Theor Appl Genet (2005) 110: 311–322
DOI 10.1007/s00122-004-1840-4

ORIGINAL PAPER

K. Isik · J. Kleinschmit

Similarities and effectiveness of test environments in selecting and deploying desirable genotypes

Received: 14 April 2004 / Accepted: 10 October 2004 / Published online: 30 November 2004

Abstract Rooted cuttings of 40 different clones (genotypes) of *Picea abies* (L.) Karst were planted in seven test environments (P, H, L, S, B, M and K) in northern Germany. Type-B genetic correlations between the pairs of test sites were estimated, and a cluster diagram based on inter-site genetic correlations for height was used to illustrate the general pattern of similarities among the test sites. Principal component analysis (PCA) showed that the first three principal components explained 97% of the variance in height, thereby providing a good summary of the relationships among the test sites. Test sites B and K accounted for a relatively high—and test sites H and M for a relatively low—proportion of genotype-environment interaction sums of squares and, consequently, for "ecovalence". "Moderator" and "instigator" test sites for interactions in different site combinations were detected. On the basis of several similarity measures, we observed two distinct groups: sites M-S-P formed one group, while sites K-L formed the other; sites H and B were closer to the first group. The empirical data indicated that a good test site should have the following features: (1) low interaction behaviour, (2) low coefficient of variation, (3) high genotypic selectivity, (4) high coefficient of determination, (5) high efficiency of expected gain. Based on these criteria, site M was determined to be the best test site for screening genotypes to be planted in other environments. We concluded that one broadly adapted vegetative propagation population of *P. abies* can be used for the potential planting environments in northern Germany.

Communicated by D.B. Neale

This revised version was published online in December 2004 with visible equation.

K. Isik (✉)
Department of Biology, Faculty of Arts and Sciences,
Akdeniz University, 07058 Antalya, Turkey
E-mail: kani@akdeniz.edu.tr
Tel.: +90-242-2262097
Fax: +90-242-2278911

J. Kleinschmit
Department of Forest Genetic Resources,
Lower Saxony Forest Research Institute,
34355 Staufenberg-Escherode, Germany

Introduction

One purpose of provenance, family and clonal tests is to obtain data on which to base deployment strategies for genetically improved (seed and/or clonal) materials. If the relative performances of the genetic entries grown on different environments are highly different, then genotype-environment interaction (G×E) becomes a major challenging factor in genetic improvement programmes (Zobel and Talbert 1984). In such cases, the breeder is faced either with developing separate breeding populations for each environment and/or with selecting genotypes that generally perform well over many environments (McKeand et al. 1990). Various genetic parameters are used to *determine genotypes* displaying high stability in a wide range of environments. It is also equally important to *determine test sites* that have a intrinsic high ability to detect and screen desirable genotypes that would be successful in a wide range of environments. When the main emphasis is on test environments, the questions often asked are:

- How similar is a given test environment to others with respect to the expression of the genotypes' characteristics?
- How do the genotypes of interest "read and interpret" the environmental variables at a given test site?
- What are the best test sites that could be used as to predict performance in other environments?

One approach to answering such questions and to determine the degree of similarity (or dissimilarity) among test environments is to estimate Type-B genetic correlations (Falconer 1952; Yamada 1962; Burdon 1977; Lu et al. 1999). These correlations show the degree

to which test sites are similar with respect to ranking genetic entries. Type-B genetic correlation has been routinely utilized in forest tree breeding to estimate the degree of similarity of genetic group ranking and level of G×E interaction among the test sites concerned (Burdon 1977; Hodge and White 1992; Adams et al. 1994; Dieters et al. 1995) and for predicting genetic gains from indirect selection (Burdon 1977; Johnson 1997).

The performance of genotypes (40 clones) of *Picea abies* (L.) Karst in seven test environments (sites) has been analysed in previous studies (St Clair and Kleinschmit 1986; Isik et al. 1995; Isik and Kleinschmit 2003). When 17 years of age, the clones were analysed for several stability-related parameters and contributions of individual clones to total interaction, and the magnitude of variance (V_{cs}) due to G×E was determined. G×E was determined to be a significant factor on the basis of traditional F tests but not so based on non-parametric tests using $S^{(m)}$ statistics (Isik and Kleinschmit 2003). In the investigation reported here, the emphasis was placed on the test sites rather than on genetic groups. Our general objective was to determine the similarities among the test environments and to subsequently draw inferences about deployment of genetic materials based on various measures of similarity. Our specific objectives were: (1) to determine Type-B genetic correlations between the test sites; (2) to estimate genetic parameters (genetic variances, genetic covariances, heritabilities) associated with Type-B genetic correlations; (3) to estimate the contribution of each test environment to total interaction; (4) to determine the genotypic selectivity of each test site; (5) to predict gains at the other test sites when selection and screening is carried out at a given test site; (6) to determine the efficiency of each specific test site with respect to serving others.

Materials and methods

Genotypes

The genotypes used in this study were rooted cuttings (ramets) from 40 different clones of Norway spruce [*Picea abies* (L.) Karst.]. The terms "genotypes" or "genetic groups" are used in the same context as "clones" in the study. The genotypes were selected at age 4 years on the basis of their growth performance from an initial vegetative propagation study started in 1968. Prior to such selection, each parent ortet was chosen from 13 different provenances of outstanding performance. The tertiary (third-cycle) cuttings from these 40 selected clones were rooted in spring 1974 and grown for 3 years in the nursery (Kleinschmit 1974). The rooted cuttings were then planted during the spring of 1977 at seven test sites. Different traits of each tree (each ramet) were observed at different ages. The data evaluated in this study are based on height in centimeters, measured at 17 years of age from rootings in the nursery.

Test sites and the experimental design

The seven test sites are located in northern Germany and were chosen to represent the range of environments where genetically desirable Norway spruce clones could be planted in the future. Three test sites [S = Syke (Hasbruch), M = Medingen (Goehrde), B = Binnen (Nienburg)] are located in the low coastal plains in the north of the country (between 39 m and 50 m a.s.l.), while the other four [P = Paderborn, K = Kattenbühl, H = Holzminden (Neuhaus), L = Lautenthal (Seesen)] are located further south and at relatively higher elevations (between 340 m and 575 m a.s.l.). For more details on the test sites see St Clair and Kleinschmit (1986) and Isik et al. (1995). The experimental design at each test site was a randomized complete block design with 20 blocks, each block containing a single-tree plot of each of the 40 clones (and nine seedlings from Westerhof origin). The performances of the seedlings compared to those of the clones have been evaluated earlier by St Clair and Kleinschmit (1986) and Isik et al. (1995). The present study includes only the clonal material. Thus, each test site contained 40 clones, each initially with 20 ramets.

Biostatistical analyses

In order to estimate genetic parameters at each test site and to estimate Type-B genetic correlations between site pairs, our first step was to perform separate analysis of variance for each test site. Other parameters were estimated at subsequent steps as described below.

