Stochastic Modelling of the Impacts of Four Generations of Pollen Contamination in Unpedigreed Gene Resources

By R.D. BURDON[*]) and S. KUMAR

New Zealand Forest Research Institute, Private Bag 3020, Rotorua, New Zealand

(Received 25th February 2002)

Abstract

Unpedigreed *ex-situ* gene resources will often incur repeated pollen contamination from intensively improved stands of narrow genetic base. This will undermine the value of the resources for their intended purposes, especially if natural and/or silvicultural selection needs to be part of the management system. The potential importance of such contamination was studied by stochastic simulation over five generations of which four incurred contamination, using the scenario for *Pinus radiata* in New Zealand as an illustration. Assumptions included: a single conceptual trait (which may represent a multi-trait function); two semi-empirical initial differences in genetic merit between gene resources and improved (contaminating) stands (Δ_1), which confer a selective advantage on offspring from contaminant pollinations; a semi-empirical heritability (h^2) (and two variations thereof); four arbitrary levels of pollen contamination (c) (1%, 2.5%, 5% and 10% fertilisation by outside pollen); c uniform among parents and constant over generations; genetic improvement per unit time in the contaminating population matching that per generation in the uncontaminated gene-resource component ($\Delta_{t+1,t}$); effective cumulative contamination (C) (% contaminant-origin genes) defined by pedigree; discrete generations; and 40 (or 160) selected seed parents with 275 offspring each. For each of 18 combinations of input parameter values 100 replicate runs were made to obtain distributions of expected outcomes. The value of C was very sensitive to both Δ_1 and c indicating major impacts of these factors, and built up progressively over successive generations, reflecting the selective advantage of individuals resulting from contaminant pollinations combined with continued fresh contamination. Impacts of likely departures from the assumptions are discussed, but results point to a need to quantify actual contamination rates. The native population concerned and choice of planting site may both influence C strongly.

Key words: gene resources, *ex-situ*, genetic management, genetic contamination, stochastic simulation, exotic species, *Pinus radiata.*

Introduction

In forest tree breeding there can be good reasons for maintaining *ex-situ* gene resources that are designed to preserve and perpetuate, for potential use by breeders, the variability represented in natural populations. Such material is designed to underpin breeding populations, as a store of greater genetic variability. Preserving such material over successive generations, however, can incur technical problems as well as needing strong political will. New Zealand's *Pinus radiata* D. Don breeding programme, involving an exotic species, is an example, for which well-documented precedents seem to be lacking. In it, *ex-situ* plantings of native-provenance material figure prominently in gene conservation (BURDON, 1986, 1988). For the Californian mainland populations it was planned to maintain unpedigreed subpopulation lots from seed collected in California in 1978. The present scheme (*op. cit.*) entails: growing material from designated population units (≥ 10 ha for any subpopulation); and conducting, by successive generations of seed collection, mild mass selection for general health and silvicultural traits (growth and form). The rationale has been several-fold (*op. cit.*). The large blocks were intended to provide 'internal' protection against significant pollen contamination. They would also allow significant 'clean-up' selection for silvicultural traits while maintaining a good population size. It was assumed that, provided population size is kept large, one can do significant selection without materially eroding the genetic variability that one wants to preserve. The 'clean-up' selection, by making living, *ex-situ* material more commercially acceptable, should reduce the opportunity costs per unit area of ground of maintaining such resources in place of intensively improved stock, and prevent an unacceptable gap in genetic quality between the back-up resource material and the breeding population. This will be in the context of some inevitable natural selection in local environments if living material is grown in the field. In turn, the reduced opportunity costs will allow larger resource plantings to be maintained, with various attendant benefits. Discrete generations are envisaged for renewing gene resources, but with relatively long generation intervals in order to reduce allele losses. In any event, the gene resources have been established as sources of new variability, in the sense of having no known or recent coancestry with the existing breeding population, but with the gap in general level of improvement to be much reduced compared with natural stands.

In some other parts of the world a number of exotic species also show excellent growth potential, but a high proportion of individuals that are imperfectly adapted and/or of poor tree form (ZOBEL and TALBERT, 1984). Therefore, the general scenario for *P. radiata*, which entails 'clean-up' selection within managed *ex-situ* gene resources, could be quite widely applicable.

Recent consideration (BURDON, 1995 Fig. 2, 1997), however, has raised concern over the possible cumulative effects of pollen contamination of such *ex-situ* gene resources, given that:

- the contamination should come increasingly from intensively improved but narrowly-based commercial stands,
- seedlings resulting from such pollen contamination should enjoy a selective advantage, deriving from the genetic merit of their pollen parents, under the proposed management scheme,
- the impact of the selective advantage of contaminant offspring will probably compound over generations,
- in addition, such pollen contamination can occur afresh each generation.

Several of these conditions are clearly met, in large degree, with *P. radiata* in New Zealand.

---

[*]) E-mail: rowland.burdon@forestresearch.co.nz

Substantial cumulative contamination will severely compromise two key intended features of gene resources: their being both sources of new variability and having no recent coancestry with the existing breeding population. Intuitively, the cumulative contamination should be very sensitive to both the contamination rate per generation and any selective advantage enjoyed by trees of contaminant descent. How sensitive it is has been uncertain, but modelling such contamination should give a basis for interpreting the longer-term significance of any pollen contamination rate.

The *ex-situ* gene resources of *P. radiata*, and their management, have become far more important now that presence of pitch canker rules out for the foreseeable future further seed imports from mainland California.

