



**HAL**  
open science

## Allocation of varietal testing efforts for implementing conifer multi-varietal forestry using white spruce as a model species

Weng, Park, Marek Krasowski, T. Mullin

► **To cite this version:**

Weng, Park, Marek Krasowski, T. Mullin. Allocation of varietal testing efforts for implementing conifer multi-varietal forestry using white spruce as a model species. *Annals of Forest Science*, 2011, 68 (1), pp.129-138. 10.1007/s13595-011-0014-1 . hal-00930710

**HAL Id: hal-00930710**

**<https://hal.science/hal-00930710>**

Submitted on 11 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Allocation of varietal testing efforts for implementing conifer multi-varietal forestry using white spruce as a model species

Yuhui Weng · Yill-Sung Park · Marek J. Krasowski ·  
T. J. Mullin

Received: 15 September 2009 / Accepted: 28 April 2010 / Published online: 19 February 2011  
© INRA and Springer Science+Business Media B.V. 2011

## Abstract

• **Introduction** Multi-varietal forestry (MVF) is the deployment of tested tree varieties in plantation forestry. Computer simulation using POPSIM Simulator identified optimal combination of numbers of families, varieties per family and ramets per variety (nf, nc and nr, respectively) yielding the largest genetic gain for a specific status number (NS) in a varietal test (VT) intended for MVF of conifers.

• **Results and discussion** Testing 40 to 80 full-sib families and 20 to 30 varieties per family would be optimal for a VT. This nf interval was insensitive to the number of candidate varieties planted, ratio of genetic variances and selection restriction. It was influenced somewhat by individual narrow-sense heritability ( $h^2$ ), required NS and mating design. Lower  $h^2$ , lower NS and designs with fewer

matings per parent tree favoured the lower range of the nf interval. The optimal nr was 6. It was not markedly affected by the required NS, family size or selection restriction but was strongly influenced by  $h^2$  and the ratio of genetic variances. Larger  $h^2$  or an introduction of non-additive genetic variance allowed planting fewer ramets per variety.

**Keywords** POPSIM Simulator · Genetic gain · Status number · Optimization

## 1 Introduction

In this paper, we define multi-varietal forestry (MVF) as the use of tested tree varieties in plantation forestry. We use the term ‘tree variety’ in a similar manner as it is used in horticulture where a ‘horticultural variety’ is an individual selected for desirable characteristics (Columbia Electronic Encyclopaedia 2003). Owing to recent refinement in somatic embryogenesis (SE) and cryopreservation techniques for conifers, the same varieties (i.e. genotypes or clones) are propagated consistently over time and commercially deployed to plantations. The main advantage of MVF is its ability to increase productivity through utilization of all available genetic variance (Park 2002). This benefit can be fully realized only through accurate varietal selection based on varietal tests (VT). Thus, VT is critical to MVF success. The term ‘VT’ is purposely used in this paper to discern it from a clonally replicated genetic test. VT focuses on varietal selection for deployment, while clonally replicated genetic tests are used primarily for genetic parameter estimation and breeding population selection.

Before establishing a genetic test, tree breeders must decide how many families and progenies per family need to be planted. Earlier relevant studies focused on maximizing

---

Handling Editor: Christophe Plomion

Y. Weng (✉)

New Brunswick Department of Natural Resources,  
Kingsclear Forest Nursery,  
E3E 1G3, Island View, NB, Canada  
e-mail: Yuhui.Weng@gnb.ca

Y.-S. Park

Natural Resources Canada,  
Canadian Forest Service-Canadian Wood Fibre Centre,  
P.O. Box 4000, E3B 5P7, Fredericton, NB, Canada

M. J. Krasowski

Faculty of Forestry and Environmental Management,  
University of New Brunswick,  
28 Dineen Drive, P.O. Box 4400, E3B 5A3, Fredericton, NB,  
Canada

T. J. Mullin

BioSylve Forest Science NZ Limited,  
45 Korokoro Road,  
Lower Hutt 5012, New Zealand

gain through accurate evaluation of family means (Cotterill and James 1984; Cotterill 1990) or through balancing between the choice (the more families are tested, the greater the choice) and accuracy of each family evaluation (Robertson 1957). Optimal numbers of progenies per family have been reported for a fixed number of families in genetic tests (Lambeth et al. 1983; Cotterill and James 1984). Likewise, optimal family numbers for forming breeding populations have been calculated (Lindgren et al. 1997; Danusevicius and Lindgren 2005). These studies provided bases for allocating testing efforts in genetic tests with seedlings aimed at an advancement of breeding populations.

Genetic testing for varietal selection results in another level of tradeoffs, i.e. how many ramets per variety must be raised. Efforts made toward this end concentrated on maximizing gain in varietal selection for deployment (Russell and Libby 1986; Russell and Loo-Dinkins 1993; Isik et al. 2005). These efforts identified two to six ramets per variety as being optimal, the actual number depending on broad-sense heritability, selection intensity and family size.

The studies cited above aimed mainly at maximizing gain. A selection and deployment strategy which maximizes gain without considering its effects on genetic diversity is not acceptable at present. The current goal for MVF is to guarantee that the genetic diversity of a varietal mixture remains above a threshold of acceptability, and under this constraint, the gain is maximized (Libby 1982; Lindgren 1993). In search of an optimal allocation of testing efforts for a VT, genetic diversity must be kept above that threshold, while numbers of families, varieties per family and ramets per variety are allowed to vary.

Comparisons of testing efforts regarding optimizing gain at a given genetic diversity are complex, but simulation techniques make such comparisons relatively straightforward (King and Johnson 1993; Mullin and Park 1995; McKeand and Bridgewater 1998). Simulation models may be classified into ‘deterministic’ and ‘stochastic’. One of the latter, the POPSIM, is especially suitable for simulating genetic systems (Mullin and Park 1995). Using the POPSIM simulator, the objectives of this study were to (1) determine optimal distribution of testing efforts in a VT and (2) evaluate the sensitivity of involved factors to changes in allocation of testing efforts.

## 2 Methods

### 2.1 Simulation model

Simulation was done using the stochastic quantitative genetic infinitesimal breeding simulator POPSIM. Theoretical basis and assumptions for POPSIM were described in

detail elsewhere (Mullin and Park 1995; Rosvall et al. 1998). In summary, the additive breeding value ( $A$ ) for a parent tree was sampled from a normal distribution  $N(\mu, V_A)$ , where  $\mu$  is the population mean and  $V_A$  is the variance due to additive effects. The dominance effects ( $D$ ) for full sibs were sampled from a normal family distribution with variance equal to  $N(0, 0.75V_D)$ ; the within-family portion of dominance variance) and with a family mean dominance effect sampled from  $N(0, 0.25V_D)$ , where  $V_D$  is the variance due to dominance effects. The epistatic effects ( $I$ ) and environmental effects ( $E$ ) were sampled from  $N(0, V_I)$  and  $N(0, V_E)$ , respectively, where  $V_I$  and  $V_E$  are epistatic and environmental variances, respectively. The phenotypic value of an individual was simulated as a sum of its  $A$ ,  $D$ ,  $I$  and  $E$ .

