1 Functional Quantitative Trait Loci (QTL) analysis for adaptive traits in a
2 three-generation Scots pine pedigree

3

4 Calleja-Rodriguez A^{1,2}, Li Z³, Hallingbäck H R^{1,4}, Sillanpää M J⁵, Wu H X¹,
 5 Abrahamsson S², García-Gil MR¹

6 ¹Department of Forest Genetics and Plant Physiology, Swedish University of
 7 Agricultural Science, SE-901 83 Umeå, Sweden

8 ² Skogforsk, Box 3, SE-91821 Sävar, Sweden

9 ³ Ecological Genetics Research Unit, Department of Biological Sciences,
 10 University of Helsinki, FI-00014 Helsinki, Finland.

11 ⁴Department of Plant Biology, Uppsala BioCenter, Linnean Centre for Plant
 12 Biology, Swedish University of Agricultural Science, SE-75007 Uppsala,
 13 Sweden

14 ⁵Department of Mathematical Sciences and Biocenter Oulu, University of Oulu,
 15 FI-90014, Oulu, Finland.

16 Corresponding author: M.Rosario Garcia Gil. E-mail: m.rosario.garcia@slu.se,
 17 0730962240

18 Running title: Functional QTL for adaptive traits

19 Word count for main text: 6 425

20 **Abstract**

21 In forest tree breeding, QTL identification aims to accelerate the breeding cycle
 22 and increase the genetic gain of traits with economical and ecological value. In
 23 our study, both phenotypic data and predicted breeding values were used in the
 24 identification QTL linked to the adaptive value in a three-generation pedigree
 25 population, for the first time in a conifer species (*Pinus sylvestris* L.). A total of
 26 11 470 open pollinated F₂-progeny trees established at three different locations,
 27 were measured for growth and adaptive traits. Breeding values were predicted
 28 for their 360 mothers, originating from a single cross of two parents. A multilevel
 29 LASSO association analysis was conducted to detect QTL using genotypes of
 30 the mothers with the corresponding phenotypes and estimated breeding values
 31 (EBVs). Different levels of genotype-by-environment (G×E) effects among sites
 32 and ages were detected for survival and height. Moderate-to-low narrow sense
 33 heritabilities and EBVs accuracies were found for all traits and all sites. We
 34 identified 18 AFLPs and 12 SNPs to be associated with QTL for one or more
 35 traits. 62 QTL were significant with percentages of variance explained ranging
 36 from 1.7 to 18.9%, mostly for traits based on phenotypic data. Two SNP-QTL
 37 showed pleiotropic effects for traits related with survival, seed and flower
 38 production. Furthermore, we detected several QTL with significant effects across
 39 multiple ages, which could be considered as strong candidate loci for early

40 selection. The lack of reproducibility of some QTL detected across sites may be
41 due to environmental heterogeneity and QTL-by-environment effects.

42 Keywords: Scots pine, climate change, quantitative trait locus, LASSO, adaptive
43 traits

44

45 **Introduction**

46 Trees are of great economic importance as there are a source of wood and fibers
 47 in industrial processes and of immense ecologic value as carbon sinks and
 48 habitats. In the last several decades, scientists and breeders have joined efforts to
 49 improve tree health and productivity in the context of climate change, which is
 50 increasingly impacting on the adaptation of local populations.

51 As a result of an increase in the average temperatures in Nordic countries,
 52 longer growth seasons have been predicted which may influence adaptive traits
 53 such as timing of growth, flowering and cold acclimation of forest trees
 54 (Linderholm, 2006; Walther and Linderholm, 2006; Khorsand Rosa et al. 2015).
 55 Given the predictions of climatic change and the effects on forest fitness, tree
 56 breeders are developing strategies to breed for trees adapted to the future climatic
 57 conditions. One breeding strategy is to translocate trees (gene donors) within the
 58 species distribution range in order to introgress new gene variants (alleles) into
 59 the existing population with the aim of accelerating population adaptation
 60 (assisted gene flow) (Aitken and Whitlock, 2013). Alternatively, the populations'
 61 standing genetic diversity could be sufficient to shift the population fitness mean
 62 towards a new optimal by artificial selection of local progenitors bearing alleles
 63 favourable to the new climatic conditions (i.e., within-population assisted
 64 selection) (Matuszewski et al. 2015). Either way, the breeding protocols involve

65 assessment of tree performance across multiple locations as a method to evaluate
66 and minimize genotype by environment interactions (GXE) and, thus, improve
67 tree adaptation to an entire breeding zone (Ivkovic et al. 2015).

68 There are two major uses of mapping Quantitative Trait Loci (QTL). The
69 first is to accelerate the breeding cycle by artificial selection of desired genotypes
70 at an earlier stage using molecular marker data: Marker Assisted Selection
71 (MAS) originally introduced by Sax (1923). The breeding cycle could be
72 shortened by the indirect selection of a causal gene (i.e., gene controlling a trait)
73 based on the selection of molecular marker(s) in linkage disequilibrium with this
74 gene (Collard and Mackill, 2008). The second use of genetic architecture is the
75 development of optimal breeding methods using species genomic models
76 (Hallingbäck et al. 2014; Wu et al. 2016).

77 Traditionally, QTL studies in tree species have been based on phenotypic
78 rather than estimated genotypic or breeding values (EBVs). The main limitation
79 of the traditional phenotypic-based QTL detection approach is the design of the
80 experiments that typically involve only one generation, often consisting of a
81 single full-sib family where environmental and genetic factors are confounded
82 (Thavamanikumar et al. 2013, Isik 2014, Hall et al. 2016). Alternatively,
83 phenotypic values can be substituted by EBVs. This method requires

84 phenotyping progenies of the target (mother) trees and posteriorly computing
85 EBVs with the purpose of ranking the target trees for the traits under study.
86 Estimation of genetic effects, such as EBVs (i.e additive genetic effects), can
87 improve the estimation of genotypic performance and also improve the genotypic
88 value estimation of candidate genotypes in breeding programs (Piepho et al.
89 2008).

90 Since the milestone article by Lander and Botstein (1989) there have been
91 a large number of QTL mapping studies in multiple plant species (Mauricio,
92 2001). In forest species, QTL analyses have either been based on progenies (full-
93 sib or half-sib families) or unrelated individuals (association mapping). Those
94 studies resulted in the identification of several QTL of small to moderate
95 individual effects mostly related to growth (e.g., Lerceteau et al. 2000, 2001;
96 Markussen et al. 2003; Yazdani et al. 2003; Bartholome et al. 2013, 2016), wood
97 quality (e.g., Li et al. 2014; Thumma et al. 2010, Freeman et al. 2013) and disease
98 resistance (Hanley et al. 2011). However, less effort has been devoted to the study
99 of traits of adaptive values (Jermstad et al. 2001a,b; Hurme et al. 2000; Yazdani
100 et al. 2003).

101 In this study, the goal was to identify QTL of economical and ecological
102 value such as growth, tree quality, frost hardiness and survival measured across
103 three different trials and multiple years in a three generation pedigree in Scots

104 pine (*Pinus sylvestris* L.). The marker information has already been published by
105 Li et al. (2014) but in our current study we used new growth and adaptive
106 phenotypic traits data and EBVs to perform a novel QTL study.

