## ORIGINAL PAPER

# Pattern of genotype by environment interaction for radiata pine in southern Australia

Miloš Ivković • Washington Gapare • Huixiao Yang • Gregory Dutkowski • Peter Buxton • Harry Wu

Received: 17 April 2014 / Accepted: 4 November 2014 / Published online: 21 November 2014 © INRA and Springer-Verlag France 2014

## Abstract

• *Key message* Based on analyses using 20 genetically connected radiata pine trials and on the pattern of trial-trial genetic correlations, current regionalisation of breeding in southern Australia seems justified. However, relationships between environmental variables and genotype by environment interaction are complex.

#### Handling Editor: Bruno Fady

**Contribution of the co-authors** Miloš Ivković designed and performed the research, analysed data and wrote the manuscript

Dr. Gapare contributed to initial (spatial) data analyses and revised the manuscript.

Dr. Yang contributed to data analyses and revised the manuscript.

Dr. Dutkowski contributed data sets and methods for data analyses.

Dr. Wu contributed ideas that initiated this work.

Mr. Buxton contributed trial data sets and performed field measurements.

**Electronic supplementary material** The online version of this article (doi:10.1007/s13595-014-0437-6) contains supplementary material, which is available to authorized users.

M. Ivković (⊠) · W. Gapare · H. Wu CSIRO Agriculture Flagship, GPO Box 1600, Canberra, ACT 2601, Australia e-mail: milosh.ivkovich@csiro.au

W. Gapare e-mail: washington.gapare@csiro.au

H. Wu e-mail: harry.wu@csiro.au

#### H. Yang

CSIRO and College of Forestry, Guangdong Academy of Forestry, Guangzhou 510520, China e-mail: yanghx1111@163.com

G. Dutkowski PlantPlan Genetics, Private Bag 55, Hobart, TAS 7001, Australia e-mail: greg.dutkowski@plantplan.com

#### P. Buxton

Southern Tree Breeding Association, PO Box 1811, Mount Gambier, SA 5290, Australia

• *Context* Current radiata pine breeding and deployment in Australia is based largely on the plantation inventory zones rather than on biological patterns of genotype by environment interaction ( $G \times E$ ), and consequently cannot deliver optimal genetic gains across the whole plantation estate.

• *Aim* This study examined patterns of  $G \times E$  to facilitate deployment of genetic stock to particular environments.

• *Methods* We used 20 genetically well-connected trials across southern Australia to obtain estimates of genetic correlations between performances at different trial sites. Extended factor analyses (XFA) were used to estimate  $G \times E$  variance and produce a matrix of sitesite genetic correlations. The patterns among these correlations were examined using a heat map and hierarchical clustering.

• *Results* The XFA captured a large proportion of both additive and non-additive  $G \times E$ . Significant  $G \times E$  for diameter growth can be expected between Tasmania and Mainland, and within Tasmania itself. The study also confirmed presence of  $G \times E$  between Murray Valley region in New South Wales and the rest of southern Australia.

• *Conclusion* The  $G \times E$  interaction at transcontinental scale can be correlated to the climate variables, primarily rainfall and temperature; however, the drivers of  $G \times E$  may also be related to smaller scale environmental variation (i.e. soil and terrain variation).

**Keywords** *Pinus radiata* · Factor analysis · Breeding zones · Genotype deployment

## **1** Introduction

Genotype by environment interaction ( $G \times E$ ) often represents a significant proportion of the overall genetic variation (Yan and Kang 2003). Benefits genetic improvement can be fully realised only if improved genotypes are matched to



environments so they can fully realise their potential. Importance of  $G \times E$  in forest trees has been emphasised by White et al. (2007). Furthermore, future adaptation of forests depends on the response of genotypes to fast-changing climate conditions (e.g. Schreiber et al. 2011). The main approach to the issue of  $G \times E$  has been to characterise test environments and identify main environmental factors driving  $G \times E$ . Breeding and deployment zones are then defined so that within these zones,  $G \times E$  is minimised. However,  $G \times E$  is often not simply related to a single environmental variable, but rather several climatic or other site characteristics (e.g. Matheson and Cotterll 1990).

Many tree improvement programmes have not resolved with the issue of  $G \times E$  because of the lack of sufficiently genetically connected trials and difficulty of interpreting  $G \times$ E. Radiata pine (*Pinus radiata* D. Don) tree improvement programme in Australia is one of the most advanced in the world. Significant improvement has been made in growth, form and wood quality traits (Wu et al. 2008). However, current radiata pine breeding and deployment is based largely on the National Plantation Inventory regions (Gavran and Parsons 2011), rather than on environmental drivers of  $G \times$ E. Therefore, current breeding and deployment zones cannot deliver optimal genetic gains across the whole estate.

Important G×E among radiata pine families were observed in several experiments in Australia. For example, the "Australia-wide Diallel" (AWD) experiment planted on 10 sites revealed significant rank changes between two highelevation sites in New South Wales (NSW) and other sites in southern Australia (Wu and Matheson 2005; Gapare et al. 2010). Raymond (2011) reported significant G×E at family level for diameter growth, with elevational differences between sites being a key driver of G×E in NSW. A recent paper by Gapare et al. (2012) also confirmed those results and divided NSW sites into high-elevation/high-rainfall and lowelevation/low-rainfall groups. Another study involving eight sites revealed significant G×E for growth and branch size between sites in Tasmania and mainland southern Australia (Baltunis et al. 2010).

This study was based on the most comprehensive, genetically well-connected and well-distributed set of genetic trials in Australia. The study used information from 20 radiata pine progeny trials established in 1996/1997 by the Southern Tree Breeding Association (STBA). Based on data from those trials with common genetic material, estimates of site-site genetic correlations were obtained. The patterns among these correlations were modelled against various environmental characteristics. The study is part of a larger effort to understand  $G \times E$  and to obtain site classifications that will capture a large proportion of  $G \times E$ . The general aim is to facilitate optimal deployment of genetic stock to particular environments and development of software tools for deployment by the STBA.

