

Y. Q. Zheng · D. Lindgren · O. Rosvall · J. Westin

Combining genetic gain and diversity by considering average coancestry in clonal selection of Norway spruce

Received: 7 April 1997 / Accepted: 9 June 1997

Abstract Genetic relationship within a population can be measured by average coancestry. This can also be expressed as an effective number which represents the relative genetic diversity of the population. The goal of breeding can be formulated to maximise genetic value minus average coancestry times a constant (the “penalty constant”). An iterative search algorithm can then be used to find the best selections for meeting this goal. Two such algorithms, one for a fixed number of selections and the other for a variable optimum number, were applied to select a mixture of field-tested Norway spruce clones with known parents. The results were compared with those from the conventional method of restricting parental contributions to the selected population as a means to control diversity. Coancestry-adjusted selection always yielded more gain than restricted selection at a given effective population size (except under circumstances where the methods were equivalent). Expressed another way, at any given level of gain, coancestry-adjusted selection maintained a larger effective population size than did restricted selection. The relative superiority of coancestry-adjusted selection declined when the effective population size approached the lowest value, that at which no penalty or restriction was applied. The method was extended by the second search algorithm to optimise the selected number of clones. The optimal number of clones can be rather large when diversity is heavily valued, but the reduction in genetic gain becomes large.

Key words *Picea abies* · Coancestry · Kinship · Diversity · Clonal selection

Introduction

Combining the objectives of genetic gain and diversity has long been a focus of breeders (Wei 1995 and references cited therein). It is generally recognised that these two contradictory goals must be compromised in selection when genetic relationships among selected individuals are considered. Classical selection methods in tree breeding have usually emphasised performance (predicted breeding value), incorporating genetic relationships only as a limiting constraint rather than as a selection goal. Even when relationship and diversity are considered (often by means of some restriction on relatives), this is never done according to a model formulated for a maximally efficient compromise.

Quinton et al. (1992) introduced the concept of comparing breeding methods at the same level of inbreeding, and it seems to be an acceptable premise to regard inbreeding as the entity that must be compromised in the pursuit of gain (e.g., Caballero et al. 1996; Klieve et al. 1994). Recently, new selection methods have been proposed in which both predicted breeding value and genetic relationship among members in the selected population are considered at the time when the selection decision is made. Brisbane and Gibson (1995) see the compromise as a way to achieve gain with minimal inbreeding, while Lindgren and Mullin (1997) regard low average coancestry (kinship) combined with high gain as the ultimate goal. In the latter approach, which considers genetic relationship in terms of average coancestry, all members of the group are involved in evaluation of the relationship within the selected population. However, by definition, the inbreeding coefficient is more relevant as a measure of individual diversity rather than that of a population. Nevertheless,

Communicated by P. M. A. Tigerstedt

Y. Q. Zheng (✉) · D. Lindgren
Department of Forest Genetics and Plant Physiology,
Swedish University of Agricultural Sciences, S-901 83, Umeå,
Sweden

O. Rosvall · J. Westin
Institute of Forestry Research of Sweden, Box 3, S-918 21, Sävar,
Sweden

in the first cycle of breeding (where most tree breeding programmes currently find themselves), the inbreeding coefficient is still zero, giving little information about the undesirable build-up of relationships because founders are usually unrelated and not inbred. Therefore, the use of average coancestry to describe the genetic relationships among all members of a population seems to be more meaningful.

Unlike classical selection methods, the coancestry-adjusted selection can be understood as a method more or less similar to group selection, which is based on the contribution of each individual to overall group performance (we refer to this as the “benefit”). The benefit of a group depends on both the members’ individual merit and the genetic relationships among them. The genetic relationship is measured by average coancestry. A penalty constant is placed on the average coancestry when the selection criterion is formulated. The theoretical development of coancestry-adjusted selection and its advantages over classical selection methods have been demonstrated by Lindgren and Mullin (1997).

Norway spruce [*Picea abies* (L.) Karst.] is an important tree species which contributes almost half of the standing volume and current growth of Swedish forests and is one of the most important commercial species in the country. The species grows under a wide range of ecological conditions and has been the most important coniferous species in Europe. In Sweden, operational tree breeding programmes of Norway spruce have been initiated, and a number of seed orchards have been established (Danell 1991). A recent development of the breeding programme is the establishment of new seed orchards and clonal mixtures for vegetative mass-propagation using field-tested clonal material. It is possible to propagate Norway spruce vegetatively by rooted cuttings (stecklings), and there are many tests with clones of such materials. The required number of clones in clonal mixtures and how to balance the number of clones against the advantage in gain are burning questions.

The aims of this paper are to apply the method of coancestry-adjusted selection to a real breeding programme where tested clonal material from a partial-diallel mating design are to be selected for clonal mixtures, to develop explicit selection criteria using the method, to further develop coancestry-adjusted selection methods to optimise selection when there is no prior constraint on the total number of selections, and to investigate the advantages of coancestry-adjusted selection over more conventional selection methods.