107

108 **Material and Method**

109 **Plant material**

110 The parents (AC3065 and Y3088) of the full-sib cross are plus-trees from north
111 Sweden (latitudes 65°08'N and 64°09'N, respectively), and belong to the
112 Swedish breeding program (see supplementary figure S4). Progeny generated
113 from this cross (F₁-generation) were planted in 1988 as one year old seedlings
114 (Abrahamsson et al. 2012). 455 fullsib individuals from this cross were included
115 in the study (trial F485: Flurkmark, see Table 1 and supplementary figure S4). In
116 2006, open pollinated (OP) seeds (F₂-generation) from 360 fullsib trees in the
117 F485 trial were collected and grown at the Skogforsk nursery in Sävar. These
118 fertile 360 trees producing female cones in 2006 were the only ones included in
119 the collection as the others were non-fertile at this time. None of the 455 trees
120 produced any male cones. The resulting OP seedlings were planted in four
121 different field trials in 2008, using a complete randomized single-tree plot design.
122 Three out of the four trials were chosen for this study, with an average number
123 of trees per OP family of 9, 10 and 11 for sites F725, F723 and F726 (Table 1

124 and supplementary material S4), respectively (the fourth trail was unusable due
125 to the construction of a new road over the trial). These trials could, from the
126 technical point of view, be treated as half-sib trials. Although it should be
127 cautioned that all F_2 -progeny were interrelated to a limited extent (single-cousins
128 at the least).

129

130 **Phenotypic measurements**

131 Measured traits and the age of the trees at time of measurement for each trial are
132 described in Table 2.

133 *Trait measurements on the F_1 fullsib cross (F485)*

134 Height (Ht) was measured annually between 1996 and 2006 using a telescopic
135 pole from ground to terminal bud. Stem diameter at breast height (DBH) was
136 estimated as the mean of DBH measured in two directions (north-south and east-
137 west direction) in 1996, 1999, 2004 and 2005. Branch diameter (BrD) was scored
138 as the horizontal diameter in two representative branches in the 10th whorl
139 counted from the top. If this branch whorl did not have any living branches,
140 branches were chosen from the 9th whorl or the next consecutive whorl with
141 living branches. The measurements were taken at a maximum of two cm from
142 the stem with a digital calliper, and were done in 1996, 1999 and 2004. Branch
143 angle (BrA) was measured, in two representative branches, between the stem and

144 the direction of the branch at the base on a scale of 360°. Branches with a
145 representative angle of all the branches present in the tree were chosen for BrA.
146 The two representative branches were not necessarily the same branches that
147 were measured for BrD, but were located in the same whorl. These measurements
148 were scored in 1996, 1999 and 2004. Ht measured between years 1996 and 1999,
149 DBH, BrA and BrD measured at years 1996 and 1999, for individuals 1-93 have
150 been previously published by Lerceteau et al. (2000) and Lerceteau et al. (2001).

151 Needles, from the trees, as close to the top as possible were collected and
152 frozen to calculate hardness (Ct) according to Nilsson and Walfridsson (1995)
153 in 1992, 1997, 1999 and 2004. Presence or absence of female flower production
154 (FP) was scored as a binary character: 1 as present and 0 as absence. Weight of
155 1000 seeds (W) in grams and number of cones per tree (CO) were counted
156 exhaustively at the time of cone collection and seed extraction in 2006. The
157 number of seeds per cone (nSC) were scored as: $nSC =$
158 *Seed number per tree / CO*.

159

160 *Trait measurements on the F₂ OP trials (F723, F725 and F726)*

161 All phenotypic measurements are shown in Table 2 and described here. For
162 simplicity we will refer to trial F723 as Site1, trial F725 as Site2 and trial F726
163 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 2). 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 measured at an additional age of three and at the same age Ct was also assessed in the manner previously described for the full-sib cross.

174

175 **DNA extraction and marker development**

CTAB (Doyle, 1991) method was used to extract DNA from vegetative buds of the F₁ 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

183 datasets were separately used for QTL analysis, for further details see Li et al.

184 (2014)

185

186 **Statistical analysis for estimating breeding values (EBVs)**

187 Progeny tests are primarily established to rank the parents according to the EBVs

188 (White and Hodge, 2013). The different EBVs and their acronyms can be found

189 in Table 2. An “e” preceding the trait acronym indicates that we are referring to

190 an EBV and a number before the underscore sign refers to the OP trial. For

191 example, eVt3_8 refers to an EBV for survival (eVt) at Site3 (eVt3) measured

192 when the trees were eight years old (eVt3_8).

193

194 *Single site spatial analysis*

195 Spatial analysis is useful to detect spatial environmental variation thereby

196 improving the quality of the adjusted data and the accuracy of the genetic

197 parameter estimates (Cappa et al. 2015). Therefore prior to any genetic analysis

198 the data was adjusted for within-trial environmental effects, using a two-

199 dimensional separable autoregressive (AR1) model where row and column

200 directions were fitted for each trial (supplementary information S2). This single

201 model was applied to each trait data (DBH, Ht, BrQ, BrA, Vt and Ct) using

202 ASReml3 (Gilmour et al. 2009). This method divides the residual variance into

203 an independent component and a two dimensional spatially autocorrelated
204 component (Dutkowski et al. 2002, 2006; Bush et al. 2013). An additional
205 random term with one level for each experimental unit was therefore fitted
206 ((Dutkowski et al. 2002; Ivkovic et al. 2015).

207 Diagnostic tools, variogram and plots of spatial residuals were used to
208 detect design, treatment, local and extraneous effects with ASReml 3.0 (Gilmour
209 et al. 2009). The predicted design effects and spatial residuals were extracted
210 from ASReml output files and used to remove all environmental effects from the
211 raw data. The adjusted data were first analysed in a single-site (univariate)
212 analysis to estimate the genetic variance components.

213

214 *Multi-environment analysis (MET): $G \times E$*

215 The spatially adjusted data were also used in a multienvironment analysis
216 (MET). In MET the use of sophisticated models is necessary because the diverse
217 environments may exhibit variance heterogeneity as well as genotype-by-
218 environment interactions. These effects may affect genetic parameter estimations
219 (Ogut et al. 2014, Isik et al. 2017). Therefore, MET was performed using a factor
220 analytic (FA) model in order to explore the additive $G \times E$ in the best possible
221 manner with the final objective of predicting EBVs.

222 FA models have become more popular and efficient in plant breeding
223 MET (Kelly et al. 2009, Cullis et al. 2014, Smith et al. 2015, Chen et al. 2017)
224 because they provide a good parsimonious approximation to unstructured (US)
225 models using few informative factors.

226 The following individual tree (or animal) linear mixed model with FA
227 variance structure was thus used:

$$228 \quad y = X\tau + Z_g u_g + e,$$

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

240 u_g the vector subjected to factor analysis.

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

248 The mixed model in factor analytic form (see supplementary information
 249 S2) was used to model the variance structure of the additive G×E effects:

$$250 \text{ } Var(u_g) = G_g \otimes A ,$$

$$251 \text{ } G_g = \Lambda_g \Lambda_g^T + \Psi_g ,$$

252 Here, G_g is the genetic covariance matrix among trials, A is the
 253 $m \times m$ numerator/additive genetic relationship matrix, Λ_g is the $s \times k$ matrix
 254 of site loadings, and, Ψ is the $s \times s$ diagonal matrix containing site-specific
 255 variance. See supplementary material for the full description of the FA model.

256 To perform MET , ASReml 4.0 (Gilmour et al. 2015) was used.

257 The variance components obtained from the FA models were used to calculate
 258 narrow sense heritability and accuracy of predicted EBVs.

259 Narrow-sense heritability (h^2) was calculated for each site separately but also
 260 overall across all trials, as:

$$h^2 = \frac{\sigma_a^2}{\sigma_p^2},$$

where σ_a^2 is the additive genetic variance and σ_p^2 is the phenotypic variance.

The accuracy of the predicted EBVs was calculated for each F₁ full-sib mother as:

$$r = \sqrt{1 - \frac{PEV}{\sigma_a^2}},$$

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

Further information on the models can be found in supplementary information S2.

Statistical analysis for QTL mapping

For all the traits, each year of measurement was considered separately as a single trait and, therefore, analysed by single trait mapping approach (described below). In addition, growth trajectories were fitted by linear regression (Figure 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 analyze longitudinal data).

