Theor Appl Genet (2005) 110: 1236–1243
DOI 10.1007/s00122-005-1957-0

ORIGINAL PAPER

Fikret Isik · Dennis D. Boos · Bailian Li

The distribution of genetic parameter estimates and confidence intervals from small disconnected diallels

Received: 10 September 2004 / Accepted: 5 February 2005 / Published online: 8 March 2005

Abstract The distributions of genetic variance components and their ratios (heritability and type-B genetic correlation) from 105 pairs of six-parent disconnected half-diallels of a breeding population of loblolly pine (*Pinus taeda* L.) were examined. A series of simulations based on these estimates were carried out to study the coverage accuracy of confidence intervals based on the usual *t*-method and several other alternative methods. Genetic variance estimates fluctuated greatly from one experiment to another. Both general combining ability variance (σ^2_g) and specific combining ability variance (σ^2_s) had a large positive skewness. For σ^2_g and σ^2_s, a skewness-adjusted *t*-method proposed by Boos and Hughes-Oliver (Am Stat 54:121–128, 2000) provided better upper endpoint confidence intervals than *t*-intervals, whereas they were similar for the lower endpoint. Bootstrap BCa-intervals (Efron and Tibshirani, An introduction to the bootstrap. Chapman & Hall, London 436 p, 1993) and Hall's transformation methods (Zhou and Gao, Am Stat 54:100–104, 2000) had poor coverages. Coverage accuracy of Fieller's interval endpoint(J R Stat Soc Ser B 16:175–185, 1954) and *t*-interval endpoint were similar for both h^2 and r_B for sample sizes $n \leq 10$, but for $n = 30$ the Fieller's method is much better.

Communicated by D.B. Neale

F. Isik (✉) · B. Li
Department of Forestry and Environmental Resources,
North Carolina State University,
Campus Box 8002, Raleigh,
NC 27695, USA
E-mail: fisik@ncsu.edu
Tel.: +1-919-5155029
Fax: +1-919-5153169

D. D. Boos
Department of Statistics, North Carolina State University,
Campus Box 8002, Raleigh, NC 27695, USA

Introduction

Confidence intervals for genetic variance components and their ratios have been determined using a variety of techniques. Öfversten (1993) developed methods for obtaining exact *F*-tests of variance components for three unbalanced mixed linear models. The methods were extended and generalized by Christensen (1996) and Fayyad et al. (1996) to test the null hypothesis that a variance component is equal to zero. The standard errors of variance components and their ratios were also approximated with chi-square distributions using the expected mean squares (Lu and Graybill 1987). Assuming standard normal distribution properties, Knapp et al. (1987) defined parametric interval estimators for heritability, a ratio of additive genetic variance to phenotypic variance. They used *F*-distributions to establish exact confidence intervals for the point estimates. Knapp et al. (1989) used Monte Carlo simulation to estimate Jackknife intervals of family-mean heritability and expected selection response. Their methods were based on simulated data for a balanced one-factor random linear model and expected mean squares.

The literature cited here-above mainly focused on the precision of variance components and their functions for specific experiments. In practice, parents of a breeding population are grouped into smaller units, such as diallels, for mating and progeny testing. Diallel mating designs have been widely used in plant breeding programs to estimate genetic variances (Baker 1978). Usually, a small number of parents (4–12) are allocated to a diallel because of logistic difficulties and the cost-associated limitations of breeding and testing (Wilson et al. 1978; Cisar et al. 1982; Boyle 1987; Xiang et al. 2003). Then, each diallel group is field-tested and analyzed separately to obtain variances. Thus, variance components may vary dramatically from one experiment to another because of the "bottleneck" from random sampling a small number of parents (Wright 1985). The question we address: when there are

numerous empirical variance components estimated by maximum likelihood methods (ML or REML) and heritability estimates from different experiments (diallels), how can statistical inferences be made about the population?

We utilized a unique empirical data set from the North Carolina State University-Industry Cooperative Tree Improvement Program to make a statistical inference about a breeding population. For the second-cycle breeding of loblolly pine (*Pinus taeda* L.), 3,800 parents were grouped into six-parent disconnected half diallels. The parents were mated, and 15 crosses were produced for each diallel group (Li et al. 1999). Progeny of two diallels were tested at four locations. Two pairs of diallels were analyzed together to estimate variance components (Li et al. 1996). In this study, we utilized 105 variance components and their ratios from 105 pairs of diallels of the Coastal breeding population of loblolly pine. The data provided an excellent opportunity to examine the distribution properties of variance components from different experiments and to make statistical inference about them.

The objectives of the study reported here were to answer the following questions. (1) How are variance components and their ratios, estimated from individual diallels, distributed? (2) Depending on the distribution properties, what is the true coverage of *t*-confidence intervals? (3) How can we improve the true coverage of confidence intervals if there are significant departures from the normal distribution?

