ORIGINAL PAPER

Optimization of combined genetic gain and diversity for collection and deployment of seed orchard crops

T. Funda · M. Lstibůrek · P. Lachout · J. Klápště · Y. A. El-Kassaby

Abstract Genetic gain and diversity of seed orchards' crops are determined by the number of parents, their breeding values and relatedness, within-orchard pollination efficiency, and level of pollen contamination. These parameters can be manipulated at establishment by varying clonal representation (e.g., linear deployment), during orchard development by genetic thinning, or by selective harvesting. Since clonal fecundities are known to vary both within and among years, then each seed crop has a unique genetic composition and, therefore, crops should be treated on a yearly basis. Here we present an optimization protocol that maximizes crop's genetic gain at any desired genetic diversity through the selection of a subset of the crop that meets both parameters. The genetic gain is maximized within the biological limit set by each clone's seed-cone production and effective population size is used as a proxy to genetic diversity whereby any relationship among clones is considered. The optimization was illustrated using 3 years' reproductive output data from a first-generation western larch seed orchard and was tested under various scenarios including actual male and female reproductive output and male reproductive output assumed to be either equal to that of female or a function of clonal representation. Furthermore, various levels of co-ancestry were assigned to the orchard's clones in supplementary simulations. Following the optimization, all solutions were effective in creating custom seedlots with different gain and diversity levels and provided the means to estimate the genetic properties of composite seedlots encompassing the remaining "unused" seed from a number of years.

Keywords Seed orchard · Optimization · Genetic gain · Diversity · Western larch

Communicated by R. Burdon

T. Funda · Y. A. El-Kassaby (✉)
Department of Forest Sciences, Faculty of Forestry,
University of British Columbia,
2424 Main Mall,
V6T 1Z4 Vancouver, BC, Canada
e-mail: y.el-kassaby@ubc.ca

T. Funda · M. Lstibůrek · J. Klápště
Department of Dendrology and Forest Tree Breeding,
Faculty of Forestry and Wood Sciences,
Czech University of Life Sciences Prague,
Kamýcká 129,
Prague 6, 165 21, Czech Republic

P. Lachout
Department of Probability and Statistics, Faculty of Mathematics
and Physics, Charles University in Prague,
Sokolovská 83,
Prague 8, 186 75, Czech Republic

P. Lachout
Institute of Information Theory and Automation,
Czech Academy of Sciences,
Pod vodárenskou věží 4, Praha 8,
182 08 Prague, Czech Republic

Introduction

Seed orchards are the main delivery vehicle for coniferous species' genetic improvement programs, in which balancing gain and diversity is a major concern. The level of genetic gain (genetic worth) of a seed orchard's crop depends primarily on the genetic superiority of the selected parents, their actual gametic contribution to the resultant seed crop, and the level of gene flow from extraneous pollen sources (pollen contamination) and their respective genetic quality (Stoehr et al. 1998; Lindgren et al. 2004; Slavov et al. 2005). The genetic diversity of a seed crop is greatly

influenced by the magnitude of parental fertility variation (unequal parental contribution to the resulting seed crops) (Xie et al. 1994; Kjær 1996; Burczyk and Chalupka 1997) as well as by the level of kinship among the orchard's parental population (Lindgren and Mullin 1998). Theoretically, maximum genetic diversity for a given number of unrelated and non-inbred genotypes is attained when all contribute equally to the seed crop while gain depends on the proportional contribution of high-breeding-value parents (note that maximum gain could only be attained from the best clone; however, this is not the objective of seed orchards).

The reproductive choreography—the time, duration, and extent of reproductive activity of orchard's parents—is of great importance because it affects the pattern of interpollination among parents within the seed orchard, as well as pollen contamination, and ultimately the genetic composition of the resulting seed crop (El-Kassaby et al. 1984; El-Kassaby and Ritland 1986; Matziris 1994, Burczyk and Chalupka 1997). Variation in gametic contribution among an orchard's parents may lead to over-representation of the most productive genotypes (Kjær 1996), accumulation of co-ancestry, and ultimately to erosion of genetic diversity (Lindgren et al. 1996). Fecundity variation among orchards' parents should therefore be taken into consideration when genetic gain and diversity are estimated (Kang and El-Kassaby 2002).

In reality, fecundity variation is very common in seed orchards and is caused by: (1) unequal parental representation during orchards' establishment (e.g., linear deployment where parents are linearly represented in the orchard population in proportion to their breeding values) (Lindgren and Matheson 1986), (2) mortality during orchards' development (El-Kassaby and Cook 1994), and (3) simple differences in parental reproductive output (Sitka spruce: El-Kassaby and Reynolds (1990); Douglas-fir: El-Kassaby and Askew (1991) and El-Kassaby et al. (1990); loblolly pine: Askew (1988); radiata pine: Griffin (1982); Japanese red pine: Kang (2000); and black pine: Kang et al. (2004)). Parental gametic contributions can be estimated using different methods ranging from low-cost/low-accuracy to relatively high-cost/high-accuracy. The former include phenotypic assessment of reproductive strobili output variability and seed production (Woods 2005), while the latter utilize variation revealed at isozymes or DNA markers (Roberds et al. 1991; Xie and Knowles 1992; Moriguchi et al. 2004; Slavov et al. 2005; Funda et al. 2008).

Genetic gain (response) and diversity of seed orchards' crops can be manipulated at orchard establishment by varying parental representation following linear deployment (Lindgren and Matheson 1986), during genetic thinning of existing orchards (i.e., reverse proportionate removal of low-breeding-value parents; Bondesson and Lindgren 1993), or by practicing selective seed-cone harvest (i.e., seed-cones are collected from parents with higher breeding values; Lindgren and El-Kassaby 1989; Kang et al. 2001). Assuming unrelated and non-inbred parents, seed crops' genetic diversity is at maximum under equal parental gametic output; however, the magnitude of female and male fertility variation and the direction as well as the strength of their correlation, strongly affect these practices (Kang and El-Kassaby 2002). Most seed orchards' fecundity variation studies have shown that male and female reproductive output is either weakly correlated (Nikkanen and Velling 1987; Caron and Powell 1989; Kjær 1996; Kang 2000) or independent (Savolainen et al. 1993; Kang and Lindgren 1998, 1999; Kang et al. 2004). Therefore, the common practice of estimating the genetic diversity based on one sex alone (e.g., seed production) is often inaccurate (Kang and Lindgren 1999). The above-mentioned review implies that genetic gain and diversity of a seed orchard's crop are unique within any given year, thus an orchard's reproductive output assessment and its manipulation are better done on a yearly basis.

