Additive and non-additive genetic variance in juvenile Sitka spruce (Picea sitchensis Bong. Carr)

Additive and non-additive genetic variance in juvenile Sitka spruce (*Picea sitchensis* Bong. Carr)

J.J. Ilska*[1], D.J. Tolhurst*[1], H. Tumas[2], P. Maclean[3], J. Cottrell[3], S.J. Lee[3], J. Mackay[2], and J.A. Woolliams[1]

* These authors should be considered as joint first authors.

1. The Roslin Institute, Royal (Dick) School of Veterinary Science, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, Scotland, UK.
2. Department of Biology, University of Oxford, South Parks Road, Oxford, OX1 3RB, UK.
3. Forest Research, Northern Research Station, Roslin, Midlothian, EH25 9SY, UK.

Corresponding author: john.woolliams@roslin.ed.ac.uk

ORCID IDs

JAW: 0000-0002-7880-9972

JJI: 0000-0002-1563-0174

DJT: 0000-0002-4787-080X

JC: 0000-0001-6355-1326

JM: 0000-0002-4883-195X

Abstract

Offspring from three full-sib Sitka spruce families were clonally replicated to produce 12 ramets from each of 1500 offspring per family. The trees were planted across three sites, with each offspring represented by one ramet in each of four blocks per site. All trees were measured for height at 2, 4, 6 and 11 years of age and, depending on site and family, for bud burst at 5 years and pilodyn penetration depth at 10 years. Genotypes were obtained by RADseq for all six parents and for a subset of 1524 offspring and a linkage map was used to impute genotypes for 1630 loci segregating in all three families. A linear mixed model was developed which combines all available phenotypic and genomic data on all genotyped and non-genotyped trees to estimate the fraction of additive genetic variance (f_a) for each trait and family. The consensus value for f_a increased with age for height from 0.60 at 2 years to 0.80 at 11 years, with non-overlapping 95% support intervals (I_{95}). The estimated value of f_a for bud burst was 0.83 (I_{95}=[0.78, 0.90]) and was 0.80 (I_{95}=[0.76, 0.92]) for pilodyn depth. There was no evidence of differences in f_a between families for bud burst and height, but there was evidence of differences for pilodyn depth ($P<0.05$). There was no evidence of inbreeding depression associated with genomic homozygosity, which would be expected if dominance variance were the major component of the non-additive variance, and the results offer no support for the development of sublines for crossing within the species.

[256 words]

Introduction

Breeding is well-established in many forest tree species but it is often hindered by the lengths of generation intervals. Current breeding cycles in conifers including spruces, pines, larches, firs, among others, commonly exceed 15 years requiring candidate trees to have reached reproductive age and have sufficiently accurate predictions of breeding value obtained from early predictors of adult performance, possibly supplemented with progeny testing, prior to selection decisions being made. The history of Sitka spruce (*Picea sitchensis* [Bong.] Carr.) in the UK is an example of a well-planned and executed breeding program which is faced with this challenge of long intervals. Sitka is a conifer species originating from the Pacific North West extending from south-eastern Alaska to northern California. It was first brought to the UK in the 1830's (Lee et al. 2013), and now accounts for over 50% of all the area planted with conifers and ~25% of all woodland area of Great Britain (IFOS-Statistics, 2022). Breeding objectives for the species relate to its primary use for construction timber and wood pulp (Lee et al. 2013), and although improvements have been achieved since the start of plus tree selection in the early 1960s, only two cycles of selection have been completed (Lee and Connelly, 2010). This time constraint along with the high cost of field evaluations, among others, has often limited the ability to fully characterise the genetic control of phenotypic traits, such as their partitioning into additive and non-additive gene effects.

An attraction of genomic prediction is the potential to transform forest breeding through reducing generation intervals while retaining sufficient accuracy of the predicted breeding values (EBVs) to obtain faster rates of improvement (Grattapaglia, 2017). This is due to a different approach to estimating breeding values using molecular data (Meuwissen et al. 2001) and genomic relationship coefficients (Van Raden, 2008), compared to using pedigree and the matrix of numerator relationship coefficients derived from it. In the pedigree approach, the breeding values are predicted from models which, beyond the base generation, rely on estimating Mendelian sampling terms of individuals, which in turn relies on obtaining phenotypic information on the candidate or offspring. In contrast, when applying the molecular approach, the breeding values are predicted from the estimated effects of (dense and genome wide) marker alleles, typically SNPs, which can be obtained for all genotyped individuals provided relevant data are available for estimating the SNP effects. With genotypes

available from ‘conception’ one barrier to reducing the generation interval and obtaining an EBV that encompasses an individual’s own genome is removed.

While most attention in tree breeding (as in other fields) has focused upon developing genomic predictions of breeding value, or additive genetic merit, and the variance of the breeding values defines the additive genetic variance. However, the total genetic variance includes contributions from non-additive genetic variation, and predicting non-additive effects can be used to improve the merit of those deployed in the wider forest population for timber. The SNP data make it possible to access the non-additive genetic effects more directly, and to predict non-additive components of the total genetic merit (Vitezica et al. 2017; Joshi et al. 2020). One benefit of using the genomic data is that it is feasible to estimate non-additive genetic variance from simpler designs than would be necessary using pedigree data. In forestry, the ease of vegetative propagation allows clonal experiments to be established which provide material to estimate the total genetic variance and broad heritability (H^2), while the genotypic data can be used to estimate genomic relationships, and hence estimate the additive genetic variance and narrow sense heritability (h^2). Consequently, the extent and potential impact of the non-additive genetic variance can be estimated.

One of the challenges of advancing the use of genomic techniques in Sitka has been the need to generate the thousands of SNP marker genotypes on selection candidates to provide a marker coverage of the genome that is sufficiently dense. There are multiple ways of obtaining SNP genotypes, e.g. through SNP chip arrays, whole genome resequencing, or reduced representation sequencing. Restriction-site associated DNA sequencing (RADseq) belongs to a group of reduced representation sequencing methods which have been particularly popular in non-model species. The benefits of RADseq are its flexibility and relatively low cost compared to whole genome resequencing (Parchman et al. 2018) but it is particularly attractive for species, including many conifers, with large repetitive genomes where the compilation of a draft reference genome is challenging (Pan et al. 2015; Fuentes-Utrilla et al. 2017; Parchman et al. 2018). While assays for RADseq have been described for Sitka (Fuentes-Utrilla et al. 2017), this was for a single family and their application and performance across multiple families is unknown. One of the drawbacks of RADseq is the stochastic

nature of the sequence reads for a given coverage, particularly when the coverage is low but this can be overcome using imputation (e.g. Li et al. 2009).

The goal of this paper was to estimate the fraction of additive genetic variance in the total genetic variance of Sitka spruce, based on SNP markers derived from RADseq data and phenotypic data collected on height, wood density and bud burst in the offspring of three full sib families. The newly developed linkage map of the Sitka spruce genome (Tumas et al. 2023) allowed us to apply an imputation procedure which enabled missing genotypes to be recovered, thereby making maximum use of the available SNP data. The tree height data was collected at different ages, and allowed the sensitivity to site and family variation to be studied as it came from three large full sib families, clonally replicated across three geographically and climatically diverse sites. The analyses employed spatial modelling to account for natural and extraneous variation within each site (Gilmour et al. 1997). To the authors' knowledge, this is the first paper where the heritability of economically important traits in conifers was estimated using analyses which simultaneously accounted for additive and non-additive genetic variance based on genomic data along with spatial modelling.

Materials & Methods

Population

The phenotypic and SNP data were based on material in a large field experiment established in 2005 by Forest Research. The experiment consisted of three full-sib families (denoted as FS1, FS2 and FS3) each with 1,500 offspring (where offspring represent a unique genotype), clonally replicated across three contrasting sites: Huntly, Llandovery and Torridge (Table 1).

The families were created by controlled pollination of maternal clones growing in the Sitka spruce clone bank of Forest Research. Each offspring was vegetatively propagated from cuttings to produce 12 ramets (clonally replicated copies of an offspring genotype), with four ramets of each genotype on each site and, within sites, one ramet of each genotype in each of four randomised blocks. In addition to 1,500 offspring trees, each block contained 46 control trees raised from open pollinated

seed collected from Haida Gwaii (formerly Queen Charlotte Islands), British Columbia. The trees of FS1 at Torridge formed the data for a previous publication (Fuentes-Utrilla et al. 2017).

Traits

All trees had their height measured after 2, 4, 6 and 11 years of age. In addition, the depth of penetration of a pilodyn pin at breast height after 10 years was recorded as an indicator trait for wood density at the Torridge site only. Trees from family FS1 were also scored for the timing of bud-burst at the start of the fifth year, using an 8-point scale according to Krutzsch (1973) at all sites. This scoring was carried out on three occasions within a three-week period (5A, 5B and 5C). A summary of the design for the measurement of traits is shown in Table 2. The number of trees available for measurement of height at each age is shown in Table 3, which also provides a guide to numbers of trees assessed for bud burst and pilodyn measurements at the intermediate ages.