Pollen contamination, apart from potentially degrading gene-resource plantings, may be a serious issue for the release of genetically-engineered pine. While transgenic pollen could not escape into natural New Zealand ecosystems, any pollen release from transgenic trees could arouse public fears, unless reproduction is totally suppressed.

Accordingly, a stochastic study was made of the cumulative impacts of pollen contamination, using a combination of some input parameters that were derived from empirical data from *P. radiata* field trials and some arbitrary variations thereof, and is reported here. A specific scenario is thus used; nevertheless, it should serve to illustrate, for application to other contexts, both how cumulative pollen contamination may behave and how various case-specific issues may be involved in predicting and managing pollen contamination.

Outline of Model

Basic conditions

Contamination rate. We assume fresh pollen contamination of a gene resource each generation from genetically improved plantations. We also assume, for simplicity, that the level of contamination, i.e., percent of seed resulting from outside pollination (c), will be the same for all seed parents, and that the level of *de-novo* contamination and the census size of the gene-resource planting will be constant over generations.

Contamination pedigrees. As generations pass, the number of possible 'contamination pedigrees' will increase rapidly. In the first generation there will, by definition, be no contamination. Next generation there will have just two categories, uncontaminated and contaminant, the latter having a frequency corresponding to the pollen contamination rate (c). The following generation can have six such pedigree classes: four representing the combinations (including reciprocals) of the two preceding classes, plus two classes from *de-novo* contamination. The fourth generation can have 42 such classes, and the fifth generation 1806. The number of possible classes in the t^{th} generation, q_t, is given by

$$q_t = q_{t-1}(q_{t-1} + 1), \text{ for } t \geq 2 \ (q_1 = 0) \qquad (1)$$

where q_{t-1} and q_1 are the numbers for Generations $t-1$ and 1 respectively.

Simulating the selection at the end of Generation t, leads to an estimate of C in Generation $t+1$.

Measure of contamination (C). This represents the percentage of genes of contaminant origin, as inferred from pedigree. For each pedigree class C was the mean of that represented in the pollen parents and seed parents respectively. Thus, in Generation 2, for individuals resulting from contaminant pollen $C = 0.5$; for Generation 3, when those individuals crossed with pure gene-resource pollen $C = 0.25$, if they crossed with like $C = 0.5$, and if they crossed with contaminant pollen $C = 0.75$. Thus overall C in the gene resource was the sum, over the various pedigree classes, of the products of the proportion of each class and its C level (see Eq. 2 below).

Population size. We assume that in each generation: 11,000 individuals are planted (e.g. 10 ha @ 1100 stems/ha); there is no mortality; the 40 best phenotypes are selected as seed parents (matching a management proposal, T.G. VINCENT pers. comm.); and each seed parent contributes seedlings with equal probability. For a subsample of other parameter combinations (all heritabilities, but only the basic values of c and the initial genetic superiority of the source of contaminating pollen), 160 seed parents out of 44,000 individuals were selected each generation. Among pollen parents 3 out of 6 phenotypic selection (selection intensity = $i_{3,6}$) (LINDGREN and NILSSON, 1985) is assumed, again with each of the selected 50 % contributing with equal probability in panmictic pollination, subject to the condition of the probability (c) of contaminant pollination. Issues of effective population size and possible inbreeding depression were disregarded as being probably negligible with the population sizes considered.

Genetic parameters. For the basic scenario, a single trait was assumed of normal distribution, phenotypic variance (σ^2_P) of 1 and an initial mean of zero for the gene resource, i.e., ($N(0,1)$), thereby standardising initial values. Thus, for this case, the additive genetic variance (σ^2_A) equalled narrow-sense heritability (h^2 ($= \sigma^2_A/\sigma^2_P$, $\sigma^2_P = \sigma^2_A + \sigma^2_e$, σ^2_e = environmental variance)). No discrimination was needed between non-additive and environmental variation. The assumed h^2 was 0.27, based on a multi-trait index (LIN and ALLAIRE, 1977) constructed from 'consensus' variance-covariance matrices (from BURDON, 1992; BURDON *et al.,* 1992b) involving the selection traits for a gene resource (growth-rate and tree-form variables). Alternative h^2 (0.15 and 0.4) within plantations (the gene resource and contaminating sources alike) were simulated by varying (σ^2_P) rather than σ^2_A.

Contaminating population. Assumed genetic parameters are as for the gene-resource population. However, the initial genetic superiority (Δ_1) was set semi-empirically; this was done using largely unpublished data from seven field trials comparing native populations and some improved seedlots (G.R. JOHNSON *et al.,* unpubl.) and numerous provenance trials (BURDON *et al.,* 1992a, 1992c, 1994, 1997). Gain per [gene-resource] generation was arbitrarily assumed to equal that expected from the prescribed selection intensity in the gene resource and its h^2 and σ^2_P, with h^2 and σ^2_P 'empirically' declining slightly in time through the Bulmer effect. Within each generation, 3 out of 12 selection in the course of thinning was assumed, at the same current h^2 as in the gene resource. As a variation, Δ_1 was assumed to be double the value initially considered.

Other assumptions. These included: discrete generations; no epistasis; and no specific selection (beyond what was implied by the simulation process) for genes of any particular origin.

