south interior of British Columbia where transfer should be guided by local microclimatic conditions. The south interior region also presents the biggest challenge in terms of paper birch gene conservation because of the large differentiation among the populations within this region and difficulties in relating the variation to the macroclimatic variables as evidenced by the variation in frost hardiness.

Acknowledgements

This study was supported by Forest Renewal British Columbia Grant number HQ96060-RE made to Y.A.E.

References

BENOWICZ, A., GUY, R. D. and EL-KASSABY, Y. A.: Geographic pattern of genetic variation in photosynthetic capacity and growth in two hardwood species from British Columbia. Oecologia **123**: 168–174 (2000). — BEVINGTON, J.: Geographic differences in the seed germination of paper birch (*Betula papyrifera*). Amer. J. Bot. **73**: 564–573 (1986). — CAMPBELL, R. K. and RITLAND, S. M.: Regulation of seed germination timing by moist chilling in western hemlock. New Phytol. **92**: 173–182 (1982). — CHING, T. M.: Activation of germination in Douglas-fir seed by hydrogen peroxide. Plant Physiol. **34**: 557–563 (1959). — COMEAU, P. G., and THOMAS, K. D. (*Editors*).: Silviculture of Temperate and Boreal Broadleaf-Conifer Mixtures. B.C. Ministry of Forests, Victoria, B. C. Land Manag. Handb. 36 (1996). — CZABATOR, F. J.: Germination value: an index combining speed and completeness of pine seed germination. For. Sci. **8**: 386–396 (1962). — DAVIDSON, R. H., EDWARDS, D. G. W., SZIKLAI, O. and EL-KASSABY, Y. A.: Genetic variation in germination parameters among populations of Pacific silver Fir. Silvae Genet. **45**: 165–171 (1996). — DEANS, J. D. and HARVEY, F. J.: Frost hardiness of 16 European provenances of sessile oak growing in Scotland. Forestry (Oxf.): **69**: 5–11 (1996). — EL-KASSABY, Y. A., EDWARDS, D. G. W. and TAYLOR, D. W.: Genetic control of germination parameters in Douglas-fir and its importance for domestication. Silvae Genet. **41**: 48–54 (1992). — ELLIOTT, D. M. and TAYLOR, I. E. P.: Germination of red alder (*Alnus rubra*) seed from several locations in its natural range. Can. J. For. Res. **11**: 517–521 (1981). — ERIKSSON, G. and JONSSON, A.: A review of the genetics of *Betula*. Scand. J. For. Res. **1**: 421–434 (1986). — FALCONER, D. S.: Introduction to Quantitative Genetics. Longman Inc., New York (1989). — FARQUHAR, G. D. and SHARKEY, T. D.: Stomatal conductance and photosynthesis. Ann. Rev. Plant Physiol. **33**: 317–45 (1982). — FOWELLS, H. A.: Silvics of Forest Trees of the USA. U.S. Dep. Agric., For. Serv., Agric. Handb. No 271 (1965). — GLERUM, C.: Frost hardiness in coniferous seedlings: principles and applications. *In:* Evaluating Seedling Quality : Principles, Procedures and Abilities of Major Tests. *Edited* by DURYEA, M. L. Forest Research Laboratory, Oregon State University, Corvallis, Oreg. pp. 107–123 (1985). — GUTTERMAN, Y.: Maternal effects on seeds during development. *In:* Seeds: The Ecology of Regeneration in Plant Communities. *Edited by* FENNER, M. CAB International, Wallingford, UK. pp. 27–59 (1992). — International Seed Testing Association: International rules for seed testing.: Seed Sci. Technol. **13**: 299–513 (1985). — LUDLOW, M. M. and JARVIS, P. G.: Photosynthesis in Sitka spruce (*Picea sitchensis* (BONG.) CARR.). I. General characteristics. J. Appl. Ecol. **8**: 925–953 (1971). — MASSIE, M. R. C., PETERSON, E. B., PETERSON, N. M. and ENNS, K. A.: An Assessment of the Strategic Importance of the Hardwood Resource in British Columbia. B. C. Ministry of Forests, Victoria, B. C. FRDA Rep. 221 (1994.). — MORGENSTERN, E. K.: Genetic Variation in seedlings of *Picea mariana* (MILL.) BSP. I. Correlation with ecological factors. Silvae Genet. **18**: 151–61 (1969). — OWENS, J. N.: Flowering and seed set. In Physiology of Trees. Edited by RAGHAVENDRA, A. S. John Wiley & Sons, New York. pp. 247–271 (1991). — RADWAN, M. A. and DEBELL, D. S.: Germination of Red Alder Seed. U.S. Dep. Agric., For. Serv. Pac. NW For. Range Exp. Sta., Portland, Oregon. Research Note No. PNW–370 (1981). — REHFELDT, G. E., YING, C. C., SPITTLEHOUSE, D. L. and HAMILTON, D. A.: Genetic responses to climate for *Pinus contorta* in British Columbia: niche breadth, climate change, and reforestation. Ecol. Monogr. **69**: 375–407 (1999). — REYES, O., CASAL, M. and TRABAUD, L.: The influence of population, fire and time of dissemination on the germination of *Betula pendula* seeds. Plant Ecol. **133**: 201–208 (1997). — SAKAI, A. and LARCHER, W.: Frost Survival of Plants: Responses and Adaptation to Freezing Stress. Springer-Verlag, Berlin (1987). — SAS.: SAS/STAT® User's Guide, Release 6.03 Edition. SAS Institute Inc., Cary, N.C. (1988). — SIMARD, S. and VYSE, A.: Ecology and Management of Paper Birch and Black Cottonwood in Southern British Columbia. Ministry of Forests, Victoria, B.C. Land Manag. Rep. **75** (1992). — THOMSON, A. J. and EL-KASSABY, Y. A: Interpretation of seed-germination parameters. New For. **7**: 123–132 (1993). — VIHERA-AARNIO, A. and VELLING, P.: Growth and stem quality of mature birches in a combined species and progeny trial. Silva Fenn. **33**: 225–234 (1999). VON CAEMMERER, S. and FARQUHAR, G. D.: Some relationships between the biochemistry of photosynthesis and gas exchange of leaves. Planta **153**: 376–387 (1981). — WANG, J. R., HAWKINS, C. D. B. and LETCHFORD, T.: Relative growth rate and biomass allocation of paper birch (*Betula papyrifera*) populations under different soil moisture and nutrient regimes. Can. J. For. Res. **28**: 44–55 (1998a). — WANG, J. R., HAWKINS, C. D. B. and LETCHFORD, T.: Photosynthesis, water and nitrogen use efficiencies of four paper birch (*Betula papyrifera*) populations grown under different soil moisture and nutrient regimes. For. Ecol. Manage. **112**: 233–244 (1998b).

