Genomic models for accurate estimation of breeding values in Sitka spruce (Picea sitchensis Bong. Carr)

Gregor Gorjanc

Gregor.Gorjanc@roslin.ed.ac.uk

University of Edinburgh https://orcid.org/0000-0001-8008-2787

John Woolliams

Mateja Janes

University of Edinburgh

Hayley Thumas

University of Oxford

Joanna Ilska

Roslin Institute, University of Edinburgh

John McLean

Forest Research

Joan Cottrell

Forest Research

John Mackay

University of Oxford

Article

Keywords:

Posted Date: July 9th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7076000/v1

Additional Declarations: There is no duality of interest

Genomic models for accurate estimation of breeding values in Sitka spruce (*Picea sitchensis* Bong. Carr)

J. A Woolliams[1], M. Janes[1†], H. Tumas[2††], J. J. Ilska[1†††], J. P. McLean[3††††], J. Cottrell[3], J. Mackay[2], and G. Gorjanc[1]

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. Northern Research Station, Forest Research, Roslin, Midlothian, EH25 9SY, UK.

Current addresses:

† Ruđer Bošković Institute. Bijenička cesta 54, 10000 Zagreb, Croatia.

†† Department of Forest and Conservation Sciences, University of British Columbia, 2424 Main Mall, Vancouver, BC Canada V6T 1Z4.

††† The Kennel Club, 10 Clarges St, London, W1J 8AB, UK.

†††† Norwegian Institute of Bioeconomy Research (NIBIO), Postboks 115, 1431 Ås, Norway.

Corresponding author: gregor.gorjanc@roslin.ed.ac.uk

ORCID IDs

JAW: 0000-0002-7880-9972

MJ: 0000-0003-3933-2056

HT: 0000-0003-1083-1687

JI: 0000-0002-1613-333X

JC: 0000-0001-6355-1326

JM: 0000-0002-4883-195X

GG: 0000-0001-8008-2787

Abstract

Genomics is advancing estimation of breeding values for selective breeding. Traditional pedigree information used in the estimation of breeding values can still be competitive to genomic information when markers do not adequately capture genetic variance or genomic datasets are not sufficiently large. Using the UK Sitka spruce breeding population as a study system, we assessed the accuracy of breeding values using pedigree or genomic information. Nine phenotypic traits associated with growth and wood quality were measured across 1295 individuals in 50 full-sib families replicated across two plantation sites. Pedigree-based estimation was compared to genome-based estimation using identity by state (IBS, across families) or identity by descent (IBD, within families) information from 4123 SNPs. Accuracy of breeding values was assessed using prediction standard errors obtained from the fitted model, which assumes the model and parameters are correct, and a less biased 12-fold cross-validation by withholding phenotypes from the model fitting. Accuracies obtained directly from the model were very similar across different types of information, but the more objective cross-validation assessment showed genomic information gave the best predictions for all but one of the traits. No bias was observed for any of the trait- information combinations. Results demonstrate that the predictive performance of genomic information is already competitive with the pedigree information, despite evidence that the marker panel did not capture the full genetic variance. Furthermore, accuracies with genomic information have the potential to significantly improve, unlike pedigree, by increases in the size of the training set and marker density.

Introduction

A major theme in the development of models for the genetic evaluation of forest tree species in the last decade has been the incorporation of genomic technology and a reduction in the reliance on pedigree. This offers the potential advantages of increasing selection intensity by obtaining genetic predictions from a wider pool of candidates, more accurate selection-particularly within families, and reducing the length of generation intervals, all beneficial to generating progress to selection goals in tree breeding programmes. The development of genomic prediction in commercially important trees has included several species of pine (*Pinus* spp.; e.g., Freeman et al. 2022; Walker et al. 2022), and spruce (*Picea* spp.; e.g., Beaulieu et al. 2014; Chen et al. 2018) as well as Japanese cedar (*Cryptomeria japonica*: Ejima et al. 2023), and Douglas-fir (*Pseudostuga menziesii*; Thistlethwaite et al. 2017). In hardwood trees, several reports have concerned Eucalyptus and its hybrids (e.g., Müller et al., 2017; Duarte 2024). These genomic predictions have targeted phenotypic traits such as growth rate and phenology, wood structure and strength, pulp yield, and stress responses such as insect resistance traits (Grattapaglia, 2022; Beaulieu et al. 2024). In these systems, identifying the factors that determine the accuracy of genomic predictions is a major question of interest with population features appearing as crucial (e.g., effective population size, relatedness among individuals, and size of the training set) along with the genetic architecture of the traits targeted (Isik, 2022; Grattapaglia, 2022; Ilska et al. 2023; Beaulieu et al. 2024).

The historical use of pedigree in BLUP evaluations relies on the concept of identity by descent (IBD), using the probabilities of IBD traced back to a base population to derive numerator relationships among the individuals in the pedigree, as these relationships describe the (co)variances of the population's breeding values. This has provided a robust model for predicting breeding values for traits under the control of many loci. In contrast, the routine use of genomics for evaluations has followed the principles of (Meuwissen et al., 2001) in using identity by state (IBS) to construct relationship matrices (VanRaden, 2008) without the need to refer to a pedigree and, in the absence of whole genome sequence, relying on population-wide linkage disequilibrium (LD) of causative variants with dense markers distributed over the whole genome. Considerable genome resources are now available to enable genomic predictions in forest trees as recently reviewed (Beaulieu et al.

2024). In conifers, whole genome sequencing initially gave highly fragmented assemblies due to the large size (often >10 Gbp) and the relative complexity of their nuclear genomes (e.g., Nystedt et al. 2013; Birol et al. 2013; Zimin et al. 2014). More recently, long-read sequencing and methods such as proximity ligation sequencing (e.g., Hi-C) have greatly improved the contiguity of genome assemblies (Scott et al. 2020; Xiong et al. 2021; Niu et al. 2022). In some species, high-density genetic maps are available to probe genome architecture when such assemblies are lacking (Liu et al. 2019; Bernhardsson et al. 2019; Tumas et al. 2024). Large-scale DNA sequencing and population surveys have enabled the development of genotyping methods in conifers, from custom chips (Pavy et al. 2013; Plomion et al. 2015) to targeted sequencing (Neves et al. 2014; Bernhardsson et al. 2019) and reduced representation whole-genome sequencing (e.g., restriction site associated DNA sequencing (Gamal El-Dien et al. 2015).

The approaches of pedigree and IBD on the one hand, and genomics and IBS on the other, use different underlying models to predict breeding values, and whilst in some breeding sectors, e.g., poultry, benefits have been established for routinely using genomics, this has yet to be realised in many tree breeding programmes (Lebedev et al. 2020; Beaulieu et al. 2024), particularly in conifers with characteristically large and complex genomes. Included among the tangible economic advantages that are regularly argued for genomic predictions are shortening the breeding cycle by several years and reducing the cost of selection for traits which are costly to measure such as wood strength (Isik, 2022; Grattapaglia, 2022; Beaulieu et al. 2024). Such potential benefits have been widely discussed and explored through deterministic and stochastic simulations (e.g., Papin et al. 2024) but as discussed by Grattapaglia et al. (2022), there are multiple reasons for a long lag-time for development of genomic prediction in trees. These include the lack of the large training sets of genotyped and phenotyped individuals required for genomic models (most tree studies had training sets <1500; Beaulieu et al. 2024), and the relatively low density of markers used (80% of studies used <20,000 markers; Beaulieu et al. 2024) to capture the LD compared to those managed species where genomics is used routinely.