Parameter notation

σ^2_{At} = additive genetic variance at Generation t ($t = 1,2,....$) ($\sigma^2_{A1} = \sigma^2_A$),

σ^2_P, σ^2_e = phenotypic and environmental variances respectively,

c_t = level of *de-novo* pollen contamination in Generation t, with the special condition that $c_1 = 0$, but for $t > 1$, $c_{t+1} = c_t$),

C_{jt} = level of contaminant ancestry in pedigree class j in Generation t (j is specified in terms of the combination of male and female parental pedigree classes),

p_{jt} = proportion of the population represented by pedigree class j in generation t,

C_t = cumulative level of pollen contamination in Generation t, with the special condition that $C_1 = 0$,

C_{0t} = expected contamination, at generation t, without genetic superiority/inferiority of contamination pollen source,

Δ_1 = initial genetic superiority of source of contaminating pollen over gene resource (Generation 1),

Δ' = genetic gain in contaminating pollen population through thinning, calculated as $i_{3,12}h^2$, where $i_{3,12}$ denotes selection intensity (standardised selection differential) corresponding to selecting the best three individuals out of every 12,

$\Delta_{(t,1)}$ = genetic gain accumulated in contaminating population from Generation 1 through Generation t, $= \Delta_{(2,1)} + \Delta_{(3,2)} + \ldots\ldots\ldots\Delta_{(t,t-1)}$, $\Delta_{(t,t-1)}$ being calculated, assuming (deterministically) same expected rate of genetic advance per gene-resource generation as in the uncontaminated gene resource, as $0.5(i_{40,\,11000} + i_{3,\,6})h^2\sigma_P$.

Expectations for cumulative contamination

Thus we have

$$C_t = \sum_j C_{jt} p_{jt} \qquad (2)$$

Assuming zero selective advantage associated with contaminant pollination ($\Delta = 0$ for all t), and deterministic behaviour, we have the expectation

$$C_{0t} = 1 - (1 - 0.5c)^{t-1} \qquad (3)$$

Variable parameters

The following parameter values were specified (bold type denoting cases that were considered exhaustively):

Contamination rate (c) –	1 %, 2 %, **5 %**, 10 % ,
Heritability (h^2) in plantations –	0.15, **0.27**, 0.4,
Phenotypic variance (σ^2_P) –	$(0.15/0.27)^{-1}$, **1**, $(0.4/0.27)^{-1}$,
Additive genetic variance (σ^2_A) –	0.27,
Superiority of contaminating population over uncontaminated gene resource in Generation 1 (Δ_1) –	**0.36**, 0.72,
Seed parents/generation (n) –	**40**, 160,
Total candidates/generation (N) –	11000, 44000.

Simulation Procedures

The basic simulation structure is shown in *Figure 1.* To recapitulate assumptions: base population of N trees for Generation 1 of gene-resource population; n seed parents selected each generation to create next generation of N trees (i.e., N/n progeny per seed parent); σ^2_A same in the gene-resource and the contaminating population; and because of selection in gene-resource population, σ^2_A will tend to decline over generations (Bulmer effect).

Simulation steps were:

1. Generate additive (A) and environmental (E) effects of 11,000 base trees drawn randomly from normal distributions with means of 0 and variances of σ^2_A and σ^2_e, i.e., ($N(0,\sigma^2_A)$ and $N(0, \sigma^2_e)$ respectively).

2. Sum A and E of each tree, to give phenotype (P), and select top n as seed parents.

3. For each of the n seed parents, generate 275 offspring with probabilities of c and $(1-c)$ for external and internal pollination respectively. For each offspring, additive genetic values were obtained as:

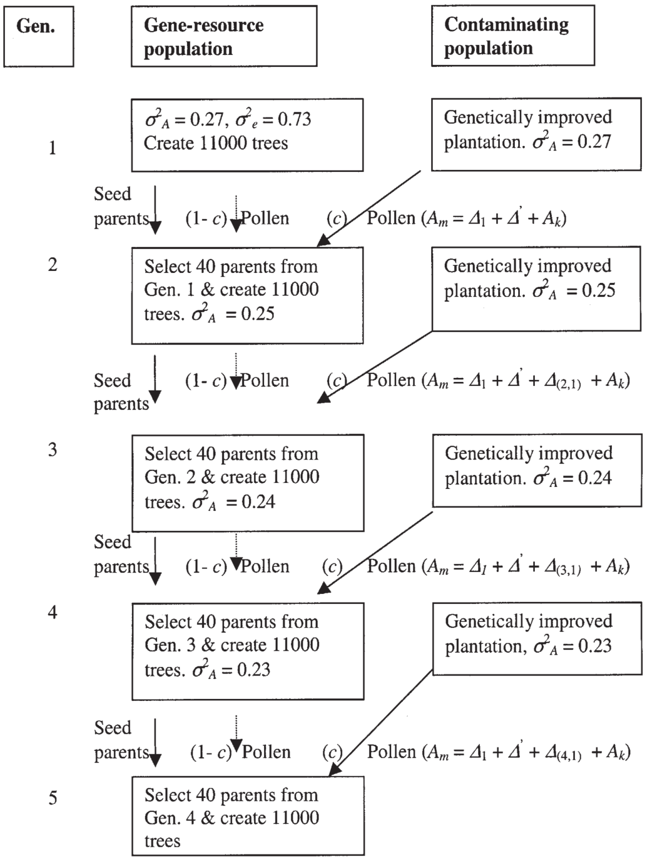

Figure 1. – Different steps for simulating gene-resource population over five generations. The 'c' represents rate of fresh pollen contamination each generation, A_k is the genetic value of k^{th} pollen parent (Eq. 5). Environmental variance (σ^2_e) remains same while additive variance (σ^2_{At}) declined over generations.