The objective of this study is to develop a model that maximizes the genetic value of a selected portion of an existing orchard crop, through the optimization of parental contributions. It is designed to account for relatedness among parents, an issue that is becoming increasingly important as advanced-generation seed orchards are established. The need to account for relatedness has been demonstrated in animal breeding (e.g., Meuwissen 1997) and it is prudent to consider it also during the deployment of reproductive material from forest-tree production populations.

Reproductive output data from a western larch (*Larix occidentalis* Nutt) seed orchard were used to illustrate the proposed approach and three different scenarios were tested: (a) equal male and female fecundities (i.e., presence of a perfect correlation between female and male reproductive output (see Lindgren et al. 2004; Prescher et al. 2006)); (b) male contribution proportional to the parental representation (number of ramets per clone); and (c) actual male and female reproductive output data (Stoehr et al. 2004). In the current study, a new approach that combines the above-mentioned elements in a single-optimization step is presented. The optimization leads to determining optimum clonal reproductive output proportions after considering actual male and female reproductive outputs, parental breeding values, co-ancestry, and the desired level of genetic diversity.

Materials and methods

Seed orchard and fecundity assessment

The proposed optimization approach is illustrated using total male and female reproductive output records collected

over three consecutive years (2004–2006) in a western larch clonal seed orchard (57 clones). The orchard is located near Vernon, BC, Canada (Kalamalka Forestry Centre: altitude 480 m, latitude 50°14' N, longitude 119°16' E) and was established in 1989 using trees selected from natural stands within south-eastern British Columbia prior to any genetic evaluation of the parental population. The genetic composition of the orchard maintained a dynamic state as a result of replacing low-breeding-value parents by higher-ranking ones based on subsequent progeny test evaluation. The orchard's parental breeding values (height at age 60) were predicted based on two series of 10-year-old progeny test trials located within the parental natural range (East Kootenay, BC). Maternal (female) and paternal (male) contributions to the orchard's total seed crop were estimated based on the number of seed-cones and volume of pollen-cones produced by each ramet within every clone, respectively (i.e., 100% survey), following the methods by Woods (2005). All parents in the orchard were assumed to be unrelated and non-inbred; however, this assumption was intentionally relaxed in several instances so that the issues concerning build-up of co-ancestry in advanced-generation seed orchards could be addressed (details are provided at the end of this section).

Pearson's product-moment correlation coefficients were calculated for male and female fecundities within and among years to illustrate the commonly observed differences in yearly reproductive output data (Kang 2000) and the difficulty of predicting future seed crops (Lindgren et al. 2007; Kroon et al. 2007). Several clones had to be excluded from the correlation analysis due to changes in the orchard's population structure over time and in some cases due to large differences in the parental representation (i.e., very small number of ramets per clone). Correlation analysis was therefore only based on 27 clones (with the average of 35.6 ramets per clone) that were present in the orchard over the three study years.

Mathematical model

Let N be the number of clones in a seed orchard with corresponding breeding values provided in a vector $\mathbf{X}$. The average breeding value (genetic response) is

$$\Delta G = \sum_{i=1}^{N} x_i p_i \tag{1}$$

where p_i denotes the contribution of clone i to the orchard crop and $0 \leq p_i \leq 1$ and

$$\sum_{i=1}^{N} p_i = 1 \tag{2}$$

The gametic contribution of clone i is split into its respective female (f_i) and male (m_i) components such that

$$p_i = \left(\frac{f_i + m_i}{2} \right) \tag{3}$$

where $0 \leq f_i \leq 1$, $0 \leq m_i \leq 1$, $\sum_{i=1}^{N} f_i = 1$ and $\sum_{i=1}^{N} m_i = 1$.

Genetic diversity of a seed orchard crop can be expressed as the effective population size, which describes the proportion of parents involved in the production of the seed crop. Considering related and/or inbred individuals, we used the concept of status number (Lindgren et al. 1996; referred to as N_e throughout this study) to calculate the effective population size as follows:

$$N_e = \frac{1}{2\Theta} \tag{4}$$

where Θ is group co-ancestry (Cockerham 1967). Group co-ancestry is defined as the average co-ancestry of all pairs of population members including individuals with themselves while co-ancestry is the probability that any two alleles sampled at random (one from each individual) are identical by descent (IBD, Malécot 1948).

Let us assume N parents contributing to the resulting gametic pool (seed orchard crop). Selecting alleles randomly from the gametic pool, with replacement, the probability that the first allele originates from genotype i is r_i, and the probability that the second originates from genotype j is r_j. The likelihood that these two alleles are IBD is c_{ij}, which is the coefficient of co-ancestry between genotypes i and j (an element in the $N \times N$ co-ancestry matrix $\mathbf{C}$). Note that this coefficient is also referred to as coefficient of kinship (Falconer and Mackay 1996), which increases with the level of relationship (in outbred populations, it equals to 1/8 for half-sibs, 1/4 for full-sibs and parent–offspring, and 1/2 for selfing). The probability (Θ) that a pair of alleles sampled from the gamete pool is IBD is given by adding over all possible probabilities (Lindgren and Mullin 1998), thus

$$\Theta = \sum_{i=1}^{N} \sum_{j=1}^{N} r_i r_j c_{ij} \tag{5}$$

Assuming that all individuals are known starting from the founder population, then matrix $\mathbf{C}$ can be recursively calculated from a given lineage (e.g. Emik and Terrill 1949). Any $\mathbf{C}$, if correctly specified and thence internally consistent, is a positive definite matrix with diagonal elements between 0.5 and 1 (the diagonal elements represent individuals' self co-ancestries), which is a prerequisite for its inclusion in the proposed optimization protocol.