🖄 Springer



The specific objectives of the study were:

- To estimate genetic variances (i.e. additive and dominance) and heritability for growth in the set of trials
- To estimate genetic correlations between trials
- To examine patterns of G×E between trial locations in southern Australia
- To identify potential geo-climatic drivers of G×E

## 2 Materials and methods

## 2.1 Field trials and measurements

Ten genetic control-pollinated progeny trials were planted in each 1996 and 1997 by the Southern Tree Breeding Association, in the "BR" trial series. The trials are listed together with their locations and dates of planting in Table 1, and the map of locations is given in Fig. 1. All trials had replication, incomplete block and row-plot design features, and were represented on a row-column grid for spatial analyses (see below). The assessments of diameter growth at breast height (DBH) at juvenile stage (i.e. 6–10 years) were used for the analyses.

One of the most important issues in undertaking a  $G \times E$ analysis is the assessment of the degree of genetic connectivity between trials. The concurrence matrix is based on the numbers of parents in common between trials is given in Online Resource 1. Number of parent trees represented in each trial ranged from 41 to 177 and the parent trees in common between trials from 31 to 165. There are no clear cut guidelines which determine an acceptable level of connectivity; however, it is obvious that disconnected or poorly connected data sets will not permit reliable estimation of the co-variance.

## 2.2 Environmental variables

Basic climate, geological and soil information for BR1996/97 trial locations are given in Table 2. Daily climate data for selected locations within Australia were extracted from the SILO Climate Database (http://www.longpaddock.qld.gov.au/ silo/index.html). The daily climate data have been constructed using observations from 4,600 locations across Australia for rainfall, maximum and minimum temperatures, evaporation and solar radiation using spatial interpolation algorithms. A simple monthly aridity index (AIX) was calculated as a ratio of monthly mean daily pan evaporation rate to the total monthly rainfall. The aridity index is related to water balance, and monthly minimum AIX value for the most arid quarter was used to rank the sites in terms of aridity. Climate data sequence from planting to DBH assessment was obtained for each trial. Table 1State, region, abbrevia-<br/>tion of National Plantation Inven-<br/>tory region (NPI), location coor-<br/>dinates, altitude and planting year<br/>for BR1996/1997 trial series

Frial_ID	State	Region	NPI	Long	Lat S	Alt (m)	Planted
BR9601	South Australia	Green Triangle	GT	140.8	37.7	60	1996
BR9604	Victoria	Green Triangle	GT	141.3	37.8	67	1996
BR9606	Victoria	Green Triangle	GT	141.2	37.9	55	1996
BR9608	Victoria	Central	CV	143.4	37.8	268	1996
BR9610	Tasmania	North-west	TAS	145.6	41.2	409	1996
BR9611	Victoria	Gippsland	GL	146.8	38.2	170	1996
BR9613	Victoria	Gippsland	GL	147.1	37.8	165	1996
BR9614	Tasmania	South East	TAS	146.9	42.8	628	1996
BR9615	Tasmania	North-east	TAS	147.4	41.2	215	1996
BR9617	NSW	Murray Valley	MV	147.7	35.8	458	1996
BR9701	West Aust	South West	WA	116.0	33.5	207	1997
BR9702	West Aust	South West	WA	116.1	33.7	232	1997
BR9703	South Aust	Green Triangle	GT	140.4	37.5	55	1997
BR9705	South Aust	Green Triangle	GT	140.9	37.8	55	1997
BR9707	South Aust	Green Triangle	GT	141.3	37.6	140	1997
BR9709	Victoria	Central	CV	143.4	37.8	268	1997
BR9710	Victoria	Otways	CV(OTW)	143.5	38.6	211	1997
BR9712	Tasmania	South East	TAS	146.7	42.7	240	1997
BR9713	Victoria	Gippsland	GL	147.1	37.8	180	1997
BR9715	Tasmania	North-east	TAS	147.7	41.1	350	1997

A seamless national coverage of outcrop and surface geology was obtained from Geoscience, Australia (http://mapconnect.ga. gov.au/MapConnect/Geology). Soil types were categorised using a technical classification system into 11 categories of parent rock code. This system was developed to group forest sites according to expected volume productivity (Turner et al. 2001). In addition, a high resolution weathering intensity index for the Australian continent, based on airborne gamma-ray spectrometry and digital terrain analysis was also obtained (Wilford 2012).

Australian Soil Resource Information System (ASRIS, www.asris.csiro.au) database was used to obtain information on soils in a consistent format across southern Australia. The information was obtained on soil depth, water storage, permeability, fertility, carbon and erodibility, with most soil information recorded at five depths. Soil profile data with fully characterised sites were also available from forestry organisations and companies.

2.3 Statistical analyses

## 2.3.1 Spatial analyses

Spatial analyses were performed using a two-dimensional separable autoregressive (AR1) model fitted to column-row grid diameter data of each trial using ASReml<sup>®</sup> (Gilmour et al. 2009). The spatial method partitioned the residual variance into an independent component and a two-dimensional spatially auto-correlated component. An additional random term with one level for each experimental unit was used so that a



Table 2 Some basic climate, geological, and soil information for locations of BR1996/97 trials: Mean Temperature of Growing Season (Sep-Apr) (TGS), Mean Temperature Driest Month (TDM), Precipitation in Growing Season (Sep-Apr) (PGS), Precipitation Driest Quarter (PDQ), Total Evaporation (TE), Aridity Index (AIX), Parent Rock Code (PRC), Weathering Index (WI), Top soil Bulk Density (BD), Top soil PH