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

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

281

282 where y_i is the phenotypic value or estimated breeding value of individual i ($i=1$
283 ... n ; n is the total number of individuals), x_{ij} is the genotypic value of individual
284 i and marker j coded as 0 and 1 for two possible marker genotypes, β_0 is the
285 population mean parameter, β_j is the effect of marker j ($j=1 \dots p$; p is the total
286 number of markers). LASSO regression shrinks the coefficients of non-
287 important markers toward zero, and only keep the important markers (i.e. those
288 strongly associated with the traits) in the model. The tuning parameter λ ($\lambda > 0$)
289 determines how many markers should be retained in the model, and it also
290 controls the degree of shrinkage of the regression parameters. A detailed
291 guideline about how to choose optimal tuning parameters in the QTL mapping
292 context can be found in Li and Sillanpää (2012). To formally judge QTL,
293 hypothesis testing to identify QTL was conducted by a de-biased LASSO
294 approach (Javanmard and Montanari 2014; Li et al. 2017). Due to its shrinkage
295 nature, the original LASSO estimator do not (asymptotically) follow any
296 parametric distribution, and therefore it is impossible to directly estimate the
297 uncertainty of LASSO estimator such as confidence intervals and P -values. The
298 De-biased LASSO approach constructed an un-biased LASSO estimator, which

299 asymptotically follow a normal distribution. The debiased LASSO approach
 300 aims to calculate the P -values for all markers in the study, instead of only the
 301 markers selected by standard LASSO. Li et al. (2017) conducted a simulation
 302 study on debiased LASSO, and the method showed better power to detect QTL
 303 and stronger ability to control false positives compared to conventional single
 304 locus QTL mapping approaches. We used the following criterion for QTL
 305 judgement: a marker was claimed as a significant or suggestive QTL, if its
 306 adjusted P -value was smaller than 0.05 or 0.2.

307

308 **Results**

309 *Accuracy of estimated breeding values (EBVs) and heritabilities*

310 We calculated narrow-sense heritabilities, EBVs and accuracy of EBVs
 311 individually for each site and in the performed MET (see Table 3).

312 Moderate to low heritabilities and EBV accuracies were found for all
 313 traits in all sites. Higher accuracies were found for Ht_8 (0.47) and DBH_8 (0.41)
 314 using the MET analysis (Table 3). From the single-site analysis point of view,
 315 Site2 showed the highest accuracies for DBH_8 (0.37) and Ht_8 (0.44). The EBV
 316 accuracy of Ht at Site3 decreased from 0.32 to 0.21 for age three and eight,
 317 respectively. The accuracy of survival ability at Site3, Vt3_8 (0.44), was the
 318 highest accuracy among all sites and it was almost equal to the accuracy

319 estimated with the MET for Vt_8 (0.43). In the case of Vt_1, the highest accuracy
320 was estimated through the MET (0.32), where single site analysis resulted in
321 slightly smaller estimates (0.18 to 0.28). EBV accuracy of Ct at Site3 age three
322 was moderate (0.44).

323 Heritabilities at Site2 (Table 3) were only higher than for the other sites
324 for DBH_8 (0.35) and Ht_8 (0.56). Among all sites, Site3 had the highest
325 heritabilities for BrQ3_8 (0.19), BrA3_8 (0.22), and Vt_3_8 (0.42). The
326 heritability of hardness measured at age three was 0.55. The heritabilities
327 obtained for the remaining traits and sites were low, ranging between 0.03 and
328 0.10 (Table 3). Heritabilities estimated with the MET analysis were low for all
329 traits, probably due to the considerable G×E-interactions detected in some cases
330 (Table 3) and are described in the following section.

331

332 *Multi-environment analysis (MET): G×E*

333 To detect G×E interaction patterns, genetic correlations across sites were used
334 and are shown in Table 4. The genetic correlations for Ht_8 and DBH_8 (0.94 in
335 both cases) between Site1 and Site2 were high, while the Ht_8 and DBH_8
336 genetic correlations between Site1 and the other sites were very low (0.26 to
337 0.32). For Vt at age one, low genetic correlations (0.24) were detected between
338 Site1 and Site2, as well as between Site2 and Site3. In the case of Vt at age eight,

339 a moderate correlation was detected between Site1 and Site3 (0.67), but
340 otherwise low correlations were observed (0.15 to 0.24). A moderate BrQ_8
341 genetic correlation was detected between Site1 and Site3, and a low correlation
342 was detected for the same trait between Site1 and Site2. However, in both cases
343 the standard errors were very high, which made us consider these correlations as
344 not significant. In the remaining traits all the genetic correlations were very low
345 (below 0.10). BrA_8 genetic correlation between Site1 and Site2 could not be
346 estimated.

347

348 *QTL detection*

349 We found 18 AFLPs and 12 SNPs associated across all phenotypic traits,
350 intercepts and EBVs (Table 5 and supplementary S3). 62 QTL were significant
351 with a range of %PVE explained from 1.7 to 18.9 (Table 5 and supplementary
352 table 1). The lowest %PVE were detected for CO (1.70), followed by Ht at
353 different ages (%PVE from 1.90 to 3.60), Ht intercept (2.00), FP (2.40), DBH
354 (2.20 - 6.40), BrA (3.4), eBrA (3.4), eVt (4), eHt (4.30), Ct (4.30), DBH (5.70),
355 Ht (5.92) and eCt (7.20).

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

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

363 QTL-GAT81 was significantly associated to Ht between ages 13 and 17,
364 QTL-AGC91 to DBH at age 8. Whilst QTL-ARG226 and -AGC205 were
365 significantly linked to BrA at ages 16 and eight respectively. QTL-TAG62 was
366 significantly linked to Ct_4 with all the *P*-values ranging from 0.014 to 0.0003.

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

370 Five AFLP (TGG57, TCG134, TTG454, GGC97 and TTG82) were
371 significantly associated with the S+A data sets. Those QTL were detected for
372 BrA_8, Ct_9, eVt1_1, eVt1_1 and eBrQ2_8, respectively.

373 Among all the detected QTL, TCG89, GGG312, 0_7009_01-358 and
374 axs_47-502, showed significant pleiotropic effects on DBH/Ht, eBrA1_8/W,
375 FP/W/Ct_16/eCt3_3 and CO/Ct, respectively. AGC106, GGG312 and
376 0_7009_01-358 were the only three QTL significantly detected in both, the full-
377 sib and in the half-sib progenies. AGG498 was the only QTL-AFLP showing

378 significant association with eVt based on both, the individual and the MET, EBV
379 estimations.

380 Multiple markers (CGT193, AAG115, TCG89, GAT81, AGC106,
381 0_7009_01-358, 2_10352_02-413, 2_9603_01-344 and 2_10212_01-241) were
382 both significantly associated to the same trait in several years for traits such as
383 Ht, DBH and Ct. Finally, the QTL AAG115 and TCG89 were both significantly
384 associated with Ht and Ht_intercept, respectively.

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

392 Multiple QTL were detected for the same trait across several ages. For
393 example, 2_10352_02-413 explains Ht variation at ages 14 to 17 (Ht_14-17).
394 Similarly, 2_9603_01-344 is associated with Ht at ages 9 and 18, AAG115 from
395 age 12 to 17, GAT81 from age 13 to 17 and TCG89 from age 12 to 18.
396 Furthermore, TCG89 also explains the variation of DBH at ages 16 and 17. The
397 SNP-QTL 0_7009_01-358 is significantly associated to Ct based on phenotypic

398 measurements in the F₁-generation trees at the age of 16 (Ct₁₆) and on EBVs
399 scored in the F₂ progenies at the age of 3 (eCt3₃). Similarly, AGC106 explains
400 the observed variation for Ct at ages 9 and 16 based on phenotypic values and for
401 age 3 based on EBVs.