RADseq Genotypes

Assay

The RADseq data used for FS1 were initially produced for Noveltree (https://cordis.europa.eu/project/id/211868), and those for FS2 and FS3 were produced for Procogen (https://cordis.europa.eu/project/id/289841). DNA was extracted from the needles of all 6 parents and from randomly-selected subsets of the 1,500 offspring in each family, with 622, 493 and 496 sampled from FS1, FS2 and FS3 respectively. The protocols for DNA extraction and RADseq digestion were fully described in Fuentes-Utrilla et al. (2017). Briefly, DNA was extracted using a Qiagen DNeasy Plant mini-kit but with the protocol modified to maximize DNA quantity. The extracted DNA was then subjected to a double-digest RADseq protocol using AlnwI and PstI enzymes. Paired-end reads were produced for parents, and single-end for offspring using Illumina HiSeq2000. While the protocol for the RADseq digestion was the same for all 3 families, the resulting average read length ranged from 45bp in offspring of FS1 to 112bp in parents of FS2 and FS3.

Bioinformatic pre-processing of RADseq data

The raw, barcoded fastq-libraries were de-multiplexed using RADtools v1.2.4 (Baxter et al. 2011). The paired-end reads of parents were then screened for PCR duplicates using a Perl script (Kerth,

2014) which removed between 22-24% reads in parents of FS1 and 43-46% reads in parents of FS2 and FS3. Offspring whose number of reads fell outside 3 standard deviations from the overall mean within each family were removed, and this resulted in the removal of 5, 4 and 8 offspring from families FS1, FS2 and FS3 respectively. Adapter sequences were removed from all reads using Scythe v.0.994 (Buffalo, 2014). The reads were then processed with the 'process radtags' package of Stacks v.2 (Rochette and Catchen, 2017) to remove reads with uncalled bases and quality scores <20, and then to truncate all reads to 45bp to allow simultaneous processing of all three families. The 'k-mer filter' option of Stacks v.2 was used to remove both abundant and rare k-mers, with the default k-mer size set to 15. The final number of reads retained for further analysis ranged between 17.4 and 20.2M for each of the six parents, and between 2.2 and 3.5M reads for each of the 1,594 remaining offspring.

SNP genotyping

SNP markers were identified and genotyped using the Stacks v.2 pipeline: 'ustacks' to build loci within a sample; 'cstacks' to construct a catalogue of loci from parental samples; 'sstacks' to match loci from all samples to the catalogue; 'tsv2bam' to transpose data to become locus oriented; 'gstacks' to call variants sites and genotyping individuals. The parameters used for the genotyping were optimised as recommended in Paris et al. (2017). The outcome was: minimum stack depth (*m*) set to 2, distance between stacks (*M*) set to 3, and distance between catalogued loci (*n*) set to 3. The resulting genotypes were exported to a vcf format using the 'populations' package of Stacks v.2, parameterised so that a locus was processed if it was detected in at least 3 populations (*p*=3), and in at least 3% of all individuals across all populations (*R*=3). The parameter values were chosen to minimise the number of Mendelian inconsistencies and missing values across the resulting genotypes.

SNP quality control

The genotypes were split into 3 within-family datasets using Plink (Purcell et al. 2007). Quality control was then applied within each family by sequentially removing individuals with call rates less than 0.6 and then removing SNPs with call rates less than 0.8 and MAF<0.15. Note that within families the

segregating SNPs are expected to have frequencies of either 0.25, 0.5 or 0.75. The resulting call rates across all retained individuals and SNPs were 0.77, 0.79 and 0.81 for FS1, FS2 and FS3 respectively. The SNP genotypes were then tested for Mendelian inconsistencies using a custom Python script (https://github.com/joannailska/Mendelian_inconsistencies), with a Bonferroni correction for multiple testing. The resulting numbers of trees and SNPs per family are reported in Table 4. Among the retained SNPs, 2,054 SNPs were segregating in all 3 families and offered an element of validation, and these are henceforth referred to as “common SNPs”.

Imputation

Among the common SNPs, 1,630 (78%) had been reliably assigned to the 12 linkage groups of the linkage map compiled by Tumas et al. (2023) and this map was used for imputation. For each of the three families used in this study, the genotyping data for these 1,630 loci were assigned to the 12 linkage groups and ordered within them. The genotypes identified were processed using AlphaPeel (v1.1.0; Whalen et al. 2018) using the multi-locus peeling option, with an additional parameter giving the map distance in Morgans spanning the loci for each linkage group. The distribution of SNPs across linkage groups and families is shown in Table 5. Imputation accuracy was assessed using posterior genotype probabilities provided by AlphaPeel for the full-sib offspring and summarised by assuming genotypes to be assigned if a genotype posterior probability was greater than p and varying p over the range 0.5 to 1. For a given value of p, the SNP call rate over offspring and offspring call rates over SNPs were calculated.

Statistical models

A site and family combination is hereafter referred to as a trial (with 9 trials in total). Each trial was designed as a randomised complete block design with four replicate blocks. Each block was allocated 46 control trees and 1,500 offspring trees. Due to topographic constraints, some blocks were spatially separated (non-contiguous) which required “master blocks” to be constructed. For trials with non-contiguous blocks, two master blocks were created and filler plots with missing phenotypes were added to ensure a continuous spatial structure within each master block. Trials

with contiguous blocks were treated as having a single master block. The number of master blocks in each trial is shown in Table 2.

All models were fitted separately for each trial, and included: (i) a preliminary spatial model without genomic data, and (ii) an extended spatial model with genomic data. The novel feature of (ii) is that phenotypic data was included on all control and offspring trees (regardless of whether they have been genotyped) while genomic data was also included on the subset of genotyped trees in each family. This preserved all data to estimate the genetic and non-genetic variance parameters, and also enabled estimates of additive and non-additive genetic variance parameters to be obtained.

Preliminary spatial model

This is a univariate BLUP model which accommodates the experimental design and spatial modelling for each trial. The linear mixed model for $\mathbf{y}$, the vector of available phenotypic data on all 46 control trees and 1,500 offspring trees in each trial, can be written as:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_u\mathbf{u} + \mathbf{Z}_v\mathbf{v} + \mathbf{e} \qquad (1)$$

where $\mathbf{b}$ is the vector of fixed effects, here only the mean, with $\mathbf{X}$ being a vector 1's, $\mathbf{u}$ is the vector of random genetic effects for all offspring trees with design matrix $\mathbf{Z}_u$, $\mathbf{v}$ is a vector of random non-genetic effects, here only blocks, with design matrix $\mathbf{Z}_v$, and $\mathbf{e}$ is the vector of residuals. The genetic effects are assumed to be distributed as $\mathbf{u}$ ~ MVN($\mathbf{0}$, $\sigma_u^2\mathbf{I}$), where σ_u^2 is the total genetic variance. The block effects are assumed to be distributed as $\mathbf{v}$ ~ MVN($\mathbf{0}$, $\sigma_v^2\mathbf{I}$), where σ_v^2 is the block variance. The residuals are assumed to be distributed as $\mathbf{e}$ ~ MVN($\mathbf{0}$, $\mathbf{R}$), where $\mathbf{R}$ is the residual variance matrix which includes a model for natural and extraneous variation, i.e. variation due to random error and correlated spatial error (Gilmour et al. 1997). The residual variance matrix is given by:

$$\mathbf{R} = \sigma_r^2\mathbf{I} + \sigma_s^2 \oplus_{k=1}^{b} \mathbf{\Sigma}_{c(k)}(\rho_c) \otimes \mathbf{\Sigma}_{r(k)}(\rho_r) \qquad (2)$$

where σ_r^2 is the random error variance and σ_s^2 is the spatial error variance, such that $\mathrm{f_r} = \sigma_r^2/(\sigma_r^2 + \sigma_s^2)$ and $\mathrm{f_s} = \sigma_s^2/(\sigma_r^2 + \sigma_s^2)$ are the fractions of random and spatial error variance. The Kronecker plus operator ($\oplus$) constructs a block-diagonal model across the b master-blocks ($b = 1$ or 2; Table 2) and the Kronecker product operator ($\otimes$) constructs a separable model between the

columns and rows in each master block. Note that the model for each master block is parameterised by different auto-correlation matrices, i.e. $\boldsymbol{\Sigma}_{c(k)}$ and $\boldsymbol{\Sigma}_{r(k)}$, but the same auto-correlation parameters, ρ_c and ρ_r. The significance of the spatial models were informally assessed by log-likelihood ratio tests and the Akaike Information Criterion, and showed considerable improvements compared to models with independent residuals, i.e. $\mathbf{e} \sim \text{MVN}(\mathbf{0}, \sigma_e^2\mathbf{I})$. The model described by Equations [1] and [2] is hereafter referred to as Model 1.

