Journal of Theoretical Biology 462 (2019) 283–292

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtb

Analysis of phenotypic- and Estimated Breeding Values (EBV) to dissect the genetic architecture of complex traits in a Scots pine three-generation pedigree design

Ainhoa Calleja-Rodriguez[a,b], Zitong Li[c,d], Henrik R. Hallingbäck[a,e], Mikko J. Sillanpää[f], Harry X. Wu[a], Sara Abrahamsson[b], Maria Rosario García-Gil[a,*]

[a] Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Science, Umeå SE-901 83, Sweden
[b] Skogforsk, Box 3, Sävar SE-91821, Sweden
[c] Melbourne Integrative Genomics and School of Mathematics and Statistics, the University of Melbourne, Parkville, Victoria 3010, Australia
[d] Organismal and Evolutionary Biology Research Programme, University of Helsinki, Helsinki FI-00014, Finland
[e] Department of Plant Biology, Uppsala BioCenter, Linnean Centre for Plant Biology, Swedish University of Agricultural Science, Uppsala SE-75007, Sweden
[f] Department of Mathematical Sciences and Biocenter Oulu, University of Oulu, Oulu FI-90014, Finland

ARTICLE INFO

Article history:
Received 20 July 2018
Revised 17 October 2018
Accepted 9 November 2018
Available online 10 November 2018

Keywords:
Scots pine
Quantitative trait locus
LASSO
Adaptive traits

ABSTRACT

In forest tree breeding, family-based Quantitative Trait Loci (QTL) studies are valuable as methods to dissect the complexity of a trait and as a source of candidate genes. In the field of conifer research, our study contributes to the evaluation of phenotypic and predicted breeding values for the identification of QTL linked to complex traits in a three-generation pedigree population in Scots pine (*Pinus sylvestris* L.). A total of 11 470 open pollinated F_2-progeny trees established at three different locations, were measured for growth and adaptive traits. Breeding values were predicted for their 360 mothers, originating from a single cross of two grand-parents. A multilevel LASSO association analysis was conducted to detect QTL using genotypes of the mothers with the corresponding phenotypes and Estimated Breeding Values (EBV). Different levels of genotype-by-environment ($G \times E$) effects among sites at different years, were detected for survival and height. Moderate-to-low narrow sense heritabilities and EBV accuracies were found for all traits and all sites. We identified 18 AFLPs and 12 SNPs to be associated with QTL for one or more traits. 62 QTL were significant with percentages of variance explained ranging from 1.7 to 18.9%. In those cases where the same marker was associated to a phenotypic or an *ebv*QTL, the *ebv*QTL always explained higher proportion of the variance, maybe due to the more accurate nature of Estimated Breeding Values (EBV). Two SNP-QTL showed pleiotropic effects for traits related with hardiness, seed, cone and flower production. Furthermore, we detected several QTL with significant effects across multiple ages, which could be considered as strong candidate loci for early selection. The lack of reproducibility of some QTL detected across sites may be due to environmental heterogeneity reflected by the genotype- and QTL-by-environment effects.

1. Introduction

Traditionally, Quantitative Trait Loci (QTL) studies in tree species have been based on phenotypic rather than genotypic values (breeding or clonal values). The main limitation of the traditional phenotypic-based QTL detection approach is the design of the experiments that typically involve only one generation, often consisting of a single full-sib family where environmental and genetic factors are confounded (Thavamanikumar et al., 2013; Isik 2014; Hall et al., 2016). Alternatively, phenotypic values can be substituted by Estimated Breeding Values (EBV). This method requires phenotyping progenies of the target (mother) trees and posteriorly computing EBV (using pedigree information) with the purpose of ranking the target trees for the traits under study. Estimation of genetic effects, such as EBV (i.e additive genetic effects), can improve the estimation of genotypic performance and also improve the genotypic value estimation of candidate genotypes in breeding programs (Piepho et al., 2008).

* Corresponding author.
E-mail address: m.rosario.garcia@slu.se (M.R. García-Gil).

https://doi.org/10.1016/j.jtbi.2018.11.007

Since the milestone article by Lander and Botstein (1989) there have been a large number of QTL mapping studies in multiple plant species (Mauricio, 2001). In forest species, QTL mapping have been successful using two very different settings of gathering and analysing data: either family- or population-based (AM, Association Mapping or Linkage Disequilibrium (LD) mapping). AM involves sampling of unrelated trees from one or multiple populations across a geographic area, which can extend across multiple degrees of latitude and longitude. AM overcomes some limitations of family-based QTL mapping by exploiting historical recombination's at population level (Neale and Savolainen 2004). However, AM studies can result in the detection of false positive QTLs due to the effect of underlying complex population structures that are the result of population history and selection (Pritchard et al., 2000). On the contrary, family-based QTL studies are not biased by population history or selection events, and are less affected by environmental heterogeneities (i.e., the progenies are often located at a single plantation with relatively even environment). The main limitation of family-based QTL is that the LD generated following parental crossings can extend across large stretches of DNA, consequently, a significant association between a trait and a molecular markers cannot necessarily be extrapolable to the entire breeding population. This situation is, however, irrelevant for significant associations involving the causal gene (e.g., the molecular markers is within or in the close proximity of the causal gene). For this reason, single-family QTL analysis are still cited as source of candidate genes, especially in the cases where a gene (or molecular marker located at a genic region) of potentially relevant function for the studied trait is underlying a significant QTL. Interestingly, despite the many expected advantages of population-based QTL analysis, the success has been marginal (Mitchell-Olds, 2010), instead both family-based QTL (e.g., Lerceteau et al., 2000; Markussen et al., 2003; Bartholomé et al., 2013) and AM (even if conducted at the whole-genome level) (e.g., Porth et al., 2013; Uchiyama et al., 2013; Bartholomé et al., 2016) studies reveal similar trait architecture with few QTL of larger effect and a larger number of QTL of minor effect.

In this study, we aimed to perform QTL identification based on phenotypic values with EBV assessed in a three generation pedigree in Scots pine. We included in the study multiple years and sites in order to study genotype-by-environment ($G \times E$) interactions. Multi-environment analysis (MET) is useful to detect $G \times E$ interactions and the use of sophisticated models is necessary because the diverse environments may exhibit variance heterogeneity as well as genotype-by-environment interactions. These effects may affect genetic parameter estimations (Ogut et al., 2014; Isik et al., 2017). Initially in plant breeding, factor analytic (FA) models have become more popular and efficient (Kelly et al., 2009; Cullis et al., 2014; Smith et al., 2015) and now in tree breeding (Ivković et al., 2015; Chen et al., 2017) because they provide a good parsimonious approximation to unstructured (US) models (in large dimensions).

We also compared our findings about the architecture of multiple adaptive traits with previous single-family based QTL and AM studies. To achieve our goals, we conducted QTL analysis for traits of economic and ecological value such as growth, tree quality, frost hardiness and survival measured across three different trials and multiple years. The marker information has already been analyzed in association with wood properties by Li et al. (2014) and consist of 508 AFLP markers and 768 SNPs.

2. Material and method

2.1. Plant material

The parents (AC3065 and Y3088) of the full-sib cross are plus-trees from north Sweden, and belong to the Swedish breeding program (see Fig. F.1). Progeny generated from this cross (F_1-generation) were planted in 1988 as one year old seedlings (Abrahamsson et al., 2012). 455 full-sib individuals from this cross were included in the study (trial F485, see Table 1 and Fig. F.1). In 2006, open pollinated (OP) seeds (F_2-generation) from 360 full-sib trees in the F485 trial were collected and grown at the Skogforsk nursery in Sävar. These fertile 360 trees producing female cones in 2006 were the only ones included in the collection as the others were non-fertile at this time. None of the 455 trees produced any male cone. The resulting OP seedlings were planted in three different field trials in 2008, using a complete randomized single-tree plot design. The trials contained an average number of trees per OP family of 9, 10 and 11 for sites F725, F723 and F726 (Table 1 and Fig. F.1), respectively. These trials could, from the technical point of view, be treated as half-sib trials, a common practice in conifer breeding programs, due to the high early inbreeding depression (i.e. the large majority of selfed individuals are aborted at the zygote, seed or young seedlings). Although it should be cautioned that all F_2-progeny were interrelated to a limited extent (single-cousins at the least).

2.2. Phenotypic measurements

Measured traits and the age of the trees at time of measurement for each trial are described here but complementary information can be found in Table C.1.