Model constraints

The optimization's objective is to maximize the genetic gain in any given seed orchard's crop (i.e., the maximization of the function given in Eq. 1). Obviously, this would be reached if seed from the best parent or a limited number of parents with the highest breeding values were collected. However, in such a case, N_e of the resulting mixture would be unreasonably low and the seed crop could not be utilized due to its genetic vulnerability. Therefore, the minimum level of N_e was set as a constraint. This was derived out of Eqs. 4 and 5 as follows:

$$\Theta \le \frac{1}{2N_{e\,min}} \Leftrightarrow \sum_{i=1}^{N} \sum_{j=1}^{N} p_i p_j c_{ij} \le \frac{1}{2N_{e\,min}} \quad (6)$$

where $N_{e\,min}$ stands for the minimum desired N_e. This value is provided as an input to the optimization. If male contributions differ from those of females, Eq. 6 can be adjusted correspondingly (using Eq. 3):

$$\Theta \le \frac{1}{2N_{e\,min}} \Leftrightarrow \frac{1}{2} \sum_{i=1}^{N} (m_i + f_i) \sum_{j=1}^{N} (m_j + f_j) c_{ij} \le \frac{1}{N_{e\,min}} \quad (7)$$

A second constraint to the solution was given in Eq. 2.

Limits to gametic contributions

Furthermore, limits to female gametic contributions were set to reflect the amount of seed produced by each clone. Let s be a fraction of the total amount of seed crop collected in a given year (reflecting the actual seed need and the minimum desired level of diversity in such a subset), where $0 \le s \le 1$ (value of 1 represents a situation where the entire seed crop is selected). We therefore set the upper bound of the female gametic contribution of clone i as the total amount of seed collected from the same clone (denoted as b_i) relative to the total amount of seed available in the entire orchard in a given year ($\sum_{i=1}^{N} b_i$):

$$0 \le f_i \le \min\left(1, \frac{b_i}{s \sum_{i=1}^{N} b_i}\right) \quad (8)$$

Optimization

One could study individual relationships encapsulated within the equations presented earlier; however, our goal was to combine all of them into a single model. This leads to a rather elaborate problem with many variables (each clone constitutes a variable) and corresponding dimensions. Due to the complexity of the problem, we decided to use a mathematical programming approach, whereby the objec-

tive function is maximized, subject to all variables' constraints.

The optimization software MOSEK® ApS (Anonymous 2002) was used to search for the optimum solution after declaring the problem in a correct mathematical programming format. Throughout the study, we used the default optimality tolerance level, meaning that the solutions presented and the actual optima can be considered identical. To assess the sensitivity to violations of various assumptions often used in similar studies (e.g. Kang et al. 2001; Kang et al. 2003; Lindgren et al. 2004; Prescher et al. 2006), we compared three scenarios. In Scenario A, male and female contributions were assumed equal (i.e., perfect correlation between male and female reproductive outputs) and only actual female seed-cone counts were used. In this case, the male component cancels out and p_i equals to the seed production of the ith clone. Note that this scenario is mathematically identical to the concept of "genetic thinning" implemented by Bondesson and Lindgren (1993). Next, under Scenario B, male contributions were considered to be a function of the clonal representation (the total number of ramets per clone) while the female contributions were based on actual seed-cone count. Finally, under Scenario C, actual estimates of both male and female gametic contribution were used. We used reproductive output data from one seed orchard of a single species to illustrate the benefits of this new optimization approach. It should be stated that although high correlation is assumed between reproductive energy (production of seed-cones) and reproductive success (production of viable seeds) in this study, the optimization protocol does not rely on this assumption at all, as it can handle any reproductive output data available such as the number of seeds, filled seeds, germinated seeds or seedlings, depending on the resources invested during the fecundity variation assessment.

In this study, $N_{e\,min}$ was set to a value of 10 and 15 for illustration only; however, N_e of 10 is generally considered to be the minimum effective population size required to capture the majority of genetic diversity in a population (Yanchuk 2001). It is up to seed orchard managers to determine any minimum level of genetic diversity they desire, but this must be considered along with the crops' genetic breadth. Since the study orchard consists of unrelated and non-inbred clones, a certain level of relatedness was assigned to 30 clones whereby a hypothetical advanced-generation orchard with co-ancestry was simulated (see the relationships below). This was conducted to demonstrate the impact of the relatedness on the clonal optimum contributions and effective population size estimates. Furthermore, the original first-generation breeding values were retained. Two examples with the total number of 44, 40, and 47 parents were tested (the seed orchard's census numbers in 2004, 2005, and 2006, respectively), 30

of which were related (various levels of co-ancestry) and the remaining ones were unrelated. The first example consisted of 26 half-sibs and five full-sibs (moderate relatedness) while the second consisted of 13 half-sibs, seven full-sibs and 15 parent–offspring (strong relatedness). All variables employed by the optimization protocol are listed in Table 1.

Results

Correlation

Over the 3 years, with the exception of female (2005–2006), all observed correlation coefficients between male, female, and male–female fecundities were positive and significant (Table 2). Correlations between male fecundity (Table 2, above diagonal) were, on average, slightly higher than those observed for female fecundity (Table 2, below diagonal). These values imply that yearly male and female reproductive outputs are similar; however, close observation of the data on the individual ramet level indicated the presence of greater variability (data not shown). In fact, the clonal total reproductive output is the main factor affecting the genetic constitution and quality of the seedlots and not the clonal size (i.e., number of ramets/clone). Male–female reproductive output correlation varied among years (Table 2), indicating that yearly crops are unique and should be evaluated on a year-by-year basis.

Genetic gain and diversity estimates

The genetic gain estimates of the orchard's crops (100% sampling) varied slightly over the 3 years, depending on the scenario applied (Fig. 1). When only the female component was considered (i.e., assuming perfect positive correlation between male and female gametic contribution), the genetic gain estimates were either under- or over-estimated. The observed differences were caused by the relatively higher production of pollen from either high- or low-breeding-value clones, respectively (Fig. 1, 100% sampling). The

effective population size of the entire orchard crop was found to be the highest in Scenario B (male reproductive output set as a function of clonal representation), ranging from 27.3 to 33.8 for the 3 years, whereas the other two scenarios that take the male component into account reached similar but substantially lower values (Fig. 1). The former approach is rather unrealistic because it assumes male reproductive output equality among ramets within clones which inflates N_e. Using Kang and Lindgren's (1998; 1999) method for estimating male fecundity variation among clones (the concept of sibling coefficient); Scenario B produced estimates of 1.23, 1.24, and 1.51 while Scenario C produced estimates of 2.50, 2.36, and 2.84 for 2004, 2005, and 2006, respectively. These estimates are high and are indicative of extensive male reproductive output variability.

