Breeding without breeding

YOUSRY A. EL-KASSABY[1]* AND MILAN LSTIBŮREK[2]
[1] *Department of Forest Sciences, University of British Columbia, Vancouver, BC, Canada V6T 1Z4*
[2] *Faculty of Forestry and Wood Sciences, Czech University of Life Sciences, Prague, 165 21 Czech Republic*

Summary

An innovative approach to tree breeding called 'breeding without breeding' (BWB) is presented. The method, as applied on the material in hand, allows the capture of 75–85% of the genetic response to selection attained through conventional programmes without the need to do any controlled pollination and simplified or possibly no experimental field testing: both considered to be the most resource-demanding activities in breeding programmes. BWB combines the use of genotypic or phenotypic pre-selection of superior individuals, informative DNA markers for fingerprinting and pedigree reconstruction of offspring to assemble naturally created full- and half-sib families resulting from mating among selected parents, and quantitative genetics analyses to identify elite genotypes for further genetic improvement or the establishment of production populations. BWB utility is demonstrated using a retrospective study of Douglas-fir (*Pseudotsuga menziesii*) progeny tests consisting of offspring produced from 150 controlled crosses among 60 parents and established over three sites. The empirical results are supported by theoretical expectations demonstrating anticipated minimum genetic response compared with conventional approaches. The method's simplicity offers an exceptional opportunity for the development of comparable breeding efforts in developing countries, advanced and new breeding programmes, and economically important and 'minor' species.

1. Introduction

Forest gene resource management is a daunting task requiring balancing utilization and conservation goals (Yanchuk, 2001) for large, long-lived organisms covering vast ecological regions (Ying & Yanchuk, 2006). While the genetic improvement of economic traits is of primary importance, the maintenance of broad genetic diversity is essential to meeting changing breeding goals (Namkoong *et al.*, 1988) and guarding against unpredictable temporal and global environmental contingencies such as climate change (Eriksson *et al.*, 1993). Most tree improvement programmes follow recurrent selection schemes, consisting of multiple populations, including base populations where initial phenotypic selections are made, breeding populations where crosses among the selected individuals are created and tested and deployment populations (e.g. seed orchards) harbouring elite genotypes for seed production (Namkoong *et al.*, 1988). Breeding arboreta are commonly established to safeguard the genetic legacy of original and subsequent selections and to facilitate breeding. Breeding populations are managed to maintain higher levels of genetic variability for sustained long-term genetic response while capturing genetic gain is the sole function of production populations (Namkoong *et al.*, 1988).

Predetermined mating designs are considered essential for creating structured, pedigreed families for genetic testing, facilitating accurate assessment of genetic parameters (breeding values, heritabilities, genetic gain, genetic variances and covariances and genotype–environment interactions), and the selection of elite genotypes for advanced breeding and seed orchards establishment (Lambeth *et al.*, 2001). Forest

* Corresponding author. Yousry A. El-Kassaby, 3030-2424 Main Mall, Department of Forest Sciences, University of British Columbia, Vancouver, BC, Canada V6T 1Z4. Tel: (604) 822 1821; Fax: (604) 822 9102. e-mail: y.el-kassaby@ubc.ca

geneticists often use rather complex schemes, such as the disconnected partial diallel mating design (Namkoong, 1979) to meet these objectives; however, the large numbers of crosses create logistical difficulties requiring years to complete and often the parental authenticity of the resulting offspring is not error-free (Adams *et al.*, 1988; Devey *et al.*, 2002). Phenotypic selection was proposed by Cotterill (1986) and its efficiency was evaluated by Andersson *et al.* (1998) to simplify and reduce tree breeding efforts. However, while simple, phenotypic selection is expected to cause cryptic build-up of coancestry along the selection cycles; thus the incorporation of genetic markers and parentage analyses would be of great value for this scheme implementation.

The recent development of highly informative DNA markers and advanced pedigree reconstruction methods made it possible to determine the genealogical relationships among offspring produced from natural mating among individuals (see Jones & Ardren, 2003 for a review). Attempts to incorporate the use of DNA markers and parentage analysis into tree breeding were made by Lambeth *et al.* (2001) and Grattapaglia *et al.* (2004); however, these were proposed to meet specific scenarios such as the complementary polycross mating design and a special case of *Eucalyptus* inter-specific hybridization, respectively (see the Results and discussion section). In the present study, we extended the pedigree reconstruction approach to open-pollinated managed forest tree breeding populations for reconstructing mating relationships without making crosses and using the result to estimate genetic parameters for parental genetic evaluation and selection of superior individuals. Our approach allows the capture of a substantial amount of the genetic gain attained through conventional tree breeding. This is accomplished without any controlled pollination and simplified testing and the approach could be extended to eliminate the use of elaborate experimental field testing altogether. The method, called 'breeding without breeding' (BWB), combines the use of phenotypic or genotypic pre-selection of superior individuals to minimize fingerprinting efforts, pedigree reconstruction of offspring to assemble naturally created full- and half-sib families and quantitative genetics analyses to identify elite genotypes. We demonstrated our method using a retrospective study of Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) consisting of progenies produced from 150 controlled crosses among 60 parents and established on three sites. These findings were supported by theoretical expectations demonstrating expected minimum genetic gain compared with conventional approaches.

This is the first large-scale, operationally oriented study demonstrating the utility of partial- or full-pedigree reconstruction to assemble naturally created full- and half-sib families for forest trees breeding and extends the results to propose not only avoiding crosses, but also to simplify or even bypass traditional progeny testing.

2. Materials and methods

(i) *Plant materials* (*retrospective study*)

Eight-year height data from a subset of large-scale Douglas-fir progeny trials described in Yanchuk (1996) was used to illustrate BWB. In brief, the data consisted of ten 6-parent disconnected partial diallels planted in replications over three test sites, where each individual (a total of 7442) is identified by site, replication and its male and female parents.