Analysis of variance (ANOVA)

Applying general linear models (GLM) procedures, we performed ANOVA for each test site to estimate genetic parameters within the test sites (SAS 1987; Sokal and Rohlf 1995) (Table 1). The ANOVA model was:

$$X_{ik} = \mu + C_i + e_{ik},$$

where X_{ik} is the observed value of the kth tree (ramet) of the ith clone; μ, the overall site mean; C_i, the effect of the ith clone ($i = 1 \rightarrow c$; where $c = 40$); e_{ik}, the error (within clone) term ($k = 1 \rightarrow n$; where n = the number of ramets per clone; initial n was equal to 20).

Block effects were not considered in the analyses due to differential deaths of clones, each of which was represented only once within each block.

Genetic variances

Expected mean square values obtained by GLM as in the last column of Table 1 were used to estimate genetic variances between groups. Since each "genetic group" in this study is a separate clone [that is, since the members

Table 1 ANOVA model[a] used to conduct a separate analysis for each test site

Source of variation[b]	Sum of squares	Degrees of freedoms	Mean square	Expected mean square
Groups = c (clones)	SS_c	$c-1$	MS_g	$V_e + kV_c$
Error	SS_e	$c(n-1)$	MS_e	V_e
Total		$c.\ n-1$		

[a] c = the number of groups (clones) = 40; n = 20 = initial number of trees (ramets) per clone; $k = 13.317$ = effective number of trees per group (clone); $V_e = MS_e$ = within-group (error) variance; V_c = variance due to group differences = $(MS_g - MS_e)/k$

[b] All effects are assumed to be random

(ramets) of a given clone are identical genotypes], it is assumed that $V_g = V_c$. Thus:

$V_c = V_g$ = genetic variance between clones (genetic groups) at a test site,
$V_p = V_c + V_e$ = phenotype variance,
$V_{\bar{c}} = V_{\bar{g}} = V_c + (V_e/k)$ = variance of the clone means.

A subscript of x or y can be added to each parameter in order to indicate to which test site the parameters belong to. For example, V_{g_x} and V_{g_y} represents genetic variances at test site x and test site y, respectively.

Heritabilities

Heritability ($h^2_{\bar{g}_x}$) of clone (group) means (or repeatability) at test site x is estimated by $h^2_{\bar{g}_x} = V_{c_x}/V_{\bar{g}_x}$. Broad-sense heritability ($h^2_{b_x}$) (or heritability on an individual ramet basis) at test site x is estimated by $h^2_{b_x} = V_{c_x}/V_{P_x}$. For estimates at test site y, corresponding values with subscript y are used.

Genetic covariance

Lu et al. (1999) reported that when data are balanced within and across test sites, it yields unbiased estimates of genetic covariance. Therefore, in the present study missing observations for a given clone due to differential mortality at different test sites were balanced (by using only equal numbers of individuals across the test sites for a given clone) (see n_{ij}, Table 2). SAS (1987) PROC GLM with the MANOVA option was used to obtain cross-product matrices [Error sum of squares cross products (E-SSCP matrix) and Type-III SSCP (H-matrix)] and then to estimate covariance components in the manner as variance components (with a model analogous to that in Table 1). Genetic covariance (COV g_{xy}) between environments x and y was estimated according to the equation:

$$\mathrm{COV}g_{xy} = (\mathrm{MCP}_g - \mathrm{MCP}_e)/k \quad (1)$$

where MCP indicates the mean cross-products.

The procedure to calculate $\mathrm{COV}g_{xy}$ by Eq. 1 is analogous to using mean square values for estimating V_c in Table 1. If site x and site y are the same site, then COV $g_{xy} = V_c = V_g$ (Table 3). Thereby, in addition to the method in Table 1, the V_c value can also be found from the corresponding values in the E- and H-matrices, just as $\mathrm{COV}g_{xy}$ is calculated.

Type-B genetic correlations (genetic correlations between pairs of test sites)

When a given trait of the same clone is measured on two different environments (x and y), the genetic correlation between the environments concerned is denoted Type-B genetic correlation. Type-B genetic correlation ($r_{g_{xy}}$) is one criterion by which to measure the level of similarities among the test environments. It is essentially a measure of G×E, and $0 \leqslant r_{g_{xy}} \leqslant 1$ (Burdon 1977; Dieters et al. 1995). A high genetic correlation (i.e., $r_{g_{xy}} \approx 1$) between environments x and y indicates that the genotypes behave nearly the same (i.e. the rank orders of genetic groups will be more or less similar) at the two environments, with no detectable G×E. It can be calculated by:

$$r_{g_{xy}} = \frac{\mathrm{COV}g_{xy}}{\sqrt{V_{g_x} V_{g_y}}}. \quad (2)$$

Burdon (1977) provided an alternative formula for the estimation of $r_{g_{xy}}$ value as:

$$r_{g_{xy}} = \frac{r_{\bar{g}_{xy}}}{(h_{\bar{g}_x} h_{\bar{g}_y})} \quad (3)$$

where: $r_{\bar{g}_{xy}}$ = correlation coefficients between clonal means at environments x and y, and

$$r_{\bar{g}_{xy}} = \frac{\mathrm{COV}g_{xy}}{\sqrt{V_{\bar{g}_x} V_{\bar{g}_y}}}.$$

Cluster analyses

Genetic correlations between pairs of test sites ($r_{g_{xy}}$) were used as a measure of similarity. When $r_{g_{xy}}$ is close to unity, then the concerned sites are assumed to be more similar. By using $r_{g_{xy}}$, a phenogram of clustering of sites (from the unweighted pair-group method using arithmetic averages, UPGMA) was employed to illustrate the general pattern of similarities among the test sites (Sneath and Sokal 1973, pp 230–234).

Principal component analysis

PCA is a multivariate technique for detecting linear relationships among several quantitative traits without designating some as independent and others as dependent. The purpose of PCA in this study was to detect relationships among the test sites. As input data, each test site in this study is represented by 40 quantitative traits (clonal means). If there were only two traits (variables), the locations of the test sites could have

Table 2 Height means (in centimeters) of 40 *Picea abies* clones 17 years of age at seven test sites

