CrossMark

ORIGINAL ARTICLE

Patterns of additive genotype-by-environment interaction in tree height of Norway spruce in southern and central Sweden

Zhi-Qiang Chen[1] · Bo Karlsson[2] · Harry X. Wu[1,3]

Abstract Genotype-by-environment (G × E) interaction for
tree height measured at ages 7 to 13 was investigated in 20
large open-pollinated progeny trials for Norway spruce (*Picea
abies* (L.) H. Karst.) in southern and central Sweden. Factor
analytic method using spatially adjusted data and a reduced
animal model was used to explore the pattern of G × E inter-
action. Extended factor analyses captured 93.0% of additive
G × E interaction variances using three factors. The mean
daily temperature less than 3.2 °C in May and June explained
27.8% G × E interaction, and it was moderately correlated
with the first factor, indicating that spring or autumn frost
weather condition could be a main driver for G × E interaction
in Norway spruce. Cluster analysis has divided 20 trials into
either 6 clusters or 3 clusters. Both sets of clusters reflected the
geography of the trials (climates) and the genetic connected-
ness among testing series, indicating that more trials with bet-
ter connectedness are required to examine whether current
delineation of breeding or seed zones is optimal. Parental sta-
bility using latent regression could be used to locate best par-
ents that have the highest breeding values and are highly sta-
ble across trials.

Communicated by J. Beaulieu

✉ Harry X. Wu
  harry.wu@slu.se; harry.wu@csiro.au

[1] Umeå Plant Science Centre, Department of Forest Genetics and Plant
  Physiology, Swedish University of Agricultural Sciences,
  SE-90183 Umeå, Sweden

[2] Skogforsk, Ekebo 2250, SE-268 90 Svalöv, Sweden

[3] CSIRO NRCA, Black Mountain Laboratory, Canberra, ACT 2601,
  Australia

Introduction

Based on the photoperiod and temperature gradients, the
Swedish breeding strategy for Norway spruce [*Picea
abies* (L.) Karst.] delineates the country into 22 breeding
zones (populations), with each including 50 founders
(Danell 1993). This multiple population strategy was consid-
ered to be effective in managing the genetic diversity and
adaptation to the future climate change while maintaining ge-
netic gain. The drawback of managing so many breeding pop-
ulations is the increased costs (Rosvall et al. 2011). Also, the
division of populations was mainly based on geo-climatic da-
ta. Similarly, the deployment of Norway spruce was also di-
vided into 9 seed orchard zones to maximize adaptation and
genetic gain.

To optimize the regionalization in Swedish breeding pro-
gram, it was suggested that all available Swedish Norway
spruce trials with an acceptable genetic connectivity should
be used to find the biological base of the current division of
breeding and seed zones. Berlin et al. (2014) analyzed 65 trials
and estimated genetic correlations across trial within each of
17 test series to explore genotype-by-environment (G × E)
interaction patterns. However, no convincing G × E interac-
tion patterns justifying the current division of breeding and
seed zones were observed.

Within the Norway spruce breeding program in Sweden,
progeny trials at 4 sites as a test series are used for each of the
22 populations. These trials are usually well connected for
genetic entries. However, genetic connectivity among test se-
ries is usually poor (Johnson 2004). Within each breeding
zone, a low or moderate G × E interaction, demonstrated with

 Springer

a high genetic correlation, was usually found using the trials from the same test series for tree height. This was attributed to the relatively small geographic area within each test series (Berlin et al. 2014). It is highly desirable to combine different test series within adjacent breeding zones for a joint analysis to explore the validity of current delineation of breeding zones using biological data.

There are many traditional ways to analyze G × E interactions and detect the patterns including (1) analysis of variance (ANOVA), (2) principal components analysis (PCA), and (3) linear regression (Freeman 1973). These methods are not always adequate to dissect a complex interaction structure with missing connections (Zobel et al. 1988). ANOVA can only test the significance of G × E and the relative size of G × E variance to genetic variance, but cannot provide any insight about its patterns; PCA only considers the multiplicative effects of G × E. The linear regression method combines additive and multiplicative components. Various stability parameters can also be estimated from such regression to examine stability of genotypes/families and test trials to infer the causes of the interactions (Finlay and Wilkinson 1963; McKeand et al. 2003; Wu and Ying 1998). For multiple site progeny trials in forestry, among-trial type B (Burdon 1977) genetic correlations were usually estimated using a mixed linear model (Baltunis et al. 2010). Since the 1980s, singular value decomposition (SVD) was employed to describing the G × E patterns (Gauch 1992), initially in agronomic crops using additive main effects and multiplicative interaction model (AMMI), and later in forestry. Recently, factorial regression using a mixed model approach (factor-analytic method—FA) was introduced to explore the G × E patterns for multiple environmental trials in crops (Burgueño et al. 2008; Fox et al. 2007; Mathews et al. 2007; Piepho 1998; Smith et al. 2001; Smith et al. 2015) to relate underlining factors to the causes of G × E interactions. Besides the linear and nonlinear fixed and mixed models using the parametric approaches to decompose the G × E interactions, there are also nonparametric methods to analyze G × E such as multivariate regression tree (MRT) (Hamann et al. 2011). For index selection with linear combination of many traits, G × E pattern for index is a function of the relative importance of each traits (weights) and traits themselves (Namkoong 1985). In forestry, multiple regression and response surfaces were also commonly used to detect a relationship between population variation and environmental gradients for inferring adaptive variation to describe nonlinear G × E (Campbell and Sorensen 1978; Rehfeldt 1983; Wang et al. 2010; Wu and Ying 2004).

Cullis et al. (2014) summarized three main advantages of FA model in multi-environment trials (MET) analysis: (1) the ability to estimate the unstructured variance-covariance matrix without the use of an excessive number of variance parameters when the number of MET increases; (2) less than full rank variance structure for the G × E effect could be fitted using an appropriate estimation algorithm; and (3) the FA model is a mixture of multiple regression and principle component analysis. It could be easily used to explain the nature and the extent of G × E interaction using graphical tools such as biplots (Kempton 1984), regression plots (Thompson et al. 2003), and heatmap for estimated genetic correlation matrix with rows and columns ordered using cluster analysis (Cullis et al. 2010).

The FA model is also becoming popular in forestry MET analysis (Costa e Silva and Graudal 2008; Costa e Silva et al. 2006; Cullis et al. 2014; Hardner et al. 2010; Ivković et al. 2015). Costa e Silva et al. (2006) successfully used FA models to analyze stem diameter at breast height (DBH) for 15 *Eucalyptus globulus* progeny trials.

Spatial analysis models have recently been widely used in estimating genetic parameters for tree species (Costa e Silva et al. 2001; Dutkowski et al. 2006; Fu et al. 1999). Ye and Jayawickrama (2008) reported that spatial analysis could generally increase heritability and improve the accuracy of estimates of genetic parameters. Smith et al. (2001) considered that using an FA model and adjustments of spatial local and global field trend will improve the results of MET analysis. The suggested approach has been used in MET analysis of radiata pine (*Pinus radiata*) and eucalypt hybrids (*Eucalyptus camaldulensis* Dehnh. × *E. globulus* Labill. and *E. camaldulensis* × *E. grandis*) in Australia (Hardner et al. 2010; Ivković et al. 2015). A clear G × E pattern and some causes of G × E patterns have been explored in the 20 full-sib radiata pine progeny trials (Ivković et al. 2015).

In this paper, we selected 20 relatively well-connected progeny trials from a total of 146 Norway spruce trials to explore G × E patterns in southern and central Sweden. The aim of this study are to (1) estimate additive genetic variance and heritability for height in 20 large trials within six test series in three seed orchard zones, (2) estimate genetic correlations between trials, (3) dissect the G × E patterns in southern and central Sweden, and (4) analyze stabilities of selected parents in these trials.