2.2.1. Trait measurements on the F_1 full-sib cross *(F485)*

Height (Ht) was measured annually between 1996 and 2006 using a telescopic pole from ground to terminal bud. Stem diameter at breast height (DBH) was estimated as the mean of DBH measured in two directions (north-south and east-west direction) in 1996, 1999, 2004 and 2005. Branch diameter (BrD) was scored as the horizontal diameter in two representative branches in the 10th whorl counted from the top. If this branch whorl did not have any living branch, branches were chosen from the 9th whorl or the next consecutive whorl with living branches. The measurements were taken at a maximum of two cm from the stem with a digital calliper, and were done in 1996, 1999 and 2004. Branch angle (BrA) was measured, in two representative branches, between the stem and the direction of the branch at the base on a scale of 360°. Branches with a representative angle of all the branches present in the tree were chosen for BrA. The two representative branches were not necessarily the same branches that were measured for BrD but were located in the same whorl. These measurements were scored in 1996, 1999 and 2004. Ht measured between years 1996 and 1999, DBH, BrA and BrD measured at years 1996 and 1999, for individuals 1–93 have been previously published by Lerceteau et al. (2000) and Lerceteau et al. (2001).

Needles, from the trees, as close to the top as possible were collected and frozen to calculate hardiness (Ct) according to Nilsson and Walfridsson (1995) in 1992, 1997, 1999 and 2004. Presence or absence of female flower production (FP) was scored as a binary character: 1 as present and 0 as absence. Weight of 1000 seeds (W) in grams and number of cones per tree (CO) were counted exhaustively at the time of cone collection and seed extraction in 2006. The number of seeds per cone (nSC) were scored as: $nSC = Seed\ number\ per\ tree/CO$.

Table 1
Field trials´description.

Trial nr	Trial id	Location	LAT (N)	LON (E)	ALT (m)	Gen	Year	N	MAT(°C)	MAP(mm)
1	F485	Flurkmark	64.03	20.5	115	F_1	1988	1000	4	762.7
2	F723 (Site1)	Hennan	62.031	15.648	375	F_2	2008	3804	2.8	704.8
3	F725 (Site2)	Sollefteå	63.197	17.272	110	F_2	2008	3532	3.8	618.5
4	F726 (Site3)	Kebbeberget	64.105	17.674	400	F_2	2008	4134	1.5	731.9
5	Sävar	Sävar	63.895	20.549	65	F_2	2007	8640	4.3	715.8

Trial nr, number of the trial; Trial id, trial identification; LAT, latitude; LON, longitud; ALT, altitude; Gen, generation; Year, plantation year; N, number of trees in each trial; MAT(°C), mean annual temperature (in Celsius degrees) averaged from years 2008 to 2016; MAP(mm), mean annual precipitation (measured in mm) averaged from years 2008 to 2016.; MAT and MAP obtained from the Swedish database http://luftweb.smhi.se.

2.2.2. Trait measurements on the F_2 OP trials (F723, F725 and F726)

All phenotypic measurements are shown in Table C.1 and described here. For simplicity we refer to trial F723 as Site1, trial F725 as Site2 and trial F726 as Site3, hereafter.

Survival ability (Vt) was scored in all the three OP trials at ages one and eight plus at age three for Site3 only. Vt was scored for each tree, according to Persson and Andersson (2003), in four categorical classes: healthy, slightly damaged, severely damaged but alive, and dead. DBH, Ht and BrA were scored at age eight in all OP trials following the same protocol described for the full-sib cross (Table C.1). At the same age, branch quality (BrQ) was scored in nine categorical classes considering the appearance of the entire crown in relation to the tree size and the neighbouring trees. Finally, in Site3 Ht was additionally measured at age three and at the same age Ct was also assessed in the manner previously described for the full-sib cross. An "e" preceding the trait acronym indicates that we are referring to an EBV and a number before the underscore sign refers to the OP trial. For example, eVt3_8 refers to an EBV for survival (eVt) at Site3 (eVt3) measured when the trees were eight years old (eVt3_8).

2.3. DNA extraction and marker development

CTAB (Doyle, 1991) method was used to extract DNA from vegetative buds of the F_1 full-sib individuals. All 455 individuals were genotyped by amplified length polymorphism (AFLP) markers, and a small subset of 90 individuals were genotyped for single nucleotide polymorphism (SNP) markers. Filtering and ordering resulted in a dataset of 153 AFLP markers genotyped for 455 individuals (abbreviated as A set), and a small mixed dataset of 153 AFLP and 166 SNP markers genotyped for the 90 individuals (abbreviated as S + A set). These two datasets were separately used for QTL analysis, for further details see Li et al. (2014).

2.4. Statistical analysis for estimating breeding values (EBV)

2.4.1. Single site spatial analysis

Prior to any genetic analysis the data was adjusted for within-trial environmental effects. We fitted a univariate single site analysis in which we used first a design model, where the environmental effects were modelled only with the experimental design features and an independent error (Dutkowski et al., 2002, 2006), with the objective to investigate the spatial distribution of the residuals. In the case that a non-random distribution was detected for any trait, a second model incorporating a two-dimensional separable autoregressive structure (AR1), where row and column directions were fitted in the model, was applied to the data. This method (Appendix A) divides the residual variance into an independent component and a two dimensional spatially autocorrelated component (Chen et al., 2018; Bian et al., 2017).

This single model was applied to each trait data (DBH, Ht, BrQ, BrA, Vt and Ct). Diagnostic tools, variogram and plots of spatial residuals were used to detect design, treatment, local and extraneous effects with ASReml 3.0 (Gilmour et al., 2009). The predicted design effects and spatial residuals were extracted from ASReml output files and used to remove all environmental effects from the raw data. The adjusted data were first analysed in a single-site (univariate) analysis to estimate the genetic variance components.

2.4.2. Multi-environment analysis (MET): G × E

The spatially adjusted data were also used in a MET. MET was performed using a factor analytic (FA) model in order to explore the additive G × E in the best possible manner with the final objective of predicting EBV. Due to convergence issues it was not possible to use a full covariance structure with heterogeneous variances (US) or a full correlation structure with heterogeneous variances (CORGH) models.

The following individual tree (or animal) linear mixed model with FA variance structure was thus used, to evaluate simultaneously the performance of grandparents (F0), parents (F1) and offspring (F2), as well as to estimate the parental breeding values that were used in the QTL analysis explained in the next section:

$$y = X\tau + Z_g u_g + e, \tag{1}$$

where y is the vector of observations for s sites (in our study site and trial are the same) and m genotypes, combined across all trials, τ is the vector of fixed effects (intercept and site) with the associated design matrix X; u_g is the random vector of genotype within environment effects ($ms \times 1$) with the associated design matrix Z_g, and e is the combined vector of random residuals from all sites. The random effects are assumed to follow a multivariate normal distribution with means and variances defined by: $u_g \sim N(0, \sigma_g^2 A)$, $e \sim N(0, \sigma_e^2 I)$, where 0 is a null vector; A is the average numerator relationship matrix that describes the additive genetic relationships among individual genotypes, based on the pedigree; I is the identity matrix, with order equal to the number of trees; σ_g^2 is the additive genetic variance; σ_e^2 is the residual variance; u_g is the vector subjected to factor analysis.

FA models are usually based on the number of multiplicative terms (k factors) in the model (Cullis et al., 2014), so we can denote a model with k factors as an FAk model. Each genotype effect in each trial is a sum of k multiplicative terms, and we want the model that can describe most accurately the observed variance-covariance relationships among and within environments (Isik et al., 2017), using as few factors as possible in the model, in our case $k = 1$. We have only 3 sites so we decided to not increase the number of k parameters in our study.

The mixed model in factor analytic form (see Appendix B) was used to model the variance structure of the additive G × E effects:

$$Var(u_g) = G_g \otimes A \tag{2}$$

$$G_g = \Lambda_g \Lambda_g^T + \Psi_g, \tag{3}$$

Here, G_g is the genetic covariance matrix among trials, A is the $m \times m$ numerator/additive genetic relationship matrix, Λ_g is the $s \times k$ matrix of site loadings, and, Ψ_g is the $s \times s$ diagonal matrix containing site-specific variances.

Narrow-sense heritability (h^2) was calculated for each site separately as,

$$h^2 = \sigma_a^2/\sigma_p^2 = \sigma_a^2/\left(\sigma_a^2 + \sigma_\varepsilon^2\right) \quad (4)$$

where σ_a^2 is the single site additive genetic variance, σ_p^2 is the phenotypic variance, σ_ε^2 is the single site residual variance, but also, overall across sites as,

$$h_i^2 = \hat{\sigma}_A^2/\hat{\sigma}_P^2 = \left(\sum_{i=1}^{s}\sigma_a^2 + 2\sum \hat{\sigma}_{gss'}/s^2\right)/\hat{\sigma}_A^2 + \hat{\sigma}_\varepsilon^2, \quad (5)$$

$\hat{\sigma}_A^2$ is the across site additive variance, $\hat{\sigma}_P^2$ is the across site phenotypic variance, $\hat{\sigma}_{gss'}$ is the pairwise covariance between sites, $\hat{\sigma}_\varepsilon^2$ is the average residual variance.