Comparison of Multivariate and Univariate Methods for the Estimation of Type B Genetic Correlations

By P. LU[1)2)], D. A. HUBER[1)] and T. L. WHITE[1)]

(Received 14th June 2000)

Summary

Univariate and multivariate statistical methods for type B genetic correlation estimation are numerically compared based on computer simulated half-sib forest genetic experimental data. Multivariate methods have demonstrated desirable properties in effectively handling unbalanced data associated with heterogeneous variances across environments and yielded empirically unbiased estimates of type B genetic correlations for various data structures. Constraining estimates of type B genetic correlations within theoretical parameter space when using multivariate methods helps improve estimation accuracy and narrow the confidence interval. While some univariate methods can also produce unbiased estimates of type B genetic correlations for unbalanced data with heterogeneous variances, multivariate methods tend to estimate type B genetic correla-

[1)] School of Forest Resources and Conservation, 118 Newins-Ziegler Hall, University of Florida, Gainesville, FL 32611, USA

[2)] Ontario Forest Research Institute, Ministry of Natural Resources, 1235 Queen Street East, Sault Ste. Marie, ON P6A 2E5, Canada
Email: pengxin.lu@mnr.gov.on.ca
dah@harm.sfrc.ufl.edu
tlw@harm.sfrc.ufl.edu

tion more stable and efficiently by simultaneously using all information from multi-site data and giving estimates for all paired environments.

Key words: Genetic correlation, multivariate method, estimation bias, unbalanced data, heterogeneous variance.

Introduction

In quantitative forest genetic data analysis, type B genetic correlations (BURDON, 1977) have routinely been estimated using the univariate methods of YAMADA (1962) and BURDON (1977) (BURDON, 1977; JOHNSON and BURDON, 1990; WOOLASTON et al., 1991; HODGE and WHITE, 1992; HODGE and PURNELL, 1993; ADAMS ET AL., 1994; DIETERS et al., 1995; DIETERS, 1996; PSAWARAYI et al., 1997). For pairs of genetic tests, these methods first estimate genetic variances and covariances using univariate linear models in one or/and two environments separately and then calculate genetic correlations according to the described procedures (YAMADA, 1962; BURDON, 1977). Although these univariate methods minimize computational demands and facilitate the application of standard computer software, it has been suggested that estimates of genetic correlations from these univariate methods are less satisfactory for some data structures (FERNANDO et al., 1984; DUTILLEUL and CARRIÈRE, 1998; LU et al., 1999).

One undesirable property of these univariate methods is their inability to yield unbiased estimates of type B genetic correlation for unbalanced experimental data accompanied with heterogeneous variances across environments. For example, theoretical considerations and empirical evidence have indicated that biases may occur with estimates of type B genetic correlation from YAMADA's (1962) methods when data are severely unbalanced and variances are heterogeneous between experiments (FERNANDO et al., 1984; ITO and YAMADA, 1990; DUTILLEUL and CARRIÈRE, 1998). Severe biases for estimates of type B genetic correlations were also found with BURDON's method when mortality caused data imbalance among genetic groups within a genetic test (LU et al., 1999). Although improvement to univariate methods may be made (DUTILLEUL and CARRIÈRE, 1998; LU et al., 1999) so that unbiased estimates of type B genetic correlations are still obtainable for unbalanced data with heterogeneous variances, some of the procedures become computationally less convenient.

Another concern with univariate methods is that type B genetic correlation estimates are frequently out of the theoretical parameter space and results are difficult to apply in practice. Both empirical evidence (KOOTS and GIBSON, 1996) and simulation study (LU, 1999) have indicated that the frequency of out-of-bound estimates increases with a decrease in true population heritabilities. Out-of-bound estimates often occur when near-to-zero estimates of genetic variance are obtained in one or two of the concerned environments, while the estimates of genetic covariance are relatively large. Although out-of-bounds estimates are primarily attributable to sampling errors of genetic variances and covariances, they may also be related to the fact that genetic variance-covariances are not estimated from a closed system with these univariate methods.

A third problem facing these commonly used univariate methods of type B genetic correlation estimation is how to properly account for genetic relatedness and inbreeding among and within genetic groups. Since the possibility of genetic relatedness among genetic entries (especially among full-sib families) increases as tree improvement programs move into advanced generations (WHITE et al., 1993; BORRALHO and DUTKOWSKI, 1998), the assumption that genetic entries are independent random samples of a large population would be violated if genetic relatedness does exist. Failure to account for genetic relatedness and inbreeding by using pedigree information may result in loss of useful information from genetic tests and causes inaccurate estimates of genetic variance and covariances, and potential bias in estimates of type B genetic correlations (NOTTER and DIAZ, 1993).

Multivariate methods can estimate genetic variances and covariances simultaneously using the restricted maximum likelihood (REML) approach with an iterative procedure (PATTERSON and THOMPSON, 1971; Schaeffer and Wilton 1978). For multivariate methods, measurements from different environments are treated as different traits with different variance and covariance structures. Consequently, the problem of heterogeneous variances facing univariate methods is properly addressed. Evidence indicates that REML approach is generally more desirable than ANOVA (analysis of variance) methods for handling unbalanced data for the purpose of variance component estimation (SEARLE et al., 1992; HUBER et al., 1994). In addition, multivariate methods can apply constraints to estimates of genetic variances and covariances so that estimates of genetic correlation stay within the theoretical parameter space (BOLDMAN et al., 1995). Furthermore, many multivariate REML approaches are specifically designed to make full use of pedigree information and can analyze data based on individual tree models so that genetic relatedness among genetic entries is properly accounted for in the process of variance component estimation (BOLDMAN et al., 1995; GILMOUR et al., 1997).

Despite several potential advantages of multivariate approaches in estimating type B genetic correlations, uncertainty remains as to whether constrained multivariate procedures can yield unbiased estimates. It is also unclear how many environments should be used in a constrained system to enhance the quality of estimates. Since theoretical prove of these properties is difficult if not impossible, the purpose of this study is to numerically compare several multivariate and univariate approaches for type B genetic correlation estimation using computer simulated forest genetic data. Specifically, we examine these estimation methods in terms of empirical bias and precision, as well as the distribution properties of the estimates under different true genetic parameter settings.

Material and Methods

Data Generation

Computer simulated data were used in numerical comparisons because underlining true type B genetic correlation is known for each data set and can be used to evaluate the qualities of estimates. Data were generated based on a randomized complete block (RCB) design with one tree per family per plot (i.e., single-tree plots) which is recommended by several studies in forest genetic testing (LAMBETH and GLADSTONE, 1983; LOO-DINKINS and TAUER, 1987; LOO-DINKINS et al., 1990; WHITE, 1996). Genetic structures of the data were simulated based on half-sib families created from a polymix mating design with 120 female parents.

For the field experimental designs, it was assumed that these 120 half-sib families were tested over 4 environments, each having 90 families and 20 blocks. It was further assumed that there were 60 half-sib families in common for any paired progeny tests, but there was no family in common across the 4 environments. The above experimental layout is intended to mimic forest genetic tests with scenarios of: 1) a large number of genetic entries need to be tested over environments but with constrained block size (around 0.1 ha.) (WHITE, 1996) in order to maintain environmental homogeneity within a block; and 2)

partially genetically disconnected tests due to the availability of seedlings and other reasons.