Materials and methods

Field trials and measurements

In six test series, 20 open-pollinated progeny trials were planted by the Swedish tree breeding organization Skogfork in southern and central Sweden from 1986 to 1996, within three seed orchard zones (Fig. 1 and Table 1). The detailed characteristics of the trials are listed in Table 1. Randomized complete block design (RCB) with single-tree plots was used in test series 1 and 2, and completely randomized design with single-tree plots was used in the other test series. These 20 large trials were selected because of their relatively good
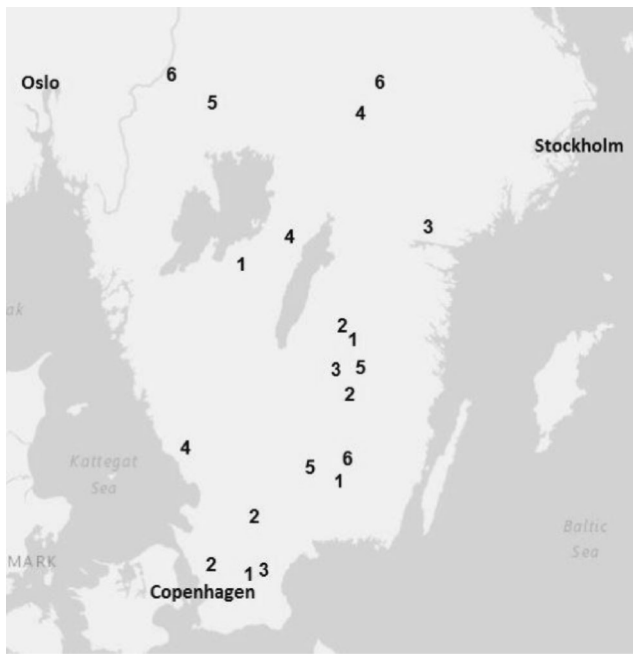
Fig. 1 Location of 20 Norway spruce progeny trials in six test series. The numbers correspond to the test series described in Table 1

parental connectedness. The number of female parents varied from 304 to 1389 among trials. The height of trees was measured at the ages of 7–13 years.

Climate data

Climate information for these 20 trials is shown in Table 2. Daily mean temperature and precipitation from the year of planting to the year of height measurement were downloaded from climate database (http://luftweb.smhi.se/). Mean annual precipitation (MAP), mean precipitation growing season (April to October) (MPGS), mean annual temperature (MAT), and mean annual heat sum (MAHT) above 5 °C were calculated. In order to estimate the influence of low temperature (related to frost damage) on bud burst and the consequent growth of Norway spruce trees, two climate indices were computed, including the mean temperature of days when daily mean temperature was below 3.2 °C in May and July (MTMJ) and the mean temperature of days when daily mean temperature was below 1.3 °C in September and October (MTSO). The MTMJ and MTSO represent spring and autumn cold indices, respectively.

Statistical analysis

Spatial analysis

Spatial analysis based on a two-dimensional separable autoregressive (AR1) model was used to fit the row and column directions for tree height data at each trial using ASReml

Table 1 Summary of trial information

Trial_ID	Site	Test series	Seed orchard Zone	Female	LE	LN	Alt	No_stems (alive)	Age	Year
F1021	Nöbbele	1	7	305	15.09	56.73	155	5550	13	1986
F1022	Hörby	1	8-9S	304	13.70	55.87	165	5733	13	1986
F1023	Remningstorp	1	7	304	13.58	58.47	110	4750	13	1986
F1024	Sund	1	7	188	15.23	57.86	235	6875	8	1986
F1059	Verum	2	8-9S	305	13.77	56.37	125	6280	13	1987
F1064	Ekebo	2	8-9S	305	13.11	55.95	80	7210	7	1987
F1067	Torpa	2	7	305	15.14	57.97	175	7032	13	1987
F1069	Alseda	2	7	304	15.25	57.40	160	6016	8	1987
F1145	Virå Bruk	3	7	1358	16.47	58.78	65	10,286	8	1990
F1146	Höreda	3	7	1358	15.04	57.61	225	12,639	7	1990
F1147	Erikstorp	3	8-9S	1360	13.93	55.90	155	11,776	7	1990
F1148	Hjuleberg	4	8-9S	1389	12.72	56.95	50	12,842	8	1990
F1149	Sätra	4	7	1387	14.33	58.70	145	12,388	12	1990
F1150	Lindesberg	4	6	1386	15.42	59.68	110	8853	7	1990
F1184	Ämtervik	5	6	732	13.12	59.76	110	7098	7	1991
F1215	Huseby	5	7	732	14.64	56.79	175	7254	7	1991
F1216	Ingatorp	5	7	730	15.42	57.63	190	6958	7	1991
F1267	Björneke	6	7	691	15.22	56.86	235	4260	11	1996
F1270	Gunnarskog	6	6	880	12.50	60.00	185	4907	11	1996
F1271	Dagarn	6	6	876	15.72	59.92	160	9925	8	1996

Female number of female parents in each trial, *LE* east longitude, *NL* north latitude, *Alt* altitude, *year* planting year

Table 2   Clusters, trial identification, test series, frost damage, age (age of frost damage measured), rotated trial loadings for first three factors, and calculated climate indices for each trial

Cluster	Trial_ID	Series	Frost damage (%)	Age (Year)	LD 1	LD 2	LD3	(mm)	MPGS (mm)	MAT (°C)	MAHS (°C)	MTMJ (°C)	MTSO (°C)
1	F1271	6	39.8	2	0.4	0.3	2.4	697	456	5.6	2279	2.3	-0.8
2	F1267	6			0.7	-5.2	2.7	831	487	7.5	2644	3.2	-0.2
	F1270	6	9.3	2	-0.7	-4.0	3.2	669	458	6.8	2496	2.8	-0.2
3	F1067	2			7.1	3.3	2.3	636	396	6.1	2301	2.6	-0.4
	F1147	3			3.5	1.1	0.8	610	373	6.4	2367	3.2	-0.6
	F1145	3			3.4	0.2	0.5	866	512	7.1	2493	3.2	0.2
	F1146	3	22.0	7	3.2	0.3	-0.4	636	421	6.6	2454	2.7	-0.4
4	F1069	2	11.6	8	3.9	2.0	-1.9	902	526	7.3	2576	3	-0.5
	F1059	2			4.6	6.2	1.2	648	387	7	2494	3.2	-0.6
	F1064	2			2.3	2.2	1.1	853	474	8.6	2883	3.2	-0.8
5	F1024	1			5.1	-0.9	1.6	920	559	5.3	2274	1.5	-1.1
	F1148	4			3.3	-0.9	1.1	637	386	6.3	2399	2.6	-0.7
	F1184	5			2.0	-1.0	1.6	767	461	5.7	2317	2.1	-1.3
	F1216	5	7	7	2.6	-1.3	2.7	790	466	6.8	2557	2.3	-1.3
	F1149	4			2.3	-0.6	1.3	744	495	5.1	2213	2.3	-1.1
	F1215	5	9	7	2.8	-0.8	2.1	749	452	6.6	2481	2.4	-0.3
6	F1021	1	5.5	8	7.0	-1.6	-1.6	744	451	6.6	2441	3.1	-0.6
	F1022	1	21.3	8	7.9	-1.4	-2.7	721	449	6.8	2498	3.2	-0.6
	F1023	1	26.6	8	4.5	-3.4	-3.4	601	383	6.8	2492	2.6	-0.6
	F1150	4	12	6	3.9	-4.0	-0.9	1089	590	7.4	2606	3.2	-0.3

Note: if cell in frost damage column is blank representing that the trials have not been recorded for frost damage. Frost damage was measured as the frequency of trees with visible frost damage (0 = no frost damage, 1 = visible frost damage)

LD1 trial loading for factor 1, *LD2* trial loading for factor 2, *LD3* trial loading for factor 3, *MAP* mean annual precipitation, *MPGS* mean precipitation growing season (April to October), *MAT* mean annual temperature, *MAHS* mean annual heat sum (>5 °C), *MTMJ* mean temperature when daily mean temperature less than 3.2 °C in May and July, *MTSO* mean temperature when daily mean temperature less than 1.3 °C in September and October

3.0 (Gilmour et al. 2009). Block, trial, and extraneous effects were estimated simultaneously and all significant block and spatial effects were removed from the raw data. The spatially adjusted data were used for the MET analysis in this paper.

Statistical model, variance components, and genetic parameters

The following reduced parental linear mixed model was used for multi-environment analysis:

$$\mathbf{y} = \mathbf{X}\tau + \frac{1}{2}\mathbf{F_p}\mathbf{u_{a_p}} + \mathbf{e} \tag{1}$$

where $\mathbf{y}$ is the vector of observations; τ is the vector of fixed effect, including the grand mean, trial, and age; $\mathbf{u_{a_p}}$ is the vector of random additive genetic effects for female parents (mother); and $\mathbf{e}$ is the vector of random residual term. $\mathbf{X}$ and $\mathbf{F}_p$ are known incidence matrices for fixed effects and female parents, respectively. Age effect was confounded with the trial effect; therefore, it was dropped out in the final model because of singularity. The random effects in the model are assumed to follow a multivariate normal distribution with means and variances defined by:

$$\mathbf{u_{ap}} \sim \mathbf{N}\left(\mathbf{0}, \sigma_\mathbf{a}^2\mathbf{A_{pp}}\right) \text{ and } \mathbf{e} \sim \mathbf{N}\left(\mathbf{0}, \sigma_\mathbf{e}^2\mathbf{I}\right)$$

where $\mathbf{0}$ is the null vector; $\mathbf{A_{pp}}$ is the numerator relationship matrix of female parents; $\mathbf{I}$ is the identity matrix, with order equal to the number of trees; and σ_a^2 and σ_e^2 are the additive and residual variances, respectively.