Clone code (i = 1 to 40)	Test site clonal value (X_{ij}[a]) ($j = 1 \rightarrow 7$)[b]							Genotypic (clonal) value		
	P	H	L	S	B	M	K	$\sum_j n_{ij}$[c]	$\bar{X}_{i.}$	Rank[d]
4	675	544	500	824	542	712	615	91	630	25
9	706	528	502	785	508	728	558	91	616	31
11	617	468	413	690	583	711	454	77	562	38
15	671	498	524	802	552	751	476	91	611	32
18	650	566	505	810	609	823	545	91	644	19
26	811	573	554	904	644	914	579	98	711	05
37	763	621	584	899	696	789	669	98	717	04
41	788	643	545	945	735	935	533	91	732	02
42	736	601	522	843	494	829	579	105	658	16
45	668	575	531	853	637	797	536	98	657	17
46	689	549	471	870	582	836	609	105	658	15
50	682	549	456	865	646	799	519	112	645	18
66	649	526	542	769	570	755	545	91	622	27
87	760	640	453	905	606	886	542	84	685	09
88	726	592	474	779	597	789	455	91	630	24
90	674	600	501	799	539	774	566	98	636	21
94	679	540	434	776	466	758	498	112	593	33
95	738	618	569	889	706	885	666	105	725	03
98	723	569	449	790	619	756	561	98	638	20
101	617	605	552	970	650	882	492	28	681	12
103	695	634	530	956	689	847	553	98	701	08
104	636	567	447	725	537	727	512	105	593	34
107	834	596	449	960	680	890	549	91	708	06
112	632	526	479	745	371	713	510	98	568	37
113	720	460	556	791	568	796	543	84	633	23
115	728	543	475	795	483	775	543	105	621	29
116	657	540	474	856	528	779	519	84	622	28
118	631	475	465	780	529	771	471	98	589	35
123	745	707	576	847	750	939	676	105	748	01
125	682	587	533	727	609	742	563	105	634	22
142	685	576	471	768	623	718	498	91	620	30
143	682	575	487	866	711	811	659	105	684	10
145	786	589	557	836	534	804	594	98	671	14
152	723	609	498	920	593	855	529	98	675	13
173	610	502	366	820	540	696	397	63	561	39
181	706	499	382	767	568	737	448	91	587	36
188	751	641	586	851	632	899	584	112	706	07
189	634	467	446	685	488	652	342	84	531	40
196	655	571	483	776	675	769	476	77	629	26
197	769	525	501	854	692	893	536	84	682	11
$\bar{X}_{.j}$[e] =	699	565	496	827	595	798	536	N = 3,731	645[f]	

[a]Mean of a clone at a given test site
[b]Test sites: P, Paderborn; H, Holzminden; L, Lautenthal; S, Syke; B; Binnen; M, Medingen; K, Kattenbühl
[c]$n_{ij} = \sum_j n_{ij}/7$ = Number of individuals observed for clone i at test site j
[d]The tallest clone has the rank of 1
[e]Site value (mean of all clones at a test site)
[f]$\bar{X}_{..}$ = Overall mean

Table 3 Variance-covariance (var-covar) matrices between site pairs in height in 17-year-old trees. Below diagonal: genetic var-covar; above diagonal: phenotypic var-covar. Bold values on the diagonals are the $V_c = V_g$ values for the test site concerned

Site[a]	P	H	L	S	B	M	K
P	3,144**1,717**	1,554	1,158	2,115	1,953	2,661	1,967
H	1,559	2,997**2,338**	1,506	2,401	2,581	2,715	2,299
L	1,245	1,447	2,795**2,117**	1,526	1,720	2,071	2,376
S	2,379	2,177	1,285	5,134**3,759**	3,408	4,075	2,269
B	2,047	2,672	1,806	3,526	6,729**5,171**	3,872	2,460
M	2,730	2,714	1,949	3,888	3,936	5,101**3,865**	2,630
K	1,873	2,262	2,412	2,249	2,692	2,602	4,854**3,546**

[a]For test site abbreviations, see Table 2

easily been plotted in a two-dimensional space. However, there are 40 variables, and thus it is impossible to plot all of the variables simultaneously to locate relative positions of the test sites within multidimensional space. That is where PCA helps the researcher. PCA reduces the number of variables to a few principal components

such that data can be summarized into two or three dimensions. Subsequently, the plots of the first few principal components are valuable tools in illustrating relationships among the groups concerned (SAS 1987; Johnson and Wichern 2002).

As input data in PCA, we used clonal means arranged in an $r \times c$ table, where r denotes the rows (test sites = 7) and c denotes the columns (clones = 40). In other words, the mean of each clone was considered to be a separate trait at each test site, so that each test site had 40 variables to be evaluated by means of PCA. The $r \times c$ table is a transposed version of the data in Table 2. At the end of PCA, 40 principal components, and the corresponding eigenvectors and eigenvalues, were computed based on correlation matrix (SAS 1987).

Type-B ecovalence

When different genotypes are tested in a range of specific environments, generally the contribution of each genotype (ecovalence) to the total interaction sum of squares is estimated (Wricke 1962; Becker and Léon 1988; Karlsson et al. 2001; Isik and Kleinschmit 2003). Ecovalence obtained this way will be designated as Type-A ecovalence (W_i). *Where the emphasis is on the role of environments* rather than of genotypes, a second type of ecovalence can also be estimated. The ecovalance thus obtained will be designated as Type-B ecovalence (W_j), which is an analogous terminology to that proposed by Burdon (1977) on genetic correlations between test sites.

Type-B ecovalence was estimated using Eq. 4 and the data in the $r \times c$ table (Table 2).

$$W_j = \sum_i (X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})^2, \quad (4)$$

where W_j is the contribution of the test site j to the total interaction sum of squares (Type-B ecovalence); X_{ij}, the mean value of clone (group) i at site j; $\bar{X}_{i.}$, the overall mean value (or genotypic value) of clone i (where: $i \rightarrow 1$ to 40); $\bar{X}_{.j}$, the overall mean value of environment j (where: $j \rightarrow 1$ to 7); $\bar{X}_{..}$, the grand mean (mean of all clones at all sites).

The W_j values were calculated in a stepwise process (STEX process). In the first step, when all seven test sites are included in the data set (i.e. when $s = 7$), we estimated the W_j values by Eq. 4. Then, the site with the highest W_j value was considered to be the most interacting site (MIS). In the second step, the MIS was dropped from the data set, so that $s = 6$. Estimations of W_j values were also made, and the MIS was detected similarly within the six-site-combination. Such stepwise exclusions of the most interacting sites were carried out until there remained the last two sites. This procedure is called the STEX (stepwise exclusion) process, and it helps to detect the interactive behaviour of test sites relative to each other and determines the combinations of various test sites that are least interacting and more similar.

Genotypic selectivity of a test site

The same data (Table 2) that we used for Type-B ecovalence was also used for genotypic selectivity. *When the main concern is the genotypes*, the general application is that the value of each genotype is plotted against the *environmental mean* (Finlay and Wilkinson 1963; Owino 1977; St Clair and Kleinschmit 1986; Li and McKeand 1989; McKeand et al. 1990). The slope of the regression line (b_i) thus obtained measures the *environmental sensitivity* of the genotype (Falconer 1981, p 123; Isik and Kleinschmit 2003). However, *when primary attention is directed towards the role of the test environments* (j) rather than that of the genotypes (i), the mean value of each genetic group at a given test environment (on the Y axis) can be plotted against the *genotypic values* (on the X axis) [the genotypic value is defined as the overall mean of a given genetic group as it is expressed in all environments (Falconer 1981, pp 123–124)]. The slope of the regression line (b_j) in such a plot measures the unit increase of site value per unit improvement in genotypic quality (Fig. 3, Table 6). We calculated b_j values for each test site and used the term *genotypic selectivity* to describe the reaction of a test site to a per unit increase in genotypic quality. Thus, sites with high regression slopes (i.e. $b_j > 1.0$) are highly selective, with a relatively high discriminating (or selective) power against poor genotypes, and relatively favor good genotypes. To test for equality among the regression slopes, we used the method described in Sokal and Rohlf (1995, pp 493–499). From the regression analyses that we used to calculate the b_j value, we also calculated the r^2 value (coefficient of determination) for each test site (SAS 1987).