Optimization

The proposed protocol was illustrated by performing the optimization in a stepwise manner that excluded the cone crop in an increment of 10% while maintaining a minimum N_e of 10 (Figs. 1 and 2a) and 15 (Fig. 2b). An appreciable increase in genetic gain was attained by sacrificing some genetic diversity through relaxing the effective population size constraint. As expected, exclusion of seed from lower-breeding-value clones was associated with a steady increase in genetic gain across the three scenarios over the study years.

Under no relatedness (Fig. 1), Scenario A produced a relatively steeper gain curve compared to Scenarios B and C. This result can be attributed to the fact that genetic gain calculations for the former were restricted to the subset of clones from which seed had been selected while male contribution from the remaining clones was not taken into account. It is important to note that N_e of 10 was attained with fewer seed parents under Scenarios B and C due to the inclusion of all pollen contributors. That means that the genetic diversity reduction associated with focusing on fewer females is counteracted by the inclusion of many parents as males. For instance, for 2006 crop, seed from as few as seven parents would secure the attainment of $N_e=10$

Table 1 List of variables used in the optimization protocol

	Variable	Symbol
Input	Vector of male gametic contributions	**M**
	Vector of upper bounds set by clonal seed production	**B**
	Vector of breeding values	**X**
	Co-ancestry matrix	**C**
	Selected fraction of total seed crop	s
	Minimum desired effective population size	$N_{e\ min}$
Output	Vector of female gametic contributions (forming custom seedlots)	**F**

N_e is expressed as status effective number (Lindgren et al. 1996)

Table 2 Pearson's correlation coefficients between female (below diagonal), male (above diagonal), and between male and female fecundities (diagonal)

Year	2004	2005	2006
2004	0.62[a]	0.50[a]	0.60[a]
2005	0.36[b]	0.48[a]	0.77[a]
2006	0.83[a]	0.28[ns]	0.62[a]

N=27, r critical value=0.330 (α=0.05) and 0.453 (α=0.01)

ns not significant

[a] Highly significant (α=0.01)

[b] Significant (α=0.05)

when 20% of the crop was selected (Fig. 1), thus demonstrating the benefits of including male reproductive output in the genetic diversity calculations.

Relatedness among the orchard's clones drastically affected the genetic gain and diversity estimates (Fig. 2a, b), although it should be noted that the genetic gain is only linked to this quantity indirectly. As expected, the higher the genetic relatedness among parents the greater the reduction in the expected diversity estimates. For instance, when strong relatedness was assigned to orchards' clones as defined in "Materials and methods" section, N_e of the entire crop (100% sampling) dropped from 22.5, 19.6, and 20.9 to 15.5, 12.8, and 14.5 for 2004, 2005, and 2006, respectively. Further reduction was associated with the crop proportion selected by the optimization and N_e eventually reached a value of 10 and 15 (Fig. 2a, b, respectively), depending on the constraint set prior to optimization. Conversely, the genetic gain estimates increased with the decrease in the crop proportion selected (i.e., the inclusion of seed from

fewer good parents). When the constraint of the effective population size was further relaxed ($N_{e\ min}$= 10), the effect of relatedness on the attained gain was only restricted to small proportions of the seed selected ($s\cong$0.1; Fig. 2a), i.e., the optimum proportions were equal regardless of relatedness present in the orchard. However, under higher N_e ($N_{e\ min}$= 15; Fig. 2b), the genetic gain reduction became apparent along the whole range of the X-axis (curves became more divergent), indicating that the increase in genetic diversity forces the inclusion of additional female parents with lower breeding values. Note that under strong relatedness, when N_e is set at 15, no solutions could be attained when more than 70% (2005) and 90% (2006) of seed crop was retained. This situation generally arises when all the options of manipulating clonal proportionate representation are attempted without reaching any combination yielding$\geq N_{e\ min}$ (i.e., the preset N_e exceeds a particular crop's biological limits; Fig. 2a). Furthermore, as expected, a gradual selection of smaller subsets of seed lead to a steady increase in gain in all cases tested, i.e., across all scenarios and years (Figs 1 and 2a, b). However, although the concurrent increase in gain and diversity is known to be antagonistic, an increase in the former does not necessarily have to be associated with a decrease in the latter (see Fig. 2a). For instance, 2004 crop shows a situation where N_e value fluctuates across different seed retention ratios. Although this may seem surprising, it can be simply explained as a result of the gametic composition of the "best" 40% of seed, where the effective population size is just a by-product of the optimization process, as long as it meets the preset constraint.

Finally, it should be stated that the above examples are strictly intended to illustrate the optimization procedure,
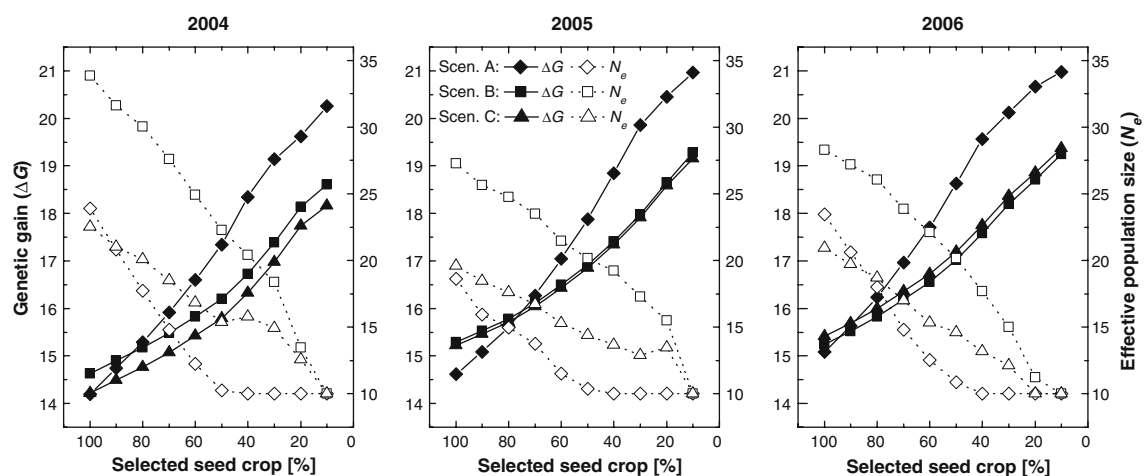
Fig. 1 Estimated genetic gain (*left axes, solid markers*) and corresponding effective population size (*right axes, open markers*) in the western larch seed orchard over 3 years (2004–2006) after a stepwise exclusion of lower-breeding-value seed constrained by $N_{e\ min}$=