Theory

A benefit criterion for a group of individuals considering both gain and relationship was formulated by

Lindgren and Mullin (1997) as

$$B_{\omega} = \bar{g}_{\omega} - c\bar{\theta}_{\omega} = \bar{g}_{\omega} - \frac{c}{2N_{\omega}} \quad (1)$$

where B_{ω} is the selection criterion; ω is the set of individuals; c is a constant (the penalty constant), \bar{g}_{ω} and $\bar{\theta}_{\omega}$ are the average genetic value and the average coancestry (or, equivalently, average kinship), respectively, of ω ; and N_{ω} is the “status effective number” of ω , defined by Lindgren et al. (1996) as half the inverse of average coancestry. The goal of breeding is defined as selecting the set ω which maximises B_{ω} .

Theoretically, B_{ω} could be maximised by an exhaustive search of all possible sets of selections, ω , but in general there are too many alternatives. There is no known general procedure to find the exact maximum or to validate a suggested maximum, although iterative search algorithms can be constructed that approach the optimal selected set. Construction of a search algorithm is facilitated by the mathematical treatments described below in which we express the benefit difference when an individual is included in or removed from the selected population.

Consider that the contribution to the benefit of the population ω by member i is $B_{i \in \omega}$, which is the difference between the benefit including i , B_{ω} , and the benefit without i , $B_{\omega - i}$:

$$B_{i \in \omega} = B_{\omega} - B_{\omega - i}$$

where ω has N members and $\omega - i$ has $N - 1$ members.

These expressions for benefit can then be developed in terms of the predicted genetic value of the individuals, \hat{g}_i , and their relationship (coancestry), θ_{ij} , as follows

$$B_{\omega} = \frac{1}{N} \sum_{i \in \omega} \hat{g}_i - \frac{c}{N^2} \sum_{i \in \omega} \sum_{j \in \omega} \theta_{ij} \quad (2)$$

and

$$B_{\omega - i} = \frac{1}{N - 1} \sum_{i \in \omega - i} \hat{g}_i - \frac{c}{(N - 1)^2} \sum_{i \in \omega - i} \sum_{j \in \omega - i} \theta_{ij} \quad (3)$$

After simplification, the expression for the difference is obtained as

$$B_{i \in \omega} = \frac{\hat{g}_i - \bar{g}_{\omega}}{N - 1} - \frac{c(\theta_{ii} + 2\sum_{i \in \omega - i} \theta_{ii} + (1 - 2N)\bar{\theta}_{\omega})}{(N - 1)^2} \quad (4)$$

This expression can be used as a criterion for removing an individual from a population. The individual in the selected population which has the lowest contribution ($B_{i \in \omega}$) to the benefit is removed, and the number of selected individuals is decreased by one. This procedure

of removing an individual (Eq. 4) is equivalent to that of maximising $B_{\omega-1}$ (Eq. 3), but is computationally simpler; \bar{g}_ω and $\bar{\theta}_\omega$ need only be calculated once, and the calculations for each individual become linear instead of quadratic with N .

A corresponding criterion can be developed for adding an individual to a selected population, and thus increase the number by one. The benefit ($B_{\omega+i}$) of the population $\omega+i$ with size of $N+1$ is defined as:

$$B_{\omega+i} = \frac{1}{N+1} \sum_{i \in \omega+i} \hat{g}_i - \frac{c}{(N+1)^2} \sum_{i \in \omega+i} \sum_{j \in \omega+i} \theta_{ij} \quad (5)$$

This can be seen as the sum of the benefit (B_ω) of population ω and the contribution ($B_{i \in \omega+i}$) to the benefit of ω by individual i , when it is added. The set ω has N members and the set $\omega+i$ has $N+1$ members. The contribution of individual i when it is added to the previously selected population ω is $B_{i \in \omega+i} = B_{\omega+i} - B_\omega$. An explicit expression for this can be obtained analogously to Eq. 4 by considering formulae 2 and 5.

$$B_{i \in \omega+i} = \frac{\hat{g}_i - \bar{g}_\omega}{N+1} - \frac{c(\theta_{ii} + 2 \sum_{i \in \omega} \theta_{ii} - (1+2N)\bar{\theta}_\omega)}{(N+1)^2} \quad (6)$$

This expression can be used as the criterion for adding an individual into the selected population.

Note that Brisbane and Gibson (1995) and Ballou and Lacy (1995) used a similar selection criterion for a similar reason, but theirs did not consider self-coancestry in the way that our algebra suggests and is thus probably less efficient. The algebra given here also considers the number of selections (N) explicitly making it possible to let the number of selections vary. The number of selections was fixed in earlier formulations (Lindgren and Mullin 1997; Brisbane and Gibson 1995; Ballou and Lacy 1995).

