allele at a locus for another;

D is the dominance genetic effect of within-locus allele effects which remain after subtracting additive effects, i.e., the intra-locus interaction between alleles;

I is the epistatic genetic effect which arises from interactions among all alleles which affect expression of the trait; and,

E is the environmental effect.

The total phenotypic variation in a population is thus described as the sum of independent variances for each of these effects:

$$\sigma_p^2 = \sigma_A^2 + \sigma_D^2 + \sigma_I^2 + \sigma_E^2$$

Simulating a base population of N trees is a matter of describing the mean and variance of each of these components in the population, and sampling from each distribution to produce the effects that together define the phenotype of each individual. If we assume that each of these effects is normally distributed, the random sampling from the distribution takes the form:

$$x_{ij} = \mu_j + r\sqrt{\sigma_{x_j}^2}$$

where x_{ij} is a normal variate for the j th effect of the i th tree, with mean of μ_j and variance $\sigma_{x_j}^2$, and r is a normal deviate drawn from a distribution of random numbers with mean of 0 and variance of 1. From the generated data for N trees in the population, we can readily calculate estimates of the population mean (\bar{X}_j) and variance ($\hat{\sigma}_{x_j}^2$) for each of the genetic and environmental effects, using standard formulae:

$$\bar{X}_j = \frac{\sum_i x_{ij}}{N} \quad \text{and} \quad \hat{\sigma}_{x_j}^2 = \frac{\sum_i x_{ij}^2 - \left(\sum_i x_{ij}\right)^2 / N}{N-1}$$

Simulating genetic recombination in offspring requires a description of the among-family and within-family inheritance for each genetic effect, as detailed below.

1. Additive effects

For additive genetic effects, the among-family component is the average additive effect of the female (f) and male (m) parents, i.e., the mid-parent value, and the additive effect for each offspring (A'_i) may be generated around this mean with a variance equivalent to the within-family portion of additive variance, $\hat{\sigma}_A^2$, adjusted by the expected reduction due to inbreeding of the parents (DEMPLE, 1990):

$$A'_i = \frac{A_f + A_m}{2} + r\sqrt{\frac{1 - \frac{1}{2}(F_f + F_m)}{2} \hat{\sigma}_A^2}$$

where F_f and F_m are the coefficients of inbreeding for the female and male parents, respectively.

2. Dominance effects

The among-family dominance effect for a given female-male parent combination is drawn randomly from a normal distribution with mean equivalent to the dominance effect in the parent population \bar{D} and a variance of $\frac{1}{4}\hat{\sigma}_D^2$. The dominance effects for individuals within this family (D'_i) are then drawn randomly from a distribution with mean equal to the within-family dominance effect and variance $\frac{3}{4}\hat{\sigma}_D^2$. The dominance effect is corrected to account for inbreeding depression, using the regression of phenotypic value on the inbreeding coefficient (DE BOER and VAN ARENDONK, 1992; BORRALHO, 1994):

$$D'_i = \bar{D} + r_{fm}\sqrt{\frac{1}{4}\hat{\sigma}_D^2} + r_i\sqrt{\frac{3}{4}\hat{\sigma}_D^2} + bF_{fm}\hat{\sigma}_P$$

where r_{fm} is a random normal deviate drawn for each given fm family combination, r_i is a random normal deviate for each individual in the family, b is a regression coefficient expressing the reduction in phenotypic values in units of phenotypic standard deviation ($\hat{\sigma}_P$), and F_{fm} is the inbreeding coefficient of the family members. The regression coefficient may be calculated from experimental data by using a linear regression of inbreeding coefficients on corresponding standard deviation units of phenotypic values. For example, inbreeding depression on the mean of selfed progeny is often 15% to 30% lower than outcrossed progeny (GRIFFIN and COTTERILL, 1988; PARK and FOWLER, 1984). If the reduction in the mean of selfed progeny ($F=0.5$) is one standard deviation below the phenotypic mean, the regression coefficient will be -2.0 .

3. Epistasis effects

All progeny epistatic effects (I'_i) are considered to be within-family, with mean and variance equal to that of the parent population. While the generation of the effect resembles the method used to assign random environmental effects, the epistasis effect remains constant for all individuals cloned from the same genotype:

$$I'_i = \bar{I} + r\sqrt{\hat{\sigma}_i^2}$$

4. Environmental effects

The simulator produces all tree populations in the same "environment", so that environmental effect (E'_{ik}) for the k th ramet of the i th genotype is drawn from a distribution with the same mean and variance as that specified for the original population:

$$E'_{ik} = \bar{E} + r\sqrt{\hat{\sigma}_E^2}$$

Inbreeding and effective population size

Changes in genetic diversity of populations are represented by the average inbreeding and estimates of effective population size. The inbreeding coefficient is the probability that 2 genes in an individual are identical by descent and describes the proportion of homozygotic loci in excess of that found in an infinitely large, random-mating population. The inbreeding coefficient for each tree in the population is determined by analyzing the pedigree of parent trees to determine the degree of common ancestry. The calculation is fairly simple in small pedigrees, but quickly becomes very cumbersome in large, complex pedigrees. Fortunately, techniques for pedigree analysis have made tremendous advances and efficient algorithms have become available in just the past few years that make the calculation of inbreeding coefficients straightforward (e.g., MEUWISSEN and LUO, 1992; TIER, 1990). The average inbreeding for the population is calculated simply as the average inbreeding coefficient for all trees.

In POPSIM, 2 approaches are used to determine the effective size of breeding and seed orchard populations, both based on the accumulation of gene correlations. The "inbreeding effective size" is defined by FALCONER (1981) as the number of individuals that would give rise to the observed rate of inbreeding, if they bred in the manner of the idealized population. In POPSIM, we estimate the inbreeding effective population size, N_e , from the rate at which inbreeding accumulates, ΔF . FALCONER (1981: equation 3.12) gives the inbreeding coefficient in any generation t , referring to the base population at t_0 :

$$F_t = 1 - (1 - \Delta F)^t$$

By rearranging this formula, we can calculate the average rate of change in inbreeding at a given generation, when the inbreeding coefficient is available from the pedigree record:

$$\Delta F = 1 - (1 - F_t)^{1/t}$$

FALCONER (1981: equation 4.1) gives the rate of inbreeding as a function of inbreeding effective size as:

$$\Delta F = \frac{1}{2N_e}$$

By rearrangement of this equation and substituting the previous equation for ΔF , the calculation of inbreeding effective size is:

$$N_e = \frac{1}{2(1 - (1 - F_t)^{1/t})}$$

Note that the inbreeding effective size of the base population is undefined and, if no inbreeding has accumulated, remains infinitely large in subsequent generations.