402

403 **Discussion**

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

408

409 *Genetic architecture (number and %PVE)*

410 This QTL architecture reveals that in both data sets (A, S+A), there are more
411 phenotypic-QTL detected than EBV-QTL. This could be explained, to some
412 extent, by the larger number of phenotypic traits assessed compared to the EBV-
413 based traits (30 phenotypic traits versus 21 EBV-based traits). Multiple studies
414 on simulated and empirical data has shown that a minimum population size is
415 required to detect QTL consistently (Hall et al. 2016), and the number of QTL
416 detected typically increases with population size and heritability (e.g., Vales et
417 al. 2005; Falke and Frisch, 2011; Wang et al. 2012). However, the number of

418 QTL detected and their effects is not only a function of population size or its
419 underlying genetic architecture (i.e., number of genes controlling the trait and
420 their effect size). Other factors, such as marker type (dominant versus co-
421 dominant) can also contribute to the number and effect of the detected QTL. In
422 our study, S+A data set has similar number of AFLP and SNP markers, however,
423 SNP-QTL represent almost the 70% (AFLP 30%) of the total number of QTL
424 detected in the S+A data set. This is possibly due to the higher informative
425 (codominant) nature of the SNPs (i.e., the heterotizgote Aa can be distinguished
426 from the dominant homozygote AA).

427 Congruent with the theoretical and empirical expectations, our
428 experimental design mainly allowed the detection of QTL with medium-to-large
429 effects. In both data sets, QTL for eCt3_3 show the highest %PVE (7.20 for the
430 A dataset and 18.90 for the S+A dataset), and exhibits the second highest h^2
431 (0.55). However, in the S+A data set, we suspect %PVEs to be overestimated due
432 to a lower number of study trees (i.e., 91 F₁ trees for the phenotypic-QTL or 61
433 F₁ trees for the EBV-QTL in the A+S data set compared to 496 F₁ trees for the
434 phenotypic-QTL or 356 F₁ trees for the EBV-QTL in the A data set).
435 Furthermore, Ht in Scots pine has been reported to be under low genetic control
436 (Abrahamsson et al. 2012), in the S+A data set, the QTL associated with Ht are
437 among the highest %PVE values.

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. For example, in *Cunninghamia lanceolata* (Lamb) Hook (Chinese 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 menziesii*, 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 Stoeher, 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 5 and 6). 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-fir (Jermstad et al. 2001b; Wheeler et al. 2005), with %PVE that varies from 0.7 to 24.9 depending on the species, age and tests performed. %PVE ranged from 1.3 to 34.9

458 for height at different ages in different species, for instance in white spruce
459 (Pelgas et al. 2011), *Quercus robur* (Gailing et al. 2008; Scotti-Saintagne et al.
460 2004), *Eucalyptus urophylla* (Gion et al. 2011) and Scots pine (Lerceteau et al.
461 2000, 2001). In the case of BrD and BrA there are not many QTL studies
462 available, but Lerceteau et al. (2001) found two and one QTL with %PVE
463 estimates of 29.5 and 16.1 respectively.

464

465 *Pleiotropic effects*

466 Pleiotropy is a complex phenomenon that refers to a single gene affecting
467 multiple traits (He and Zhang, 2006). We have detected two SNP-QTL that may
468 have pleiotropic effects. 0_7009_01-358 is significantly associated with FP, Ct
469 and W and axs_47_502 is significantly associated with CO and Ct. Pleiotropic
470 effects have previously been described in the conifer literature. For example,
471 Pelgas et al. (2011), reported one possible common gene for budset, bud flush
472 and height growth, and another for budset and height. Lerceteau et al. (2001)
473 found one gene in common between DBH and BrD. Hurme et al. (2000) also
474 found one gene with a possible pleiotropic effect between budset and frost
475 hardiness. In the present work, it is however, not possible to discern between
476 pleiotropic effects and the action of two or more loci in the proximity. It has
477 been described in animal breeding literature, that the possibility of pleiotropic

478 effects are hiding other effects such as linkage of separate loci (Johnsson et al.
479 2014; Wright et al. 2010). However, in our case we did not detect any significant
480 correlation between the phenotypic traits for which we have detected pleiotropic
481 effects.

482

483 *QTL age stability and QTL $\times$ E*

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

493 In our study, we could test the effect of environmental heterogeneity
494 between-sites (i.e., different trials for the same families) and we could also
495 compare the results of individual site versus MET analysis. The lack of
496 reproducibility of the QTL across sites could be a consequence of environmental
497 heterogeneity and G $\times$ 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).

506

507 *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-apirose/xylose synthase. This enzyme catalyzes the NAD⁺-dependent conversion of UDP-

518 glucuronic acid to UDP-D-apiose and UDP-Dxylose (Grisebach and Schmid,
519 1972) and it is involved in the plant cell wall formation. UDP-Dxylose
520 involvement in the pleiotropic control of CO and Ct may be the result of multiple
521 genes in linkage disequilibrium within this single QTL. For example, fine QTL-
522 mapping in rice identified multiple genes associated to the same QTL region
523 (Lim et al. 2014). The 2_10212_01-241 encodes for a glutathione transferase
524 located at the chloroplast. The involvement of this gene in stress regulation
525 (Sappl et al. 2009) seems compatible with this SNP-QTL' association with frost
526 tolerance (Ct). The a3ip2_387 encodes for the ABI3-interacting protein 2 (ABI3),
527 a transcription factor of the abscisic acid signal transduction pathway (Koornneef
528 et al. 1984). In conifers, dormancy and frost tolerance are interconnected
529 processes that share common molecular mechanisms with *Arabidopsis* (Welling
530 and Palva, 2006). Considering the role of ABI3 in frost tolerance (Tamminen et
531 al. 2001) and dormancy (Nambara et al. 1995) in *Arabidopsis*, it could be
532 postulated that a possible involvement of ABI3 in frost tolerance also occurs in
533 conifers, a hypothesis which is reinforced by the known association of this gene
534 in seed development and dormancy in conifers (Zeng et al. 2013). The
535 2_10352_02-413 QTL' best BLAST is a nucleotide sugar epimerase, which is
536 involved in photosynthesis membrane biogenesis and its overexpression leads to
537 growth acceleration in *Arabidopsis* (Li et al. 2011), thus supporting its role in

538 conifer growth. The best BLAST hit result for QTL 2_6731_01-230 was a
539 putative F-box protein GID2. F-box protein, gibberellin-insensitive dwarf2
540 (GID2), mediates the action of phytohormone gibberellin (GA) in the control
541 of growth and development in plants (Gomi et al. 2004). This gives credibility to
542 the SNP-QTL association with growth (Ht) found in this study. We also found a
543 significant association between GI_f1_334 that encodes for gigantean (GI) with
544 vitality (Vt). GI is a very well characterized gene involved in multiple
545 physiological processes including stress responses to cold and drought in
546 *Arabidopsis* (reviewed by Mishra and Panigrahi, 2015). This could be interpreted
547 as a validation of GI involvement in vitality mediating processes in conifers.
548 0_11919_01-122, 2_9603_01-344, 0_17247_02-266 and CL2495Contig1_03-
549 101 were not annotated at the time of this study.
550

551 **Conclusions**

552 The identification of QTL associated with different growth, quality and adaptive
 553 traits through phenotypic measurements and EBVs, has resulted in the detection
 554 of QTL linked to 18 AFLPs and 12 SNPs. 62 significant QTL were detected for
 555 phenotype based traits, probably due to the larger number of phenotype based
 556 traits measured compared with the EBV-based traits. The levels of %PVE of the
 557 QTL ranged from 1.7 to 18.9. Multiple QTL were significantly associated to the
 558 same trait across ages, therefore those QTL would be considered as the most
 559 promising candidates for early selection. We detected environmental
 560 heterogeneity between sites and genotype-by-site interactions, which could be
 561 the main reason behind the absence of reproductibility of some QTL across sites.
 562 We detected Ht and Ct QTL stable across several tree ages, thus indicating that
 563 some of the QTL detected in this study could be considered strong candidates for
 564 early selection for certain traits based on their cross-age stability.