Extension to include genomic data

Model 1 was then extended to genomic BLUP using a genomic relationship matrix, $\mathbf{G}$, derived from RADseq data (described below). This model included phenotypic data on all 46 control trees and 1,500 offspring trees, while genomic data were also included on the 572, 470 and 482 genotyped trees in FS1, FS2 and FS3 respectively. Let the vector of genetic effects be partitioned as $\boldsymbol{u} = (\boldsymbol{u}_1, \boldsymbol{u}_2)$ where $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are the vectors for the non-genotyped and genotyped trees, respectively. Since there is clonal replication, the genetic effects for the genotyped trees can be further partitioned into additive and non-additive effects, where $\boldsymbol{u}_2 = \boldsymbol{u}_{2(a)} + \boldsymbol{u}_{2(d)}$. The design matrix is partitioned conformably with $\boldsymbol{u}$, where $\mathbf{Z} = [\mathbf{Z}_1 \ \mathbf{Z}_2]$.

The linear mixed model for $\mathbf{y}$ can now be written as:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_1\mathbf{u}_1 + \mathbf{Z}_2[\mathbf{u}_{2(a)} + \mathbf{u}_{2(d)}] + \mathbf{Z}_v\mathbf{v} + \mathbf{e} \qquad (3)$$

where the non-genetic and residual terms are as described for Model 1. The genetic effects for the non-genotyped trees are assumed to be distributed as $\mathbf{u}_1 \sim \text{MVN}(\mathbf{0}, \sigma_u^2\mathbf{I})$. The additive genetic effects (breeding values) for the genotyped trees are assumed to be distributed as $\mathbf{u}_{2(a)} \sim \text{MVN}(\mathbf{0}, \sigma_a^2\mathbf{G})$, where σ_a^2 is the additive genetic variance and $\mathbf{G}$ is the genomic relationship matrix. The non-additive effects for the genotyped trees are assumed to be distributed as $\mathbf{u}_{2(d)} \sim \text{MVN}(\mathbf{0}, \sigma_d^2\mathbf{I})$, where σ_d^2 is the non-additive genetic variance. The model described by Equations [2] and [3] is hereafter referred to as Model 2. Model 2 was repeated with heterozygosity included as a covariate: this extension is of interest in describing -1x inbreeding depression, but potentially removes non-additive variance and results are given in Supplementary File 1.

Model 2 provides a direct estimate of the total genetic variance from the non-genotyped trees (σ_u^2) and an indirect estimate from the genotyped trees, which is a function of the additive (σ_a^2) and non-additive (σ_d^2) genetic variances. In terms of the additive genetic variance, it should be noted that the model parameter σ_a^2 is *not* the additive genetic variance of each family since it corresponds to a population with markers in Hardy-Weinberg equilibrium. In a full-sib family there is negative assortment of alleles, i.e. increased heterozygosity, when selfing is excluded. The true additive genetic variance of each family was therefore estimated by $k\sigma_a^2$, where $k = \overline{diag(\mathbf{G})} - \bar{\mathbf{G}}$ with the bar denoting the mean value, and $k =$ 0.669, 0.686 and 0.603 for FS1, FS2, and FS3 respectively. Note that scaling was not necessary for σ_d^2 and σ_u^2 since $\overline{diag(\mathbf{I})} - \bar{\mathbf{I}} \approx 1$ when the number of genotyped and non-genotyped trees is large. The total genetic variance for the non-genotyped and genotyped trees was therefore constrained as:

$$\sigma_u^2 = k\sigma_a^2 + \sigma_d^2 \qquad (4)$$

This constraint was applied when fitting Model 2 (see below).

Genomic relationship matrix

The genomic relationship matrix, $\mathbf{G}$, was constructed separately for each family following Van Raden's Model 1 (Van Raden, 2008), using the alternative allele dosages for each locus for each genotyped tree provided by AlphaPeel (Whalen et al. 2018) following imputation. The dosage is the expected number of alternative alleles accounting for the genotypic probabilities, and takes values between 0 and 2. For example, if the genotype probabilities for locus j of tree i are 0.01, 0.99 and 0.00 for allele counts 0, 1 and 2 the dosage is 0.99 $(0 \times 0.01 + 1 \times 0.99 + 2 \times 0.00)$. The allele frequencies (p_i for locus i) used for centring the dosages and calculating the scaling factor ($\sum_i 2p_i\,(1 - p_i)$) were calculated from the full-sib parents for each family and were therefore either 0.25, 0.50 or 0.75, as each parent had been genotyped to high coverage, and additionally had been imputed from the large full-sib family.

Model fitting

Models 1 and 2 were fitted separately for each trait and trial in *ASReml-R* (Butler et al., 2018), which obtains REML estimates of the variance parameters and empirical BLUPs of the random effects. The spatial model in Equation [2] was constructed by fitting a separate model for each master block,

with the sets of auto-correlation and variance parameters constrained to be equal across master blocks using the *vcc* argument (see Tolhurst et, 2019). Model 2 was fitted with the constraint in Equation [4] using the 'own' function, which constructs user specified variance models. The variance models for the genotyped trees were constructed as $\text{var}(\mathbf{u}_{2(a)}) = \sigma_u^2 \text{f}_\text{a} \mathbf{G}/k$ and $\text{var}(\mathbf{u}_{2(d)}) = \sigma_u^2 \text{f}_\text{d} \mathbf{I}$, where $\text{f}_\text{a} = k\sigma_a^2/\sigma_u^2$ and $\text{f}_\text{d} = \sigma_d^2/\sigma_u^2$ are the fractions of additive and non-additive genetic variance and σ_u^2 is constrained to equal the total genetic variance of the non-genotyped trees, i.e. $\text{var}(\mathbf{u}_1) = \sigma_u^2 \mathbf{I}$.

Model summaries

Sample variograms showing the residual semi-variance between plots were constructed using a custom Python script (https://github.com/joannailska/Sitka_variogram/blob/main/variogram_Sitka.py). The variograms were used to informally assess the spatial models and detect any additional extraneous variation in the column or row directions. An example is presented and summarised in Supplementary File 2.

Model 2 provided estimates of H^2 and h^2. These were defined for the observed populations of full-sibs, as $H^2 = \sigma_u^2/(\sigma_u^2 + \sigma_v^2 + \sigma_r^2 + \sigma_s^2)$ and $h^2 = f_a \sigma_u^2/(\sigma_u^2 + \sigma_v^2 + \sigma_r^2 + \sigma_s^2)$. The confidence intervals for the estimated fraction of additive genetic variance were obtained from likelihood profiles calculated by constraining f_a in Model 2 to take values over relevant ranges in the interval [0,1], most densely around the REML estimates. The 95% confidence intervals were defined by the interval for which the drop in 2logL was less than 3.84, the 95% point for χ_1^2. Estimates of the parameters were pooled across families, and sometimes sites as described in the results. For the estimates of spatial parameters and heritabilities, this was done by weighting estimates by the reciprocal of their sampling variance. For the fraction of additive genetic variance, this was done by summing the likelihood profiles.

Results

Imputation

The cumulative distribution functions of the call rates for SNPs over offspring are shown in Figure 1a, and those for offspring over SNPs in Figure 1b, for no imputation and for different thresholds (*p)* for the posterior probability required to call a genotype following imputation by AlphaPeel. All such functions will tend to 1 as the call rate tends to 1, and if all genotypes were known with certainty the function would be a step function, where f(x)=0 for x<1 and f(x)=1 for x=1. The distribution functions will asymptote towards this step function as the number of genotypes called increases, and the sensitivity of the distribution functions to the value of *p* decreases as confidence in the imputation increases. When the threshold was set to *p*=0.9, 96% of SNPs had call rates exceeding 95% over all offspring (from Figure 1a), and 94% of offspring had call rates exceeding 95% over all SNPs (from Figure 1b). Without imputation, only 50% of SNPs and 67% of offspring had call rates exceeding 95%.

Spatial Parameters

The spatial parameters described in Equation [2] are treated in this study as nuisance parameters and are summarised below in less detail than the genetic parameters of interest. The sample variograms presented in Supplementary Information 2 shows an example outcome from fitting Model 2 and illustrate the residual semi-variance between plots *x* rows and *y* columns apart. The variograms peak at the spatial (σ_s^2) and total error variance ($\sigma_r^2 + \sigma_s^2$), with the discontinuity at zero displacement reflecting the random error variance (σ_r^2). The general shape of the variograms is determined by the auto-correlation parameters ρ_c and ρ_r. Table 6 summarises the fraction of random error variance (f_r) and the auto-correlation parameters for height at all four ages. Since the 'column' and 'row' labels were arbitrarily assigned for each site, the values for ρ_r and ρ_c have been pooled into a common value ρ.