The concept of inbreeding effective size was developed originally with random-mating, unstructured populations in mind and is not readily expanded to subdivided populations due to differences in the accumulation of gene correlations among and within breeding groups. When the subdivision is complete with no genetic exchange, the correlation of genes

from random individuals from different groups is zero. In this case, the N_e of the population is calculated as the N_e within groups, multiplied by the total number of groups, although this probably overestimates the actual inbreeding effective size (CHESSER *et al.*, 1993).

A second measure of effective population size calculated by the program is based on the concept of coancestry, which is defined as the probability that genes sampled from parents will be identical by descent and equivalent to the coefficient of inbreeding that would occur *after* mating. The accumulation of gene correlations is altered when parents are established in a seed orchard and progeny result from matings among individuals from different breeding groups, so it makes sense to evaluate the size of the orchard based on the coancestry of the parents, rather than their accumulated inbreeding. The average inbreeding coefficient of the progeny (F_{t+1}) from all possible crosses between parents, including self matings, is equivalent to the average coefficient of coancestry for the parents (f_t). Using a similar derivation as used for inbreeding effective size (N_e) we can then calculate the number of individuals that would give rise to the calculated average coefficient of inbreeding in the next generation if bred in the manner of an idealized population (N_s):

$$N_s = \frac{1}{2f_t}$$

This is similar to an effective size parameter proposed by BURROWS (1984) based on average pair-wise coancestry; whereas BURROWS excluded self mating, we have included all possible pair-wise matings, including selfs. LINDGREN has suggested calling this parameter "status effective number," referring to the status of a population at a particular moment (Dr. DAG LINDGREN, pers. comm.) and has recommended its suitability as a measure of diversity in *any* population (not just seed orchards) since its value is not affected by population structure, it is never undefined, and is always less-than or equal to the census number, N . In POPSIM, effective size of orchard populations is always expressed as N_s , although the user is given the option to calculate the effective size of breeding populations as either N_s or the more traditional N_e .

Breeding and selection plans

The remainder of the simulation logic is a matter of mimicking the mating designs and evaluation procedures used in a given breeding plan and describing the distribution of effects in the resulting populations. In its current version, POPSIM can simulate several operations and approaches used in typical breeding plans (e.g., FOWLER, 1986; PARK *et al.*, 1993), as outlined below.

1. Sublining of the breeding population

Breeding populations may be subdivided into "subline" groups, with full-sib mating for generation advancement, i.e., to produce progeny for selection of the next-generation breeding population, restricted to crosses among parents within sublines (VAN BUIJTENEN and LOWE, 1979).

Since the accumulation of inbreeding is restricted to within sublines, crosses among orchard parents selected from different sublines will always be totally outcrossed. Subdividing the population also lessens the impact of genetic drift, as sampling of gametes within each subline is independent of that in the remainder of the population and different alleles become fixed in the different groups (CABALLERO, 1994; LACY, 1987).

When sublining is requested, parent trees are assigned to sublines in the base generation, *after* production of all seed

orchard populations, but *before* any breeding is carried out for generation advancement. In order to ensure genetic equality, the assignment of trees to sublimes is normally random; however, when positive assortative mating is requested, the user can elect to first rank the parents by breeding value and then assign them uniformly across all sublimes. Once formed, the sublimes are managed independently and maintained at the same size over subsequent generations.

2. Polycross testing for ranking of parents on breeding values

If a breeding strategy calls for selection of parents for a "tested" orchard, or if assortative mating is to be used for generation advancement, the program will simulate a polycross test of all parents in the breeding population. The user declares the number of male parents represented in the pollen "mix" and the number of progeny to be tested for each tree in the breeding population. The pollen mix remains constant through all generations of breeding and is drawn from the same population distribution as the base population. Male parents are drawn at random from the mix with equal probability for each offspring, and genetic and environmental effects are generated as described earlier.

The family means from the polycross test estimate the breeding value of each tree in the population. Trees in each subline are ranked for selection as a tested orchard population, or for assortative mating.

3. Full-sib mating for generation advancement

The simulation assumes that all breeding for generation advancement will be based on full-sib crosses, using the same mating design across all sublimes and all generations. Considerable flexibility is permitted in the description of the mating design, as virtually any number and combination of parents may be used, subject only to the array storage limitations imposed in the user's copy of the program. The simulator uses a mating design "template" that resembles the grid representation used by many authors to describe proposed mating designs. The user will be warned if any parents are omitted from the proposed test design, although this may be desired in some strategies to exclude the lowest ranking parents in assortative mating schemes. Warnings are also issued if attempts are made to produce duplicate or reciprocal crosses, as the simulation will not generate correct genetic effects for such crosses.

4. Positive assortative mating for generation advancement

When the user requests assortative mating, the parents in each subline are first sorted in descending order based on breeding value estimated from the simulated polycross test. The mating design template can then be used to increase the number of crosses among the better parents, or to cross parents with similar or dissimilar breeding values, or both.

5. Field progeny testing with specified family size, and the option for clonal replication

The program assumes that all progeny testing is performed in the same environment, with the variance of environmental effects identical to that of the base population, and that all selection for generation advancement will be carried out in these test plantations. The user declares the number of test genotypes to be evaluated. The user may also elect to replicate test genotypes by cloning, in order to increase the efficiency of individual genotype selection (LIBBY, 1964; MULLIN and PARK, 1992; SHAW and HOOD, 1985), and for selection of genotypes based on clone means for deployment of clonal mixtures. In

this case, clonal replicates (ramets) will be generated with identical genotypes, but independent environmental effects.

6. Selection options

The program simulates three selection methods, and different methods may be applied to the selection of the next breeding population and each of the various production populations. The simplest of these is referred to as "mass selection", where selection is based solely on the individual tree's phenotype or, in the case where clonal replication is used in the progeny tests, on the clone mean performance. So-called "2-stage" selection is performed by identifying a fixed number of best phenotypes from each of the best families. "Combined-index selection" is based on an individual index value (*CI*) that weights family and individual performance by their respective heritabilities (FALCONER, 1981):

$$CI = h_w^2 (\bar{X}_{ifm} - \bar{X}_{fm}) + h_F^2 (\bar{X}_{fm} - \bar{X})$$

where

\bar{X}_{ifm} is the mean (or individual observation) for the *i*th genotype in the *f*mth family;

\bar{X}_{fm} is the mean performance of the *f*mth full-sib family of size *n*;

\bar{X} is the overall mean;

h_F^2 is the heritability of full-sib family means, and calculated as

$$h_F^2 = h^2 \frac{1 + (n-1)r}{1 + (n-1)t}$$

h_w^2 is the heritability of within-family deviations, calculated as

$$h_w^2 = h^2 \frac{(1-r)}{(1-t)}$$

h^2 is the narrow-sense heritability of individual observations;

r is the genotypic intraclass correlation, which is $1/2$ for full-sib families;

t is the phenotypic intraclass correlation; and

n is the family size.