Trial_id	Region	TGS °C	TDM ℃	PGS mm	PDQ mm	TE mm	AIX	PRC	WI	BD g/cm <sup>3</sup>	PH
BR9601	GT	15.2	24.4	316	89	1,313	14	022	72	1.5	5.1
BR9604	GT	15.4	24.3	319	92	1,336	13.4	022	85	1.3	5.3
BR9606	GT	15.0	24.2	338	83	1,311	10.1	022	85	1.4	6.1
BR9608	CV	14.8	23.4	311	119	1,300	12.6	112	52	1.4	4.6
BR9610	TAS	10.7	18.5	808	234	899	2.4	051	67	1.0	6.0
BR9611	GL	15.3	21.5	391	143	1,180	5.5	021	64	1.3	4.6
BR9613	GL	15.4	19.8	443	170	1,143	4.1	021	64	1.3	5.1
BR9614	TAS	11.2	17.2	489	163	936	4.1	111	48	1.0	6.4
BR9615	TAS	13.0	21.6	575	183	1,015	7.2	091	66	0.8	4.8
BR9617	MV	15.0	24.0	484	156	1,169	10	051	77	N/A	N/A
BR9701	WA	17.5	30.2	295	41	1,414	310	071	71	1.5	4.8
BR9702	WA	17.5	30.7	267	45	1,357	213	071	71	1.5	4.8
BR9703	GT	15.1	24.0	317	82	1,332	11.3	042	81	1.4	6.1
BR9705	GT	15.2	24.9	322	92	1,314	13.8	042	81	1.5	4.7
BR9707	GT	15.1	24.3	314	85	1,321	9.9	041	55	1.4	5.2
BR9709	CV	14.6	22.4	319	119	1,294	7.1	112	52	1.4	4.6
BR9710	CV(OTW)	13.3	20.6	740	219	1,163	2.7	051	71	1.0	4.4
BR9712	TAS	12.6	21.0	558	178	1036.4	4.3	051	67	1.0	5.1
BR9713	GL	15.4	19.8	443	170	1,143	4.1	021	64	1.3	5.1
BR9715	TAS	12.2	19.7	547	187	954	23.4	051	63	N/A	N/A

second (so called "units") error term was fitted (Dutkowski et al. 2002).

All terms were fitted in a single model, where the spatial, extraneous (e.g. assessor) and treatment effects are estimated simultaneously. Diagnostic tools, variogram and plots of spatial residuals as recommended by Gilmour et al. (2009) were used to detect extraneous effects. Spatial adjustment to raw data was done for surface sum experimental design (i.e. replicate and plots) and/or spatial variation. Using the adjusted data, each trial was first analysed as univariate (i.e. single-site) in order to estimate the genetic variance components, before proceeding with multisite analyses.

# 2.3.2 Statistical model, variance components and genetic parameters

The following reduced (i.e. parental) linear mixed-effects model was used for multi-environment analyses:

## $y = X\tau + Zu_p + Zu_f + e$

where y is a vector of observations formed by stacking the data for each trial  $j \mathbf{y} = (\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_i^T)^T, \mathbf{b}$  is a vector of fixed effects (i.e. site mean),  $u_p$  is a vector of random additive genetic effects (i.e. female parent),  $u_f$  is a vector of random non-additive effects (i.e. full-sib family), and e is a vector of random residual terms. X and Z are known incidence matrices relating





the observations in y to effects in b, and p, f, respectively. The random effects in the model were assumed to follow a multivariate normal distribution with means and variances defined by:

$$u_p \sim N(\mathbf{0}, \sigma_p^2 A), u_f \sim N(\mathbf{0}, \sigma_f^2 I) \text{ and } e \sim N(\mathbf{0}, \sigma_e^2 I)$$

where  $\mathbf{0}$  is a null vector;  $\mathbf{A}$  is the numerator relationship matrix, which describes the additive genetic relationships among individual genotypes; I is identity matrix, with order equal to the number of full-sib families or number of trees;  $\sigma_p^2$  is additive parental variance;  $\sigma_f^2$  is the nonadditive variance between full-sib families; and  $\sigma_e^2$  is the residual variance.

For all models, Restricted Maximum Likelihood (REML) derived variance and covariance estimates using the ASReml program (Gilmour et al. 2009) were constrained to fall within the theoretically possible range; variance components estimates were constrained to be greater than zero. Two separate analyses were used to test for the significance of variance components and the difference between the 2×log-likelihoods (i.e. likelihood ratio) of a full model that included a term and a reduced model without the term was obtained, and tested against a one degree of freedom chi-squared distribution to estimate p values.

Individual heritability  $(h^2)$  and proportion of dominance variance  $(d^2)$  were calculated following Costa e Silva et al. (2004):

$$h^2 = \frac{\sigma_a^2}{\sigma_p^2} = \frac{4 \times \sigma_p^2}{2 \times \sigma_p^2 + \sigma_f^2 + \sigma_e^2}$$

$$d^2 = \frac{\sigma_d^2}{\sigma_P^2} = \frac{4 \times \sigma_f^2}{2 \times \sigma_p^2 + \sigma_f^2 + \sigma_e^2}$$

where  $\sigma_a^2$  is additive genetic variance component (4\*general combining ability),  $\sigma_d^2$  is non-additive variance component (4\*specific combining ability),  $\sigma_P^2$  is phenotypic variance, and the other parameters as defined previously. Standard errors of genetic correlations were derived based on Taylor series approximation using the R pin function (White 2013). Considering regionalisation of breeding in southern Australia, the numerator of heritability becomes  $\sigma_a^2 + \sigma_{a \times r}^2$  as opposed to only  $\sigma_a^2$  with no regionalisation, where and  $\sigma_{a \times r}^2$  is genotype by region variance (Atlin et al. 2000).

The site-site genetic correlations were confounded with the age-age correlations, but the assumption was that the betweensite genetic correlations were not significantly influenced by the age of assessments, because the age-age correlations for a narrow range of ages are usually very high (e.g. Li and Wu 2005).