Two trends for height were observed in Table 6: (i) f_r diminished from 2 to 11 years of age, indicating stronger spatial (positive) associations in height with neighbours as the trees grew; and (ii) the auto-correlations differed between sites, indicating that the observable associations extended over longer distances at Torridge, and conversely smallest at Llandovery. The value of f_r and ρ for pilodyn depth averaged across families at Torridge was 0.64 (range [0.39, 0.77]) and 0.92 (range [0.79, 0.98]),

354 respectively. For the three measurements of bud burst at 5 years of age (5A, 5B and 5C) for FS1 at
355 the three sites, f_r was comparatively high (mean 0.81; range [0.73, 0.94]) and ρ was also
356 comparatively high (mean 0.78; range [0.61, 0.97]). Taken together, although the common
357 environmental component of variance among neighbours decays slowly for all traits, there is
358 substantial environmental variance independent of neighbours for these ages.
359
360 *Pilodyn depth at 10 years*
361 Pilodyn depth was measured at 10 years in all three families at Torridge only, with the results shown
362 in Table 7. The total genetic variance, σ_u^2, was considerable in all families, although the broad sense
363 heritability, H^2, differed widely between families (range [0.112, 0.349]). These differences were
364 largely due to the differing environmental variances. Considerable additive genetic variance, σ_a^2,
365 was detected in all families with differences in h^2 that reflected the differences in H^2. This
366 correspondence was due to a relative constancy in the fraction of additive genetic variance, f_a. Figure
367 2 shows the likelihood profile for f_a in each family, together with the consensus profile pooled across
368 families. The consensus value for f_a was estimated as 0.80 with 95% confidence interval of [0.76,
369 0.92]; although this estimate was within the 95% confidence intervals for each family, the hypothesis
370 of a common value across families was rejected by the chi-squared test ($P<0.05$; $X^2 = 6.16$ c.f. χ^2_2).
371
372 *Bud burst at 5 years*
373 Bud burst at five years was only measured in FS1 and Table 8 focuses on the first measurement
374 (5A), and the results for the other two measurements are given in the Supplementary Information 3.
375 Estimates of H^2 differed between sites (range [0.276, 0.476]) and these differences were, again,
376 reflected in the estimates of h^2 for the sites. However, f_a was very similar across sites. Figure 3 shows
377 the likelihood profiles for f_a and the consensus profile pooled across sites. The consensus value for
378 f_a was estimated as 0.83 with 95% confidence interval of [0.78, 0.90]. There was no evidence to
379 reject the hypothesis of a common value across sites ($P>0.05$; $X^2 = 0.617$ c.f. χ^2_2). Similar results
380 were obtained for measurements 5B and 5C, which had consensus values of 0.91 (s.e. 0.03) and
381 0.89 (s.e. 0.04) respectively.
382

Height at 2, 4, 6 and 11 years

Table 9 shows the broad sense heritabilities, H^2, and phenotypic variances, σ_P^2, for height measured at 2, 4, 6 and 11 years in all four families at all three sites. For each site, σ_P^2 increased with age but there was no clear trend in the changes in H^2 with age. The estimates of H^2 for Llandovery were generally smaller than for Huntly and Torridge, which were associated with generally larger σ_P^2 at a given age compared to the other sites. There were no clear trends in H^2 or σ_P^2 between families across ages or sites.

The consensus values of f_a pooled across families are given in Table 10 for each age and site. There was no evidence of differences in f_a between families for each age and site ($P<0.05$). However, there was evidence of increasing f_a with age at all sites, in particularly f_a was largest at 11 years and smallest at 2 years. This trend is particularly evident in the consensus values of f_a after pooling across sites with f_a increasing from 0.60 at 2 and 4 years to 0.80 at 11 years, with 95% confidence intervals that do not overlap. The confidence interval for the consensus value of f_a at 11 years does not include 1, i.e. not all genetic variance is additive. However, some individual families at some individual sites do include 1 in their confidence intervals, which are wider, and the best point effort for FS2 at 11 years at Llandovery was 1. There was no evidence of differences between families across sites from the goodness of fit tests.

Discussion

This study combines all available phenotypic and genomic data from a multi-site, clonally replicated experiment with large full-sib families produced by controlled crossing to partition the genetic variance observed between clones for height, bud burst and pilodyn penetration depth into additive and non-additive components. The additive genetic variance formed the largest fraction of total genetic variation for all traits, with estimates of 0.60 for height at 2 years of age increasing to 0.80 at 11 years, 0.80 for pilodyn penetration depth at 10 years, and ranging from 0.83 to 0.91 for the 3 measures of bud burst at 5 years. This partition is possible as the model underlying the Van Raden relationship matrix, **G**, is a ridge regression model on marker allele counts and therefore only

412 describes what is observed as an additive sum of effects over loci, whereas the total genetic variance
413 obtained from the clonal replication includes dominance and epistasis. The experimental design had
414 several aspects that made the study feasible, or more powerful, beyond the clonal replication of the
415 offspring. Firstly, the experiment's large full-sib families made it possible to consolidate genotypes
416 obtained from RADseq by imputation, using the recent availability of a molecular map for Sitka
417 spruce (Tumas et al. 2023). Secondly, the measurement of traits across sites, or across families, or
418 both, allowed for a degree of replication estimates of f_a, and the estimates were found to be very
419 largely consistent, subject to their sampling errors.
420
421 The majority of studies identified by the authors that partition the genetic variance in forest species
422 have used models based on pedigree combined with clones. An important theoretical perspective to
423 consider when comparing the current results with those from published studies is that here the
424 partitioning has been carried out entirely within full-sib families. Therefore, the estimates presented
425 here are partitions of the Mendelian sampling variance and not the full genetic variance for a random-
426 mating population. The expectations for the additive and non-additive components can be scaled up
427 to the corresponding variance for a full random-mating population, and based on these expectations,
428 the fraction f_a would increase. Although half the additive variation lies within families, a greater portion
429 (3/4) of the dominance and the additive x additive epistatic variation is within families, and more than
430 3/4 for higher order epistatic terms (Falconer & Mackay, 1996). Assuming that any non-additive
431 variation observed within families is explained by dominance or additive by additive, then the
432 expectation is that f_a in this study corresponds to $3 f_a /(2+ f_a)$ in a random mating population; e.g. f_a
433 = 0.6 and 0.8 corresponds to 0.69 and 0.86. While only three full-sib families were sampled, the
434 consensus values for f_a estimated for the traits measured on all families is important.
435
436 Among previous studies, Weng et al. (2008) estimated partitions of genetic variance in white spruce,
437 a close relative to Sitka spruce, for a similar range of ages for height, and also for pilodyn depth.
438 Their results show comparable estimates of f_a ranging from ~ 0.4 at 4 years to 0.8 at 14 years,
439 despite the large s.e.s found in their data. The study of Nguyen et al. (2022) in Norway spruce
440 covered a range of ages for height between 6 and 12 years and their results also appear to suggest

441 that f_a decreases between these ages, however examination of the results show large s.e.'s and
442 negative estimates which f_a seriously limit interpretation. Results for Norway spruce were also
443 reported by Chen et al. (2019) using genomic analysis: for height at 17 years, f_a ~ 0.4 and 0.6 at
444 two sites, and for pilodyn depth at 30 years of age f_a ~ 1 at both sites. Among other studies of height,
445 Isik et al. (2003) assessed four ages between 1 and 6 years and Baltunis et al. (2007) at 2 years,
446 both in loblolly pine, Baltunis et al. (2013) at 12 years in yellow cypress but for the large sampling
447 errors limit comparability. Few studies have examined pilodyn depth, but those that have are in
448 agreement with the findings here that the fraction of additive genetic variance is very high, with
449 estimates of 0.90 (s.e. 0.18) at 26 years of age in white spruce (Nguyen et al. 2022); ~0.8 in
450 *Eucalyptus globulus* at 4 years derived from the results of Costa de Silva et al. (2004). There are no
451 comparable results for bud burst in other published studies. Each trait should be expected to have
452 its own architecture, but too few results are available to attempt generalisation particularly given the
453 substantial standard errors of many estimates (stem diameter in Norway spruce (Nguyen et al. 2022;
454 Berlin et al. 2019), *Eucalyptus globulus* (Costa de Silva et al. 2004) and radiata pine (Baltunis et al.
455 2009); wood quality traits in white spruce (Nguyen et al. 2022) and Norway spruce (Chen et al. 2019).
456
457 This study partitions the genetic variance in Sitka spruce into additive and non-additive components
458 using an approach similar to that of de Almeida Filho et al. (2019), which used the classical ridge
459 regression model to estimate the fraction of additive genetic variance and the clonal variance to
460 estimate the total genetic variance. However, their approach requires all trees to be genotyped, and
461 removes any non-genotyped trees. In this paper, a linear mixed model was developed which
462 combines all available phenotypic and genomic data on all trees, regardless of whether they have
463 been genotyped. In particular, the fraction of additive genetic variance was estimated using the
464 subset of genotyped trees and the total genetic variance was estimated using all genotyped and
465 non-genotyped trees. This approach preserves all available data to estimate the genetic and non-
466 genetic variances, which is particularly important for spatial modelling (as it requires a continuous
467 spatial structure). This approach also bears similarities to single-step GBLUP (Legarra et al. 2009),
468 but without the need for pedigree or the need to construct a **H** matrix. The distinguishing feature here
469 is that the primary goal of this study was to obtain reliable estimates of the fraction of additive genetic

variance, rather than obtaining predictions of additive genetic merit for genomic selection. In fact, this approach is equivalent to setting the non-genotyped trees as diagonal (independent) in the genomic relationship matrix within Model 2, so that the additive component for these trees would not be well defined.