Materials and methods

Data

The North Carolina State University-Industry Cooperative Tree Improvement Program used a six-parent disconnected-half-diallel mating design (referred to as 'diallel' hereafter) to generate progenies for the second-generation breeding program of loblolly pine (Fig. 1). Parent trees of the Coastal breeding population were allocated to diallels for mating according to their flower production. For each diallel, six parents were mated and 15 crosses (full-sib families) were produced. Progenies of 30 crosses from two diallels were tested together using a randomized complete blocks design with six blocks. Each cross was represented by six-tree row plots in each six block. The experiments were replicated on four sites (Xiang et al. 2003). Each parent was represented by approximately 720 progeny across four sites. Height, volume, disease infection and stem straightness of each tree were recorded at age 6 years.

A pair of diallels tested at the same locations were analyzed together. A linear mixed model was fitted to the data (Xiang et al. 2003) to estimate variance components using individual tree measurements.

$$Y_{ijoklm} = \mu + T_i + B_{j(i)} + D_o + G_{k(o)} + G_{l(o)} + S_{kl(o)} + \mathrm{TG}_{ik(o)} + \mathrm{TG}_{il(o)} + \mathrm{TS}_{ikl(o)} + P_{ijokl} + E_{ijoklm} \quad (1)$$

where Y_{ijoklm} is the mth observation of the jth block for the klth cross in the oth diallel in the ith test; μ is the overall mean; T_i is the ith fixed test environment (location) effect; $i = 1$ to t; $B_{j(I)}$ is the fixed effect of the jth block within the ith test; $j = 1$ to b; D_o is the oth fixed diallel effect; $o = 1$; 2, $G_{k(o)}$ or $G_{l(o)}$ is the random general combining ability (GCA) effect of the kth female or the lth male within the oth diallel, assumed to be normally and independently distributed (NID) (0, σ^2_g); k, $l = 1$ to p, and $k < l$; $S_{kl(o)}$ is the random specific combining ability (SCA) effect of the kth and the lth parents within the oth diallel ($k \neq l$) ~NID (0, σ^2_s); $TG_{ik(o)}$ or $TG_{il(o)}$ is the random interaction effect of the ith test by the kth female or the lth male GCA ~NID (0, σ^2_{gt}); $TS_{ikl(o)}$ is the random interaction effect of the ith test by the klth SCA ~NID (0, σ^2_{st}); P_{ijokl} is the random plot effect for the klth cross within the oth diallel in the jth block ~NID (0, σ^2_p); $l = 1$ to n; E_{ijoklm} is the random within-plot error term ~NID (0, σ^2_e).

The model was analyzed with the SAS MIXED procedure by creating dummy variables for GCA effects (Xiang and Li 2001). The variance components of random effects were estimated using the default REML model fitting option. In the data analysis, it was assumed that six parents were randomly assigned into each diallel and that the parent trees in a diallel represent a random sample of the whole breeding population. Diallel effects were negligible and were dropped from the mixed model. Individual-tree narrow-sense heritability (h^2) for each diallel was estimated as the ratio of additive genetic to pheno-

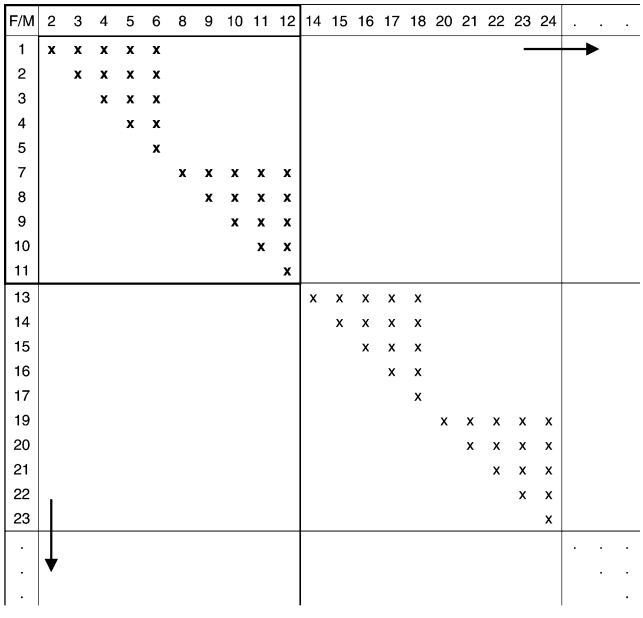

Fig. 1 The disconnected half-diallel mating design used for the second-generation loblolly pine improvement program in the southern USA. Two disconnected half-diallels (parents 1–12) were field-tested and analyzed together to estimate variance components

typic variance: $h^2 = 4\sigma^2_g/(2\sigma^2_g + \sigma^2_s + 2\sigma^2_{gt} + \sigma^2_{st} + \sigma^2_p + \sigma^2_e)$. Type-B genetic correlations (r_B) were calculated as $r_B = \sigma^2_g/(\sigma^2_g + \sigma^2_{gt})$, which is a ratio of family variance components over the sum of the family and genotype-by-environment interaction variance (Yamada 1962). This ratio has been used commonly to quantify the level of genotype-by-environment interaction in progeny trials (Burdon 1977). In this study, we used 105 variance components, heritabilities and type-B genetic correlations estimated from 105 pairs of diallels for 6-year height growth to make inferences about the Coastal breeding population of loblolly pine in the southern USA.