Narrow-sense heritability was calculated as follows:

$$\hat{h}_i^2 = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_p^2} = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_e^2} \tag{2}$$

where σ_a^2 is the estimate of additive genetic variance, σ_p^2 is the phenotypic variance, and σ_e^2 is the residual variance. All variance components were estimated from FA model.

Factor analysis

Smith et al. (2001) have found that a FA variance model provides a good parsimonious approximation to the unstructured genotype-by-environment covariance matrix using a few informative factors. The FA model can be viewed as arising from a multiplicative model for the genetic effect in each trial. The additive genetic effect i at trial j can be expressed as

$$u_{a_{pij}} = \lambda_{a_{1j}}f_{a_{1i}} + \lambda_{a_{2j}}f_{a_{2i}} + \ldots + \lambda_{a_{kj}}f_{a_{ki}} + \delta_{a_{ij}} \tag{3}$$

which includes a sum of k multiplicative terms. Each term is the product of a genetic effect ($f_{a_{ri}}$), which is known as a factor

score, and an environment effect ($\lambda_{a_{rj}}$), which is known as a loading. The k of the FA models is the number of factors (multiplicative terms), and we denote a FA model with k factors as a FAk model. The last term of $\delta_{a_{ij}}$ represents a lack of fit of the regression model, and so will be termed as a genetic residual.

The model 3 can be written in vector notation as

$$u_{a_p} = (I_m \otimes \Lambda_a)f_a + \delta_a \tag{4}$$

where Λ_a is the t (site) $\times k$ matrix of environment loadings, f_a is the $mk \times 1$ vector of additive genetic scores, and δ_a is $mt \times 1$ vector of genetic regression residuals with t and m representing number of trials and additive effects, respectively. f_a and δ_a are assumed to be independent and be distributed as multivariate Gaussian distribution with zero means and variance matrices given by

$$Var(f_a) = A_{pp} \otimes I_k \text{ and } var(\delta_a) = A_{pp} \otimes \psi_a$$

where ψ_a is a $t \times t$ diagonal matrix with a variance for each environment, which is known as a specific variance. These assumptions lead to

$$Var(u_a) = A_{pp} \otimes \left(\Lambda_a \Lambda_a^T + \psi_a\right) \tag{5}$$

So that genetic covariance matrix among trials: $G_a = \left(\Lambda_a \Lambda_a^T + \psi_a\right)$.

To determine the appropriate value of k, Cullis et al. (2014) suggested a measure similar to an R^2 goodness of fit value for a multiple regression. Thus, it can define, for each trial, a percentage of additive genetic variance accounted for by the k multiplicative terms as follows:

$$v_{a_j} = 100 \sum_{r=1}^{k} \lambda_{a_{rj}}^2 \bigg/ \left(\sum_{r=1}^{k} \lambda_{a_{rj}}^2 + \psi_{a_j}\right) \tag{6}$$

In addition, an overall percentage variance accounted for can be calculated as

$$v_a = 100 \text{ trace}\left(\Lambda_a \Lambda_a^T\right) \bigg/ \text{trace}(G_a) \tag{7}$$

The FA models with different levels of r factors were tested using residual maximum likelihood ratio test (REMLRT).

Parental stability

Empirical best linear unbiased predictions (EBLUPs) of breeding value for parent i at trial j could be expressed as

$$\hat{u}_{a_{pij}} = \hat{\lambda}_{a_{1j}}\hat{f}_{a_{1i}} + \hat{\lambda}_{a_{2j}}\hat{f}_{a_{2i}} + \ldots + \hat{\lambda}_{a_{kj}}\hat{f}_{a_{ki}} + \hat{\delta}_{a_{ij}}$$

$$= \hat{\beta}_{ij} + \hat{\delta}_{a_{ij}} \tag{8}$$

where $\hat{\beta}_{ij}$ is the predicted regression component for parent i at trial j. Equation 8 could be used to investigate parental

stability (Cullis et al. 2010; Smith et al. 2015). When k in the FA model is greater than 1, using rotated trial loading ($\hat{\lambda}^*_{a_{rj}}$) and scores ($\hat{f}^*_{a_{ri}}$) to model $\hat{\beta}_{ij}$ as $\sum_{r=1}^{k} \hat{\lambda}^*_{a_{rj}} \hat{f}^*_{a_{ri}}$, it may explain more of biological meaning of the interaction. We could build k plots for each parent using the predicted parental regression component ($\hat{\beta}_{ij}$) on each trial and their corresponding trial loading ($\hat{\lambda}^*_{a_{rj}}$) for k factors. Therefore, for parent i, the plot for the first factor has the y- and x-axes corresponding to the values of $\hat{\beta}_{ij}$ and $\hat{\lambda}^*_{a_{1j}}$, respectively, with a slope representing rotated factor score of $\hat{f}^*_{a_{1i}}$ as shown in the plot. The plot for factor 2 was similar, but with adjusted values from factor 1. Generally, the y-axis for plot b ($b = 2,…,k$) corresponds to the adjusted value of $\hat{\beta}_{ij}$ as ($\hat{\beta}_{ij} - \sum_{r=1}^{b-1} \hat{\lambda}^*_{a_{rj}} \hat{f}^*_{a_{ri}}$), and the x-axis values are $\hat{\lambda}^*_{a_{bj}}$. Similarly, the line drawn on plot b for parent i has a slope given by $\hat{f}^*_{a_{bi}}$.

Heatmap and hierarchical clustering

There are several tools for exploring G × E interaction based on the FA model (Cullis et al. 2010). We used heatmap to represent genetic correlation matrix. Cluster analysis was employed using the dissimilarity of genetic correlation between trials, and the dendrogram of cluster analysis is presented in the same heatmap.

All the models were fitted using ASReml-R package (Butler et al. 2009) within the R statistical environment (R Core Team 2014). Heatmap.2 is used to build the heatmap for genetic correlation matrix in gplots package. The dissimilarity used for cluster analysis is calculated by hierarchical clustering algorithm in hclust function within the R package.

Approximate standard errors of genetic correlations were estimated in ASReml 4.1 stand alone (Gilmour et al. 2015). Since Swedish Norway spruce tree breeding program used breeding value (BV) for volume at age 55 years as the standard age for selection, all BV estimates need to be projected to the BV at age 55. Breeding values of female parents at age 55 for volume shown in the 20 trials were predicted using TREEPLAN (McRae et al. 2004).

Results

The parental concurrence between trials is presented in Fig. 2 along with test series. The diagonal elements of the matrix are the number of female parents used in the trials, and the off-diagonal elements are the number of the female parents in common among trials. Test series 4 (F1148-F1150) and 5 (F1184, F1215, and F1216) had the best and the second best connection with all other trials, respectively. Test series 3 had good connection with series 2, 4, and 5. However, test series 1 (F1021-F1024) had no connections with test series 2 and 3 (F1059-F1147), and test series 6 (F1267-F1271) also had weak connection with all other test series (i.e., less than 13 common parents with all other test series).

Based on REMLRT, $k = 4$ was the best model to fit G_a (Table 3). However, the models beyond $r = 3$ did not show