Obtaining reasonable precision on the fraction of additive genetic variance using pedigree alone has proved challenging as it typically involves scaling up and calculating linear functions of the estimated pedigree components. The models used here are parsimonious in that no attempt has been made to partition the non-additive genetic variance into dominance and epistatic components to avoid overfitting. The further partition is in general feasible, without assuming Hardy-Weinberg equilibrium, as shown by Vitezica et al. (2017), and exemplified in *Nile tilapia* by Joshi et al. (2020). This involves using the markers to calculate orthogonal relationship matrices for the dominance and epistatic components (e.g. the additive by additive relationship matrix is proportional to the Hadamard product of G with itself). This was attempted in the study of Chen et al. (2019) in Norway spruce but assumed Hardy-Weinberg equilibrium. Furthermore, no attempt has been made here to estimate genetic variance across families (as distinct from pooling the results within families) for two reasons: (i) the number of parents is small, and (ii) the number of markers are too few for satisfactory estimation across families, but more than adequate within families (Lillehammer et al. 2013). This leads to limitations in interpretation of this study, e.g. there is no assessment of whether additive marker effects in one family are similar to those from another, despite the consistency of the f_a observed within families. The different sets of anonymous markers segregating within the small number of families would make this uninterpretable.

In conclusion, the evidence suggests that the fraction of additive genetic variance increases with age for height towards the high fractions observed for pilodyn and bud burst. The results of Supplementary Information 1 show no evidence of inbreeding depression for any of the traits and therefore no evidence on the form of the non-additive genetic variance e.g. in the study of Joshi et al. (2020) in *Nile tilapia* the extra genetic variance observed in full-sibs aligned with additive by additive epistasis and not dominance. While the form of the non-additive genetic variance may be

less relevant for deployment strategies using clones, it does influence the form of the breeding program, as additive by additive fractions become converted to additive variance under selection and little benefit is expected from establishing sub-lines for crossing, as in reciprocal recurrent selection.

Competing Interests

The authors have no competing interests.

Acknowledgments

The authors gratefully acknowledge funding for the RADseq assays from European Commission Framework 7 programme through the projects Noveltree (Grant Agreement ID 211868) and Procogen (Grant Agreement ID 289841). The authors are also grateful to Dr S. O'Hara for the DNA extraction in both Noveltree and Procogen projects, and Rob Sykes and the Technical Support Unit at Forest Research for organising and collecting phenotypes and field samples. The bioinformatic and genetic analyses were part of the Sitka Spruced project funded by Biotechnology and Biological Science Research Council under grants BB/P018653/1 and BB/P020488/1.

Data Archiving Statement

All data will be made available on reasonable request. Some data is already in data repositories and remaining data will be lodged on publically accessible data repositories.

References

Baltunis BS, Huber DA, White TL, Goldfarb B, Stelzer HE (2007) Genetic gain from selection for rooting ability and early growth in vegetatively propagated clones of Loblolly pine. Tree Genet Genomes 3:227-238. https://doi.org/10.1007/s11295-006-0058-9

Baltunis B, Russell J, Van Niejenhuis A, Barker J, El-Kassaby Y (2013) Genetic analysis and clonal stability of two yellow cypress clonal populations in British Columbia. Silvae Genet, 62:173-187. https://doi.org/10.1515/sg-2013-002

Baltunis BS, Wu HX, Dungey HS, Mullin T, Brawner JT (2009) Comparisons of genetic parameters and clonal value predictions from clonal trials and seedling base population trials of radiata pine. Tree Genet Genomes, 5:269-278. https://doi.org/10.1007/s11295-008-0172-y

Baxter SW, Davey JW, Johnston JS, Shelton AM, Heckel DG et al. (2011) Linkage mapping and comparative genomics using next-generation RAD-sequencing of a non-model organism. PloS ONE 6:e19315. https://doi.org/10.1371/journal.pone.0019315

Berlin M, Jansson G, Högberg K-A, Helmersson A (2019) Analysis of non-additive genetic effects in Norway spruce. Tree Genet Genomes, 15:42. https://doi.org/10.1007/s11295-019-1350-9

Buffalo V (2014) Scythe – a Bayesian adapter trimmer (version 0.994 beta). https://github.com/vsbuffalo/scythe

Costa e Silva J, Borralho NMG, Potts BM (2004) Additive and non-additive genetic parameters from clonally replicated and seedling progenies of *Eucalyptus globulus*. Theor App Genet 108:1113-1119. https://doi.org/10.1007/s00122-003-1524-5

Chen, ZQ, Baison J, Pan J, Westin J, Gil MRG, Wu HX. (2019) Increased Prediction Ability in Norway Spruce Trials Using a Marker X Environment Interaction and Non-Additive Genomic Selection Model. J Hered. 110:830-843. http://doi.org:10.1093/jhered/esz061.

de Almeida Filho JE, Guimarães JFR, Fonsceca e Silva F, Vilela de Resende, MD, Muñoz P et al. (2019) Genomic prediction of additive and non-additive effects using genetic markers and pedigrees. *G3 Genes Genomes Genet* 9:2739-2748. https://doi.org/10.1534/g3.119.201004

Falconer DS and Mackay T (1996) Introduction to quantitative genetics. 4th Edition. Addison Wesley Longman, Harlow

Fuentes-Utrilla P, Goswami C, Cottrell JE, Pong-Wong R, Law A et al. (2017) QTL analysis and genomic selection using RADseq derived markers in Sitka spruce: the potential utility of within family data. Tree Genet Genomes. 13:33. https://doi.org/10.1007/s11295-017-1118-z

Grattapaglia D (2017) Status and Perspectives of Genomic Selection in Forest Tree Breeding. In: Varshney RK, Roorkiwal M, Sorrells ME (eds). Genomic Selection for Crop Improvement. Springer, Cham, pp 199–249. http://doi.org/10.1007/978-3-319-63170-7_9

Kerth C (2014) 'purge_PCR_duplicates.pl'. https://github.com/claudiuskerth/scripts_for_RAD

Krutzsch P (1973) Norway spruce development of buds. Royal College of Forestry.

Li Y, Willer C, Sanna S, Abecasis G. (2009) Genotype imputation. Annu Rev Genomics Hum Genet. 10:387-406. https://doi.org/10.1146/annurev.genom.9.081307.164242.

Lee SJ, Connolly T (2010) Finalizing the selection of parents for the Sitka spruce (*Picea sitchensis* (Bong.) Carr) breeding population in Britain using mixed model analysis. Forestry 83:423–31. https://doi.org/10.1093/forestry/cpq024

Lee S, Thompson D, Hansen JK (2013) Sitka Spruce (*Picea sitchensis* (Bong.) Carr. In: Pâques LE (ed) Forest Tree Breeding in Europe: Current State-of-the-Art and Perspectives. Springer, Dordrecht, pp 177–227. https://doi.org/10.1007/978-94-007-6146-9_4

Lee SJ, Woolliams J, Samuel CJA, Malcolm DC (2002) A study of population variation and inheritance in Sitka spruce II. Age trends in genetic parameters for vigour traits and optimum selection ages. Silvae Genet 50:55-65

Legarra A, Aguilar I, Misztal I (2009) A relationship matrix including full pedigree and genomic information. J Dairy Sci. 92:4656–4663. https://doi.org/10.3168/jds.2009-2061

Lillehammer M, Meuwissen THE, Sonesson AK (2013. A low-marker density implementation of genomic selection in aquaculture using within-family genomic breeding values. Genet Sel Evol. 45:39. https://doi.org/10.1186/1297-9686-45-39

IFOS-Statistics (2022) Forestry Statistics 2022. https://cdn.forestresearch.gov.uk/2022/09/Ch1_Woodland_2022.pdf

Isik F, Li B, Frampton J (2003) Estimates of additive, dominance and epistatic genetic variances from a clonally replicated test of Loblolly pine. For Sci 49:77-88. https://doi.org/10.1093/forestscience/49.1.77

Joshi R., Meuwissen THE, Woolliams JA, Gjøen HM (2020) Genomic dissection of maternal, additive and non-additive genetic effects for growth and carcass traits in *Nile tilapia*. Genet Sel Evol 52:1-13. https://doi.org/10.1186/s12711-019-0522-2

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Nguyen HTH, Chen Z-Q, Fries A, Berlin M, Hallingbäck HR et al. (2022) Effect of additive, dominant and epistatic variances on breeding and deployment strategy in Norway spruce. Forestry 95:416-427. https://doi.org/10.1093/forestry/cpab052

Pan J, Wang B, Pei Z-Y, Zhao W, Gao J, Mao J-F et al. (2015) Optimization of the genotyping-by-sequencing strategy for population genomic analysis in conifers. Mol Ecol Resour 15:711–722. http://doi.org/10.1111/1755-0998.12342

Parchman TL, Jahner JP, Uckele KA, Galland LM, Eckert AJ (2018) RADseq approaches and applications for forest tree genetics. Tree Genet Genomes 14:39. https://doi.org/10.1007/s11295-018-1251-3

Paris JR, Stevens JR, Catchen JM (2017) Lost in parameter space: a road map for Stacks. Methods Ecol Evol. 8:1360–1373. http://doi.org/10.1111/2041-210X.12775

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575. http://doi.org/10.1086/519795

Rochette NC, Catchen JM (2017) Deriving genotypes from RAD-seq short-read data using Stacks. Nat Protoc 12:2640–2659. https://doi.org/10.1038/nprot.2017.123

Tumas H, Ilska JJ, Maclean P, Cottrell J, Lee SJ, Woolliams JA, Mackay J (2023) High-density genetic linkage mapping in Sitka spruce supports integration of genomic resources. G3. Submitted.