The accuracy of the predicted EBV was calculated for each F_1 full-sib mother as:

$$r = \sqrt{1 - \left(PEV/\sigma_A^2\right)}, \quad (6)$$

where PEV is the prediction error variance derived from the elements of the inverse of the coefficient matrix of the mixed model equations.

Maternal EBV across sites were estimated as the average performance across the set of environments included in the study, $\hat{Y}_m = \bar{\hat{S}} + \bar{\hat{u}}_{msg}$ (Isik et al., 2017).

To study the G × E interactions, type-B genetic correlations were used and obtained as

$$r_B = \sigma_{(s1,s2)}/\left(\sqrt{\sigma_{s1}^2\sigma_{s2}^2}\right), \quad (7)$$

where $\sigma_{(s1,\,s2)}$ is the covariance of additive effects of the trait t between sites s_1 and s_2, σ_{s1}^2 and σ_{s2}^2 are the genetic variances of trait t at site s_1 and site s_2.

Further information on the models can be found in the appendices A and B.

To perform MET, r_B, h^2, h_i^2 and EBV predictions, ASReml 4.0 (Gilmour et al., 2015) was used.

2.4.3. Statistical analysis for QTL mapping

For all the traits, each year of measurement was considered separately as a single trait and, therefore, analysed by the single trait mapping approach (described below). In addition, growth trajectories were fitted by linear regression (Fig. 1) to the 11 years data points (1996–2006) for Ht, and the intercept and slope parameters were taken as latent traits to be used in the subsequent QTL analysis (referred to as the two-stage approach in Li et al. (2014) to analyse longitudinal data).

For each single trait (including each latent trait), the QTL analysis was conducted by solving the LASSO regression (Tibshirani, 1996) problem defined by:

$$\min_{(\beta_0\beta_j)} \frac{1}{2n}\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}x_{ij}\beta_j\right)^2 + \lambda\sum_{j=1}^{p}\left|\beta_j\right|, \quad (8)$$

where y_i is the phenotypic value or EBV of individual i ($i = 1 \ldots n$; n is the total number of individuals), x_{ij} is the genotypic value of individual i and marker j coded as 0 and 1 for two possible marker genotypes, β_0 is the population mean parameter, β_j is the effect of marker j ($j = 1 \ldots p$; p is the total number of markers). LASSO regression shrinks the coefficients of non-important markers toward zero, and only keep the important markers (i.e. those strongly associated with the traits) in the model. The tuning parameter λ ($\lambda > 0$) determines how many markers should be retained in the model, and it also controls the degree of shrinkage of the regression parameters. A detailed guideline about how to choose optimal tuning parameters in the QTL mapping context can be found in Li and Sillanpää (2012). To formally judge QTL, hypothesis testing to identify QTL was conducted by a de-biased LASSO approach (Javanmard and Montanari 2014; Li et al., 2017). Due to its shrinkage nature, the original LASSO estimator does not (asymptotically) follow any parametric distribution, and therefore it is impossible to directly estimate the uncertainty of LASSO estimator such as confidence intervals and P-values. The de-biased LASSO approach constructs an un-biased LASSO estimator, which asymptotically follow a normal distribution. The de-biased LASSO approach aims to calculate the P-values for all markers in the study, instead of only the markers selected by standard LASSO. Li et al. (2017) conducted a simulation study on de-biased LASSO, and the method showed better power to detect QTL and stronger ability to control false positives compared to conventional single locus QTL mapping approaches. We used the following criterion for QTL judgement: a marker was claimed as a significant QTL, if its adjusted P-value was smaller than 0.05.

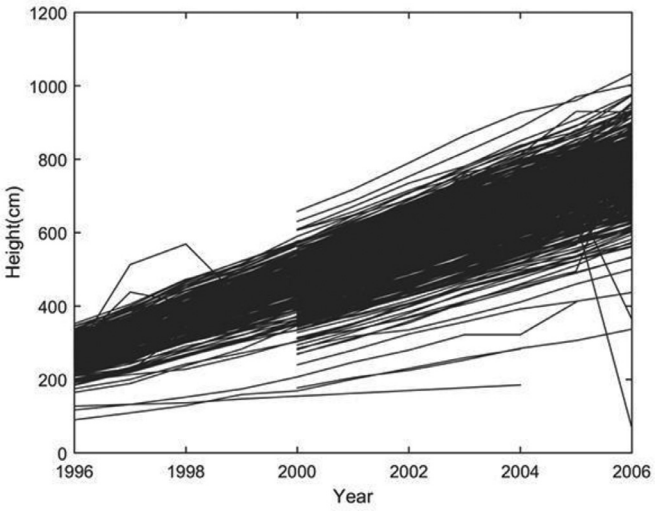

Fig. 1. Growth trajectories of the height (measured in cm) of the 500 trees from year 1996 to 2006.

3. Results

3.1. Accuracy of Estimated Breeding Values (EBV) and heritabilities

Narrow-sense heritabilities, EBV and accuracy of EBV individually for each site and in the performed MET (see Table 2 and Fig. F.2), were computed using adjusted values with the environmental effects having been removed.

Moderate to low heritabilities and EBV accuracies were found for all traits in all sites. Higher accuracies were found for Ht_8 (0.47) and DBH_8 (0.41) using the MET (Table 2b and c). From the single-site analysis point of view, Site2 showed the highest accuracies for Ht_8 (0.44) and DBH_8 (0.37). The EBV accuracy of Ht at Site3 decreased from 0.32 to 0.21 for age three and eight, respectively. The accuracy of survival ability at Site3, Vt3_8 (0.44), was the highest accuracy among all sites and it was almost equal to the accuracy estimated with the MET for Vt_8 (0.43). In the case of Vt_1, the highest accuracy was estimated through the MET (0.32), where single site analysis resulted in slightly smaller estimates (0.18 to 0.28). EBV accuracy of Ct at Site3 age three was moderate (0.44).

Table 2
(a) Average of EBV centered to the mean (a_EBV), (b) heritabilities and their standard errors in between brackets, and (c) accuracy of the EBV (r_EBV), for the traits used both in the MET and single-site analysis.

(a) a_EBV									
Site/Trait	eVt_1	eVt_3	eHt_3	eCt_3	eDBH_8	eBrQ_8	eBrA_8	eHt_8	eVt_8
Site1	2.25				29.56	4.43	4.62	253.98	2.09
Site2	2.79				33.64	5.27	5.05	302.64	2.67
Site3	2.88	2.88	32.56	15.08	21.06	4.19	4.42	219.04	2.18
MET	2.64				28.09	4.63	4.70	258.58	2.31

(b) h^2									
Site/Trait	Vt_1	Vt_3	Ht_3	Ct_3	DBH_8	BrQ_8	BrA_8	Ht_8	Vt8
Site1	0.15 (0.07)				0.17 (0.09)	0.09 (0.09)	0.03 (0.09)	0.18 (0.09)	0.11 (0.07)
Site2	0.10 (0.07)				0.35 (0.08)	0.10 (0.08)	0.04 (0.07)	0.56 (0.09)	0.12 (0.07)
Site3	0.05 (0.06)	0.06 (0.06)	0.21 (0.07)	0.55 (0.09)	0.07 (0.08)	0.19 (0.08)	0.22 (0.08)	0.10 (0.08)	0.42 (0.07)
MET	0.10				0.16	0.06	0.06	0.24	0.11

(c) r_EBV									
Site/Trait	eVt_1	eVt_3	eHt_3	eCt_3	eDBH_8	eBrQ_8	eBrA_8	eHt_8	eVt_8
Site1	0.28				0.25	0.19	0.12	0.27	0.25
Site2	0.23				0.37	0.22	0.14	0.44	0.25
Site3	0.18	0.19	0.32	0.44	0.18	0.29	0.31	0.21	0.44
MET	0.32				0.41	0.29	0.28	0.47	0.43

Heritabilities at Site2 (Table 2b) were only higher than those at the other sites for DBH_8 (0.35) and Ht_8 (0.56). Among all sites, Site3 had the highest heritabilites for BrQ3_8 (0.19), BrA3_8 (0.22), and Vt 3_8 (0.42). The heritability of hardiness measured at age three was 0.55. The heritabilities obtained for the remaining traits and sites were low, ranging between 0.03 and 0.10. Heritabilities estimated with the MET were low for all traits, probably due to the considerable G × E-interactions detected in some cases (Tables 2 and 3) and are described in the following section.