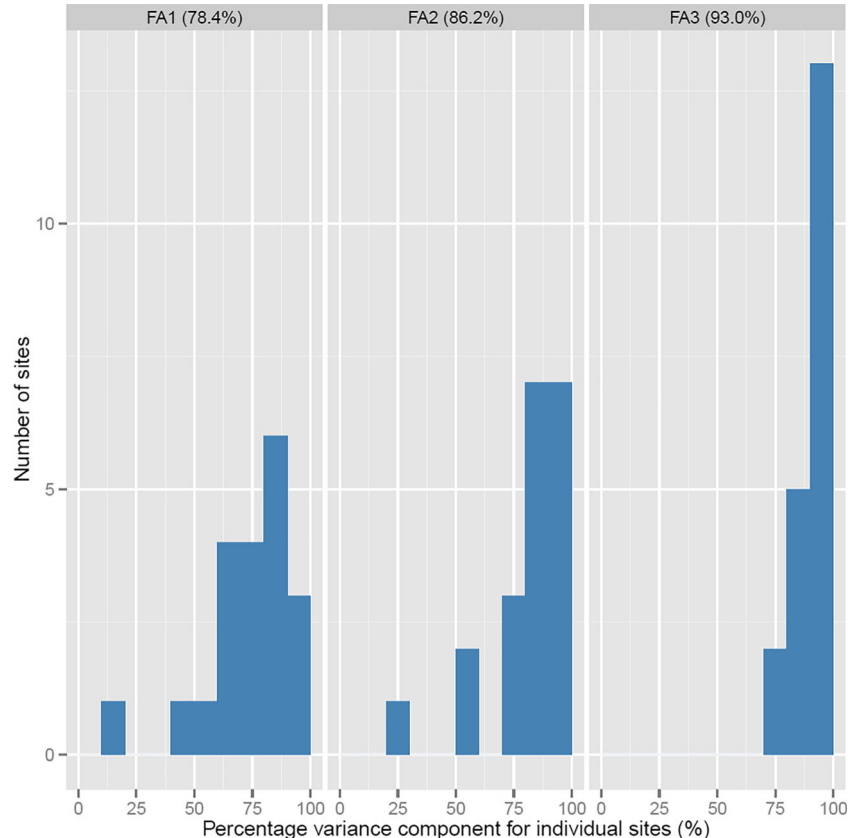| Seed Zones	7	8	7	7	8	8	7	7	7	7	8	8	7	6	6	7	7	7	6	6
Test Series	1	1	1	1	2	2	2	2	3	3	3	4	4	4	5	5	5	6	6	6
	F1021	F1022	F1023	F1024	F1059	F1064	F1067	F1069	F1145	F1146	F1147	F1148	F1149	F1150	F1184	F1215	F1216	F1267	F1270	F1271
F1021	305	304	304	303	0	0	0	0	0	0	0	74	74	74	2	2	2	12	12	12
F1022	304	304	304	303	0	0	0	0	0	0	0	74	74	74	2	2	2	12	12	12
F1023	304	304	304	303	0	0	0	0	0	0	0	74	74	74	2	2	2	12	12	12
F1024	303	303	303	303	0	0	0	0	0	0	0	74	74	74	2	2	2	12	12	12
F1059	0	0	0	0	305	305	305	304	16	16	16	20	20	20	4	4	4	1	1	1
F1064	0	0	0	0	305	305	305	304	16	16	16	20	20	20	4	4	4	1	1	1
F1067	0	0	0	0	305	305	305	304	16	16	16	20	20	20	4	4	4	1	1	1
F1069	0	0	0	0	304	304	304	304	15	15	15	19	19	19	4	4	4	1	1	1
F1145	0	0	0	0	16	16	16	15	1358	1357	1357	967	966	966	393	393	393	4	4	4
F1146	0	0	0	0	16	16	16	15	1357	1358	1358	966	965	965	393	393	393	4	4	4
F1147	0	0	0	0	16	16	16	15	1357	1358	1360	966	965	965	395	395	395	4	4	4
F1148	74	74	74	74	20	20	20	19	967	966	966	1389	1387	1386	352	353	352	9	9	9
F1149	74	74	74	74	20	20	20	19	966	965	965	1387	1387	1384	352	353	352	9	9	9
F1150	74	74	74	74	20	20	20	19	966	965	965	1386	1384	1386	352	353	352	9	9	9
F1184	2	2	2	2	4	4	4	4	393	393	395	352	352	352	732	731	729	6	6	6
F1215	2	2	2	2	4	4	4	4	393	393	395	353	353	353	731	732	729	6	6	6
F1216	2	2	2	2	4	4	4	4	393	393	395	352	352	352	729	729	730	6	6	6
F1267	12	12	12	12	1	1	1	1	4	4	4	9	9	9	6	6	6	691	691	691
F1270	12	12	12	12	1	1	1	1	4	4	4	9	9	9	6	6	6	691	880	874
F1271	12	12	12	12	1	1	1	1	4	4	4	9	9	9	6	6	6	691	874	876
6 Cluster	VI	VI	VI	V	IV	IV	III	IV	III	III	III	V	V	VI	V	V	V	II	II	I
3 Cluster	III	III	III	III	II	II	II	II	II	II	II	III	III	III	III	III	III	I	I	I

Fig. 2 The common parents between all pairs of trials within six test series in three seed zones, two cluster distributions of either 3 or 6 clusters each, respectively. The value on diagonal of the matrix show the number of parents used in each of 20 trials

Table 3 Residual log-likelihoods for different models (DIAG, heterogeneous additive variances for each of trials, FAk—factor analytic with k factors) of additive variance-covariance matrix, AIC and BIC, and percentages of variance accounted for

Model	Parameter	Residual LogL	AIC	BIC	%
DIAG	64	−381,444.5	763,017.0	763,655.3	
FA1	64	−379,896.0	759,920.0	760,558.3	78.3%
FA2	78	−379,824.6	759,805.2	760,583.2	86.1%
FA3	91	−379,802.6	759,787.2	760,694.8	93.0%
FA4	105	−379,789.3	759,788.6	760,835.9	96.4%

greater improvement for the Bayesian and Akaike information criteria (BIC and AIC, respectively). Therefore, the FA3 model was selected based on the improvement of AIC and the overall percentage variance accounted for. The distribution of individual trials based on their percentage variances accounted for the first three factors is shown in Fig. 3. FA1 accounted for about 80% variance for six trials while FA2 accounted for about 80−100% for 7 trials. We also observed that all 20 trials have an individual trial variance explained greater than 70% in FA3 model. The FA1, FA2, and FA3 models explain overall 78.3, 86.1, and 93.0% of the variances, respectively, for all 20 trials combined.

Regression interpretation of the FA model is often expressed using loadings that have been rotated to a principal component solution (Cullis et al. 2014). In FA3 model, the three loadings totally explained 93.0% of the additive genetic variance. The rotated loadings for the first three factors account for 56.2, 23.7, and 13.1% of the additive genetic variance, respectively, with each pair of loadings orthogonal.

The distribution of narrow-sense heritabilities for 20 trials is shown in Fig. 4. The mean (median) of narrow-sense heritability for the 20 trials was 0.33 (0.31) with a range from 0.11 to 0.57, based on spatial adjusted height data.

The trial-trial additive genetic correlation matrix for tree height was also obtained using the FA3 model and is represented by a heatmap with dendrograms added to the left and to the top of the heatmap (Fig. 5). Additive trial-trial correlations varied from −0.76 to 0.98 with the mean (median) of additive trial-trial correlations of 0.48 (0.58). Test series 6 (F1267, F1270, and F1271) with two most northern trials had particularly low additive trial-trial genetic correlations with other trials (Fig. 5). All the three trials had very low connection with all other trials. Excluding the three trials, mean (median) of additive trial-trial correlations increases to 0.54 (0.65).

The relationships between parents in common and estimates of type B genetic correlations (as well as their standard errors) for all pairs of trials are shown in Fig. 6. The estimates

Fig. 3 Distribution of percentage variance accounted for by individual trials in FA models (FA1, FA2, and FA3). Overall percentage for each FA model is given in *parenthesis*
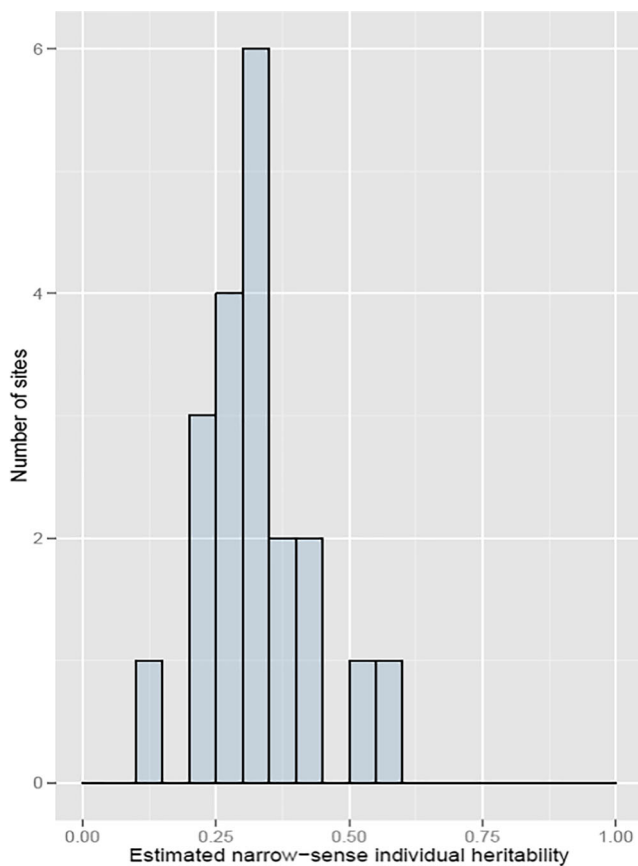
Fig. 4 Distribution of estimated narrow-sense heritabilities for 20 trials

of type B genetic correlations were generally higher and less varied when the common parents were more than 75 parents. The standard errors were smaller when common parents were more than 75 parents.