VanRaden PM (2008) Efficient methods to compute genomic predictions. J Dairy Sci, 91:4414-4423. http://doi.org/10.3168/jds.2007-0980.

Vitezica ZG, Legarra A, Toro MA, Varona L (2017) Orthogonal estimates of variances for additive, dominance, and epistatic effects in populations. Genetics 206:1297-1307. http://doi.org/10.1534/genetics.116.199406

Weng YH, Park YS, Krasowski MJ, Tosh KJ, Adams G (2008) Partitioning of genetic variance and selection efficiency for alternative vegetative deployment strategies for white spruce in Eastern Canada. Tree Genet Genomes, 4:809-819. http://doi.org/10.1007/s11295-008-0154-0

Whalen A, Ros-Freixedes R, Wilson DL, Gorjanc G, Hickey JM (2018). Hybrid peeling for fast and accurate calling, phasing, and imputation with sequence data of any coverage in pedigrees. Genet Sel Evol 50:67. https://doi.org/10.1186/s12711-018-0438-2

Table 1. Geographic and climatic characteristics of the three sites. The accumulated temperature is defined by the number of days above 5°C using historical data from the UK Meteorological Office over the 30-year recording period 1961-1990.

Characteristic	Units	Huntly	Llandovery	Torridge
		Scotland	Wales	England
Latitude	°N	57.58	51.97	50.82
Longitude	°W	2.82	4.12	4.37
Height above sea level	m	140	230	120
Accumulated temperature	°days	1106	1450	1828

Table 2. The design of trials showing which traits were measured at which sites and in which families (FS: 1, 2 or 3). Measurements were made for all trials other than those shaded grey. Trials shaded green are those which have one master block while trials shaded yellow have two master blocks (described in text).

Trait	Age (years)	Site		
		Huntly	Llandovery	Torridge
Height	2	1 2 3	1 2 3	1 2 3
	4	1 2 3	1 2 3	1 2 3
	6	1 2 3	1 2 3	1 2 3
	11	1 2 3	1 2 3	1 2 3
Bud Burst	5A	1 2 3	1 2 3	1 2 3
	5B	1 2 3	1 2 3	1 2 3
	5C	1 2 3	1 2 3	1 2 3
Pilodyn	10	1 2 3	1 2 3	1 2 3

Table 3. Numbers of trees measured for height according to site, age and family. The trials were allocated 6000 trees per family per site prior to planting.

	Huntly			Llandovery			Torridge		
Age (years)	FS1	FS2	FS3	FS1	FS2	FS3	FS1	FS2	FS3
2	5854	5903	5685	5786	5805	5821	5988	5981	5861
4	5698	5861	5414	5220	5288	5180	5987	5875	4219
6	5676	5857	5414	4716	4900	4539	5982	5829	4100
11	5673	5856	5410	4326	4760	3958	5982	5829	4087

Table 4. The number of SNP markers and trees retained within each family (FS1, FS2 or FS3) following quality control, together with the percentage of these SNPs that were found in all 3 families.

	Family		
	FS1	FS2	FS3
SNPs	15,452	17,915	13,176
% present in all FS	13.2	11.5	15.6
Trees	572	470	482

Table 5. The number of informative loci used for imputation according to family (FS1, FS2 and FS3) and linkage group and the group's map length.

Linkage Group	Length	Number of loci		
	(cM)	FS1	FS2	FS3
1	218	159	158	154
2	194	163	163	154
3	201	153	153	141
4	194	146	145	141
5	165	109	109	103
6	174	139	139	138
7	203	149	149	135
8	199	129	129	122
9	164	125	124	112
10	157	129	129	127
11	128	108	107	101
12	146	121	121	117
Total	2143	1630	1626	1545

Table 6. The fraction of random error variance (f_r) and the auto-correlation pooled across columns and rows (ρ) for height measured at four ages at all three sites (see Equation [2]). The values presented have been averaged across all three families with the range given in parentheses.

Age	Huntly		Llandovery		Torridge	
(years)	f_r	ρ	f_r	ρ	f_r	ρ
2	0.85 (0.80,0.92)	0.85 (0.75,0.90)	0.88 (0.87,0.89)	0.77 (0.67,0.86)	0.74 (0.67,0.77)	0.91 (0.79,0.98)
4	0.76 (0.70,0.85)	0.82 (0.72,0.91)	0.70 (0.62,0.76)	0.70 (0.62,0.81)	0.60 (0.43,0.76)	0.90 (0.82,0.96)
6	0.72 (0.63,0.80)	0.84 (0.79,0.91)	0.61 (0.57,0.68)	0.73 (0.64,0.81)	0.54 (0.40,0.72)	0.91 (0.84,0.95)
11	0.52 (0.45,0.58)	0.91 (0.86,0.97)	0.53 (0.49,0.58)	0.81 (0.74,0.87)	0.60 (0.39,0.82)	0.93 (0.88,0.96)

Table 7. The total genetic (σ_u^2) and phenotypic (σ_P^2) variances, broad (H^2) and narrow (h^2) sense heritabilities and the fraction of additive genetic variance (f_a) for pilodyn depth measured at 10 years in all three families at Torridge. The associated s.e.s are given in parentheses.

Family	σ_P^2	σ_u^2	H^2	f_a	h^2
FS 1	2.994 (0.125)	1.046 (0.060)	0.349 (0.019)	0.912 (0.043)	0.319 (0.024)
FS 2	4.673 (0.368)	1.283 (0.073)	0.275 (0.024)	0.751 (0.064)	0.206 (0.026)
FS 3	4.657 (0.657)	0.520 (0.053)	0.112 (0.019)	0.918 (0.199)	0.102 (0.027)

Table 8. The total genetic (σ_u^2) and phenotypic (σ_P^2) variances, broad (H^2) and narrow (h^2) sense heritabilities and the fraction of additive genetic variance (f_a) for the first measurement of bud burst at 5 years in family one at all three sites. The associated s.e.s are given in parentheses.

Site	σ_P^2	σ_u^2	H^2	f_a	h^2
Huntly	0.251 (0.010)	0.085 (0.005)	0.333 (0.019)	0.836 (0.056)	0.278 (0.026)
Llandovery	0.499 (0.021)	0.138 (0.010)	0.276 (0.024)	0.831 (0.069)	0.229 (0.025)
Torridge	0.699 (0.020)	0.326 (0.018)	0.476 (0.019)	0.851 (0.039)	0.405 (0.025)

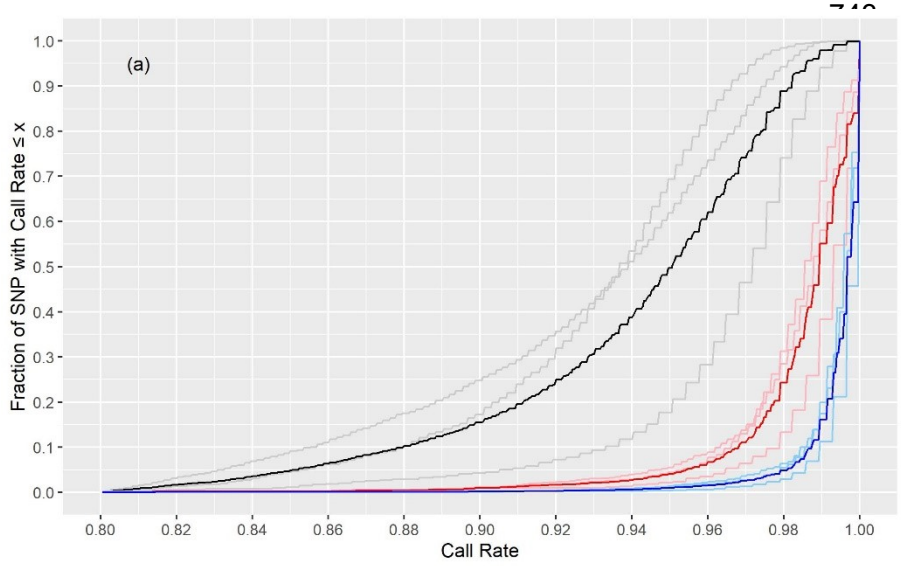
Table 9. The broad sense heritability (H^2) and phenotypic variance (σ_P^2) for height measured at four ages in all three families at all three sites. The associated s.e.s are given in parentheses.