(ii) *Simulation of BWB*

The original data set (both parents known: full-sib model (FS model)) is regarded as the benchmark for comparisons. To illustrate the utility of BWB, two strategies were modelled, the first (half-sib model (HS model)) where all male parents' information was ignored, thus resembling the classical half-sib progeny trial (Namkoong, 1966). While in the second (phenotypic-selection model (PH model)), all parents were considered unknown, representing completely unstructured field progeny trial. Modelling BWB's strategies preceded the following three steps:

Step 1 – Data analysis. Data analysis under all models (FS, HS and PH) utilized the mixed model fitted with the ASReml software (Gilmour *et al.*, 2006), featuring the restricted maximum likelihood to partitioning the phenotypic variance into causal components. Under the FS model, the evaluation followed the diallel mating scheme (Griffing, 1956) where the best linear unbiased prediction (BLUP; Henderson, 1975) utilizing the animal genetic evaluation model resulted in prediction of additive genetic (=breeding) values (BV_{FS}), optimally combining measurements on all individuals in the pedigree (Lynch & Walsh, 1998). Genetic evaluation under the HS model followed the same animal genetic model, while assuming simple half-sib genetic evaluation (only female parents were revealed in forming the additive relationship matrix) and individuals' breeding values (BV_{HS}) were predicted. Under the PH model, ANOVA was used to produce site-adjusted phenotypic values by removing site and replication effects, thus increasing the power of detecting genetic differences.

Step 2 – Genotypic or phenotypic pre-selection. To substantially reduce the fingerprinting efforts under the HS and PH models, we applied the phenotypic pre-selection method described in El-Kassaby & Lindgren (2007) and extended the concept to include genotypic pre-selection. Simply, all individuals were

ranked based on either their BV_{HS} or site-adjusted phenotypic values followed by truncation selection. The selection intensity applied varied among the studied scenarios (see below).

Step 3 – Pedigree reconstruction. Pedigree reconstruction was simulated on truncated, pre-selected populations consisting of 2000 individuals for the PH model and 500, 1000 and 2000 for the HS model. We assumed 100% success rate of the pedigree reconstruction; thus male parents were revealed for all individuals in the truncated HS population(s) or both parents for the PS model.

Step 4 – Forward selection. We used a mathematical programming framework to optimize genotypic contributions to the deployment population (seed orchard) and thereby the response to selection is maximized. The objective function is formulated as

$$\max \rightarrow \sum_{i=1}^{n} \hat{x}_i p_i, \tag{1}$$

where $\hat{x}$ is the $n \times 1$ vector of predicted breeding values and p is the corresponding vector of genotypic contributions, subject to the optimization, where $\sum_{i=1}^{n} p_i = 1$ and $0 \leqslant p_i \leqslant 1$, and n is the total number of selection candidates (both parents known), which depends on the strategy used (Step 3).

To maintain a desired level of gene diversity, the solution is constrained by the following function that sets a minimum effective population size value:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} A_{i,j} p_i p_j \leqslant \theta_{\max}, \tag{2}$$

where A is the $n \times n$ additive relationship matrix (Henderson, 1984) and $\theta_{\max}$ is the maximum group coancestry, converted from the minimum effective population size (status number, N_S, Lindgren *et al.*, 1996)

$$\theta_{\max} = \frac{1}{2N_S}. \tag{3}$$

Note that when A is a scalar matrix of the form $0.5I$ (where I is the identity matrix), the solution is equal to the Linear Deployment method of Lindgren & Matheson (1986), and in such a case, the resulting contributions are in linear association with corresponding breeding values. In other cases (including scenarios of the current study), the association deviates from the linearity.

The optimization was conducted using the MOSEK$^{\circledR}$ software (Anonymous, 2002). It solves large-scale mathematical optimization problems, including linear, quadratic and quadratically constrained ones and general convex nonlinear problems, covering all scenarios presented in the current study. Further, optimality tolerance level was controlled and purposely set high (using the default value of 10E-07), suggesting that the optimization output is within the

0.00001% range of the true optimum. From the practical point of view, we therefore regard all results as optimum.

Variance of the genetic response (eqn 1) is

$$\sigma^2 \left\{ \sum_{i=1}^{n} \hat{x}_i p_i \right\} = \sum_{i=1}^{n} \sigma^2 \{\hat{x}_i\} p_i^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sigma\{\hat{x}_i, \hat{x}_j\} p_i p_j, \tag{4}$$

where $\sigma^2(\hat{x}_i)$ is the variance associated with the selection criterion prediction in the ith individual, and $\sigma\{\hat{x}_i, \hat{x}_j\}$ is the expected genetic covariance between individuals i and j. In this study, we assumed that $\sigma\{\hat{x}_i, \hat{x}_j\} = r_{i,j}\hat{\sigma}_A^2$, where $\hat{\sigma}_A^2$ is the additive genetic variance and $r_{i,j}$ is the fraction of the additive genetic variance appearing in the covariance (Falconer & Mackay, 1996). The contribution of additional genetic effects to the covariance was assumed negligible (Hill *et al.*, 2008).

Confidence intervals for the genetic response prediction were constructed as follows:

$$\sum_{i=1}^{n} \hat{x}_i p_i \pm t^* \frac{\sigma\left\{\sum_{i=1}^{n} \hat{x}_i p_i\right\}}{\sqrt{N_S}}, \tag{5}$$

where t^* is the critical value for the t distribution.