Genetic gains

Based on the principles of correlated response (Falconer 1981, p 286; Burdon 1977), gain in environment y based on selection in environment x (i.e. $\Delta G_{y.x}$) can be predicted by:

$$\Delta G_{y.x} = i.h_{\bar{g}_x}.h_{\bar{g}_y}.r_{g_{xy}}.\sqrt{V_{\bar{g}_y}} \quad (5)$$

where i = selection intensity (in this study, $i = 1.365$, which corresponds to 20% selection).

Gain can also be estimated by (Burdon 1977):

$$\Delta G_{y.x} = ib\sqrt{V_{\bar{g}_x}} \quad (6)$$

where $b = \mathrm{COV}g_{xy}/V_{\bar{g}_x}$. Both Eqs. 5 and 6 were applied—and produced identical results in this study.

Efficiency of a site for selection

Assuming that the same intensity of selection was applied on both sites (i.e. $i_x = i_y$), the efficiency of selection at site x for planting at environment y ($E_{y.x}$) can be estimated by (Falconer 1981; Burdon 1977):

$$E_{y.x} = \frac{\Delta G_{y.x}}{\Delta G_{y.y}} = r_{g_{xy}}.(h_{\bar{g}_x}/h_{\bar{g}_y}). \quad (7)$$

According to Eq. 7 the efficiency will be equal to unity when x and y are the same site. $E_{y.x}$ is greater than unity only when $h_{\bar{g}_x}.r_{g_{xy}} > h_{\bar{g}_y}$. This is possible in certain combinations when heritability at test site x is substantially higher than that in planting environment, and/or when $r_{g_{xy}}$ is relatively large.

Results and discussion

Similarities among the test environments

In practice, similarities or dissimilarities among environments are generally measured by certain environmental factors expressed in various climatic, edaphic and/or geographic variables. For breeding and selection purposes, however, the main question to be asked is how the group of genotypes of interest "read and interpret" the environments concerned. In view of this question, it is possible to measure the extent of similarities among test environments as they are interpreted by common genotypes grown on them. We have estimated and discussed five criteria below as a means of measuring the level of similarities among the test sites.

Genetic correlations

Type-B genetic correlations in this study were always greater than the corresponding phenotypic correlations (i.e. correlation of clonal means) (Table 4). This is often observed in such studies (see Lambeth et al. 1983; Riemenschnieder 1988; Gwaze et al. 2001). Out of 21 combinations of site pairs, 19 showed genetic correlations greater than 0.60. The remaining (and the lower values) were of the test site pairs L-S (0.46) and L-B (0.55). Site L showed relatively low $r_{g_{xy}}$ values with all of the other test sites except site K ($r_{g_{KL}} = 0.88$). On the other hand, site M showed a rather high similarity with the other test sites, the lowest level being with site L ($r_{g_{ML}} = 0.68$), and the second lowest with site K ($r_{g_{MK}} = 0.70$).

In 2 out of 21 cases, genetic correlation coefficients were found to slightly exceed the theoretical parameter space (i.e., $r_{g_{MP}} = 1.06$, and $r_{g_{MS}} = 1.02$) (Table 4). Such out-of-bound values often occur when very low estimates of genetic variance are obtained in one or two of the environments concerned (Lu et al. 2001). They also occur when population heritabilities are low (Koots and Gibson 1996). In our study, site P had a low genetic variance (less than half of that for site M) (Table 3) and, consequently, relatively low heritability (Table 6). Similarly, each of site M and site S had lower genetic variances than their common covariance.

The Type-B genetic correlations obtained in our study are relatively high compared to those reported in other studies. Johnson (1997) reported that Type-B genetic correlations for each breeding zone for Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco] ranged from 0.42 to 0.84, indicating that values 0.7 or higher between test sites are reasonable within a breeding zone. Average genotypic correlations across 11 test sites for total height growth in *Picea abies* in Denmark and Sweden were estimated to be 0.54 (Karlsson et al. 2001). In slash pine (*Pinus elliottii* Engelm.) Type-B genetic correlations between site pairs for height and diameter ranged from 0.46 to 0.67 (Hodge and White 1992). The relatively high Type-B genetic correlations in our study may partly be due to the fact that our test sites represent a relatively narrow geographical range compared to the test sites of the species investigated in the other studies mentioned above.

Table 4 Correlation coefficients (corr coeff) between site pairs with respect to height at age 17 years. Below diagonal, genetic corr coeff; above diagonal, Pearson corr coeff

Site[a]	P	H	L	S	B	M	K
P		0.51	0.39	0.53	0.43	0.66	0.50
H	0.78		0.52	0.61	0.58	0.69	0.60
L	0.65	0.65		0.40	0.40	0.55	0.65
S	0.94	0.73	0.46		0.58	0.80	0.46
B	0.69	0.77	0.55	0.80		0.66	0.43
M	1.06	0.90	0.68	1.02	0.88		0.53
K	0.76	0.79	0.88	0.62	0.63	0.70	
AGC[b]	0.81	0.77	0.64	0.76	0.72	0.87	0.73

[a]For test site abbreviations, see Table 2
[b]Average genetic correlation with the other sites

Cluster analysis

To illustrate the general pattern of similarities among the test sites, we drew up a phenogram based on UPGMA cluster analysis of genetic correlation coefficients between pairs of test sites (Fig. 1). At the 0.80 similarity level there were three distinct groups: sites P-M-S-H formed the first; site B, the second; sites K-L, the third group. At the 0.78 level, there were two groups (P-M-S-H-B, and K-L); at the 0.64 level, all of the sites were similar.

Principal component analysis

The eigenvalues from PCA in this study indicated that the first three principal components provided a rather good summary of the data: the first component accounted for 91.9% of the variance, two components accounted for 95.2% and three components explained 96.8% (Table 5). The first six components accounted for all (100%) of the variance.

The first PC can be interpreted as a measure of overall means of the clones, since the first eigenvector shows approximately equal loadings on all variables The loadings remained between a narrow range of 0.142 (the lowest) and 0.163 (the highest). In other words, the first eigenvector shows that the means of all the clones are

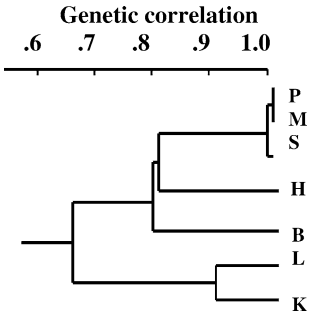

Fig. 1 Cluster tree (UPGMA) for the test sites based on clonal heights at age 17 years. The test sites were: *P* Paderborn, *H* Holzminden, *L* Lautenthal, *S* Syke, *B* Binnen, *M* Medingen, *K* Kattenbühl

equally weighted. The second eigenvector has relatively high positive loadings on the means of five clones (clones nos. 112, 145, 42, 115 and 9), and relatively high negative loadings on one clone (clone no. 196). The common features of the positively loading clones were that each had the lowest rank of mean height at site B than they did at all the other sites. The second PC, then, can be interpreted as a measure of the clones that showed their lowest relative performances at site B (thus they could be called "anti-site-B clones"). The third component (which explained only 1.6% of the variance) can be interpreted as a measure of clones that showed their highest relative performance at site K (clones 143, 4 and 46, which may be denoted "pro-site K" clones). Six of the nine clones that are impacting on the second and the third PCs in this study were among the 16 clones that exhibited relatively larger ecovalence values than those clones with average (i.e. 2.5 %) ecovalence values (Isik and Kleinschmit 2003).