10. The three pairs of curves in each year represent results from the three different scenarios applied (*A* male and female reproductive equality, *B* male as a function of clone size, and *C* actual data)
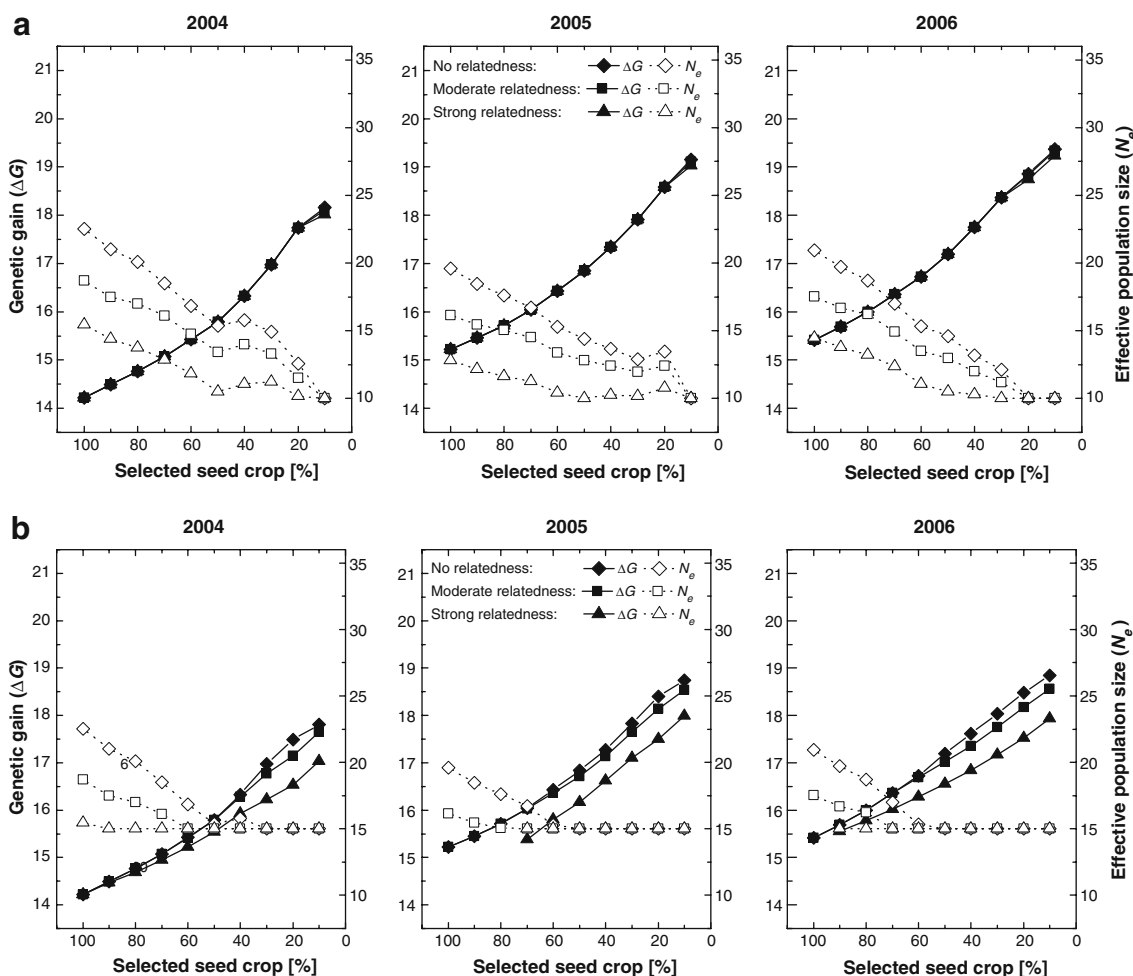
Fig. 2 a Estimated genetic gain (*left axes*, *solid markers*) and corresponding effective population size (*right axes*, *open markers*) in a western larch seed orchard over 3 years (2004–2006) after a stepwise exclusion considering breeding values and relatedness following Scenario C and constrained by $N_{e\ min}=10$. The three pairs of curves represent no, moderate, and strong relatedness among clones as specified in the "Materials and methods" (hypothetical scenario). **b** Estimated genetic gain (*left axes*, *solid markers*) and corresponding effective population size (*right axes*, *open markers*) in a western larch seed orchard over 3 years (2004–2006) after a stepwise exclusion considering breeding values and relatedness following Scenario C and constrained by $N_{e\ min}=15$. The three pairs of curves represent no, moderate, and strong relatedness among clones as specified in the "Materials and methods" (hypothetical scenario)

and the results are merely a representation of this particular orchard with its simulated co-ancestry, its observed reproductive output for the three studied crops, and the assumptions made (i.e., scenarios applied).

Discussion

The simultaneous optimization of genetic gain and diversity is the subject of intense research in animal breeding circles (e.g., Meuwissen 1997). Forest geneticists recognized this concept and efforts have been made to simultaneously optimize gain and diversity of seed orchards' crops (Son et al. 2003; Lindgren et al. 2004). These efforts include linear deployment of parents proportionate to their breeding value

in newly established orchards (Lindgren and Matheson 1986) or the application of genetic thinning of lower-breeding-value parents from established orchards in a manner that approximates linear deployment (Bondesson and Lindgren 1993; Prescher et al. 2008).