(iii) *Theoretical efficiencies of HS and PH strategies*

Three possible sources of information enter the genetic evaluation, i.e. the phenotypic measurement of an individual (P_1, all strategies), the average record of m respective half-sibs (P_2, HS and FS), and that of n full-sibs (P_3, FS). Correspondingly, we constructed a selection index (Smith, 1936; Hazel, 1943)

$$I = b_1 P_1 + b_2 P_2 + b_3 P_3, \tag{6}$$

where the b's are weighting factors, optimizing the three sources of information relative to the index value (selection criterion). The solution is provided by

$$b = P^{-1} G, \tag{7}$$

where (using the rules described in Rönningen & Van Vleck (1985) and assuming a single measurement per individual)

$$P = \begin{pmatrix} 1 & 0.25h^2 & 0.5h^2 \\ 0.25h^2 & \dfrac{1+(m-1)0.25h^2}{m} & 0.25h^2 \\ 0.5h^2 & 0.25h^2 & \dfrac{1+(n-1)0.5h^2}{n} \end{pmatrix} \tag{8}$$

and

$$G = \begin{pmatrix} h^2 \\ 0.25h^2 \\ 0.5h^2 \end{pmatrix} \sigma_P^2, \tag{9}$$

where h^2 refers to the narrow-sense heritability, and σ_P^2 is the phenotypic variance. The first, second and third rows in G (and corresponding columns in P) refer to the inclusion of respective information under PH, HS and FS strategies. The accuracy of selection (correlation between the index and true breeding value) is then

$$r_{\mathrm{IA}} = \sqrt{\frac{b'G}{\sigma_A^2}}, \tag{10}$$

where σ_A^2 is the additive genetic variance.

Under the PH strategy, $r_{\mathrm{IA(PH)}} = h$, considering only the first row (column) of the G (P) matrices. Inclusion of the second row (column) leads to the accuracy of selection under the HS strategy $r_{\mathrm{IA(HS)}}$, and in the same way, inclusion of all three rows (columns) leads to the accuracy of selection under the FS strategy $r_{\mathrm{IA(FS)}}$.

Predicted genetic response to selection (Falconer & Mackay, 1996) is

$$R = i r_{\mathrm{AI}} \sigma_A, \tag{11}$$

where i is the selection intensity.

Assuming i and σ_A remain constant across strategies, relative efficiencies of the PH and HS over FS are $E_{\mathrm{PH/FS}} = r_{\mathrm{AI(PH)}}/r_{\mathrm{AI(FS)}}$, and $E_{\mathrm{HS/FS}} = r_{\mathrm{AI(HS)}}/r_{\mathrm{AI(FS)}}$, respectively. Despite the assumption made to accommodate the current material, the comparison methodology is general to include unequal heritabilities or selection intensities.

3. Results and discussion

Compared with the FS analysis, the HS strategy approximately captured 85% of the genetic gain across all N_S values (Fig. 1). Furthermore, all HS truncations (500, 1000 and 2000) produced the same genetic response irrespective of the applied pre-selection intensity. The PH strategy produced even more surprising results, capturing 75% of the genetic gain available under the FS model across all the diversity range (N_S values, Fig. 1). As expected, under the HS and PH strategies, any increase in genetic diversity was associated with a steady decrease in genetic gain, mirroring that of the original FS analysis. Selection intensity played an important role in determining the relative efficiency magnitude of BWB's partial (HS) and full (PH) pedigree reconstruction methods to that of the conventional breeding (FS) (Fig. 2). Additionally, the observed relative efficiency showed steady improvement with increase in the heritability estimate of the trait under question (Fig. 2).

Unlike crop improvement programmes where the best genotypes are commonly released, forestry deployment concurrently considers balancing genetic gain and diversity, even if gain is intentionally
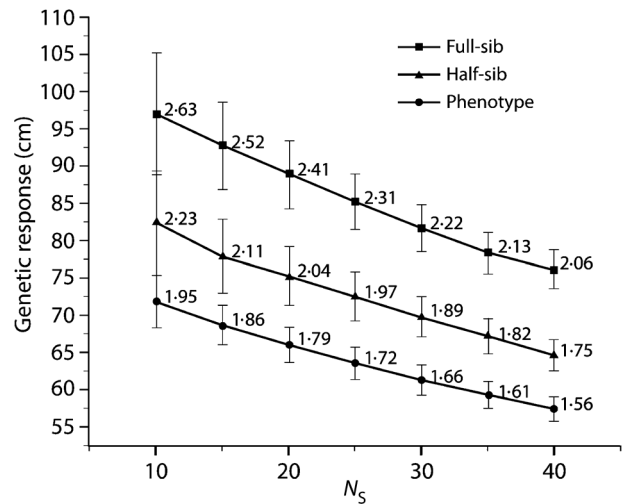

Fig. 1. Eight-year height average breeding value (genetic response) comparison for the HS and PH models relative to the diallel (FS) model for a range of effective population sizes ($N_S = 10–40$), values are plotted along with 95% confidence intervals. Data labels indicate genetic response in units of standard deviation from a candidate population mean of (0·0).

sacrificed (El-Kassaby & Askew, 2004). Therefore, the observed decline in average genetic gain with increased diversity is not a cause for concern and in fact most deployment populations (seed orchards) harbour a high number of parents, so the commonly observed attrition in reproductive output among orchard's parents is buffered (El-Kassaby et al., 1989a; Roberds et al., 1991; Savolainen et al., 1993; Lindgren et al., 2004).