Material and Methods

Field experiments and mating design

Full-sib families of Norway spruce were used, consisting of seeds from a partial diallel crossing scheme performed in 1983 at the Hissjö clonal seed orchard (latitude 63°56'N, longitude 20°09'E, elevation 85 masl) and those from a clonal archive at Sävar (latitude 63°53'N, longitude 20°35'E, elevation 10 masl). The seed orchard was established with plus-trees selected in mature stands, primarily on the basis of superior height growth. The plus-trees used in the crossings originated within a latitudinal range of 59.5°–67°N, with most of the plus-trees originating within latitude 63°–65°N. Among the plus-trees, 26 were used as seed parents, 26 were used as pollen parents, and 5 were used in reciprocal crossings. The resulting 144 full-sib families were sown in the spring of 1986. Depending on the number of seeds available for each family, up to 105 seeds were used; these were divided into 15 replications with 7 seeds in each replication. The seedlings were grown in a greenhouse and nursery beds

according to standard procedures. Depending on the number of seedlings available in each family, from 1 to 7 seedlings in each full-sib family were selected in 1988 for clonal propagation. In most of the families, 3 seedlings were selected at random while the remainder were selected primarily for superior height growth.

Cuttings from the selected 2-year-old ortets were struck for rooting in 1991. In order to reduce the influence of shoot quality on steckling size and to produce stecklings with even size, we followed a protocol for producing stecklings based on two growing seasons in 1 year. The shoots were cut in October and stored at -2°C until rooting. The rooting phase, with soil temperatures around 20°C and air temperatures around $5\text{--}10^\circ\text{C}$, started in late December and continued for approximately 8 weeks. After the rooting phase several phases followed in sequence: a growth phase (6–8 weeks); a bud-setting phase with short-day conditions (4–6 weeks); a chilling phase at $+4^\circ\text{C}$ to break bud dormancy (4 weeks); a second growth phase, and a second bud-setting phase. After hardening under short-day conditions and low temperatures, the stecklings were then stored over winter in a freezer.

The clones were established in field trials at four locations between 62.9°N and 64.2°N , in the spring of 1992. The tests were designed as completely randomised trials with single-tree plots. Two of the field trials represented relatively warm locations and the other two relatively cold locations. If available, two ramets per clone were planted in each of the four field trials. When fewer ramets were available, two ramets were first distributed to the climatically most favourable field trial and thereafter to the other trials. If more than eight ramets were available, the extra ramets were distributed randomly to all four field trials. Height measurements were performed in all field trials in the autumn of 1996, after five growing seasons.

Statistical model

The statistical analysis was conducted with the MIXED procedure of SAS (SAS Institute 1996). Models used for estimation of variance components were:

$$Y_{ijklm} = \mu + b_i + m_j + f_{ik} + c_l(m_{ij}, f_{ik}) + e_{ijklm} \quad (7)$$

where

Y = phenotypic value of clone l ;

μ = overall mean;

b = fixed effect of trial location i ($i = 1, 2, 3, 4$);

m = random effect of seed parent j at trial location i ($j = 1, 2, \dots, 31$);

f = random effect of father k trial i ($k = 1, 2, \dots, 31$);

c = random effect of clone l nested within family with seed parent j and pollen parent k ($l = 1, 2, \dots, 636$); and

e = residual effect over m ramets ($0, \sigma^2$).

Clone values (BLUP) were calculated as: overall mean + seed parent effect + pollen parent effect + clone effect.

Algorithms

Restricted selection for a fixed number of selections using a conventional algorithm (algorithm 1)

1) Decide the number of selections (n) and the restriction limit on the number of selections per parent (r).

2) Rank all genotypes on their predicted genetic values. Define a counter for each parent to record their contributions to the selected population and set the counters initially to 0.

3) Identify parents of each genotype in turn, beginning with the highest ranked. If both parental contribution counters are smaller than the restriction number, the genotype is selected and its parents' counters increased by 1. Otherwise, the genotype is discarded.

4) Repeat step 3 until n genotypes are selected. Average genetic value and average coancestry are calculated for the selected population.

This algorithm was repeated for different restriction numbers.

Coancestry-adjusted selection requires the evaluation of benefit (group performance) for all possible alternatives forming the group ω , and this presents practical difficulties as there may be an extremely large number of combinations if the number of candidates for selection is large. Thus, efficient search algorithms are required, and in this study we considered two approaches.

In the first, we fix the number of selections (the size of ω). The algorithm for coancestry-adjusted selection is based on the logic that the final selected group should have the largest value of benefit. Selections are made by screening individuals on their individual contribution to the benefit. The individual contribution is calculated with Eq. 4 (removing individual) or Eq. 6 (adding individual), as follows.

Coancestry-adjusted selection for a fixed number of selections (algorithm 2)

- 1) Decide the number of selections (n) and the value of the penalty constant (c).
- 2) Define a set of selected genotypes (ω , which is empty at the start) and a set of candidates for selection (it includes all genotypes at the start).
- 3) Calculate the contribution ($B_{i \in \omega +}$) to benefit of each candidate (i) when it is included in the selected group using Eq. 6; add the genotype which has the highest contribution into the selected group.
- 4) Repeat step 3 until n genotypes are selected. Record the benefit for the selected group.
- 5) Select one more genotype in the same way so there are $n + 1$ selections.
- 6) Calculate the contribution ($B_{i \in \omega}$) to benefit of each selected genotype (i) using Eq. 4; remove the genotype which has the lowest contribution from the selected group so that n genotypes remain. Calculate the benefit for this new group.
- 7) If the benefit is larger than that obtained in step 4, retain the new group and repeat from step 5. Otherwise, retain the group found in step 4 and terminate the algorithm. Benefit, average genetic value, and average coancestry are recorded for the final selected group.