## 2.3.3 Extended factor analyses

The availability of closely related trees planted across a wide range of environments provided genetic links and allowed estimation of across-site variance and covariance components. Flexible (i.e. unstructured) covariance can account for both scale and rank interactions, but with many environments, the estimation of an unstructured covariance matrix is not feasible (Cullis et al. 2014). Mixed model analysis with an extended factor analytic (XFA) variance structure for the  $G \times E$  effects and separate variance for the errors for each trial was used here (Gilmour et al. 2009). The XFAk form is a sparse formulation that requires an extra k levels to be inserted into the mixed model equations for the k factors.

The reduced (i.e. parental) mixed model in extended factor analytic model was used to efficiently model variance structure of the additive (and non-additive)  $G \times E$  effects (e.g. Beeck et al. 2010):

$$var(\boldsymbol{u}_p) = \boldsymbol{A} \otimes \boldsymbol{G}_t$$

 $var(\boldsymbol{u}_f) = \boldsymbol{I} \otimes \boldsymbol{G}_{tf}$ 

$$\boldsymbol{G}_{ts} = \boldsymbol{\Lambda}_{ts}\boldsymbol{\Lambda}_{ts}^{T} + \boldsymbol{\psi}_{ts}, s = p, f$$

where: A is the relationship matrix with 1+ inbreeding coefficient as diagonal and coefficient of parentage as off-diagonal elements,  $G_{jp}$  is additive genetic variance-covariance matrix for parental effect in trial j,  $G_{jf}$  is non-additive genetic variance-covariance matrix for full-sib family effect in trial j, I is an identity matrix,  $\Lambda$  is a  $t \times k_s$  matrix of trial loadings (t being number of trials k being the number of factors included in the model), and  $\Psi$ =diag ( $\psi_j$ ) and where  $\psi_j$  is the specific variance for the jth trial.

#### 2.3.4 Heat map and hierarchical clustering

Cullis et al. (2010) recommended tools for exploring  $G \times E$ interaction based on a factor analytic model. A heat map representation of the estimated genetic correlation matrix was used, with rows and columns ordered appropriately. A useful ordering is obtained by cluster analysis of trials. The cluster analysis used dissimilarity computed as  $1-\rho_{ij}$  (where  $\rho_{ij}$ is the estimated genetic correlation between trials i and j) as a measure of distance between trials, and hierarchical clustering algorithm (i.e. complete linkage method with Euclidean distance measure), that is implemented within the *hclust* package in R (R Development Core Team 2011). The relationship between-site clustering and environment was examined by correlating site loadings for XFA1 and XFA2 with climate and soil variables (Costa e Silva et al. 2006).

Principal component (PCA) ordination was used to plot the variables of environmental data matrix in two-dimensional representations. The BiplotGUI package provides a graphical user interface for the construction and manipulation of ordination plots in R (La Grange et al. 2009). The trials are represented as points and climate variables as axes, with coordinates on principal component scales.

## **3 Results**

The heritability  $(h^2)$  and proportion of dominance variance  $(d^2)$  for DBH for the 20 trials are given in Table 3. Trials BR9610 (a Spring Needle Cast trial in north-west Tasmania), BR9709 (Central Victoria) and BR9712 (in north-east Tasmania) had no additive variance, and several other trials had insignificant, low dominance variance, presumably due to leaf diseases and/or measurement errors. In trials where the sample size was sufficient to obtain precise estimates (e.g. BR9601, BR9604 and BR9608), the  $d^2$  was similar in magnitude to  $h^2$ .

The relative magnitude of the variance components for genotypes, regions, sites within regions are given in Table 4. Overall, the SCA was higher in magnitude than GCA, but SCA by environment interaction was similar in magnitude GCA by environment variances. The Site(Region) by GCA



**Table 3** Single-site analyses of variance component for DBH, including trial ID, number of parents and progeny trees tested, narrow sense heritability  $(h^2)$  and proportion of dominance variance  $(d^2)$ , and standard errors (se) of the estimates

Trial_ID	Region	Parents	Progeny	$h^2$	±se	$d^2$	±se
BR9601	GT	128	2,883	0.29	0.09	0.30	0.07
BR9604	GT	54	2,357	0.39	0.12	0.42	0.10
BR9606	GT	54	2,581	0.37	0.10	0.29	0.07
BR9608	CV	148	3,124	0.29	0.08	0.30	0.07
BR9610	TAS	43	766	0.04	0.04	0.04	0.09
BR9611	GL	134	2,757	0.17	0.06	0.24	0.07
BR9613	GL	55	2,341	0.12	0.05	0.12	0.05
BR9614	TAS	55	2,310	0.15	0.05	0.11	0.05
BR9615	TAS	123	2,529	0.22	0.08	0.18	0.07
BR9617	MV	66	1,843	0.28	0.9	0.13	0.06
BR9701	WA	171	2,110	0.20	0.07	0.12	0.06
BR9702	WA	172	1,768	0.12	0.06	0.17	0.07
BR9703	GT	42	926	0.12	0.06	0.17	0.07
BR9705	GT	103	630	0.27	0.10	0.00	0.00
BR9707	GT	41	992	0.19	0.05	0.02	0.05
BR9709	CV	101	723	0.08	0.06	0.01	0.08
BR9710	CV(OTW)	44	1,096	0.29	0.11	0.00	0.00
BR9712	TAS	44	891	0.07	0.08	0.09	0.10
BR9713	GL	145	1,049	0.15	0.07	0.00	0.00
BR9715	TAS	152	994	0.16	0.09	0.06	0.11

The estimates that were significantly different from 0 at p=0.05 level of probability are given in **bold** 

interaction was roughly 75 % in magnitude of the additive variance (i.e. GCA) and Site(Region) by SCA was 70 % of non-additive variance (i.e. SCA). The region by GCA variance suggests that regions are capturing only a significant amount (i.e. ~55 %) of G×E interaction relative to site within region by GCA interaction. Heritability within NPI regions increased relative to across NPI regions from  $0.09\pm0.03$  to  $0.13\pm0.03$  and further to  $0.16\pm0.03$  for heritability based on hierarchical site clustering based on additive genetic correlation matrix.