### 2.2 The reference MVF plan and alternatives

The following symbols are used:  $n_f$ , number of families;  $n_c$ , number of varieties per family;  $n_r$ , number of ramets per variety;  $N_C$ , total number of tested candidate varieties ( $n_f \times n_c$ );  $N_F$ , family size, total trees per family planted ( $n_c \times n_r$ );  $N_S$ , census number of varieties so that its diversity reaches a specified level; and  $N$ , total number of trees planted.

A complete MVF plan involves mating, testing and selection (Table 1). In the reference plan, 40 unrelated trees were selected as parent trees and subjected to a double-pair mating (DPM) scheme (each individual parent tree being mated twice), producing 40 full-sib families. Parent pairs were formed randomly, and selfing was excluded. For each family, 30 individuals were designed as candidate varieties, and for each variety, five ramets were vegetatively propagated to establish a VT. Selection was applied to a polygenic trait or an index of several traits and was based on the varietal genetic value, which was calculated on the basis of a combined index, where the performance of an individual and its relatives was weighted by their respective heritabilities (Mullin and Park 1995). Selection was made under a constraint of diversity level expressed as status number ( $N_S$ ), which is defined as half of the inverse of group co-ancestry (Lindgren et al. 1996). Incidentally, the status number used here is identical to the founder genome equivalents (Lacy 1995) in conservation biology. A  $N_S=15$  was designed as the diversity threshold. No restriction on the full-sib family contributions (each full-sib family could provide unlimited number of varieties) to the varietal mixture was applied in the selections.

Alternative options were set to test their influence on testing effort distribution, which is the combination of  $n_f$ , of  $n_c$  and of  $n_r$ . The alternatives included various (1) parent tree numbers, (2)  $N_S$  thresholds, (3) mating designs (single-pair mating (SPM) and quadric-pair mating (QPM), representing each individual parent tree being mated once

**Table 1** The reference and alternative parameters for the simulation of breeding, testing and selection for implementing multi-varietal forestry (MVF)

Parameters	Reference	Alternative options
<b>Breeding</b>		
No. of parent trees	40	20, 60, 80, 120, 200, 240
Allocation of mates	Random	
Mating frequency per parent tree	Double-pair	Single-pair, quadric-pair
<b>Testing</b>		
Number of families ( $n_f$ )	40	20, 60, 80, 120, 200
Number of varieties per family ( $n_c$ )	30	6, 10, 15, 20, 60
Number of ramets per variety ( $n_r$ )	5	2, 4, 6, 8, 10, 12, 20
Number of varieties tested ( $N_C$ )	1,200	300, 600, 1,800, 2,400
Number of trees per family ( $N_F$ )	150	60, 120, 180, 240
Total no. of tested trees ( $N$ )	6,000	3,600–13,200
<b>Selection</b>		
Status number ( $N_S$ ) threshold	15	10, 20
Max no. of varieties per full-sib family	Unlimited	2, 4
Other restrictions on relatedness	None	

and four times, respectively), (4) selection restrictions on family contributions to varietal mixtures and (5)  $N_C$  or  $N_F$ . It was assumed that there was no biological limit to creating families and producing candidate varieties or ramets.

### 2.3 Parameter input values

Table 2 lists the input values of the population parameters. In all simulations, the population mean (100),  $V_A$  (80) and standard rate of inbreeding depression (0) were kept constant, while  $V_D$ ,  $V_I$  and  $V_E$  were varied to represent an assortment of narrow-sense individual heritabilities ( $h^2$ ) and ratios among  $V_A$ ,  $V_D$  and  $V_I$ . The  $h^2$  (ranging from 0.05 to 0.5) was varied by changing the  $V_E$  while keeping the genetic variances constant. The  $V_A:V_D:V_I$  ratio was varied by changing  $V_D$  or  $V_I$  while keeping  $h^2$  constant through changes to its associated  $V_E$ . The  $V_D$  or  $V_I$  was assumed to be in the range between zero and  $V_A$ . The genetic parameters,  $V_A:V_D:V_I=80:2:18$  and  $h^2=0.2$ , which were observed in a clonally replicated test series of white spruce (*Picea glauca* (Moench) Voss.; Weng 2009), were used as the baseline parameter scenario.

**Table 2** Input parameter values for the baseline and alternative scenarios

Parameter	Fixed/variable	Baseline	Alternative values
Population mean value of trait	Fixed	100	
Rate of inbreeding depression	Fixed	0	
Additive variance ( $V_A$ )	Fixed	80	
Dominance variance ( $V_D$ )	Varied	2	0, 40, 80
Epistatic variance ( $V_I$ )	Varied	18	0, 40, 80
Environmental variance ( $V_E$ )	Varied	300	60, 1,500
Narrow-sense heritability ( $h^2$ )	Varied	0.20	0.05, 0.50

### 2.4 Optimization

In simulations, tradeoffs in gain at a specified  $N_S$  between  $n_f$  and  $n_c$  or between  $n_c$  and  $n_r$  were investigated under the same total testing effort, expressed as the same  $N$ . Under a model of fixed  $N$ ,  $N_C$  is also fixed in the optimization between  $n_f$  and  $n_c$ , or  $N_F$  is fixed if the objective is to optimize  $n_c$  and  $n_r$ . In the optimization between  $n_f$  and  $n_c$ , we varied  $n_f$  and  $n_c$  at a given  $n_r$ , whereas in the optimization between  $n_c$  and  $n_r$ , we varied  $n_c$  and  $n_r$  at a given  $n_f$ . Such methods of optimizing the distribution of testing efforts under a fixed  $N$  were used by other investigators (Shaw and Hood 1985; Russell and Libby 1986; Rosvall et al. 1998). The optimal  $n_c$  was also investigated by varying  $n_c$  while keeping  $n_f$  and  $n_r$  constant. This model represented different total testing efforts and was used in other studies (Cotterill and James 1984; Russell and Libby 1986).