The linear model used in data generation across 4 testing environments is given in matrix notation as:

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_4 \end{bmatrix} = \begin{bmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_4 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_4 \end{bmatrix} + \begin{bmatrix} Z_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Z_4 \end{bmatrix} \begin{bmatrix} g_1 \\ \vdots \\ g_4 \end{bmatrix} + \begin{bmatrix} Z_{m1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Z_{m4} \end{bmatrix} \begin{bmatrix} g_{m1} \\ \vdots \\ g_{m4} \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_4 \end{bmatrix} \quad (1)$$

where y_i is a n_i *x 1* vector of phenotypic observations in environment *i, i=1,...4;* $\mathbf{X_i}$ is the incidence matrix relating to fixed effects (vector β_i) in environment *i;* $\mathbf{Z_i}$ is the incidence matrix relating to female parental genetic effects (vector $\boldsymbol{g_i}$) tested in environment *i;* $\mathbf{Z_{mi}}$ is the incidence matrix relating to the genetic effects of MENDELian sampling (vector $\boldsymbol{g_{mi}}$) in environment *i;* e_i is the n_i *x 1* vector of residuals in environment *i.* Covariance between random effects (i.e., female parent and residual) in the model was assumed nil, such that

$E(y_i) = \mathbf{X_i}\,\beta_i$, $E(g_i) = \mathbf{0}$, $E(g_{mi}) = \mathbf{0}$, and $E(e_i) = \mathbf{0}$;

$$G = Var\begin{bmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} \sigma^2_{a1} & \sigma_{a12} & \sigma_{a13} & \sigma_{a14} \\ \sigma_{a21} & \sigma^2_{a2} & \sigma_{a23} & \sigma_{a24} \\ \sigma_{a31} & \sigma_{a32} & \sigma^2_{a3} & \sigma_{a34} \\ \sigma_{a41} & \sigma_{a42} & \sigma_{a43} & \sigma^2_{a4} \end{bmatrix} \qquad R = Var\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix} = \begin{bmatrix} \sigma^2_{e1} & 0 & 0 & 0 \\ 0 & \sigma^2_{e2} & 0 & 0 \\ 0 & 0 & \sigma^2_{e3} & 0 \\ 0 & 0 & 0 & \sigma^2_{e4} \end{bmatrix}$$

where σ_{ai}^2 is the additive genetic variance in environment *i,* σ_{aij} is the additive genetic covariance between environment i and j, and σ_{ei}^2 is error variance in environment *i.*

The phenotypic value of each individual was determined as the summation of all genetic and environmental effects in the model, i.e., $y_{ijk} = N(0, \sigma^2_{Bi}) + N(0, \frac{1}{4}\sigma^2_{ai}) + N(0, \sigma^2_W)$.

The levels of each effect were assumed to be a random sample from a large normal population. Correlated additive genetic effects among the 4 testing environments were created as:

$$\mathbf{A = B'C} \quad (2)$$

where A is the matrix of additive genetic effects, **B** is a matrix of the square root (Cholesky decomposition) of designed genetic variance-covariance matrix **G**, and **C** is a column vector of independent standard normal random deviates, such that

$$\text{Var}(\mathbf{A}) = \mathbf{B'}\text{Var}(\mathbf{C})\mathbf{B} = \mathbf{B'B} = \mathbf{G} \quad (3)$$

Heterogeneous genetic variances among progeny tests are reflected by the genetic variances and covariances in matrix G. The designed population genetic parameters such as heritability and type B genetic correlations were intentionally simulated to have a relatively large variation among the 4 environments aimed to represent a wide range of situations that may exist among real forest genetic data sets (*Table 1*). Without losing generality, phenotypic variance within each progeny test was set to 1.0 because data standardization is highly recommended in forest genetic data analysis to remove scale effects (WU, 1993; WHITE, 1996; DIETERS, 1996) and data standardization can always adjust phenotypic variance to 1.0 within a single environment.

Block effects were treated as fixed effects in this study for three reasons: (1) variance component estimation for random female genetic effect is not affected under customary linear models (i.e., no covariance assumed between random effects) whether block is treated as random or fixed effect; (2) treating block as fixed effect facilitates the application of multivariate methods in this study; and (3) the assumption of fixed block effects can help remove block effects from confounding genetic effects when data are unbalanced. However, for convenience in data generation, levels of block were created with the assumption that they were random samples from a normally distributed population ($\mu_b = 0$, $\sigma_B^2 = 2$), and that the observed variation among blocks was twice as large as the phenotypic variation within a block (i.e., $\sigma_B^2 = 2$). After data were generated, 30% mortality was simulated in all progeny tests by random deletion. A total of 300 independently simulated data sets, each containing 4 environments and 20 blocks, were generated and analyzed in this study.

Estimation of Type B Genetic Correlations

a. Multivariate methods

Multivariate computer programs MTDFREML (BOLDMAN et al., 1995) and ASREML (GILMOUR et al., 1997) were used to analyze each of the 300 data sets by treating the measurements of an arbitrary continuous trait from different environments as different traits. MTDFREML is a computer software

Table 1. – Designed heritabilities and type B genetic correlations for an arbitrary continuous trait among four simulated environments.

	Environment			
	1	2	3	4
	Heritability and type B genetic correlations			
1	0.40	0.90	0.80	0.70
2		0.30	0.70	0.60
3			0.20	0.50
4				0.10
	Designed genetic variance and covariance matrix (G)			
1	0.1000	0.0779	0.0566	0.0350
2		0.0750	0.0429	0.0260
3			0.0500	0.0177
4				0.0250

Shaded areas are narrow sense heritabilites of the four environments.

which uses a simplex algorithm to maximize the log likelihood function for variance component estimation and constrains estimates of genetic correlations within the theoretical parameter space. ASREML is, on the other hand, a computer program which uses an average information algorithm (GILMOUR et al., 1997) and sparse matrix techniques to efficiently solve large mixed models; it does not, however, automatically constrain estimates of genetic correlations within the theoretical parameter space. Therefore, results from MTDFREML and ASREML in this study were used to represent, respectively, constrained and unconstrained multivariate estimates of type B genetic correlations.

For both multivariate methods, input data structures were modified (*Table 2*) in order to estimate genetic variances and covariances for the same trait measured in different environments (type B) rather than for different traits measured on the same individuals (type A). Convergence criteria were set for MTDFREML as MVFV (Minimum Variance of Function Values in Simplex) $\leq 10^{-9}$, and for ASREML, $|(-2L_{n+1} - (-2L_n)| \leq 0.002$, respectively (BOLDMAN et al., 1995; GILMOUR et al., 1997). True genetic variance-covariance components were used as priors for starting the iterative processes. For both multivariate methods, two different groupings of environments were used to estimate variance components based on the assumption that estimates of genetic variances and covariances are system dependent so that estimates of type B genetic correlations are not necessarily the same if they are estimated from pair-wise or over all sites. The small grouping was paired environments while the larger grouping contained all four environments and estimated all pair-wise genetic correlations simultaneously.

Table 2. – Illustration of data structure used in multivariate analysis to estimate type B genetic correlations. Experimental design is assumed as a randomized complete block design with 3 environments, each having 3 blocks with one tree per family per block. Observations from different environments are treated as different traits.