The observed trend of genetic gain, over the range of diversity, followed theoretical quantitative genetics expectations (Falconer & Mackay, 1996) with reduced efficiency from FS to HS to PH (Fig. 1). The genetic gain captured by either the HS or PH strategy relative to the FS model is striking, given the reduced effort associated with the former methods. Conventional breeding programmes are structured around systematic repeated cycles of breeding, testing and selection requiring substantial planning and resources. Our BWB strategies require much less effort and capitalize on the assembly of natural crosses among selected parents (see Fig. 3, for methods' timeline comparison). This approach could be extended to existing operational regeneration plantings such as single or mixed family blocks (McKeand et al., 2003) as a substitute for classical progeny testing trials. In the case of single family blocks where maternal parents are known, individuals are ranked based on their BV_{HS} and partial pedigree reconstruction is only required for the truncated portion to identify the paternal parent. The efficiency of selecting within single family blocks is proportional to the number of deployed families. In cases where few
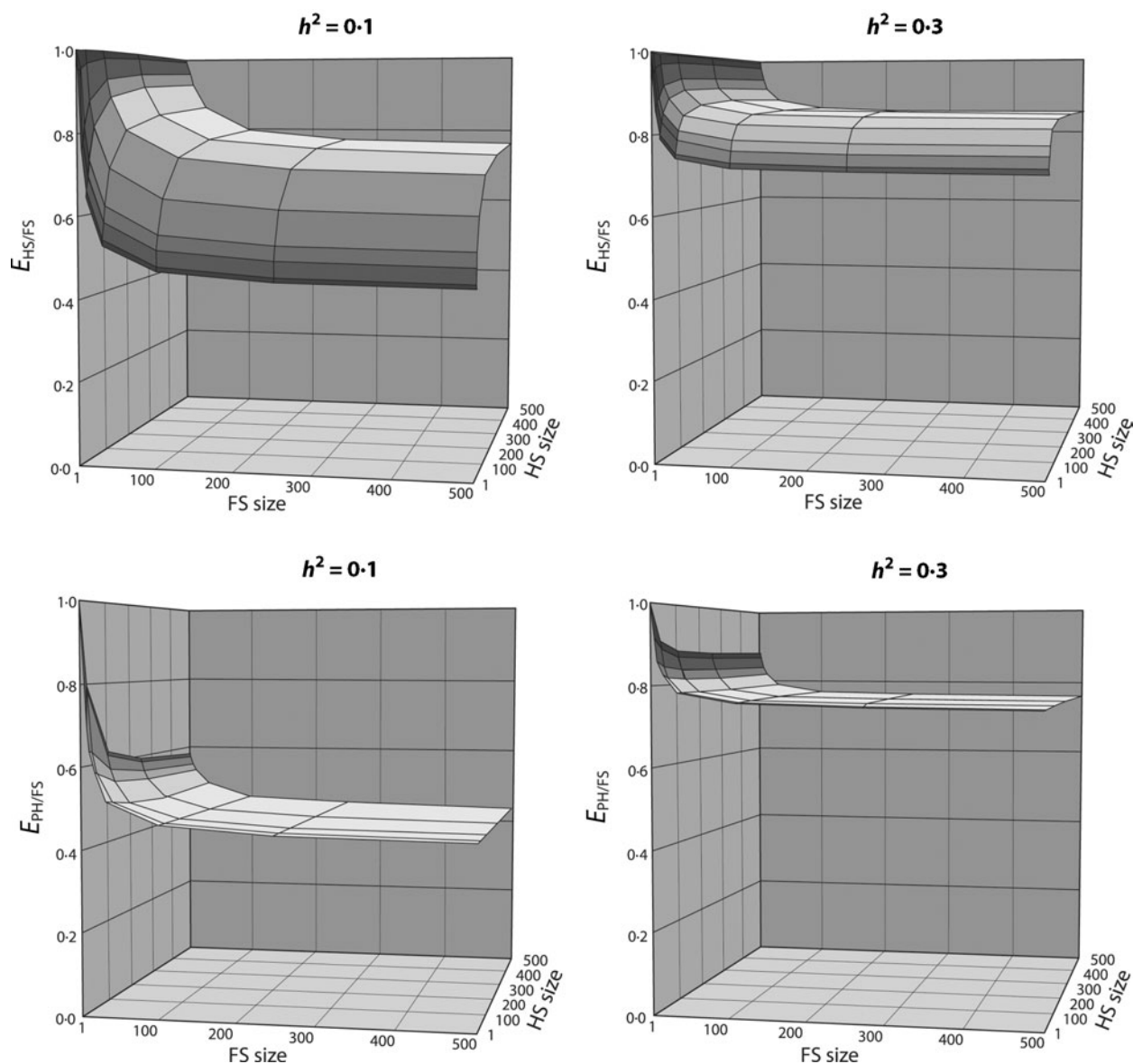
Fig. 2. BWB's theoretical relative efficiency for the HS and PH models to the diallel (FS) model for heritability values of 0·1 and 0·3 over ranges of different HS and FS family sizes.