3.2. Multi-environment analysis (MET): G× E

To detect G × E interaction patterns, type-B genetic correlations across sites were used and are shown in Table 3. The genetic correlations for Ht_8 and DBH_8 (0.94 in both cases) between Site1 and Site2 were high, while the Ht_8 and DBH_8 genetic correlations between the other pairs of sites were very low (0.26 to 0.32). For Vt at age one, low genetic correlations (0.24) were detected between Site2 and the remaining sites, while high genetic correlations were detected between Site1 and Site3, however the latest showed a high standard error. In the case of Vt at age eight, a moderate correlation was detected between Site1 and Site3 (0.67), but otherwise low correlations were observed (0.15 to 0.24). A moderate BrQ_8 genetic correlation was detected between Site1 and Site3, and a low correlation was detected for the same trait between Site1 and Site2. Nevertheless in both cases the standard errors were very high, which made us consider these correlations less reliable. In the remaining traits all the genetic correlations were considerably low (below 0.10). BrA_8 genetic correlation between Site1 and Site2 could not be estimated.

3.3. QTL detection

We found 18 AFLPs and 12 SNPs associated across all phenotypic traits, intercepts and EBVs (Tables 2a, 4 and D.1). 62 QTL were significant with a range of percentage of variance explained (%PVE) from 1.70 to 18.90 (Table 4 and Table D.1). The lowest %PVE were detected for CO (1.70), followed by Ht at different ages (%PVE from 1.90 to 3.60), Ht intercept (2.00), FP (2.40), DBH (2.20 - 6.40), BrA (3.40), eBrA (3.40), eVt (4.00), eHt (4.30), Ct (4.30), DBH (5.70), Ht (5.92) and eCt (7.20).

Most of the QTL were detected for Ht phenotypic traits at different ages, followed by Ct, eVt, eCt, eBrQ, eBrA, BrA, W, CO, eHt, DBH, BrD and FP.

Nine QTL-AFLPs (AAG115, TCG89, GAT81, AGC91, ATG226, AGC205, TAG62, GGC205 and AGG498) were detected only in the A data set. The QTL-TCG89 was associated to DBH at ages 16 and 17, Ht intercept and Ht between ages 12 and 18 (with *P*-values that vary from 0.011 to 0.0002).

QTL-GAT81 was associated to Ht between ages 13 and 17, QTL-AGC91 to DBH at age 8, whilst QTL-ARG226 and -AGC205 were linked to BrA at ages 16 and eight respectively. QTL-TAG62 was linked to Ct_4 with a *P*-value of 0.014.

QTL-GGC205 was associated to eHt2_8 (*P*-value of 0.008) and QTL-AGG498 was significant for eVt3_8 and eVt_8, with *P*-values of 0.004 and 0.008, respectively.

Five AFLP (TGG57, TCG134, TTG454, GGC97 and TTG82) were associated with at least one trait when using the S + A dataset. Those QTL were detected for BrA_8, Ct_9, eVt1_1, eVt1_1 and eBrQ2_8, respectively.

Among all the detected QTL, TCG89, GGG312, 0_7009_01-358 and axs_47-502, showed significant pleiotropic effects on DBH/Ht, eBrA1_8/W, FP/W/Ct_16/eCt3_3 and CO/Ct, respectively. AGC106, GGG312 and 0_7009_01-358 were the only three QTL detected in both, the full-sib and in the half-sib progenies. AGG498 was the only QTL-AFLP showing association with eVt based on both, the individual and the MET, EBV estimations.

The QTL detected for the EBV-based traits were only detected at one of the sites or shared between one of the sites and the MET. For the EBV of Vt, we found four QTL. The QTL-AGG498 was the only case where a QTL was detected both in the single site and MET (eVt_8 at Site3 and MET). The remaining three QTL were detected only for eVt_1 at Site1 (QTL-TTG454, QTL-GGC97 and QTL-GI_F1_334). Two QTL were detected for the EBV of BrQ_8 only in Site2 (QTL-TTG82 and QTL-LP2_625).

Multiple QTL were detected for the same trait across several ages. For example, 2_10352_02-413 explains Ht variation at ages 14 to 17 (Ht_14-17). Similarly, 2_9603_01-344 is associated with Ht at ages 9 and 18, AAG115 from age 12 to 17, GAT81 from age 13 to 17 and TCG89 from age 12 to 18. Furthermore, TCG89 also explains the variation of DBH at ages 16 and 17. The SNP-QTL 0_7009_01-358 was associated to Ct based on phenotypic measurements in the F_1-generation trees at the age of 16 (Ct_16) and on EBV scored in the F_2 progenies at the age of 3 (eCt3_3). Similarly, AGC106 explains the observed variation for Ct at ages 9 and 16 based on phenotypic values and for age 3 based on EBV.

Table 3
Across site genetic correlations (and standard errors) for traits used in the MET.

	Vt_1		Vt_8		DBH_8		Ht_8		BrQ_8		BrA_8	
	Site2	Site3	Site2	Site3	Site2	Site3	Site2	Site3	Site2	Site3	Site2	Site3
Site1	0.24 ± 0.36	0.99 ± 0.73	0.24 ± 0.29	0.67 ± 0.24	0.94 ± 0.30	0.28 ± 0.36	0.94 ± 0.25	0.32 ± 0.28	0.15 ± 0.54	0.43 ± 0.45	n.e	0.09 ± 0.33
Site2		0.24 ± 0.34		0.15 ± 0.20		0.26 ± 0.34		0.30 ± 0.26		0.07 ± 0.33		0.09 ± 0.32

n.e.: not estimable.

4. Discussion

To our knowledge, this is the first study performing an EBV-based QTL mapping in conifers. The study was also designed as a multi-environment analysis *in sensu stricto*, therefore, EBV were computed based on pedigree adjusted values with the environmental effects having been removed.

4.1. Genetic architecture (number and %PVE)

Multiple studies on simulated and empirical data have shown that a minimum population size is required to detect QTL consistently (Hall et al., 2016), and the number of QTL detected typically increases with population size and heritability (e.g., Vales et al., 2005; Falke and Frisch, 2011; Wang et al., 2012; Stange et al., 2013). However, in the context of LASSO type of multimarker QTL approaches, increased marker density will do harm for the variable selection because of increased collinearity between markers and therefore some reduction of number of markers is anyway needed (Xu 2013). Moreover, the number of QTL detected and their effects is not only a function of population size or its underlying genetic architecture (i.e., number of genes controlling the trait and their effect size). Other factors, such as marker type (dominant versus co-dominant) can also contribute to the number and effect of the detected QTL. In our study, S + A data set has similar numbers of AFLP and SNP markers, however, SNP-QTL represent almost the 70% (AFLP 30%) of the total number of QTL detected in the S + A data set. This is possibly due to the higher informative (codominant) nature of the SNPs (i.e., the heterozygote Aa can be distinguished from the homozygote AA).

Congruent with the theoretical and empirical expectations, our experimental design mainly allowed the detection of QTL with medium-to-large effects. In both data sets, QTL for eCt3_3 show the highest %PVE (7.20 for the A dataset and 18.90 for the S + A dataset) and exhibits the second highest h^2 (0.55). However, in the S + A data set, we suspect %PVEs to be overestimated due to a lower number of trees in the study (i.e., 91 F_1 trees for the phenotypic-QTL or 61 F_1 trees for the EBV-QTL in the S + A data set compared to 496 F_1 trees for the phenotypic-QTL or 356 F_1 trees for the EBV-QTL in the A data set). Furthermore, Ht in Scots pine has been reported to be under low genetic control (Abrahamsson et al., 2012), while in the S + A data set, the QTL associated with Ht are among the highest %PVE values, again suggesting an overestimation.

Ct was previously known to be under moderate to high genetic control, while traits like Ht, DBH and BrA are typically described to have moderate-to-low h^2 in multiple conifer species (Wu et al., 2008). For example, in *Cunninghamia lanceolata* (Lamb) Hook (Chineese fir), DBH and Ht were described to have low h^2 (0.14 and 0.20, respectively) (Bian et al., 2014). In *Pinus elliottii* (Slash pine), low narrow-sense heritabilities were also reported for Ht (0.03), DBH (0.02) and BrA (0.06 and 0.17 depending on the population) (Pagliarini et al., 2016). Similarly, in *Pinus banksiana* Lamb (Jack pine) h^2 for BrA was also reported to be low (0.16) (Weng et al., 2017). In *Pinus pinaster* (Ait) h^2 for Ct estimate was higher than 0.6 (Prada et al., 2014). In *Pseudotsuga menziesi*, Ht h^2 ranged from 0.03 to 0.09, while Ct was shown to be under low to moderate genetic control (0.16 - 0.37) (Hawkins and Stoehr, 2009). In Scots pine, Ht and Ct were reported to be under low (0.16) and moderate (0.37) genetic control, respectively (Persson et al., 2010; Abrahamsson et al., 2012).