The algorithm was repeated for different values of penalty factor c .

In the second approach to coancestry-adjusted selection, the size of ω is variable and is determined by the algorithm so that the benefit is maximised for a given penalty on average coancestry. The size of ω that maximises the benefit is considered optimal because it represents the best weighted balance between genetic gain and diversity, as follows.

Coancestry-adjusted selection with flexible number of selections (algorithm 3)

- 1) Decide the value of the penalty constant (c); start the number of selections (n) from 1.
- 2) Perform steps 2–7 of the algorithm 2.
- 3) If the benefit of n is smaller than the benefit of $n - 1$ ($B_n < B_{n-1}$) the algorithm terminates, and $n - 1$ is regarded as the optimal number of selections. Otherwise increase the number of selections (n) by one ($n + 1$); repeat from step 2.

The algorithm is repeated for different values of the penalty constant c .

Computation

An essential requirement for coancestry-adjusted selection is the calculation of the coancestry matrix for all individuals. We make use of the technique of additive relationship matrix (Tier 1990; Lindgren et al. 1997) to compute the coancestry matrix. All parents of candidate clones are assumed to be unrelated and non-inbred. A generalised computer programme was written to carry out the coancestry-adjusted selection and the restricted selection. It was designed to suit a range of situations in a real breeding programme. The size of the selected population could be pre-set to any value or determined by the computer programme. A range of penalty factors and restrictions on parental contribution was tested. For coancestry-adjusted selection, the penalty factor c was set to 0, 10, 20, 30, 40, 50, 100...infinity (Fig. 1). For restricted selection, limits on parental contributions were set from 1 to infinity (no restriction).

Genetic gains from clonal selection ($\Delta G = SH^2$, where S is the selection differential and H^2 the broad-sense heritability) were calculated for the two different selection methods.

Results

Predicted genetic value

The overall mean of clone values for tree height was 484 mm with a range from 391 mm to 632 mm (Table 1). The broad- and narrow-sense heritabilities for tree height were estimated as 0.39 ± 0.037 and 0.33 ± 0.034 , respectively.

Selection with fixed selection number

Selection with restriction on parental contribution

The maximum parental contribution for the selection of 20 clones was found to be 9, i.e., when no restriction was applied, the largest number of progeny selected from a single parent was 9 clones. There are two

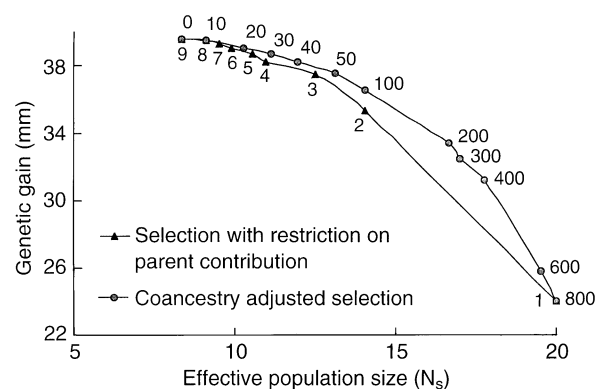


Fig. 1 Genetic gain and effective population size (N_e) for the 20 clones selected by coancestry-adjusted selection or with restriction on parental contributions. Labels of points on the upper line are penalty constants for coancestry-adjusted selection, and those for the lower line are the restriction numbers set for parental contributions under restricted selection

Table 1 Full list of predicted genetic values for height (mm) of each clone. In each cell, the number is the clone ID, and the two letters after the number denote the male and female parents of that clone, respectively. Clonal code in each column increases by 60