The striking features of the graphical plots of PCs are that site B was an outlier based on the second component and that site K was similar for the third component (Fig. 2a,b). The former may be due to five anti-site-B clones that made their lowest relative performance at site B, and the latter may be due to three pro-site-K clones. Also, both sites exhibited relatively high within-site CV values (Table 6). The patterns of proximity among the test sites on the first three PCs are quite similar to the relationships obtained from the cluster analyses in Fig. 1.

Table 5 Some parameters related to the first three principal components (PCs)

Test site[a]	PC-I	PC-II	PC-III
Coordinates of the test sites on PCs			
P	2.54	0.78	−0.98
H	−3.78	0.01	−0.40
L	−6.95	0.68	−0.56
S	8.42	0.21	0.64
B	−2.38	−2.48	0.08
M	7.10	−0.08	−0.17
K	−4.95	0.87	1.39
Eigenvalues and their weights on PCs			
Eigenvalue	36.75	1.34	0.63
Proportion	0.919	0.033	0.016
Cumulative	0.919	0.952	0.968

[a]For test site abbreviations, see Table 2

Contribution of test sites to interaction sum of squares

When all of the test sites ($s = 7$) were included in the analyses, the proportion of V_{cs} equal 2.57% of the total variance (V_{cs}, which is the variance due to G×E, was estimated according to ANOVA model described in Isik et al. 1995, Table 2a). In our seven-site combination, the site with the highest W_j value [i.e. the most interacting site (MIS)] was site B, being responsible for 23.4% of the total interaction sum of squares (Table 6). The sequence of the other sites, in decreasing order, was K, S, L, P, M and H. When the STEX process was applied by dropping MIS (site B) in the first step, the sequence of sites with respect to their interacting behaviour did not alter. However, at the second step (i.e. when K is dropped), the sequence of sites in their contribution to total interactions changed, with the new order being L,

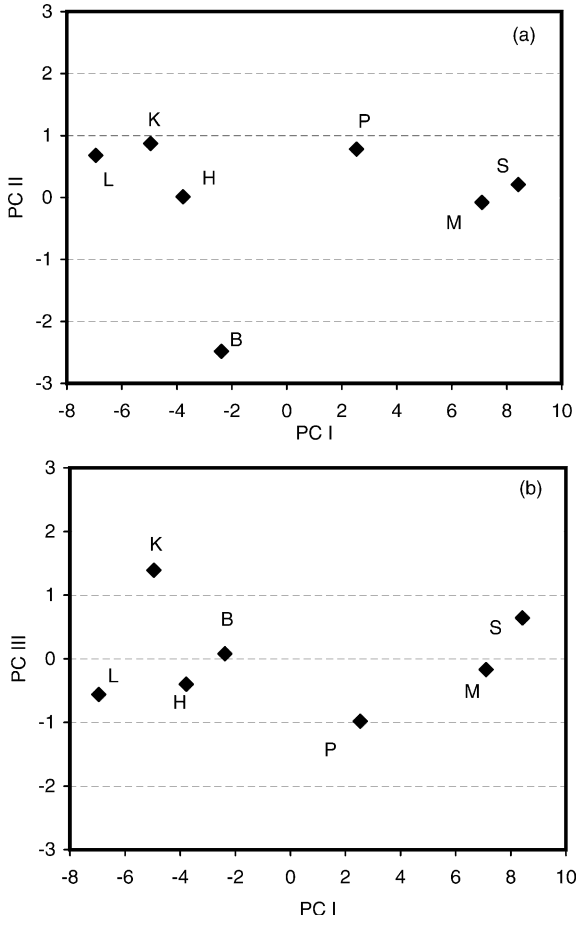

Fig. 2 Plots of first three principal components as a means of illustrating relationships among the seven test sites. **a** PC-II vs. PC-I, **b** PC-III vs. PC-I. For test site abbreviations, see Fig. 1

Table 6 Test site means ($\bar{X}_{.j}$), within-site coefficient of variations (CV), regression statistics (b, r^2), heritabilities (h^2) and Type-B ecovalence (W) values for the test sites

Test site[a]	$\bar{X}_{.j}$	CV_1^b	CV_2^b	b_j^c	r_j^{2d}	h_g^2	$h^2{}_b$	Type-B ecovalence[e]	
								W_j absolute	W_j %
P	699	8.02	18.14	0.788 cd	0.520	0.584	0.096	63,555	13.28
H	565	9.69	15.90	0.868 bcd	0.665	0.793	0.223	40,823	8.53
L	496	10.67	17.58	0.713 d	0.479	0.787	0.217	65,363	13.65
S	827	8.66	13.89	1.131 abc	0.658	0.792	0.222	70,100	14.65
B	595	13.79	26.18	1.228 ab	0.593	0.740	0.176	112,168	23.43
M	798	8.96	15.05	1.252 a	0.810	0.780	0.211	44,330	9.26
K	536	12.96	22.99	1.019 abcd	0.565	0.752	0.185	82,358	17.20
Mean	645						$\sum W_j$	47,8697	100.00

[a]For test site abbreviations, see Table 2
[b]CV_1 and CV_2 are based on clonal site means and individual observations, respectively
[c]Regression slopes with the same letter are not significantly different from each other at the 5% level
[d]Coefficient of determination (the square of correlation coefficients)
[e]As calculated when all seven test sites are included (in seven-site combination)

S, P, H and M. It appears that the presence of of K had a masking effect on the interactive behaviour of sites L and H, since following the dropping of K their relative rank orders increased. At the same time, the presence of site K provoked the interaction contributions of sites S and M. Similarly, when L was dropped in the next STEX process, the sequence of the remaining sites changed and took the order of P, S, H and M. In this latter case, site L had a masking effect on P, since following the dropping of L, site P became the most interacting site. The presence of L in the group also provoked the interaction contribution of site S, since following the dropping of site L, site S exhibited a relatively lower interaction. When P was dropped in the next step, the order of the sites was H, S and M, with site H being the most interacting within this combination. The two sites that remained in the final step (i.e., S and M) were the least interacting site combination.