Simulations were made by modifying one factor in the baseline parameter scenario or the reference plan at a time and, as much as possible, to keep the level of an effort constant. In each case, the interpretation was based on

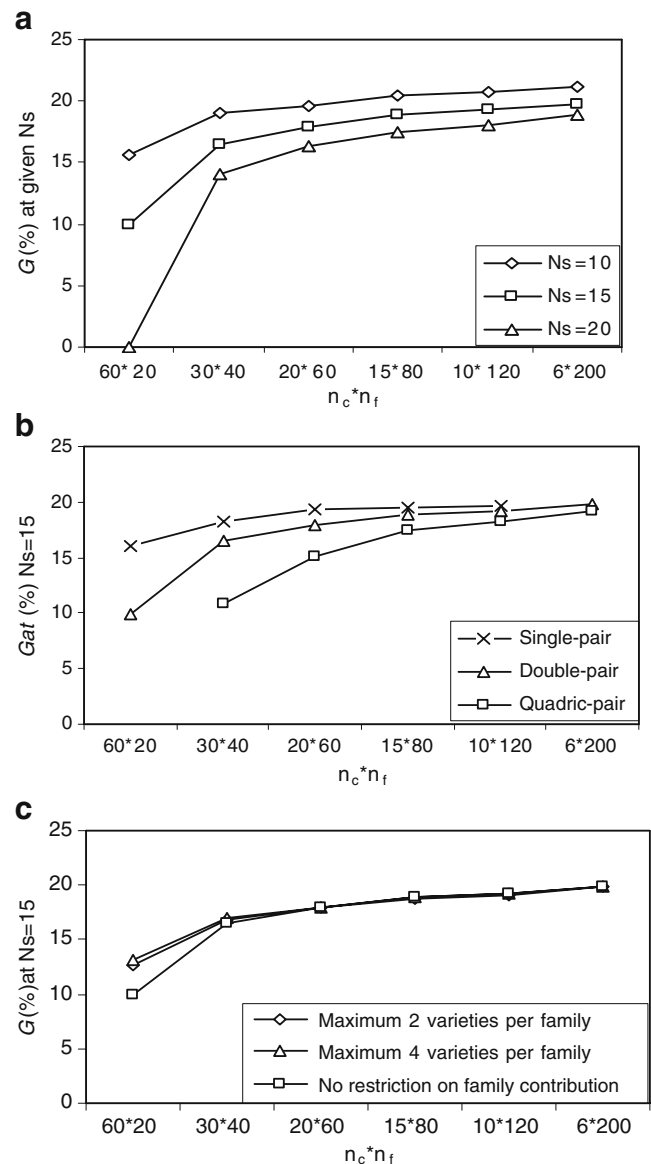
average result of 100 runs. The simulations were done in two consecutive steps. The first step was to find a  $N_V$  for a varietal mixture so that a specified  $N_S$  could be achieved for a combination of testing efforts. This was done using a linear approximation (Li and Lindgren 2006). The  $N_V$  found in step 1 was then used as the input for that specific combination so that gains obtained from various effort combinations could be compared at the same  $N_S$  level. The optimal allocation of testing efforts for a set of conditions is defined as that combination among  $n_f$ ,  $n_c$  and  $n_r$  which yields the largest gain at a specified  $N_S$  or, failing that, as the combination after which further numerical increases in  $n_f$ ,  $n_c$  and  $n_r$  increase gain at a specified  $N_S$  negligibly.

### 3 Results

#### 3.1 Tradeoffs between numbers of families ( $n_f$ ) and of varieties per family ( $n_c$ )

Under the baseline parameter scenario and the reference plan but re-distributing  $n_c$  and  $n_f$ , gain from selecting 20 top varieties increased quickly with  $n_c \times n_f$  combinations of  $60 \times 20$  to  $30 \times 40$ , but after that, it stabilized. On the other hand, there were substantial differences among the combinations with respect to  $N_S$ ; larger  $n_f$  paired with smaller  $n_c$  made  $N_S$  larger. The functional differences in gain and  $N_S$  among the combinations reflected relative efficiency of gain at a specified  $N_S$  (Fig. 1a). Overall, testing more families with fewer varieties per family resulted in higher gain at a specified  $N_S$ . However, the increase in gain was not linear; it increased quickly with  $n_f=20$  to 40; the increase slowed down when  $n_f=40$  to 80 and gain stabilized thereafter. For example, over 95% of gain at  $N_S=15$  obtained from the  $6 \times 200$  combination could be realized by the  $15 \times 80$  combination. Changing the  $N_S$  threshold did not change the optimal  $n_f$  interval but did change the relative efficiency in gain between the combinations (Fig. 1a). For example, the  $30 \times 40$  combination captured 90% of the gain obtained in the  $6 \times 200$  combination when  $N_S=10$  but reduced it to 74% when  $N_S=20$ . Thus, testing 40 to 80 families with corresponding  $n_c$  ranging from 15 to 30 could be an optimal interval for a VT of 1,200 candidate varieties. The lower bound of the optimal  $n_f$  interval was preferred when lower  $N_S$  was required.

Operationally, a given  $n_f$  can be created by various mating designs when required parent trees are available. For a given  $n_f$ , gain at a specified  $N_S$  was higher when the  $n_f$  was created by designs in which parent trees mated fewer times, although this effect gradually vanished for combinations of larger  $n_f$  paired with smaller  $n_c$  (Fig. 1b). In the  $n_c \times n_f$  combination of  $30 \times 40$ , gain at  $N_S=15$  was 18.2% when the 40 families were created by SPM, compared to 16.5%



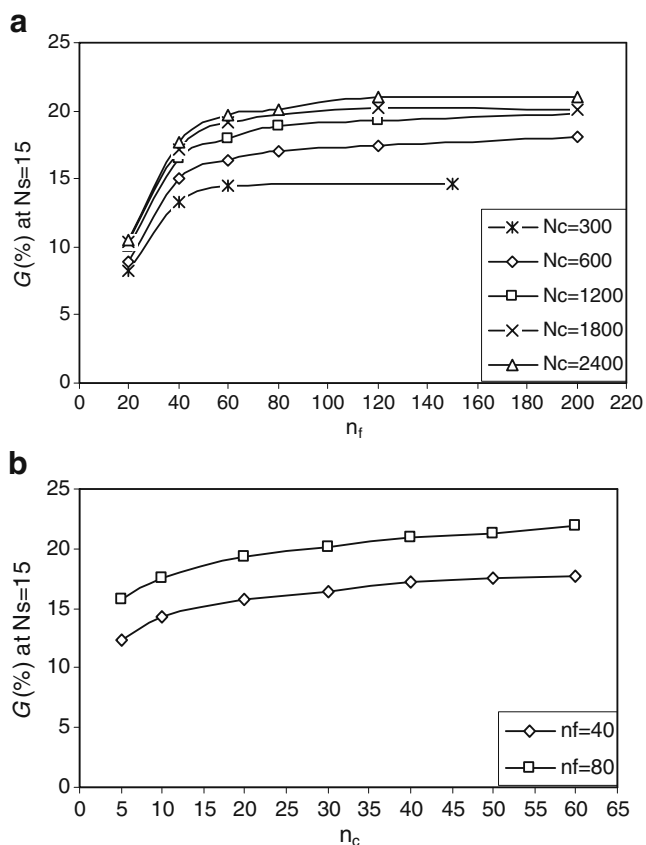
**Fig. 1** Tradeoffs in gain at a specific status number ( $N_S$ ) between numbers of varieties per family ( $n_c$ ) and of families ( $n_f$ ) under various  $N_S$  levels (the  $60 \times 20$  combination selected all varieties to reach a  $N_S=20$ , thus resulting in zero  $G$ ), (b) mating designs (the  $N_S$  failed to reach 15 in the  $60 \times 20$  combination by quadric-pair mating) and (c) restrictions on family contributions