Environment	Block	Family	Trait 1	Trait 2	Trait 3
1	1	1	10.51	.	.
1	1	2	9.83	.	.
1	1	3	7.78	.	.
1	2	1	8.39	.	.
1	2	2	7.67	.	.
1	2	3	6.78	.	.
1	3	1	12.34	.	.
1	3	2	11.23	.	.
1	3	3	10.98	.	.
2	4	1	.	8.65	.
2	4	2	.	8.21	.
2	4	3	.	7.67	.
2	5	1	.	9.69	.
2	5	2	.	8.76	.
2	5	3	.	8.65	.
2	6	1	.	6.67	.
2	6	2	.	7.43	.
2	6	3	.	5.89	.
3	7	1	.	.	12.34
3	7	2	.	.	13.45
3	7	3	.	.	10.56
3	8	1	.	.	13.45
3	8	2	.	.	12.56
3	8	3	.	.	11.98
3	9	1	.	.	12.17
3	9	2	.	.	13.64
3	9	3	.	.	15.48

Note that blocks 1 to 3 are from environment 1, blocks 4 to 6 are from environment 2 and blocks 7 to 9 are from environment 3. Dots stand for missing values, which are necessary and intentionally given in this data structure.

b. Univariate methods

Univariate methods used in this study included the traditional methods of YAMADA (1962) and BURDON (1977), and a GCA approach (LU et al., 1999). A previous comparative study among univariate methods suggested that the GCA approach improves type B genetic correlation estimation when data are highly unbalanced and heterogeneous variances exist among environments (LU et al., 1999).

For the YAMADA I method, the type B genetic correlation is estimated as:

$$\hat{r}_B = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_I^2 - \frac{(\hat{\sigma}_{g1} - \hat{\sigma}_{g2})^2}{2}} \qquad (4)$$

where $\hat{r}_B$ is the estimated type-B genetic correlation, $\hat{\sigma}_g^2$ is the genetic variance component estimated from a two-way analysis of variance involving data from two environments assuming homogeneous variance between them, $\hat{\sigma}_I^2$ is the estimate of variance component for the effect of G x E interaction, $\hat{\sigma}_{g1}^2$ and $\hat{\sigma}_{g2}^2$ are, respectively, the estimates of genetic variance components within environment 1 and 2. Often used in forest genetic studies is a simplified formula of Eq.4, YAMADA II, which is:

$$\hat{r}_B = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_I^2} \qquad (5)$$

where elements in Eq.5 are the same as in Eq.4.

With BURDON's method, the type B genetic correlation is estimated as:

$$\hat{r}_B = \frac{\hat{r}_{xy}}{\hat{h}_{\bar{x}} \hat{h}_{\bar{y}}} \qquad (6)$$

where $\hat{r}_{xy}$ is the phenotypic correlation between genetic group (i.e. half-sib families) means in environments **x** and **y**, and $\hat{h}_{\bar{y}}$ and $\hat{h}_{\bar{y}}$ are square-roots of the heritabilities of the genetic group means in environments **x** and **y**, respectively.

For the GCA approach, parental GCA effects (defined as ½ parental breeding values predicted by BLUP) are first predicted using the technique of univariate best linear unbiased prediction (BLUP) in each environment and these are adjusted to calculate type B genetic correlation as:

$$\hat{r}_B = \frac{Cov\left(\frac{\hat{g}_{xi}}{\hat{r}_{axi}}, \frac{\hat{g}_{yi}}{\hat{r}_{ayi}}\right)}{\overline{\hat{r}_{ax}\hat{r}_{ay}}\sqrt{Var\left(\frac{\hat{g}_{xi}}{\hat{r}_{axi}}\right) Var\left(\frac{\hat{g}_{yi}}{\hat{r}_{ayi}}\right)}} = \frac{\hat{r}^*}{\overline{\hat{r}_{ax}\hat{r}_{ay}}}, \qquad (7)$$

where $\hat{r}_B$ is the estimate of type B genetic correlation, r^* is a PEARSON correlation coefficient between adjusted parental GCA predictions in environments x and y using BLUP, $\hat{g}_{xi}$ and $\hat{g}_{yi}$ are predicted parental GCA effects in environments x and y respectively, and $\overline{\hat{r}_{ax}\hat{r}_{ay}}$ is the mean products of adjusted 'prediction accuracy' (LU et al., 1999) in the two environments.

Criteria for Comparisons

After the type B genetic correlations were estimated for each pair of environments within each of the 300 simulated data

sets using the above univariate and multivariate methods, three criteria were used to evaluate the estimation methods. First, empirical bias was calculated as the difference between means of estimated and true type B genetic correlations over 300 random samples for each pair of environments, i.e., $Bias = \bar{r}_{\hat{B}} - \bar{r}_{tB}$,

where $\bar{r}_{\hat{B}} = \frac{1}{N}\sum_{i}^{N} \hat{r}_{Bi}$ with $\hat{r}_{Bi}$ being the estimated type B genetic correlation of the ith sample; $\bar{r}_{tB} = \frac{1}{N}\sum_{i}^{N} r_{tBi}$ with r_{tBi} being the true type B genetic correlation of the ith sample. N (=300) is the total number of random samples. The statistical differences of the empirical biases from zero were tested by one-way analysis of variance (ANOVA).

The second criterion was the mean-distance (MD) between the estimated and true type-B genetic correlations which was calculated as: $MD = \frac{1}{N}\sum_{i}^{N} | \hat{r}_{Bi} - r_{t_{Bi}} |$. The smaller the MD, the closer the estimates to their true values, and consequently the higher estimation precision.

The third criterion was the PEARSON correlation between the estimates of type B genetic correlations and the true type B genetic correlations. Higher correlation reflects the better response of estimated type B to the changes of true type B genetic correlation and thus shows better quality of the related estimation methods. Outliers were excluded if their distances to the true values exceeded three times the MD.

Results

Bias

Among the univariate methods, as applied to pairs of environments, the GCA approach yielded empirically unbiased estimates of type B genetic correlations for all balanced and unbalanced data sets (*Table 3*). BURDON's method yielded unbiased estimates when data were balanced within an environment but yielded severely biased estimates when data were unbalanced due to missing values. YAMADA I (Eq. 4) also produced nearly unbiased estimates for almost all data sets. YAMADA II (Eq. 5) tended to give slightly downward bias, with severe bias for a few environmental pairs when ratios of genetic variances in two environments were greater than 2.0.

The unconstrained multivariate method ASREML and constrained multivariate method MTDFREML both yielded empirically unbiased estimates of type B genetic correlations when whether all four environments or only two environments were at a time analyzed (*Table 3*). MTDFREML tended to produce slightly downward bias but the magnitude was negligible. For a given pair of environments, the large and small groupings produced different estimates of type B genetic correlations (data not shown). This implied that, with multivariate methods, information from sites included in an analytical system is combined so that adding data from a third environment can influence the estimates of type B genetic correlation between the previous two environments.

Comparison of the best results from univariate methods with those of multivariate methods in terms of bias produced no obvious differences. The univariate methods, such as the GCA-

Table 3. – Empirical biases of type B genetic correlation estimates from different estimation methods for simulated half-sib data tested in a randomized complete block experimental design with single-tree plots.