BAKER (1986) showed that for families of large size, $t = rh^2$, so in the case of full-sib families the relevant heritability formulae become:

$$h_F^2 = h^2 \frac{1 + \frac{1}{2}(n-1)}{1 + \frac{1}{2}(n-1)h^2} \quad \text{and} \quad h_w^2 = h^2 \frac{\frac{1}{2}}{1 - \frac{1}{2}h^2}$$

7. Restrictions on relatives

Restrictions on the use of relatives may be imposed both during selection of the breeding population, and assembly of the various kinds of production populations. Candidates for selection may be compared with those previously selected and the number of full-sib and half-sib relatives can be limited independently, as specified by the user. If it happens that the user's restrictions can not be satisfied at a particular point in the simulation, the restrictions are relaxed only as much as necessary to complete the selection process, and a warning is issued to the user. The restriction limits are returned to those originally specified by the user.

8. Seed orchard populations

Seed orchard populations for production of improved seeds are assembled by selecting trees from each subline; the total number of trees in the orchard must be divisible by the number

of sublimes. As breeding occurs within sublimes, selections from different sublimes will not be related. If more than one tree is selected from each subline, coancestry restrictions may be applied to control relatedness of trees in the orchard. The average inbreeding coefficient of the selected orchard parents is given. The effective population size and genetic gain from the orchard are also reported, assuming equal contributions to the gamete pool and random mating among the orchard parents.

9. Deployment of full-sib family mixtures

The program also allows the user to select a mixture of the best full-sib families from those established in the progeny test. Such a mixture could then be reproduced by repeating the cross, and perhaps by using vegetative propagation to "bulk-up" sufficient numbers for field deployment, as suggested by MULLIN and PARK (1992). As for seed orchards, the selected mixture must contain an equal number of families from each subline. As all of tested crosses are made within sublimes, inbreeding will accumulate within the families, although the maximum number of related families selected can be specified by the user. The program does not, at this point, permit crosses made between sublimes for the purpose of testing outcrossed families.

10. Deployment of clonal mixtures

The user may also choose to simulate the selection of a clonal mixture, consisting of an equal number of genotypes from each subline, of those included in the progeny test. If clonal replication is used as part of the progeny test, the selection will be carried out on clone means, otherwise selection is based on individual phenotypes. Again, these clones are selected from among crosses made within sublimes where inbreeding will accumulate. The user may specify the maximum number of related clones to be included in the mixture.

Using the POPSIM Program

Although POPSIM was designed to be "user friendly", a detailed manual is available that explains installation and the various options (MULLIN and PARK, 1995). The program is written in FORTRAN and compiled with the Microsoft® FORTRAN PowerStation (Microsoft, 1993) for execution as a 32-bit DOS-extended program under the DOS or Windows operating systems. The computer must have a 32-bit processor, i.e., an Intel 80386 or better, and a minimum of 4Mb of RAM is recommended.

First-time users and those who are only running a single scenario will probably enter their simulation control parameters through the user interface prompts. An opportunity is given to confirm and edit choices before the simulation is actually started. Inappropriate responses and those which exceed the capability of the program will normally be trapped and hints are given on how to make a legal response. The selected parameters are stored in a so-called "response file" and the simulation started once all of the scenarios have been described. Experienced users will normally prefer to use a text editor or word processor to create their own response file directly, especially if multiple scenarios are to be run while the machine is unattended.

The simulation uses pseudo-random number generators to drive the many stochastic events over uniform and normal distributions; ran2 and gasdev, respectively, by PRESS *et al.* (1992). A valid integer must be specified as a "seed" number to start the pseudo-random sequence. Repeating a simulation

with the same seed number on this, or any other Intel-based personal computer, will generate identical results.

The mating design routine in POPSIM uses a "template" representing all of the crosses to be carried out within each subline. The trees in the design template are numbered 1 through however many trees are in each subline. If assortative mating is requested, the trees in each subline are sorted from best to worst, i.e., "positive assortative mating", using the estimates of breeding value generated by the polycross. In this case, the parent numbers on the design template correspond to the ranking of their breeding values; otherwise, the trees are assigned to the template in random order. Examples of templates for both non-assortative (random order) and assortative mating are given in *figure 1*. There is flexibility permitted in the declaration of mating design templates so that virtually any mating design can be accommodated, including unbalanced designs, within the array storage limitations established during compilation of the program, although the same template must be applied to all sublimes in the breeding population.

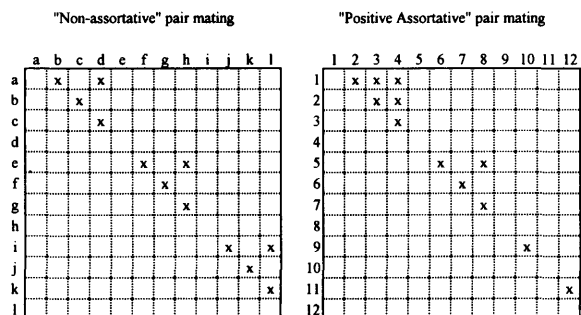


Figure 1. - Sample mating design templates for "non-assortative" and "assortative" mating designs.

Progress will be reported on the user's screen, with the level of detail requested. When all simulations are completed, the output is directed to the file or printer specified, and all temporary disk files are deleted. When output is directed to a disk file, the user must use a word processor or text editor to view the results. POPSIM stores the output from multiple scenarios in separate output files; that way, if the system is halted for any reason, the user can resume execution after the last scenario completed. POPSIM will ask what scenario number to start at, so it is not necessary to modify the response file to resume execution if the initial run was interrupted. In addition, POPSIM can be directed to write output data to a tab-delimited text file, for import into a spreadsheet or statistical analysis package.

The mean and variance of each genetic and environmental effect for the breeding population will be very close to those specified by the user for the base population. Similar statistics are reported for each of the production populations requested: untested and tested orchards, selected full-sib families and tested clone mixtures. The average inbreeding coefficient, inbreeding effective population size, and cumulative genetic gain in production populations are also reported.

Simulations of complex breeding plans with large numbers of test progeny and spanning several generations require large amounts of computing time. Unlike a deterministic model that relies on the solution of a series of equations to produce a "fixed" prediction, a stochastic model requires the generation of simulated data, including genetic and environmental effects,

for each and every tree produced by the breeding plan. These data are evaluated using the same ranking, sorting and pedigree analysis techniques that would be employed in the evaluation of real-life data. Furthermore, since random variation is a key feature of the technique, the simulation will normally be repeated several times with reporting of average results together with standard errors. Even on today's powerful PCs, the assessment of a given scenario will normally be measured in terms of minutes or even hours, as opposed to milliseconds.