$$A_i = 0.5(A_f + A_m) + N(0{,}0.5\sigma^2_A) \qquad (4)$$

where A_f and A_m are the additive values of seed- and pollen parents respectively. The values of A_f for the top n seed parents are known from Generation 1. With internal pollination, each pollen parent was selected randomly from individuals of Generation 1 with known A_m. With outside pollen, i.e., from genetically improved plantations, the A_m value was of the composition

$$A_m = \Delta_1 + \Delta_{(t,1)} + \Delta' + A_k \qquad (5)$$

where A_k is drawn from a normal distribution with mean zero and variance σ^2_A, i.e. from $N(0,\sigma^2_{A(t-1)})$. This represents the sum of the genetic level of the contaminating population is the sum of (1) the initial differential (Δ_1), (2) the gain in that population since Generation 1, (3) the gain in pollen population via the thinning (Δ'), and the additive genetic effect of the pollen parent within the contaminating population. The within-family additive genetic component related to each offspring was given by the second term of the expression for A_i. The environmental effect, drawn from $N(0,\sigma^2_e)$, was added to A_i to obtain each offspring phenotype. This was specified as Generation 2.

4. For each of the two pedigree classes, obtain the mean additive genetic (A) value for the class, and retain count of individuals.

5. Calculate cumulative level of genetic contamination (C_t) (Eq. 2), for Generation 2.

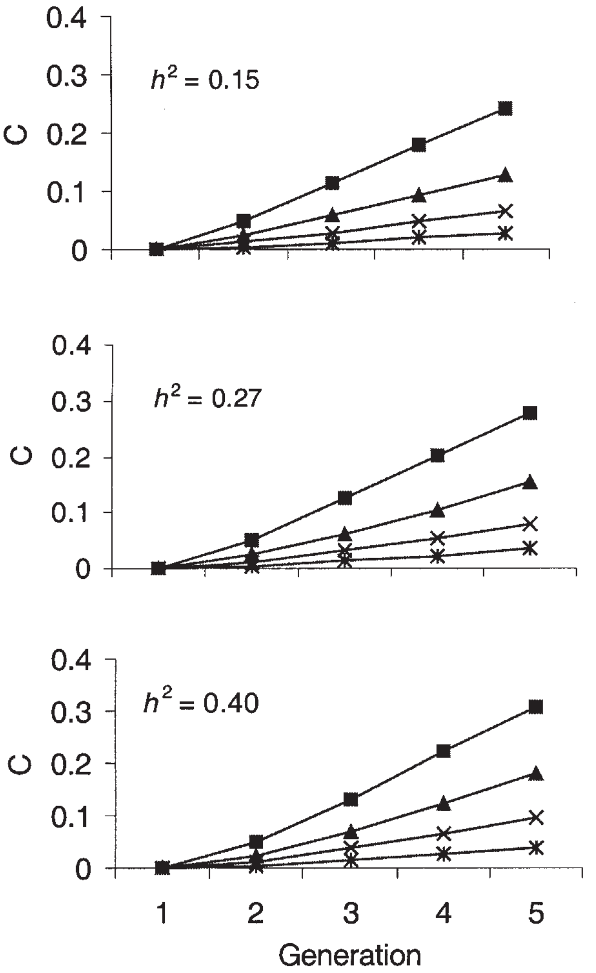

Figure 2. – Cumulative levels of genetic contamination (C) at various levels of c (0.1: –■–; 0.025: –▲–; 0.05: –×–; 0.01: –*– over different generations. Results are shown for three different levels of heritability.

6. Select the top n phenotypes (seed parents) out of N, recording the pedigree classes of each selected seed parent. Generate 275 offspring with probabilities of c and $(1 - c)$ for external and internal pollination respectively. Record the pedigree class of the pollen parent also. The pedigree class of each offspring was determined from those of the seed and pollen parents. For each offspring, the additive genetic value was generated as in Step 3 but taking account of the improved genetic level of the contaminating population (as shown in *Figure 1*). The environmental effect, drawn from $N(0,\sigma^2_e)$, was added to the additive genetic value (A_i) to obtain the phenotype of each offspring. This was specified as Generation 3.

7. For each pedigree classes generated in Step 6, derive mean additive genetic values and count of occurrence (cf Step 4).

8. Derive cumulative level of genetic contamination (C_t), as per Eq. (2), for Generation 3.

9. Derive cumulative level of genetic contamination (C_t) for Generations 4 and 5, following the same argument as in Generation 3 and taking account of the improved genetic level of contaminating population (as shown in *Figure 1*).

10. Repeat Steps 1–9 for 100 times.

11. Repeat Steps 1–10 for each of the remaining sets of input parameter values.

Results

Statistics for the basic case of 40 seed parents per generation and $\Delta_1 = 0.36$ are shown in *Table 1*. Mean values of C were very strongly dependent on c for all heritabilities (see also *Figure 2*). At the higher heritabilities the increase in C with generation (t) was appreciably curvilinear, the increase per generation tending to become greater with time up till Generation 4, noting that an asymptotic approach to 1 would eventually be expected. The dependence of mean C on h^2 was less than on c, within the ranges of parameter values considered, but was still appreciable (see also *Figure 2*).

Table 1. – Average cumulative contamination levels (based on 100 runs for each set of parameter values) together with coefficient of variation (CV%) and skewness statistics (Skew) over successive generations for three levels of h^2 and four levels of c, assuming $\Delta_1 = 0.36$. Forty seed parents are selected based on phenotype in each generation.