		Environment Pairs					
Mortality	Estimation Method	1-2	1-3	1-4	2-3	2-4	3-4
0%	MTDFREML 1†	-0.011	-0.011	-0.014	0.015	0.017	0.019
	MTDFREML 2	-0.003	0.001	0.015	0.003	0.009	0.015
	ASREML 1	0.005	0.008	0.067*	0.008	0.069*	0.053*
	ASREML 2	0.002	0.009	0.058	0.036	0.047	0.044
	GCA-approach	0.002	0.008	0.031	0.000	0.019	0.009
	Yamada I	0.000	0.008	0.031	0.006	0.023	0.017
	Yamada II	-0.018*	-0.051*	-0.148*	-0.019	-0.082*	-0.034
	Burdon	0.002	0.008	0.031	0.000	0.019	0.009
30%	MTDFREML 1	-0.023*	-0.019	-0.039	0.017	-0.004	0.020
	MTDFREML 2	-0.009	-0.001	-0.012	0.003	-0.036	0.022
	ASREML 1	0.006	0.020	0.110*	0.022	0.074*	0.098*
	ASREML 2	0.006	0.021	0.134*	0.024	0.085*	0.104*
	GCA-approach	0.004	0.022	0.049	0.018	0.047	0.033
	Yamada I	-0.004	0.015	0.047	0.022	0.041	0.038
	Yamada II	-0.027*	-0.053*	-0.153*	-0.013	-0.091*	-0.031
	Burdon	-0.232*	-0.192*	-0.130*	-0.161*	-0.077*	-0.022

True genetic parameters for different environments are given in *table 1.* MTDFREML (or ASREML) 1 and 2 refer to the 4-environment and 2-environment grouping, respectively.
*Biases are significantly different from zero at the probability level $\alpha \leq 0.05$.

approach and YAMADA I, produced estimates that are as unbiased as those from multivariate methods for almost all data structures simulated in this study. Other univariate methods, BURDON's method and YAMADA II, yielded empirically biased estimates of type B genetic correlations when data were either unbalanced due to missing values or when heterogeneity of variances among environments was large. Such results were consistent with previous theoretical considerations and empirical studies about these univariate methods (FERNANDO et al., 1984; ITO and YAMADA, 1990; DUTILLEUL and CARRIÈRE, 1998; LU et al., 1999).

Precision

The mean distance (MD) between the estimated and true type B genetic correlations from different estimation methods showed consistent differences among the estimation methods (*Table 4*). Constrained estimation methods (i.e., MTDFREML and YAMADA II) had smaller MDs than those unconstrained methods (i.e., ASREML, YAMADA I, BURDON and the GCA approach). For the constrained multivariate method MTDFREML, the 4-environment grouping consistently had smaller MD than 2-environment grouping. But this was not always true for unconstrained multivariate ASREML. Among all estimation methods, MTDFREML with a 4-environment system consistently yielded the smallest MD for a given pair of environments. This was followed by the MTDFREML with two-environment system, YAMADA II, and then the GCA approach and the ASREML (*Table 4*).

Regardless of estimation methods, the mean distance between estimated and true type B genetic correlations became steadily larger as heritabilities in two environments became low (*Table 4*). For example, for all methods except BURDON's, MDs were not greater than 0.1 between environments 1 and 2, which had heritabilities of 0.4 and 0.3, respectively. In contrast, MDs were greater than 0.2 for almost all methods between environments 3 and 4, which had true heritabilities of only 0.2 and 0.1, respectively.

Correlations Between the Estimated and True Type B Genetic Correlations

PEARSON correlation coefficients (calculated based on 300 randomly simulated data samples) between the estimated and true type B genetic correlations for a given pair of environments were generally low for all estimation methods (*Table 5*). Considerable differences, however, existed among estimation methods. For the constrained multivariate methods, the grouping containing only two environments yielded higher PEARSON correlation coefficient between the true and estimated type B genetic correlations than the grouping that included all 4 environment (*Table 5*). Among all estimation methods, the correlation coefficients were highest for the univariate GCA approach, which was followed by YAMADA II and then multivariate method MTDFREML with 2-environment grouping. BURDON's method was equal to the GCA approach when data were balanced within an environment, but inferior to the GCA-approach when data were unbalanced within an environment. YAMADA I

Table 4. – Mean-distance (MD) between estimates of type B genetic correlation and their true values for different estimation methods using simulated half-sib families tested in a randomized complete block experimental design with single-tree plots.

Mortality	Estimation Method	Environment Pairs					
		1-2	1-3	1-4	2-3	2-4	3-4
0%	MTDFREML 1	0.060	0.092	0.137	0.119	0.150	0.187
	MTDFREML 2	0.067	0.100	0.166	0.124	0.188	0.205
	ASREML 1	0.076	0.108	0.217	0.126	0.226*	0.232*
	ASREML 2	0.080	0.108	0.209	0.126	0.205	0.218
	GCA-approach	0.077	0.106	0.192	0.124	0.197	0.210
	Yamada I	0.071	0.108	0.193	0.126	0.193	0.208
	Yamada II	0.069	0.104	0.176	0.122	0.166	0.180
	Burdon	0.077	0.106	0.192	0.124	0.197	0.210
30%	MTDFREML 1	0.069	0.107	0.176	0.132	0.184	0.230
	MTDFREML 2	0.081	0.125	0.212	0.124	0.200	0.293*
	ASREML 1	0.099	0.146	0.331*	0.157	0.283*	0.306*
	ASREML 2	0.096	0.148	0.350*	0.157	0.282*	0.316*
	GCA-approach	0.100	0.149	0.316*	0.156	0.300*	0.332*
	Yamada I	0.087	0.142	0.304*	0.155	0.279*	0.303*
	Yamada II	0.084	0.129	0.218	0.144	0.207	0.239
	Burdon	0.245*	0.255*	0.305*	0.239*	0.313*	0.293*

True genetic parameters for different environments are given in *table 1.* Out-of-bound estimates were accepted with their original values. † MTDFREML (or ASREML) 1 and 2 refer to the 4-environment and 2-environment grouping, respectively. *MD from a estimation method is significantly different from the MD from the commonly used method of YAMADA II.

Table 5. – PEARSON correlation coefficients between estimates of type B genetic correlations and their underlying true values for different estimations methods using simulated half-sib families tested in a randomized complete block experimental design with single-tree plots.

Mortality	Estimation Method	Environment Pairs					
		1-2	1-3	1-4	2-3	2-4	3-4
0%	MTDFREML 1	0.264	0.278	0.279	0.228	0.399	0.207
	MTDFREML 2	0.314	0.336	0.319	0.336	0.424	0.361
	ASREML 1	0.282	0.292	0.254	0.273	0.358	0.272
	ASREML 2	0.252	0.321	0.319	0.193	0.236	0.271
	GCA-approach	0.397	0.405	0.381	0.411	0.454	0.363
	Yamada I	0.311	0.302	0.354	0.354	0.407	0.347
	Yamada II	0.313	0.296	0.360	0.360	0.424	0.365
	Burdon	0.397	0.405	0.381	0.411	0.454	0.363
30%	MTDFREML 1	0.128	0.236	0.162	0.219	0.233	0.075
	MTDFREML 2	0.204	0.290	0.242	0.367	0.269	0.203
	ASREML 1	0.160	0.234	0.082	0.262	0.256	0.165
	ASREML 2	0.181	0.260	0.100	0.257	0.214	0.141
	GCA-approach	0.298	0.329	0.336	0.339	0.360	0.235
	Yamada I	0.205	0.254	0.192	0.280	0.289	0.197
	Yamada II	0.226	0.268	0.270	0.303	0.332	0.247
	Burdon	0.237	0.181	0.162	0.197	0.046	0.067

True genetic parameters for different environments are given in *table 1.* Out-of-bound estimates were accepted with their original values. † MTDFREML (or ASREML) 1 and 2 refer to the 4-environment and 2-environment grouping, respectively.

and ASREML performed intermediately to the GCA approach and BURDON's method.