and 10.9% when they were created by DPM and QPM, respectively. The optimal  $n_f$  (40 to 80) was similar among the three mating designs, although the designs with more matings per parent tree favoured the upper bound of  $n_f$  interval. Imposing restriction on selection differently affected gain and  $N_S$ ; a selection allowing fewer varieties per family resulted in lower gain but higher  $N_S$  (data not shown). As shown in Fig. 1c, restriction on family contributions, however, did not change the optimal  $n_f$  interval. Gains at  $N_S=15$  were similar among the three restriction levels for all  $n_c \times n_f$  combinations except that of

60×20, where the unrestricted selection had a lower gain than the other options. It can be seen that imposing family restrictions was effective when  $n_f$  was small (i.e. 20).

Increasing  $N_C$  ( $n_f \times n_c$ ) in a VT will increase gain and  $N_S$ , thus increasing gain at a specified  $N_S$ . When  $N_C$  was increased from 300 to 2,400, the optimal  $n_f$  interval was fairly stable, but the corresponding  $n_c$  increased considerably (Fig. 2a). Although gain at a specified  $N_S$  increased with  $N_C$ , the additional gain from planting more than 1,200 candidate varieties at a given  $n_f$  would have been small. These results suggested that planting more than 30 varieties per family in a VT under the baseline parameter scenario was not worth the effort. This optimal  $n_c$  was further confirmed by a model with an increasing  $n_c$  at a given  $n_f$  (Fig. 2b). The additional gain at a specified  $N_S$  from testing more than 30 varieties per family was small. For example, at  $n_f=40$  or 80, over 92% of gain obtained at  $N_S=15$  from testing 60 varieties per family would be captured by testing only 30 varieties per family, and this would decrease the total testing effort by 50%.

As expected, gain at a specified  $N_S$  increased with larger  $h^2$  or with the introduction of non-additive genetic variance

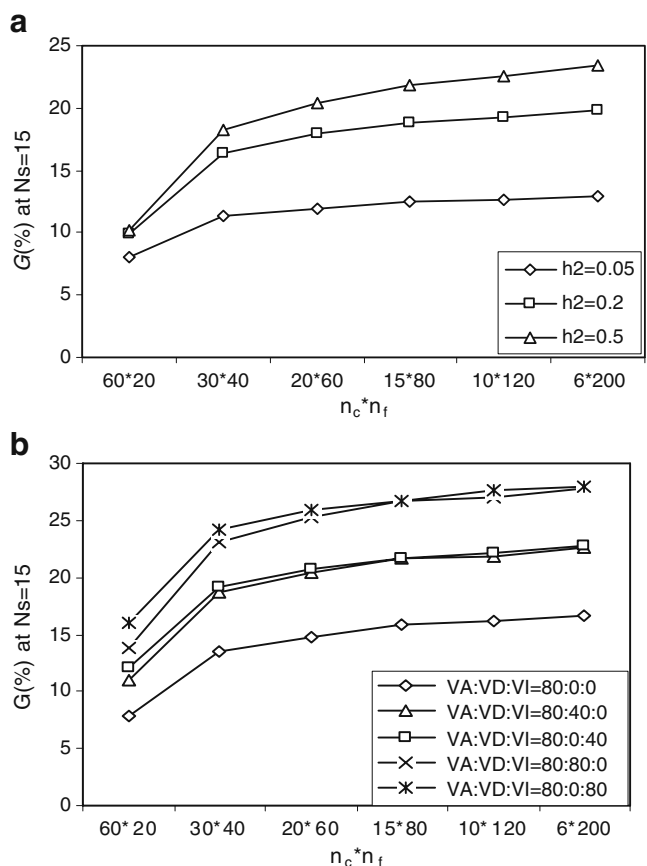


**Fig. 2** Tradeoffs in gain at a status number ( $N_S$ ) of 15 between numbers of varieties per family ( $n_c$ ) and of families ( $n_f$ ) under various candidate varietal numbers ( $N_C=n_c \times n_f$ ) by (a) varying  $n_f$  and  $n_c$  combinations and (b) increasing  $n_c$  at fixed  $n_f$

(Fig. 3). The optimal  $n_c \times n_f$  combination was somewhat influenced by  $h^2$ ; the smaller  $h^2$  preferred the lower optimal  $n_f$  bound (Fig. 3a). For example, the 30×40 combination could capture 88% of gain at  $N_S=15$  obtained in the 6×200 combination when  $h^2=0.05$ , but it would decline to 77% when  $h^2=0.5$ . Unlike with  $h^2$ , the introduction of  $V_D$  or  $V_I$  did not significantly change the optimal combination of  $n_c$  with  $n_f$  (Fig. 3b). An introduction of  $V_I$  resulted in a slightly higher gain at  $N_S=15$  than when an equal amount of  $V_D$  was introduced, although this advantage gradually diminished with increasing  $n_f$  (Fig. 3b).