Age (years)	Family	Huntly		Llandovery		Torridge	
		H^2	σ_P^2	H^2	σ_P^2	H^2	σ_P^2
2	FS 1	0.202 (0.014)	0.013 (0.000)	0.101 (0.014)	0.015 (0.001)	0.256 (0.015)	0.019 (0.001)
	FS 2	0.099 (0.018)	0.014 (0.002)	0.043 (0.011)	0.025 (0.001)	0.286 (0.014)	0.012 (0.001)
	FS 3	0.194 (0.014)	0.012 (0.000)	0.101 (0.012)	0.012 (0.001)	0.198 (0.021)	0.009 (0.001)
4	FS 1	0.155 (0.013)	0.087 (0.002)	0.088 (0.012)	0.140 (0.006)	0.239 (0.015)	0.130 (0.004)
	FS 2	0.165 (0.014)	0.075 (0.004)	0.056 (0.012)	0.178 (0.005)	0.178 (0.015)	0.075 (0.004)
	FS 3	0.152 (0.013)	0.097 (0.003)	0.062 (0.011)	0.128 (0.005)	0.063 (0.011)	0.085 (0.007)
6	FS 1	0.149 (0.013)	0.341 (0.010)	0.061 (0.012)	0.590 (0.027)	0.225 (0.012)	0.402 (0.014)
	FS 2	0.174 (0.013)	0.214 (0.005)	0.065 (0.012)	0.544 (0.016)	0.222 (0.015)	0.324 (0.017)
	FS 3	0.155 (0.012)	0.289 (0.009)	0.096 (0.014)	0.357 (0.033)	0.081 (0.016)	0.367 (0.017)
11	FS 1	0.089 (0.012)	1.416 (0.134)	0.056 (0.011)	1.786 (0.055)	0.214 (0.015)	0.779 (0.029)
	FS 2	0.100 (0.028)	1.895 (0.505)	0.062 (0.011)	2.156 (0.136)	0.334 (0.027)	1.282 (0.097)
	FS 3	0.202 (0.014)	1.622 (0.068)	0.076 (0.017)	1.344 (0.254)	0.183 (0.015)	1.470 (0.080)

Table 10. The fraction of additive genetic variance (f_a) for height measured at four ages at all three sites. The values presented are pooled across families using likelihood profiles, and the consensus value is obtained by pooling the resulting profiles across sites. The associated 95% confidence intervals are given in parentheses.

Age (years)	Huntly	Llandovery	Torridge	Consensus
2	0.53 (0.38,0.68)	0.50 (0.25,0.75)	0.60 (0.49,0.71)	0.60 (0.47,0.65)
4	0.60 (0.46,0.74)	0.50 (0.25,0.75)	0.68 (0.56,0.80)	0.60 (0.54,0.72)
6	0.57 (0.49,0.75)	0.60 (0.36,0.90)	0.72 (0.62,0.82)	0.70 (0.61,0.76)
11	0.70 (0.58,0.85)	1.00 (0.68,1.00)	0.80 (0.70,0.90)	0.80 (0.72,0.84)

Figure 1. A summary of imputation success rates for full-sib offspring obtained from AlphaPeel: **(a)** cumulative distribution function for SNP call rate over offspring, and **(b)** cumulative distribution function for offspring call rate over SNPs. These are shown for no imputation (black), with genotypes assigned with probability >0.7 (blue), and genotypes assigned with probability >0.9 (red), and where light colours are for each family and the dark colour is their average.



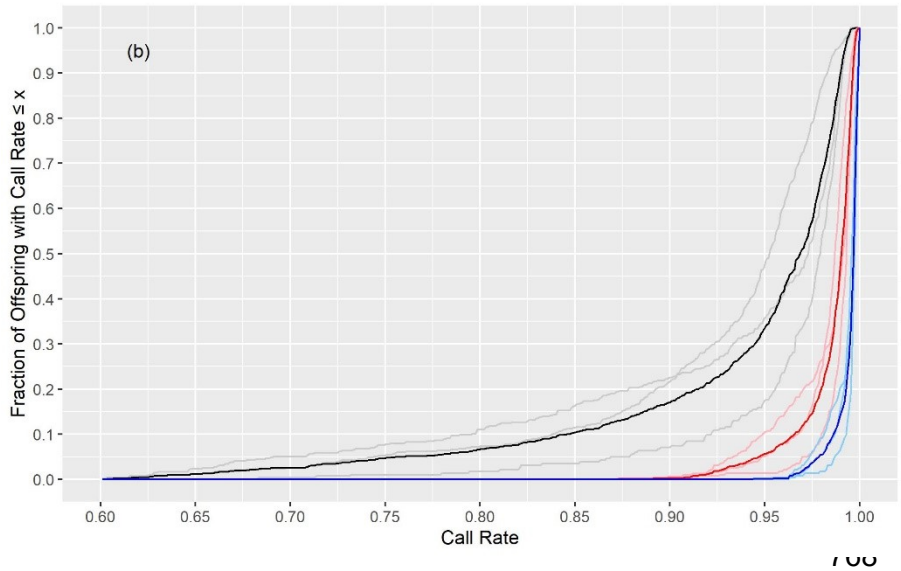


Figure 2. The profiles of -2logL for pilodyn depth measured at 10 years in all three families (grey lines) according to the constrained fraction of additive genetic variance, and the profile pooled across families. Each profile has been adjusted by subtracting its minimum value and therefore the junction with the solid red line y=0 indicates the maximum likelihood estimate, and the interval below the dashed red line y=3.84 indicates the 95% confidence interval.

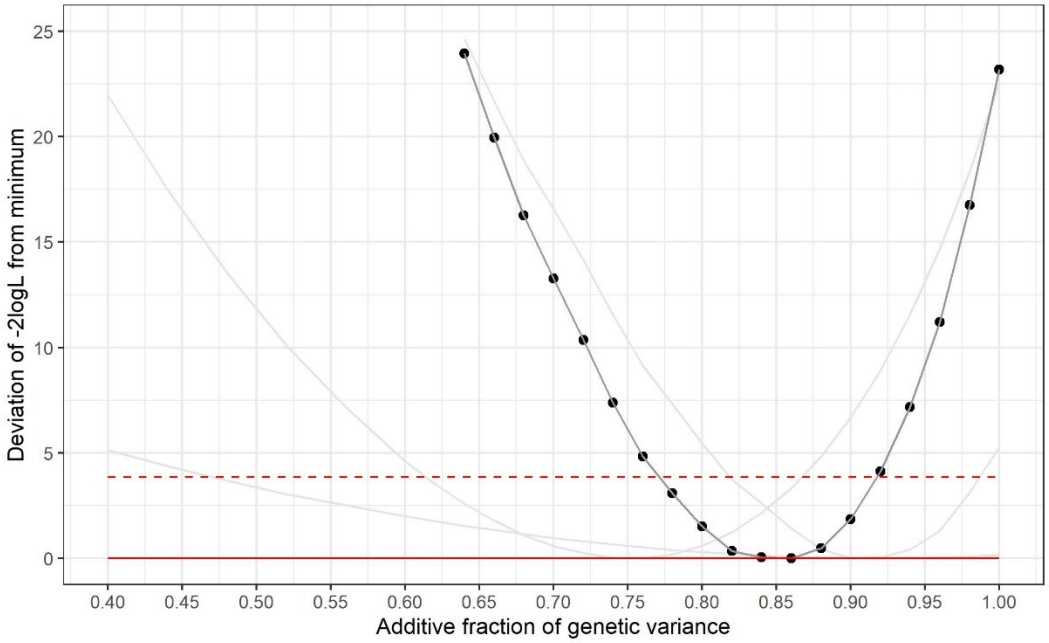
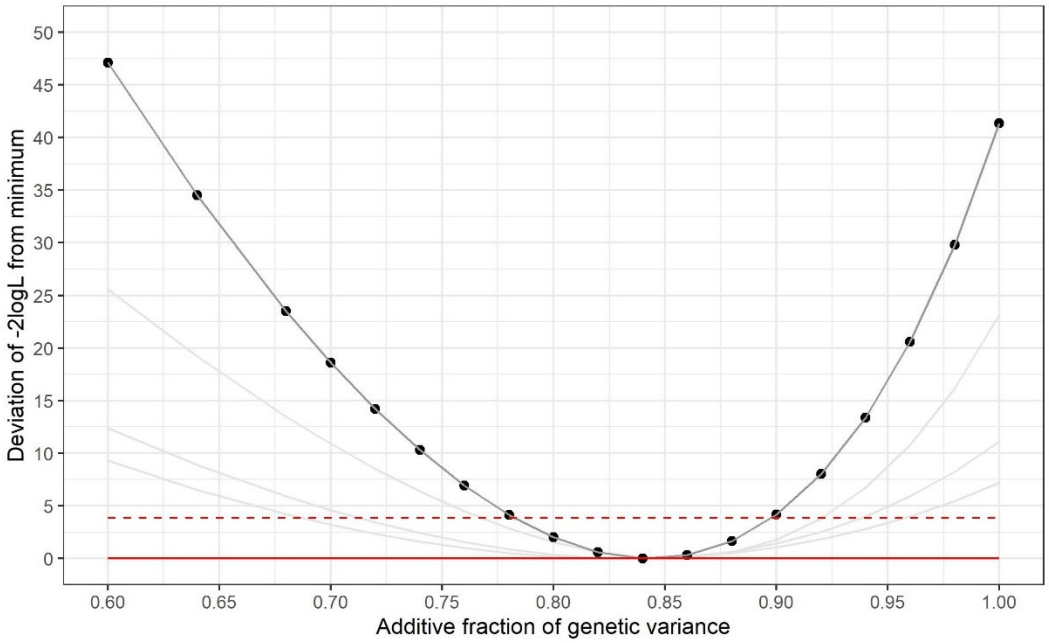


Figure 3. The profiles of -2logL for the first measurement of bud burst at 5 years in family one at all three sites (grey lines) according to the constrained fraction of additive genetic variance, and the profile pooled across sites. Each profile has been adjusted by subtracting its minimum value and therefore the junction with the solid red line y=0 indicates the maximum likelihood estimate, and the interval below the dashed red line y=3.84 indicates the 95% confidence interval.