Statistical analysis

Since each diallel mating group was assumed to be a random sample of parents from the breeding population, the estimated variance components from the diallels may be viewed as a random sample derived from the same population. We also assumed that the average of a given variance component from all diallels in the population is a 'true' parameter of the population. The descriptive statistics of variance components were produced and appear in Table 1. General combining ability (σ^2_g) and specific combining ability (σ^2_s) genetic variances from 105 pairs of diallels had a large positive skewness. In plant breeding experiments, a 2.5% error rate for one-sided *t*-intervals is considered to be a critical cut-off. In other words, if the lower and upper intervals are appropriate, then two-sided error rates should be close to 5%. A series of simulations were carried out to determine the true error rate for one-sided *t*-intervals for each variance component. Since genetic variances departed from normality because of skewed distributions, the coverage accuracy of Hall's transformation method (Zhou and Gao 2000), bootstrap BCa-intervals (Efron and Tibshirani 1993), and the skewness-adjusted *t*-intervals (Boos and Hughes-Oliver 2000) were also estimated to compare with *t*-intervals. For simulation, random samples with replacement of *n* size were drawn from the population of a variance component, where $n = 2, 3, 4, \ldots, N$ ($N = 105$). For each *n*, the mean and 95% intervals based on the mean were calculated. The process was repeated 1,000 times for each sample size.

The intervals from Boos and Hughes-Oliver (2000) (hereafter to be called *BH*-intervals) may provide better coverage than nominal *t*-intervals for skewed data and for small sample size. Using the simulated data, we estimated the miss left (ML) and miss right (MR) error rates for one-sided intervals for each sample size *n*. When the confidence interval falls to the left of the "true" mean, the interval misses the "true" mean on the left and it is called "miss left". If the confidence interval falls to right of the "true" mean, then it is called a "miss right." The following two regression equations given by Boos and Hughes-Oliver (2000) were inverted to get skewness adjusted α values to be used in the *t*-interval formula:

$$\mathrm{MR} = \frac{\alpha}{2} + (-0.73 + 0.71\ \exp(-\alpha/2))\frac{\sqrt{\beta_1}(X)}{\sqrt{n}} \tag{2}$$

$$\mathrm{ML} = \frac{\alpha}{2} + (0.19 + 0.026\ \ln(\alpha/2))\frac{\sqrt{\beta_1}(X)}{\sqrt{n}}, \tag{3}$$

where $\alpha/2$ is the acceptable one-sided miss rate (0.025 in our case), $\sqrt{\beta_1}(X)$ is the estimated Pearson skewness coefficient of the distribution X, and n is the sample size. The proportion of times out of 1,000 that an interval contains the "true" value was calculated. The average length of *t*-intervals (L_t) and *BH*-intervals (L_{BH}) for a given sample size were estimated by taking the average over the 1,000 estimated intervals.

$$L_t = (2t_{n-1})\frac{\sum_i s \Big/ 1,000}{\sqrt{n}} \tag{4}$$

$$L_{\mathrm{BH}} = \frac{1}{1,000}\sum_i (U_i - L_i) \tag{5}$$

where, U_i and L_i are the upper endpoint and lower endpoint for the *i*th data set. i.e., the average length over

Table 1 Descriptive statistics of variance components, heritability and type-B genetic correlation for 6-year height for the Coastal breeding population of loblolly pine ($N = 105$)

Variable[a]	Mean	Standard deviation	Range	Coefficient of variation (%)	Skewness	Kurtosis
σ^2_g	0.23	0.17	0.0–1.11	74	2.09	6.75
σ^2_s	0.09	0.17	0.0–1.41	188	5.88	40.21
σ^2_{gt}	0.05	0.03	0.0–1.15	65	1.13	1.70
σ^2_{st}	0.02	0.03	0.0–1.14	155	1.89	3.03
σ^2_p	0.31	0.17	0.0–0.69	53	−0.40	−0.33
σ^2_e	3.69	1.04	1.96–6.49	28	1.10	0.72
h^2_b	0.19	0.12	0.0–0.62	60	1.34	2.36
r_{B_b}	0.79	0.17	0.0–1.00	21	−1.93	5.56

[a] σ^2_g is the general combining ability variance; σ^2_s, the specific combining ability variance; σ^2_{gt}, the general combining ability and environment interaction variance; σ^2_{st}, the specific combining ability and environment interaction variance; σ^2_p, the plot-to-plot variance; σ^2_e, the residual variance; h^2_b; the biased narrow-sense individual-tree heritability; $r_{B\text{-}b}$, the biased type-B genetic correlation

the simulated samples, s is sample standard deviation averaged over 1,000 samples and n is the sample size. The narrower the length the more accurate it is for a given sample size.