The STEX process helps to detect the interactive behaviour of test sites relative to each other. The presence of certain sites in the data set either moderates or provokes the interactive behaviour of certain other sites. For example, site K was a *moderator site* for L and H, but at the same time it was an *instigator site* for M and S. Similarly, site L (which has a coalition with K) was also an instigator site for S and moderator for P. Site P was a moderator site for site H. In summary, the STEX process showed that when site M was within a combination, the least interacting site combinations and the corresponding amounts of V_{cs} were:

1. the seven-site combination B-K-L-P-H-S-M (V_{cs} = variance due to G×E = 2.57%);
2. the six-site combination K-L-P-H-S-M (V_{cs} = 1.98%);
3. the five-site combination L-S-P-H-M (V_{cs} = 1.72%);
4. the four-site combination P-S-H-M (V_{cs} = 1.15%);
5. the three-site combination H-S-M (V_{cs} = 1.17%);
6. the two-site combination S-M (V_{cs} = 0.00%).

Shelbourne (1972) proposed that, as an approximate rule of thumb, when the ratio of the interaction variance to the group variance (i.e. here V_{cs}/V_c) exceeds 0.5, there is a serious G×E, and if it is over 1.0, the breeder may have to develop separate breeding zones. For the seven-site combination in this study, this ratio was about 0.27 for height, and 0.30 for diameter at breast height (dbh) (Isik et al. 1995). For the other site combinations, the ratios were even smaller, all of which reflect the rather minor amount of clone × site interaction variance. Lindgren (1993) stated, along this line, that the easiest and perhaps the best strategy is to ignore interaction and use clones based on their average performance only.

The patterns of similarity among the test sites based on three methods (i.e. genetic correlations, PCs and W_j values) were closely related to each other (Table 4, Figs. 1, 2). This is expected because each of these three methods utilizes the same covariance matrix. These patterns also showed close associations with the soil and geographic variables (see Test sites and the experimental design in Materials and methods) of the test environments: two low-elevation sites (S, M) and one mid-elevation site (P), each with medium to very good soil conditions, were quite similar in their reactions to the genotypes tested. One high-elevation site (H) with medium to good soil conditions and one low-elevation site (B) with poor soil conditions were relatively closer to the P-S-M group, with site B somewhat outlying on the second PC. The remaining two sites (K-L) were higher elevation sites; these were quite similar in their interactive behaviours with the other test sites and also in their reactions with the tested genotypes. These two higher elevation sites thus formed a distinct group from the rest.

Genotypic selectivity of test sites

A good test site for screening genotypes must allow sufficient efficiency of selection through the expression of good genotypes. Thus, sites with high b_j values exhibit high *genotypic selectivity* and may be considered to be

good test sites for detecting and selecting good genotypes. Test site M exhibited the highest value for genotypic selectivity, followed by sites B and S (Table 6, Fig. 3). According to our tests for equality of the regression slopes, these three sites were significantly different from site L. Furthermore, regression slopes of sites M and B were also different from that of site P (Table 6). Test sites L, P and H, had a relatively low genotypic selectivity. Test sites with a low genotypic selectivity do not "encourage" good genotypes and, at the same time they have a tendency to upgrade poor genotypes. Therefore, such sites with $b_j < 1.0$ are not desirable and may be misleading as testing sites. It should also be noted that test sites with higher r^2 values are more reliable with respect to their genotypic selectivity and thus more desirable than those with low r^2 values.

The following relationships between genotypic selectivity and a number of other site variables also need to be noted: genotypic selectivity (b) is highly correlated with genetic variation (i.e. V_g = bold values on diagonals in Table 3) of sites ($r = 0.91$, $P < 0.0039$). It also tends to be positively correlated with average genetic correlation with the other sites (i.e. AGC = values in the last row in Table 4) ($r = 0.42$, $P < 0.34$). These two variables (i.e. V_g and AGC) together predict genotypic selectivity with an $r^2 = 0.98$. The prediction equation obtained from our study was: $b = -0.37131 + 0.0001587\ V_g + 1.1341$ AGC.

It should also be noted that there is a trend for large b's with taller test sites ($r = 0.54$, $P < 0.21$). This trend appears to be common because genetic gain tends to be better stated as a percentage gain rather than an absolute gain. When any two sites have similar percentage gains, the taller site has more absolute gain; therefore it has a larger b value.

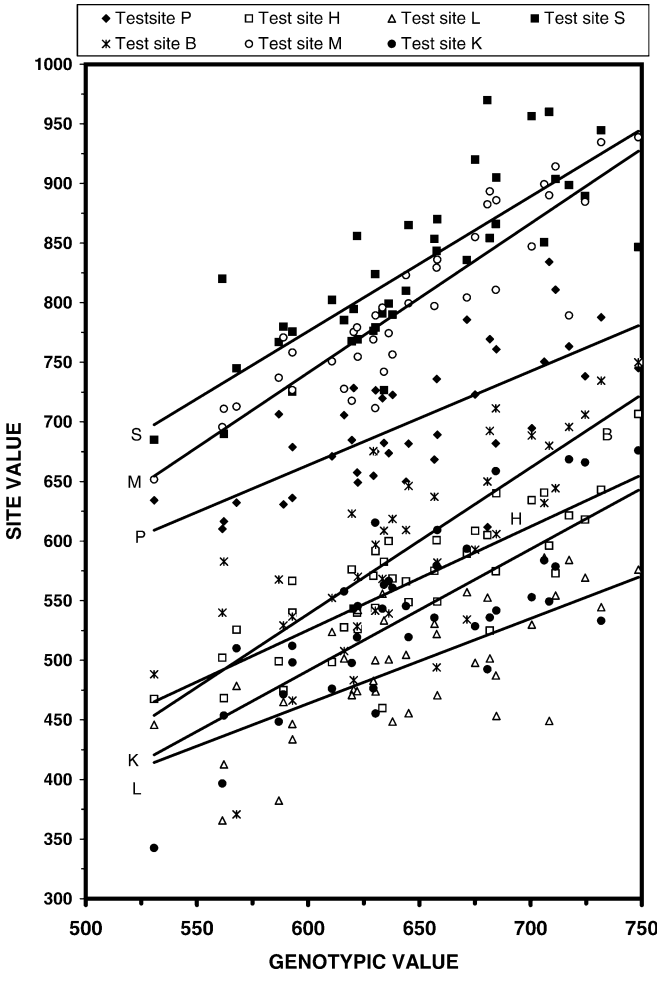

Fig. 3 The regressions of test site means on genotypic value (in centimeters). See Fig. 1 for test site abbreviations

Genetic gains and efficiency levels

The usual application of test results in plant breeding is that the selection of phenotypes is made directly at a test environment that is representative of planting sites to be planted in the future. This is a form of direct selection (Falconer 1981). By considering the principles of correlated response between pairs of environments, the selection of phenotypes can also be made at testing site x while planting is done at environment y (indirect selection). We estimated the expected gains in planting environments (y) for each testing site (x) (Table 7). Expected gains at planting environments are usually the greatest (or close to the greatest) when testing and selection is carried out within the same environment [values along the diagonal (in italics) in Table 7]. A notable exception to this rule-of-thumb was the gain in environment P (when site M is used as the testing site), where both heritability and variance were low.