	0 +	60 +	120 +	180 +	240 +	300 +	360 +	420 +	480 +	540 +	600 +	Reference
1	456 AB	416 OJ	479 SQ	438 äL	545 ÅV	511 cJ	528 fG	510 kV	439 oT	505 ög	448 uK	A Y4203
2	444 AB	475 OJ	466 TV	498 äL	518 ÅV	501 cJ	488 fG	510 kV	414 oT	491 öe	496 uL	B AC100
3	428 AB	481 OJ	462 TV	471 äL	514 ÅV	509 cJ	492 fG	529 kV	424 oY	433 öe	493 uL	C Z1000
4	453 AB	426 OJ	459 TV	488 ÄT	501 ÅV	543 cJ	460 fG	529 kV	436 oY	537 ön	475 uL	D Z4010
5	415 AB	436 OJ	462 TW	449 ÄT	522 ÅP	537 cJ	490 fG	571 kV	431 oY	479 ön	480 uM	E Y3008
6	435 AB	435 OJ	458 TW	491 ÄT	566 ÅP	501 cJ	513 fd	549 kP	447 oY	505 sd	486 uM	F Y3002
7	474 AB	447 OJ	432 TW	465 ÄT	515 ÅP	536 cC	491 fd	506 kP	432 oa	511 sd	486 uM	G AC102
8	430 AC	463 OP	532 TW	462 ÄT	517 ÅP	498 cC	536 fd	521 kP	469 oa	526 sd	463 NG	H Z3004
9	432 AC	455 OP	414 TW	456 ÄT	517 ÅP	468 cC	445 fd	523 kP	460 oa	532 sd	530 NG	I Z3012
10	383 AC	466 OP	504 TW	504 ÅP	497 ÅP	539 cC	511 fd	519 kP	436 pV	493 sd	482 NG	J Z2014
11	472 AC	486 OP	512 TW	443 ÅP	552 ÅP	522 cC	543 fd	523 kP	541 pV	514 sd	511 NG	K Y3017
12	448 AC	473 OP	484 TX	496 ÅP	616 ÅU	535 cC	518 fd	546 kP	555 pV	551 sW	497 NG	L Z4001
13	421 AD	433 OP	434 TX	508 ÅP	563 ÅU	494 cK	492 fg	562 kU	523 pV	557 sW	435 vw	M AC103
14	413 AE	423 OK	489 TX	481 ÅP	565 ÅU	531 cK	575 fg	544 kU	516 pV	494 sW	468 vw	N AC202
15	448 AE	418 OK	443 TY	501 ÅP	584 ÅU	469 cK	580 fg	545 kU	579 pV	536 sg	438 vw	O Y4200
16	451 AE	407 OK	408 TY	488 ÅP	572 ÅU	487 cK	499 fg	571 kX	508 pW	489 sg	484 vw	P Y2001
17	425 AE	425 OK	463 TY	559 ÅU	632 ÅU	515 cK	529 fg	517 kX	547 pW	488 sg	517 xW	Q Z2005
18	435 AF	439 OK	425 TY	562 ÅU	539 ÅX	489 cK	493 fh	533 kX	485 pW	484 sg	496 xW	R AC300
19	464 BG	445 OK	452 TY	578 ÅU	560 ÅX	516 cL	484 fh	536 kY	537 pW	522 sg	493 xW	S Z4006
20	451 BG	484 OL	451 TZ	565 ÅU	565 ÅX	566 cL	492 fh	525 kR	494 pg	536 sg	483 xW	T AC420
21	490 BG	442 OL	460 TZ	540 ÅU	525 ÅX	583 cL	491 fe	496 kR	505 pg	575 sg	478 xg	U Y2003
22	466 BG	453 OL	415 ÖB	576 ÅU	536 ÅX	552 cL	498 fe	529 kR	482 pg	434 sH	489 xg	V AC103
23	433 BG	468 OL	447 ÖJ	521 ÅM	492 ÅY	490 cL	491 fe	484 lB	495 pg	428 sH	482 xg	W AC205
24	479 BG	456 OL	448 ÖJ	522 ÅM	477 ÅY	563 cL	498 fe	445 lB	506 pg	469 sH	496 xe	X Y300
25	480 BG	443 OL	474 ÖJ	464 ÅM	541 ÅY	579 cM	509 fe	416 lB	521 pg	460 se	447 xe	Y Y300
26	472 BH	454 OM	496 ÖJ	472 ÅM	523 Åa	520 cM	448 hT	411 lB	495 pZ	536 se	521 xe	Z AC204
27	458 BH	422 OM	413 ÖC	460 ÅM	498 Åa	559 cM	435 hT	435 lB	493 pZ	517 se	555 yV	Ö AC101
28	477 BH	446 OM	487 ÖC	484 ÅX	509 Aa	505 cM	421 hT	414 lJ	493 pZ	513 tv	544 yV	Å AC103
29	437 IJ	440 OM	456 ÖC	536 ÅX	442 bT	509 cM	473 hT	411 lJ	474 pq	520 tV	556 yV	Ä Z3000