The heritability (h^2) and the type-B genetic correlation (r_B) estimated from a pair of diallels are biased because they are not random samples derived from the same population but are based on ratios of specific variance components. Since h^2 and r_B are ratios of variance components, the "true" population values in our simulation study are based on the means of the 105 estimated variance components:

$$h^2_{-t} = \frac{4\sigma^2_{g(m)}}{2\sigma^2_{g(m)} + \sigma^2_{s(m)} + 2\sigma^2_{gt(m)} + \sigma^2_{st(m)} + \sigma^2_{p(m)} + \sigma^2_{e(m)}} \quad (6)$$

$$r_{B-t} = \frac{\sigma^2_{g(m)}}{\sigma^2_{g(m)} + \sigma^2_{gt(m)}}, \quad (7)$$

where the subscript "m" means that we have averaged over the 105 estimates. Because heritability and type-B genetic correlation are ratios of variance components, they may have different distribution properties. To examine the distributions of h^2 and r_B, 1,000 random samples with replacement were carried out for each n sample size, and approximately unbiased estimates for h^2_{-unb} and $r_{B\text{-}unb}$ were obtained by using the above formulas (Eqs. 6 and 7) using sample averages of variance component estimates. The biases for h^2 and r_B were estimated as the difference between the "true" values and the values based on n sample means. As the sample size increases, the biases for both h^2 and r_B are expected to approach zero.

Since the estimates h^2_{-unb} and r_{B-unb} are the ratio of the sample means, we used Fieller's (1954) method to estimate their confidence intervals and compared these to nominal t-intervals. Fieller's method is a clever way to get a confidence interval for the ratio of two means for random pairs. The idea is to define a new variable $Z = Y\text{-}(\mu_y/\mu_x)X$, using the mean μ_y of the numerator variable Y and the mean μ_x of the denominator variable X. The Z variable has mean 0 and a variance that is estimated from the sample Z values. Then, the Fieller's confidence interval finds all values of μ_y/μ_x such that $t^2(Z) < t^2_{(0.975,n\text{-}1)}$, where $t(Z)$ is the t-statistic on the Z-variable and $t^2_{(0.975,n\text{-}1)}$ is the usual t-interval cut-off.

For planning purposes, how large should n be for estimating σ^2_g, σ^2_s, h^2 and r_B? The sample size for a desired length of interval can be obtained from the standard formula, $n = [t_{\alpha 2,n\text{-}1} * s/d]^2$, where $\alpha/2$ is the error rate (typically 0.025), s is the sample standard deviation (which must be guessed), and d is the margin of error. Recall that the t-interval has length of $2t_{\alpha 2,n\text{-}1}\ s/\sqrt{n}$ or margin of error $t_{\alpha 2,n\text{-}1}\ s/\sqrt{n}$. Later we give a numerical example for specific standard errors of heritability and type-B genetic correlation.

Results

Descriptive statistics

Variance components and their ratios fluctuated considerably among experiments within the breeding population (Table 1). General combining ability variance (σ^2_g) ranged from 0.0 to 1.11, whereas narrow-sense heritability (h^2) ranged from 0.0 to 0.62. Type-B genetic correlation (r_B) fluctuated between the lower (0.0) and upper (1.0) theoretical limits. High variation in variance components among diallels is reflected in high coefficients of variation (CV): r_B had the smallest CV (21%), whereas specific combining ability variance (σ^2_s) had the highest (188%) CV. σ^2_g and σ^2_s had a positive skewness of 2.09 and 5.88, respectively, indicating a significant departure from normality. The distribution of σ^2_g and σ^2_s estimates also had large kurtosis values. Plot-to-plot (σ^2_p) and within-plot (σ^2_e) error variances had the lowest skewness and kurtosis among the variance components.

Interval coverage of genetic variances (simulation results)

The one-sided coverages of t-intervals as a function of sample size for σ^2_g and σ^2_s from simulations are presented in Fig. 2, and numeric results of t- and BH-intervals for four arbitrary selected sample size (5, 10, 30, and 100) are given in Table 2. ML and MR error rates are not equal for both genetic variances, with more intervals missing left (ML) than right (MR). For example, at $n = 30$, a nominal 95% t-interval for σ^2_s would fall 25% of the time to the left of σ^2_s, whereas it would be about less than 1% times to the right of σ^2_s (Table 2). Although σ^2_g had similar MR error rates to σ^2_s, the ML error rates of σ^2_g were substantially smaller than the ML error rates of σ^2_s. At $n = 30$, the t-interval for σ^2_g would fall 9% of the time to left of σ^2_g. The ML and the MR error rates are expected to converge to 0.025 as the sample size increases. Although the trend of error rates approached 0.025, we did not see convergence for σ^2_s genetic variance, even for large sample sizes (Fig. 2). However, we observed convergence for σ^2_g, particularly for BH-intervals at $n = 100$.