Example

Use of the POPSIM software can best be demonstrated with examples, and we will consider 2 test scenarios here. We will use these to illustrate the program control and printed results from POPSIM, and how the program can be used to compare 2 breeding strategies.

In this example, we consider a hypothetical case where a breeding program is being designed to improve a trait with moderate heritability, say 0.20. We assume that the mean value of the trait under improvement is 100 "units." The base population is composed of 160 unrelated trees, and the population variance structure and rate of inbreeding depression for the trait under improvement are as presented in *table 1*. Control data were prepared to simulate 2 contrasting strategies, based roughly on the breeding strategies presented in recent papers by WEIR and TODD (1994) and PARK *et al.* (1993).

Table 1. – Population and selection parameters for 2 sample breeding strategies.

Parameter	Strategy	
	1	2
Population		
Number of trees in breeding population	160	160
Number (size) of sublimes	40 (4)	8 (20)
Effective population size	N_e	N_e
Mean value of trait in base population	100	100
Standardized rate of inbreeding depression	-2.0	-2.0
Additive genetic variance	100	100
Dominance genetic variance	25	25
Epistatic genetic variance	75	75
Environmental variance	300	300
Selection		
Polycross test family size (offspring per parent)	NIL	60
Assortment of trees in sublimes	NIL	positive uniform
Initial assignment to sublimes	random	uniform
Number of full-sib crosses in progeny test	240	192
Progeny test family size (offspring per cross)	120	100
Breeding population selection method	combined index	combined index
Maximum number selections per FS family	1	2
Restriction on number of half-sib relatives	NIL	3
Seed orchard selection method	combined index	polycross
Maximum number selections per FS family	1	1
Restriction on number of half-sib relatives	NIL	3

The first strategy would see the breeding population divided into very small sublimes of only 4 trees each, while the other would use much larger sublimes of 20 trees. In each generation, a seed orchard consisting of 40 parents, one from each subline, is selected phenotypically from among trees in the breeding population. No polycross is used with the 4-tree sublimes and parents are not sorted before mating according to a complete half-diallel (*Figure 2* – Strategy #1), producing a total of 240 full-sib families for progeny testing. The tests are established with 120 offspring per family. Selection of 160 trees, 4 per subline, is performed using the combined-index of family and individual performance, with no controls on inter-tree relatedness within sublimes.

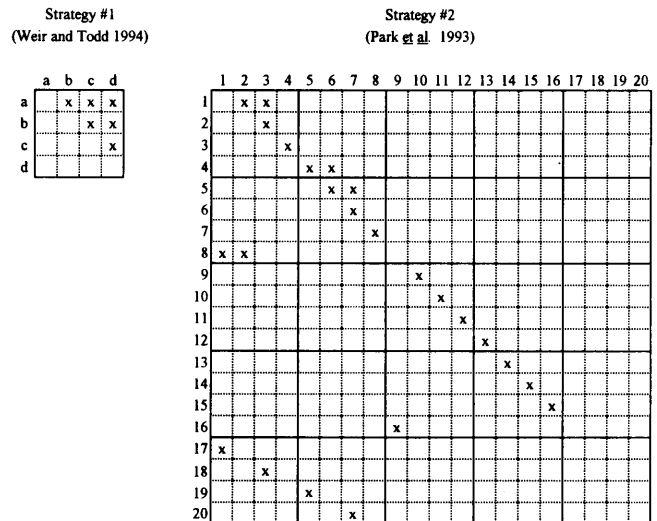


Figure 2. – Mating designs used for each of the 2 sample breeding plans.

In the second strategy, the population is subdivided into 8 sublimes of 20 trees. A polycross is performed using a 20-tree pollen mix, producing 60 offspring per parent. Breeding values estimated from the polycross are used to selected 40 trees that are grafted into a seed orchard; 5 trees are selected for the orchard from each subline, with a maximum of 1 tree from any given full-sib family and up to 3 half-sib relatives. The breeding values from the polycross are also used to rank the trees for positive assortative mating, using an unbalanced mating design where the best trees are used in more crosses than the lower ranked trees (*Figure 2* – Strategy #2). This mating design produces a total of 192 crosses among trees in the breeding population. In order to make the combined polycross and progeny testing effort equivalent to the 28,800 offspring tested in the first strategy, the family size for the progeny test in the second strategy was reduced to 100 trees. The combined-index was again used for advance-generation selection, although limits were imposed of 2 selections per full-sib family, with up to 3 half-sib relatives.

Program control and printed results

As mentioned earlier, the necessary information to run the simulator can be entered interactively, as the program will guide the user through the various choices with appropriate prompts; however, for this illustration we modified a program-generated response file which gives descriptive information about the data presented on each line, a portion of which is reproduced in *figure 3*. The first block in the response file (to the line of asterisks) contains general program control information, such as the number of scenarios to be run, and locations of temporary data storage and output devices. Then for each scenario, a separate block describes simulator settings, structure of the base population, testing and mating designs, selection methods for generation advancement and deployment options. In response files that follow the format of those generated by the program, the portion of the line to the left of the equals sign contains the data that are actually read and used by the program, while the portion to the right gives the corresponding variable name and descriptive information that may be useful when making modifications.