 Table 4
 Combined-site variance component analysis: REML estimates of variance and estimates as percentage of total genetic variance

Model term	Estimate	Percent of total genetic variance
GCA	20.9	22.4 %
SCA	23.9	25.6 %
Region×GCA	8.7	9.3 %
Site (Region)×GCA	15.9	17.0 %
Region×SCA	7.2	7.7 %
Site (Region)×SCA	16.2	18.0 %

The analyses excluded trials BR9610, BR9709 (low genetic variance) and trial BR9617 (single trial representing Murray Valley region)

Description Springer



Therefore, the genetic gain from regionalisation of breeding can be expected to be significant.

The trial-trial additive and non-additive correlation matrix for DBH obtained using extended factor analytic with three factors (XFA3) analysis is given in Table 5. The matrix for additive and non-additive correlations is represented by heat maps with a dendrograms added to the left side and to the top (Fig. 2). Rows and columns (i.e. trial IDs) were reordered based on row or column means within restrictions imposed by the dendrogram. The dendrogram with a cut-off at height of 0.5 divided the trials into clusters to maintain high estimated genetic correlations for all pairs of trials within each cluster. At the additive level, this resulted in the formation of 4 clusters with 2, 2, 5 and 10 trials, and 1 singleton (i.e. BR9710). At the non-additive level, this resulted in the formation of 3 clusters with 2, 11, and 6 trials, and 1 singleton (i.e. BR9701). The heat map shows that the pair-wise genetic correlation between trials within clusters was greater than between clusters.

Although, the XFA3 adequately modelled the G×E (i.e. explained 90.5 and 87.0 % genetic variance for additive and non-additive effects, respectively) interpretation of the effects may be problematic for sites that exhibited a low or zero proportion of genetic variance (i.e.  $h^2=0$  or  $d^2=0$ ). For example, the sites BR9705, BR9707, BR9709, BR9710 and BR9713, with relatively small non-additive variance, had 100 % variance accounted for from the XFA3 model and correlation fixed at  $r_g=1$ . The first latent variable (i.e. XFA1 loading) separated the sites with low SCA variance, whereas the second latent variable contrasted TAS (BR9712, BR9615, BR9614) with sites on Mainland (BR9703, BR9613, BR9604).

Information including the identification and location of the trials in each cluster are given in Table 6. At the additive level, cluster 1 consisted of three TAS trials and one trial representing MV region (i.e. BR9617). The trial BR9710 from CV(OTW) was a singleton. Cluster 4 consisted of two trials from WA BR9701 and BR9702, two TAS trials and an inconsistent (i.e. water-logged) trial in GT (BR9604). Finally, cluster five consisted mainly of trials in GT, two trials in CV and two trials in GL (Table 6).

At the non-additive level, cluster 1 consisted again of trial BR9610 in TAS and trial BR9617 in MV. The large cluster 3 included trials with low SCA variance (i.e. BR9705, BR9707, BR9709, BR9710 and BR9713), and some other trials with high correlations with that group. Trial BR9701 from WA was the singleton. Cluster 4 consisted of trials from five different regions. Overall, the pattern of  $G \times E$  for the non-additive effects was more complex than for the additive effects (Table 6).

To further examine  $G \times E$  patterns, average additive pairwise correlations within and between current breeding zones (i.e. NPI regions) are presented in Table 7. Two trials with low