		Cumulative contamination per generation (C)											
h^2	Gen.	.01			.025			.05			.10		
		Mean	CV%	Skew	Mean	CV%	Skew	Mean	CV%	Skew	Mean	CV%	Skew
0.15	1	0	-	-	0	-	-	0	-	-	0	-	-
	2	0.005	8.94	-0.17	0.0125	6.42	-0.18	0.025	3.92	0.49	0.050	3.0	-0.04
	3	0.012	45.2	1.29	0.028	28.3	1.12	0.058	17.5	0.64	0.114	13.3	0.42
	4	0.019	44.8	1.07	0.047	28.9	0.98	0.093	18.9	0.25	0.178	12.4	0.23
	5	0.026	42.7	1.29	0.064	24.3	0.55	0.129	17.2	0.04	0.242	11.7	0.32
0.27	1	0	-	-	0	-	-	0	-	-	0	-	-
	2	0.005	9.36	-0.02	0.0125	6.3	-0.13	0.025	3.54	-0.18	0.050	2.9	0.18
	3	0.014	45.9	1.06	0.032	28.2	0.97	0.063	18.8	0.34	0.125	13.9	0.20
	4	0.023	45.0	0.93	0.055	28.0	0.92	0.105	19.5	0.54	0.202	12.8	-0.09
	5	0.035	40.9	0.85	0.081	28.8	0.80	0.154	17.8	0.18	0.278	12.3	-0.21
0.40	1	0	-	-	0	-	-	0	-	-	0	-	-
	2	0.005	8.7	0.26	0.0125	5.8	-0.01	0.025	3.7	0.05	0.050	3.1	0.49
	3	0.014	44.2	1.00	0.037	26.9	0.78	0.071	16.2	-0.28	0.132	13.3	0.29
	4	0.026	45.5	0.92	0.066	26.9	0.62	0.124	16.4	0.02	0.222	13.1	0.08
	5	0.040	44.1	1.00	0.096	25.9	0.43	0.180	15.6	0.04	0.309	12.1	0.34

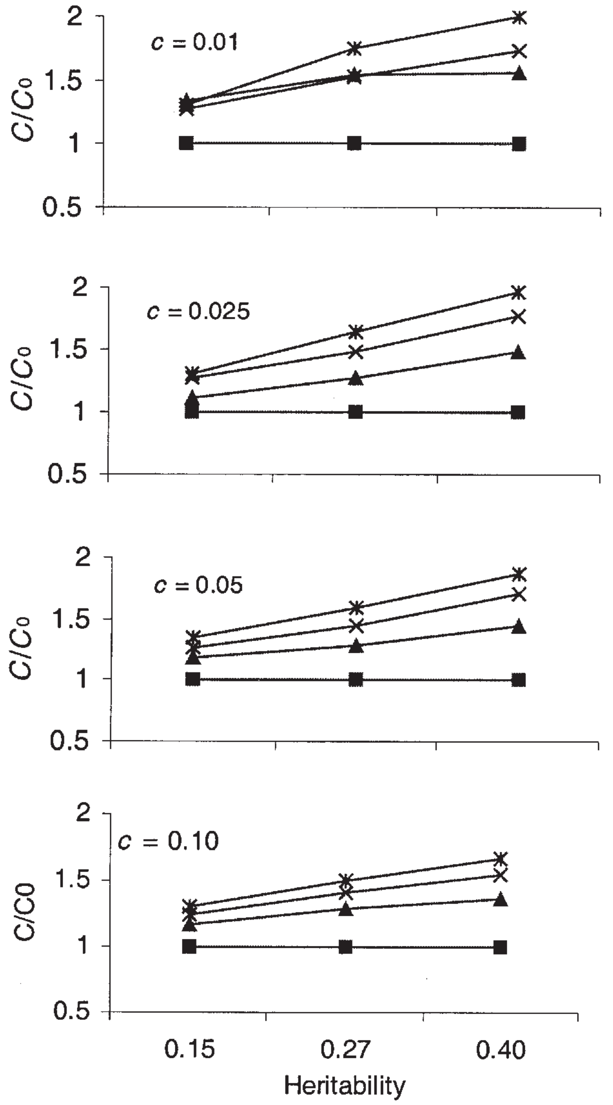

Figure 3. – Plot of the ratio (C/C_0) versus heritability in different generations (Gen. 2: –■–; Gen. 3: –▲–; Gen. 4: –×–; Gen. 5: –*–. Plots are shown for different c levels.