3.2 Tradeoffs between varietal numbers per family ( $n_c$ ) and ramet numbers per clone ( $n_r$ )

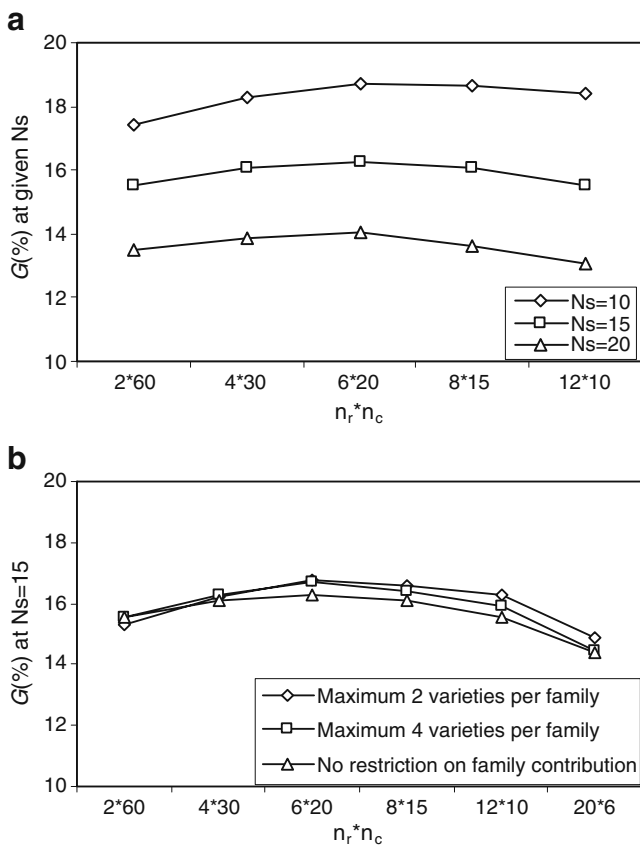
While adhering to the baseline parameter scenario and the reference plan but redistributing  $n_r$  and  $n_c$  for a VT of  $N_F=120$ , gain from selecting 20 varieties maximized at  $n_r=6$ , while  $N_S$  values were fairly stable, around 11 (data not shown). Consequently, the optimal  $n_r$  for maximizing gain



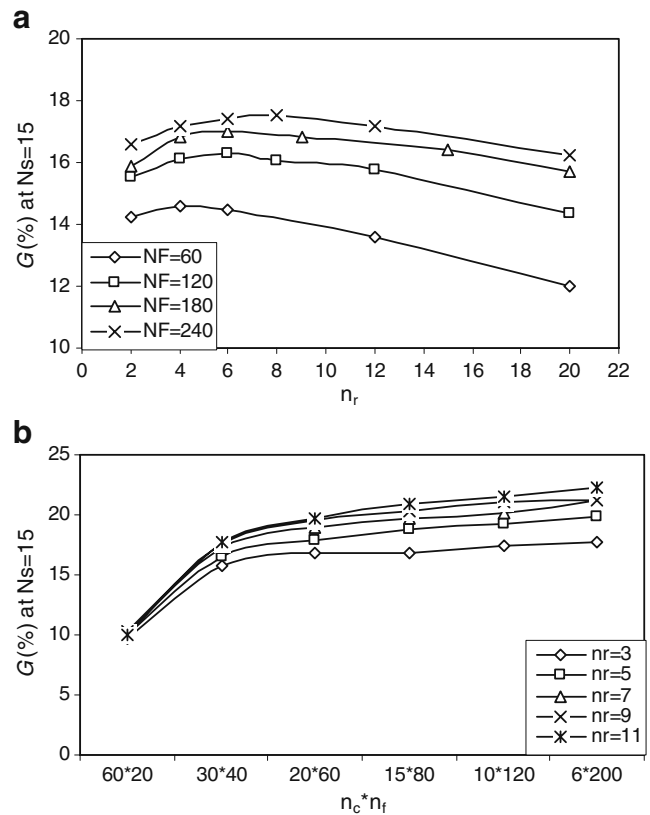
**Fig. 3** Tradeoffs in gain at a status number ( $N_S$ ) of 15 between numbers of varieties per family ( $n_c$ ) and of families ( $n_f$ ) under various genetic parameters: (a) individual heritability ( $h^2$ ) and (b) ratio of genetic variances. Note that in b, the  $h^2$  was kept constant at 0.2; thus, the respective broad-sense heritabilities were 0.20 (80:0:0), 0.30 (80:40:0 and 80:0:40) and 0.4 (80:80:0 and 80:0:80)

at a given  $N_S$  was around 6, and the corresponding  $n_c$  was 20. The optimal  $n_r$  and  $n_c$  were not changed by the assigned  $N_S$ , although there was a tendency to require planting of more ramets with a smaller required  $N_S$  (Fig. 4a). Imposing selection restriction on family contributions changed gain and  $N_S$  but did not markedly change the optimal allocation between  $n_r$  and  $n_c$  (Fig. 4b).

Tradeoffs between  $n_r$  and  $n_c$  were also examined for different  $N_F$  (Fig. 5a). Under the baseline parameter scenario, the optimal  $n_r$  was 4 when  $N_F=60$  and increased to 6 when  $N_F=120-180$  and to around 8 when  $N_F=240$ , with the corresponding  $n_c$  increasing from 15 to 30. For a given  $n_r$ , increasing  $N_F$  (through increasing  $n_c$ ) increased gain at a specified  $N_S$ . However, the additional gain at a specified  $N_S$  increased at a diminishing rate as  $N_F$  increased. This was demonstrated by the curves in Fig. 5a getting progressively closer together as  $N_F$  was rising. It can be seen that the interval of  $N_F$  between 120 and 180 would be a reasonable choice. Taking all factors into consideration, testing 120 to 180 trees per family with six ramets per variety seemed to be appropriate and corresponds to an



**Fig. 4** Effects of redistribution between numbers of varieties per family ( $n_c$ ) and of ramets per variety ( $n_r$ ) on genetic gain ( $G$ ) at a specific status number ( $N_S$ ) level under various (a)  $N_S$  levels and (b) selection restrictions on family contributions



**Fig. 5** Effects of numbers of ramets per variety ( $n_r$ ) on genetic gain ( $G$ ) at a status number ( $N_S$ ) of 15 under various tree numbers per family ( $N_F$ ) by varying (a) numbers of varieties per family ( $n_c$ ) and  $n_r$  and (b) numbers of families ( $n_f$ ),  $n_c$  and  $n_r$ . Note that  $N_F=n_c \times n_r$

optimal interval of  $n_c$  between 20 and 30. The optimal  $n_r$  interval was further investigated under various  $n_c \times n_f$  combinations (Fig. 5b). For each  $n_c \times n_f$ , gain at  $N_S=15$  increased with increasing  $n_r$ , but this increase was much reduced with larger  $n_r$ . In general, combinations of larger  $n_c$  with fewer  $n_f$  required planting of fewer  $n_r$ . In the combination of  $60 \times 20$ , the additional gain at  $N_S=15$  from testing more than three ramets per variety was negligible. The corresponding  $n_r$  increased to 5 for the combinations of  $30 \times 40$  and to 7 thereafter. Overall, these results suggest that it is unnecessary to plant more than seven ramets per variety.

As expected,  $h^2$  had significant effects on optimal distribution of  $n_c$  versus  $n_r$ ; the larger  $h^2$  favoured combinations of larger  $n_c$  with smaller  $n_r$  (Fig. 6a). Gain at  $N_S=15$  maximized at the  $n_r \times n_c$  combination of  $2 \times 60$  when  $h^2=0.5$  but of  $20 \times 6$  when  $h^2=0.05$ . Introduction of non-additive genetic variances also affected the optimal distribution between  $n_c$  and  $n_r$ . Overall, small ratios of additive to non-additive variance favoured planting fewer ramets per variety and yielded more gain at a specified  $N_S$  than cases where the ratio was large, in particular when the  $V_I$  was the main source of non-additive variance (Fig. 6b).