Because of high ML error rates, the true coverage of nominal 1-α intervals was much lower than desired. For example, at $n = 30$, the true coverage of 95% t-intervals for σ^2_s would be only 75%. In another words, σ^2_s will be not be in the nominal 95% t-interval for about 25% of the time. The total coverage of t-intervals for σ^2_g was better than the total coverage for σ^2_s intervals. For the same sample size, the true coverage of nominal 95% interval for σ^2_g would be about 91%. This total coverage is only 4% smaller than the nominal 95% interval coverage. As the sample size increases, the accuracy of t-intervals for both genetic variances increased. For example at $n = 100$, the true coverage of nominal 1-α

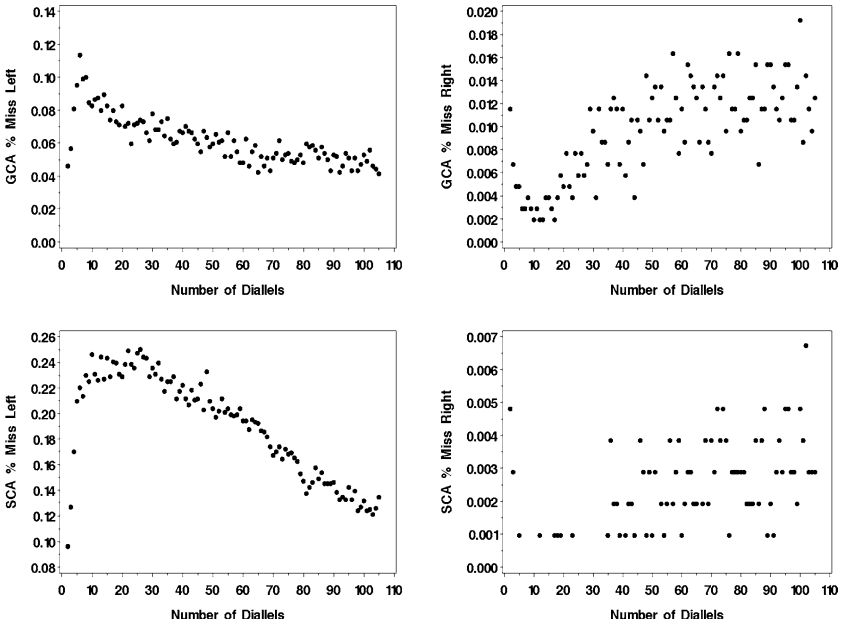

Fig. 2 The one-sided coverage of *t*-intervals as a function of sample size *n* for genetic variances σ^2_g and σ^2_s. MR and ML proportions are expected to convergence to 0.025 as the sample size further increases. We can not see this convergence for either σ^2_g or σ^2_s, even for $k > 100$ because of the skewness in the data

intervals was approximately the same as the desired level of 95% interval. The interval lengths of both genetic variances were also considerably smaller (0.07) than the interval lengths at $n = 30$.

The *BH*-intervals improved the coverage probability for σ^2_s, whereas the improvement over the *t*-intervals for σ^2_g was modest (Table 2). Bootstrap BCa and Hall's transformed intervals had poorer coverages than *BH*-intervals and were thus not reported. The true coverage of *BH* nominal 95% *t*-intervals for sample size $n = 30$ was around 80% for σ^2_s and 93% for σ^2_g. The improvement over the *t*-intervals was 5% for σ^2_s and 2% for σ^2_g for the same sample size. The total coverage of *BH*-intervals for σ^2_g is only 2% smaller than 95% interval coverage. For a smaller sample size, the *BH*-intervals had a greater interval length than the *t*-intervals. However, as the sample size increases, the interval length difference between *t* and *BH* diminished although their improvement of total coverage remained about the same.

Interval coverage of h^2 and r_B

Nominal 95% *t*-intervals and Fieller's intervals, the one-sided error rates and intervals length for heritability and type-B genetic correlation are presented in Table 3. In general, the Fieller's method had better coverage than *t*-intervals for h^2 and r_B. The two-sided coverage of nominal 95% *t*-intervals for h^2 did not change much as the sample size increased. The coverage of *t*-intervals for h^2 was as good as Fiellers' intervals for all three sample sizes. For example, at $n = 30$, the total coverage of Fieller's and *t*-intervals were 0.92 and 0.89, respectively, and the improvement of Fieller's over the *t*-interval was only 3%. However, the Fieller's method had much better coverage of the 'true' r_B mean than did the *t*-intervals for large samples, i.e., 0.95 vs. 0.82 for $n = 30$ and 0.94 vs. 0.31 for $n = 100$. These low coverages are due to the fact that the sample mean of the r_B values is biased for estimating the true value given in Eq. 7. This bias becomes more apparent as the sample size is increased to $n = 100$.