Distribution of Estimates

For a given true type B genetic correlation, various estimates were obtained from random data samples due to sampling errors. The scatter plots of estimated type B genetic correlations against their true values were affected by both the estimation methods and the true genetic parameters. The multivariate method MTDFREML and the univariate method of YAMADA II (Eq.5) constrained estimates of type B genetic correlations to be no greater than 1 and, consequently, skewed the distribution of estimates when true type B genetic correlation was close to 1 (*Figure 1*). The multivariate method ASREML and all other univariate methods, on the other hand, allowed for out-of-bound estimates which made the distribution of estimates nearly symmetric against the true values. These univariate and the unconstrained multivariate methods, however, produced a few estimates with very high magnitude. In a separate simulation, the frequency of out-of-bound estimates steadily increases as heritabilities in one or both of the environments became smaller (*Figure 2*). From the simulation results of this study, between environments 2 and 3, which had heritability 0.3 and 0.2, respectively, there were only 21 (out of 300) estimates greater than 1.0. In contrast, between environments 1 and 4, which had heritabilities of 0.4 and 0.1, respectively, more than 60 (out of 300) estimates were greater than 1.0, although the designed mean true parameters for both environment pairs were 0.7.

Discussion

Results of simulations in this study have shown that estimates of type B genetic correlations using multivariate methods were empirically as unbiased as the best results from univariate methods (*Table 3*) for both balanced and balanced data with heterogeneous genetic and error variances. Although a tendency for downward bias was detected for the constrained multivariate method MTDFREML, the magnitudes of such biases were negligible. This tendency of downward bias was likely caused by constraining estimates within theoretical parameter space, some of which would otherwise be out-of-bounds. The small magnitudes of such downward biases were due to: (1) the relatively small proportion of estimates of type B genetic correlations which would be out-of-bounds between environments having higher heritabilities and, consequently, causing little changes to the mean of estimates when they were constrained within parameter space; and (2) potential compensation by upward biases of estimates of type B genetic correlations for pairs of environments having low heritabilities.

The empirically unbiased estimates of type B genetic correlations from the univariate methods of YAMADA I and II were likely due to the specific data structures simulated in this study. Previous numerical studies (FERNANDO et al., 1984;

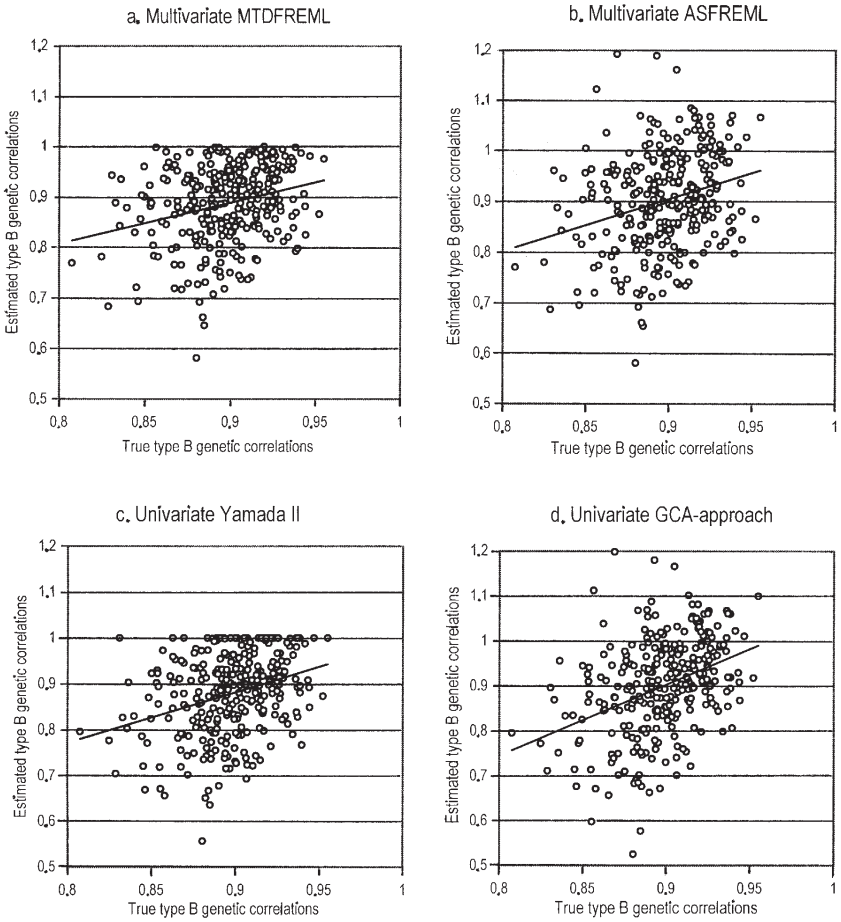

Figure 1. – Scatter plots of estimates of Type B genetic correlations from multivariate and univariate methods against the actual true type B genetic correlations for the 300 random samples. The mean true parameters are: $h_1^2 = 0.4$, $h_2^2 = 0.3$ and $r_B = 0.9$. Experimental designs in both environments are randomized complete block design with single-tree plots. There are 90 families in each environment.

DUTILLEUL and CARRIÈRE, 1998; LU et al., 1999) indicated that estimates of type B genetic correlations from YAMADA's methods were subject to bias when heterogeneity of variances was severe and data were highly unbalanced between environments in terms of the relative size of the experiments. In this simulation study, data samples had about the same sizes across all environments although the genetic and environmental variances were heterogeneous. As a result, biases from YAMADA's methods became less severe or negligible.

Because each estimation method was used to analyze the same data sets, differences among estimation methods in MD reflected their differential estimation precision. The smaller MDs (*Table 4*) obtained from constrained estimation methods (MTDFREML and YAMADA II) than from unconstrained methods (ASREML, YAMADA I, BURDON, and the GCA approach) were expected because, by theory, the true values of type B genetic correlations cannot be located outside the parameter space. Higher estimation precision can, therefore, be achieved by constraining estimates of genetic correlations within the parameter space.

The smaller MD for a multivariate method MTDFREML which used data from all four environments to estimate the pair-wise type B genetic correlation compared to the two-environment grouping may possibly be due to the more stable estimates of genetic variances from the four-environment grouping than from the two-environment grouping. Evidence supporting this reasoning is that the standard deviation of estimates of genetic variance among the 300 random data samples for a given environment was slightly smaller for the four-environment grouping than for the two-environment grouping (data not shown). This may suggest that, for multivariate methods, additional information in a larger grouping system helped improve the quality of estimates of genetic variance components so that the sampling error for type B genetic correlation is smaller. Comparison of MDs between multivariate methods can, however, be complicated by the search parameters being used in each run because the search parameters can influence the solution, especially in an analysis with a large number of parameters to be estimated.