The output consists of a page summarizing the control parameters for the particular scenario (*figure 4* gives the first page for the first strategy), followed by a page of results for

```

2 = NSCEN - number of scenarios to be run
E:\PROGENY = PROGENY - path\filename for temporary progeny data
E:\SUMMARY = SUMMARY - path\filename for temporary summary data
0 = IPRN - printer output (0=No, 1=LPT1, 2=LPT2)
1 = IFILE - file output (0=No, 1=Summaries, 2=Series, 3=Both)
EXAMPLE = ROOT - output file ROOT name
1 = IS - detailed timing reports (0=No, 1=Yes)
*****
1 = KSCEN - scenario number
45678 = IISEED - random seed number
5 = NGEN - number of generations
50 = NITER - number of simulation iterations
160 = NT - number of trees in breeding pop'n
40 = NS - number of sblines
1 = ISIZE - effective size: 0 = Ne, 1 = Ns
100.00000 = EFFBAR(4) - mean of trait in base pop'n
-2.00000 = BID - rate of inbreeding depression
100.00000 = EFFVAR(1) - additive variance
25.00000 = EFFVAR(2) - dominance variance
75.00000 = EFFVAR(3) - epistatic variance
300.00000 = EFFVAR(4) - environmental variance
0 = NP - number of pollen parents
0 = NX - number of polycross progeny per tree
0 = ISORT - PAM: 0=none, 1=random, 2=uniform
6 = NXS - number of controlled crosses per subline
1 2 = MATEF and MATEM for cross # 1
1 3 = MATEF and MATEM for cross # 2
1 4 = MATEF and MATEM for cross # 3
2 3 = MATEF and MATEM for cross # 4
2 4 = MATEF and MATEM for cross # 5
3 4 = MATEF and MATEM for cross # 6
120 = NC - number of genotypes tested per cross
1 = NR - number of ramets tested per genotype
3 = IBPSEL - selection method for breeding pop'ns
32000 = NBPFS - (maximum) number per FS cross for BP selection
32000 = NBPFS - maximum number HS relatives for BP selection
32000 = NBPCC - maximum number progeny/parent for BP selection
1 = IUNTST - select untested orchard = 1, else 0
0 = IPXTST - select PX-tested orchard = 1, else 0
3 = ISOSEL - selection method for untested orchards
40 = NSSO - total number of selections in seed orchards
1 = NSOFS - maximum number per FS cross for SO selection
0 = NSOHS - maximum number HS relatives for SO selection
32000 = NSOPC - maximum number progeny/parent for SO selection
0 = IFAMSEL - select FS family mixture = 1, else 0
0 = NFAM - number of families in deployed mixture
0 = NFAMHS - maximum number HS relatives in mixture
0 = NFAMPC - maximum number families with common parent
0 = ICLNSEL - select clonal mixture = 1, else 0
0 = NCLN - number of clones in deployed mixture
0 = NCLNFS - maximum number clones from any given cross
0 = NCLNHS - maximum number HS relatives in mixture
0 = NCLNPC - maximum number clones with common parent
*****

```

Figure 3. - A portion of the response file generated by POPSIM and used to describe the first example scenario.

each generation. The simulated mean and variance of each genetic and environmental effect for the breeding population in generation 0 will, over 50 iterations, be very close to those specified by the user for the base population. That is, the mean of the trait in the base population is 100 with environmental variance of 300, and variances for additive, dominance and epistasis effects being 100, 25 and 75, respectively. The simulation was run under MS-DOS 6.0 on a 486/50DX2 personal computer with 8Mb of RAM, and using a RAM-disk for interim storage of progeny-test data. Completion of the 50 iterations for this first scenario required about 2 hours.

A sample generation summary page is illustrated in figure 5, which shows the average results and their standard deviations for generation 5 of the first scenario, after 50 iterations. The means of additive, epistasis and environmental effects in the breeding population have increased to 38.02, 34.95 and 138.4, respectively, and dominance decreased to -15.56, yielding an average phenotypic effect, i.e., sum of all effects, of 195.8. Inbreeding has accumulated within the sblines with an average $F=0.5198$ and the effective population size (Ns) has decreased to 31.48. The additive variance has increased slightly over the 5 generations of breeding.

When requested by the user, the output also includes descriptions of assembled seed orchard populations, such as

FOREST TREE BREEDING POPULATION SIMULATOR
Version 2.0.15

Copyright (C) 1994-1995, Genesis Forest Science Canada Inc.

CONTROL DATA FOR SCENARIO # 1 OF 2

The user has supplied control data in file: EXAMPLE.RSP
Progeny and summary data written to drive E:
Output will NOT be sent to the printer.
Scenario summaries WILL be written to file EXAMPLE.001
Series data will NOT be written to a file.
Detailed progress and times WILL be displayed.

MAIN SIMULATOR SETTINGS ...

```

1. Seed for random number generator . . . . . 45678
2. Number of generations for simulation . . . . . 5
3. Number of iterations . . . . . 50
STRUCTURE OF BASE POPULATION (Generation 0) ...
4. Number of trees in breeding population . . . . . 160
5. Number of sblines in breeding population . . . . . 40
6. Effective size of breeding populations expressed as . . . Ns
7. Mean of trait in base population . . . . . 100.0
8. Standardized rate of inbreeding depression . . . . . -2.000
9. Additive variance in base population . . . . . 100.0
10. Dominance variance in base population . . . . . 25.00
11. Epistatic variance in base population . . . . . 75.00
12. Environmental variance . . . . . 300.0
Heritabilities: Narrow-sense: .200 Broad-sense: .400

```

DESIGN FOR POLY-CROSS TESTING ...

```

13. Number of males represented in polymix . . . . . 0
14. Number of polycross progeny tested per parent . . . . . 0

```

DESIGN FOR GENERATION ADVANCEMENT PLANTATIONS ...

```

15. NON-ASSORTATIVE mating design totalling 240 crosses.
16. Number of genotypes tested per cross . . . . . 120
17. Number of ramets cloned per genotype . . . . . 1

```

SELECTION METHODS ...

```

18. Selection of breeding population:
    Combined index selection
    with up to 120 trees per FS family and 32000 HS relatives,
    and no restriction on number of progeny per parent.
19. Selection of seed orchard populations:
    UNTESTED orchards selected by combined index selection,
    to select 40 trees,
    with up to 1 per cross and 0 HS relatives,
    and no restriction on number from a common parent.
20. Seeding mixture of selected FS families will NOT be deployed.
21. A mixture of selected clones will NOT be deployed.

```

TIMING:

First iteration started: 1995.02.20 at 09:09:16.06
Last iteration completed: 1995.02.20 at 11:08:44.94

Figure 4. - Printout of simulation control parameters for Example 1.

RESULTS FOR SCENARIO # 1 AFTER 50 ITERATION(S)

GENERATION 5

BREEDING POPULATION of 160 trees:

Effect	Mean	(Std. Dev.)	Variance	(Std. Dev.)
Additive effects	38.02	(5.753)	109.8	(29.76)
Dominance effects	-15.56	(3.227)	22.24	(5.227)
Epistasis effects	34.95	(4.267)	41.91	(9.128)
Environmental effects	138.4	(1.166)	119.0	(16.53)
Total phenotype	195.8	(6.744)	145.7	(26.81)
Average Inbreeding (F)	.5198	(.8017E-02)	.2447E-02	(.5783E-03)
Effective pop'n size	31.48	(.3710)		

UNTESTED Seed Orchard Population (40 best untested phenotypes):

Effect	Mean	(Std. Dev.)	Variance	(Std. Dev.)
Additive effects	38.56	(5.856)	113.4	(37.36)
Dominance effects	-15.39	(3.271)	21.90	(7.754)
Epistasis effects	35.64	(4.879)	42.63	(13.64)
Environmental effects	143.9	(1.890)	120.0	(28.27)
Total phenotype	202.8	(6.964)	147.4	(35.27)
Average Inbreeding (F)	.5199	(.8244E-02)	.2450E-02	(.6420E-03)
Effective pop'n size	26.32	(.1426)		
Cumulative gain %	38.56	(5.856)		

Figure 5. - Printout of population structure of generation 5 for the first example scenario.

untested and polycross-tested seed orchards. Statistics similar to those for the breeding population are given, i.e., the means and variances of effects. Additionally, for these seed orchards, the output gives the "cumulative gain" compared with the trait mean in the base population, which represents the cumulative increase in the mean of additive effects, as the progeny from the orchard will capture gain only from additive effects. Had deployment of a full-sib family mixture been requested, the cumulative genetic gain would include the mean increase in

both additive and dominance effects, since the full-sib progeny will reflect selection for both general and specific combining ability effects. Similarly, for deployment of a clonal mixture, the cumulative gain includes the mean increase in all genetic effects, i.e., the sum of additive, dominance and epistasis effects.