565

566

567

568 **Acknowledgements**

569 We are thankful to Ross Whetten for guidance in the development and
570 genotyping of AFLP markers, to Outi Savolainen for the SNP data, to Sonja
571 Kujala for the information on SNP annotation, and to Skogforsk that planted and
572 maintained the field material, and also for the help with field measurements. We
573 also grateful to Anders Fries and René Fernández for the field sampling, and, Dr.
574 Fikret Isik, Dr. Zhi-Qiang Chen and Dr. Arthur Gilmour for their assistance in
575 the statistical analysis. We are thankful to John Baison for his help with the
576 English language review. This study was supported by the Research School in
577 Forest Genetics and Breeding at the Swedish University of Agricultural Sciences,
578 Knut och Alice Wallenbergs foundation, EVOLTREE EU project and
579 PROCOGEN EU project.

580

581 **Data archiving**

582 Phenotypic data and pedigree information are available in supplementary
583 materials (S1). Genotype datasets are available in files S3 and S4 at Li et al.
584 (2014). Sorting and mapping of the marker data are available at file S5 of Li et
585 al. (2014). The authors state that all the other data necessary to support the
586 conclusions of this work are fully represented within the article.

587 **References**

- 588 Abrahamsson S, Nilsson JE, Wu H, García-Gil MR, Andersson B (2012).
589 Inheritance of height growth and autumn cold hardiness based on two generations
590 of full-sib and half-sib families of *Pinus sylvestris*. *Scandinavian Journal of*
591 *Forest Research* **27**: 405-413.
- 592 Aitken SN, Whitlock MC (2013). Assisted gene flow to facilitate local adaptation
593 to climate change. *Annual Review of Ecology, Evolution, and Systematics* **44**:
594 367-388.
- 595 Bartholome J, Bink MC, van Heerwaarden J, Chancerel E, Boury C, Lesur I, et
596 al. (2016) Linkage and association mapping for two major traits used in the
597 maritime pine breeding program: height growth and stem straightness. *PLoS*
598 *ONE* **11**: e0165323.
- 599 Bartholome J, Salmon F, Vigneron P, Bouvet JM, Plomion C, Gion JM (2013).
600 Plasticity of primary and secondary growth dynamics in *Eucalyptus* hybrids: a
601 quantitative genetics and QTL mapping perspective. *BMC Plant Biology* **13**: 120.
602

- 603 Bian L, Shi J, Zheng R, Chen J, Wu HX (2014). Genetic parameters and
604 genotype–environment interactions of Chinese fir (*Cunninghamia lanceolata*) in
605 Fujian Province. *Canadian Journal of Forest Research* **44**: 582-592.
- 606 Bush D, Marcar N, Arnold R, Crawford D (2013). Assessing genetic variation
607 within *Eucalyptus camaldulensis* for survival and growth on two spatially
608 variable saline sites in southern Australia. *Forest Ecology and Management* **306**:
609 68-78.
- 610 Cappa EP, Muñoz F, Sanchez L, Cantet RJ (2015). A novel individual-tree mixed
611 model to account for competition and environmental heterogeneity: a Bayesian
612 approach. *Tree Genetics & Genomes* **11**: 120.
- 613 Chen ZQ, Karlsson B, Wu HX. (2017). Patterns of additive genotype-by-
614 environment interaction in tree height of Norway spruce in southern and central
615 Sweden. *Tree Genetics & Genomes* **13**: 25.
- 616 Collard BCY, Mackill DJ (2008). Marker-assisted selection: an approach for
617 precision plant breeding in the twenty-first century. *Philosophical Transactions*
618 *of the Royal Society B-Biological Sciences* **363**: 557-572.

- 619 Cullis BR, Jefferson P, Thompson R, Smith AB (2014). Factor analytic and
620 reduced animal models for the investigation of additive genotype-by-
621 environment interaction in outcrossing plant species with application to a *Pinus*
622 *radiata* breeding programme. *Theoretical and Applied Genetics*, **127**: 2193-
623 2210.
- 624 Doyle J (1991). DNA protocols for plants. In *Molecular Techniques in*
625 *Taxonomy* (pp. 283-293). Springer Berlin Heidelberg.
- 626 Dutkowski GW, Silva JCE, Gilmour AR, Lopez GA (2002). Spatial analysis
627 methods for forest genetic trials. *Canadian Journal of Forest Research* **32**: 2201-
628 2214.
- 629 Dutkowski GW, Costa e Silva J, Gilmour AR, Wellendorf H, Aguiar A (2006).
630 Spatial analysis enhances modelling of a wide variety of traits in forest genetic
631 trials. *Canadian Journal of Forest Research* **36**: 1851-1870.
- 632 Falke KC, Frisch M (2011). Power and false-positive rate in QTL detection with
633 near-isogenic line libraries. *Heredity* **106**: 576-584

634 Freeman JS, Potts BM, Downes GM, Pilbeam D, Thavamanikumar S,
635 Vaillancourt RE (2013). Stability of quantitative trait loci for growth and wood
636 properties across multiple pedigrees and environments in *Eucalyptus*
637 *globulus*. *New Phytologist* **198**: 1121-1134.

638 Frewen BE, Chen TH, Howe GT, Davis J, Rohde A, Boerjan W, et al. (2000)
639 Quantitative trait loci and candidate gene mapping of bud set and bud flush in
640 *Populus*. *Genetics* **154**: 837-845.

641 Gailing O, Langenfeld-Hayser R, Polle, A, Finkeldey R (2008). Quantitative trait
642 loci affecting stomatal density and growth in a *Quercus robur* progeny:
643 implications for the adaptation to changing environments. *Global Change*
644 *Biology* **14**: 1934-1946.

645 Gilmour AR, Gogel BJ, Cullis BR, Thompson R. (2009). ASReml User Guide
646 Release 3.0. *VSN International Ltd*, Hemmel Hempstead, UK.

647 Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thompson R (2015). ASReml
648 user guide release 4.1 structural specification. *Hemel Hempstead: VSN*
649 *International Ltd*, Hemmel Hempstead, UK.

- 650 Gion JM, Carouché A, Deweer S, Bedon F, Pichavant F, Charpentier JP, et al.
651 (2011) Comprehensive genetic dissection of wood properties in a widely-grown
652 tropical tree: *Eucalyptus*. *BMC Genomics* **12**: 301.

- 653 Gomi K, Sasaki A, Itoh H, Ueguchi-Tanaka M, Ashikari M, Kitano H, et al.
654 (2004) GID2, an F-box subunit of the SCF E3 complex, specifically interacts
655 with phosphorylated SLR1 protein and regulates the gibberellin-dependent
656 degradation of SLR1 in rice. *The Plant Journal* **37**: 626-634.

- 657 Grisebach H, Schmid R (1972). Chemistry and biochemistry of branched-chain
658 sugars. *Angewandte Chemie International Edition* **11**: 159-173.

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

- 662 Hall D, Hallingbäck HR, Wu HX (2016). Estimation of number and size of QTL
663 effects in forest tree traits. *Tree Genetics & Genomes* **12**: 110.

664 Hallingbäck HR, Sánchez L, Wu HX (2014). Single versus subdivided
665 population strategies in breeding against an adverse genetic correlation. *Tree*
666 *Genetics & Genomes* **10**: 605-617.

667 Hanley SJ, Pei MH, Powers SJ, Ruiz C, Mallott MD, Barker JH, et al. (2011)
668 Genetic mapping of rust resistance loci in biomass willow. *Tree Genetics &*
669 *Genomes* **7**: 597-608.