For example, gain at  $N_S=15$  was maximized at the  $n_r \times n_c$  combination of  $2 \times 60$  when  $V_A:V_D:V_I=80:80:0$  or  $80:0:80$ .

## 4 Discussion

### 4.1 Optimum testing effort allocation

Simulations were done not so much to estimate genetic gains at specific diversity levels but to compare efficiency among various effort allocations. The baseline parameter scenario was specific for white spruce, but alternative options (Table 2) were deliberately selected to cover other conifers. Extensive studies on conifers have demonstrated that additive effects are the main contributor to genetic variance, whereas non-additive genetic variance is unlikely to contribute more than half of the total genetic variance (see Weng (2009) and references therein). A partitioning pattern of  $V_A:V_D:V_I=80:20:0$  has been usually assumed for coniferous species and used in other simulations (Russell

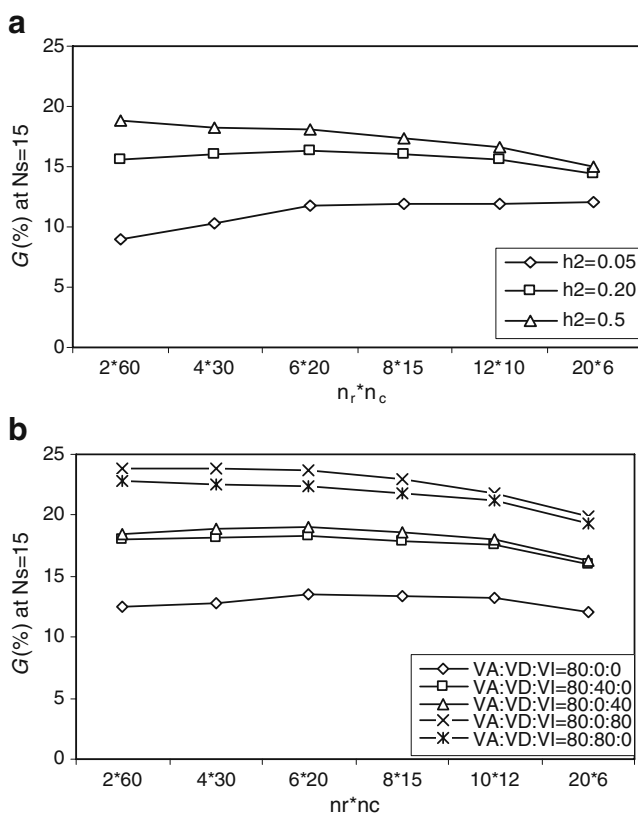
and Loo-Dinkins 1993; Rosvall et al. 1998). The optimal allocation of VT efforts under the ‘80:20:0’ scenario (data not shown) was virtually the same to the recommendation found in this study.

The minimum varietal number for a MVF plantation is debatable, but theoretical studies generally suggest that planting more than ten unrelated varieties in a plantation should suffice for protection against unpredicted biotic and environmental stresses (Lindgren 1993; Park et al. 1998). Biologically,  $N_S$  can be interpreted as the size of a non-inbred population in which all selections are unrelated (White et al. 2007). Thus, a  $N_S$  ranging from 10 to 20 for a varietal mixture corresponds well to the above theoretical recommendation. Further increases in  $N_S$  may be preferred for evolutionary purposes but may contribute little to productivity (Lindgren 1993).

Thus, results of this study provide insights into optimum effort allocation for establishing a VT at reasonable diversity levels for conifers. If families can be created at low cost, it will be wise to include more families. For most conifers, producing fewer crosses is always attractive operationally because it increases the likelihood that mating work can be completed within a reasonable time. Our study suggests that a VT should include 40 to 80 families, 20 to 30 varieties per family and about six ramets per variety in field environments. Tests made under more uniform environments (e.g. in a greenhouse) may need fewer ramets per variety than tests made under more variable conditions. The recommended  $n_f$  interval is a prudent compromise between gain and feasibility and should be treated as an optimal approximation.

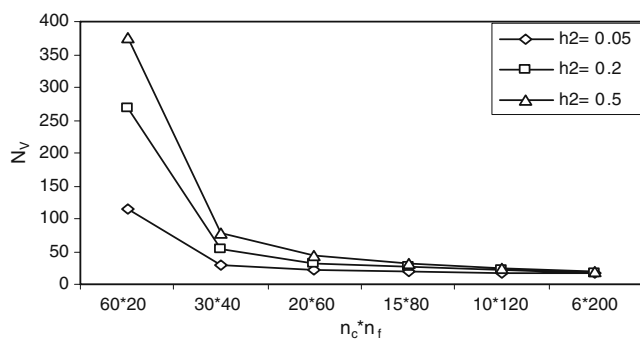
Other than maximizing gain at a specific  $N_S$ , an optimization should consider the operational feasibility of cloning genotypes. SE is one of the key techniques which make the implementation of MVF possible. Operationally, raising trees from fewer varieties is always desirable, and this should be another criterion in optimizing effort allocation. To obtain a specific  $N_S$ , combinations of fewer  $n_f$  paired with more  $n_c$  required including more varieties in the mixture (larger  $N_V$ ). Under the baseline parameter scenario, a varietal mixture of  $N_S=15$  needed to include as many as 250 varieties ( $N_V$ ) under the  $n_c \times n_f$  combination of  $60 \times 20$ . The corresponding  $N_V$  would decrease to about 50 under the  $30 \times 40$  combination and to 30 or less when 60 or more families were involved (Fig. 7). In this regard, a  $n_f$  of 40 or larger seems indispensable, which is consistent with the optimum  $n_f$  required for maximizing gain at a specific  $N_S$ . Although the number of varieties actually required for a specific combination of  $n_f$  and  $n_c$  may vary with heritability (Fig. 7), mating design and selection restriction (data not shown), the conclusion found here seems to hold.