Average genetic correlation coefficients (standard errors) for tree height among and within 6 test series, derived from factor-analytic model, are shown in Table 4. Average additive genetic correlations within test series (average of 0.76 ± 0.05) were higher than those among test series (average of 0.44 ± 0.18), except that genetic correlation (0.69 ± 0.08) between test series 1 and 4 was slightly higher than that within the test series 4 (0.67 ± 0.05). Test series 6 had average correlation of 0.60 ± 0.11 within the test series, but had the lowest (a negative) correlation of -0.19 ± 0.34 with the test series 2 and had a correlation close to 0 with the test series 1 and 3.

Cluster analysis could divide the 20 trials into six clusters as shown in Figs. 2 and 5 and Table 2 if a low cutoff value (1.4) is used. There was only one trial (F1271) in cluster I. Cluster II had two trials (F1270 and F1267). The three trials for clusters I and II were all from the same test series 6 with two of the most northern trials. Cluster III had four trials that were from the test series 3 and one trial F1067 from test series 2. Cluster IV had three trials from test series 2. Cluster V

included all trials in test series 5, plus one trial from test series 1 (F1024) and two trials from test series 4 (F1148 and 1149). Cluster VI had three trials from test series 1 plus one trial F1150, which is from test series 4. It seemed that the six clusters represented more of same series plus some trials which had good connection with the test series. For example, cluster V included all trials in test series 5 plus two trials of test series 4 which had 352/353 common families with series 5.

Three clusters could be derived if a higher value was used as cutoff value (2.3) (Fig. 2): cluster I including all trials in test series 6, cluster II including all trials of test series 2 and 3, and cluster III including all trials of test series 1, 4, and 5. Again, three clusters more or less reflected trial geography and trial connectedness. Cluster I had two trials (F1270 and F1271) in the northern seed zone (zone 6) and one trial (F1267) in the middle seed zone 7. The reason that these three trials being in cluster I could be explained by geography (F1270 and F1271 in the seed zone 6) and good connection with the third trials F1267 (e.g., same series of seed zone 6). For cluster III, test series 1 had good connection (74 common families) with the test series 4 and the best connection (352 common families) between test series 4 and 5, and plus the geography (6 trials in the seed zone 7).

The correlations between climate variables and the trial rotated loadings for each of three factors may reveal the contribution level of the climatic variable to the observed different performance between trials (Table 2). Under the FA3 model, the trial loadings of the three factors were moderately to highly correlated to mean annual temperature (MAT) (0.41, 0.42, and -0.28), mean annual heat sum (MAHT) (0.30, 0.33, and -0.20), conditional mean daily temperature of May and June (MTMJ) (0.65, 0.44, and -0.46), and conditional mean daily temperature of September and October (MTSO) (0.55, 0.27, and -0.50). Both mean annual precipitation (MAP) and mean precipitation growing season − (April to October) (MPGS) showed low correlations with trial loadings for each of 3 factors (-0.27, 0.02, and 0.17 and -0.30, -0.10, and 0.11, respectively). For geographical variables, only latitude showed significant correlation with trial loadings for each of 3 factors (-0.45, -0.36, and 0.21).

Ten trials were recorded for frost damage (Table 2). The trial F1271 in the northeast had the heaviest frost incidence (39.8%). Pearson correlation between dissimilarity (computed as 1 -genetic correlation across trial) and absolute difference of frost damage across trial is 0.33. Absolute differences of MTMJ, MTSO, AMT, and North latitude (NL) showed significant Pearson correlations with the relatively additive genetic correlations across trial (0.49, 0.35, 0.25, and 0.23, respectively). Stepwise regression was used to select the variables to predict the variation of additive genetic correlation matrix. We only found that MTMJ and MTSO had significant effect on G × E and they explained 27.8% of variation.

Fig. 5 Heat map and dendrogram of cross-trial additive genetic correlations for height
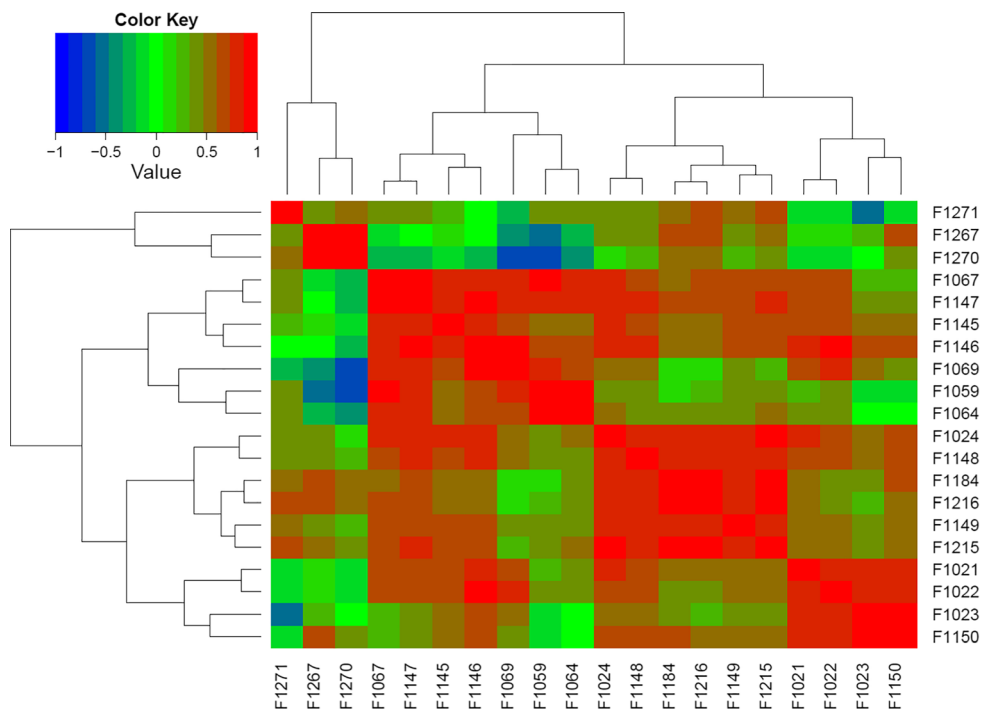
Parental stability may be best viewed using latent regression plots which show genetic responses to each of trial loadings. Twelve parents with the highest breeding values of volume at age 55 estimated by TREEPLAN (McRae et al. 2004) were selected to show their latent regression plots responding to the first three factors in Figs. 7 and 8. If the progeny of the associated parents were planted in the corresponding trials, the score points on the plot were colored as blue; otherwise, they were colored as red dots. For example, progeny of the parents P1, P3, and P4 was tested in four, seven, and six trials, respectively, as shown in Fig. 7. The line on each plot is a regression slope estimated by the predicted factor scores (rotated) for the parent. Table 5 shows the 3 factor scores for the 12 parents and the percentage representation of each parent planted in the 20 trials.

The estimated trial loadings for the first factor were all positive, except for trial F1270 (Table 2). In addition, the first factor explained a large proportion of additive G × E variation (56.2%), indicating that the latent regression on the first factor had the greatest impact on predicted breeding values. The large positive rotated factor scores (slope) for the first factor indicated positive correlation between tree height and trial loading score (Fig. 7, on the left). For example, P5 had the largest slope for factor 1, the predicted breeding values were all positive in all trials (Table 2, see loadings for factor 1), and it had larger breeding values in those trials with higher trial loadings. All 12 parents had positive slopes in Fig. 7, which means that these parents had positive response to trial loadings in the first factor. However, P1, P2, P7, P11, and P12 had more spread around the lines (slope) than other parents.
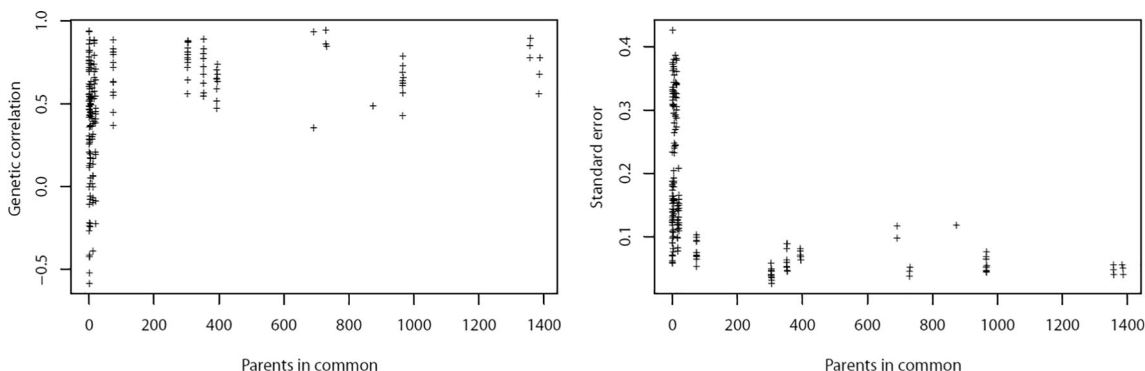
Fig. 6 Estimated genetic correlations for tree height on the *left* and their approximated standard errors on the *right*