670 Hawkins BJ, Stoehr M (2009). Growth, phenology, and cold hardiness of 32
671 Douglas-fir full-sib families. *Canadian Journal of Forest Research* **39**: 1821-
672 1834.

673 He X, Zhang J (2006). Toward a molecular understanding of
674 pleiotropy. *Genetics* **173**: 1885-1891.

675 Holliday JA, Ritland K, Aitken SN (2010). Widespread, ecologically relevant
676 genetic markers developed from association mapping of climate-related traits in
677 Sitka spruce (*Picea sitchensis*). *New Phytologist* **188**: 501-514.

- 678 Hurme P, Sillanpää MJ, Arjas E, Repo T, Savolainen O (2000). Genetic basis of
679 climatic adaptation in Scots pine by Bayesian quantitative trait locus
680 analysis. *Genetics* **156**: 1309-1322.
- 681 Isik F (2014). Genomic selection in forest tree breeding: the concept and an
682 outlook to the future. *New Forests* **45**: 379-401.
- 683 Isik F, Holland J, Maltecca C (2017). Genetic Data Analysis for Plant and Animal
684 Breeding. Vol 1. Springer International Publishing, New York.
- 685 Ivković M, Gapare W, Yang H, Dutkowski G, Buxton P, Wu HX (2015). Pattern
686 of genotype by environment interaction for radiata pine in southern
687 Australia. *Annals of Forest Science* **72**: 391-401.
- 688 Jansson G (2007). Gains from selecting *Pinus sylvestris* in southern Sweden for
689 volume per hectare. *Scandinavian Journal of Forest Research* **22**: 185-192.
- 690 Javanmard A, Montanari A (2014). Confidence intervals and hypothesis testing
691 for high-dimensional regression. *Journal of Machine Learning Research* **15**:
692 2869–2909.

693 Jermstad KD, Bassoni DL, Jech KS, Wheeler NC, Neale DB. (2001). Mapping
694 of quantitative trait loci controlling adaptive traits in coastal Douglas-fir. I.
695 Timing of vegetative bud flush. *Theoretical and Applied Genetics* **102**: 1142-
696 1151.

697 Jermstad KD, Bassoni DL, Wheeler NC, Anekonda TS, Aitken SN, Adams WT,
698 et al. (2001) Mapping of quantitative trait loci controlling adaptive traits in
699 coastal Douglas-fir. II. Spring and fall cold-hardiness. *Theoretical and Applied*
700 *Genetics* **102**: 1152-1158.

701 Johnsson M, Rubin CJ, Höglund A, Sahlqvist AS, Jonsson KB, Kerje S, et al.
702 (2014) The role of pleiotropy and linkage in genes affecting a sexual ornament
703 and bone allocation in the chicken. *Molecular Ecology*, **23**: 2275-2286.

704 Kelly AM, Cullis BR, Gilmour AR, Eccleston JA, Thompson R (2009).
705 Estimation in a multiplicative mixed model involving a genetic relationship
706 matrix. *Genetics Selection Evolution* **41**: 33.

707 Khorsand Rosa R, Oberbauer SF, Starr G, Parker La Puma I, Pop E, Ahlquist L,
708 et al. (2015) Plant phenological responses to a long-term experimental extension

709 of growing season and soil warming in the tussock tundra of Alaska. *Global*
710 *Change Biology* **21**: 4520-4532.

711 Koornneef M, Reuling G, Karssen CM (1984). The isolation and characterization
712 of abscisic acid-insensitive mutants of *Arabidopsis thaliana*. *Physiologia*
713 *Plantarum* **61**: 377-383.

714 Lai M, Sun X, Chen D, Xie Y, Zhang S (2014). Age-related trends in genetic
715 parameters for *Larix kaempferi* and their implications for early selection. *BMC*
716 *Genetics* **15**: S10.

717 Lander ES, Botstein D (1989). Mapping mendelian factors underlying
718 quantitative traits using RFLP linkage maps. *Genetics* **121**: 185-199.

719 Lerceteau E, Plomion C, Andersson B (2000). AFLP mapping and detection of
720 quantitative trait loci (QTL) for economically important traits in *Pinus sylvestris*:
721 a preliminary study. *Molecular Breeding* **6**: 451-458.

722 Lerceteau E, Szmidt AE, Andersson B (2001). Detection of quantitative trait loci
723 in *Pinus sylvestris* L. across years. *Euphytica* **121**: 117-122.

724 Li C, Wang Y, Liu L, Hu Y, Zhang F, Mergen S, et al. (2011) A rice plastidial
725 nucleotide sugar epimerase is involved in galactolipid biosynthesis and improves
726 photosynthetic efficiency. *PLoS Genetics* **7**: e1002196.

727 Li Z, Sillanpää MJ (2012a). Overview of LASSO-related penalized regression
728 methods for quantitative trait mapping and genomic selection. *Theoretical and*
729 *Applied Genetics* **125**: 419–435.

730

731 Li Z, Hallingbäck HR, Abrahamsson S, Fries A, Gull BA, Sillanpää MJ, et al.
732 (2014) Functional multi-locus QTL mapping of temporal trends in scots pine
733 wood traits. *G3 (Bethesda)* **4**: 2365-2379.

734 Li Z, Guo B, Yang J, Herczeg G, Gonda A, Balázs G, et al. (2017) Deciphering
735 the genomic architecture of the stickleback brain with a novel multilocus gene-
736 mapping approach. *Molecular Ecology* **26**: 1557-1575.

737 Lim CC, Krebs SL, Arora R (2014). Cold hardiness increases with age in juvenile
738 *Rhododendron* populations. *Frontiers in Plant Science* **5**: 542

- 739 Lim JH, Yang HJ, Jung KH, Yoo SC, Paek NC (2014). Quantitative trait locus
740 mapping and candidate gene analysis for plant architecture traits using whole
741 genome re-sequencing in rice. *Molecules and Cells* **37**: 149.
- 742 Linderholm HW (2006). Growing season changes in the last
743 century. *Agricultural and Forest Meteorology* **137**: 1-14.
- 744 Markussen T, Fladung M, Achere V, Favre JM, Faivre-Rampant P, Aragones A,
745 et al. (2003). Identification of QTL controlling growth, chemical and physical
746 wood property traits in *Pinus pinaster* (Ait.). *Silvae Genetica* **52**: 8-14.
- 747 Matuszewski S, Hermisson J, Kopp M (2015). Catch me if you can: adaptation
748 from standing genetic variation to a moving phenotypic optimum. *Genetics* **200**:
749 1255-1272.
- 750 Mauricio R (2001). Mapping quantitative trait loci in plants: uses and caveats for
751 evolutionary biology. *Nature Reviews Genetics* **2**: 370-381.
- 752 McNamara S, Pellett H (2000). Cold hardiness of *Phellodendron sachalinense*
753 Friedr. Schmidt seedlings increases with age. *HortScience* **35**: 304-305.

- 754 Mishra P, Panigrahi KC (2015). GIGANTEA—an emerging story. *Frontiers in*
755 *Plant Science*, **6**: 8.
- 756 Mrode RA (2014). Linear Models for the Prediction of Animal Breeding Values.
757 Cabi. UK
- 758 Nambara E, McCourt P, Naito S (1995). A regulatory role for the ABI3 gene in
759 the establishment of embryo maturation in *Arabidopsis*
760 *thaliana*. *Development* **121**: 629-636.
- 761 Nilsson JE, Walfridsson EA (1995). Phenological variation among plus-three
762 clones of *Pinus sylvestris* (L.) in Northern Sweden. *Silvae Genetica* **44**: 20-28.
- 763 Ogut F, Maltecca C, Whetten R, McKeand S, Isik F (2014). Genetic analysis of
764 diallel progeny test data using factor analytic linear mixed models. *Forest*
765 *Science* **60**: 119-127.
- 766 Pagliarini MK, Kieras WS, Moreira JP, Sousa VA, Shimizu JY, Moraes MLT, et
767 al. (2016) Adaptability, stability, productivity and genetic parameters in slash
768 pine second-generation families in early age. *Silvae Genetica* **65**: 71-82.