higher-ranking families are deployed to maximize the response to selection (i.e. high selection intensity), then selection within these families is not recommended as it leads to a substantially reduced effective population size (N_e), and it would minimize our ability to attain high gain through the identification of superior offspring. This can be explained by the relatively small variation that exists among these elite families, as opposed to that of within-family (Lstibůrek *et al.*, 2006). So, we recommend this approach only when a large number of families exist or selection from the simplified HS testing after revealing the full pedigree of the candidate population. On the other hand, when mixed family blocks are used, full pedigree reconstruction is required on the phenotypically selected individuals to determine their maternal and paternal parents. If 'operational' plantings will be considered as a substitute for 'groomed' progeny testing trials, then it is recommended that site spatial variation removal techniques should be implemented to improve the efficiency of either genotypic and/or phenotypic pre-selection and thus increase the effectiveness of truncation selection. Site spatial variation removal techniques are commonly used as a tool for detecting global site heterogeneity or environmental gradients in agriculture and forestry (Dutkowski *et al.*, 2002; Cappa & Cantet, 2007). When unstructured populations are used, it should be expected that the precision of the estimated genetic parameters would be lower than expected from the FS option (see below). The PH selection method is presented to highlight the range of options available with the use of
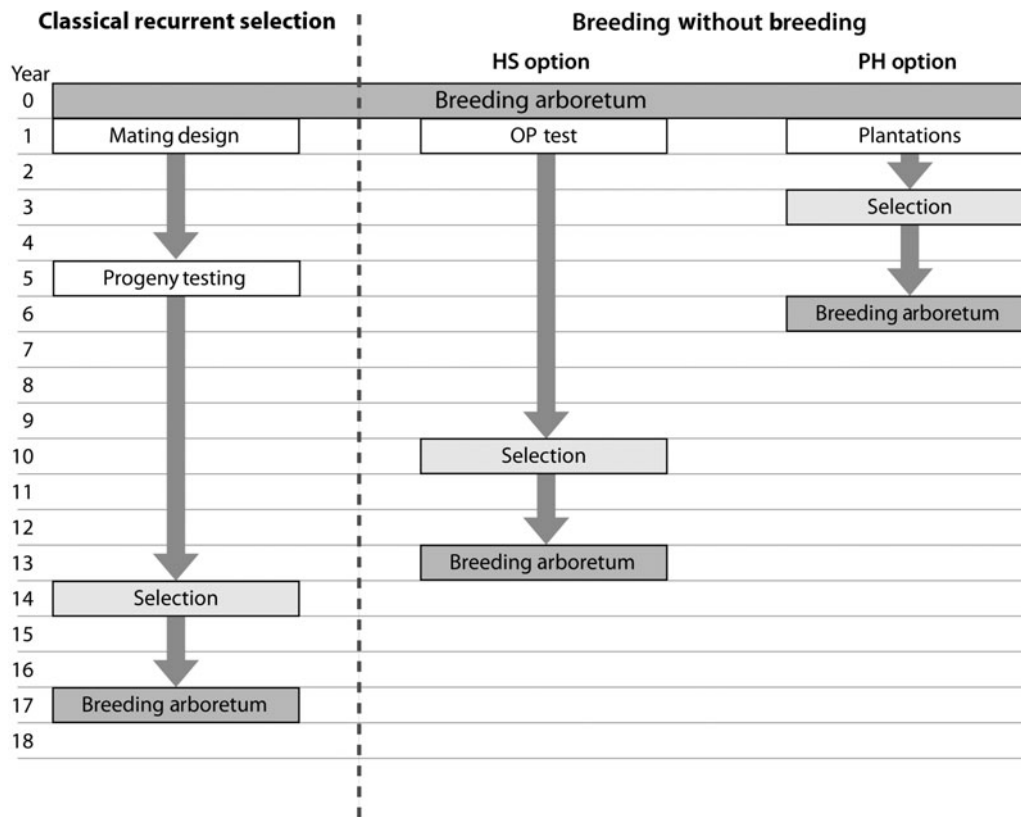
Fig. 3. Flowchart depicting the timeline for a 'classical' recurrent selection scheme and the two BWB proposed options (HS: simplified OP testing; PH: no testing (mass selection)). This comparison assumes that the initial breeding arboretum is sexually mature, with already established plantations of mixed families, and progeny evaluation is conducted 8 years after planting followed by selection and the establishment of the next generation breeding arboretum 2 years after selection. Both BWB options assume that phenotypic assessment, site variability removal, genetic and/or phenotypic ranking and truncation are conducted in 1 year and DNA fingerprinting is conducted the following year. Infusion through the use of supplemental-mass-pollination and introduction of new tested and/or untested parents is assumed for all breeding arboreta.

DNA fingerprinting and pedigree reconstruction. This option should be considered in situations where modest gain increments are better than no gain. For example, when breeding resources are distributed among multiple species and/or among several breeding zones within a species, programmes are always restricted to species or zones with the highest return on investment (ROI). The use of the PH option would offer an opportunity to include some of the economically lesser species and/or extend testing to zones with lower ROI.

Our method provided the means to estimate the genetic parameters, particularly breeding values and coancestry, needed for capturing genetic gain and managing diversity. Genetic parameters obtained from structured mating designs and conventional testing, while potentially more accurate, are associated with a substantial amount of added efforts and costs that do not necessarily contribute much to the captured gain. The retrospective study demonstrated the usefulness and applicability of BWB; however, it should be pointed out that all the presented genetic gain estimates were relative to and dependent on the available information from the FS model. Accordingly, we derived a simplified method to predict the relative efficiency of partial (HS) and full pedigree (PH) reconstruction strategies over that of the commonly practised FS evaluation (Fig. 2). The theoretical method developed is solely based on the original breeding or phenotypic values produced from the pre-selected individuals for either the HS or PH strategy; thus it represents minimum baseline gain estimates. According to our theoretical quantitative evaluation, narrow-sense heritability and the within-population genetic relatedness (size of FS and HS families) are the two key factors affecting the relative genetic efficiency of BWB.

Using the 8-year height's narrow sense heritability estimate from the FS analysis ($h^2 = 0.19$; S.E.M. = 0.045), our minimum relative efficiency estimations of BWB compared with the FS model are 84.4 and 61.9% for the HS and PH, respectively (Fig. 2). Comparing these theoretical expectations with the results obtained for the BWB simulation (Fig. 1)

indicates that the HS strategy is close to attaining its expected efficiency estimate, whereas the PH was higher than expected. Since we used the 8-year height narrow sense heritability estimate produced from the diallel conventional testing, it is safe to assume that it is higher than what is expected under 'less than ideal' situation (i.e. after pedigree reconstruction family sizes would be different). It should be stated that even with the balanced data used as the initial 'set-up' in the present study, a 'less than ideal' scenario is expected to emerge after ranking and truncation selection (i.e. a subset of the original data with unequal representation of the tested parents). The theoretical impact of this genetic and statistical imbalance on estimating the genetic parameters (i.e. prediction of breeding values) by BLUP using the animal (individual tree) models will be unbiased if all known additive relationships are accounted for in matrix **A** even in the presence of selection. This situation remains unchanged even in the presence of fertility variation among the breeding arboretum parents (Henderson, 1988; Kennedy *et al.*, 1988).