Table 4   Average genetic correlation coefficients (standard errors) for height among the six test series, derived from factor-analytic model

	S1 (F1021-24)	S2 (F1059-69)	S3 (F1145-47)	S4 (F1148-50)	S5 (F1184-16)	S6 (F1267-71)
S1	0.75 (0.04)	0.45 (0.15)	0.71 (0.09)	0.69 (0.08)	0.54 (0.11)	0.04 (0.33)
S2		0.79 (0.04)	0.76 (0.12)	0.35 (0.14)	0.38 (0.17)	−0.19 (0.34)
S3			0.86 (0.05)	0.64 (0.05)	0.62 (0.07)	0.01 (0.34)
S4				0.67 (0.05)	0.72 (0.06)	0.34 (0.30)
S5					0.89 (0.04)	0.58 (0.28)
S6						0.60 (0.11)

The values of diagonal are the average of genetic correlations within the test series

Latent regression plots for factor 2 are shown in Fig. 7 (on the right). P1, P2, and P12 had strong positive responses to trial loadings while P6, P9, and P10 showed small positive responses to trial loadings. P3, P4, P5, and P11 showed negative responses to trial loadings and P8 showed quite stable responses to trial loadings.

Latent regression plots for factor 3 are shown in Fig. 8. P1, P2, P4, P5, and P12 showed moderate to strong positive response to trial loadings for factor 3. P3, P9, P10, and P11 had moderate to strong negative responses to trial loadings. P6, P7, and P8 were almost stable across trials. The differential parental response to FA2 and FA3 factors in Figs. 7 and 8 may indicate potential genotype by FA factor (trial loading) interaction.

Discussion

Connectedness

Prior to estimating the genetic parameters and exploring $G \times E$ interaction in MET analysis, genetic connectedness between trials should be examined and analyzed (Apiolaza 2012; Ivković et al. 2015; Smith et al. 2015). Poor connectedness between trials may bias the estimates of genetic correlations (Apiolaza 2012; Dutkowski et al. 2016). Cullis et al. (2014) reported that in an unpublished simulation study, when an FA model was used for estimation of $G \times E$ effect, there was little bias in the estimated genetic correlation for a pair of trials with poor concurrence (even zero concurrence), if there was
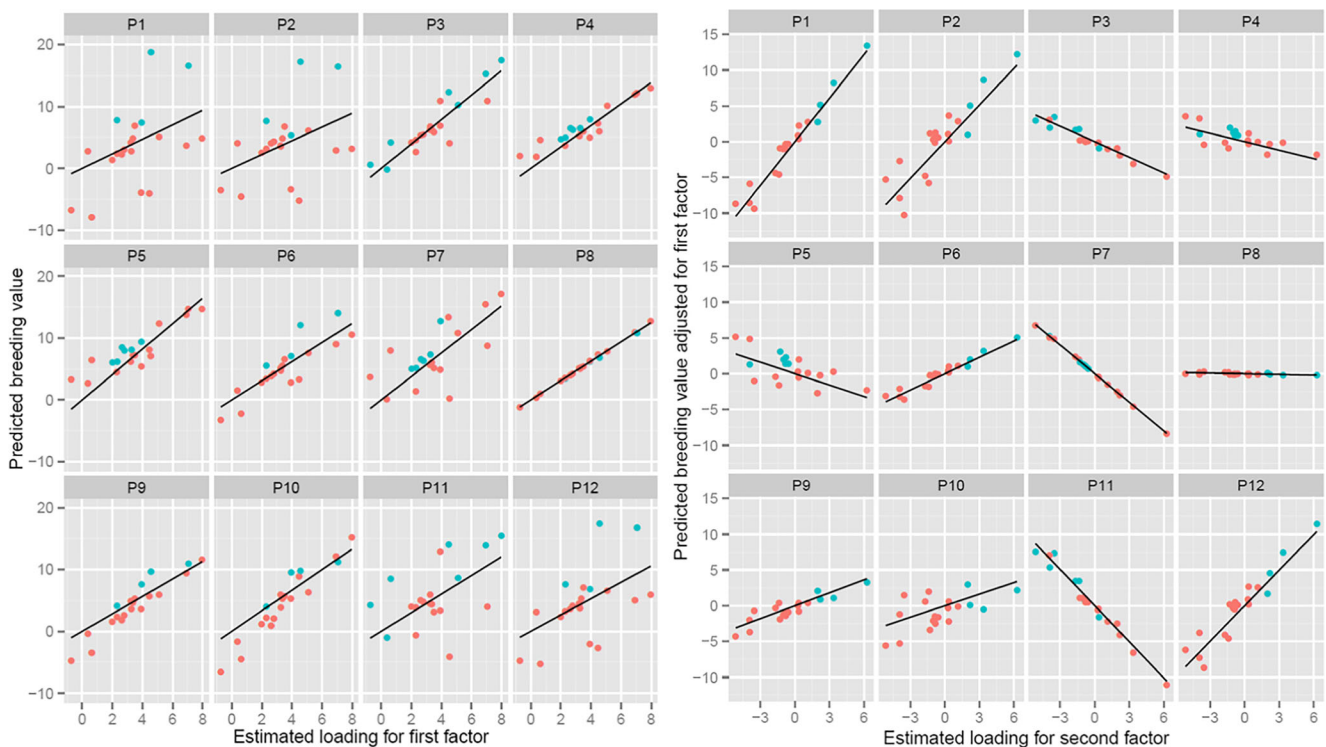
Fig. 7 Latent additive genetic regression plots using the first factor (on the *left*) and the second factor (on the *right*) for 12 parents with the largest breeding value for volume in age 55 estimated in TREEPLAN

Fig. 8 Latent additive genetic regression plot for the third factor for 12 parents with the largest breeding value for volume in age 55 estimated in TREEPLAN
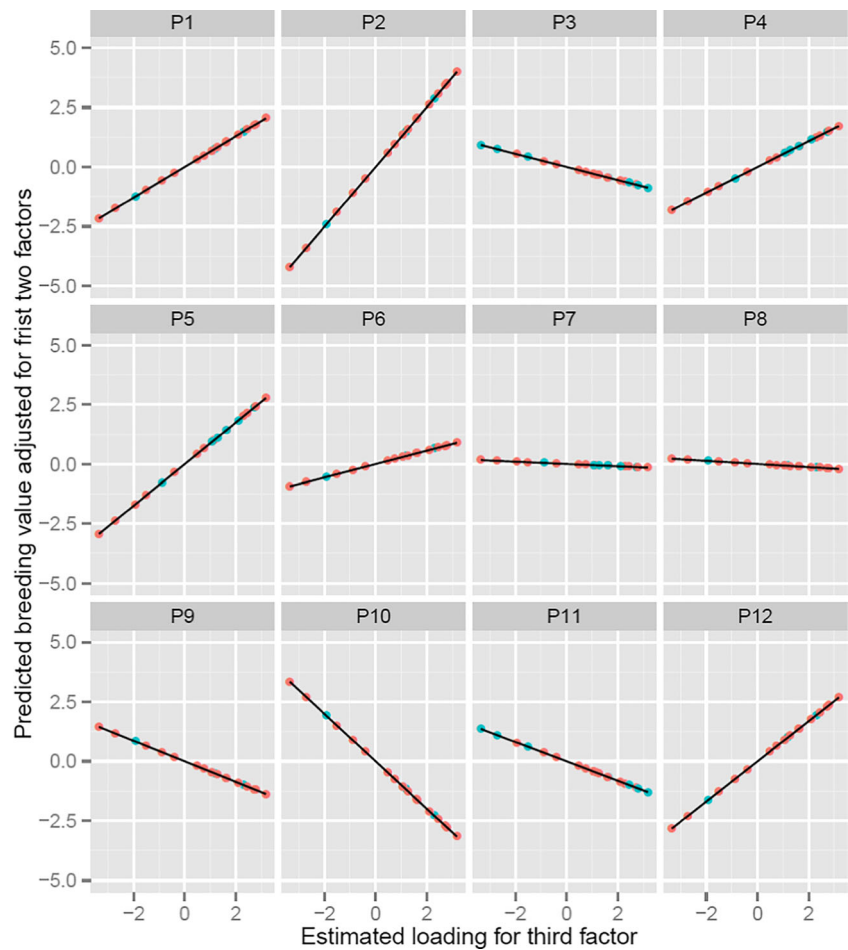
sufficient linkage through other trials. In our study, genetic correlation within test series was generally higher than that among test series. Genetic correlations were generally low among two series if there were a few direct common parents, and in addition, there were only a few common parents in the third series that connected with the both series. For example, testing series 6 had poor correlations and large standard errors with other series because it has no good connections directly