Discussion

Variation

Partitioning of parents into small groups of diallels for breeding is commonly employed for monoecious woody

Table 2 Simulation results for σ_g^2 and σ_s^2 genetic variances. MR, ML proportions, total coverage and the length of coverage for nominal 95% *t*-intervals and *BH*-intervals for sample sizes $n = 5$, 10, 30, and 100. Entries are based on 1,000 Monte Carlo replications. The standard errors of the miss rates, coverage estimates, and interval lengths are all approximately 0.01

Variable	*n*	Miss right (MR)		Miss left (ML)		Total coverage		Interval length	
		BH	*t*	BH	*t*	BH	*t*	BH	*t*
σ_g^2	5	0.01	**0.01**	0.11	**0.11**	0.89	**0.88**	0.43	**0.37**
	10	0.01	**0.01**	0.07	**0.09**	0.92	**0.91**	0.24	**0.22**
	30	0.01	**0.01**	0.06	**0.08**	0.93	**0.91**	0.13	**0.12**
	100	0.02	**0.01**	0.03	**0.05**	0.96	**0.95**	0.07	**0.07**
σ_s^2	5	0.00	**0.00**	0.19	**0.22**	0.81	**0.78**	0.31	**0.25**
	10	0.00	**0.00**	0.17	**0.23**	0.83	**0.78**	0.19	**0.17**
	30	0.00	**0.00**	0.19	**0.25**	0.80	**0.75**	0.11	**0.11**
	100	0.01	**0.00**	0.11	**0.15**	0.88	**0.85**	0.07	**0.06**

Table 3 Simulation results for nominal 95% Fieller and *t*-confidence intervals for heritability (h^2) and type-B genetic correlation (r_B) based on sample sizes $n = 5$, 10, 30, and 100. Entries are based on 1,000 Monte Carlo replications. The standard errors of the miss rates, coverage estimates, and interval lengths are all about 0.01

Variable	*n*	Miss right (MR)		Miss left (ML)		Total coverage		Interval length	
		Fieller	*t*	Fieller	*t*	Fieller	*t*	Fieller	*t*
h^2	5	0.00	**0.00**	0.09	**0.10**	0.91	**0.90**	0.27	**0.25**
	10	0.00	**0.00**	0.09	**0.10**	0.91	**0.90**	0.17	**0.16**
	30	0.01	**0.00**	0.07	**0.11**	0.92	**0.89**	0.09	**0.09**
	100	0.01	**0.00**	0.04	**0.12**	0.95	**0.88**	0.05	**0.05**
r_B	5	0.02	**0.03**	0.02	**0.02**	0.96	**0.96**	0.49	**0.35**
	10	0.02	**0.02**	0.03	**0.04**	0.96	**0.94**	0.20	**0.22**
	30	0.02	**0.00**	0.04	**0.18**	0.95	**0.82**	0.10	**0.12**
	100	0.02	**0.00**	0.03	**0.69**	0.94	**0.31**	0.05	**0.07**

perennials such as conifers (Yanchuk 1996; Yeh and Heaman 1987; Johnson and King 1998). Despite the several advantages in using diallels, genetic sampling of parents from a large population for a given diallel is subject to random genetic drift. Genetic differences among diallel groups may cause significant variation in estimates of genetic variances (Hill 1985). In our study, over 100 variance components estimates of six-parent disconnected half-diallels were used to make inferences about a breeding population. Thus, the observed large variation of variance components among individual diallels was not surprising. Similar large differences in genetic variance estimates between diallels were also observed in Douglas-fir (*Pseudotsuga menziessi* var. *menziessii*) (Yeh and Heaman 1987) and in another population of loblolly pine (Balocchi et al. 1993).

In this study, among the variance components studied, σ^2_s had a wider range and higher coefficient of variation than any other estimates. This is partly due to the nature of σ^2_s estimation from the diallels. The estimation of σ^2_s is based on specific crosses of one parent and is subject to larger standard errors (Yanchuk 1996). On the other hand, σ^2_g is estimated from many more progenies in a diallel with a smaller variance, as shown by the smaller coefficient of variation. SCA by environment interaction variance (σ^2_{st}) followed σ^2_s in having a large variance. The error variance components (σ^2_p, σ^2_e) can be approximated with a normal distribution as they had a smaller skewness and kurtosis than the genetic variances.

In general, the pattern of distribution of h^2 followed the distribution pattern of σ^2_g. This is expected because h^2 is mainly the product of σ^2_g. However, h^2 had a smaller kurtosis value than σ^2_g. Although r_B is also the product of σ^2_g, we observed a negative skewness for r_B. This is mainly due to the high frequency of large σ^2_{gt} values and sporadic zero σ^2_{gt} values. The kurtosis of r_B was similar to the σ^2_g value.