The PH option results, although lower and less precise than the other options, confirm earlier breeding strategy comparisons that highlighted the increased long-term efficiency of phenotypic selection as compared with other breeding schemes (Cotterill, 1986; Andersson *et al.*, 1998). However, as indicated in the Introduction section (above), while phenotypic (mass) selection is tempting for its simplicity, this approach requires DNA fingerprinting and pedigree reconstruction augmentation to detect the unnoticeable build-up of coancestry that will ultimately negate the method's advantages. It is also important to mention that when the relative efficiency comparisons were made, both the size and genetic structure of the selected individuals were instrumental in both options' success (Fig. 2). The ideal situation arises when many small full-sib families are available within half-sib families (Fig. 2).

The proposed method is anchored around pedigree reconstruction, which, in the present study, was assumed to be error-free. Critical factors in pedigree reconstruction are the number of loci used, their polymorphic information content (Botstein *et al.*, 1980) and the rate of genotyping error (Vandeputte, *et al.*, 2006; Kalinowski *et al.*, 2007; Wang, 2004, 2007). While most pedigree reconstruction methods account for genotyping error, the ability to estimate its magnitude and adjusting its impact are essential so individuals selected for either future breeding or production populations are all authentic. A limited number of informative loci have proved to be very effective in pedigree reconstruction (Gerber *et al.*, 2000; El-Kassaby *et al.*, 2006); however, increasing the number of informative loci has proved to be instrumental in reducing type I error (false assignment) and type II error (false exclusion) (Vandeputte *et al.*, 2006). Additionally, the level of immigration into the breeding arboreta is expected to proportionally increase the fingerprinting efforts (Pompanon *et al.*, 2005). However, we expect that most of the individuals resulting from immigration/contamination (unselected parents) will rank low during either the genotypic or phenotypic pre-selection as a result of the selection deferential between the breeding arboretum and contamination source and thus most of the individuals sired by outside pollen will be eliminated even under the known high migration rate observed in many seed orchards (Friedman & Adams, 1985; El-Kassaby *et al.*, 1989*b*). Highly ranked individuals, even if they are generated from pollen migration, should be retained in the testing population and their inclusion in further breeding and/or production selection could be considered as an infusion. In a parallel western larch (*Larix occidentalis* Nutt.) field testing experiment consisting of 14 open-pollinated families collected from a 41-parent seed orchard with 22 % pollen contamination, we successfully constructed the pedigree and several FS families were identified within either the tested 14 females or the orchard's 41 males. The pedigree reconstruction resulted into the emergence of a mating design with females sired by an average of 16 males and males mated with an average of five females, producing a testing population with ample mating bridges for the prediction of breeding values using the BLUP analysis (Funda *et al.*, 2008). Additionally, the rate of pollen contamination into the orchard/breeding arboretum could be reduced with pollen augmentation techniques (e.g. supplemental-mass-pollination: Wakeley *et al.*, 1966). This technique has proved to be effective in reducing pollen contamination (El-Kassaby & Davidson, 1990) and successfully siring up to 8 % of the offspring using only one pollen application (El-Kassaby *et al.*, 1993). Thus pollen augmentation could also be viewed as a means to expand the size of the breeding arboreta (i.e. the number of parents) through the introduction of new parents.

The cost associated with BWB implementation is mainly dependent on the number of individuals requiring genotyping (i.e. the intensity of pre-selection). Considering the lack of observed differences in the genetic response among the three intensities (HS 500, 1000 and 2000) used in this study, it could be argued that further reduction in the number of individuals for genotyping could be possible. However, a balance between the number of genotyped individuals and the accuracy of pedigree reconstruction should be considered. The primary role affecting the pre-selection intensity is limiting false assignments and improvement of the exclusion probabilities of pedigree reconstruction.

Cost comparisons between FS breeding and either of the two proposed options should be viewed in general terms to avoid any species- or case-specific scenarios. Fair comparison should account for all important cost and time components (see Fig. 3). In the present study, we have only focused on presenting a general methodology to calculate genetic-based efficiency, which by definition is expected to be lower than 1, as the BWB strategy is based on a less informative evaluation scheme. Additionally, net present value of the BWB and FS strategies should be calculated and their ratio should be reported as the 'relative efficiency of the BWB'.

Incorporating the DNA fingerprinting and partial pedigree reconstruction into tree breeding was first attempted by Lambeth *et al.* (2001) in the complementary polycross mating design and followed by Grattapaglia *et al.* (2004) in a unique case of *Eucalyptus* inter-specific hybridization. While DNA fingerprinting and parentage assignment were restricted to the identification of a limited number of male parents forming the polymix (i.e. pollen mix donors) following controlled crosses, Lambeth *et al.*'s (2001) method successfully transformed the polycross liabilities (less genetic gain potential and limited potential for forward selection) into assets (simplicity and complete pedigree information). This approach has gained increased attention and evidence for unequal paternal contributions was reported (Wheeler *et al.*, 2006; Kumar *et al.*, 2007; Doerksen & Herbinger, 2008). Grattapaglia *et al.*'s (2004) inter-specific hybridization work, on the other hand, only involved six *E. urophylla* males and a single *E. grandis* female clones. These males were effectively further reduced to only three due to an incident of parental labelling error and a very limited reproductive output from two other males (lack of (0/149) or extremely reduced (1/149) contribution), thus reducing the number of tested males to three. Notwithstanding the originality of the two approaches listed above, it should be stated that they differ substantially from our proposed BWB approach, which is targeted at operational tree breeding programmes aiming at capturing general combining ability with unrestricted number of parents, possibly in much larger population sizes (thousands of selection candidates), and the obvious lack of control crosses.