105 The use of genomics does not need to rely on IBS to provide information. Genomic relationships of
106 genotyped individuals can be embedded within a larger pedigree population in single-step
107 evaluations resulting in EBVs that have used both IBD and IBS information (Legarra et al. 2009;
108 Misztal et al. 2009; Aguilar et al. 2010; Christensen and Lund, 2010) and this approach has been
109 widely applied in breeding populations, e.g., in *Eucalyptus dunnii* (Jurcic et al. 2021) and *Eucalyptus*
110 *grandis* (Duarte et al, 2024). A different idea is to use the dense markers to track IBD inheritance of
111 individual genome segments from offspring to parent recursively back to a pedigree base. In doing
112 so, averaging over the segments distinguishes full-sibs, which would otherwise be regarded as
113 having identical relationships and these differences are carried through when considering the
114 relationships among their offspring (i.e., the sets of cousins). Individuals that are unrelated in the
115 pedigree relative to the assumed base remain unrelated. Using this linkage analysis approach, a
116 genomic relationship matrix based on IBD can be calculated, which is a genome wide extension to
117 the single-locus relationship matrix of Fernando and Grossman (1989). Luan et al. (2012) showed
118 that in dairy cattle after four generations this approach gave approximately equivalent accuracies
119 when predicting breeding values compared to those based on IBS using VanRaden matrices. Jucjic
120 et al. (2021) considered this use of markers to track IBD in their study of *Eucalyptus dunnii.*
121
122 The linkage analysis approach may offer some advantages particular to conifer breeding. Firstly,
123 conifer breeding has the potential to generate large sib families for all parents. This in turn offers the
124 possibility of estimating model parameters and training predictions from LD within each family, while
125 using the family means to differentiate between families, as considered by (Lillehammer et al., 2013)
126 in aquaculture. Secondly, as recombination is a relatively rare event along a chromosome, the
127 density of markers required to localise a recombination in a parent's gamete need not be large and
128 so the relatively low marker density across large conifer genomes is potentially less important.
129 Therefore, this study investigates a Sitka spruce (*Picea sitchensis* Bong. Carr) breeding population
130 used in productive forestry in the UK (Lee et al. 2013) in which additive and non-additive effects in
131 growth and wood phenotypic traits were recently dissected using two large full-sib families (Fuentes-
132 Utrilla et al. 2017; Ilska et al. 2023). Here, we compare the accuracy of prediction obtained for a
133 range of timber quality and growth traits in a much wider sample of the UK Sitka breeding population

over one generation, using both IBD and IBS. The accuracy of prediction is considered both from assuming the model and its parameters are correct, which are obtained from the prediction standard errors and by using cross-validation where no such assumptions are made.

Materials & Methods

Study design

The study population consisted of controlled crosses of parent trees in the UK elite population (EP) and open pollinated controls from a Sitka spruce (*Picea sitchensis* Bong. Carr) provenance originating on Haida Gwaii (HG). Two forest experiments containing the study populations were destructively sampled in 2017. The individual trees were a sample from two larger, genetically linked experimental trials (A and B) established in 1996/97 with each trial replicated at two sites: Brecon (Lat: 51.89º, Long: -3.33º, Alt. 375 m a.s.l., Wales, UK) and at Kintyre (Lat: 55.36 º, Long:-5.64 º, Alt. 60 m a.s.l., Scotland, UK). Trials A and B involved 56 and 52 distinct full-sib (FS) families respectively from EP, bred from 59 parents (42 sires, 52 dams, and 35 parents used as both sire and dam) and planted in 30 blocks/trial following an incomplete block design with an HG individual as a control in each block.

Nine to 16 trees were sampled from each of 25 full-sib EP families from each trial at each site for a total of 1295 trees and an HG control was sampled from 16-20 blocks from each trial at each site for a total of 73 trees (Table 1). EP families were selected for connectedness of parentage, with the selected families bred from 45 distinct parents with several families sharing a parent (38 sires, 36 dams, and 29 parents were used as both sire and dam). The blocks were incomplete replications due to fractional replication design of the trials, hence the families were sampled unevenly and 36 (Kintyre) to 38 (Brecon) blocks were required to be sampled at each site. Each block contained an individual from some of the full-sib families in addition to other individuals not sampled for this study, and each block contained an HG control individual. Table 1 shows the resulting number of trees sampled each site and in total alongside family and block information.

Table 1. The numbers of experimental trees, full-sib (FS) families and blocks represented classified by elite population (EP) or Haida Gwaii (HG), and Brecon or Kintyre site.

	Brecon		Kintyre		Total
Trial	A	B	A	B	
Number of individuals					
- total	338	344	342	344	1368
- from EP	318	326	323	328	1295
- from HG (Controls)	20	18	19	16	73

EP FS families					
- number represented	25	25	25	25	50
- minimum family size	9	13	11	13	20
- maximum family size	14	14	16	14	27
Blocks					
- number represented	20	18	19	17	74
- with EP individuals	14	14	19	17	64
- with HG individuals	20	18	19	16	73

Measurements

The measurements used in the model testing included both juvenile traits measured before the current study and traits measured on the felled trees as part of the study.

Juvenile Indicator Traits. All trees had their height measured at 6 years of age (H_{06}). At 11 years of age the following traits were measured: diameter at breast height (D_{11}), branching score (B_{11}), stem straightness score (S_{11}) and, as a proxy for wood density, the penetration depth of a pilodyn pin at breast height (P_{11}) using a Pilodyn 6J Forest device (Proceq, Switzerland). The scores for branching and stem straightness were obtained by visual appraisal using a 6-point scale where '1' equates to the best and '6' the worst, as described by Cameron et al. (2012).

Merchantable Volume at Felling. Merchantable volume (V_{23}) was the portion of felled tree stem that would be typically used for forest products, defined as the volume for the length of the stem >7cm in diameter. Material with lower diameter is not marketed. V_{23} was estimated by measuring stem diameter at 1m intervals using girthing callipers. To account for oval stems, two measurements of diameter were taken at each assessment position and averaged, and halved to give the radius of section *i*, expressed in m. The volume was calculated for each of these *n* sections by treating them as 1m high (*h*) frustums of cones, unless the 7cm diameter point was intermediate when the height of the last section (h_n) was obtained by linear interpolation. V_{23} was then calculated as:

$$V_{23} = \sum_{i=0,1,2\ldots n}^{n} \frac{1}{3} \cdot \pi \cdot h_i \cdot \left(r_{i+1}^2 + r_{i+1} \cdot r_i + r_i^2\right) \quad \text{[Eqn 1]}$$

Cross-Sectional Averages of flexural strength*,* flexural stiffness, *and wood density at Felling.* Cross-sectional averages (CSA) of flexural strength (MOR_{23} or modulus of rupture), flexural stiffness (MOE_{23} or modulus of elasticity), and wood density (DEN_{23}) are area-weighted averages of the three

properties that are used in the European grading system of softwoods (CSN 338:2016). CSA provide a means of accounting for variation within trees (Zhou et al., 2020). Small clear specimens (SCS, British Standards Institute, 1957), free of knots and other defects that may influence the measured properties, measuring 20x20x300mm were produced from 2m long logs with their base approximately 1.5m from the ground. First a board spanning the diameter of each log was cut out using a bandsaw. The board was examined to find sections free from knots and where the wood grain was perpendicular to the primary axis, and multiple SCS were cut per section using a circular saw and the surfaces smoothed with a planer-thicknesser. The pattern of SCS was centred on the pith (sample 1) and followed two opposing radii so sampling the rings in the section in pairs either side of the pith. The number of samples obtained from a section varied depending on the diameter of the tree and ranged from 3 to 13. In some sections with greater asymmetry about the pith, the number of samples along one radius was one greater than along the other. SCS were stabilised in a climate-controlled environment to nominally 12% moisture content. Following stabilisation, the sample mass and dimensions were recorded and each sample was tested in a three-point flexural bending test using a Tinius Olsen HK5-S (Redhill, UK) with a flexural testing rig for MOE_{23}, and MOR_{23} of the individual SCS. DEN_{23} was calculated as bulk density from SCS mass and volume at 12% moisture content. Finally, the CSA for each of these three traits for each individual was calculated as a weighted average of the SCS by assuming the series of SCS from an individual were representative of a concentric series of rings whose areas could be calculated as annuli and expressed as relative to the area of a solid circle with a radius that was the product of SCS dimensions in the radial dimension (20mm) and the total number of SCS produced in a radius. Note the first sample was considered to have a radial dimension of 10mm as it was centred on the midpoint. These area weighted values for each trait could then be summed to produce the CSA. As each induvial had nominally two radii of SCS, the arithmetical mean of both radii was used as the CSA.Using this system, it follows that weights for outer SCS were greater than those for inner CSA reflecting a greater area fraction of the CSA. The weights are given in Supplementary Information 1.