769 Pelgas B, Bousquet J, Meirmans PG, Ritland K, Isabel N (2011). QTL mapping
770 in white spruce: gene maps and genomic regions underlying adaptive traits across
771 pedigrees, years and environments. *BMC Genomics* **12**: 145.

772 Persson T, Andersson B (2003). Genetic variance and covariance patterns of
773 growth and survival in northern *Pinus sylvestris*. *Scandinavian Journal of Forest*
774 *Research* **18**: 332-343.

775 Persson T, Andersson B, Ericsson T (2010). Relationship between autumn cold
776 hardiness and field performance in northern *Pinus sylvestris*. *Silva Fennica* **44**:
777 255-266.

778

779 Piepho HP, Möhring J, Melchinger AE, Büchse A (2008). BLUP for phenotypic
780 selection in plant breeding and variety testing. *Euphytica* **161**: 209-228.

781 Plomion C, Durel CE, O'Malley DM (1996). Genetic dissection of height in
782 maritime pine seedlings raised under accelerated growth conditions. *Theoretical*
783 *and Applied Genetics* **93**: 849-858.

784 Prada E, Alia R, Climent J, Diaz R (2014). Seasonal cold hardiness in maritime
785 pine assessed by different methods. *Tree Genetics & Genomes* **10**: 689-701.

- 786 Sappl PG, Carroll AJ, Clifton R, Lister R, Whelan J, Harvey Millar A, et al.
787 (2009) The Arabidopsis glutathione transferase gene family displays complex
788 stress regulation and co-silencing multiple genes results in altered metabolic
789 sensitivity to oxidative stress. *The Plant Journal* **58**: 53-68.
- 790 Sax K (1923). The association of size differences with seed-coat pattern and
791 pigmentation in *Phaseolus vulgaris*. *Genetics* **8**: 552-560.
- 792 Scotti-Saintagne C, Bodénès C, Barreneche T, Bertocchi E, Plomion C, Kremer
793 A (2004). Detection of quantitative trait loci controlling bud burst and height
794 growth in *Quercus robur* L. *Theoretical and Applied Genetics* **109**: 1648-1659.
- 795 Smith AB, Ganesalingam A, Kuchel H, Cullis BR (2015). Factor analytic mixed
796 models for the provision of grower information from national crop variety testing
797 programs. *Theoretical and Applied Genetics* **128**: 55-72.
- 798 Tamminen I, Makela P, Heino P, Palva ET (2001). Ectopic expression of ABI3
799 gene enhances freezing tolerance in response to abscisic acid and low
800 temperature in *Arabidopsis thaliana*. *The Plant Journal* **25**: 1-8.

801 Thavamanikumar S, Southerton SG, Bossinger G, Thumma BR (2013).
802 Dissection of complex traits in forest trees—opportunities for marker-assisted
803 selection. *Tree Genetics & Genomes* **9**: 627-639.

804 Thumma BR, Southerton SG, Bell JC, Owen JV, Henery ML, Moran GF (2010).
805 Quantitative trait locus (QTL) analysis of wood quality traits in *Eucalyptus*
806 *nitens*. *Tree Genetics & Genomes* **6**: 305-317.

807 Tibshirani R (1996). Regression shrinkage and selection via the lasso. *Journal of*
808 *the Royal Statistical Society: Series B* **58**: 267–288.

809 Vales MI, Schön CC, Capettini F, Chen XM, Corey AE, Mather DE, et al. (2005)
810 Effect of population size on the estimation of QTL: a test using resistance to
811 barley stripe rust. *Theoretical and Applied Genetics* **111**: 1260-1270.

812 Wakeley PC (1961). Results of the southwide pine seed source study through
813 1960-61. Southern Forest Experiment Station. In: *Proceedings of the 6th*
814 *Southern Conference on Forest Tree Improvement*. 1961 June 7-8. Gainesville,
815 FL. pp. 10-24.

- 816 Walther A, Linderholm HW (2006). A comparison of growing season indices for
817 the Greater Baltic Area. *International Journal of Biometeorology* **51**: 107-118.

- 818 Wang H, Smith KP, Combs E, Blake T, Horsley RD, Muehlbauer GJ (2012).
819 Effect of population size and unbalanced data sets on QTL detection using
820 genome-wide association mapping in barley breeding germplasm. *Theoretical*
821 *and Applied Genetics* **124**: 111-124.

- 822 Welling A, Palva ET (2006). Molecular control of cold acclimation in
823 trees. *Physiologia Plantarum* **127**: 167-181.

- 824 Weng Y, Ford R, Tong Z, Krasowski M (2017). Genetic parameters for bole
825 straightness and branch angle in jack pine estimated using linear and generalized
826 linear mixed models. *Forest Science* **63**: 111-117.

- 827 Wheeler NC, Jermstad KD, Krutovsky K, Aitken SN, Howe GT, Krakowski J, et
828 al. (2005) Mapping of quantitative trait loci controlling adaptive traits in coastal
829 Douglas-fir. IV. Cold-hardiness QTL verification and candidate gene
830 mapping. *Molecular Breeding* **15**: 145-156.

- 831 White TL, Adams WT, Neale DB (2007). Forest Genetics. *CABI*. UK.

832 White TL, Hodge GR (2013). Predicting Breeding Values with Applications in
833 Forest Tree Improvement (Vol. 33). Springer Science & Business Media.

834 Woldendorp G, Hill MJ, Doran R, Ball MC (2008). Frost in a future climate:
835 modelling interactive effects of warmer temperatures and rising atmospheric
836 [CO₂] on the incidence and severity of frost damage in a temperate evergreen
837 (*Eucalyptus pauciflora*). *Global Change Biology* **14**: 294-308.

838

839 Wright D, Rubin CJ, Martinez Barrio A, Schütz K, Kerje S, Brändström H, et al.
840 (2010) The genetic architecture of domestication in the chicken: effects of
841 pleiotropy and linkage. *Molecular Ecology* **19**: 5140-5156.

842

843 Wu HX, Hallingbäck HR, Sánchez L (2016). Performance of seven tree breeding
844 strategies under conditions of inbreeding depression. *G3 (Bethesda)* **6**: 529-540.

845

846 Yazdani R, Nilsson JE, Plomion C, Mathur G (2003). Marker trait association for
847 autumn cold acclimation and growth rhythm in *Pinus sylvestris*. *Scandinavian*
848 *Journal of Forest Research* **18**: 29-38.

849

850 Zeng Y, Zhao T, Kermode AR (2013). A conifer ABI3-interacting protein plays
851 important roles during key transitions of the plant life cycle. *Plant*
852 *Physiology* **161**: 179-195.

853

854 Zhang X, Gou M, Liu CJ. (2013). Arabidopsis Kelch repeat F-box proteins
855 regulate phenylpropanoid biosynthesis via controlling the turnover of
856 phenylalanine ammonia-lyase. *The Plant Cell* **25**: 4994-5010.

857

858 Table 1 Trials description.

Trial nr	Trial id	Location	LAT (N)	LON (E)	ALT		Gen	Year	N
					(m)				
1	F485	Flurkmark	64.03	20.5	115		F ₁	1988	1000
2	F723 (Site1)	Hennan	62.031	15.648	375		F ₂	2008	3804
3	F725 (Site2)	Sollefteå	63.197	17.272	110		F ₂	2008	3532
4	F726 (Site3)	Kebbeberge	64.105	17.674	400		F ₂	2008	4134
5	Sävar	Sävar	63.895	20.549	65		F ₂	2007	8640

859 Trial nr, number of the trial; Trial id, trial identification; LAT, latitude; LON, longitud; ALT,

860 altitude; Gen, generation; Year, plantation year; N, number of trees in each trial.