The number of QTL and their %PVE agree with what has previously been reported in the literature (Tables 4, Table D.1 and Table E.1). Several QTL have been detected for hardiness in different species such as Sitka spruce (Holliday et al., 2010), Scots pine (Hurme et al., 2000: Yazdani et al., 2003) and Coastal Douglas-

Table 4
Description of the significant QTL.

Trait	Total nQTL	nQTL A dataset	nQTL S + A dataset	*P*-value	PVE (%) range
BrA_8	2	1	1	0.006	4.90 - 15.00
BrA_16	1	1		0.002	3.40
eBrA1_8	1	1		0.021	3.40
BrD_11	1		1	0.042	11.40
eBrQ2_8	2		2	0.005	16.60 – 17.30
CO	2		2	0.0005	1.70 - 2.80
Ct_4	2	1	1	0.012	5.50 - 12.10
Ct_9	3		3	0.0001	4.30 – 10.00
Ct_11	2		2	0.02	7.50 - 9.10
Ct_16	3	2	1	0.0001	0.0001 - 0.01
eCt3_3	2	1	1	0.0003	7.20 - 18.90
DBH_8	1	1		0.0003	6.40
DBH_16	1	1		0.006	2.20
DBH_17	1	1		0.028	2.20
FP	1		1	0.01	2.40
Ht intercept	2	2		0.007	2.60
Ht_8	1		1	0.0005	10.00
Ht_9	1		1	0.032	11.20
Ht_12	2	2		0.016	2.00
Ht_13	3	3		0.006	2.00 - 2.20
Ht_14	4	3	1	0.004	1.90 – 10.00
Ht_15	4	3	1	0.003	2.10 - 10.10
Ht_16	5	3	2	0.001	2.40 - 11.40
Ht_17	5	3	2	0.001	2.20 - 13.10
Ht_18	2	1	1	0.0002	3.60 - 11.30
eHt2_8	1	1		0.008	4.30
eVt_8	1	1		0.004	4.00
eVt1_1	3		3	0.002	6.50 - 13.90
eVt3_8	1		1	0.008	3.80
W	2		2	0.028	11 - 11.90

nQTL: number of QTL

fir (Jermstad et al. 2001b; Wheeler et al., 2005), with %PVE that varied from 0.70 to 24.90 depending on the species, age and tests performed. %PVE ranged from 1.30 to 34.90 for height at different ages in different species, for instance in white spruce (Pelgas et al., 2011), *Quercus robur* (Gailing et al., 2008; Scotti-Saintagne et al., 2004), *Eucalyptus urophylla* (Gion et al., 2011) and Scots pine (Lerceteau et al., 2000, 2001). In the case of BrD and BrA there are not many QTL studies available, but Lerceteau et al. (2001) found two and one QTL with %PVE estimates of 29.50 and 16.10 respectively.

Interestingly, family-based QTL analysis for the dissection of complex traits have revealed similar trait architecture (e.g., Lerceteau et al., 2001; Markussen et al., 2003; Bartholomé et al., 2013) than population-based QTL studies (even if conducted at the whole-genome level) (e.g., Porth et al., 2013; Uchiyama et al., 2013; Bartholomé et al., 2016). Both types of strategies, typically detect less than 10 QTL and reveal a trait architecture where the majority of the QTL explain a small proportion of the variance (Table E.1). This provides evidence that even AM strategies are more likely to reach significant associations at the single-gene level, population-based mapping has not yet proven advantageous in dissecting trait architecture and estimation of the magnitude of the QTL (Mitchell-Olds, 2010). In other words, due to the extended LD that characterizes single-family progenies, less markers are required to identify the most significant QTL, whereas, historical recombination has resulted in a substantial decrease in LD extension that can only be compensated by increasing substantially the number of markers. Moreover, in both type of strategies, population/progeny size may be a main limiting factor to achieve a fine dissection of the genetic architecture of complex traits (Lynch and Walsh, 1998; Sham et al., 2000), in other words, to find the missing heritability, while other factors such as rare alleles, epistasis and epigenetics may also play an important role (Brachi et al., 2011). In conifers, single-family QTL studies have resulted in the identification of several QTL of small to moderate individual effects mostly related to growth (e.g., Lerceteau et al., 2000, 2001; Markussen et al., 2003; Yazdani et al., 2003; Bartholomé et al., 2013, 2016), wood quality (e.g., Li et al., 2014; Thumma et al., 2010, Freeman et al., 2013) and disease resistance (Hanley et al., 2011). However, less effort has been devoted to the study of traits of adaptive values (Jermstad et al., 2001a,b; Hurme et al., 2000; Yazdani et al., 2003).

4.2. Phenotypic-based QTL *versus* EBV-based QTL (ebvQTL)

Ekine et al. (2014) observed that the use of EBV for QTL detection should be avoided in those experimental designs where the family structure includes information from several relatives and a small population size, with the exception of progeny tests. This is not the case in our study, where our experimental design includes a single family of larger size, which contains exclusively pedigree information from the 11470 half-sib offspring trees.

In the case of cold hardiness we were able to compare the performance of the phenotypic versus EBV-based QTL (*ebv*QTL). In those cases, we observed that when the same marker was associated to a phenotype-based QTL and to an *ebv*QTL, the *ebv*QTL always explained higher proportion of the variance, which could be expected given the more accurate nature of EBV.

4.3. Pleiotropic effects

Pleiotropy is a complex phenomenon that refers to a single gene affecting multiple traits (He and Zhang, 2006). We have detected two SNP-QTL that may have pleiotropic effects. 0_7009_01-358 is significantly associated with FP, Ct and W and axs_47_502 is significantly associated with CO and Ct. Pleiotropic effects have previously been described in the conifer literature. For example,

Pelgas et al. (2011), reported one possible common gene for budset, bud flush and height growth, and another for budset and height. Lerceteau et al. (2001) found one gene in common between DBH and BrD. Hurme et al. (2000) also found one gene with a possible pleiotropic effect between budset and frost hardiness. In the present work, it is however, not possible to discern between pleiotropic effects and the action of two or more loci in the proximity. It has been described in animal breeding literature, that the possibility of pleiotropic effects are hiding other effects such as linkage of separate loci (Johnsson et al., 2014; Wright et al., 2010). However, in our case we did not detect any significant correlation between the phenotypic traits for which we have detected pleiotropic effects.

4.4. QTL age stability and QTLxE

We detected Ht and Ct QTL significantly associated to the same trait across several tree ages. This indicates that some of the QTL detected in this study could be considered as candidates for early selection to certain traits based on their cross-age stability. High age-age correlations have been detected for DBH and Ht, in *Larix kaempferi* (Lai et al., 2014). In the literature, Ct has been shown to exhibit a moderate age-age correlation in some *Rhododendron* populations (Lim, C.C. et al., 2014; Lim, J.H. et al., 2014), whilst in *Phellodendron sachalinense* a high correlation of 0.91 was observed between 34 months old seedlings and 35 year old parent (McNamara and Pellett, 2000).

In our study, we could test the effect of environmental heterogeneity between-sites (i.e., different trials for the same families) and we could also compare the results of individual site versus MET. The lack of reproducibility of the QTL across sites could be a consequence of environmental heterogeneity and G × E interactions observed among our study sites. Jansson (2007) and Persson et al. (2010) reported that the heritabilities of height and survival could vary significantly at different Scots pine progeny tests. Freeman et al. (2013) and Bartholomé et al. (2013) also found clear evidences of QTL × E interaction for growth traits in *Eucalyptus globulus* and *Eucalyptus* hybrids respectively. Rae et al. (2008) have shown that the environment influences the detection of QTL in three different locations in Europe, for bioenergy traits in *Populus* hybrids. In conifers, evidence of QTL × E interactions has been found for wood specific gravity in *Pinus taeda* L. (Groover et al., 1994).