Genotyping and imputation

DNA was extracted from all felled trees and the 45 parent trees for genotyping with a SNP chip containing 12911 markers (Tumas et al., 2024). Following chip manufacture (Illumina, San Diego, CA, USA), genotyping was conducted by the Centre d'expertise et de services Génome Québec (Montréal, QC, Canada). Genotype calling was performed using the GenomeStudio v2.0.5 software (Illumina, San Diego ,CA, USA), and genotype clusters were visually examined and manually curated when necessary to reject monomorphic and failed polymorphisms. Genotypes were obtained for 1341 (98%) of the felled trees and the 45 parents. Excel files obtained from GenomeStudio were formatted for PLINK v1.90b4 (Purcell et al. 2007) using R v4.0.2 (R Core Team 2022). The SNP data were then reduced to 6062 SNP that were informative in the experimental population (EP+HG) and had a call rate across the genotyped individuals >0.88. Using these SNPs the pedigree and genomic data were checked for consistency by comparing relationships obtained from pedigree with the VanRaden Model 1 relationships obtained from SNP genotypes (VanRaden, 2008). Genotypes for two of the 45 recorded parents were inconsistent with all of their genotyped offspring, and two dummy parents were introduced to the pedigree - this affected four FS families. In a further family of nominal FS, one of the recorded parents had only 6 offspring, whereas the remaining 20 members had a further ungenotyped parent, which was assigned a third dummy identity.

To obtain more complete information for the EP group, missing genotypes were imputed using the 21.42 Morgan linkage map for Sitka spruce with 12 linkage groups (average length per group of 1.78 Morgans) developed by Tumas et al. (2024). To use this map only a subset of 4123 SNPs on the chip previously assigned to the 12 linkage groups could be used (n_{loci}=4123). Imputation was carried out by multi-locus peeling (Whalen et al. 2018) using AlphaPeel (https://github.com/AlphaGenes/AlphaPeel). The impact of imputation was assessed following Ilska et al. (2023). Prior to imputation, 84% of trees had a call rate > 95% and 75% of SNPs had a call rate > 95%, after imputation if assignment was judged by posterior probability for a genotype > 0.90, then 96% of trees had a call rate >95%, and 91% of SNPs had a call rate >95%. A second assessment, estimated the overall probability of error in a genotype after imputation to be 0.022. Further description of these methods is in Supplementary Information 2.

Statistical methods

Models I, II, and III. Three relationship matrices were constructed and the accuracy of the EBVs for the 9 traits described above obtained from using each relationship matrix was assessed from (i) the model-based standard errors or (ii) from cross validation. Mixed linear models were fitted to the data using ASReml-R (Butler et al. 2017). The general model fitted was a bivariate model to encompass genotype by environment interactions between Brecon (*j*=1) and Kintyre (*j*=2) sites:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wc} + \mathbf{e} \qquad \text{[Eqn 2]}$$

where **y** denotes a vector of trait measurements for individuals ordered by site *j*=1,2. The term **b** denotes a vector of fixed effects for trait means at each site (2 d.f.), trial within site (2 d.f.), and the genetic group effect between HG controls and the experimental population at each site (2 d.f.), with **X** their design matrix relating observations to effects. The term **a** denotes a vector of breeding values for each individual represented in the relationship matrix for each site, assumed randomly distributed MVN(**0**, $\mathbf{V}_G \otimes \mathbf{G}$), where $\mathbf{V}_G$ is a 2x2 matrix of genetic (co)variances for sites *j*=1,2, denoted $v_{g,1}$, $v_{g,2}$ and $r_g\sqrt{v_{g,1}v_{g,2}}$ where r_g is the genetic correlation, and **G** is a genetic relationship matrix (described below), and **Z** the design matrix. The term **c** denotes a vector of block effects sorted by sites *j*=1,2, assumed random and distributed MVN($\mathbf{0}, v_{c,1}\mathbf{I}_1 \oplus v_{c,2}\mathbf{I}_2$), where $\mathbf{I}_j$ is an identity matrix of size appropriate to the number of blocks at site *j*, and **W** is their design matrix. The residual **e** was assumed random and distributed MVN($\mathbf{0}, v_{e,1}\mathbf{I}_1 \oplus v_{e,2}\mathbf{I}_2$), where $\mathbf{I}_j$ is an identity matrix of size appropriate to the number of individuals at site *j*. The relationship matrix **G** incorporated into this model was one of 3 giving rise to Models I, II, and III, calculated as described below with further theoretical details given in Supplementary Information 3.

Models I and II use well-established and widely used methods. Model I incorporated Wright's pedigree (numerator) relationship matrix (**A**), which is based upon expected IBD (Wright, 1922; Henderson, 1976). The inverse of **A** was calculated within ASReml-R from the pedigree described above. Model II incorporated VanRaden's genome relationship matrix ($\mathbf{G}_{VR1}$), which is based upon IBS among markers (VanRaden, 2008). The matrix $\mathbf{G}_{VR1}$ was calculated using the imputed gene

dosages obtained from AlphaPeel, using the observed allele frequencies to centre the dosages and calculate the scaling factor involved in its calculation.

Model III incorporated the whole-genome version of the marker relationship matrix ($\mathbf{G}_{FG}$) based upon realised IBD (Fernando & Grossman, 1989; Meuwissen et al. 2011). $\mathbf{G}_{FG}$ was calculated from the '.seg' files produced by AlphaPeel that contain the probabilities (q) that individual i's alleles at locus k (k=1,…,n_{loci}) originate from the (i) paternal grand-sire and maternal grand-sire ($q_{i,k}^{pp}$), (ii) paternal grand-sire & maternal grand-dam ($q_{i,k}^{pm}$), (iii) paternal grand-dam & maternal grand-sire ($q_{i,k}^{mp}$), or (iv) paternal grand-dam & maternal grand-dam ($q_{i,k}^{mm}$). With *s(i)* and *d(i)* denoting the sire and dam of i, the probability of inheriting a specific parental allele at locus k can be calculated; the probabilities of having inherited the sire's paternal allele is $q_{s(i),k}^{p} = q_{i,k}^{pp} + q_{i,k}^{pm}$ or its maternal allele is $q_{s(i)k}^{m} = q_{ik}^{mp} + q_{ik}^{mm}$; similarly probabilities of the dam's paternal or maternal allele are $q_{d(i)k}^{p} = q_{ik}^{pp} + q_{ik}^{mp}$ and $q_{d(i)k}^{m} = q_{ik}^{pm} + q_{ik}^{mm}$. When the set of parents form the base generation, the probability of IBD in two offspring i and j depend solely on receiving the same gamete from a common parent, and the coefficient of kinship is $\phi_{ij} = \frac{1}{4}\sum_{k=1}^{n_{loci}} \sum_{u \varepsilon \{s(i),d(i)\}} \sum_{v \varepsilon \{s(j),d(j)\}} (q_{uk}^{p} q_{vk}^{p} + q_{uk}^{m} q_{vk}^{m})\delta_{u,v} / n_{loci}$ where $\delta_{uv} = 1$ if u and v are identical or 0 otherwise. The (i,j) element of $\mathbf{G}_{FG}$ is then $\mathbf{G}_{FG}[i, j] = 2\phi_{ij}$, analogous to the correspondence between Wright's numerator relationship and the coefficient of kinship.