Another option to balance genetic gain and diversity is the practicing of selective seed-cone harvesting (Lindgren and El-Kassaby 1989; Kang et al. 2001). This approach has a positive impact on the genetic gain attained in the resulting seed crops and it is also expected to maintain high diversity levels since no control is imposed on pollen donors. When selective seed-cone harvesting is implemented to mimic linear deployment, the amount of seed collected from each clone is linearly proportional to its breeding value, provided the following four assumptions

are met: (1) all clones are unrelated and non-inbred (i.e., the co-ancestry matrix is of the form $0.5\mathbf{I}$, where $\mathbf{I}$ is an identity matrix), (2) mating is random, (3) reproductive output is proportional to clonal representation, and (4) clonal male and female gametic contributions are equal. In other scenarios (parents related and/or inbred, unequal clonal male/female gametic contributions, and presence of constraints due to variation in seed production), the optimum contributions become non-linear, and the deviation from linearity is progressively more accentuated with the added effects of these factors. The generalized deployment method presented in this study thus covers both linear and non-linear solutions. Bila et al. (1999) and Kang and El-Kassaby (2002) suggested constructing crops with equal female contribution (i.e., equal harvesting) to maximize the genetic diversity. It should be noted, however, that a crop's genetic diversity is at maximum only if clones are unrelated and non-inbred, or strictly speaking, if there is no variation in self and/or pair-wise co-ancestries. Furthermore, this method is limited by the seed production of the least fertile clone, thus making it inappropriate since substantial amount of seed from the more fertile clones is abandoned (Kang et al. 2003). It should be stated that both the selective seed-cone harvesting of Lindgren and El-Kassaby (1989) and attempts to equalize clonal female reproductive contributions (Bila et al. 1999; Kang and El-Kassaby 2002) can be considered as special cases of the proposed generalized optimization approach. However, neither of these methods, although expected to be effective in increasing the genetic gain and diversity, takes advantage of the crop's reproductive output assessment information.

As mentioned earlier, the commonly observed parental imbalance among an orchard's clones (Eriksson et al. 1973; Griffin et al. 1982; Askew 1988; El-Kassaby et al. 1989; El-Kassaby and Reynolds 1990; El-Kassaby and Cook 1994; Bila et al. 1999; Kjær and Wellendorf 1998; Bilir et al. 2002; Kang et al. 2004; Hansen and Kjær 2006) which was dubbed as conforming to the 80/20 rule (Anonymous 1976: asserting that 80% of the seed-cone crop is produced by as little as 20% of the parents), coupled with the observed fecundity variation over the orchard's developmental stage, environmental conditions, and management practices (El-Kassaby et al. 1989; Lindgren et al. 2004), lead us to the conclusion that the genetic composition of any orchard's crop is genetically unique. The observed among-years reproductive output correlations and their variation over time in the present study are consistent with Kang's (2000) findings in a Japanese red pine seed orchard and show that reproductive output assessments made on existing orchard for any particular year or even over multiple years are unreliable in predicting the genetic composition of future seed crops.

To demonstrate the concept of the presented optimization model and its features, we used the reproductive output data from the studied orchard (2004–2006) and produced new optimum seedlots (portions of the entire crop) which encompass the maximum possible genetic gain given the genetic diversity constraints under three modes of relatedness (no, moderate, and strong relatedness). Note that while the optimization produces various proportional contributions on the female parents (i.e., the amount of seed from a particular clone entering the seedlot); the male component is also simultaneously considered. Furthermore, although the proposed approach results in the utilization of only a subset of the seed crop based on a preset culling criterion, the unused portion of the seed crop is still usable and its properties in terms of the genetic gain and diversity could be estimated. For instance, the remaining seed after selective harvesting through optimization attained the required minimum N_e of 10 in nearly all possible cases (results not shown), allowing the capture of approximately 95% of the genetic diversity of the seed orchard population (Yanchuk 2001). If seed selection is done to attain high genetic gain relative to the original crop, then the remaining seed portion will be substantial and its genetic diversity will be high due to its large number of contributing female and male parents. As to comparing the three approaches tested in the present study, there seems to be a substantial difference between Scenario A on one hand and Scenarios B and C on the other (see Fig. 1 for details). Notwithstanding the significant male–female reproductive output correlation, it appears that consideration of the male and female gametic contribution into the optimization is relevant for obtaining more reliable estimates of genetic gain and diversity. The observed small differences between Scenarios B and C in 2005 and 2006 may suggests that estimating pollen production solely based on clone size seems to be sufficient and relatively accurate for these 2 years; however, it can be misleading as in 2004, highlighting the caveat of this assessment method (Fig. 1).

The presented optimization is based on reproductive output data and uses this assessment as a proxy for clonal gametic contribution. The assumption of congruence between reproductive energy and reproductive success is accepted for neither female (Reynolds and El-Kassaby 1990; El-Kassaby and Cook 1994) nor male (Schoen and Stewart 1986; 1987) components. Since the optimization protocol is developed to maximize gain at a desired diversity when crops are harvested by clones, then any discrepancy between reproductive energy and success is confined to the male component. Evidences of correspondence between pollen-cone production and siring success are becoming available when genetic markers are used (e. g., Roberds et al. 1991; unpublished data). Furthermore, the protocol assumes no pollen contamination. However, even

if the rate of pollen contamination is known, then its consideration as input into optimization only appears relevant if clones show large phenological differences. Xie et al. (1994) evaluated the effect of reproductive attributes in seed orchards on estimating N_e and concluded that male and female outputs have greater importance than other attributes, including phenology. The authors also point out that the role of reproductive phenology becomes pronounced only in small orchards and under substantial asynchrony. When pollen contamination is expected to be uniform across orchard's clones, the resultant optimum contributions to any designed seedlot remain unchanged and the genetic gain is calculated a posteriori. It must be stated, however, that the genetic gain estimate obtained by the optimization protocol is still at the maximum possible value under any pollen contamination rate.