4.5. Significant SNP molecular functions

Several significant SNP-QTL may encode for known proteins. The best BLAST hit for 0_7009_01-358 is a kelch-motif containing protein. The kelch motif is an evolutionarily-widespread sequence motif of 44–56 amino acids in length involved in protein-protein interactions and the proteins that contain this motif are involved in the regulation of the circadian clock, in brassinosteroid modulation or in phenylpropanoid biosynthesis, among other functions (Zhang et al., 2013). The molecular and functional nature of such motif seems congruent with the pleiotropic nature of 0_7009_01-358. Another SNP-QTL involved in the control of CO and Ct is axs_47_502 that encodes for an UDP-apiose/xylose synthase. This enzyme catalyzes the NAD + -dependent conversion of UDPD-glucuronic acid to UDP-D-apiose and UDP-Dxylose (Grisebach and Schmid, 1972) and it is involved in the plant cell wall formation. UDP-Dxylose involvement in the pleiotropic control of CO and Ct may be the result of multiple genes in linkage disequilibrium within this single QTL. For example, fine QTL-mapping in rice identified multiple genes associated to the same QTL region (Lim, C.C. et al., 2014; Lim, J.H. et al., 2014). The 2_10212_01-241 encodes for a glutathione transferase located at the chloroplast. The involvement of this gene in stress regulation (Sappl et al., 2009) seems compatible with this SNP-QTL' association with Ct. The a3ip2_387 encodes for the ABI3-interacting protein 2 (ABI3), a transcription factor of the abscisic acid signal transduction pathway (Koornneef et al., 1984). In conifers, dormancy and frost tolerance are interconnected processes that share common molecular mechanisms with *Arabidopsis* (Welling and Palva, 2006). Considering the role of ABI3 in frost tolerance (Tamminen et al., 2001) and dormancy (Nambara et al., 1995) in *Arabidopsis*, it could be postulated that a possible involvement of ABI3 in frost tolerance also occurs in conifers, a hypothesis which is reinforced by the known association of this gene in seed development and dormancy in conifers (Zeng et al., 2013). The 2_10352_02-413 QTL' best BLAST is a nucleotide sugar epimerase, which is involved in photosynthesis membrane biogenesis and its overexpression leads to growth acceleration in *Arabidopsis* (Li et al., 2011), thus supporting its role in conifer growth. The best BLAST hit result for QTL 2_6731_01-230 was a putative F-box protein GID2. F-box protein, gibberellin-insensitive dwarf2 (GID2), mediates the action of phytohormone gibberellin (GA) in the control of growth and development in plants (Gomi et al., 2004). This gives credibility to the SNP-QTL association with growth (Ht) found in this study. We also found a significant association between GI_f1_334 that encodes for gigantean (GI) with Vt. GI is a very well characterized gene involved in multiple physiological processes including stress responses to cold and drought in *Arabidopsis* (reviewed by Mishra and Panigrahi, 2015). This could be interpreted as a validation of GI involvement in vitality mediating processes in conifers. 0_11919_01-122, 2_9603_01-344, 0_17247_02-266 and CL2495Contig1_03-101 were not annotated at the time of this study.

5. Conclusions

Our study was congruent with the a typical genetic architecture where the majority of the significant QTL have only a small contribution to the variance. In agreement with theoretical expectations, *ebv*QTL showed a higher percentage of explained variance for cold hardiness, which may indicate that EBV are more accurate than phenotypic in the cases of progeny test, where the pedigree contains only information from the direct offspring, and no information from other relatives is included in the pedigree. We detected QTLs for Ht and Ct (including SNPs) that were stable across several tree ages, thus indicating that some of the QTL detected in this study could be considered as strong candidates for early selection for certain traits based on their cross-age stability. We detected environmental heterogeneity between sites and genotype-by-environment interactions, which could be the main reason behind the absence of reproducibility of some QTL across sites.

We acknowledge the limitations of utilizing dominant, anonymous markers such as AFLPs for the purpose of identification of molecular markers for their application in Assisted Selection. Moreover, it is important to highlight that in non-model species and laboratories with limited resources, AFLP and SSRs are still expected to contribute to the study of traits complexity and the identification of QTL. This could still be justified considering that AFLP loci can be cloned and turned into single nucleotide polymorphisms (SNPs) (Brugmans et al., 2003).

Declaration of interests

none.

Acknowledgements

We are thankful to Ross Whetten for guidance in the development and genotyping of AFLP markers, to Outi Savolainen and

Sonja Kujala for the information on SNP annotation, and to Skogforsk that planted and maintained the field material, and also for the help with filed measurements. We are also grateful to Anders Fries and René Fernández for the field sampling, Dr. John Baison for proof reading of the manuscript, and, Dr. Fikret Isik, Dr. Zhi-Qiang Chen and Dr. Arthur Gilmour for their assistance in the statistical analysis. This study was supported by the Research School in Forest Genetics and Breeding at the Swedish University of Agricultural Sciences, Knut och Alice Wallenbergs foundation, EVOLTREE EU project and PROCOGEN EU project.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jtbi.2018.11.007.

References

Abrahamsson, S., Nilsson, J.E., Wu, H.X., García-Gil, M.R., Andersson, B., 2012. Inheritance of height growth and autumn cold hardiness based on two generations of full-sib and half-sib families of Pinus sylvestris. Scand. J. For. Res. 27, 405–413. doi:10.1080/02827581.2012.663403.

Bartholomé, J., Bink, M.C., van Heerwaarden, J., Chancerel, E., Boury, C., Lesur, I., et al., 2016. Linkage and association mapping for two major traits used in the maritime pine breeding program: height growth and stem straightness. PLoS ONE 11, e0165323.

Bartholomé, J., Salmon, F., Vigneron, P., Bouvet, J.M., Plomion, C., Gion, J.M., 2013. Plasticity of primary and secondary growth dynamics in *Eucalyptus* hybrids: a quantitative genetics and QTL mapping perspective. BMC Plant Biol. 13, 120. doi:10.1186/1471-2229-13-120.

Bian, L., Shi, J., Zheng, R., Chen, J., Wu, H.X., 2014. Genetic parameters and genotype–environment interactions of Chinese fir (Cunninghamia lanceolata) in Fujian Province. Can. J. For. Res. 44, 582–592. doi:10.1139/cjfr-2013-0427.

Bian, L., Zheng, R., Xiao, H., Su, S., Gan, Z., Wu, H.X., Shi, J., 2017. Spatial analysis increases efficiency of rogeny testing of Chinese fir. J. For. Res. 28, 445–452. doi:10.1007/s11676-016-0341-z.

Brachi, B., Geoffrey, P.M., Justin, O.B., 2011. Genome-wide association studies in plants: the missing heritability is in the field. Genome Biol. 12, 232. doi:10.1186/gb-2011-12-10-232.

Brugmans, B., van der Hulst, R.G., Visser, R.G., Lindhout, P., van Eck, H.J., 2003. A new and versatile method for the successful conversion of AFLP™ markers into simple single locus markers. Nucleic Acids Res 31 e55-e55.

Chen, Z.Q., Karlsson, B., Wu, H.X., 2017. Patterns of additive genotype-by-environment interaction in tree height of Norway spruce in southern and central Sweden. Tree Genet. Genomes. 13, 25. doi:10.1007/s11295-017-1103-6.

Chen, Z.Q., Helmersson, A., Karlsson, B., Westin, J., Wu, H.X., 2018. Efficiency of using spatial analysis for Norway spruce field tests in Sweden. Ann. For. Sci. 75, 2. doi:10.1007/s13595-017-0680-8.

Cullis, B.R., Jefferson, P., Thompson, R., Smith, A.B., 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.

Doyle, J., 1991. DNA protocols for plants. In: Molecular Techniques in Taxonomy. Springer, Berlin Heidelberg, pp. 283–293.

Dutkowski, G.W., Silva, J.C.E., Gilmour, A.R., Lopez, G.A., 2002. Spatial analysis methods for forest genetic trials. Can. J. For. Res. Res. 32, 2201–2214 doi:10.1139/x02-111.

Dutkowski, G.W., Costa e Silva, J., Gilmour, A.R., Wellendorf, H., Aguiar, A., 2006. Spatial analysis enhances modelling of a wide variety of traits in forest genetic trials. Can. J. For. Res. Res. 36, 1851–1870 doi:10.1139/x06-059.

Ekine, C.C., Rowe, S.J., Bishop, S.C., de Koning, D.J., 2014. Why breeding values estimated using familial data should not be used for genome-wide association studies. G3 (Bethesda) 4, 341–347. doi:10.1534/g3.113.008706.

Falke, K.C., Frisch, M., 2011. Power and false-positive rate in QTL detection with near-isogenic line libraries. Heredity 106, 576–584. doi:10.1038/hdy.2010.87.