When we wish to compare the expected gains within and between planting environments, the efficiency concept as presented in Table 8 is a much better parameter than the absolute amounts of gains as in Table 7. The efficiency of indirect selection relative to that of direct selection may be expressed as the ratio of expected gains at the sites concerned. The efficiency of a testing site is equal to unity (values along diagonal in Table 8) when testing and selection is made within the same environment . The highest value in a row in the table identifies the best testing site for the planting site in the row heading. For example, for environment P the best testing site was not itself, but site M. When selection is done at a single site for planting at all environments, the best site for testing is site M (average gain = 9.3%, and average $E_{y.M} = 0.918$). Site L had the lowest efficiency level.

The efficiency level of a testing site (in a given column in Table 8) also reflects its similarity level to the corresponding planting environment. This is expected because genetic correlation is a major component of efficiency (Eq. 7). Based on efficiency, the degree of similarity of testing site M to other environments was in the order of P, S, B, H, K and L. Testing site S also exhibited the same order of similarity level: P, M, B, H, K, L.

Table 7 Expected genetic gains (in centimeters)[a] at planting site y based on selection at testing site x

Planting site, y[b]	Testing site, x[b]						
	P	H	L	S	B	M	K
P	*43.2* (*6.2*)	39.2 (5.6)	32.8 (4.7)	47.1 (6.7)	33.4 (4.8)	52.9 (7.6)	37.2 (5.3)
H	39.3 (6.9)	*58.8* (*10.4*)	38.1 (6.7)	43.1 (7.6)	43.6 (7.7)	52.6 (9.3)	44.9 (8.0)
L	31.4 (6.3)	36.4 (7.3)	*55.7* (*11.2*)	25.2 (5.1)	29.4 (5.9)	37.8 (7.6)	47.9 (9.7)
S	59.9 (7.2)	54.7 (6.6)	33.8 (4.1)	*74.5* (*9.0*)	57.6 (6.9)	75.4 (9.1)	44.7 (5.4)
B	51.5 (8.7)	67.1 (11.3)	47.4 (8.0)	69.8 (11.7)	*84.4* (*14.2*)	76.3 (12.8)	53.5 (9.0)
M	68.7 (8.6)	68.2 (8.5)	51.3 (6.4)	77.1 (9.6)	64.2 (8.0)	*75.0* (*9.4*)	51.7 (6.5)
K	47.2 (8.8)	56.8 (10.6)	63.4 (11.8)	44.6 (8.3)	43.9 (8.2)	50.4 (9.4)	*70.5* (*13.1*)
Mean	48.74 (7.5)	54.46 (8.6)	46.07 (7.6)	54.50 (8.3)	50.94 (8.0)	60.08 (9.3)	50.07 (8.1)

[a] Values in parenthesis are percentage gains. $\%G_{y.x} = (\Delta G_{y.x} \times 100)/\bar{X}_{.y}$, where $\bar{X}_{.y}$ = the overall mean value at site y

[b] For test site abbreviations, see Table 2

The efficiency of a testing site was larger than unity in 4 out of 49 cases (Table 8) ($E_{P.M} = 1.223$, $E_{P.S} = 1.089$, $E_{S.M} = 1.012$ and $E_{M.S} = 1.027$). In the first and second cases, heritability in site P was much lower than heritabilities in sites M and S. In addition, genetic correlations between the two respective pairs of sites were closer to unity. In the third and the fourth cases, heritabilities at sites M and S were nearly the same, and genetic correlations between the two sites were greater than one.

Based on the results discussed above, site M was identified as the best testing site, followed by sites H and S. When the trees were tested at age 10 years with the same experiments, the best site for testing was site M, followed by S and B (St Clair and Kleinschmit 1986). Apparently, as the trees get older, site B exhibited higher interactions with the other test sites. This is also reflected on the relatively low genetic correlations of site B with most of the other sites (Table 4). Although site B has a very high genotypic selectivity, it has very high interaction contribution (64% more than average), high CV value (Table 6) and relatively low $E_{y.B}$ (Table 8). Therefore, it is not preferred as a reference testing site. Site K had an average genotypic selectivity ($b_j = 1.0$) but high W_j, high CV, low r^2_j and relatively low $E_{y.K}$. In addition, site K has been an instigator site for the sites M and S. All these features made site K undesirable as a testing site.

Table 8 Efficiency of test sites ($E_{y.x}$) when testing and selection of genotypes are performed at site x for planting at site y

Planting site, y[a]	Testing site, x[a]						
	P	H	L	S	B	M	K
P	1.0	0.906	0.757	1.089	0.772	1.223	0.861
H	0.667	1.0	0.647	0.733	0.742	0.894	0.764
L	0.562	0.653	1.0	0.456	0.528	0.678	0.860
S	0.804	0.734	0.453	1.0	0.773	1.012	0.600
B	0.609	0.795	0.561	0.826	1.0	0.903	0.633
M	0.916	0.909	0.683	1.027	0.857	1.0	0.689
K	0.669	0.806	0.899	0.632	0.622	0.715	1.0
Mean	0.747	0.829	0.715	0.823	0.756	0.918	0.773

[a] For test site abbreviations, see Table 2

How frequently do the genotypes selected on test site M occur on the other sites? When site M is used as the testing site, the best eight clones and their rank orders are given in Table 9. The orders of the best eight clones selected according to the CSP_{i5} index (Isik and Kleinschmit 2003) are also presented in Table 9 [CSP_{i5} is an index that considers selection for both high stability and high performance simultaneously (Nassar et al. 1994)]. Of the best eight clones detected by the CSP_{i5} criteria, all were also among the best eight clones in overall means (Table 9). Of these, six (i.e. 75%) were also among the best eight at testing site M. It appears that selection by the GSP_{i5} index and phenotypic selection at testing site M produce nearly similar results. The frequency of occurrence at other sites of the best eight clones detected in Table 9 is presented in Table 10. Lindgren (1993) suggested that in case "certain clones perform better on a well-defined type of site, these clones could be deployed on site types where they perform best". With selection intensity $i = 1.365$, the average gain from selecting at site M and planting at all suitable environments would be 9.3 % (Table 7). The corresponding gain based on selection from the GSP_{i5} index was estimated to be 10.0 %. These values represent gains only from additional clonal selection from previously selected clones, and it does not include gains already achieved from earlier truncation selection nor from gain obtained relative to seedlings (see Isik et al. 1995).