	BR9601	BR9604	BR9606	BR9608	BR9610	BR9611	BR9613	BR9614	BR9615 1	BR9617	BR9701	BR9702	BR9703	BR9705	BR9707	BR9709	BR9710	BR9712	BR9713	BR9715
BR9601	1	0.52	0.87	0.72	0.16	0.75	0.52	0.54	0.37	0.24	0.68	09.0	0.68	0.87	0.89	0.89	0.20	0.60	0.65	0.54
BR9604	0.77	1	0.53	0.39	-0.12	0.53	0.27	0.53	0.25	0.08	0.51	0.60	0.33	0.54	0.59	0.59	0.20	0.57	0.29	0.52
BR9606	0.75	0.72	1	0.86	0.34	0.82	0.56	0.63	0.46	0.35	0.69	0.51	0.71	0.95	0.98	0.98	0.07	0.53	0.69	0.48
BR9608	0.64	0.49	0.57	1	0.41	0.68	0.46	0.53	0.42	0.34	0.54	0.31	0.69	0.81	0.82	0.82	0.00	0.34	0.63	0.31
BR9610	0.30	0.05	0.46	0.43	1	0.14	0.06	0.20	0.22	0.36	-0.07	-0.45	0.29	0.16	0.19	0.19	-0.52	-0.38	0.21	-0.33
BR9611	0.74	0.80	0.64	0.50	-0.04	1	0.46	0.59	0.37	0.22	0.65	0.6	0.61	0.81	0.84	0.84	0.17	0.6	0.56	0.54
BR9613	0.62	0.62	0.38	0.51	-0.31	0.73	1	0.08	0.14	0.13	0.42	0.38	0.59	0.64	0.59	0.59	0.31	0.4	0.67	0.33
BR9614	0.66	0.57	0.59	0.59	0.29	0.58	0.56	1	0.48	0.23	0.40	0.45	0.16	0.45	0.61	0.61	-0.27	0.41	-0.09	0.43
BR9615	0.70	0.61	0.61	0.61	0.28	0.62	0.61	0.60	1	0.22	0.23	0.13	0.23	0.35	0.41	0.41	-0.20	0.14	0.11	0.15
BR9617	0.17	-0.19	0.12	0.49	0.68	-0.11	0.08	0.30	0.29	1	0.09	-0.09	0.22	0.25	0.27	0.27	-0.23	-0.07	0.14	-0.04
BR9701	0.43	0.57	0.62	0.13	0.34	0.38	-0.16	0.21	0.22	-0.41	1	0.73	0.59	0.78	0.77	0.77	0.31	0.71	0.63	0.62
BR9702	0.75	06.0	0.71	0.39	-0.08	0.83	0.60	0.52	0.56	-0.39	0.68	1	0.4	0.66	0.68	0.68	0.61	0.94	0.46	0.82
BR9703	0.78	0.62	0.82	0.71	0.76	0.54	0.29	0.65	0.67	0.47	0.54	0.54	1	0.83	0.77	0.77	0.29	0.44	0.83	0.36
BR9705	06.0	0.85	0.84	0.72	0.33	0.82	0.69	0.74	0.77	0.19	0.48	0.83	0.86	1	0.98	0.98	0.32	0.67	0.85	0.58
<b>BR9707</b>	0.90	0.85	0.84	0.72	0.33	0.82	0.69	0.74	0.77	0.19	0.48	0.83	0.86	1.00	1	1.00	0.23	0.68	0.74	0.61
<b>BR9709</b>	0.90	0.85	0.84	0.72	0.33	0.82	0.69	0.74	0.77	0.19	0.48	0.83	0.86	1.00	1.00	1	0.23	0.68	0.74	0.61
<b>BR9710</b>	0.00	0.85	0.84	0.72	0.33	0.82	0.69	0.74	0.77	0.19	0.48	0.83	0.86	1.00	1.00	1.00	1	0.58	0.51	0.46
<b>BR</b> 9712	0.73	0.86	0.68	0.42	-0.06	0.80	0.63	0.53	0.57	-0.30	0.58	0.94	0.53	0.81	0.81	0.81	0.81	1	0.5	0.77
<b>BR</b> 9713	06.0	0.85	0.84	0.72	0.33	0.82	0.69	0.74	0.77	0.19	0.48	0.83	0.86	1.00	1.00	1.00	1.00	0.81	1	0.39
BR9715	0.56	0.59	0.50	0.40	0.03	0.58	0.52	0.44	0.47	-0.04	0.31	0.61	0.45	0.63	0.63	0.63	0.63	0.59	0.63	1
Genetic	correlatic antly diffe	ons above srent from	diagonal a unity at $p$ =	tre for add =0.01 leve	litive (GC el of proba	A) and be (hility)	low diag	onal non-	additive (S	CA) effe	cts. Corre	lations gi	ven in bol	ld were th	rree stand	ard deviati	ions away	y from co	rrelation o	f 1 (i.e.

Table 5 Trial-trial additive correlation matrix for DBH, obtained using XFA3 analysis



🖄 Springer





heritability (i.e. BR9709, BR9610) and five trials with low dominance proportion (BR9705, BR9707, BR9709, BR9710 and BR9713) were excluded. It was evident that the correlations between GT and CV regions and between CV and GL regions are not lower than the within region correlations. TAS had low correlation both with trials in other regions and within TAS. The trial BR9617 in MV region had very low or negative correlation with all the other regions. The trial BR9710 in OTW region had low correlation with GT and a negative correlation with MV trials.

Climate variables for the 20 sites of the BR 1996/97 progeny trial series were summarised using PCA-based ordination (Fig. 3). The first two principal components of the biplot explained 78 % of the variation. Trials from TAS, CV and GT regions were separated on the axes related to precipitation growing season (PGS), annual precipitation (AP), and relative humidity at high temperature (RHHT). The two trials in WA were separated from the other trials on the axes related to aridity, i.e. high evaporation (TE), solar radiation (MR), and high evaporation to precipitation ratio (AIX).

Additive					Non-Additi	ve	
Cluster	Trial_ID	State	Region	Cluster	Trial_ID	State	Region
1	BR9610	TAS	TAS	1	BR9610	TAS	TAS
	BR9617	NSW	MV		BR9617	NSW	MV
2	BR9710	VIC	CV(OTW)	2	BR9701	WA	WA
3	BR9614	TAS	TAS	3	BR9608	VIC	CV
	BR9615	TAS	TAS		BR9614	TAS	TAS
4	BR9604	VIC	GT		BR9615	TAS	TAS
	BR9701	WA	WA		BR9601	SA	GT
	BR9715	TAS	TAS		BR9713	VIC	GL
	BR9702	WA	WA		BR9710	VIC	CV(OTW)
	BR9712	TAS	SE		BR9709	VIC	CV
5	BR9613	VIC	GL		BR9705	SA	GT
	BR9703	SA	GT		BR9707	SA	GT
	BR9713	VIC	GL		BR9606	VIC	GT
	BR9608	VIC	CV		BR9703	SA	GT
	BR9601	SA	GT	4	BR9702	WA	WA
	BR9611	VIC	GL		BR9611	VIC	GL
	BR9606	VIC	GT		BR9604	VIC	GT
	BR9705	SA	GT		BR9712	TAS	SE
	BR9707	SA	GT		BR9613	VIC	GL
	BR9709	VIC	CV		BR9715	TAS	NE

**Table 6** Identification and location of the trials in clusters in heatmap (Fig 2a, b) from left to rightand from top to bottom

 $\underline{\textcircled{O}}$  Springer



**Table 7** Average genetic correlation coefficients for diameter at the regional level: Green Triangle (GT), Central Victoria (CV), Gippsland (GL), Tasmania (TAS) and Western Australia (WA), and single trials in the Murray Valley (MV) and Otways (OTW) regions

	GT	CV	GL	TAS	WA	MV	OTW
GT	0.82\0.86	0.85	0.76	0.58	0.72	0.23	0.39
CV	0.78	0.72\0.77	0.83	0.59	0.73	0.3	0.56
GL	0.73	0.71	0.75\0.77	0.57	0.71	0.16	0.69
TAS	0.60	0.57	0.52	0.37\0.46	0.69	0.14	0.50
WA	0.64	0.46	0.49	0.42	0.68\0.89	0.00	0.51
MV	0.16	0.34	0.05	0.19	-0.40	NA	-0.23
OTW	0.91	0.86	0.84	0.66	0.66	0.19	NA