Freeman, J.S., Potts, B.M., Downes, G.M., Pilbeam, D., Thavamanikumar, S., Vaillancourt, R.E., 2013. Stability of quantitative trait loci for growth and wood properties across multiple pedigrees and environments in Eucalyptus globulus. New Phytol. 198, 1121–1134. doi:10.1111/nph.12237.

Gailing, O., Langenfeld-Hayser, R., Polle, A., Finkeldey, R., 2008. Quantitative trait loci affecting stomatal density and growth in a Quercus robur progeny: implications for the adaptation to changing environments. Glob. Change Biol. 14, 1934–1946. doi:10.1111/j.1365-2486.2008.01621.x.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Thompson, R., 2009. ASReml User Guide Release 3.0. VSN International Ltd, Hemmel Hempstead, UK.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J., Thompson, R., 2015. ASReml User Guide Release 4.1 Structural Specification. Hemel Hempstead: VSN International Ltd, Hemmel Hempstead, UK.

Gion, J.M., Carouché, A., Deweer, S., Bedon, F., Pichavant, F., Charpentier, J.P., et al., 2011. Comprehensive genetic dissection of wood properties in a widely-grown tropical tree: eucalyptus. BMC Genom. 12, 301. doi:10.1186/1471-2164-12-301.

Gomi, K., Sasaki, A., Itoh, H., Ueguchi-Tanaka, M., Ashikari, M., Kitano, H., et al., 2004. GID2, an F-box subunit of the SCF E3 complex, specifically interacts with phosphorylated SLR1 protein and regulates the gibberellin-dependent degradation of SLR1 in rice. Plant J. 37, 626–634.

Grisebach, H., Schmid, R., 1972. Chemistry and biochemistry of branched-chain sugars. Angew. Chem. 11, 159–173. doi:10.1002/anie.197201591.

Groover, A., Devey, M., Fiddler, T., Lee, J., Megraw, R., Mitchel-Olds, T., et al., 1994. Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. Genetics 138, 1293–1300.

Hall, D., Hallingbäck, H.R., Wu, H.X., 2016. Estimation of number and size of QTL effects in forest tree traits. Tree Genet. Genomes. 12, 110. doi:10.1007/s1129.

Hanley, S.J., Pei, M.H., Powers, S.J., Ruiz, C., Mallott, M.D., Barker, J.H., et al., 2011. Genetic mapping of rust resistance loci in biomass willow. Tree Genet. Genomes. 7, 597–608. doi:10.1007/s11295-010-0359-x.

Hawkins, B.J., Stoehr, M., 2009. Growth, phenology, and cold hardiness of 32 Douglas-fir full-sib families. Can. J. For. Res. 39, 1821–1834. doi:10.1139/X09-092.

He, X., Zhang, J., 2006. Toward a molecular understanding of pleiotropy. Genetics 173, 1885–1891. doi:10.1534/genetics.106.060269.

Holliday, J.A., Ritland, K., Aitken, S.N., 2010. Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (Picea sitchensis). New Phytol. 188, 501–514. doi:10.1111/j.1469-8137.2010.03380.x.

Hurme, P., Sillanpää, M.J., Arjas, E., Repo, T., Savolainen, O., 2000. Genetic basis of climatic adaptation in Scots pine by Bayesian quantitative trait locus analysis. Genetics 156, 1309–1322.

Isik, F., 2014. Genomic selection in forest tree breeding: the concept and an outlook to the future. New For 45, 379–401. doi:10.1007/s11056-014-9422-z.

Isik, F., Holland, J., Maltecca, C., 2017. Genetic Data Analysis for Plant and Animal Breeding, 1. Springer International Publishing, New York.

Ivković, M., Gapare, W., Yang, H., Dutkowski, G., Buxton, P., Wu, H.X., 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.

Jansson, G., 2007. Gains from selecting Pinus sylvestris in southern Sweden for volume per hectare. Scand. J. For. Res. 22, 185–192. doi:10.1080/02827580701330894.

Javanmard, A., Montanari, A., 2014. Confidence intervals and hypothesis testing for high-dimensional regression. J. Mach. Learn. Res. 15, 2869–2909.

Jermstad, K.D., Bassoni, D.L., Jech, K.S., Wheeler, N.C., Neale, D.B., 2001a. Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir. I. Timing of vegetative bud flush. Theor. Appl. Genet. 102, 1142–1151.

Jermstad, K.D., Bassoni, D.L., Wheeler, N.C., Anekonda, T.S., Aitken, S.N., Adams, W.T., et al., 2001b. Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir. II. Spring and fall cold-hardiness. Theor. Appl. Genet. 102, 1152–1158.

Johnsson, M., Rubin, C.J., Höglund, A., Sahlqvist, A.S., Jonsson, K.B., Kerje, S., et al., 2014. The role of pleiotropy and linkage in genes affecting a sexual ornament and bone allocation in the chicken. Mol. Ecol. 23, 2275–2286. doi:10.1111/mec.12723.

Kelly, A.M., Cullis, B.R., Gilmour, A.R., Eccleston, J.A., Thompson, R., 2009. Estimation in a multiplicative mixed model involving a genetic relationship matrix. Genet. Sel. Evol. 41, 33. doi:10.1186/1297-9686-41-33.

Koornneef, M., Reuling, G., Karssen, C.M., 1984. The isolation and characterization of abscisic acid-insensitive mutants of Arabidopsis thaliana. Physiol. Plant 61, 377–383. doi:10.1111/j.1399-3054.1984.tb06343.x.

Lai, M., Sun, X., Chen, D., Xie, Y., Zhang, S., 2014. Age-related trends in genetic parameters for Larix kaempferi and their implications for early selection. BMC Genet. 15, S10. doi:10.1186/1471-2156-15-S1-S10.

Lander, E.S., Botstein, D., 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121, 185–199.

Lerceteau, E., Plomion, C., Andersson, B., 2000. AFLP mapping and detection of quantitative trait loci (QTL) for economically important traits in Pinus sylvestris: a preliminary study. Mol. Breed. 6, 451–458. doi:10.1023/A:1026548716320.

Lerceteau, E., Szmidt, A.E., Andersson, B., 2001. Detection of quantitative trait loci in Pinus sylvestris L. across years. Euphytica 121, 117–122. doi:10.1023/A:1012076825293.

Li, C., Wang, Y., Liu, L., Hu, Y., Zhang, F., Mergen, S., et al., 2011. A rice plastidial nucleotide sugar epimerase is involved in galactolipid biosynthesis and improves photosynthetic efficiency. PLoS Genet. 7, e1002196. doi:10.1371/journal.pgen.1002196.

Li, Z., Sillanpää, M., 2012. Overview of LASSO-related penalized regression methods for quantitative trait mapping and genomic selection. Theor. Appl. Genet. 125, 419–435. doi:10.1007/s00122-012-1892-9.

Li, Z., Hallingbäck, H.R., Abrahamsson, S., Fries, A., Gull, B.A., Sillanpää, M.J., et al., 2014. Functional multi-locus QTL mapping of temporal trends in scots pine wood traits. G3 (Bethesda) 4, 2365–2379. doi:10.1534/g3.114.014068.

Li, Z., Guo, B., Yang, J., Herczeg, G., Gonda, A., Balázs, G., et al., 2017. Deciphering the genomic architecture of the stickleback brain with a novel multilocus gene-mapping approach. Mol. Ecol. 26, 1557–1575 10.1111/mec.14005.

Lim, C.C., Krebs, S.L., Arora, R., 2014. Cold hardiness increases with age in juvenile Rhododendron populations. Front. Plant Sci. 5, 542. doi:10.3389/fpls.2014.00542.

Lim, J.H., Yang, H.J., Jung, K.H., Yoo, S.C., Paek, N.C., 2014. Quantitative trait locus mapping and candidate gene analysis for plant architecture traits using whole genome re-sequencing in rice. Mol. Cells. 37, 149. doi:10.14348/molcells.2014.2336.