Model IIIA. A further model was fitted as a modification to Model III, analogous to the idea of within-family selection of (Lillehammer et al., 2013), which argues that the family means (the parental average breeding value) are well estimated without genomic information when sib-families are large, and that estimating the genomic marker effects within families might be done with a smaller training set (see Discussion). This was implemented by constructing the genomic matrix of only Mendelian sampling terms based on IBD, $\mathbf{M}_{FG}$. Consider the partition of $\mathbf{G}_{FG}$ into the parents (P) and offspring (O) so $\mathbf{G}_{FG}$ = ($\mathbf{G}_{FG}$[P,P] $\mathbf{G}_{FG}$[P,O] | $\mathbf{G}_{FG}$[O,P] $\mathbf{G}_{FG}$[O,O]) then $\mathbf{M}_{FG}$ can be calculated by conditioning on the parents so that:

$$\mathbf{M}_{FG} = \mathbf{G}_{FG}[O,O] - \mathbf{G}_{FG}[O,P]\ \mathbf{G}_{FG}[P,P]^{-1}\ \mathbf{G}_{FG}[P,O], \quad \text{[Eqn 3]}$$

where $\mathbf{G}_{FG}[P,P]$ is the identity matrix, $\mathbf{G}_{FG}[P,O] = \mathbf{G}_{FG}[O,P]^T$, and the *(i,j)* element of $\mathbf{G}_{FG}[P,O]$ is 0.5 if *i* is a parent of *j* and 0 otherwise (there was no selfing). Therefore, in this population, the (*i,j*) element of $\mathbf{M}_{FG}[i,j] = 0.5$ if *i=j*, and $\mathbf{M}_{FG}[i,j] = \mathbf{G}_{FG}[i,j] - \mathbf{A}[i,j]$ otherwise. Therefore, the Model IIIA was:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_1\mathbf{t} + \mathbf{Z}_2\mathbf{m} + \mathbf{Wc} + \mathbf{e}\,, \quad \text{[Eqn 4]}$$

where **t** is a vector of transmitting abilities of parents and $\mathbf{Z}_1$ the design matrix relating the parents to offspring assumed distributed MVN(**0**, $\mathbf{V}_T \otimes \mathbf{I}$), **m** a vector of Mendelian sampling terms assumed distributed MVN(**0**, $\mathbf{V}_M \otimes \mathbf{M}_{FG}$), and $\mathbf{V}_T$ and $\mathbf{V}_M$ are 2x2 (co)variance matrices for the two sites fitted in the bivariate model. The EBV of offspring *i* is then given by $\hat{g}_i = \hat{t}_{s(i)} + \hat{t}_{d(i)} + \hat{m}_i$.

Accuracies. To obtain the accuracy of EBVs, assuming the model and the estimated parameters were correct, the model was fitted to the complete set of data to obtain the square root of prediction error variance (S.E.) for individual i at site j (s_{ij}) and the accuracy of this EBV (r_{ij}) was calculated as $r_{ij} = \sqrt{1-(s_{ij}^2 / \hat{v}_{g,j} g_{ii})}$, where g_{ii} is the element on the leading diagonal of the relevant **G** for individual *i*. This was done for Models I, II and III, but not IIIA.

The second assessment of accuracy was from 12-fold cross-validation carried out for all models. Accuracy was assessed within blocks, within sites. For this purpose, the blocks within each site were randomly split into 12 sets (*s*=1,…,12), with the only restriction applied that at least 100 EP trees across both sites were represented in each set. The distribution of elite and control trees and blocks across sets and sites is shown in Table 2. The models were fitted in 12 runs and in each run the records from the individuals of the EP (not the HG controls) in set *k* were removed to form a validation set and their EBVs after fitting the model were saved. After all the runs the EBVs obtained for the EP were compared to their phenotypes within blocks, i.e., block by block. For each block, the accuracy of EBV was estimated as the correlation of the EBV and the phenotype ($r_{EBV,P}$) and the bias of EBV was estimated as the slope of the regression of phenotype on EBV ($b_{EBV,P}$). The estimates were pooled over blocks, using Fisher's z-transformation for $r_{EBV,P}$. Pooling followed the meta-analysis of (DerSimonian & Laird, 1986), whereby a weighted mean is estimated from the individual

block estimates, with weights given the reciprocal of the sampling variance, i.e., 1/ (S.E.)^2, and an estimate of variation over blocks beyond sampling error is calculated, together with a χ^2 value quantifying the evidence for heterogeneity, and finally the weights and weighted mean are re-calculated accounting for any heterogeneity observed. The degrees of freedom for χ^2 were 37 for Brecon and 35 for Kintyre.

Table 2. The numbers of elite and control trees and blocks classified by set and site in the randomised validation trial.

	Set												
	1	2	3	4	5	6	7	8	9	10	11	12	Sum
n elite													
Brecon	72	63	52	47	45	52	53	47	55	50	46	50	632
Kintyre	49	41	52	54	55	54	51	54	54	55	60	60	639
n controls													
Brecon	4	3	3	3	3	3	3	3	3	2	3	3	36
Kintyre	3	3	2	3	3	3	3	2	3	3	3	3	34
n blocks													
Brecon	4	4	3	3	3	3	3	3	3	3	3	3	38
Kintyre	3	3	3	3	3	3	3	3	3	3	3	3	36

In addition to the estimates of accuracy and bias, the estimates of heritability within the EP were calculated by scaling $\hat{v}_{g1}$ and $\hat{v}_{g2}$ by $k_G = \overline{diag(\mathbf{G}[e])} - \overline{\mathbf{G}[e]}$ where $\mathbf{G}[e]$ is sub-matrix the relationship matrix used to generate the estimates corresponding to the EP individuals, and bar denotes the average value. Therefore, for site j $h_j^2 = k_G \hat{v}_{gj} / (k_G \hat{v}_{gj} + \hat{v}_{cj} + \hat{v}_{ej})$. Scaling of $\hat{v}_c$ and $\hat{v}_e$ was not carried out as the scaling factors were >0.98, and did not differ among the models.

Results

Genetic parameters. The log-likelihoods (logL) and genetic correlations are shown in Table 3, and the heritabilities obtained for each site with each relationship matrix are shown in Table 4. For all traits, excepting branching B_{11}, the estimated r_g values were very large and positive. Supplementary Information 4 considers the evidence for genotype by environment (GxE) interaction in more detail, testing the deviation of r_g from 1, and the magnitudes of the genetic variance at the two sites. Phenotypic variances are given in Supplementary Information 5. The best models for each trait, as judged by logL, were spread across all three relationship matrices but **A** (expected IBD information according to pedigree) only provided the best model for straightness S_{11}, whereas $\mathbf{G}_{VR1}$ (IBS information according to genomic markers) and $\mathbf{G}_{FG}$ (realised IBD information according to genomic markers within the pedigree) each provided the best model for 4 of the other traits. Summing logL over all traits, and comparing to **A**, the total logL for $\mathbf{G}_{VR1}$ and $\mathbf{G}_{FG}$ were +18.81, and +32.01 respectively, i.e., greatest for $\mathbf{G}_{FG}$.

Table 3. The log-likelihood (logL) for the full model and the estimated r_g for the 9 traits using the three relationship matrices based on pedigree (**A**), genomic IBS ($\mathbf{G}_{VR1}$), and genomic IBD ($\mathbf{G}_{FG}$) information. The greatest (favourable) logL values for each trait are shaded grey. The S.E.'s for r_g are shown in parentheses; values shown as 1 with no S.E. were bound in the model fitting.

	A		$\mathbf{G}_{VR1}$		$\mathbf{G}_{FG}$	
Trait	r_g	logL	r_g	logL	r_g	logL
H_{06}	0.979 (0.144)	294.50	1 (-)	295.69	0.980 (0.150)	294.84
D_{11}	1 (-)	-1390.85	1 (-)	-1387.33	1 (-)	-1390.97
P_{11}	0.955 (0.077)	-2036.52	0.880 (0.120)	-2035.07	0.920 (0.108)	-2037.68
B_{11}	0.395 (0.242)	-856.00	0.290 (0.225)	-854.66	0.421 (0.219)	-853.525
S_{11}	0.839 (0.182)	-808.70	0.726 (0.214)	-814.259	0.816 (0.216)	-809.089
V_{23}	0.846 (0.108)	2373.00	0.815 (0.116)	2377.61	0.815 (0.112)	2376.752
MOE_{23}	0.989 (0.027)	-344.16	1 (-)	-343.49	0.980 (0.036)	-335.949
MOR_{23}	0.916 (0.058)	-3184.94	0.863 (0.082)	-3182.08	0.900 (0.061)	-3177.57
DEN_{23}	0.968 (0.030)	-5227.90	0.932 (0.050)	-5219.2	0.943 (0.042)	-5216.39