Table 5 Predicted (rotated) three factor scores for 12 parents with the highest breeding value of volume in age 55

Parent_ID	Parent	Factor 1	Factor 2	Factor 3	No of trials	Representation (%)
S21K8120405	P1	1.18	2.03	0.65	4	20
S21K8120402	P2	1.12	1.71	1.26	4	20
S21K8120835	P3	1.98	−0.72	−0.28	7	35
S21K8120451	P4	1.74	−0.40	0.54	6	30
S21K8120840	P5	2.05	−0.53	0.87	6	30
S21K8220629	P6	1.55	0.76	0.28	4	20
S21K8120535	P7	1.90	−1.34	−0.05	6	30
S21K8220073	P8	1.56	−0.03	−0.07	4	20
S21K8220063	P9	1.41	0.60	−0.43	4	20
S21K8120407	P10	1.67	0.54	−1.00	4	20
S21K8120412	P11	1.51	−1.69	−0.41	7	35
S21K8220633	P12	1.33	1.66	0.85	4	20

Note that these breeding values were estimated in TREEPLAN Platform in 2014, based on all available data in field trials. Also shown is the number of trials in which the progeny of those parents were planted and the percentage representations in 20 trials

and indirectly with other series. On the other hand, test series 1 had no direct connections with test series 2 and 3, but had more than 74 common parents with test series 4, while the test series 4 had about 20 and 965 common parents with the test series 2 and series 3, respectively. From this, we observed a medium correlation of 0.45 between test series 1 and 2 and a medium standard error of 0.15, and a high correlation of 0.71 between test series 1 and 3 and a low standard error of 0.09. Therefore, we could estimate genetic correlation between series that have no or low direct connection, but have good connection with a common third series.

Estimate of G × E interaction and cluster analysis

To understand and estimate the amount of G × E interaction, genetic correlations between trials are commonly used to explore G × E interaction (Berlin et al. 2014; Burdon 1977; Hannrup et al. 2008; Smith et al. 2015; Wu and Matheson 2005). In Sweden, there are a few published papers that discuss patterns of G × E interaction and their significance for Norway spruce breeding. Bentzer et al. (1988) reported high genotypic correlations of height at age 5 for two studies with six and three clonal trials, respectively. Average genotypic correlations of 0.66 and 0.91 were observed, respectively. Karlsson and Högberg (1998) observed a medium type B genotypic genetic correlation of 0.54 for tree height in several clonal trials at age 11. In another study, Karlsson et al. (2001) showed a similar result using 11 clonal trials in southern Sweden and Demark. However, a univariate model was used to estimate the genotypic genetic correlations in their studies. Kroon et al. (2011) reported that a high average of 0.80 for additive genetic correlation in a study using a bivariate model, indicating a low G × E interaction in Norway spruce. Berlin et al. (2014) used a widely accepted multivariate model to estimate genetic correlations and found a low to moderate G × E interaction with an average of 0.72 trial-by-trial genetic correlation in 65 Norway spruce trials in southern Sweden. Estimation of type B genetic correlation between two trials using conventional approach may be inefficient, and often result in estimates of genetic correlation greater than 1 or less than −1 (Cullis et al. 2014).

To improve MET data analysis, Smith et al. (2001) recommended a new approach combining a FA model and spatial analysis; thus, in this study, we used the more advanced approach to estimate trial-by-trial genetic correlation. The substantial G × E interaction was found with an average (median) of 0.48 (0.58) for genetic correlations between the 20 trials in southern and middle Sweden. Our study indicated stronger G × E than the most studies of Norway spruce except for the study by Karlsson and Högberg (1998) in which they observed a median genotypic correlation of 0.54. Such large G × E interaction indicates that exploring patterns and the drivers of G × E is important in current Norway spruce breeding.

In previous studies for G × E interaction of Swedish spruce, estimates of genetic correlations were all within the same test series; so, exploring G × E pattern is relatively difficult because the area of a breeding test series is not large enough to cover a large area. In this study, we used 20 trials in 6 test series which cover several major southern and central Sweden's seed orchard zones. Thus, a cluster analysis is recommended to detect if the clusters follow the current seed zones. The clusters partially reflected the three seed orchard zones but with many exceptions that were similarly observed by Karlsson et al. (2001) using 11 clonal trials in Denmark and southern Sweden. The discrepancy between the existing seed zone delineation and our clustering results using biological data may indicate that current delineation of seed zone may not be optimal. The fact that uneven trial distribution of same series among the seed zones, low connectedness between series made the clustering of trials based on type B genetic correlation more complicated, and it may have only partially reflected the real G × E patterns for the region tested. We recommend that more connections among testing series should be used in future progeny testing. For example, St Clair and Kleinschmit (1986) found a simple geographic pattern for Norway spruce in northern Germany with trials having more connections. Their study area was divided into two zones based on the altitude of trials, which indicated effect of spring or autumn frost.

Environment drivers

In a previous study, Berlin et al. (2014) examined 15 climate variables to investigate the possibility of predicting genetic correlations across trials using a linear regression model. However, none of climate variables could explain more than 5% G × E interaction variation. However, they observed that frost damage had a significant influence on across-trial genetic correlation. In almost all the published papers involving the G × E discussion for Norway spruce, frost damage was considered as the main cause in southern and central Sweden (Berlin et al. 2014; Karlsson and Högberg 1998; Karlsson et al. 2001). Thus, in this study, we firstly tried to use the frost damage to explain across-trial genetic correlations and found that it could explain 11% of additive genetic correlations. However, the frost damage was only recorded in ten trials out of the total 20 trials and was measured at different ages. To relate potential frost severity at different trials with G × E interactions, we constructed two climate indices (MTMJ and MTSO) related to the temperature at spring and autumn frost seasons. We observed that MTMJ and MTSO could explain 27.8% variation of additive genetic correlation matrix and these two variables showed significant correlation with the first and second (rotated) trial loadings ($r = 0.65$, and 0.44 for MTMJ for the first and second loading, respectively, and $r = 0.55$ and 0.37 for MTSO for the first and second loading,

respectively). These may indicate that the climate indices that reflect spring or autumn frost seasons could more reliably predict the G × E interaction. We recommend use of the minimum and maximum daily temperatures in spring and autumn to further examine possible accuracy in predicting the G × E interaction for Norway spruce. Since temperature loggers are routinely placed in all field trials now, the local temperature records should be more reliable.

Stability of parents

In general, tree breeding programs select individual trees with the highest breeding values, which are stable across trials. In this study, 12 parents selected with the highest breeding values from a total of 3160 parents. P3, P4, P5, and P8 had both high stability and high breeding values. They are high valued for breeding and deployment in these 20 trials (three seed zones). In contrast, most of other parents that had the higher breeding value in many trials showed very high incidence of frost damage on frost prone trials, and therefore showed poor stability.

In conclusion, we find a relatively large amount of G × E interactions for tree height of Norway spruce in southern and central Sweden. Revision of the delineation of breeding and seed orchard zones was not recommended at moment due to the limited number of test trials and less satisfactory genetic connection among several series in the study. It is recommended that more trials with better connection among test series should be used to improve delineation of breeding and seed orchard in the future. The climate indices of spring and autumn temperature have been shown to account for a large amount of G × E interaction and frost damage could be an important cause of G × E interaction in Norway spruce of southern and central Sweden.

Data archiving statement All raw data are archived in DataPlan for access.