30	435 IJ	420 OM	440 ÖE	502 äB	483 bT	548 cM	438 hT	417 lJ	474 pq	596 tU	534 yW	ö AC204
31	439 IJ	479 OM	432 ÖE	459 äB	427 bT	505 cM	447 hT	439 lJ	539 pq	586 tU	529 yW	ä Y4204
32	487 IJ	425 OQ	407 ÖE	481 äB	421 bT	499 cQ	450 hP	408 lJ	490 rd	544 tU	524 yW	ä Z4002
33	419 IJ	443 OQ	435 äB	511 äB	424 bT	506 cQ	442 hP	404 lJ	516 rd	543 tU	508 yW	a AC202
34	424 IJ	469 OR	485 äB	478 äB	427 bT	492 cQ	465 hP	433 lC	520 rd	592 tU	482 yY	b AC103
35	463 IJ	499 OR	448 äB	474 äB	414 bT	520 cQ	532 hL	434 lC	516 rd	537 tU	466 yY	c AC101
36	420 IC	428 OR	474 äB	476 äB	490 bP	500 cF	449 hL	408 lC	560 rd	575 tW	480 yY	d Z2015
37	421 IC	527 OR	459 äB	453 äD	475 bP	514 cF	451 hL	422 lK	477 rd	580 tW		e Y4201
38	436 IC	483 OR	426 äB	509 äD	474 bP	540 cF	452 hL	424 lK	489 rW	518 tX		f AC1023
39	427 IC	476 ST	411 äJ	476 äD	463 bP	460 DG	473 hL	402 lK	509 rW	514 tX		g Y2006
40	466 IC	485 ST	433 äJ	440 äD	449 bP	501 DG	464 hL	394 lL	517 rW	546 tX		h Z2013
41	464 IK	431 ST	424 äJ	453 äE	459 bP	475 DG	428 hM	437 lL	499 rg	515 tY		i BD2001
42	447 IK	459 ST	416 äJ	521 äE	482 bP	473 DG	437 hM	439 lL	520 rg	464 tY		j BD1020
43	430 IK	524 ST	437 äJ	475 äE	477 bU	428 DG	478 hM	428 lL	509 rg	505 tY		k AC1031
44	447 IK	468 ST	423 äJ	483 äE	523 bU	497 DG	443 hM	478 mG	506 rg	542 tY		l Y4202
45	466 IK	449 ST	443 äJ	474 äE	519 bU	479 DG	485 hM	496 mg	500 rg	481 tY		m Y2012
46	453 IK	523 SP	461 äC	496 äE	514 bU	519 Dd	442 hM	480 mg	547 re	485 tZ		n Y2004
47	467 IL	558 SP	441 äC	480 äö	514 bU	542 Dd	468 ij	501 mg	536 re	445 tZ		o AC1037
48	472 IL	529 SP	394 äC	539 äö	467 bX	494 Dd	468 ij	527 mH	471 re	455 tZ		p Z2011
49	426 IL	449 SP	429 äC	526 äö	465 bX	574 Dd	464 ij	459 mH	554 rZ	497 uJ		q AC3010
50	454 IL	532 SU	441 äC	591 äö	450 bX	466 Dd	487 ij	426 mH	468 rZ	450 uJ		r Z3006
51	450 IL	528 SU	446 äC	546 äö	473 bX	547 Dd	461 ij	457 me	508 rZ	460 uJ		s Z3011
52	486 IM	532 SU	465 äC	450 ÄT	452 bX	527 Dd	471 ij	508 me	515 öG	472 uJ		t AC2053
53	501 IM	530 SU	451 äE	478 ÄT	500 bX	490 DH	525 ij	510 me	421 öG	454 uJ		u Y2007
54	463 IM	576 SU	455 äE	468 ÄT	463 bX	541 DH	473 kT	493 me	457 öd	481 uJ		v AC1002
55	485 IM	561 SU	460 äE	485 ÄT	405 bY	418 DH	468 kT	466 mn	462 öd	477 uJ		w AC1008
56	443 IM	491SM	430 äE	498 ÄT	442bY	526 De	498 kT	583 oT	456 öd	445 uC		x Z3005
57	443 IM	510 SM	391 äE	472 ÄT	444 bY	557 De	439 kT	487 oT	495 öd	459 uC		y Y3027
58	509 IN	486 SM	406 äE	524 ÅV	468 bR	544 De	487 kT	450 oT	433 öd	450 uC		
59	483 IN	495 SM	452 äE	515 ÅV	446 bR	497 De	488 kT	483 oT	448 ög	474 uC	Mean	
60	457 IN	456 SQ	478 äL	531 ÅV	459 bR	540 De	510 kV	449 oT	473 ög	447 uK	484	