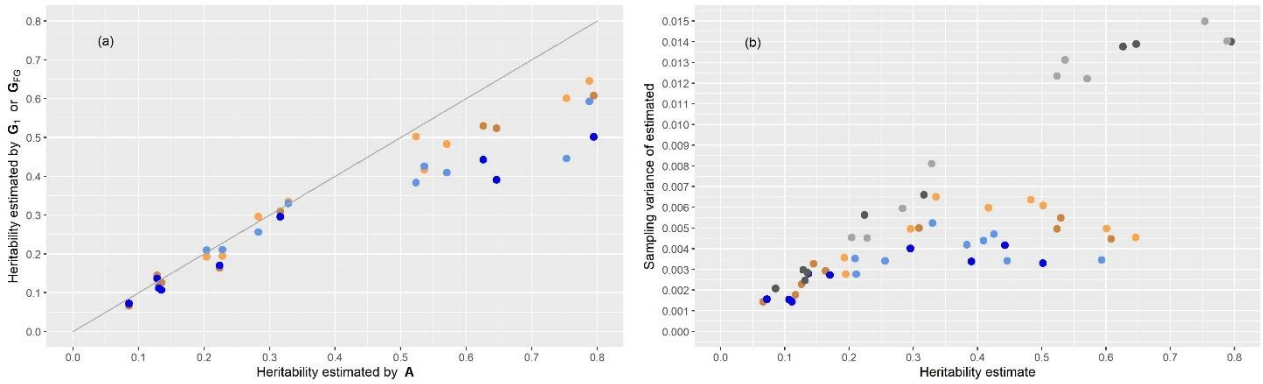
The heritabilities at Brecon were lower than at Kintyre for all juvenile traits, most notably for S_{11}, but were more similar at the two sites for the quality traits and merchantable volume. Figure 1 shows the tendency for the heritability estimated using $\mathbf{G}_{VR1}$ to be less than that estimated using **A**, with the estimates from using $\mathbf{G}_{FG}$ being intermediate. Figure 1 also illustrates that S.E.s of the estimates

were consistently smaller for $\mathbf{G}_{VR1}$ than for **A** and showed much weaker dependence on the magnitude of the estimate. The S.E.s for $\mathbf{G}_{FG}$ showed similar trends as $\mathbf{G}_{VR1}$ but were less extreme.

Table 4. The heritabilities for the 9 traits at Brecon and Kintyre using the three relationship matrices based on pedigree (**A**), genomic IBS information ($\mathbf{G}_{VR1}$), and genomic IBD information ($\mathbf{G}_{FG}$). The S.E.s are shown in parentheses.

	Brecon			Kintyre		
Trait	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$
H_{06}	0.135 (0.053)	0.107 (0.048)	0.126 (0.048)	0.204 (0.067)	0.209 (0.059)	0.193 (0.060)
D_{11}	0.131 (0.050)	0.111 (0.038)	0.117 (0.042)	0.228 (0.067)	0.211 (0.053)	0.195 (0.053)
P_{11}	0.224 (0.075)	0.170 (0.052)	0.164 (0.054)	0.536 (0.115)	0.425 (0.069)	0.417 (0.077)
B_{11}	0.128 (0.055)	0.137 (0.053)	0.145 (0.057)	0.329 (0.090)	0.330 (0.072)	0.335 (0.081)
S_{11}	0.085 (0.045)	0.072 (0.039)	0.066 (0.038)	0.571 (0.111)	0.409 (0.066)	0.483 (0.080)
V_{23}	0.317 (0.081)	0.295 (0.063)	0.309 (0.071)	0.283 (0.077)	0.256 (0.058)	0.296 (0.070)
MOE_{23}	0.637 (0.118)	0.390 (0.058)	0.524 (0.070)	0.753 (0.122)	0.446 (0.058)	0.601 (0.070)
MOR_{23}	0.626 (0.117)	0.442 (0.064)	0.530 (0.074)	0.523 (0.111)	0.383 (0.065)	0.502 (0.078)
DEN_{23}	0.795 (0.118)	0.501 (0.057)	0.608 (0.067)	0.789 (0.118)	0.593 (0.059)	0.646 (0.063)

Figure 1. (a) The relationship of h^2 estimated using **A** and those from using $\mathbf{G}_{VR1}$ (in blue shades) and $\mathbf{G}_{FG}$ (in tan shades), and (b) the relationship between the sampling variance of the estimates and the estimate itself from using different relationship matrices (pedigree **A**, grey shades; genomic IBS $\mathbf{G}_{VR1}$, blue shades; and genomic IBD $\mathbf{G}_{FG}$, tan shades). In both (a) and (b) the darker and lighter shades of each colour denote estimates from Brecon and Kintyre respectively



373 *Model-derived accuracies.* The model-derived accuracies of the EBVs assume that both the model
374 structure and the estimated parameters are correct. Table 5 gives the accuracies for each trait for
375 performance at Brecon and Kintyre. Model-derived accuracies were only obtained for Models I, II,
376 and III, but not IIIA. Given the potential GxE (Supplementary Information 4) the accuracies of an EBV
377 for a tree will depend on where its own recording took place, although the bivariate model ensures
378 information across both sites is collated appropriately.
379
380 Some broad trends are evident. Firstly, the accuracies are lower for traits with lower heritabilities as
381 records have low information for EBV. Secondly, lower genetic correlations across sites promote
382 greater accuracy of EBV within sites compared to between sites. None of the relationship matrices
383 show a clear superiority, but $\mathbf{G}_{FG}$ is typically intermediate. There is a trend for **A** to have greater
384 accuracy for the quality traits, and for $\mathbf{G}_{VR1}$ to offer greater accuracy with the remaining traits. In
385 some cases, such as DEN_{23}, the difference can be quite substantial.
386
387 **Table 5.** The model-derived accuracy of EBVs for 9 traits at Brecon (top) and Kintyre (bottom)
388 according to the site of origin of the tree, obtained when using the full bivariate model (Eqn 1) and
389 three different relationship matrices based on pedigree (**A**), genomic IBS ($\mathbf{G}_{VR1}$), and genomic IBD
390 ($\mathbf{G}_{FG}$) information. For a given site and trait, green shading indicates relationship matrix with the
391 highest accuracy given the tree origin and red shading the lowest.

Tree origin	Brecon			Kintyre		
Rel. matrix	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$
Trait	Brecon					
H_{06}	0.630	0.656	0.639	0.655	0.691	0.662
D_{11}	0.646	0.674	0.651	0.698	0.717	0.691
P_{11}	0.692	0.698	0.675	0.772	0.730	0.725
B_{11}	0.572	0.599	0.600	0.529	0.548	0.558
S_{11}	0.592	0.584	0.588	0.702	0.622	0.668
V_{23}	0.704	0.726	0.714	0.663	0.688	0.677
MOE_{23}	0.841	0.794	0.813	0.883	0.811	0.831
MOR_{23}	0.834	0.798	0.811	0.766	0.753	0.770
DEN_{23}	0.907	0.831	0.845	0.876	0.826	0.826
Trait	Kintyre					
H_{06}	0.630	0.656	0.639	0.662	0.692	0.668
D_{11}	0.646	0.674	0.651	0.698	0.717	0.691
P_{11}	0.691	0.708	0.681	0.798	0.785	0.764
B_{11}	0.591	0.637	0.617	0.709	0.736	0.724
S_{11}	0.637	0.668	0.651	0.807	0.770	0.783
V_{23}	0.667	0.689	0.677	0.695	0.715	0.714

MOE_{23}	0.835	0.794	0.806	0.891	0.812	0.841
MOR_{23}	0.794	0.758	0.776	0.795	0.779	0.801
DEN_{23}	0.886	0.813	0.822	0.896	0.852	0.851

392

393 *Accuracies derived by cross-validation*. The accuracies obtained by cross-validation as judged by
394 the correlations of EBV with phenotype are shown in Table 6. The accuracies have S.E.s but are
395 necessarily restricted to EBVs for trees at the site for which they have a phenotype. Differences are
396 small compared to the S.E.s but some suggestive trends are apparent: the performance of models
397 using **A** (expected IBD information according to pedigree) is comparatively less effective, with an
398 advantage for incorporating the genomic models, but not necessarily using $\mathbf{G}_{VR1}$ (IBS information
399 according to genomic markers) over $\mathbf{G}_{FG}$ or $\mathbf{M}_{FG}$ (realised IBD information according to genomic
400 markers within the pedigree). This is also evident in Table 7 in which the correlations have been
401 pooled over sites. Direct comparison with Table 5 requires dividing the values in Tables 6 and 7
402 through by the true $\sqrt{h^2}$ at each site using the same value for each model (as it is a rescaling of the
403 phenotype with no dependence on the model used). For example, with P_{11} assuming the true h^2 to
404 be 0.20 and 0.45 at Brecon and Kintyre respectively, the scaling factors are 0.45 and 0.67, and the
405 cross-validation accuracies of EBV are approximately 0.67 and 0.73, marginally lower than the
406 model-derived accuracies of approximately 0.69 and 0.78. Evidence for extra block variation in the
407 estimates of accuracy from the meta-analysis was restricted to S_{11} and was observed for all the
408 models.