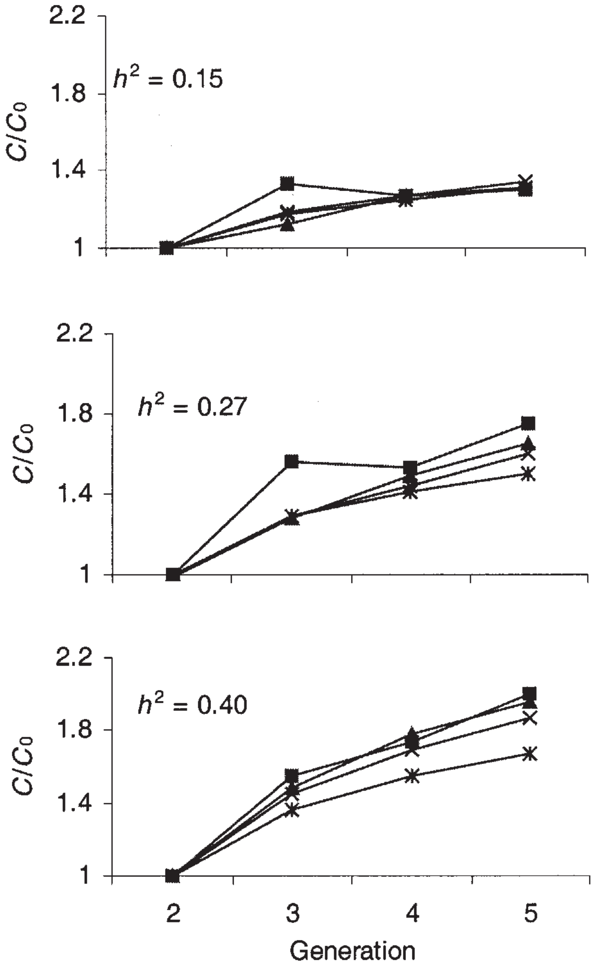

Figure 4. – Plot of the ratio (C/C_0) for different levels of c (0.01: –■–; 0.025: –▲–; 0.05: –×–; 0.1: –*– in different generations. Three heritability levels were used.

Distributions of C also depended strongly on c (*Table 1*). The coefficients of variation, which rose sharply to relatively constant values after Generation 2, dropped sharply with increasing c. They also decreased slightly with increasing h^2. Skewness was generally positive, but apparently tended to fall with increasing c.

The relationship between mean C and C_0 (C_0 = deterministic expected value of C in the absence of any selective advantage of contaminants) was complex. At the lowest h^2, C was only slightly greater than C_0 (*Figures 3, 4*), but with increasing h^2 this difference rose markedly in both relative and absolute terms. The ratio of mean $C : C_0$, which generally rose curvilinearly with t (*Figure 4*), tended to increase strongly with increasing h^2 in the later generations (*Figures 3, 4*), but to fall slightly with increasing c (*Figure 3*). Interestingly, the ratio showed a short-term maximum at Generation 3 for the two lower heritabilities (*Figure 4*).

Effects of increasing n each generation from 40 to 160 are shown in *Table 2*, for $c = 5\%$. Not surprisingly, the coefficients of variation were consistently reduced. Skewness also tended to be reduced. Contamination levels were increased marginally in Generation 3 but were marginally decreased in Generation 5, relative to the case of 40 seed parents, for the two higher heritabilities (*Table 2*).

Doubling Δ_1 to 0.72, leaving other variables constant, led to substantial and continuing increases in mean C (*Table 3*). This was especially marked at higher h^2. Coefficients of variation tended to be lower, especially at the higher heritabilities and in the later generations. No strong pattern was evident for skewness.

Discussion

Importance of assumptions

The following may be noted concerning express or implied assumptions, partly in respect of the scenario for the particular species:

Normal distributions, constant over time. Departures undoubtedly occur, but non-normality should be less for the multi-trait selection criterion than for the component traits. Changes in variance with time, while they may affect the different populations unequally, seem unlikely to cause major bias in predicted C, partly because the visual selection is not necessarily based on normality or constant variances.

Three-out-of-six selection among potential pollen parents in the gene resource. This allows arbitrarily for how thinning, natural or silvicultural, will be strongly constrained by spacing. This thinning intensity is conservative, but the net impact of that on C will reflect a trade-off between additional upgrading of the gene resource (tending to be negative) and potentially more contaminant offspring among pollen parents (tending to be positive, mainly at higher levels of c).

Four-out-of-12 selection among pollen parents in the contaminating population. Again, this allows for selection during thinning, with arbitrary accommodation of constraints on spacing. It seems realistic for the current generation.

Essentially equal contributions among seed parents and surviving pollen parents. An assumption of convenience, it implies simple truncation selection, but departures seem unlikely to invalidate our main conclusions.

Table 2. – Comparison between results for 40 versus 160 seed parents per generation, with c (rate of fresh pollen contamination each generation) = 0.05, for C (cumulative contamination). Coefficient of variation (CV%) and skewness are also shown.

		Mean*		CV%		Skewness	
h^2	Gen	40	160	40	160	40	160
0.15	1	0	0	-	-	-	-
	2	0.025	0.025	3.92	1.8	0.49	0.05
	3	0.058	0.061	17.5	8.5	0.64	0.07
	4	0.093	0.097	18.9	9.8	0.25	0.25
	5	0.129	0.131	17.2	8.7	0.04	0.22
0.27	1	0	0	-	-	-	-
	2	0.025	0.025	3.54	2.0	-0.18	0.30
	3	0.063	0.067	18.8	8.7	0.34	0.39
	4	0.105	0.111	19.5	9.2	0.54	0.03
	5	0.154	0.151	17.8	9.3	0.18	0.11
0.40	1	0	0	-	-	-	-
	2	0.025	0.025	3.7	2.1	0.05	-0.10
	3	0.071	0.075	16.2	8.5	-0.28	0.21
	4	0.124	0.126	16.4	9.5	0.02	-0.08
	5	0.180	0.174	15.6	9.5	0.04	-0.04

* The corresponding values of C_0 (Eq. 3) were 0.025, 0.049, 0.073, and 0.096 for Generations 1, 2, 3, 4 and 5 respectively.

Table 3. – Effect of varying selective advantage differential (Δ_1) on C (cumulative contamination). The rate of fresh pollen contamination each generation (c) was kept at 0.05 and 40 parents are selected each generation. Coefficient of variation (CV%) and skewness are also shown.