Most studies concerned with estimating genetic gain and diversity of seed orchards' crops have been limited to first-generation seed orchards. However, it has recently become obvious that omitting genetic relatedness among clones is no longer sustainable in seed orchard management and that an improved theoretical basis for dealing with advanced-generation material is needed (Prescher 2007; Lindgren and Danusevicius 2008). For instance, Bila et al. (1999) demonstrated rapid build-up in group co-ancestry over generations in a simulation study on a teak stand where fertility variation differed across scenarios but was kept constant over generations. As expected, the steepest increase in group co-ancestry was observed when both female and male fertilities varied, while the slowest was observed when fertilities were assumed equal across all parents in the stand, resulting into a considerably steeper decline in effective population size in the former instance. This optimization approach reduces the number of assumptions employed by combining genetic gain and diversity of actual rather than just predicted seed orchard crops. It is perfectly suitable for both first-generation and advanced-generation seed orchards, as it is capable of limiting individual female and male gametic contributions, while accounting for relatedness. Owing to its ability of creating custom seedlots with multiple gain levels, we expect that it will be particularly advantageous in situations where seed crops exceed need or where more extensive site management will be applied (matching specific sites productivity). The surplus unused seed can still be utilized if it meets acceptable minimum genetic diversity and in situations where mixing of surpluses from multiple years is possible. In these cases, the same optimization protocol could be used to design crops with desired genetic gain and diversity. Furthermore, this protocol can be applied in situations where a given seed orchard's crop (100% of seeds) does not meet the minimum desired level of genetic diversity. This is likely to occur in rogued and/or advanced-generation seed orchards where the number of parents is limited. This situation is further exacerbated by the build-up of co-ancestry and fertility variation. An added advantage of this approach is to design seedlots with levels of genetic diversity exceeding that of the total seed crop (N_e 10 vs. 15, Fig. 2a, b). This is accomplished through sacrificing specific proportions of seed from particular clones to maximize the gain while meeting the diversity constraint. For instance, under strong relatedness, the entire seed crop's N_e of 12.8 was increased to 15 after excluding as little as 30% of the crop (Fig. 2b). Finally, this optimization could extend the lifespan of older orchards through the formation of seed crops with genetic gain exceeding orchards' averages.

References

Anonymous (2002) The MOSEK.NET API manual. Version 4.0 (Revision 60). Mosek ApS, Denmark, p 341

Anonymous (1976) Twentieth annual report on cooperative tree improvement and hardwood research program. North Carolina State University, Raleigh

Askew GR (1988) Estimation of gamete pool compositions in clonal seed orchards. Silvae Genet 37:227–232

Bila AD, Lindgren D, Mullin TJ (1999) Fertility variation and its effect on diversity over generations in a teak plantation (*Tectona grandis* L.f.). Silvae Genet 48:109–114

Bilir N, Kang K-S, Ozturk H (2002) Fertility variation and gene diversity in clonal seed orchards of *Pinus brutia*, *Pinus nigra* and *Pinus sylvestris* in Turkey. Silvae Genet 51:112–115

Bondesson FL, Lindgren D (1993) Optimal utilization of clones and genetic thinning of seed orchards. Silvae Genet 42:157–163

Burczyk J, Chalupka V (1997) Flowering and cone production variability and its effect on parental balance in a Scots pine clonal seed orchard. Ann Sci For 54:128–144

Caron GE, Powell GD (1989) Patterns of seed-cone and pollen-cone production in young *Picea mariana* trees. Can J For Res 19:359–364

Cockerham C (1967) Group inbreeding and coancestry. Genetics 56:89–104

El-Kassaby YA, Askew GR (1991) The relation between reproductive phenology and output in determining the gametic pool profile in a Douglas-fir seed orchard. For Sci 37:827–835

El-Kassaby YA, Cook C (1994) Female reproductive energy and reproductive success in a Douglas-fir seed orchard and its impact on genetic diversity. Silvae Genet 43:243–246

El-Kassaby YA, Reynolds S (1990) Reproductive phenology, parental balance and supplemental mass pollination in a Sitka spruce seed orchard. Forest Ecol Manag 31:45–54

El-Kassaby YA, Ritland K (1986) The relation of outcrossing and contamination to reproductive phenology and supplemental mass pollination in a Douglas-fir seed orchard. Silvae Genet 35:240–244

El-Kassaby YA, Fashler AMK, Crown M (1989) Variation in fruitfulness in a Douglas-fir seed orchard and its effect on crop-management decisions. Silvae Genet 38:113–121

El-Kassaby YA, Fashler AMK, Sziklai O (1984) Reproductive phenology and its impact on genetically improved seed production in a Douglas-fir seed orchard. Silvae Genet 33:120–125

Emik LO, Terrill CE (1949) Systematic procedures for calculating inbreeding coefficients. J Heredity 40:51–55

Eriksson G, Johnsson A, Lindgren D (1973) Flowering in a cone trial of *Picea abies* Karst. Studia Forestalia Suecica 110:45

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th edn. Harlow, Longman

Funda T, Chen C, Liewlaksaneeyanawin C, Kenawy A, El-Kassaby YA (2008) Pedigree and mating system analyses in a western larch (*Larix occidentalis* Nutt.) experimental population. Ann For Sci 65:705

Griffin AR (1982) Clonal variation in radiata pine. I. Some flowering, cone and seed production traits. Australian Forest Research 12:295–302

Hansen OK, Kjær ED (2006) Paternity analysis with microsatellites in a Danish *Abies nordmanniana* clonal seed orchard reveals dysfunctions. Can J For Res 36:1054–1058

Kang K-S (2000) Clonal and annual variation of flower production and composition of gamete gene pool in a clonal seed orchard of *Pinus densiflora*. Can J For Res 30:1275–1280

Kang K-S, El-Kassaby YA (2002) Considerations of correlated fertility between genders on genetic diversity: the *Pinus densiflora* seed orchard as a model. Theor Appl Genet 105:1183–1189

Kang K-S, Lindgren D (1998) Fertility variation and its effect on the relatedness of seeds in *Pinus densiflora*, *Pinus thunbergii* and *Pinus koraiensis* clonal seed orchards. Silvae Genet 47:196–201