409 **Table 6.** The correlation of the EBV with the recorded phenotype obtained by cross-validation using
410 four different relationship matrices based on pedigree (**A**), genomic IBS ($\mathbf{G}_{VR1}$), and genomic IBD
411 ($\mathbf{G}_{FG}$ and $\mathbf{M}_{FG}$) information. The S.E.s are given in parentheses. The green shading indicates the
412 best model for a given trait within site and the red shading the poorest.

Site	Brecon				Kintyre			
Trait	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$	$\mathbf{M}_{FG}$	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$	$\mathbf{M}_{FG}$
H_{06}	0.275 (0.041)	0.264 (0.041)	0.278 (0.041)	0.267 (0.043)	0.313 (0.042)	0.328 (0.040)	0.315 (0.042)	0.318 (0.042)
D_{11}	0.270 (0.041)	0.272 (0.041)	0.278 (0.041)	0.260 (0.043)	0.379 (0.037)	0.406 (0.036)	0.377 (0.037)	0.394 (0.039)
P_{11}	0.309 (0.040)	0.309 (0.040)	0.302 (0.040)	0.289 (0.045)	0.483 (0.033)	0.496 (0.033)	0.483 (0.033)	0.490 (0.035)
B_{11}	0.201 (0.042)	0.208 (0.042)	0.212 (0.042)	0.211 (0.044)	0.379 (0.037)	0.375 (0.037)	0.388 (0.037)	0.373 (0.039)
S_{11}	0.164 (0.043)	0.161 (0.043)	0.157 (0.043)	0.143 (0.045)	0.542 (0.038)	0.533 (0.037)	0.547 (0.037)	0.557 (0.038)

V_{23}	0.421 (0.037)	0.435 (0.036)	0.426 (0.037)	0.428 (0.038)	0.374 (0.038)	0.383 (0.037)	0.393 (0.039)	0.434 (0.037)
MOE_{23}	0.509 (0.033)	0.542 (0.031)	0.537 (0.031)	0.531 (0.033)	0.574 (0.029)	0.578 (0.031)	0.587 (0.028)	0.602 (0.029)
MOR_{23}	0.525 (0.032)	0.542 (0.031)	0.541 (0.031)	0.541 (0.032)	0.499 (0.037)	0.507 (0.035)	0.520 (0.034)	0.528 (0.036)
DEN_{23}	0.594 (0.028)	0.603 (0.028)	0.608 (0.028)	0.603 (0.029)	0.593 (0.028)	0.628 (0.026)	0.621 (0.027)	0.634 (0.027)

Table 7. The correlation of the EBV with the recorded phenotype obtained by cross-validation using four different relationship matrices based on pedigree (**A**), genomic IBS ($\mathbf{G}_{VR1}$), and genomic IBD ($\mathbf{G}_{FG}$ and $\mathbf{M}_{FG}$) information when pooled over Brecon and Kintyre. The S.E.s are given in parentheses. The green shading indicates the best model within site and the red shading the poorest.

Trait	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$	$\mathbf{M}_{FG}$
H_{06}	0.294 (0.029)	0.297 (0.028)	0.296 (0.029)	0.293 (0.030)
D_{11}	0.326 (0.028)	0.341 (0.027)	0.329 (0.028)	0.329 (0.029)
P_{11}	0.401 (0.026)	0.408 (0.027)	0.397 (0.026)	0.395 (0.029)
B_{11}	0.294 (0.028)	0.295 (0.028)	0.304 (0.028)	0.295 (0.030)
S_{11}	0.367 (0.036)	0.361 (0.035)	0.366 (0.036)	0.365 (0.039)
V_{23}	0.397 (0.026)	0.409 (0.026)	0.409 (0.026)	0.431 (0.027)
MOE_{23}	0.543 (0.022)	0.561 (0.021)	0.562 (0.021)	0.567 (0.022)
MOR_{23}	0.511 (0.023)	0.524 (0.022)	0.530 (0.022)	0.534 (0.023)
DEN_{23}	0.594 (0.020)	0.616 (0.019)	0.614 (0.019)	0.619 (0.020)

Bias. All estimates of the slope of the regression of phenotype on EBV, pooled over validation sets, were within 1.5 S.E.s of 1.0. This was true within sites and pooled across the sites, for all three relationship matrices. The detailed results are shown in Supplementary Information 6. Therefore, overall, there was no evidence of bias in the EBVs. However, within Kintyre the meta-analysis indicated some variation among blocks particularly for H_{06} and S_{11}.

Discussion

The study's main objective was to compare the accuracy of predicting breeding values of Sitka spruce using different relationship matrices with a particular emphasis on whether, at its current stage of development, the genomics approach provides more accurate predictors than reliance on pedigree. An additional avenue of interest was the comparison between the use of IBS and IBD, since the application of genomics does not pre-suppose a reliance on IBS as the markers can be used to track locus-specific IBD through pedigrees back to the base population (Fernando & Grossman, 1989). This places more emphasis on the Mendelian sampling from the base rather than assuming population wide disequilibrium between a marker and a QTL. In Sitka spruce, where marker densities are currently low and where available full- and half-sib families are large, the intensity of selection applied to within-family variation arising from the Mendelian sampling may constitute a significant part of the total intensity.. The findings indicate that in terms of accuracy and bias, the use of 4123 SNP in the study's family structure was competitive with pedigree information, with (for this marker density) a slight benefit in using realised/genomic IBD information. This does not consider cost or feasibility, and it should be noted the study population was rich in half- and full-sib relationships. Further findings showed that estimates of heritability among different relationship matrices were broadly comparable, though for traits with greater heritability the genomic models produced lower estimates indicating a gap due to the low marker density. The genome-based estimates had smaller standard errors and showed weaker dependence on the magnitude of the estimate than pedigree-based estimates (Swiger et al. 1964; Hill, 1971)

While the observed differences in accuracy were not large, there were subtle changes in the pattern of best prediction models coming from model-based or validation methodologies. In model-based accuracy the outcomes assume that both the model, with its inherent assumptions, and its parameterisation are correct, whereas in validation accuracies the predictions are tested against new phenotypes formed with the full complexity of the true underlying influences and interactions. The stronger objectivity of the validation accuracies in avoiding self-reference makes validation accuracies more valuable during phases of model development. Validation accuracies in Tables 6 and 7 indicate Model I (using pedigree relationship matrix **A**) showed a trend for lower accuracies

compared to genomic models, which was not evident in model-based accuracies in Table 5. One distinction between the validation accuracies and model-based accuracies is that with Model I the predicted EBV in the validation accuracies are entirely based on parental averages with no estimate of the Mendelian sampling for an offspring (see Supplementary Information 3). In Model I the only information on which to predict an offspring's Mendelian sampling in this data set would come from its own phenotype, but this is excluded when generating the prediction in the validation set. Therefore, in Model 1, EBVs for two full-sibs in a validation set are identical, unlike the corresponding model-based accuracies where all the offspring phenotypes are included. In contrast, all the genomic models provide some estimate of the Mendelian sampling and differentiate full-sibs in both validation and model-based accuracies. It is this benefit of predicting the Mendelian sampling in the absence of a phenotype or offspring which underlies many of the potential benefits of genomic models. It is difficult to compare the results of this study with other tree breeding assessments of accuracy in using genomics because of this is the first study in Sitka spruce, the differing stages of genomic development across species, and the differing sizes of training sets and marker panels.