Comparing the two breeding strategies

The changes in genetic structure of the breeding populations over the 5 generations of breeding are summarized in *table 2*, where it can be seen that there were large differences between the 2 strategies. Although the increase in the mean additive effect is less, the strategy using the small, 4-tree sublines has retained the additive genetic variance originally present in the base population. Dominance and epistasis variance have remained relatively constant in both strategies. The small sublines have maintained a larger effective population size, despite the much greater accumulation of inbreeding within the sublines.

In *table 3* we can see that while seed orchards for both breeding plans produced considerable genetic gain, that achieved by the plan using larger sublines, positive assortative mating and polycross-tested orchards was about double that of the small-subline strategy in each of the first 3 generations, but this was at the expense of effective population size and genetic variance has been seriously eroded. From this small example, it is easy to see how many “what-if” questions can be addressed by runs of the simulator, in an attempt to identify a combination of breeding plan design parameters which will maximize genetic gains, while maintaining acceptable levels of genetic diversity in both the breeding and production populations.

Discussion

POPSIM is currently in version 2.0, but is considered a “work-in-progress”. Suggestions from several researchers and operational tree breeders have been incorporated during its development, and additional features will likely be added as

the software is applied to evaluate real-life breeding proposals. POPSIM was designed to give insight into the effectiveness of various population-management procedures and to guide the design of operational breeding programs that must balance the quest for genetic gain against the requirement to maintain diversity in plantations of improved stock. However, all simulation models have limitations and the desire to produce a flexible tool is hampered by some important practical problems.

POPSIM addresses some of the issues raised by KING and JOHNSON (1993), but is still limited by current quantitative theory describing the recombination of genetic effects. This is particularly true for the simulation of nonadditive effects which must be considered when strategies incorporating deployment of full-sib families or clones are to be evaluated. Many users will only be interested in the simulation of additive effects in strategies that deploy material produced in seed orchards, and the algorithms presented here for inheritance of additive effects are considered somewhat more reliable than those for nonadditive effects. Unlike the model employed by KING and JOHNSON (1993), POPSIM calculates family variance components directly and performs an adjustment to within-family additive variance to account for the effects of inbreeding. Even so, the rapid erosion of additive variance observed over several generations under some selection schemes raises some doubt about the validity of current quantitative theory when extended over several breeding cycles, in the absence of mutation or other mechanisms that might generate “new” genetic variance. Users should interpret simulations of such long-term scenarios with caution.

The breeding population in our model is closed and generations are discrete. Furthermore, all sublines are managed in an identical fashion. In a real-life program, sublines may be out-of-phase with each other and are often managed differently under a 2-tiered “nucleus” breeding scheme (COTTERILL *et al.*, 1989) with some identified as “elite” breeding groups. Some genetic exchange may occur among sublines and new material may be introduced. Selections established in production

Table 2. – Structure of the breeding population up to the fifth generation for 2 example strategies.

	Generation							
	0 (base)		1		2		5	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
<i>Strategy #1- 4-parent sublines:</i>								
Additive effects	-0.1	100.8	10.9	81.0	18.9	87.2	38.0	109.8
Dominance effects	0.0	25.4	3.5	23.8	2.8	23.9	-15.6	22.2
Epistasis effects	-0.1	75.5	8.5	65.0	16.1	56.6	35.0	41.9
Environmental effects	100.0	295.6	133.3	137.8	135.2	129.6	138.4	119.0
Total phenotype	99.7	500.4	156.2	105.4	172.9	121.2	195.8	145.7
Average inbreeding (F)	0.000	0.000	0.000	0.000	0.211	0.004	0.520	0.002
Effective population size (N_e)	160.0		69.2		49.9		31.5	
<i>Strategy #2 - 20-parent sublines:</i>								
Additive effects	-0.1	98.9	11.1	79.7	20.7	71.9	43.4	58.9
Dominance effects	0.0	25.4	3.2	23.3	5.7	22.9	9.8	21.2
Epistasis effects	0.1	76.2	8.9	65.9	16.8	57.0	35.7	41.2
Environmental effects	100.2	307.4	135.1	135.2	135.7	130.8	137.4	110.8
Total phenotype	100.1	503.3	158.4	108.6	178.8	101.8	226.5	95.1
Average inbreeding (F)	0.000	0.000	0.000	0.000	0.022	0.003	0.121	0.004
Effective population size (N_e)	160.0		78.4		49.0		23.1	

Table 3. – Structure of seed orchard populations up to the fifth generation for 2 example strategies.

	Generation							
	0 (base)		1		2		5	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
<i>Strategy #1 - Orchard with 40 trees selected by combined-index:</i>								
Additive effects	5.5	83.1	11.6	80.1	19.5	85.8	38.6	113.4
Dominance effects	1.5	24.2	3.8	23.7	3.2	24.9	-15.4	21.9
Epistasis effects	4.5	65.6	9.8	63.1	17.3	56.0	35.6	42.6
Environmental effects	116.3	157.2	138.9	136.0	140.4	128.6	143.9	120.0
Total phenotype	127.8	116.7	164.1	109.8	180.4	127.3	202.8	147.4
Average inbreeding (F)	0.000	0.000	0.000	0.000	0.211	0.004	0.520	0.002
Effective population size (N_e)	40.0		40.0		33.0		26.3	
Cumulative genetic gain %	5.5		11.6		19.5		38.6	
<i>Strategy #2 - Orchard with 40 trees selected by Polycross test:</i>								
Additive effects	10.7	40.4	19.5	46.3	28.1	44.9	49.0	44.0
Dominance effects	0.0	25.9	2.8	22.2	5.0	23.5	9.5	22.2
Epistasis effects	0.1	81.1	8.2	66.2	16.1	56.7	35.0	38.6
Environmental effects	100.0	325.2	131.7	138.0	132.6	136.8	135.0	111.9
Total phenotype	110.9	487.9	162.1	117.1	181.8	107.3	228.4	102.8
Average inbreeding (F)	0.000	0.000	0.000	0.000	0.030	0.004	0.129	0.004
Effective population size (N_e)	40.0		32.6		25.8		16.7	
Cumulative genetic gain %	10.7		19.5		28.1		49.0	

orchards will usually be the known best parents and may come from different generations. While some of these features could be incorporated in a model, the simplifications used by POPSIM are thought to be reasonable and still permit the comparison of management strategies in a meaningful way. Similarly, POPSIM is limited to only a few of the many approaches that could be employed in selection and only one trait (or multiple-trait index) is considered. Alternative selection schemes could be added to the software, as required.