extremes for restricted selection: (1) no restriction, and (2) only one selection per parent. The greatest genetic gain (40 mm) was obtained with no restriction, where a larger number of genotypes were selected from the best performing parents (t and U). This maximum gain was achieved at great expense of genetic diversity, and the status number (effective population size) at this extreme was only 8.3. At the other extreme, selection with maximum restriction produced the lowest genetic gain (24 mm), but the highest genetic diversity (effective population size of 20). In this case, the effective population size equals the census number of the selected population and all selections are from different parents.

These results follow the expected tendency that as effective population size increases, genetic gain decreases. The genetic gain dropped sharply when the effective population size approached the census number.

Coancestry-adjusted selection

The relationship between genetic gain and diversity values for coancestry-adjusted selection was generally similar to that for restricted selection. In the two extreme cases, with zero and infinite values of the penalty constant, results from both selection methods were identical (Fig. 1). In all other cases, i.e., when the penalty constant was greater than zero and less than infinity, the effective population size under coancestry-adjusted selection was larger than that for selection with restriction on parent contribution for any given genetic gain. Coancestry-adjusted selection retained greater genetic diversity in the selected population than did restricted selection. For any given effective population size, coancestry-adjusted selection resulted in greater genetic gain.

The relative superiority of coancestry-adjusted selection over the restricted selection (except for the two extreme cases) was generally greater when larger effective population size was maintained (Fig. 1). At a smaller effective population size, the relative difference between the two methods was less.

Selection with flexible selection number

When the size of ω was determined by the selection algorithm (algorithm 3) and the penalty constant set to zero, the optimal number of selections was determined to be 1, selecting the single genotype with the largest genetic value (Fig. 2). If there were varying degrees of inbreeding in the studied genotypes, the genotype with the highest genetic value would not be necessarily selected. As the penalty increased, the optimal number of selections determined by algorithm 3 also increased. When the penalty was set to infinity, selection was driven mainly on minimising the relationship (average coancestry). The number of selections determined by

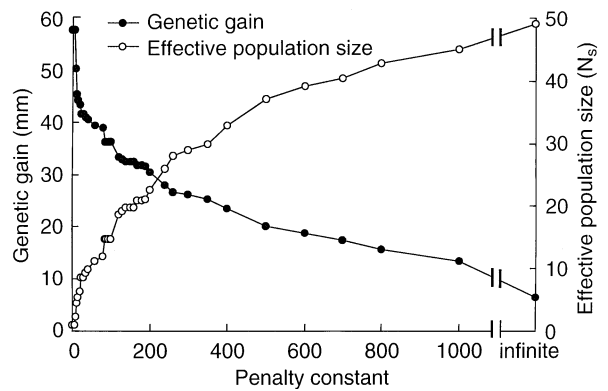


Fig. 2 Coancestry-adjusted selection with a variable selected number. The number of selections giving the best balance of genetic gain and diversity for a given penalty on average coancestry in the selected population is determined by the selection algorithm

the algorithm was 231, yielding an effective population size of 49 and selecting more than one-third of the total number of 636 clones. In this case, virtually no genetic gain (6 mm) was obtained. The genetic gain for each number of selections decreased with increasing penalty value. The diversity measure (status number) varied with the penalty in the opposite way. However, gain and diversity are believed to be best balanced by this algorithm, as the number of selections was determined as that giving the maximum benefit at a given penalty constant.

Discussion

Advantage of coancestry-adjusted selection

The present results from the application to real materials demonstrated that coancestry-adjusted selection is never less effective than selection with restriction on parental contributions. This agrees well with results from a study by Lindgren and Mullin (1997) on simulated and symmetric data, in which coancestry-adjusted selection was compared with restricted selection and found to be substantially superior, even after five generations of selection.

A major advantage of coancestry-adjusted selection is that it is based on group performance rather than individual merit and that it considers genetic relationship candidates with all other members of the selected population. It combines individual merits and relationships with others into a single value, the group performance (benefit), as the selection criterion.

In contrast, restricted selection is based on individual performance (predicted genetic value) with simple limits applied to the number of progeny from the same parent (parental contribution). It does not try to

minimise the average relationship but just to restrict it by minimising the influence of ancestors in a rather simple way. Rather than fixing the restriction number for all parents, the restriction on parental contributions could be allowed to vary so that superior parents can contribute more to the selected population. However, criteria for doing this will be required, and at present there exists no theory for that.

Coancestry-adjusted selection also takes into account relatedness in parents, if any exists, while selection with restriction on parental contributions does not. In this study, the parents were assumed to be unrelated; however, were there relatedness in the parent population, the superiority of coancestry-adjusted selection would be even more evident. The relative superiority over restricted selection is also affected by variation between and among families in genetic value, by the mating scheme, and by the population structure (such as number and size of families). It would be of interest to further explore these influences on coancestry-adjusted selection.

Coancestry-adjusted selection is equivalent to classical selection in several special cases. When the penalty on average coancestry is set to negative infinity, zero, and positive infinity, it is equivalent to family selection, combined-index selection, and within-family selection, respectively. Another advantage of the method is that it uses status number to describe effective population size. As status number is derived from average coancestry, it is a better measure of the genetic diversity in a population than the inbreeding coefficient or the conventional effective population size (Lindgren et al. 1996).

Algorithm 2 is similar to that of Lindgren and Mullin (1997), but we used the individual contribution to the group benefit as the criterion for adding or removing genotype from the selected group while they used the benefit. This improvement largely reduced the computing time required by the algorithm. Algorithm 3 presented here allows the number of selections to vary, while algorithms by Brisbane and Gibson (1995) and Lindgren and Mullin (1997) did not have this feature. This feature enables the algorithm to fit a wide range of purposes in a breeding programme. In practice, breeders have difficulty in deciding how large the breeding population should be. Generally, there is no explicit quantity for balancing genetic gain and diversity available that would enable breeders to make selection decisions. Our algorithm can be used for this purpose. It provides breeders with a useful tool for choosing the optimal size of the breeding population giving the best balance of genetic gain and diversity.