		Mean*		CV%		Skewness	
				Δ_1			
h^2	Gen.	0.36	0.72	0.36	0.72	0.36	0.72
0.15	1	0	0	-	-	-	-
	2	0.025	0.025	3.92	4.4	0.49	0.02
	3	0.058	0.067	17.5	18.7	0.64	0.18
	4	0.093	0.112	18.9	18.8	0.25	0.29
	5	0.129	0.167	17.2	18.1	0.04	0.19
0.27	1	0	0	-	-	-	-
	2	0.025	0.025	3.54	4.2	-0.18	0.41
	3	0.063	0.078	18.8	17.2	0.34	0.40
	4	0.105	0.145	19.5	18.5	0.54	0.00
	5	0.154	0.214	17.8	16.7	0.18	-0.10
0.40	1	0	0	-	-	-	-
	2	0.025	0.025	3.7	4.7	0.05	0.24
	3	0.071	0.089	16.2	19.9	-0.28	0.20
	4	0.124	0.176	16.4	15.8	0.02	-0.35
	5	0.180	0.258	15.6	14.2	0.04	-0.39

* The corresponding values of C_0 (Eq. 3) were 0.025, 0.049, 0.073, and 0.096 for Generations 1, 2, 3, 4 and 5 respectively.

Uniform level of contamination among seed parents. Variation among seed parents in c will surely occur, especially according to location within a stand, and will reduce the effective number of seed parents represented in the next generation, through promoting differential representation of seed parents among the top phenotypes (and genotypes). This would tend to reduce the genetic base of the gene resource, especially with respect to the population to be preserved.

Single-trait model. While a combination of traits can in principle be addressed as a composite trait with its own h^2 the situation is will be complicated by how natural selection in *P. radiata* in New Zealand often favours the 'multinodal' or 'short-internode' branching (BURDON, 1992; BURDON *et al.*, 1992a, 1997) for which there has also been strong artificial selection. This will tend to favour contaminants for the next generation. Deliberate selection for long 'internodes' will tend to avoid contaminants, but genetic upgrading of the gene resource will be hampered by the adverse genetic correlations in *P. radiata* between internode length, on the one hand, and the remaining growth and form traits, on the other. How selection in the gene resource fits or conflicts with the pattern of genetic correlations, will influence effective heritability for the multi-trait selection criterion, largely through varying the effective value of σ^2_A.

Same rate of genetic advance in contaminating population as in uncontaminated gene resource. This is very arbitrary, but was governed by the existence of likely countervailing effects. On the one hand, the shorter generation time and more precise selection that would apply in the breeding population would tend to increase $\Delta_{(t,t-1)}$ with time in contaminating stands, thereby increasing the 'improvement gap' between the two classes of material. On the other hand, there will be some selection, for producing the commercial stands, for quality-related traits that cannot be readily addressed in a gene resource, and such selection will be pursued at the expense of gain in the traits under selection in the gene resource; moreover, selection in a breeding population may be approaching the limits of response in the silvicultural traits. Overall, this assumption remains an important area of uncertainty, especially in view of the evident importance of Δ_1. However, it emerged fortuitously (details not shown) that expected C was notably insensitive to an appreciable error in calculating $\Delta_{(t,1)}$.

Same values of σ^2_A in both contaminating population and uncontaminated gene resource. An arguable assumption of convenience, this works from the semi-empirical decline in σ^2_A with time in uncontaminated gene resource (*Figure 1*). While σ^2_A may be less in individual improved seedlots, contamination may come from a varied mix of such seedlots which would inflate σ^2_A in the contaminating pollen source. The limited impact of varying σ^2_A and h^2 on C suggests that this aspect was unimportant.

Discrete generations. For the gene-resource population this condition is inherent in the management scheme. However, the contaminating stands may represent a mosaic of different improvement levels; that could inflate σ^2_A, as above.

No allowance for inbreeding in gene-resource material. Genotypic variance will be inflated somewhat by inbreeding, but that should not greatly affect what seed parents would be selected.

No specific heterotic effects, representing hybrid superiority, resulting from crossing between contaminant pollen and the gene resource. Depending on the population being kept as an unpedigreed gene resource, appreciable heterosis could accentuate the selective advantage of the contaminant offspring (BURDON, 1992). Such offspring, however, may be readily identifiable.

No specific effect of selection on proportion of alleles from any pollen source, C being determined entirely by pedigree as it varies in the simulated model. Effects such as minority advantage are speculative, and even if they did occur their impact is unclear.

The findings

Only a few generations were involved, but they still represent a long timespan, particularly if generations are deliberately extended to delay the build-up of C.

Increasing N and n, while it greatly increased computing time, had a modest impact on mean C. Not surprisingly, it reduced the mean C slightly for the parameter combinations tested. Given this, and the computing demands, exhaustive comparisons for 160 seed parents were not pursued.

Values of C were very sensitive to variation in Δ_1 (*Table 3*), suggesting that the intensive selection of seed parents in the gene resource can be very important, yet obtaining a realistic value for Δ_1 is not easy. Considering the case of *P. radiata*, it has been difficult to relate values of Δ_1 at all precisely to differences in the scale of measuring genetic improvement (Vincent, 1987), which purports only to being monotonic, rather than linear. Also there is some *a priori* basis, including complications from inbreeding depression, for expecting an inflated Δ_1 where native-stand collections are involved. Indications are (Burdon *et al.,* 1997, 1999) that Δ_1 can vary widely according to both the native population concerned and the planting site, as well as the contaminating pollen source(s), since the superiority of land-race stocks can vary widely among sites. Volcanic plateau sites in New Zealand might be chosen for their proximity to major growing operations, and the strong resolution of genetic differences that help upgrading, through selection, of gene-resource stands. However, on these sites pollen contamination pressures will be intense and the superiority of improved local stock is probably best expressed.