Particularly when compared to deterministic models, POPSIM is quite slow and generating reliable results requires several iterations. The program must process realistic quantities of simulated test data, performing all of the evaluation and selection calculations that would be carried out in a real-life breeding program. Furthermore, the simulation of stochastic processes requires the generation of random numbers and POPSIM spends most of its execution time performing this procedure. Pseudo-random number generators vary tremendously in speed and randomness, and we have opted for a very reliable, but rather slow algorithm. If speed were critical and a user could tolerate less reliable randomness, substituting other generators could perhaps make POPSIM up to four times faster.

Enquiries on software distribution, capabilities of the current version, and possibilities for modifications to meet specific requirements should be directed to the senior author.

Acknowledgments

We are grateful to the many colleagues who have tested earlier, and at times rather unstable, versions of the POPSIM software and who shared their ideas and suggestions with us. The development of this software was performed by Genesis Forest Science Canada Inc. as part of a research contract supported through the Canada/Nova Scotia Cooperation Agreement on Forest Development. We would also like to thank Dr. DAG LINDGREN, Dr. PAUL JEFFERSON, Dr. JOHN KING, Dr. KRIS MORGENSTERN, Dr. HANS VAN BULJTENEN and one anonymous reviewer for their helpful reviews and corrections to the manuscript.

Literature

- BAKER, R. J.: Selection indices in plant breeding. CRC Press, Boca Raton, FL. 218 pp (1986). — BORRALHO, N. M. G.: Heterogeneous selfing rates and dominance effects in estimating heritabilities from open-pollinated progeny. *Can. J. For. Res.* **24**: 1079–1082 (1994). — BURROWS, P. M.: Inbreeding under selection from unrelated families. *Biometrics* **40**: 357–366 (1984). — CABALLERO, A.: Developments in the prediction of effective population size. *Heredity* **73**: 657–679 (1994). — CHESSER, R. K., RHODES, O. E., JR., SUGG, D. W. and SCHNABEL, A.: Effective sizes for subdivided populations. *Genetics* **135**: 1221–1232 (1993). — COTTERILL, P. P.: Genetic gains expected from alternative breeding strategies including simple low cost options. *Silvae Genet.* **35**: 212–223 (1986). — COTTERILL, P. P. and JACKSON, N.: Gains expected from clonal orchards under alternative breeding strategies. *For. Sci.* **35**: 183–196 (1989). — COTTERILL, P. P., DEAN, C., CAMERON, J. and BRINDBERGS, M.: Nucleus breeding: a new strategy for rapid improvement under clonal forestry. In: *Breeding tropical trees: Population structure and genetic improvement strategies in clonal and seedling forestry*. Proc. IUFRO Conference, Pattaya, Thailand, November 1988. (Edited by G. I. GIBSON, A. R. GRIFFIN and A. C. MATHESON). Oxford Forestry Institute, Oxford, U. K. and Winrock International, Arlington, Virginia, U.S.A. pp. 39–51 (1989). — DE BOER, I. J. M., and VAN ARENDONK, J. A. M.: Prediction of additive and dominance effects in selected or unselected populations with inbreeding. *Theor. Appl. Genet.* **84**: 451–459 (1992). — DEMPFFLE, L.: Problems in the use of the relationship matrix in animal breeding. In: *Advances in statistical methods for genetic improvement of livestock*. (Edited by D. GIANOLA and K. HAMMOND). Springer Verlag, Berlin, Germany. pp. 454–473 (1990). — DE ROO, G.: A stochastic model to study breeding schemes in a small pig population. *Agric. Syst.* **25**: 1–25 (1987). — DE VRIES, A. G., VAN DER STEEN, H. A. M. and DE ROO, G.: Multi-stage selection in a closed dam line of pigs. *Livest. Prod. Sci.* **24**: 161–180 (1990). — FALCONER, D. S.: *Introduction to quantitative genetics*. 2nd ed. Longman, London and New York. 340 pp (1981). — FOWLER, D. P.: Strategies for the genetic improvement of important tree species in the Maritimes. *Inf. Rep. M-X-156*, Can. For. Serv. - Maritimes, Fredericton, NB, 30 pp (1986). — FRASER, A. S.: Simulation of genetic systems by automatic digital computers. I. Introduction. *Aust. J. Biol. Sci.* **19**: 484–491 (1957). — GRIFFIN, A. R. and COTTERILL, P. P.: Genetic variation in growth of outcrossed, selfed and open-pollinated progenies of *Eucalyptus regnans* and some implications for breeding strategy. *Silvae Genet.* **37**: 124–131 (1988). — KEMPTHORNE, O.: An overview of the field of quantitative genetics. In: *Proceedings of the Second International Conference on Quantitative Genetics*, North Carolina State University, Raleigh, NC. (Edited by B. S. WEIR, E. J. EISEN, M. M.