In conclusion, we believe that BWB is a competitive strategy to current breeding programmes, providing an effective and economic method to breed outcrossing species. In the same way, the method is fully applicable within zones of lesser economic value that do not warrant the establishment of a full structured programme such as minor species and respective gene management/conservation programmes. As the cost of DNA fingerprinting is steadily declining due to the availability of commercially independent high-throughput laboratories, the ease of accessing numerous sources of primers for many species in the Internet and/or scientific journals (e.g. *Molecular Ecology Research* is the default depository for most developed markers), and the cross amplification of conspecific, congeners and confamilial markers, all reduce the need for a species-specific primer, thus making BWB a viable option for developing countries, where the resources are too limited to support classical breeding programmes. More importantly, breeding programmes are in themselves an efficient means of gene conservation and the development of a low-cost but efficient approach to tree improvement would provide an incentive for countries that would have not otherwise developed breeding/conservation programmes.

References

Adams, W. T., Neale, D. B. & Loopstra, C. A. (1988). Verifying controlled crosses in conifer tree-improvement programs. *Silvae Genetica* **37**, 147–152.

Andersson, E. W., Spanos, K. A., Mullin, T. J. & Lindgren, D. (1998). Phenotypic selection compared to restricted combined index selection for many generations. *Silva Fennica* **32**, 111–120.

Anonymous (2002). *The MOSEK.NET API Manual, Version 4.0 (Revision 60)*. Denmark: Mosek. 341 pp.

Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. (1980). Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. *The American Journal of Human Genetics* **32**, 314–331.

Cappa, E. P. & Cantet, R. J. C. (2007). Bayesian estimation of surface to account for spatial trend using penalized splines in an individual-tree mixed model. *Canadian Journal of Forest Research* **37**, 2677–2688.

Cotterill, P. P. (1986). Breeding strategy: don't underestimate simplicity. In *Proceedings of the Joint IUFRO Meeting, Working Parties on Breeding Theory, Progeny Testing and Seed Orchards (Williamsburg, VA)*, pp. 8–23.

Devey, M. E., Bell, J. C., Uren, T. L. & Moran, G. F. (2002). A set of microsatellite markers for fingerprinting and breeding applications in *Pinus radiata*. *Genome* **45**, 984–989.

Doerksen, T. K. & Herbinger, C. M. (2008). Male reproductive success and pedigree error in red spruce open-pollinated and polycross mating systems. *Canadian Journal of Forest Research* **38**, 1742–1749.

Dutkowski, G. W., Costa-Silva, J., Gilmour, A. R. & Lopez, G. A. (2002). Spatial analysis methods for forest genetic trials. *Canadian Journal of Forest Research* **32**, 2201–2214.

El-Kassaby, Y. A. & Askew, G. R. (2004). A method for calculating expected genetic diversity of vegetatively propagated populations produced from commonly used mating designs. In *Plantation Forest Biotechnology for the 21st Century* (eds C. Walter & M. Carson), pp. 129–146. Trivandrum, India: Research Signpost.

El-Kassaby, Y. A. & Davidson, R. (1990). Impact of crop management practices on the seed crop genetic quality in a Douglas-fir seed orchard. *Silvae Genetica* **39**, 230–237.

El-Kassaby, Y. A. & Lindgren, D. (2007). Increasing the efficiency of breeding without breeding through phenotypic preselection in open pollinated progenies. In *29th Southern Forest Tree Improvement Conference and the Western Forest Genetics Association, Galveston, TX*, pp. 15–19.

El-Kassaby, Y. A., Fashler, A. M. K. & Crown, M. (1989*a*). Variation in fruitfulness in a Douglas-fir seed orchard and its effect on crop management decisions. *Silvae Genetica* **38**, 113–121.

El-Kassaby, Y. A., Rudin, D. & Yazdani, R. (1989*b*). Levels of outcrossing and contamination in two Scots pine seed orchards. *Scandinavian Journal of Forest Research* **4**, 41–49.

El-Kassaby, Y. A., Barnes, S., Cook, C. & MacLeod, D. A. (1993). Supplemental-mass-pollination success rate in a mature Douglas-fir seed orchard. *Canadian Journal of Forest Research* **23**, 1069–1099.

El-Kassaby, Y. A., Lstibůrek, M., Liewlaksaneeyanawin, C., Slavov, G. T. & Howe, G. T. (2006). *Low Input Breeding and Genetic Conservation of Forest Tree Species*. Antalya, Turkey: IUFRO. http://www.akdeniz.edu.tr/english/iufro/2007.pdf

Eriksson, G., Namkoong, G. & Roberds, J. H. (1993). Dynamic gene conservation for uncertain futures. *Forest Ecology and Management* **62**, 15–37.

Falconer, D. S. & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*, 4th edn. New York: Longman.

Friedman, S. T. & Adams, W. T. (1985). Estimation of gene flow into seed orchards of loblolly pine (*Pinus taeda* L.). *Theoretical and Applied Genetics* **69**, 609–615.

Funda, T., Chen, C., Liewlaksaneeyanawin, C., Kenawy, A. & El-Kassaby, Y. A. (2008). Pedigree and mating system analyses in a western larch (*Larix occidentalis* Nutt.) experimental population. *Annals of Forest Science* **65**, 705.

Gerber, S., Mariette, S., Streiff, R., Bodénès, C. & Kremer, A. (2000). Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. *Molecular Ecology* **9**, 1037–1048.

Gilmour, A. R., Gogel, B. J., Cullis, B. R. & Thompson, R. (2006). *ASReml User Guide, Release 2.0*. Hemel Hempstead, UK: VSN International.