Supplementary Information 1. Regression on observed heterozygosity.

Methods. Model 2 was modified to include the heterozygosity of the offspring as a linear covariate, and it was fitted separately to all site by family by trait combinations. For tree heights and pilodyn penetration depth the coefficients were pooled across families within sites following Dersimian and Laird (1986). The pooled value for height at each age, and the estimates for bud burst observations in Family 1 were then pooled over sites using DerSimian and Laird (1986).

Results. The regression coefficients are shown in Table S1.1. Since the covariate was heterozygosity positive values represent deleterious inbreeding depression. There was no evidence of heterogeneity when pooling across families for height or pilodyn. Similarly there was no evidence of heterogeneity when pooling across sites. The magnitudes of the consensus estimates rarely exceeded 1 s.e. and were always <1.2 s.e. In summary there was no evidence for an effect of heterozygosity and inbreeding depression in these data.

Table S1.1. The regression coefficient for the fraction of heterozygous marker loci for height, bud burst and pilodyn. Coefficients were pooled across families, and consensus values across sites were calculated following DerSimian and Laird (1986).

		Site							
Trait	Age	Huntly		Llandovery		Torridge		Concensus	
Height	2	-0.01	(0.10)	0.03	(0.12)	0.01	(0.10)	0.01	(0.06)
	4	-0.39	(0.37)	-0.08	(0.46)	0.05	(0.33)	-0.13	(0.22)
	6	-0.72	(0.65)	0.34	(0.67)	-0.44	(0.95)	-0.25	(0.42)
	11	-1.43	(1.01)	0.80	(1.85)	-0.43	(1.03)	-0.71	(0.67)
Bud Burst *	5A	1.52	(0.91)	-1.70	(1.38)	2.56	(1.62)	0.79	(1.18)
	5B	0.39	(1.12)	-0.18	(1.32)	2.75	(2.14)	0.51	(0.79)
	5C	1.19	(0.95)	-0.62	(1.10)	1.51	(1.96)	0.55	(0.68)
Pilodyn	10					2.84	(2.44)		

Reference

DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. Control Clin Tr 7:177-188. http://doi.org/10.1016/0197-2456(86)90046-2

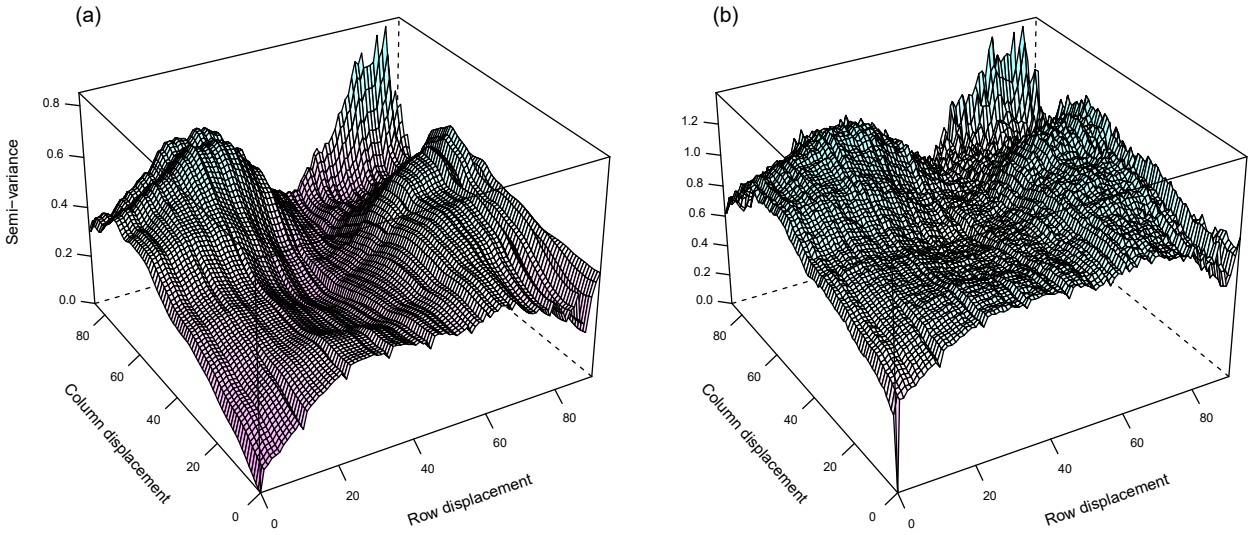
Supplementary Information 2. Example of the sample variograms obtained from fitting the residual spatial model in Model 2.

Fig S2.1 shows the sample variograms obtained for height measured at 11 years in family FS1 at Huntly from fitting Model 2, which includes the residual spatial model in Eqn 2. This example was chosen as the fraction of random error variance was low (f_r = 0.45) and the auto-correlation pooled across columns and rows was high (ρ=0.95). The sample variograms are constructed using the residual semi-variance between plots *x* rows and *y* columns apart, and are hence different to a theoretical variogram constructed directly from the column and row auto-correlation parameters, ρ_c and ρ_r (Gilmour et al. 1997). There are three important features of the sample variograms shown. Firstly, the variograms have properties determined by the parameters: Fig. 2.1 (a) peaks at ~0.7 which reflects the spatial error variance (σ_s^2); and Fig. 2.1 (b) peaks at ~1.3 which reflects the total error variance ($\sigma_r^2 + \sigma_s^2$). The discontinuity at zero displacement in (b) is ~0.6 which reflects the random error variance (σ_r^2). Secondly, both variograms demonstrate a substantial increase in semi-variance between plots either 60 columns or 60 rows apart. Random column and row terms were fitted to account for this variation, but these proved unsuccessful. Lastly, there is a noticeable decrease in semi-variance between plots 11, 22, 33, … rows apart. Unfortunately, the source of this variation could not be identified. In models the standard random terms with effects for rows 1 to the number of rows were supplemented with terms cycled through 1 to 11 over the rows, but this proved unsuccessful.

References

Gilmour AR, Cullis BR, Verbyla AP (1997) Accounting for natural and extraneous eariation in the analysis of field experiments. J Agric Biol Environ Stat 2:269-293. https://doi.org/10.2307/1400446

838 **Figure S2.1** Sample variograms for height measured at 11 years in FS1 at Huntly. Plot (a) shows
839 the variogram for the correlated spatial error and (b) shows the variogram for the correlated spatial
840 error plus random error. The z-axis shows the residual semi-variance between pairs of plots *x* rows
841 and *y* columns apart. Only semi-variances based on more than 50 pairs are shown.
842



843
844

Supplementary Information 3. Results for bud burst observations 5B and 5C.

The 5B and 5C measurements of bud burst in FS1 were subject to the same analyses as 5A, and the results for 5A are presented in the main text. The Tables S3.1 and S3.2 show the results for 5B and 5C corresponding to Table 8 for 5A. The consensus estimates across sites, obtained using Dersiminian & Laird (1986), were 0.909 (s.e. 0.027) and 0.891 (s.e. 0.037) for 5B and 5C respectively.

Table S3.1 Ortet (σ_u^2) and phenotypic (σ_P^2) variances, the broad (H^2) and narrow (h^2) heritabilities and the fraction of additive genetic variance (f_A) for measurement 5B of bud burst in FS1. The associated s.e.s are given in parentheses.

Site	σ_P^2	σ_u^2	H^2	f_A	h^2
Huntly	0.475 (0.029)	0.077 (0.006)	0.163 (0.015)	0.867 (0.092)	0.141 (0.020)
Llandovery	0.499 (0.019)	0.142 (0.010)	0.285 (0.018)	0.939 (0.063)	0.268 (0.024)
Torridge	1.288 (0.044)	0.628 (0.034)	0.488 (0.018)	0.907 (0.032)	0.442 (0.024)

Table S3.2 Ortet (σ_u^2) and phenotypic (σ_P^2) variances, the broad (H^2) and narrow (h^2) heritabilities and the fraction of additive genetic variance (f_A) for measurement 5C of bud burst in FS1. The associated s.e.s are given in parentheses.

Site	σ_P^2	σ_u^2	H^2	f_A	h^2
Huntly	0.385 (0.025)	0.050 (0.005)	0.129 (0.014)	0.901 (0.116)	0.117 (0.019)
Llandovery	0.381 (0.009)	0.061 (0.006)	0.159 (0.015)	0.846 (0.116)	0.135 (0.021)
Torridge	1.267 (0.190)	0.408 (0.023)	0.322 (0.050)	0.895 (0.041)	0.289 (0.047)

Reference

DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. Control Clin Tr 7:177-188. http://doi.org/10.1016/0197-2456(86)90046-2