Field testing is often constrained by limitations in resources and logistics. The optimal distribution effort



**Fig. 6** Effects of redistribution between numbers of varieties per family ( $n_c$ ) and of ramets per variety ( $n_r$ ) on genetic gain ( $G$ ) at a status number ( $N_S$ ) of 15 under various genetic parameters: (a) individual heritability ( $h^2$ ) and (b) ratios of genetic variances. Note that in b, the  $h^2$  was kept constant at 0.2; thus, the respective broad-sense heritabilities were 0.20 (80:0:0), 0.30 (80:40:0 and 80:0:40) and 0.4 (80:80:0 and 80:0:80)





**Fig. 7** Effects of redistribution between numbers of families ( $n_f$ ) and of varieties per family ( $n_c$ ) on the required number of varieties ( $N_v$ ) in mixtures so that their status numbers ( $N_S$ ) reach 15 under various individual heritability ( $h^2$ )

(40–80 families and 20–30 varieties per family with six ramets) results in a testing size of 4,800 to 14,400 trees, which is much larger than the typical genetic testing size of 2,500 trees (i.e. 1 ha planted at 2×2 m spacing). One solution to this situation is to establish testing series over time, planting about 20 families (2,400 to 3,600 trees) each year. In such a way, each series can stand alone, and reasonable gain can still be obtained by selection from it (Table 3). Whenever data are available for all testing series, they should be analyzed together and selections be made across the series for maximizing gain at a given  $N_S$ .

4.2 Effects of other factors on the effort distribution

Other factors investigated in this study represent those that are under breeder's control or those that can be determined in advance from earlier studies. Generally, the optimal  $n_f$  interval (40 to 80) is insensitive to the tested  $N_C$ , introduction of non-additive variance and selection restriction on family contributions. Consideration of the following factors may require the lower bound of the optimal  $n_f$  when (1) a lower  $N_S$  threshold is required, (2)  $h^2$  is lower or (3) a mating design is used in which each parent tree is mated fewer times. In these cases, effects of relatedness among the varieties (or constraint of  $N_S$ ) on gain were smaller, resulting in an optimal  $n_f$  closer to that (40) with no

**Table 3** Genetic gain (percent) and status number ( $N_S$ ) of variety mixtures consisting of 20–40 selected varieties ( $N_v$ ) from a varietal test of 20 families ( $n_f$ ) paired with 20–30 varieties per family ( $n_c$ ) and six ramets per variety according to the baseline parameter scenario

$n_c \times n_f$	$N_v=20$		$N_v=30$		$N_v=40$	
	Gain	$N_S$	Gain	$N_S$	Gain	$N_S$
20×20	16.8	10.0	15.4	12.7	14.7	14.4
30×20	17.5	9.5	16.3	11.8	15.3	13.4

constraint on  $N_S$ . At lower  $h^2$ , more varieties per family are needed so that the precision of family evaluation can be improved (Lindgren et al. 1997).

Overall comparison among the three mating designs suggests that SPM would be preferred due to its simplicity, smaller optimal  $n_f$  and higher gain at a given  $N_S$  (Fig. 1b), although it requires more available parent trees. The parent trees should be selected from the ongoing long-term tree improvement programs. More complicated mating designs, such as complete diallel, were not evaluated in this study because they are not likely to be used in future MVF programs.

Under fixed  $N_F$ , the tradeoff between  $n_c$  and  $n_r$  reflects the balance of selection accuracy and intensity. Increasing  $n_r$  may increase the precision of varietal selection but may limit  $N_C$  that can be tested, thus decreasing selection intensity. The balance between  $n_c$  and  $n_r$  is not likely to be greatly changed by the required  $N_S$  level,  $N_F$  and selection restriction, at least not under the ranges tested, but it may vary markedly with  $h^2$  and an introduction of non-additive variance. Lower  $h^2$  means that larger  $n_r$  is required from a smaller  $n_c$  (Fig. 6a; White et al. 2007). The introduction of non-additive variances forced a lower  $V_E$  under a constant  $h^2$ , resulting in a higher broad-sense heritability (Fig. 6b). Varietal genetic values can be accurately predicted even if few ramets per variety are planted in a VT with high broad-sense heritability (Fig. 6b; Shaw and Hood 1985; Russell and Libby 1986).

Since the optimal  $n_f$  and  $n_r$  intervals are fairly insensitive to the factors investigated in this study (except effects of  $h^2$  and ratio of genetic variances on  $n_r$ ), the optimal  $n_c$  is virtually a function of  $N_C$  or  $N_F$ . Given there is no limit on available funds and land, testing larger  $N_C$  or  $N_F$  by increasing  $n_c$  would be appropriate and recommended. The additional gain resulting from testing more than 1,200 candidate varieties or 180 trees per family, however, was small (Figs. 2a and 5a), not encouraging such an effort economically. The optimal  $n_c$  can be much larger than those when  $h^2$  is higher or when more non-additive genetic variance is introduced (Fig. 6).

4.3 Comparison to other studies

In MVF, we face two critical issues that must be considered simultaneously: (a) developing varietal lines that maximize genetic gain through best utilization of the available number of genetic entries (families and varieties within them) and (b) maintaining the acceptable level of genetic diversity in MVF plantations. Our study addresses these two issues by focussing on optimization of testing efforts in VT that maximize gain under a constraint on  $N_S$ . The importance of sufficient  $N_S$  in MVF has been emphasized in the literature (Roberds and Bishir 1997), and some

jurisdictions defined the minimum for that number (e.g. Alberta Sustainable Resource Development 2003; British Columbia Ministry of Forests and Range 2004).

Several investigators have studied numbers of families and progeny per family required in genetic tests. Robertson (1957) recommended testing as few as five individuals per family so greater efforts could be directed to testing more families. Robertson's solution is a tradeoff between the availability of families for selection and accuracy of comparing family means. Cotterill (1990) recommended planting five to ten families per provenance and 20 to 40 trees per family if the objective was to evaluate accurately provenance means. In a family test, planting 10–20 progeny per family would suffice if the objective were to accurately calculate the family mean (Cotterill and James 1984). However, unlike our study, these studies did not consider genetic diversity and focused more on the accuracy of estimating means than on optimizing genetic gain.

For establishment of breeding populations, optimum number of families was investigated considering the genetic gain, diversity and cost efficiency (Lindgren et al. 1997; Danusevicius and Lindgren 2005). In their models, however, the genetic diversity was treated as a penalty factor reducing gain. The objectives of these studies were relevant to managing breeding populations; therefore, they were different from our objectives aiming at maximizing genetic gain at a defined level of diversity in MVF plantations.

Most early VTs were established following breeder's intuition, i.e. planting many clones from few elite families (White et al. 2007) or creating families by complicated mating designs, to maximize genetic gain. Varietal selection from such tests might result in a high gain but increase genetic relatedness among selections (NBTIC 1999). As demonstrated in our study, such high gains may not be attainable due to loss of diversity when a  $N_S$  is imposed. Our study also shows that single-pair mating is better than the structured mating designs, for example, diallel mating, to reduce relatedness among families and varieties.