The magnitudes of MDs were large for estimates of type B genetic correlations between pairs of environments with traits of low heritabilities. In the simulation, the MDs increased substantially from balanced data with 20 single-tree blocks to unbalanced data with an average of 14 single-tree blocks. With such large sampling errors, the biases of the estimates become less meaningful. The results imply that to have reliable estimates of type B genetic correlations relatively large numbers of replications and families in the field experimental designs are required (NOTTER and DIAZ, 1993).

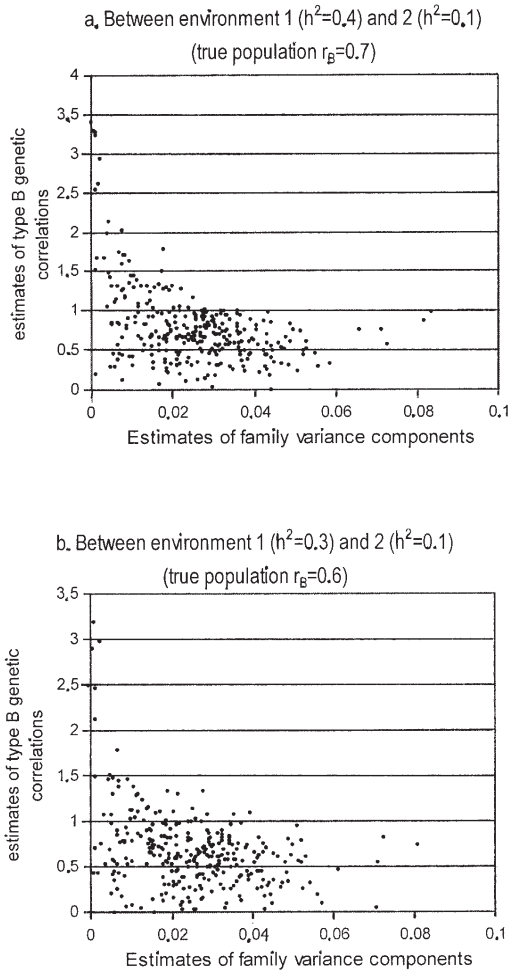

Figure 2. – Effects of genetic variance component estimates on the number of out-of-bound estimates of type B genetic correlation.

Besides the properties of unbiasedness and precision, the correlation between estimated and true type B genetic correlation could be important because it reflects the response of estimates to changes on the underlying true values. In this study, the generally low correlation between the estimated and true type B genetic correlation was attributable to the large sampling errors of type B genetic correlations as indicated by the range of estimates for a given true value (*Figure 1*). For constrained multivariate method, a smaller system containing only two environments seemed to be more desirable in this regard than a larger system containing more environments in type B genetic correlation estimation. Interferences by information from unconcerned environments may have reduced the consistence of estimates of type B genetic correlation to their true values. This trend was clearly demonstrated by MTDFREML with the higher correlation coefficients from the two-environment system than from the four-environment system for various genetic backgrounds. However, the two-environment system had larger MDs as compared with the four-environment system. The reason for such contradicting results was not clear, probably due to differences between the two grouping systems in variance component estimation.

While constraining estimates of type B genetic correlation within the theoretical parameter space have increased the overall estimation precision, it has also skewed up the distribution of estimates and cause downward biases. For example, scatter plots of estimates of type B genetic correlations against their true values indicated that for the unconstrained methods (multivariate ASREML and univariate GCA approach) the estimates were nearly symmetrically distributed around a given true value (*Figure 1b* and *1d*). In contrast, for constrained multivariate method MTDFREML and YAMADA II, the distribution of estimates (*Figure 1a* and *1c*) was not symmetric against the true values due to the limitation of the theoretical boundary.

The choice of constrained or unconstrained estimates of type B genetic correlations may be objective specific. For example, for theoretical studies of the distribution and sampling errors of type B genetic correlation, unconstrained estimates may be more desirable to show the original distributional pattern of estimates, so that confidence intervals of estimates can be investigated. In practical genetic data analysis, however, constrained estimates of type B genetic correlations may be easier to interpret and more reasonable to apply when they are involved in indirect selections. For example, for a given set of values of heritabilities, phenotypic variance and selection intensities, genetic response from indirect selection is theoretically less or equal to the gain from direct selection (FALCONER, 1989). However, if estimates of genetic correlations greater than 1 were used, the above theoretical rule would be violated, yielding greater predicted genetic gains from indirect selection than from direct selection.

Although univariate methods, such as the GCA approach, can achieve comparable estimates of type B genetic correlations to those from unconstrained multivariate methods for balanced and unbalanced data, it may be practically more convenient and efficient to use multivariate methods if computer software is available. Simulation in this study has only demonstrated the desirable properties of multivariate methods with a simple half-sib genetic structure containing no genetic relatedness. More desirable properties of multivariate methods would have manifested had complex genetic structures been involved. It would only be prudent to point out the potential advantages of multivariate methods in analyzing data of complex genetic structures, including (1) making use of pedigree information so that type B genetic correlations can be estimated between two environments that have only indirect genetic connection; (2) addressing genetic relatedness among genetic groups within and between environments so that the assumption of independence among genetic groups within an environment may be relaxed (NOTTER and DIAZ, 1993); and (3) including data from experiments for mixture of mating designs and different generations. As tree improvement programs progress into advanced generations and data structures become more complex (WHITE, 1996; BORRALHO and DUTKOWSKI, 1998), multivariate methods are expected to be more appropriate in estimating type B genetic correlations.

Conclusion

Although some univariate methods can yield unbiased estimates of type B genetic correlation for unbalanced data with heterogeneous variances, advantages associated with multivariate methods make them viable options in the estimation of type B genetic correlations. Estimates of type B genetic correlations from multivariate methods are empirically unbiased for unbalanced data with heterogeneous variances. Constraining estimates within theoretical parameter space help improve estimation precision and practical application. Placing more environments in an analytical system with multivariate methods not only increases computational efficiency, but may also enhance the quality of estimates of genetic variances,

resulting in smaller sampling errors of the estimates of type B genetic correlations.