Considering native populations of *P. radiata*, the genetic superiority of the improved local stock is likely to be less over Monterey than over other populations – on some sites it could even be negative (cf Burdon *et al.,* 1997). For the Cambria population, especially, its susceptibility to Dothistroma needle blight could place purebred trees at a major selective disadvantage on some major New Zealand site categories that are of high disease hazard, thus increasing Δ_1. Locating gene-resource plantings on sites where the superiority of local stocks is less strongly expressed, while it should reduce C, is likely to reduce the genetic improvement from selection in gene resources for adaptive and silvicultural traits. Also, the fact that the breeding population is managed on sites that are conducive to obtaining rapid genetic gain may lead to $\Delta_{(t+1,t)}$ being greater in relation to the corresponding genetic advance within the gene-resource planting. Overall, the two values of Δ_1 appear to represent a realistic range for Californian mainland populations, considering the different populations and different planting sites. With the Guadalupe Island population, offspring from contaminant pollen are likely to be markedly superior, but they should be readily identifiable. Moreover, there is a commitment to maintain purebred Guadalupe material as control-crosses between select parents. With the Cedros Island population establishment of an unpedigreed gene-resource planting in New Zealand failed. However, if it had succeeded pollen-contaminant offspring in the next generation would be easily recognisable because of much faster growth (Burdon, 1992).

Where contamination is likely it may be important to quantify it and, if possible, identify offspring of contaminant pollinations. This will probably require 'Type 2' genetic markers, which can differentiate sensitively between populations by being private or semi-private alleles within the various populations of interest. That may be a major challenge. It is likely to be most effective in first-generation contaminants, which may be candidate seed parents for perpetuating the managed gene resourecs.

Conclusions

- The results strongly support the notion that the pressures of contamination can be greatly reinforced by the selective advantage of individuals fathered by the contaminating population.
- The selective advantage conferred by paternity from the contaminating population can often be unclear, but it can evidently depend strongly on both site and the provenance represented in a gene-resource planting.
- Siting may therefore be an important aid to reducing cumulative contamination pressures by reducing the selective advantage of trees of paternity from improved local stocks.
- However, choice of sites that reduce the selective advantage may slow the improvement of gene resources for silvicultural traits.
- Development of 'Type 2' markers, which serve to differentiate populations by being private or semi-private alleles within native populations, may be highly desirable.
- Such markers would not only reveal the magnitude of a contamination problem but also offer a means of countering the problem by allowing individuals resulting from contaminant pollen to be identified among the candidate seed parents.

Acknowledgment

We are grateful for critical input from an anonymous referee.

References

Burdon, R. D.: Gene resource management plan. Ch. 4 (pp. 51–69) *in* Shelbourne, C. J. A. et al. "Development Plan for Radiata Pine Breeding". NZ For. Serv., For. Res. Inst., Rotorua, New Zealand (1986). — Burdon, R. D.: Recruitment for breeding populations: objectives, genetics and implementation. Pp. 555–572 *in* Weir, B. S. et al. (Eds) Proc. 2nd Internat. Conf. on Quantitative Genetics. Sinauer, Sunderland, MA, USA (1988). — Burdon, R. D.: Genetic survey of *Pinus radiata*. 9: General discussion and implications for genetic management. NZ J. For Sci. **22**: 274–298 (1992). — Burdon, R. D.: Future directions in tree breeding: some questions of what to seek and how to manage the genetic resource. Pp. 24–39 *in* Lavereau, J. (Ed.) "Evolution and Tree Breeding", Proc. Can. Tree Impr. Ass./Western For. Genet. Ass. Conf., Victoria, BC, Canada (1995). — Burdon, R. D.: Genetic diversity for the future: Conservation or creation and capture? Pp. 237–246 *in* Burdon, R. D. and Moore, J. M. "IUFRO '97 Genetics of Radiata Pine", Proc. NZFRI-IUFRO Conf. 1–4 Dec. and Workshop 5 Dec., Rotorua , New Zealand. FRI Bull. No. 203 (1997). — Burdon, R. D., Bannister, M. H. and Low, C. B.: Genetic survey of *Pinus radiata*. 2: Population comparisons for growth rate, disease resistance, and morphology. NZ J. For. Sci. **22**: 138–159 (1992a). — Burdon, R. D., Bannister, M. H. and Low, C. B.: Genetic survey of *Pinus radiata*. 5: Between-trait and age-age correlations for growth rate, morphology, and disease resistance. NZ J. For. Sci. **22**: 211–227 (1992b). — Burdon, R. D., Bannister, M. H., Madgwick, H. A. I and Low, C. B.: Genetic survey of *Pinus radiata*. 1: Introduction, description of experiment, and basic methodology. NZ J. For. Sci. **22**: 119–127 (1992c). — Burdon, R. D., Firth, A., Low, C. B. and Miller, M. A.: Native provenances of *Pinus radiata* in New Zealand. NZ [J.] For. **41**(4): 32–36. (1997). — Lin, C. Y. and Allaire, F. R.: Heritability of a linear combination of traits. Theor. Appl. Genet. **51**: 1–3 (1977). — Lindgren, D. and Nilsson, J. E.: Calculations concerning selection intensity. Swedish University of Agricultural Sciences, Umeå, Dept. For. Genet. & Plant Physiol. Report No. 5. (1985). — Vincent, T. G.: Certification system for forest tree seed and planting stock. NZ Ministry of Forestry, For Res. Inst., FRI Bull. 134 (1987). — Zobel, B. J. and Talbert. J.T.: Applied Forest Tree Improvement. John Wiley & Sons. (1984).