Grattapaglia, D., Ribeiro, V. J. & Rezende, G. D. S. P. (2004). Retrospective selection of elite parent trees using paternity testing with microsatellite markers: an alternative short term breeding tactic for *Eucalyptus*. *Theoretical and Applied Genetics* **109**, 192–199.

Griffing, B. (1956). Concept of general and specific combing ability in relation to diallel crossing system. *Australian Journal of Biological Sciences* **9**, 463–493.

Hazel, L. N. (1943). The genetic basis for constructing selection indexes. *Genetics* **28**, 476–490.

Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* **31**, 423–447.

Henderson, C. R. (1984). *Applications of Linear Models in Animal Breeding*. Guelph, ON: University of Guelph.

Henderson, C. R. (1988). Theoretical basis and computational methods for a number of different animal models. *Journal of Dairy Science* **71** (Suppl. 2), 1–16.

Hill, W. G., Goddard, M. E. & Visscher, P. M. (2008). Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics* **4**, 1–10.

Jones, A. G. & Ardren, W. R. (2003). Methods of parentage analysis in natural populations. *Molecular Ecology* **12**, 2511–2523.

Kalinowski, S. T., Taper, M. L. & Marshall, T. C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* **16**, 1099–1106.

Kennedy, B. W., Schaeffer, L. R. & Sorensen, D. A. (1988). Genetic properties of animal models. *Journal of Dairy Science* **71** (Suppl. 2), 17–26.

Kumar, S., Gerber, S., Richardson, T. & Gea, L. (2007). Testing for unequal paternal contributions using nuclear and chloroplast SSR markers in polycross families of radiata pine. *Tree Genetics and Genomes* **3**, 207–214.

Lambeth, C., Lee, B. C., O'Malley, D. & Wheeler, N. (2001). Polymix breeding with parental analysis of progeny: an alternative to full-sib breeding and testing. *Theoretical and Applied Genetics* **103**, 930–943.

Lindgren, D. & Matheson, A. C. (1986). Increasing the genetic quality of seed from seed orchards by using the better clones in higher proportions. *Silvae Genetica* **35**, 173–177.

Lindgren, D., Gea, L. D. & Jefferson, P. A. (1996). Loss of genetic diversity monitored by status number. *Silvae Genetica* **45**, 52–59.

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

Lstibůrek, M., Mullin, T. J. & El-Kassaby, Y. A. (2006). The impact of differential success of somatic embryogenesis on the outcome of clonal forestry programs. I. Initial comparison under multitrait selection. *Canadian Journal of Forest Research* **36**, 1376–1384.

Lynch, M. & Walsh, B. (1998). *Genetics and analysis of quantitative traits*. Sunderland, MA: Sinauer Associates.

McKeand, S., Mullin, T., Byram, T. & White, T. (2003). Deployment of genetically improved loblolly and slash pines in the South. *Journal of Forestry* **101**, 32–37.

Namkoong, G. (1966). Inbreeding effects on estimation of genetic additive variance. *Forest Science* **12**, 8–13.

Namkoong, G. (1979). Introduction to quantitative genetics in forestry. Washington, DC: US Department of Agriculture, Forest Service. Technical Bulletin No. 1588.

Namkoong, G., Kang, H. C. & Brouard, J. S. (1988). *Tree Breeding: Principles and Strategies*. New York: Springer-Verlag. Monograph, Theoretical and Applied Genetics 11.

Pompanon, F., Bonin, A., Bellemain, E. & Taberlet, P. (2005). Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics* **6**, 847–859.

Rönninger, K. & Van Vleck, L. D. (1985). *General and Quantitative Genetics*, pp. 187–225. New York: World Animal Science, Elsevier Science Publishing Company.

Roberds, J. H., Friedman, S. T. & El-Kassaby, Y. A. (1991). Effective number of pollen parents in clonal seed orchards. *Theoretical and Applied Genetics* **82**, 313–320.

Savolainen, O., Karkkainen, K., Harju, A., Nikkanen, T. & Rusanen, M. (1993). Fertility variation in *Pinus sylvestris*: a test of sex allocation theory. *American Journal of Botany* **80**, 1016–1020.

Smith, H. (1936). A discriminant function for plant selection. *Annals of Eugenics* **7**, 240–250.

Vandeputte, S., Mauger, S. & Dupont-Nivet, M. (2006). An evaluation of allowing for mismatches as a way to manage genotyping errors in parentage assignment by exclusion. *Molecular Ecology Notes* **6**, 265–267.

Wakeley, P. C., Wells, O. O. & Campbell, T. E. (1966). Mass production of shortleaf × slash pine hybrids by pollinating unbagged female flowers. In *Proceedings of the 2nd Genetics Workshop of the Society of American Foresters and 7th Lake States for Tree Improvement Conference*. Research Paper NC-6. pp. 78–79.

Wang, J. L. (2004). Sibship reconstruction from genetic data with typing errors. *Genetics* **166**, 1963–1979.

Wang, J. L. (2007). Parentage and sibship exclusions: higher statistical power with more family members. *Heredity* **99**, 205–217.

Wheeler, N., Payne, P., Hipkins, V., Saich, R., Kenny, S. & Tuskan, G. (2006). Polymix breeding with paternity analysis in *Populus*: a test for differential reproductive success (DRS) among pollen donors. *Tree Genetics and Genomes* **2**, 1614–2950.

Yanchuk, A. D. (1996). General and specific combining ability from disconnected partial diallels of coastal Douglas-fir. *Silvae Genetica* **45**, 37–45.

Yanchuk, A. D. (2001). A quantitative framework for breeding and conservation of forest tree genetic resources in British Columbia. *Canadian Journal of Forest Research* **31**, 566–576.

Ying, C. C. & Yanchuk, A. D. (2006). The development of British Columbia's tree seed transfer guidelines: purpose, concept, methodology, and implementation. *Forest Ecology and Management* **227**, 1–13.