Error rates and the coverage of the intervals

The dependence of coverage probabilities on the skewness have been covered extensively in the literature (Chaffin and Rhiel 1993; Chen 1995; Boos and Hughes-Oliver 2000). In the present study, very high ML error rates of intervals observed for σ^2_g and σ^2_s were mainly the effect of a large skewness in the distributions of these two genetic variances. Among the variance components studied, σ^2_s had a larger skewness and greater standard deviation than any other variance components. Thus, the high error rates of two-sided intervals for σ^2_s are not surprising. If breeders basically average σ^2_s variances from different experiments to obtain a mean with a nominal 95% confidence interval, then the 'true' σ^2_s mean would likely be out of that interval at least 25% of the time. We suggest using the *BH* skewness-adjusted *t*-intervals for variance components such as σ^2_s and σ^2_g. Although the *BH* intervals were longer than the *t*-intervals, the improvement can be substantial for smaller samples. The length of coverage of *BH*-intervals approaches the length of nominal *t*-intervals as the sample size increases. The drawback of the *BH* intervals is that the right-side coverages are not perfect. In fact, the *t* and *BH* right side error rates are similar, and they converge slowly to 0.025, but even $n = 100$ is not large enough to see the convergence.

Knapp et al. (1987, 1989) found similarities between nominal interval estimates of point heritability and variance components. In their study, confidence interval widths of heritability decreased slightly as sample size increased. Their heritability estimates and precision statistics were based an analysis of variance and expected mean squares but not maximum likelihood estimation procedures. When the error rates were taken into consideration, Fieller's intervals for h^2 and r_B were not superior to *t*-intervals, particularly for small sample size *n*. However, one needs to consider the length of the interval along with the coverage. Fieller's intervals were superior to *t*-intervals as shown by narrower interval lengths for both h^2 and r_B. The decrease in length of the Fieller's intervals and *t*-intervals in our study was clearly pronounced as the sample size increased. We observed a sharp drop in the coverage of the *t*-interval for r_B at $n = 30$ and $n = 100$. This is due to the bias of averaging r_B values.

Accuracy of the population statistics and the sample size *n*

Plant breeders may ask about the sample size (number of estimates) to obtain a reliable sample statistics that is a

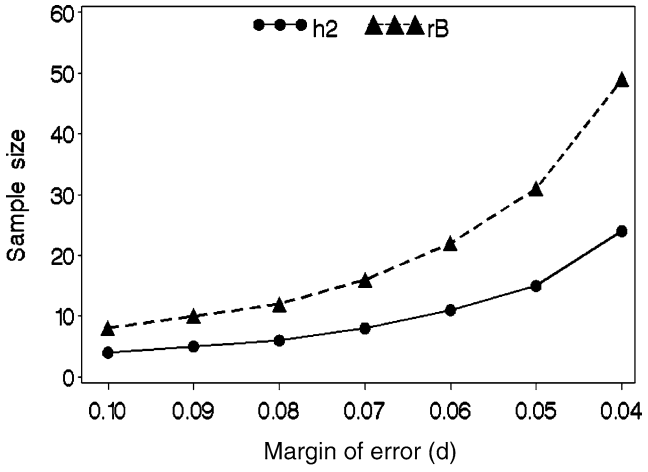

Fig. 3 Sample size as a function of error margins for heritability (h^2) and type-B genetic correlation (r_B) for a specific situation. The sample size required for any given error of margins were based on the assumptions that the population standard deviation for h^2 and r_B are 0.12 and 0.17, respectively. Larger sample sizes are required for r_B than for h^2 for the same error margins because of a greater population standard deviation of r_B

representative of the population? The answer to the above question varies depending on the accuracy sought. For the Coastal breeding population of loblolly pine, let us say the true population standard deviation of h^2 is $\sigma = 0.12$. If we are satisfied with a $d = 0.05$ margin of error, then for a 95% confidence level we need to have at least $n = (t \times \sigma/d)^2 = (1.96 \times 0.12/0.05)^2 = 22$ diallels, where 1.96 is the Z-value for the 95% level. The difference between a Z -value and a t-value is not very large. We used the Z-value 1.96 as an approximation to the t-value because of simplicity. For the same margin of error, the number of diallels required for r_B would be double (about 44), because the population standard deviation for r_B (0.17) is greater than h^2. If $d = 0.10$ is an acceptable margin of error, then the sample size needed for h^2 and r_B would be 6 and 11 diallels, respectively. For this specific population, assuming that the population standard deviations for h^2 and r_B are 0.12 and 0.17, respectively, the sample size as a function of a series of error margins is given in Fig. 3. For planning proposes, population standard deviations of h^2 and r_B can be used to estimate an approximate sample size. Here, we used t-intervals to calculate approximate sample size for h^2 and r_B because the actual coverage of t-intervals was as good as Fieller's intervals for small samples.