Lynch, M., Walsh, B., 1998. Genetics and Analysis of Quantitative Traits. Sinauer, Sunderland, MA Volumen 1.
Markussen, T., Fladung, M., Achere, V., Favre, J.M., Faivre-Rampant, P., Aragones, A., et al., 2003. Identification of QTL controlling growth, chemical and physical wood property traits in Pinus pinaster (Ait.). Silvae Genet 52, 8–14.
Mauricio, R., 2001. Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. Nat. Rev. Genet. 2, 370–381. doi:10.1038/35072085.
McNamara, S., Pellett, H., 2000. Cold hardiness of Phellodendron sachalinense Friedr. Schmidt seedlings increases with age. HortScience 35, 304–305.
Mishra, P., Panigrahi, K.C., 2015. GIGANTEA–an emerging story. Front. Plant Sci. 6, 8. doi:10.3389/fpls.2015.00008.
Mitchell-Olds, T., 2010. Complex-trait analysis in plats. Genome Biol. 11, 113. doi:10.1186/gb-2010-11-4-113.
Nambara, E., McCourt, P., Naito, S., 1995. A regulatory role for the ABI3 gene in the establishment of embryo maturation in Arabidopsis thaliana. Development 121, 629–636.
Neale, D.B., Savolainen, O., 2004. Association genetics of complex traits in conifers. Trends Plant Sci. 9, 325–330. doi:10.1016/j.tplants.2004.05.006.
Nilsson, J.E., Walfridsson, E.A., 1995. Phenological variation among plus-three clones of Pinus sylvestris (L.) in Northern Sweden. Silvae Genet. 44, 20–28.
Ogut, F., Maltecca, C., Whetten, R., McKeand, S., Isik, F., 2014. Genetic analysis of diallel progeny test data using factor analytic linear mixed models. Forest Sci. 60, 119–127. doi:10.5849/forsci.12-108.
Pagliarini, M.K., Kieras, W.S., Moreira, J.P., Sousa, V.A., Shimizu, J.Y., Moraes, M.L.T., et al., 2016. Adaptability, stability, productivity and genetic parameters in slash pine second-generation families in early age. Silvae Genet. 65, 71–82. doi:10.1515/sg-2016-0010.
Pelgas, B., Bousquet, J., Meirmans, P.G., Ritland, K., Isabel, N., 2011. QTL mapping in white spruce: gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments. BMC Genom. 12, 145. doi:10.1186/1471-2164-12-145.
Persson, T., Andersson, B., 2003. Genetic variance and covariance patterns of growth and survival in northern Pinus sylvestris. Scand. J. For. Res. 18, 332–343. doi:10.1080/02827580310003993.
Persson, T., Andersson, B., Ericsson, T., 2010. Relationship between autumn cold hardiness and field performance in northern Pinus sylvestris. Silva Fenn. 44, 255–266.
Piepho, H.P., Möhring, J., Melchinger, A.E., Büchse, A., 2008. BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161, 209–228. doi:10.1007/s10681-007-9449-8.
Porth, I., Klapšte, J., Skyba, O., Hannemann, J., McKown, A.D., Guy, R.D., et al., 2013. Genome-wide association mapping for wood characteristics in Populus identifies an array of candidate single nucleotide polymorphisms. New Phytol. 200, 710–726. doi:10.1111/nph.12422.
Prada, E., Alia, R., Climent, J., Diaz, R., 2014. , Seasonal cold hardiness in maritime pine assessed by different methods. Tree Genet. Genomes. 10, 689–701. doi:10.1007/s11295-014-0714-4.
Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
Rae, A.M., Pinel, M.P., Bastien, C., Sabatti, M., Street, N.R., Tucker, J., et al., 2008. QTL for yield in bioenergy Populus: identifying G× E interactions from growth at three contrasting sites. Tree Genet. Genomes. 4, 97–112. doi:10.1007/s11295-007-0091-3.
Sappl, P.G., Carroll, A.J., Clifton, R., Lister, R., Whelan, J., Harvey Millar, A., et al., 2009. The Arabidopsis glutathione transferase gene family displays complex stress regulation and co-silencing multiple genes results in altered metabolic sensitivity to oxidative stress. Plant J. 58, 53–68. doi:10.1111/j.1365-313X.2008.03761.x.
Scotti-Saintagne, C., Bodénès, C., Barreneche, T., Bertocchi, E., Plomion, C., Kremer, A., 2004. Detection of quantitative trait loci controlling bud burst and height growth in Quercus robur L. Theor. Appl. Genet. 109, 1648–1659. doi:10.1007/s00122-004-1789-3.
Sham, P.C., Zhao, J.H., Curtis, D., 2000. The effect of marker characteristics on the power to detect linkage disequilibrium due to single or multiple ancestral mutations. Ann. Hum. Genet. 64 (2), 161–169. doi:10.1017/S0003480000008046.
Smith, A.B., Ganesalingam, A., Kuchel, H., Cullis, B.R., 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.
Stange, M., Utz, H.F., Schrag, T.A., Melchinger, A.E., Würschum, T., 2013. High density genotyping: an overkill for QTL mapping? Lessons learned from a case study in maize and simulations. Theor. Appl. Genet. 126, 2563–2574. doi:10.1007/s00122-013-2155-0.
Tamminen, I., Makela, P., Heino, P., Palva, E.T., 2001. Ectopic expression of ABI3 gene enhances freezing tolerance in response to abscisic acid and low temperature in Arabidopsis thaliana. Plant J. 25, 1–8.
Thavamanikumar, S., Southerton, S.G., Bossinger, G., Thumma, B.R., 2013. Dissection of complex traits in forest trees—opportunities for marker-assisted selection. Tree Genet. Genomes. 9, 627–639. doi:10.1007/s11295-013-0594-z.
Thumma, B.R., Southerton, S.G., Bell, J.C., Owen, J.V., Henery, M.L., Moran, G.F., 2010. Quantitative trait locus (QTL) analysis of wood quality traits in Eucalyptus nitens. Tree Genet. Genomes 6, 305–317. doi:10.1007/s11295-009-0250-9.
Uchiyama, K., Iwata, H., Moriguchi, Y., Ujino-Ihara, T., Ueno, S., Taguchi, Y., et al., 2013. Demostration of genome-wide association studies for identifying markers for wood property and male strobili in Cryptomeria japonica. PLoS ONE 8, e79866. doi:10.1371/journal.pone.0079866.
Vales, M.I., Schön, C.C., Capettini, F., Chen, X.M., Corey, A.E., Mather, D.E., et al., 2005. Effect of population size on the estimation of QTL: a test using resistance to barley stripe rust. Theor. Appl. Genet. 111, 1260–1270. doi:10.1007/s00122-005-0043-y.
Wang, H., Smith, K.P., Combs, E., Blake, T., Horsley, R.D., Muehlbauer, G.J., 2012. Effect of population size and unbalanced data sets on QTL detection using genome-wide association mapping in barley breeding germplasm. Theor. Appl. Genet. 124, 111–124. doi:10.1007/s00122-011-1691-8.
Welling, A., Palva, E.T., 2006. Molecular control of cold acclimation in trees. Physiol. Plant. 127, 167–181. doi:10.1111/j.1399-3054.2006.00672.x.
Weng, Y., Ford, R., Tong, Z., Krasowski, M., 2017. Genetic parameters for bole straightness and branch angle in jack pine estimated using linear and generalized linear mixed models. Forest Sci 63, 111–117 doi:10.5849/forsci.16-039.
Wheeler, N.C., Jermstad, K.D., Krutovsky, K., Aitken, S.N., Howe, G.T., Krakowski, J., et al., 2005. Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir. IV. Cold-hardiness QTL verification and candidate gene mapping. Mol. Breed. 15, 145–156. doi:10.1007/s11032-004-3978-9.
Wright, D., Rubin, C.J., Martinez Barrio, A., Schütz, K., Kerje, S., Brändström, H., et al., 2010. The genetic architecture of domestication in the chicken: effects of pleiotropy and linkage. Mol. Ecol. 19, 5140–5156. doi:10.1111/j.1365-294X.2010.04882.x.
Wu, H.X., Ivković, M., Gapare, W.J., Matheson, A.C., Baltunis, B.S., Powell, M.B., McRae, T.A., 2008. Breeding for wood quality and profit in radiata pine: a review of genetic parameters and implication for breeding and deployment. N. Z. J. For. 38, 56–87.
Xu, S., 2013. Genetic mapping and genomic selection using recombination breakpoint data. Genetics 195, 1103–1115. doi:10.1534/genetics.113.155309.
Yazdani, R., Nilsson, J.E., Plomion, C., Mathur, G., 2003. Marker trait association for autumn cold acclimation and growth rhythm in Pinus sylvestris. Scand. J. For. Res 18, 29–38. doi:10.1080/02827581.2003.10383135.
Zeng, Y., Zhao, T., Kermode, A.R., 2013. A conifer ABI3-interacting protein plays important roles during key transitions of the plant life cycle. Plant Physiol. 161, 179–195. doi:10.1104/pp.112.206946.
Zhang, X., Gou, M., Liu, C.J., 2013. Arabidopsis Kelch repeat F-box proteins regulate phenylpropanoid biosynthesis via controlling the turnover of phenylalanine ammonia-lyase. Plant Cell 25, 4994–5010. doi:10.1105/tpc.113.119644.