Over and above the small apparent benefit from using genomics, there were also slight differences observed between the two distinct approaches of either using IBS (Model II with $\mathbf{G}_1$) or IBD (Models III and IIIA with $\mathbf{G}_{FG}$ and $\mathbf{M}_{FG}$ respectively). The model-based accuracies tended to be greater using Model II rather than Model III, whereas validation accuracies tended to be greatest using Model IIIA (with $\mathbf{M}_{FG}$). One perspective on Model III using $\mathbf{G}_{FG}$ is that it builds directly on $\mathbf{A}$ by adding the discrimination among the sibs, so leading to a hypothesis that the accuracy should increase, but this was not observed in the model-based accuracies. An explanation for this is that underlying Model III is the estimation of a genetic value for all the distinct fragments of the base generation genomes that were inherited by the offspring and this greatly increases the potential number of parameters to be estimated, of the order of twice the number of observed recombinations among the offspring, i.e., 2 x 21.42 x 1341=57475, where 21.42 is the genome length in Morgans (Tumas et al. 2024). This is more than in $\mathbf{A}$ where only the full-genome breeding values for the parents (45) and the full-genome Mendelian sampling term for each offspring (1341) are required to be estimated, or $\mathbf{G}_1$ with 4123 marker effects (see Supplementary Information 3). When using $\mathbf{G}_{FG}$ rather than $\mathbf{A}$ for validation the

accuracy benefits from predictions of the Mendelian sampling term in the validation set, but also has the same implicit constraint as **A** that the Mendelian sampling variance among the offspring is half the genetic variance in the parental generation. This is released in Model IIIA using $\mathbf{M}_{FG}$ by letting the pedigree account for the between-family variance, and this appeared to have a small benefit in improving the validation accuracy. Comparisons with Jurcic et al. (2021), where the use of markers to capture IBD was also explored, are problematical as the authors used the same training phenotypes both to develop the genomic predictions and the EBVs with which these predictions were correlated, so introducing a bias.

There are several threads of evidence that the accuracy of predictions using Model II with $\mathbf{G}_1$ from these 45 parents is reduced by the marker density (4123 markers over 21.42 M). Firstly, inspection of Supplementary Information 4 shows a syndrome of lower estimates of genetic variance and lower phenotypic variance for traits estimated using **A**. The phenotypic variance is typically the most accurately estimated variance in the models but when the additive genetic variance is not captured the positive (co)variances among sibs that are unaccounted for fall into and reduce the residual, and overall reduce, the phenotypic variance. These reductions would be expected to be most visible in the quality traits with high h^2, which is what is observed. This contributes to the lower heritabilities observed in Figure 1 for traits with high h^2. Genomic estimates of h^2 will benefit from the rich sib-structure as genomic relationship matrices will exhibit this, albeit diffusely, and when h^2 is low this will be the dominant source of information and so estimates agree closely with **A**. As h^2 increases the genomic information from the Mendelian sampling variation will have more weight compared to family means in determining the h^2 estimate, and if the genetic variation is not being fully captured the estimate of h^2 will be reduced. Model IIIA gave an opportunity to examine this in more detail as the Mendelian sampling variance ($\frac{1}{2}v_M$ in Eqn 4) was estimated independently of the variance of family means ($2v_T$ in Eqn 4), and the ratio estimated from the markers was 0.666 (S.E. 0.121), which is significantly different from the classical value of 1 with random mating. The historical design of the trial was for random mating among the parents available in the EP (S. Lee, pers. comm) and the subset chosen for this study made no reference to performance, so no large deviation from 1

would be anticipated. Therefore, this suggests the markers captured only 83% of the full variance (0.833=(1+0.666)/(1+1)). The final thread of evidence is an empirical comparison with livestock breeding, where populations have effective population sizes close to 50 and the breeders have had access to a wide range of densities up to 30,000 markers/M, and where the standard has settled at ~2,500 markers/Morgan, i.e., 10-fold greater than used here.

The benefits of genomics will come from re-structuring breeding programmes through, for example, genomic selection within families where phenotypes are unavailable: in untested seeds or saplings for early selection decisions, or where phenotyping is destructive and prevents breeding, or recruiting new individuals to the breeding population that were not recorded. In these situations, the candidate's phenotype is not available, and so the validation accuracies for predicting the breeding value are critical. The results suggest that with the current 1341 records, the accuracies of predicting the breeding value are typically in the range from 0.6 to 0.7 across the traits, with some estimates as low as 0.5 for S_{11} at Brecon. Using a framework provided by Dekkers (2007), Daetwyler at al. (2008), and Gorjanc et al. (2015), Supplementary Information 6 explores how accuracies might change as the size of the training sets increase and suggest it would require an 8-fold increase in records for genomic predictions to reach 0.9 for all traits studied, close to the maximum accuracy with this number of markers (i.e., $0.913 = \sqrt{0.833}$). One might speculate that more would be needed if the population lacked the structure used here as predictions would rely more strongly on population-wide LD. With pedigree **A** Supplementary Information 6 shows an 8-fold increase in records per parent would only increase a validation accuracy from 0.5 to 0.667, approaching the maximum possible of 0.707 when relying on parent averages for new generations of non-phenotyped individuals. These projections are extrapolations from the data presented here, but they demonstrate that once the genomic model is competitive for validation as found here, then increasing the resources will result in increasing benefits of genomics over pedigree.

Some authors (e.g., Beaulieu et al. 2022) have hypothesised that the pedigree predictions are biased compared to genomic predictions, but the results presented here show no evidence of bias in over- or under-estimating EBVs in any of the models used (see Supplementary Information 7). Theory

would support that given the appropriate model for the data there should be no bias in predictions. It is arguable that pedigree is more open to cryptic confounding with intractable or unmodelled sources of variation in field data (e.g., maternal or geographical sources) although these problems also beset QTL discovery. Furthermore, heritability estimates using the variance parameter estimated by a model do not provide an estimate of heritability within a population or sub-population, and require further scaling (Legarra, 2016) or post-processing model results (Lara et al., 2022).

In conclusion, the study has shown that the accuracy of prediction using genomics in this structured population is already competitive with that from pedigree with no evidence of bias. Progress beyond this experimental population will require (i) a denser SNP chip to ensure the IBS approaches using VanRaden (2008) matrices are capable of capturing the full genetic variance in less structured populations, and (ii) a greater training set of $O(10^4)$, rather than $O(10^3)$ as here, to allow accuracies to increase significantly further towards the theoretical maximum is 1 for an unphenotyped individual from the elite population. Achieving accurate prediction with genomics in tree breeding, especially in conifers with characteristically large and complex genomes, would revolutionize the field and practice, allowing for selections without the need to wait decades for phenotyping that would result in more rapid economic gain for forestry on a global scale.

Acknowledgements

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, which included an industry contribution from the Conifer Breeding Co-operative, and ISP grants to The Roslin Institute (BBS/E/D/30002275, BBS/E/RL/230001A, and BBS/E/RL/230001C), all of which are gratefully acknowledged. The authors are grateful to staff at the Northern Research Station particularly Stuart O'Hara for the DNA extractions, and all the staff involved with the field work, including Adam Ash, Glyndawr Jones, Ruben Manso, Andrew Price, and Steve Lee. Steve Lee was also very active in the initiation and development of the project.

Author Contributions

JAW contributed to project initiation, development, and management, led the data analysis, and wrote the initial draft of the manuscript and contributed to its subsequent development.
MJ, HT, and JJI contributed to data preparation and analysis.
JPM supervised the field work and associated sampling and led all aspects of the measurement of wood quality on the samples post-felling.

JC contributed to project initiation and development, and discussing results.
JM contributed to project initiation, development and management, and discussing results.
GG contributed to project management, discussing results, and writing parts of the manuscript.
All authors have read the draft manuscript and provided edits in revision.

Conflict of Interest

The authors state no conflict of interest.

Data Archiving

Data is available on a research collaboration basis by approaching the corresponding author.

Research Ethics Statement

Not applicable.