Conclusions

1. Each of the three separate methods (i.e. Type-B genetic correlations, PCA and Type-B ecovalence) identified a closely related pattern of similarity among the test sites. There were two distinct groups: sites M-S-P formed one group, and sites K-L formed the other. Sites H and B were closer to the first group. This pattern also showed close associations with the soil and geographic variables of the test environ-

Table 9 Best eight clones and their rank orders based on different selection criteria

Selection based on:	Rank orders and clone ID numbers (tallest the 1st)							
	1	2	3	4	5	6	7	8
Test site M	123	41	26	188	197	107	87	95
GSP_{i5} index	123	95	37	188	41	103	26	107
Overall mean	123	41	95	37	26	107	188	103

Table 10 Frequency of occurrence of clones in Table 9 as being among the best eight clones at each of the other test sites

Selection based on:	Test sites[a]							Mean
	P	H	L	S	B	M	K	
Test site M	6/8	5/8	4/8	4/8	5/8	8/8	3/8	35/56
GSP_{i5} index	5/8	6/8	5/8	5/8	6/8	6/8	4/8	37/56

[a] For test site abbreviations, see Table 2

ments. The efficiency level of a given testing site also reflected its similarity level to the corresponding planting environments.

2. Empirical data indicated that a desirable test site should have the following features: (1) low interaction behaviour, (2) low coefficient of variation, (3) high genotypic selectivity, (4) high coefficient of determination and (5) high efficiency of expected gain. In other words, overall gain is maximized when a test site: (1) has a high genetic variance (which relates to genotypic selectivity and coefficient of determination); (2) has a low environmental variance (which relates to coefficient of variation); (3) is well correlated with the deployment zone (which relates to low interaction behaviour and a high efficiency of expected gain). Based on these criteria, site M (Medingen) was determined to be the best test site for screening genotypes to plant at other environments. The present low level of G×E can be ignored in selection and deployment decisions. One broadly adapted vegetative propagation population of *Picea abies* will serve the potential planting environments in northern Germany.
3. The criteria indicated in item 2 above would help to identify good test sites for screening good genotypes. However, these criteria could be obtained only *after* the tests are performed. The practical question that remains to be answered is how to identify such sites *before* the tests are established.

Acknowledgements The contributions from the following persons and institutions are gratefully acknowledged: Jochen Schmidt and Joseph Svolba helped at the initial stages of the clonal propagation and at the establishment and maintenance of the test sites; Dr. Tony (C.J.A.) Shelbourne, Dr. Steve McKeand and an anonymous reviewer provided constructive criticisms on the manuscript; the Alexander von Humboldt Foundation supported Kani Isik during the evaluation of the study; Dr. Wilfried Steiner provided a peaceful and friendly working environment at the LSFRI, Department of Forest Genetic Resources, Escherode, Germany, where the study has been performed.

References

Adams WT, White TL, Hodge GR, Powell GL (1994) Genetic parameters for bole volume in longleaf pine—large sample estimates and influences of test characteristics. Silvae Genet 43:357–366

Becker HC, Léon J (1988) Stability analysis in plant breeding. Plant Breed 101:1–23

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

Dieters MJ, White TL, Hodge GR (1995) Genetic parameter estimates for volume from full-sib tests of slash pine (*Pinus elliottii*). Can J For Res 25:1397–1408

Falconer DS (1952) The problem of environment and selection. Am Nat 86:293–298

Falconer DS (1981) Introduction to quantitative genetics, 2nd edn. Longman, London, New York

Finlay KW, Wilkinson GN (1963) The analyses of adaptation in a plant breeding programme. Aust J Agric Res 14:742–754

Gwaze DP, Bridgewater FE, Byram TD, Lowe WJ (2001) Genetic parameter estimates for growth and wood density in loblolly pine (*Pinus taeda* L.). For Genet 8(1):47–55

Hodge GR, White TL (1992) Genetic parameter estimates for growth traits at different ages in slash pine and some implications for breeding. Silvae Genet 41:252–262

Isik K, Kleinschmit J (2003) Stability related parameters and their evaluation in a 17-year old Norway spruce clonal test series. Silvae Genet 52:133–138

Isik K, Kleinschmit J, Svolba J (1995) Survival, growth trends and genetic gains in 17-year old Picea abies clones at seven test sites. Silvae Genet 44:116–128

Johnson GR (1997) Site-to-site genetic correlations and their implications on breeding zone size and optimum number of progeny test sites for coastal Douglas-fir. Silvae Genet 46:280–285

Johnson RA, Wichern DW (2002) Applied multivariate statistical analysis, 5th edn. Prentice Hall, New York

Karlsson B, Wellendorf H, Roulund H, Werner M (2001) Genotype × trial interaction and stability across sites in 11 combined provenance and clone experiments with *Picea abies* in Denmark and Sweden. Can J For Res 31:1826–1836

Kleinschmit J (1974) A program of large-scale cutting propagation of Norway spruce. N Z J For Sci 4:359–366

Koots KR, Gibson JB (1996) Realized sampling variances of estimates of genetic parameters and the difference between genetic and phenotypic correlations. Genetics 143:1409–1416

Lambeth CC, van Biujtenen JP, Duke SD (1983) Early selection is effective in 20-year-old genetic tests of loblolly pine. Silvae Genet 32:210–215

Li B, McKeand SE (1989) Stability of loblolly pine families in the southeastern US. Silvae Genet 38:96–101

Lindgren D (1993) The population biology of clonal deployment. In: Ahuja MR, Libby WJ (eds) Clonal forestry I. Genetics and biotechnology. Springer, Berlin Heidelberg New York, pp 34–49

Lu P, White TL, Huber DA (1999) Estimating type B genetic correlations with unbalanced data and heterogeneous variances for half-sib experiments. For Sci 45:562–572

Lu P, Huber DA, White TL (2001) Comparison of multivariate and univariate methods for the estimation of type B genetic correlations. Silvae Genet 50:13–22

McKeand SE, Li B, Hatcher AV, Weir RJ (1990) Stability parameter estimates for stem volume for loblolly pine families growing in different regions in the southeastern United States. For Sci 26:10–17

Nassar R, Léon J, Huehn M (1994) Tests of significance for combined measures of plant stability and performance. Biometric J 36:109–123

Owino F (1977) Genotype-environment interaction and genotypic stability in loblolly pine: II Genotypic stability comparisons. Silvae Genet 26:21–25

Riemenschnider DE (1988) Heritability, age-age correlations, and inferences regarding juvenile selection in jack pine. For Sci 34:1076–1082

SAS Institute (1987) SAS/STAT guide for personal computers, 6th edn. Cary, N.C.

Shelbourne CJA (1972) Genotype-environment interaction: its study and its implications in forest tree improvement. In: IUFRO Genet-SABRAO Joint Symp, Tokyo, p 27

Sneath PHA, Sokal RR (1973) Numerical taxonomy. WH Freeman, San Francisco

Sokal RR, Rohlf FJ (1995) Biometry: the principles and practice of statistics in biological research, 3rd edn. WH Freeman, New York

St Clair JB, Kleinschmit J (1986) Genotype-environment interaction and stability in ten-year height growth on Norway spruce clones (*Picea abies* Karst.). Silvae Genet 35:177–186

Wricke G (1962) Ueber eine Methode zur Enfassung der Oekologischen Streubreite in Feldversuchen. Z Pflanzenzuecht 47:92–96

Yamada Y (1962) Genotype by environment interaction and genetic correlation of the same trait under different environments. Jpn J Genet 37:498–509

Zobel BJ, Talbert J (1984) Applied forest tree improvement. Wiley, New York, p 505