The optimal  $n_r$  for a test with clonal replicates has been well investigated both for breeding population (Shaw and Hood 1985; Russell and Loo-Dinkins 1993; Rosvall et al. 1998) and production population (Russell and Loo-Dinkins 1993; Isik et al. 2005), using a model of maximizing genetic gain but without considering genetic diversity. Russell and Libby (1986) presented a model based on selecting and deploying a fixed number of clones without considering relatedness among clones. Our model, which includes a constraint of  $N_S$  and genetic relationships, would be more flexible and closer to reality. The result of our study that planting around six ramets per variety was comparable to numbers recommended by all of the studies cited above, but some differences should be noted. Unlike

in Shaw and Hood's study (1985), our results showed that the restriction on family contribution did not affect the optimal  $n_r$ . Under similar model parameters, the optimal  $n_r$  was slightly larger than that reported by Russell and Libby (1986). They found that the optimal  $n_r$  was 2 or 3 when 10 or 20 unrelated clones were selected from a clonal test with its broad-sense heritability of 0.24.

There are other aspects of VT which are not addressed here. Optimizing effort allocation of a VT should consider  $G \times E$  interaction. Russell and Loo-Dinkins (1993) found that the extra gains from planting more than two sites were small when the  $G \times E$  interaction was less than 25% of the genetic variance. It should also be noted that the optimal effort allocation reported here targeted to candidacy testing, with many varieties and few ramets per variety. These selected varieties should be further tested across more landscape environments and by different experimental designs (i.e. large block of monoclonal or mixed clonal plots) to get more detailed productivity information on these varieties. The number of varieties at this level would be modest, and the number of ramets per variety would be large.

Another important but unaddressed aspect is the cost of optimization. In forest operations, the same  $N_C$  created by varying  $n_f$  and  $n_c$ , or the same  $N_F$  created by varying  $n_c$  and  $n_r$  may involve different expenses. As suggested by Lindgren (1985), a cost–benefit analysis is needed to fully determine optimal distribution among  $n_f$ ,  $n_c$  and  $n_r$  for a genetic test with clonal replicates. The optimal  $n_r$  for a VT does not seem to be affected by including a cost analysis (Isik et al. 2004).

## References

- Alberta Sustainable Resource Development (2003) Standards for tree improvement in Alberta. Available at <http://www3.gov.ab.ca/srd/forests/fmd/manuals/index.html>. Accessed 18 May 2008
- British Columbia Ministry of Forests and Range (2004) Chief forester's standards for seed use. Available at <http://www.for.gov.bc.ca/code/cfstandards/html/>. Accessed 18 May 2008
- Columbia Electronic Encyclopaedia 6th edn. (2003) Available at <http://encyclopedia2.thefreedictionary.com>. Accessed 18 June 2009
- Cotterill PP (1990) Numbers of families and progenies required for provenance testing. *Silvae Genet* 39:82–83
- Cotterill PP, James JW (1984) Number of offspring and plot sizes required for progeny testing. *Silvae Genet* 35:203–209
- Danusevicius D, Lindgren D (2005) Optimization of breeding population size for long-term breeding. *Scand J For Res* 20:18–25
- Isik F, Li B, Frampton J, Goldfarb B (2004) Efficiency of seedlings and rooted cuttings for testing and selection in *Pinus taeda*. *For Sci* 50:44–53
- Isik F, Goldfarb B, LeBude A, Li B, McKeand S (2005) Predicted genetic gains and testing efficiency from two loblolly pine clonal trials. *Can J For Res* 35:1754–1766

- King JN, Johnson GR (1993) Monte Carlo simulation models of breeding-population advancement. *Silvae Genet* 42:68–78
- Lacy RC (1995) Clarification of genetic terms and their use in the management of captive populations. *Zoo Biol* 14:565–577
- Lambeth CC, Gladstone WT, Stonecypher RW (1983) Statistical efficiency of row and non-contiguous family plots in genetic tests of loblolly pine. *Silvae Genet* 32:24–28
- Li H, Lindgren D (2006) Comparison of phenotype and combined index selection at optimal breeding population size considering gain and gene diversity. *Silvae Genet* 55:13–19
- Libby WJ (1982) What is a safe number of clones per plantation? In: Heybroek HM, Stephan BR, von Weissenberg K (eds) Resistance to diseases and pests in forest trees. Center Agricultural Pub & Document, Wageningen, pp 342–360
- Lindgren D (1985) Cost-efficient number of test sites for field trials. *Biometrics* 41:887–893
- Lindgren D (1993) The population biology of clonal deployment. In: Ahuja MR, Libby WJ (eds) Clonal forestry I: genetics and biotechnology. Springer, Berlin, pp 34–49
- Lindgren D, Gea LD, Jefferson PA (1996) Loss of genetic diversity monitored by status number. *Silvae Genet* 45:52–59
- Lindgren D, Wei RP, Lee SJ (1997) How to calculate optimum family number when starting a breeding program. *For Sci* 43:206–212
- McKeand SE, Bridgwater FE (1998) A strategy for the third breeding cycle of loblolly pine in southern U.S. *Silvae Genet* 47:223–234
- Mullin TJ, Park YS (1995) Stochastic simulation of population management strategies for tree breeding: a new decision-support tool for personal computers. *Silvae Genet* 44:132–141
- New Brunswick Tree Improvement Council (NBTIC) (1999) Black spruce clonal testing series. New Brunswick Tree Improvement Council Report #57. Department of Natural Resources, Fredericton, New Brunswick, Canada
- Park YS (2002) Implementation of conifer somatic embryogenesis in clonal forestry: technical requirements and deployment considerations. *Ann For Sci* 59:651–656
- Park YS, Bonga JM, Mullin TJ (1998) Clonal forestry. In: Mandal AK, Gibson GL (eds) Forest genetics and tree breeding. CBS Publishers & Distributors, Daryaganj, pp 143–167
- Roberds JN, Bishir JW (1997) Risk analyses in clonal forestry. *Can J For Res* 27:425–435
- Robertson A (1957) Optimum group size in progeny testing and family selection. *Biometrics* 13:442–450
- Rosvall O, Lindgren D, Mullin TJ (1998) Sustainability robustness and efficiency of a multi-generation breeding strategy based on within-family clonal selection. *Silvae Genet* 47:307–321
- Russell JH, Libby WJ (1986) Clonal testing efficiency: the trade-offs between clones tested and ramets per clone. *Can J For Res* 16:925–932
- Russell JH, Loo-Dinkins JA (1993) Distribution of testing effort in cloned genetic tests. *Silvae Genet* 42:98–104
- Shaw DV, Hood JV (1985) Maximizing gain per effort by using clonal replicates in genetic tests. *Theor Appl Genet* 71:392–399
- Weng YH (2009) Developing “multi-varietal forestry” strategy for white spruce in New Brunswick. Ph.D. dissertation, University of New Brunswick, Fredericton, p 146
- White TL, Adams WT, Neale DB (2007) Forest genetics. CABI Publishing, Cambridge, p 682