Genetic correlations above diagonal are additive (GCA) and below diagonal non-additive (SCA) effects. On diagonal are within-region SCA\GCA correlations

The correlations between climatic attributes and the latent variates indicated the contributions of the individual climatic variables to differences between trials. At the additive level, XFA1 and XFA2, respectively, were moderately to highly correlated to total evaporation (TE) (r=-0.55, r=-0.55), annual precipitation (AP) (r=0.62, r=0.58), PGS (r=0.48, r=0.74), mean monthly temperature (MT) (r=-0.31, r=0.59), mean temperature growing season (MTGS) (r=-0.19, r=-0.59), mean monthly maximum temperature (MAXT) (r=-0.19, r=-0.60), mean monthly minimum temperature

(MINT) (r=0.48, r=-0.50), vapour pressure (VPP) (r=-0.37, r=-0.63). On the other hand, the correlations with soil characteristics were significant only for soil depth (r=0.63, r=-0.08), bulk density of layer 3 (r=-0.56, r=0.05), available water content in the top layer (r=0.18, r=0.50).

At the non-additive level, the factor analytic site loadings on correlation scale for XFA1 and XFA2, respectively, were moderately to highly correlated to PGS (r=-0.52, r=-0.41), AP (r=-0.57, r=-0.32), PDQ (r=-0.41, r=-0.45), aridity index (AIX) (r=-0.07, r=0.52), MINT (r=0.42, r=-0.39)

Fig. 3 PCA ordination based on climate variables and grouping of 20 trial site locations as NPI regions: Western Australia (WA), Green Triangle (GT), Murray Valley (MV), Central Victoria (CV), Otways (OTW), Central Gippsland (GL), and Tasmania (TAS). Axes represent of most distinguishing climate variables: annual precipitation (AP), precipitation growing season (PGS), relative humidity at highest temperature (RHHT), relative humidity low temperature (RHLT), mean solar radiation (MTGS), and aridity index (AIX) (i.e. ratio of evaporation to precipitation)





and VPP (r=0.41, r=0.49). The correlations with soil characteristics were significant only for soil pH (r=0.40, r=-0.08).

## 4 Discussion

For the 20 trials from the STBA BR1996/97 series, heritability  $(h^2)$  of DBH was within the range observed for radiata pine in Australia (Wu et al. 2008). The proportion of dominance variance  $(d^2)$  was similar in magnitude as  $h^2$ . In some previous studies of radiata pine, the magnitude of  $d^2$  was less than  $h^2$  (e.g. Wu and Matheson 2005; Gapare et al. 2010). However, relative magnitude of  $d^2$  and  $h^2$  similar to the one observed in the current study was observed in other pines (e.g. in 26-year-old in Scots pine in Waldmann et al. 2008). The result suggests that non-additive components genetic variance can be as important as the additive component (Wu and Matheson 2004).

The relative magnitudes of the variance components also suggest that the current deployment regions are capturing significant amount of  $G \times E$  interaction of total  $G \times E$  variance. Nevertheless, crossover interaction of breeding values can be expected between trials within a certain region and between different regions, even for regions in different states. For example, at both additive and non-additive level, the correlations between trials within TAS were lower than between trails in TAS and trials in other regions (except one trial in MV region). Nevertheless, our analyses of DBH at the regional level confirmed the results of Baltunis et al. (2010) that there is a significant  $G \times E$  between Tasmania and mainland Australia.

This study also confirmed findings of Wu and Matheson (2005) in regards to  $G \times E$  between two trials in NSW and those in other parts of southern Australia. A more recent study by Gapare et al. (2012) grouped trials according to rainfall and altitude. High-altitude sites had also higher rainfall and lower temperature. Those trials formed two groups on based on additive and provenance site-site correlations. This grouping was also consistent with the results of Raymond (2011), where higher altitude, higher rainfall sites were distinctly different from warmer, drier sites within NSW.

At the additive level, the trials within TAS, one trial in MV region in NSW and one trial in the OTW region in Victoria had either low or negative correlations with other trials. Nevertheless,  $G \times E$  was present even between two sites (i.e. BR9604 and BR9608) within the otherwise relatively uniform GT region, and such interaction may be caused by microenvironmental causes (in this case, water-logging). Generally, as site-site correlations within regions increase to the same level as that between regions, the usefulness of regionalisation decreases (Pederick 1990). Genetic response to selection can be calculated as:  $G=i\sigma h^2$ , where *i* is selection intensity,  $\sigma$  is phenotypic standard deviation, and  $h^2$  is

Dispringer



heritability. As heritability increased significantly for withinregion selection regionalisation seems to be justified.

Generally, the results presented here confirm the results previously found for additive gene effects. However, at the non-additive level, the pattern of site-site genetic correlations was less clear than at the additive level. Nevertheless, the overall pattern of regional (i.e. average) non-additive genetic correlations was similar to additive. Although non-additive genetic variation can be very important (e.g. Waldmann et al. 2008), there are no examples in forestry literature comparing environmental causes of both additive and non-additive G×E.

In some cases, certain climate, soil or other variables can be identified, which may then be utilised to delineate specific deployment regions. The BR1996/97 series of trials covered a wide geographic range in southern Mainland Australia and Tasmania, and covered a wide range of climates and soils. When geo-climatic variables for pairs of trials were used as predictor variables, at the large transcontinental scale, XFA1 and XFA2 site loadings were moderately correlated to climate variables, in particular annual rainfall. The results linked G×E to aridity with regards to the two Western Australian trials. Further modelling of genetic correlations based on environmental variables is underway. The modelling is involving all genetically connected pairs of STBA trials in addition to the BR trial series analysed here, and the results from those analyses will be used for re-defining breeding and deployment regions in southern Australia.