861

862 Table 2 Traits measured and their acronyms by trial.

Trait	Acronym	Site			
		F485	F723 (Site1)	F725(Site2)	F726(Site3)
Flower production	FP	FP			
Number of cones per tree	CO	CO			
Weight of 1000 seeds	W	W			
Number of seeds per cone	nSC	nSC			
Height	Ht	Ht_8, Ht_9, Ht_10, Ht_11, Ht_12, Ht_13, Ht_14, Ht_15, Ht_16, Ht_17, Ht_18	eHt1_8	eHt2_8	eHt3_3, Ht3_8
Stem diameter at breast height	DBH	DBH_8, DBH_11, DBH_16, DBH_17	eDBH1_8	eDBH2_8	eDBH3_8

Hardiness	Ct	Ct_4, Ct_9, Ct_11, Ct_16			eCt3_3
Branch diameter	BrD	BrD_8, BrD_11, BrD_16			
Branch quality	BrQ		eBrQ1_8	eBrQ2_8	eBrQ3_8
Branch angle	BrA	BrA_8, BrA_11, BrA_16	eBrA1_8	eBrA2_8	eBrA3_8
Survival ability	Vt		eVt1_1, eVt1_8	eVt2_1, eVt2_8	eVt3_1, eVt3_3, eVt3_8

863 The first column in the table is the trait description, followed by the trait acronym used for each trait and site of measurement. The number
864 before the underscore indicates the site where the trait was measured (except for the FS and Sävar sites where no number was used), and the
865 number after the score indicates the age of measurement. An “e” preceding the name acronym indicates that the trait was used to calculate
866 EBV.

867 Table 3 (a) Average of EBV centered to the mean (a_EBV), (b) heritabilities and their standard errors in between
 868 brackets, and (c) accuracy of the EBV (r_EBV), for the traits used both in the MET analysis 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

869

(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
-----	------	--	--	--	------	------	------	------	------

870

(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

871

872 Table 4 Across site genetic correlations (and standard errors) for traits used in the MET analysis.

	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

873 n.e.: non estimable.

875 Table 5 Description of the significant QTL.

Trait	Total nQTL	nQTL A dataset	nQTL A+S dataset	<i>P</i> -value	PVE (%) range
BrA_8	2	1	1	0.006	4.90 - 15.00
BrA_16	1	1		0.002	3.4
eBrA1_8	1	1		0.021	3.40
BrD_11	1		1	0.042	11.4
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
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	6.4
DBH_16	1	1		0.006	2.2
DBH_17	1	1		0.028	2.2
FP	1		1	0.01	2.4
Ht intercept	2	2		0.007	2.6

Ht_8	1		1	0.0005	10
Ht_9	1		1	0.032	11.2
Ht_12	2	2		0.016	2
Ht_13	3	3		0.006	2.00 - 2.20
Ht_14	4	3	1	0.004	1.90 - 10
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	4.3
eVt_8	1	1		0.004	4
eVt1_1	3		3	0.002	6.50 - 13.90
eVt3_8	1		1	0.008	3.8
W	2		2	0.028	11 - 11.90

876

877

878

879

880

881 Table 6. QTL detected in the literature for growth and adaptive traits

Species	Trait	n QTL	%PVE	Reference
White spruce	Ht test 1_pop1_year 1	4	9.3 - 10.5	Pelgas et al. 2011
	Ht test 1_pop2_year 1	5	3.4 - 6.6	
	Ht test 1_pop1_year 2	0		
	Ht test 1_pop2_year 2	12	3 - 8.2	
	Ht test 1_pop2_year 3	8	2.5 - 9.2	
	Ht test 2_pop1_year 1	0		
	Ht test 2_pop2_year 1	8	2.8 - 7.5	
	Ht test 2_pop1_year 2	0		
	Ht test 2_pop2_year 2	10	2.9 - 5.2	
	Ht test 2_pop2_year 3	4	3.3 - 6.2	
<i>Quercus robur</i>	Ht year 1	8	2.7 - 8.6	Gailing et al. 2008
	Ht year 2	10	2.1 - 11.9	
	Ht year 3	9	2.7 - 12	
Sitka spruce	Hardiness	9	0.7 - 5.4	Holliday et al. 2010
<i>Eucalyptus urophylla</i>	Ht 59 months	3	4.1 - 7.9	Gion et al. 2011
Scots pine	Ht	3	25.8	Lerceteau et al. 2000
Scots pine	DBH	2	9.3 - 15	Lerceteau et al. 2000
Scots pine	Ht year 9	3	32.2	Lerceteau et al. 2001
	Ht year 10	3	33.6	

	Ht year 11	3	33.9	
	Ht year 12	3	34.9	
	DBH year 9	1	15	
	DBH year 12	3	27.3	
	BrD year 12	2	29.5	
	BrA year 9	1	16.1	
<i>Quercus robur</i>	Ht	4	1.3 - 3.6	Scotti-Saintagne et al. 2004
Scots pine	Hardiness	4	24.9	Hurme et al. 2000
Coastal Douglas-fir	Hardiness	8	2.8 - 7.5	Wheeler et al. 2005
Maritime pine	Seedling Ht	6	5.5 - 12	Plomion et al. 1996
Scots pine	Hardiness test 1 year 1	2	9 - 15.2	Yazdani et al. 2003
	Hardiness test 2 year 1	2	6 - 12.9	
	Autumn hardiness test 1 year 1	1	12.9	
	Autumn hardiness test 2 year 1	2	10.8 - 11.7	
	Hardiness test 1 year 2	2	12 - 16.5	
	Autumn hardiness test 1 year 2	2	10.4 - 19.5	
Coastal Douglas-fir	Lateral bud flush site 1 year 1	6	1.2 - 5.6	Jermstad et al. 2001 a
	Lateral bud flush site 1 year 2	6	3.0 - 6.2	
	Lateral bud flush site 1 year 3	6	2.3 - 6.6	

	Terminal bud flush site 1 year 1	5	2.0 - 6.6	
	Terminal bud flush site 1 year 2	3	3.4 - 4.8	
	Terminal bud flush site 1 year 3	1	6.8	
	Bud flush site 1 year 4	8	2.1 - 5.2	
	Terminal bud flush site 2 year 2	2	2.4 - 8.4	
	Bud flush site 2 year 4	6	4.1 - 11.5	
	Bud flush QTL_x_year interaction	8	1.6 - 7.6	
	Bud flush QTL x site 1 interaction	2	4.8 - 6.5	
	Bud flush QTL x site 2 interaction	2	5.7 - 8.4	
	Bud flush QTL x site 3 interaction	5	1.3 - 11.6	
Coastal Douglas-fir	Cold hardiness trait 1	2	4.3 - 4.4	Jermstad et al. 2001 b
	Cold hardiness trait 2	1	6.8	
	Cold hardiness trait 3	5	2 - 4.4	
	Cold hardiness trait 4	3	1.4 - 5.8	
	Cold hardiness trait 5	5	3.8 - 9.8	
	Cold hardiness trait 6	4	4.5 - 8.7	

882 Table describe in the first column de specie in which de QTL were detected; Trait, refers to the
883 treatment in which the QTL were detected; n QTL, is the number of QTL detected; %PVE, is the

884 percentage of variance explained; reference is the published article in which the information was

885 found.

886

887

888

889

890

891 Figure 1. Growth trajectories of the height (measured in cm) of the 500 trees from

892 year 1996 to 2006.

893

894