Conclusions

The genetic variance component estimates σ^2_g and σ^2_s from disconnected diallels for 6-year height growth had a considerable large positive skewness. The results are based on empirical estimates of variance components from a breeding population of loblolly pine. Because of this skewness, the true coverage of nominal 95% t-intervals for genetic variances would likely be less than the nominal coverage. The *BH*-intervals adjusted for skewness provided better coverage than regular t-intervals. However, *BH*-intervals are not optimal. Their miss right error rates were as poor as t-intervals. In contrast to genetic variances, h^2 and r_B had smaller skewness, but because they are ratios of variance components, it is wise to use Fieller's intervals over the t-intervals for h^2 and r_B although regular t-intervals are not much worse in small samples. When statistics (mean, standard deviation, and intervals) are estimated from a sample of genetic variances and their functions, the distribution properties should be taken into account for more reliable inferences.

References

Baker RJ (1978) Issues in diallel analysis. Crop Sci 18:533–536

Balocchi CE, Bridgwater FE, Zobel BJ, Jahromi S (1993) Age trends in genetic parameters for tree height in non-selected population of loblolly pine. For Sci 39:231–251

Boos DD, Hughes-Oliver J (2000) How large does n have to be for Z and t intervals? Am Stat 54:121–128

Boyle TA (1987) A diallel cross in black spruce. Genome 29:180–189

Burdon RD (1977) Genetic correlation as a concept for studying genotype-environment interaction in forest tree breeding. Silvae Genet 26:168–175

Chaffin WW, Rhiel GS (1993) The effect of skewness and kurtosis on the one-sample t test and the impact of knowledge of the population standard deviation. J Stat Comput Simul 46:79–90

Chen L (1995) Testing the mean of skewed distributions. J Am Stat Assoc 90:767–772

Christensen R (1996) Exact tests for variance components. Biometrics 52:309–314

Cisar G, Brown CM, Jedlinksi H (1982) Diallel analyses for tolerance in winter wheat to the barley yellow dwarf virus. Crop Sci 22:328–333

Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman and Hall, London

Fayyad R, Graybill FA, Burdick RK (1996) A note on exact tests for variance components in unbalanced random and mixed linear models. Biometrics 52:306–308

Fieller EC (1954) Symposium on interval estimation: some problems in interval estimation. J R Stat Soc Ser B 16:175–185

Hill WG (1985) Effect of population size on response to short term and long term selection. J Anim Breeding Genet 102:161–173

Johnson GR, King JN (1998) An analysis of half-diallel mating designs. Silvae Genet 47:74–79

Knapp SJ, Ross WM, Stroup WW (1987) Precision of genetic variance and heritability estimates from sorghum populations. Crop Sci 27:265–268

Knapp SJ, Bridges WC Jr, Yang MH (1989) Nonparametric confidence interval estimators for heritability and expected selection response. Genetics 121:891–898

Li B, McKeand SE, Weir RJ (1996) Genetic parameter estimates and selection efficiency for the loblolly pine breeding in the south-eastern US. In: Tree improvement for sustainable tropical forestry. In: Dieters MJ, Matherson DG, Nikles DG, Harwood CE, Walker SM (eds) Proc QFRI-IUFRO Conf. Caloundra, Queensland, Australia, pp 164–168

Li B, Weir R, Hatcher A, McKeand SE (1999) Third cycle loblolly pine breeding plan. Technical Report 99–2. NC State University-Industry Cooperative Tree Improvement Program, 16 p

Lu TFC, Graybill FA (1987) Confidence intervals on the ratio of expected mean squares theta-1 plus d-theta-2-theta-3. Biometrics 43:535–544

Öfversten FA (1993) Matrices with applications in statistics, 2nd edn. Belmont, Calif.
Wilson ND, Weibel DE, Mcnew RW (1978) Diallel analysis of grain yield, percent protein, and yield in gain sorghum. Crop Sci 18:491–494
Wright AJ (1985) Diallel designs, analyses, and reference populations. Heredity 54:307–311
Xiang B, Li B (2001) A new mixed analytical method for genetic analysis of diallel data. Can J For Res 31:2252–2259
Xiang B, Li B, Isik F (2003) Time trend of genetic parameter estimates in growth traits of *Pinus Taeda* L. Silvae Genet 52:114–121
Yamada Y (1962) Genotype by environment interaction and genetic correlation of the same trait under different environments. Jpn J Genet 6:498–509
Yanchuk AD (1996) General and specific combining ability from disconnected partial diallels of coastal Douglas fir. Silvae Genet 45:37–44
Yeh FC, Heaman JC (1987) Estimating genetic parameters of height growth in seven year-old coastal Douglas-fir from disconnected diallels. For Sci 33:946–957
Zhou XH, Gao S (2000) One-sided confidence intervals for means of positively skewed distributions. Am Stat 54:100–104