## **5** Conclusions

- 1. Based on previous studies and this study, significant G×E for diameter growth can be expected between Tasmanian and Mainland sites, and within Tasmania itself.
- There were indications that the sites in Murray Valley region in NSW and Otway region in Victoria may exhibit G×E interaction with other regions; however, this is based only on individual trials.
- 3. Heritability increased significantly for within-region selection and regionalisation seems to be justified.
- 4. The G×E interaction at transcontinental scale can be correlated to the climate variables, primarily to rainfall and temperature. However, the drivers may also be related to smaller scale environmental variation (i.e. soil and terrain variation).
- 5. The results presented here can be used as evidence in favour of reconsidering the current breeding and deployment zones. However, further work, using other pairs of genetically connected trials, will give more robust results on which to base  $G \times E$  regionalisation.

Acknowledgments The authors thank the Southern Tree Breeding Association, Australia, in particular Dr. Tony McRae, Peter Buxton and

David Pilbeam, for access to trial data. The authors thank Colin Matheson, David Bush and Xinguo Li for the initial review of the manuscript. Funding for this research was provided by the Forest and Wood Products Australia, Radiata Pine Breeding Company, New Zealand, and Timberlands Pacific Southern Tree Breeding Association, Australia.

## References

- Atlin GN, Baker RJ, McRae KB, Lu X (2000) Selection response in subdivided target regions. Crop Sci 40:7–13
- Baltunis B, Gapare WJ, Wu HX (2010) Genotype by environmental interaction for DBH, wood density, branch angle, branch size, and stem straightness in eight young *Pinus radiata* D. Don trials in Australia. Silvae Genet 59:113–123
- Beeck C, Cowling WA, Smith AB, Cullis BR (2010) Analysis of yield and oil from a series of canola breeding trials. Part I: Fitting factor analytic mixed models with pedigree information. Genome 53:992– 1001
- Costa E Silva J, Borralho NMG, Potts BM (2004) Additive and nonadditive genetic parameters from clonally replicated and seedling progenies of *Eucalyptus globulus*. Theor Appl Genet 108:1113– 1119
- 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
- Cullis BR, Smith AB, Beeck C, Cowling WA (2010) Analysis of yield and oil from a series of canola breeding trials. Part II: exploring VxE using factor analysis. Genome 53:1002–1016
- 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
- Dutkowski GW, Costa E Silva J, Gilmour AR, Lopez GA (2002) Spatial analysis methods for forest genetic trials. Can J For Res 32:2201– 2214
- Gapare WJ, Ivković M, Baltunis BS, Matheson AC, Wu HX (2010) Genetic stability of wood density and diameter in Pinus radiata D. Don plantation estate across Australia. Tree Genet Genomes 6:113–125
- Gapare WJ, Ivkovich M, Dutkowski GW, Spencer DJ, Buxton P, Wu HX (2012) Genetic parameters and provenance variation of *Pinus radiata* D. Don. 'Eldridge collection' in Australia 1: growth and form traits. I Tree Genet Genomes 8:391–407
- Gavran M, Parsons M (2011) Australian plantation statistics 2011. Sciences, Bureau of Agricultural and Resources Economics and

- Gilmour AR, Gogel BJ, Cullis BR, Thompson R. (2009) ASReml User Guide Release 3.0 VSN International Ltd, Hemel Hempstead, UK
- La Grange A, Le Roux N, Gardner-Lubbe S (2009) BiplotGUI: interactive biplots in R. J Stat Software 30:1–37
- Li L, Wu HX (2005) Efficiency of early selection for rotation-aged growth and wood density traits in Pinus radiata. Can J For Res 35: 2019–2022
- Matheson AC, Cotterll PP (1990) Utility of genotype  $\times$  environment interactions. For Ecol Manag 30:159–174
- Pederick L (1990) Family × site interactions in Pinus radiata in Victoria, Australia and Implications for breeding strategy. Silvae Genet 39: 134–140
- Raymond CA (2011) Genotype by environment interactions for *Pinus* radiata in NSW, Australia. Tree Genet Genomes 7:819–833
- R Development Core Team (2011) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna
- Schreiber SG, Ding C, Hamann A, Hacke UG, Thomas BR, Brouard JS (2011) Frost hardiness vs. growth performance in trembling aspen: an experimental test of assisted migration. J App Ecol 50:939–949
- Turner J, Lambert MJ, Hopmans P, McGrath J (2001) Site variation in Pinus radiata plantations and implications for site specific management. New For 21:249–282
- Waldmann P, Hallander J, Hoti F, Sillanpa MJ (2008) Efficient Markov Chain Monte Carlo Implementation of Bayesian Analysis of Additive and Dominance Genetic Variances in Noninbred Pedigrees. Genet 179:1101–1112
- White I (2013) The R pin function. Available from: http://www. homepages.ed.ac.uk/iwhite//asreml/useofpin.pdf Accessed: 11/2013
- White TL, Adams TW, Neale DB (2007) Forest Genetics, CAB International, 682 pp
- Wilford J (2012) A weathering intensity index for the Australian continent using airborne gamma-ray spectrometry and digital terrain analysis. Geoderma 183/184:124–142
- Wu HX, Matheson AC (2004) General and specific combining ability from partial diallels of radiata pine: implications for utility of SCA in breeding and deployment populations. Theor App Genet 108:1503– 1512
- Wu HX, Matheson AC (2005) Genotype by environment interaction in an Australia-wide radiata pine diallel mating experiment: implications for regionalised breeding. For Sci 51:29–40
- Wu HX, Ivkovic M, Gapare WJ, Matheson AC, Baltunis BS, Powell MB, McRae TA (2008) Breeding for wood quality and profit in radiata pine: a review of genetic parameters and implications for breeding and deployment. NZ J For Sci 38:56–87
- Yan W, Kang MS (2003) GGE biplot analysis: a graphical tool for breeders, geneticists and agronomists. CRC Press, Washington, 271pp