Literature Cited

ADAMS, W. T., WHITE, T. L., HODGE, G. R. and POWELL, G.L.: Genetic parameters for bole volume in longleaf pine: large sample estimates and influences of test characteristics. Silvae Genet. **43**: 357–366 (1994). — BOLDMAN, K. G., KRIESE, L. A., VAN VLECK, L. D., VAN TASSELL, C. P. and KACHMAN, S. D.: A manual for use of MTDFREML. A set of programs to obtain estimates of variances and covariances [draft]. U.S. Department of Agriculture, Agricultural Research Service (1995). — BORRALHO, N. M. G. and DUTKOWSKI, G. W.: Comparison of rolling front and discrete generation breeding strategies for trees. Can. J. For. Res. **28**: 987–993 (1998). — BURDON, R. D.: Genetic correlation as a concept for studying genotype-environment interaction in forest tree breeding. Silvae Genet. **26**: 168–175 (1977). — DIETERS, M. J. J.: Genetic parameters for slash pine (*Pinus elliottii*) grown in south-east Queensland, Australia:growth, stem straightness and crown defects. Forest Genetic **3**(1): 27–36 (1996). — DIETERS, M. J., WHITE, T. L. and HODGE, G. R.: Genetic parameter estimates for volume from full-sib tests of slash pine (*Pinus elliottii*). Can. J. For. Res. **25**: 1397–1408 (1995). — DUTILLEUL, P. and CARRIÈRE, Y.: Among-environment heteroscedasticity and the estimation and testing of genetic correlation. Heredity **80**: 403–413 (1998). — FALCONER, D. S.: Introduction to quantitative genetics. John Wiley and Sons, New York (1989). — FERNANDO, R. L., KNIGHTS, S. A. and GIANOLA, D.: On a method of estimating the genetic correlation between characters measured in different experimental units. Theor. Appl. Genet. **67**: 175–178 (1984). — GILMOUR, A. R., THOMPSON, R., CULLIS, B. R. and WELHAM, S. J.: ASREML, NSW Agriculture, Orange, 2800, Australia (1997). — HILL, W. G. and THOMPSON, R.: Probability of non-positive define between-group or genetic covariance matrices. Biometrics **34**: 429–439 (1978). — HODGE, G. R. and PURNELL, R. C.: Genetic parameter estimates for wood density, transition age and radial growth in slash pine. Can. J. For. Res. **23**: 1881–1891 (1993). — HODGE, G. R. and WHITE, T. L.: Genetic parameter estimates fro growth traits at different ages in slash pine and some implications for breeding. Silvae Genet. **41**: 252–262 (1992). — HUBER, D. A., WHITE, T. L. and HODGE, G. R.: Variance component estimation techniques compared for two mating designs with forest genetic architecture through computer simulation. Theor. Appl. Genet. **88**: 236–242 (1994). — ITOH, Y. and YAMADA, Y.: Relationship between genotype x environment interaction and genetic correlation of the same trait measured in different environments. Theor. Appl. Genet. **80**: 11–16 (1990). — JOHNSON, G. R. and BURDON, R. D.: Family-site interaction in *Pinus radiata*: implications for progeny testing strategy and regionalized breeding in New Zealand. Silvae Genet. **39**: 55–62 (1990). — KOOTS, K. R. and GIBSON, J. P.: Realized sampling varainces of estimates of genetic parameters and the difference between genetic and phenotypic correlations. Genetics **143**: 1409–1416 (1996). — LAMBETH, C. C. and GLADSTONE, W. T.: Statistical efficiency of row and noncontiguous family plot in genetic test of loblloly pine. Silvae Genet. **32**: 24–28 (1983). — LOO-DINKINS, J. A. and TAUER, C. G. : Statistical efficiency of six progeny test field designs on three loblloly pine site types. Can. J. For. Res. **17**: 1066–1070 (1987). — LOO-DINKINS, J. A., TAUER, C. G. and LAMBETH, C. C.: Selection system efficiencies for computer simulated progeny test field designs in loblloly pine. Theor Appl Genet **79**: 89–96 (1990). — LU, P.: Biases if incomplete linear models in forest genetic data analysis and optimal methods for estimating type B genetic correlations. Ph. D. dissertation. School of Forest Resources and Conservation, University of Florida (1999). — LU, P., WHITE, T. L. and HUBER, D. A.: Estimating type B genetic correlations with unbalanced data and heterogeneous variances for half-sib experiments. For. Sci. **45**: 562–572 (1999). — NOTTER, D. R. and DIAZ, C.: Use of covariances between predicted breeding values to assess the genetic correlation between expressions of a trait in 2 environments. Genet Sel Evol **25**: 353–372 (1993). — PATTERSON, H. D. and THOMPSON, R.: Recovery of interblock information when block sizes are unequal. Biometrika **58**: 545–554 (1971). — PSWARAYI, I. Z., BARNES, R. D., BIRKS, J. S. and KANAWSKI, P. J.: Genotype-environment-interaction in a population of *Pinus elliottii* ENGELM. var *elliottii.* Silvae Genet. **46**: 35–40 (1997). — SAS Institute: SAS language, SAS Institute Inc., Cary, NC 27513 (1990). — SCHAEFFER, L. R. and WILTON, J. W.: Simultaneous estimation of variance and covariance components for multtrait mixed model equations. Biometrics **34**: 199–208 (1978). — SEARLE, S. R., CASELLA, G. and MCCULLOCH, C. E.: Variance Components. John Wiley & Sons, Inc., New York. 501 p. (1992). — WHITE, T. L.: Genetic parameter estimates and breeding value prediction: issues and implications in tree improvement programs. In: DIETERS, M. J., MATHESON, A. C., NIKLES D. G., HARWOOD, C. E. and WALKER, S. M. (eds): Tree improvement for sustainable tropic forestry. Prox. QFRI-IUFRO Conf., Caloundra, Queensland, Australia. pp. 110–117 (1996). — WHITE, T. L., HODGE, G. R. and POWELL, G. L.: An advanced-generation tree improvement plan for slash pine in the southeastern United States. Silvae Genet. **42**: 359–371 (1993). — WOOLASTON, R. R., KANAWSKI, P. J. and NIKLES, D. G.: Genotype-environment interaction in *Pinus caribaea* var. *hondurensis* in Queensland, Australia. II. Family x site interaction. Silvae Genet. **40**: 228–232 (1991). — WU, X. M.: Early selection in Lodgepole pine. Ph.D. thesis, University of Alberta, Edmonton, Alberta, Canada T6G 2H1. (1993). — YAMADA, Y.: Genotype by environment interaction and genetic correlation of the same trait under different environments. Jap. J. Genet. **37**: 498–509 (1962).

Genetic Variation of Oaks (*Quercus* spp.) in Switzerland

2. Genetic Structures in "Pure" and "Mixed" Forests of Pedunculate Oak (*Q. robur* L.) and Sessile Oak (*Q. petraea* (MATT.) LIEBL.)

By R. FINKELDEY

Swiss Federal Research Institute WSL, Zürcherstraße 111, CH-8903 Birmensdorf, Switzerland
Tel.: +41 - 1 - 7392489; Fax.: +41 - 1 - 7392215; E-mail: reiner.finkeldey@wsl.ch

(Received 5th July 2000)

Abstract

Sessile oak (*Quercus petraea*) and pedunculate oak (*Q. robur*) are two closely related, interfertile taxa. They are the most frequent oak species in Switzerland. Allelic and genotypic structures at 17 isozyme gene loci were observed in 21 populations from Switzerland. Twelve populations of *Q. petraea,* six populations of *Q. robur*, and three "mixed" populations (*Q. petraea* and *Q. robur*) were investigated. The species status of the populations was confirmed by Principal Component Analysis (PCA) based on leaf morphological traits. All populations are highly variable at enzyme gene loci. Differentiation among the taxa is reflected at allelic structures at several enzyme gene loci (*ACP-C, GDH-A, IDH-B, NDH-A, PGM-A*). An excess of homozygotes relative to corresponding HARDY-WEINBERG structures was observed in all populations. Moderate levels of inbreeding are likely to contribute to these genotypic structures, but heterogeneity of inbreeding coefficients among loci suggests that deviations from random mating are not the only cause of the homozygote excess at particular loci (*AAP-A, PGM-A*). On average, expected heterozygosity is highest in the "mixed" populations, but observed heterozygosity of the "mixed" stands is in-between *Q. petraea* and *Q. robur*. A plausible