GOODMAN, and G. NAMKOONG). Sinaeur Assoc. Sunderland MA. pp. 47–56 (1988). — KING, J. N. and JOHNSON, G. R.: Computer simulation of clonal selection gains. In: Proceedings: FRI/NZFP Forests Ltd clonal forestry workshop. 1–2 May 1989, Rotorua, NZ. (Edited by J. T. MILLER). FRI Bulletin No. 160, Forest Research Institute and NZFP Forests Ltd., Rotorua, NZ. pp. 51–52 (1991). — KING, J. N. and JOHNSON, G. R.: Monte Carlo simulation models of breeding-population advancement. *Silvae Genet.* **42**: 68–78 (1993). — LACY, R. C.: Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conserv. Biol.* **1**: 143–158 (1987). — LEVIN, B. R.: Simulation of genetic systems. In: Computer applications in genetics. (Edited by N. E. MORTON). Proceedings of the International Conference on Computer Applications in Genetics, Univ. Hawaii Press, Honolulu, HI. pp. 38–46 (1969). — LIBBY, W. J.: Clonal selection, and an alternative seed orchard scheme. *Silvae Genet.* **13**: 32–40 (1964). — MAHALOVICH, M. F.: Modeling positive assortative mating and elite populations in recurrent selection programs for general combining ability. Ph. D. thesis, NC State Univ. 129 pp (1990). — MATHESON, A. C. and LINDGREN, D.: Gains from the clonal and the clonal seed-orchard options compared for tree breeding programs. *Theor. Appl. Genet.* **71**: 242–249 (1985). — MEUWISSEN, T. H. E. and LUO, Z.: Computing inbreeding coefficients in large populations. *Genet. Sel. Evol.* **24**: 305–313 (1992). — MEYER, K. and SMITH, C.: Comparison of theoretical and simulated equilibrium genetic response rates with progeny testing in dairy cattle. *Anim. Prod.* **50**: 207–212 (1990). — MICROSOFT: Microsoft® FORTRAN PowerStation Professional Development System Version 1.0 for MS-DOS® and Windows™ Operating Systems. Microsoft Corporation, Redmond, WA (1993). — MULLIN, T. J.: Evaluating the economics of alternative breeding and deployment strategies for northeastern conifers. In: Proceedings of the 24th Meeting of the Canadian Tree Improvement Association. Part 2. 9–12 August 1993. Fredericton, NB. (Edited by J. LAVEREAU). Natural Resources Canada, Fredericton, NB. pp. 82–105 (1994). — MULLIN, T. J. and PARK, Y. S.: Estimating genetic gains from alternative breeding strategies for clonal forestry. *Can. J. For. Res.* **22**: 14–23 (1992). — MULLIN, T. J. and PARK, Y. S.: Stochastic simulation of forest tree breeding populations: a users guide for POPSIM version 2.0. Inf. Rep. M-X-195, Canadian Forest Service – Maritimes, Fredericton, NB (1995). — NAMKOONG, G., SNYDER, E. B. and STONECYPHER, R. W.: Heritability and gain concepts for evaluating breeding systems such as seedling orchards. *Silvae Genet.* **15**: 76–84 (1966). — PARK, Y. S. and FOWLER, D. P.: Inbreeding in black spruce (*Picea mariana* (MILL.) B.S.P.): self-fertility, genetic load, and performance. *Can. J. For. Res.* **14**: 17–21 (1984). — PARK, Y. S., SIMPSON, J. D., ADAMS, G. W., MORGENSTERN, E. K. and MULLIN, T. J.: An updated breeding strategy for black spruce (*Picea mariana* (MILL.) B.S.P.) in New Brunswick. In: Papers presented at the Workshop on Breeding Strategies of Important Tree Species in Canada, Fredericton, New Brunswick. August 18, 1993. (Edited by Y. S. PARK and G. W. ADAMS). Inform. Rep. M-X-186E, Natural Resources Canada, Fredericton, NB. pp. 41–54 (1993). — PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T. and FLANNERY, B. P.: Numerical recipes in FORTRAN: the art of scientific computing. (2nd ed.). Cambridge University Press, Cambridge, UK. 963 pp (1992). — RUANE, J. and THOMPSON, R.: Comparison of simulated and theoretical results in adult MOET nucleus schemes for dairy cattle. *Livest. Prod. Sci.* **28**: 1–20 (1991). — SHAW, D. V. and HOOD, J. V.: Maximizing gain per effort by using clonal replicates in genetic tests. *Theor. Appl. Genet.* **71**: 392–399 (1985). — SHELBOURNE, C. J. A.: Tree breeding methods – A review of tree breeding strategies in relation to classical plant breeding methods with quantitative genetic expectations of gain, and predictions using data from *Pinus radiata*. Technical Report No. 55, New Zealand Forest Service, Wellington. 44 pp. (1969). — SHELBOURNE, C. J. A.: Genetic gains from different kinds of breeding population and seed or plant production population. Paper presented at the IUFRO symposium "Intensive forestry: the role of eucalypts", held in Durban, South Africa, in September 1991. *South African Forestry Journal* **160**: 49–65 (1992). — TIER, B.: Computing inbreeding coefficients quickly. *Genet. Sel. Evol.* **22**: 419–430 (1990). — van BUIJTENEN, J. P. and LOWE, W. J.: The use of breeding groups in advanced generation breeding. In: Proceedings of the 15th Southern Forest Tree Improvement Conference. pp. 56–65 (1979). — WEIR, R. J. and TODD, D.: Third-cycle breeding strategy: a description and economic appraisal for the North Carolina University – Industry Cooperative Tree Improvement Program. In: Proceedings of the 24th Canadian Tree Improvement Association, Part 2. 9 to 12 Aug. 1993. Fredericton, NB. Edited by J. LAVEREAU. Natural Resources Canada, Fredericton, NB. pp. 41–51 (1994).

Callus Induction and Haploid Plant Regeneration from Anther Culture of Two Poplar Species

By A. MOFIDABADI¹), J. KISS¹), K. MÁZIK-TÓKEI¹), E. GERGÁCZ²) and L. E. HESZKY¹)³)

(Received 9th February 1995)

Summary

Anthers from 5 different clones of *Populus nigra* and *Populus deltoides* were isolated from dormant flower buds before catkin elongation and exposition, in order to establish haploid and doubled haploid poplar plants in vitro. Isolated anthers were cold treated and placed on MS medium supplemented with 2,4-D (0.5 mg/l to 2.0 mg/l) and Kinetin (0.1 mg/l to 1.0 mg/l). The optimal hormone concentration for callus induction was 1.0 mg/l of 2,4-D and 0.1 mg/l of Kinetin. The callus initiation response of the genotypes ranged from 24% to 75%. The genotypes with the best callus initiation response were *P. nigra* N-90 (59%) and *P. deltoides* D-29 (75%). Calli were subcultured on MS or WPM medium supplemented with BA (0 mg/l to 2.5 mg/l) and NAA (0 mg/l to 0.2 mg/l) for plant regeneration.

The highest shoot regeneration frequency (79%) was obtained after 2 subcultures on MS medium supplemented with BA (1.0 mg/l) and NAA (0.2 mg/l), and on WPM medium supplemented with BA (2.5 mg/l). The rate of shoot regeneration and number of shoots/calli ranged from 4% to 79% and 1 to 9, respectively.

Key words: Androgenesis, *Populus*.

FDC: 165.442; 176.1 *Populus nigra*; 176.1 *Populus deltoides*.

Abbreviations

MS: Murashige and Skoog; WPM: Woody Plant Medium; 2,4-D: 2,4-dichlorophenoxyacetic acid; BA: benzyladenine; NAA: naphthaleneacetic acid.

Introduction

In most tree species recurrent inbreeding to increase homozygosity is not successful due to long generation cycles, high initial levels of heterozygosity and inbreeding depression

¹) Gödöllő University of Agricultural Sciences, Gödöllő, Hungary

²) Research Station of Forest Research Institute, Sárvár, Hungary

³) Corresponding author