The implementation of coancestry-adjusted selection in a practical breeding programme does not require any extra information compared with the normal selection methods. All calculations are made easily, provided the pedigrees of the individuals are clear. From our experience, coancestry-adjusted selection does not take unreasonable computing time; however, memory

requirements grow when the number of candidates is large, and computing time increases quickly with the number of selections.

Applications

Coancestry adjustment may be applied to situations in which the pedigree is partially known. For example, in the case of open pollination or polymix pollination, the maternal parent is known and the paternal parent is unknown. Coancestry-adjusted selection would enforce a penalty on the relatedness of half sibs in the selected population. Even if the pedigree is unknown, for instance, for plus-trees selected from a wild population, coancestry can be assumed to be equally distributed among individuals. Therefore the coancestry-adjusted selection becomes the selection on breeding values only. By definition, the average inbreeding coefficient in a population is the average coancestry of the parent population after random mating (Falconer 1989). Although one may think that the inbreeding coefficient for a population can be estimated by using allozyme or DNA markers under certain assumptions, it merely reflects the mating system history of the population and cannot be used to predict the genetic relationship in the selected population. It therefore has limited use in selection decisions.

An important application of coancestry-adjusted selection is its use in optimally allocating genetic resources in a breeding programme to balance the short- and long-term genetic gain with change in genetic diversity. For instance, it can be used to choose an appropriate time to introduce new material into the breeding population. As generations advance, inbreeding in the breeding population accumulates and diversity decreases. The introduction of new material into the breeding population has been well recognised as an important issue, but in practice when and how such new material is needed to achieve an optimal balance of genetic gain and diversity remains unanswered. It is possible to apply the method of coancestry-adjusted selection to solve this problem.

Our algorithm can also be extended to select for maximising diversity while disregarding gain. This can be said to correspond to the special case when c equals infinity, but it is better to formulate a selection of genotypes so that $B_\omega = -\bar{\theta}_\omega = -1/2N_\omega$ is maximised. In this case, the selection becomes the minimum coancestry selection as discussed by Askew and Burrows (1983). This can be achieved in a similar way as the ranking of individual mean kinship suggested by Ballou and Lacy (1995), but our selection criteria (Eqs. 4 and 6 disregarding the gain) are slightly more relevant. Ballou and Lacy (1995) suggested selecting individuals on the basis of their individual mean kinship, while we suggest selecting individuals based on their impact on the group of selected individuals.

The choice of an appropriate value for the penalty constant depends on the relative importance of the genetic relationship (coancestry) and the genetic value of the population. There could be much debate on this, but we will not discuss it in this study. Like sensitivity analysis in economics, it is useful to apply a series of penalty values and to analyse the corresponding outcomes. At the moment, a practical technique may be to first decide an acceptable value for the effective population size (N_s) and then maximise group benefit with this target effective size. At later stages, varying the c -value within a reasonable range could lead to a more flexible determination of N_s . This remains the task of further studies.

Acknowledgements Authors thank Tim Mullin for valuable comments on previous version of the manuscript. The Kempe foundation has supported this study.

References

- Askew GR, Burrows PM (1983) Minimum coancestry selection. I. A *Pinus taeda* population and its simulation. *Silvae Genet* 32:125–131
- Ballou JD, Lacy RC (1995) Identifying genetically important individuals for management of genetic variation in pedigreed populations. In: Ballou, JD, Gilpin, M, Foose TJ (eds) Population management for survival and recovery. Columbia University Press, New York, pp 76–111
- Brisbane JR, Gibson JP (1995) Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions. *Theor Appl Genet* 91:421–431
- Caballero A, Santiago D, Toro MA (1996) Systems of mating to reduce inbreeding in selected populations. *Anim Sci* 62:431–442
- Danell Ö (1991) Survey of past, current and future Swedish forest tree breeding. *Silva Fen* 25:241–247
- Falconer DS (1989) Introduction to quantitative genetics, 3rd edn. Longman. London New York
- Klieve HM, Kinghorn BP, Barwick SA (1994) The joint regulation of genetic gain and inbreeding under mate selection. *J Anim Breed Genet* 111:81–88
- Lindgren D, Mullin T (1997) Selection for coancestry-adjusted performance to combine gain and effective population size. *Silvae Genet* (in press)
- Lindgren D, Gea LD, Jefferson PA (1996) Loss of genetic diversity monitored by status number. *Silvae Genet* 45:52–59
- Lindgren D, Gea LD, Jefferson PA (1997) Status number for monitoring genetic diversity. *For Genet* 4(2):69–76
- Quinton M, Smith C, Goddard ME (1992) Comparison of selection methods at the same level of inbreeding. *J Anim Sci* 70:1060–1067
- SAS Institute (1996) SAS/STAT software: changes and enhancements through release 6.11. SAS Institute Inc, Cary, N.C.
- Tier B (1990) Computing inbreeding coefficients quickly. *Genet Sel Evol* 22:419–430
- Wei P (1995) Predicting genetic diversity and optimising selection in breeding programmes. PhD thesis, Swedish University of Agricultural Sciences, Department of Forest Genetics and Plant Physiology, Umeå, Sweden