Kang K-S, Lindgren D (1999) Fertility variation among clones of Korean pine (*Pinus koraiensis* S. et Z.) and its implications on seed orchard management. For Genet 6:191–200

Kang K-S, Kjær ED, Lindgren D (2003) Balancing gene diversity and nut production in *Corylus avellana* collections. Scand J For Res 18:118–126

Kang K-S, Lindgren D, Mullin TJ (2001) Prediction of genetic gain and gene diversity in seed orchard crops under alternative management strategies. Theor Appl Genet 103:1099–1107

Kang K-S, Lindgren D, Mullin TJ (2004) Fertility variation, genetic relatedness, and their impacts on gene diversity of seeds from a seed orchard of *Pinus thunbergii*. Silvae Genet 53:202–206

Kjær ED (1996) Estimation of effective population number in a *Picea abies* (Karst.) seed orchard based on flower assessment. Scand J For Res 11:111–121

Kjær ED, Wellendorf H (1998) Studies on the effect of unequal flowering on the effective population number in Danish seed orchard crops. Forest Tree Improvement 26:1–9

Kroon J, Prescher F, Wennström U, Lindgren D (2007) Cone set over time for different clones in a Scots pine seed orchard. In: Prescher F (ed) Seed orchards—genetic considerations on function, management and seed procurement. Doctoral thesis No. 2007:75, The Swedish University of Agricultural Sciences, Umea, Sweden, p 22

Lindgren D, Danusevicius D (2008) Deployment of clones to seed orchards when candidates are related. In: Lindgren D (ed.)

Proceedings of a seed orchard conference. Umeå, Sweden, pp 135–141

Lindgren D, El-Kassaby YA (1989) Genetic consequences of combining selective cone harvesting and genetic thinning in clonal seed orchards. Silvae Genet 38:65–70

Lindgren D, Matheson AC (1986) Increasing the genetic quality of seed from seed orchards by using the better clones in higher proportions. Silvae Genet 35:173–177

Lindgren D, Mullin TJ (1998) Relatedness and status number in seed orchard crops. Can J For Res 28:276–283

Lindgren D, Cui J, Son SG, Sonesson J (2004) Balancing seed yield and breeding value in clonal seed orchards. New Forests 28:11–22

Lindgren D, Gea LD, Jefferson PA (1996) Loss of genetic diversity monitored by status number. Silvae Genet 45:52–59

Lindgren D, Tellalov Y, Prescher F (2007) Seed set for Scots pine grafts is difficult to predict. In: Isik F (ed.) Proceedings of the IUFRO Division 2 Joint Conference: Low input breeding and conservation of forest genetic resources. Antalya, Turkey, pp 139–141

Malécot G (1948) Les mathématiques de l'hérédité. Masson, Paris, p 65

Matziris D (1994) Genetic variation in the phenology of flowering in black pine. Silvae Genet 43:321–328

Meuwissen THE (1997) Maximizing the response of selection with a predefined rate of inbreeding. J Anim Sci 75:934–940

Moriguchi Y, Taira H, Tani N, Tsumura Y (2004) Variation of paternal contribution in a seed orchard of *Cryptomeria japonica* determined using microsatellite markers. Can J For Res 34:1683–1690

Nikkanen T, Velling P (1987) Correlations between flowering and some vegetative characteristics of grafts of *Pinus sylvestris*. Forest Ecol Manag 19:35–40

Prescher F (2007) Seed orchards—genetic considerations on function, management and seed procurement. Doctoral Dissertation No. 2007:75. Swedish University of Agricultural Sciences, Umeå.

Prescher F, Lindgren D, El-Kassaby YA (2006) Is linear deployment of clones optimal under different clonal outcrossing contributions in seed orchards? Tree Genetics & Genomes 2:25–29

Prescher F, Lindgren D, Karlsson B (2008) Genetic thinning of clonal seed orchards using linear deployment may improve both gain and diversity. Forest Ecol Manag 254:188–192

Roberds JH, Friedman ST, El-Kassaby YA (1991) Effective number of pollen parents in clonal seed orchards. Theor Appl Genet 82:313–320

Reynolds S, El-Kassaby YA (1990) Parental balance in a Douglas-fir seed orchard: cone vs. seed production. Silvae Genet 39:40–42

Savolainen O, Karkkainen K, Harju A, Nikkanen T, Rusanen M (1993) Fertility variation in *Pinus sylvestris*: a test of sexual allocation theory. Am J Bot 80:1016–1020

Schoen DJ, Stewart SC (1986) Male reproductive investment and male reproductive success in white spruce. Evolution 40:1109–1120

Schoen DJ, Stewart SC (1987) Variation in male fertilities and pairwise mating probabilities in *Picea glauca*. Genetics 116:141–152

Slavov GT, Howe GT, Adams WT (2005) Pollen contamination and mating patterns in a Douglas-fir seed orchard as measured by simple sequence repeat markers. Can J For Res 35:1592–1603

Son SG, Kang K-S, Lindgren D, Hyun JO (2003) Qualification for the value of seed orchard considering breeding value and seed productivity. J Korean Forest Assoc 91:601–608

Stoehr M, Webber J, Woods J (2004) Protocol for rating seed orchard seedlots in British Columbia: quantifying genetic gain and diversity. Forestry 77:297–303

Stoehr MU, Orvar BL, Vo TM, Gawley JR, Webber JE, Newton CH (1998) Application of a chloroplast DNA marker in seed orchard management evaluations of Douglas-fir. Can J For Res 28:187–195

Woods JH (2005) Methods for estimating gamete contributions to orchard seed crops and vegetative lots in British Columbia. B.C. Ministry of Forests and Range, Technical report 025. p 17

Xie CY, Knowles P (1992) Male fertility variation in an open-pollinated plantation of Norway spruce (Picea abies (L.) Karst.). Can J For Res 22:1463–1468

Xie CY, Woods J, Stoehr M (1994) Effects of seed orchard inputs on estimating effective population size of seedlots—a computer simulation. Silvae Genet 43:145–154

Yanchuk AD (2001) A quantitative framework for breeding and conservation of forest tree genetic resources in British Columbia. Can J For Res 31:566–576