References

Apiolaza LA (2012) Basic density of radiata pine in New Zealand: genetic and environmental factors. Tree Genet Genomes 8:87–96. doi:10.1007/s11295-011-0423-1

Baltunis BS, Gapare WJ, Wu HX (2010) Genetic parameters and genotype by environment interaction in radiata pine for growth and wood quality traits in Australia. Silvae Genet 59:113–124

Bentzer BG, Foster GS, Hellberg AR, Podzorski AC (1988) Genotype × environment interaction in Norway spruce involving three levels of genetic control: seed source, clone mixture, and clone. Can J For Res 18:1172–1181. doi:10.1139/x88-180

Berlin M, Jansson G, Högberg K-A (2014) Genotype by environment interaction in the southern Swedish breeding population of *Picea abies* using new climatic indices. Scand. J For Res 30:112–121. doi:10.1080/02827581.2014.978889

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

Burgueño J, Crossa J, Cornelius PL, Yang R-C (2008) Using factor analytic models for joining environments and genotypes without crossover genotype × environment interaction. Crop Sci 48. doi:10.2135/cropsci2007.11.0632

Butler DG, Cullis BR, Gilmour AR, Gogel BJ (2009) ASReml-R reference manual. The State of Queensland, Department of Primary Industries and Fisheries, Brisbane

Campbell RK, Sorensen FC (1978) Effect of test environment on expression of clines and on delimitation of seed zones in Douglas-fir. Theor Appl Genet 51:233–246. doi:10.1007/bf00273770

Costa e Silva J, Graudal L (2008) Evaluation of an international series of *Pinus kesiya* provenance trials for growth and wood quality traits. Forest Ecol Manag 255:3477–3488. doi:10.1016/j.foreco.2008.02.027

Costa e Silva J, Dutkowski GW, Gilmour AR (2001) Analysis of early tree height in forest genetic trials is enhanced by including a spatially correlated residual. Can J For Res 31:1887–1893. doi:10.1139/x01-123

Costa e Silva J, Potts B, Dutkowski G (2006) Genotype by environment interaction for growth of *Eucalyptus globulus* in Australia. Tree Genet Genomes 2:61–75. doi:10.1007/s11295-005-0025-x

Cullis BR, Smith AB, Beeck CP, Cowling WA (2010) Analysis of yield and oil from a series of canola breeding trials. Part II. Exploring variety by environment interaction using factor analysis. Genome 53:1002–1016. doi:10.1139/g10-080

Cullis BR, Jefferson P, Thompson R, Smith AB (2014) Factor analytic and reduced animal models for the investigation of additive genotype-by-environment interaction in outcrossing plant species with application to a *Pinus radiata* breeding programme. Theor Appl Genet 127:2193–2210. doi:10.1007/s00122-014-2373-0

Danell Ö (1993) Breeding programmes in Sweden. 1. General approach. Progeny testing and breeding strategies: Proceedings of the Nordic Group of Tree Breeders, Edinburgh:6–10

Dutkowski GW, Costa e Silva J, Gilmour AR, Wellendorf H, Aguiar A (2006) Spatial analysis enhances modelling of a wide variety of traits in forest genetic trials. Can J For Res 36:1851–1870

Dutkowski G, Ivković M, Gapare WJ, McRae TA (2016) Defining breeding and deployment regions for radiata pine in southern Australia. New For 47:783–799. doi:10.1007/s11056-016-9544-6

Finlay K, Wilkinson G (1963) The analysis of adaptation in a plant-breeding programme. Aust J Agric Res 14:742–754. doi:10.1071/AR9630742

Fox GP, Bowman J, Kelly A, Inkerman A, Poulsen D, Henry R (2007) Assessing for genetic and environmental effects on ruminant feed quality in barley (*Hordeum vulgare*). Euphytica 163:249–257. doi:10.1007/s10681-007-9638-5

Freeman GH (1973) Statistical methods for the analysis of genotype-environment interactions. Heredity 31:339–354

Fu Y-B, Yanchuk AD, Namkoong G (1999) Spatial patterns of tree height variations in a series of Douglas-fir progeny trials: implications for genetic testing. Can J For Res 29:714–723. doi:10.1139/x99-046

Gauch HG (1992) Statistical analysis of regional yield trials: AMMI analysis of factorial designs. Elsevier Science Publishers

Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2009) ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead

Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thompson R (2015) ASReml user guide release 4.1. VSN International Ltd, Hemel Hempstead

Hamann A, Gylander T, Chen P-y (2011) Developing seed zones and transfer guidelines with multivariate regression trees. Tree Genet Genomes 7:399–408. doi:10.1007/s11295-010-0341-7

Hannrup B, Jansson G, Danell Ö (2008) Genotype by environment interaction in *Pinus sylvestris* L. in southern Sweden. Silvae Genet 57: 306–311

Hardner CM, Dieters M, Dale G, DeLacy I, Basford KE (2010) Patterns of genotype-by-environment interaction in diameter at breast height at age 3 for eucalypt hybrid clones grown for reafforestation of lands affected by salinity. Tree Genet Genomes 6:833–851. doi:10.1007/s11295-010-0295-9

Ivković M, Gapare W, Yang H, Dutkowski G, Buxton P, Wu H (2015) Pattern of genotype by environment interaction for radiata pine in southern Australia. Ann For Sci 72:391–401. doi:10.1007/s13595-014-0437-6

Johnson G (2004) Common families across test series − how many do we need? For Genet 11:103–112

Karlsson B, Högberg K (1998) Genotypic parameters and clone x site interaction in clone tests of Norway spruce (*Picea abies* (L.) karst.). For Genet 5:21–30

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

Kempton RA (1984) The use of biplots in interpreting variety by environment interactions. J Agric Sci 103:123–135

Kroon J, Ericsson T, Jansson G, Andersson B (2011) Patterns of genetic parameters for height in field genetic tests of *Picea abies* and *Pinus sylvestris* in Sweden. Tree Genet Genomes 7:1099–1111. doi:10.1007/s11295-011-0398-y

Mathews KL, Chapman SC, Trethowan R, Pfeiffer W, Ginkel M, Crossa J, Payne T, DeLacy I, Fox PN, Cooper M (2007) Global adaptation patterns of Australian and CIMMYT spring bread wheat. Theor Appl Genet 115:819–835. doi:10.1007/s00122-007-0611-4

McKeand S, Amerson H, Li B, Mullin T (2003) Families of loblolly pine that are the most stable for resistance to fusiform rust are the least predictable. Can J For Res 33:1335–1339

McRae T, Dutkowski GW, Pilbeam D, Powell MB, Tier B (2004) Genetic evaluation using the TREEPLAN system. In: Li B, McKeand SE (eds) Forest genetics and tree breeding in the age of genomics: progress and future. IUFRO Joint Conference of Division 2, Charleston, SC, USA, 1–5 November 2004 2004. pp 388–399

Namkoong G (1985) The influence of composite traits on genotype by environment relations. Theor Appl Genet 70:315–317. doi:10.1007/bf00304918

Piepho HP (1998) Methods for comparing the yield stability of cropping systems. J Agron Crop Sci 180:193–213. doi:10.1111/j.1439-037X.1998.tb00526.x

R Core Team (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Rehfeldt GE (1983) Adaptation of *Pinus contorta* populations to heterogeneous environments in northern Idaho. Can J For Res 13:405–411. doi:10.1139/x83-061

Rosvall O, Ståhl P, Almqvist C, Anderson B, Berlin M, Ericsson T, Eriksson M, Gregorsson B, Hajek J, Hallander J (2011) Review of the Swedish tree breeding programme. Arbetsrapport, Skogforsk, Uppsala Science Park, Uppsala:114

Smith A, Cullis B, Thompson R (2001) Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. Biometrics 57:1138–1147. doi:10.1111/j.0006-341X.2001.01138.x

Smith AB, Ganesalingam A, Kuchel H, Cullis BR (2015) Factor analytic mixed models for the provision of grower information from national crop variety testing programs. Theor Appl Genet 128:55–72. doi:10.1007/s00122-014-2412-x

St. Clair J, Kleinschmit J (1986) Genotype-environment interaction and stability in ten-year height growth of Norway spruce clones (*Picea abies* karst.). Silvae Genet 35:177–185

Thompson R, Cullis B, Smith A, Gilmour A (2003) A sparse implementation of the average information algorithm for factor analytic and reduced rank variance models. Aust N Z J Stat 45:445–459. doi:10.1111/1467-842X.00297

Wang T, O'Neill GA, Aitken SN (2010) Integrating environmental and genetic effects to predict responses of tree populations to climate. Ecol Appl 20:153–163. doi:10.1890/08-2257.1

Wu HX, Matheson AC (2005) Genotype by environment interactions in an Australia-wide radiata pine diallel mating experiment: implications for regionalized breeding. For Sci 51:29–40

Wu HX, Ying CC (1998) Stability of resistance to western gall rust and needle cast in lodgepole pine provenances. Can J For Res 28:439–449. doi:10.1139/x98-009

Wu HX, Ying CC (2004) Geographic pattern of local optimality in natural populations of lodgepole pine. For Ecol Manag 194(1–3):177–198

Ye T, Jayawickrama KS (2008) Efficiency of using spatial analysis in first-generation coastal Douglas-fir progeny tests in the US Pacific northwest. Tree Genet Genomes 4:677–692. doi:10.1007/s11295-008-0142-4

Zobel RW, Wright MJ, Gauch HG (1988) Statistical analysis of a yield trial. Agron J 80:388–393