References

Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S et al. (2010) Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science* 93:743–752. https://doi.org/10.3168/jds.2009-2730

Beaulieu J, Doerksen T, MacKay J, Rainville A, Bousquet J (2014) Genomic selection accuracies within and between environments and small breeding groups in white spruce. *BMC Genomics* 15:1048. https://www.biomedcentral.com/1471-2164/15/1048

Beaulieu J, Lenz P, Bousquet J. (2022) Metadata analysis indicates biased estimation of genetic parameters and gains using conventional pedigree information instead of genomic-based approaches in tree breeding. *Scientific Reports* 12:3933. https://doi.org/10.1038/s41598-022-06681-y

Beaulieu J, Lenz PRN, Laverdière J-P, Nadeau S, Bousquet J (2024) A meta-analysis on the effects of marker coverage, status number, and size of training set on predictive accuracy and heritability estimates from genomic selection in tree breeding. *Tree Genetics & Genomes* 20:1-14. https://doi.org/10.1007/s11295-024-01653-x

Bernhardsson C, Vidalis A, Wang X, Scofield DG, Schiffthaler B et al. (2019). An ultra-dense haploid genetic map for evaluating the highly fragmented genome assembly of Norway spruce (*Picea abies*). *G3 Genes|Genomes|Genetics* 9:1623–32. https://doi.org/10.1534/g3.118.200840

Birol I, Raymond A, Jackman SD, Pleasance S, Coope R (2013) Assembling the 20 Gb white spruce (Picea glauca) genome from whole-genome shotgun sequencing data. *Bioinformatics* 29:1492-1497. https://doi.org/10.1093/bioinformatics/btt178

British Standards Institute (1957) Methods of testing small clear specimens of timber. British Standard Institute, London, p 31 BS 373:1957

Butler DG, Cullis BR, Gilmour AR, Gogel BJ, Thompson R (2017) ASReml-R Reference Manual Version 4. https://vsni.co.uk/software/asreml-r

Cameron AD, Kennedy SG. Lee SJ (2012) The potential to improve growth rate and quality traits of stem straightness and branching habit when breeding *Picea sitchensis* (Bong. Carr). *Annals of Forest Science* 69:363–371. https://doi.org/10.1007/s13595-011-0167-y

Chen ZQ, Baison J, Pan J, Karlsson B, Andersson B et al. (2018) Accuracy of genomic selection for growth and wood quality traits in two control-pollinated progeny trials using exome capture as the genotyping platform in Norway spruce. *BMC Genomics* 19:946. https://doi.org/10.1186/s12864-018-5256-y

Christensen OF, Lund MS (2010). Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution* 42:1-8. https://doi.org/10.1186/1297-9686-42-2

Cuyabano, B.C.D., Boichard, D. & Gondro, C. Expected values for the accuracy of predicted breeding values accounting for genetic differences between reference and target populations. Genet Sel Evol 56, 15 (2024). https://doi.org/10.1186/s12711-024-00876-9

Daetwyler HD, Villanueva B, Woolliams JA (2008) Accuracy of Predicting the Genetic Risk of Disease Using a Genome-Wide Approach. PLoS One 3:e3395. https://doi.org/10.1371/journal.pone.0003395

Dekkers JCM (2007) Prediction of response to marker-assisted and genomic selection using selection index theory. *Journal of Animal Breeding and Genetics* 124:331-341. https://doi.org/10.1111/j.1439-0388.2007.00701.x

Dekkers JCM, Su H, Cheng J (2021) Predicting the accuracy of genomic predictions. Genetics Selection Evolution 53:55. https://doi.org/10.1186/s12711-021-00647-w

DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Controlled Clinical Trials* 7:177-188. https://doi.org/10.1016/0197-2456(86)90046-2

Duarte D, Jurcic EJ, Dutour J, Villalba PV, Centurión C, Grattapaglia D and Cappa EP (2024) Genomic selection in forest trees comes to life: unraveling its potential in an advanced four-generation *Eucalyptus grandis* population. *Frontiers in Plant Science* 15-2024. https://doi.org/10.3389/fpls.2024.1462285

Ejima A, Uchiyama K, Mori H, Tsumura Y (2023). Genomic prediction of tree height, wood stiffness, and male flower quantity traits across two generations in selected individuals of *Cryptomeria japonica* D. Don (Japanese cedar). *Forests* 14:2097. https://doi.org/10.3390/f14102097

Elsen J-M (2016) Approximated prediction of genomic selection accuracy when reference and candidate populations are related. *Genetics Selection Evolution* 48:18. https://doi.org/10.1186/s12711-016-0183-3

Elsen J-M (2017) An analytical framework to derive the expected precision of genomic selection. *Genetics Selection Evolution* 49:95. https://doi.org/10.1186/ s12711-017-0366-6

Fernando RL, Grossman M (1989) Marker assisted selection using best linear unbiased prediction. *Genetics Selection Evolution* 21:467-477. https://doi.org/10.1051/gse:19890407

Freeman JS, Slavov GT, Butler JB, Frickey T, Graham NJ et al. (2022) High density linkage maps, genetic architecture, and genomic prediction of growth and wood properties in *Pinus radiata*. BMC Genomics 23:731. https://doi.org/10.1186/s12864-022-08950-6

Gamal El-Dien O, Ratcliffe B, Klápště J, Chen C, Porth I et al. (2015). Prediction accuracies for growth and wood attributes of interior spruce in space using genotyping-by-sequencing. *BMC Genomics* 16:370. https://doi.org/10.1186/s12864-015-1597-y

Goddard ME, Hayes BJ. (2007) Genomic selection. *Journal of Animal Breeding and Genetics* 124:323-330. https://doi.org/10.1111/j.1439-0388.2007.00702.x

Gorjanc G, Bijma P, Hickey JM (2015) Reliability of pedigree-based and genomic evaluations in selected populations. *Genetics Selection Evolution* 47:65. https://doi.org/10.1186/s12711-015-0145-1

Grattapaglia D. (2022) Twelve years into genomic selection in forest trees: climbing the slope of enlightenment of marker assisted tree breeding. *Forests* 13:1554. https://doi.org/10.3390/f13101554

Grattapaglia D, Resende MDV (2011) Genomic selection in forest tree breeding. *Tree Genetics and Genomes* 7:241–255. https://doi.org/10.1007/ s11295-010-0328-4

Henderson CR (1976) A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32:69-83. https://doi.org/10.2307/2529339

Hill WG (1971). Design and efficiency of selection experiments for estimating genetic parameters. *Biometrics* 27:293-311. https://doi.org/10.2307/2528996

Ilska, J., Tolhurst, D., Tumas, H. Maclean JP, Cottrell J et al. (2023) Additive and non-additive genetic variance in juvenile Sitka spruce (Picea sitchensis Bong. Carr). *Tree Genetics & Genomes* 19:53. https://doi.org/10.1007/s11295-023-01627-5

Isik F (2022) Genomic prediction of complex traits in perennial plants: a case for forest trees. In *Complex Trait Prediction: Methods and Protocols*; Ahmadi N, Bartholomé, J, Eds; Humana: New York, NY, USA. pp. 493–520.

Jurcic EJ, Villalba PV, Pathauer PS, Palazzini DA, Oberschelp GPJ et al. (2021) Single-step genomic prediction of *Eucalyptus dunnii* using different identity-by-descent and identity-by-state relationship matrices. *Heredity* 127:176–189. https://doi.org/10.1038/s41437-021-00450-9

Lara LA, Pocrnic I, Oliveira T, Gaynor RC, Gorjanc G (2022) Temporal and genomic analysis of additive genetic variance in breeding programmes. Heredity 128:21. https://doi.org/10.1038/s41437-021-00485-y

Lebedev VG, Lebedeva TN, Chernodubov AI, Shestibratov KA (2020) Genomic selection for forest tree improvement: methods, achievements and perspectives. *Forests* 11:1190. https://doi.org/10.3390/f11111190

Legarra A (2016) Comparing estimates of genetic variance across different relationship models. *Theoretical Population Biology* 107:26-30. https://doi.org/10.1016/j.tpb.2015.08.005

Legarra A, Aguilar I, Misztal I (2009) A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science* 92:4656-4663. https:\\doi.org\10.3168/jds.20092061

Liu J-J, Schoettle AW, Sniezko RA, Yao F, Zamany A et al. (2019) Limber pine (Pinus flexilis James) genetic map constructed by exome-seq provides insight into the evolution of disease resistance and a genomic resource for genomics-based breeding. *The Plant Journal* 98:745–758. https://doi.org/10.1111/tpj.14270

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

Luan T, Woolliams JA, Ødegård J, Dolezal M, Roman-Ponce SI et al. (2012) The importance of identity-by-state information for the accuracy of genomic selection. *Genetics Selection Evolution* 44:28. doi:10.1186/1297-9686-44-28

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

Meuwissen THE, Luan T, Woolliams JA (2011) The unified approach to the use of genomic and pedigree information in genomic evaluations revisited. *Journal of Animal Breeding and Genetics* 128:429-439. https://doi.org/10.1111/j.1439-0388.2011.00966.x

Misztal I, Legarra A, Aguilar I (2009) Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *Journal of Dairy Science* 92:4648-4655. https://doi.org/10.3168/jds.2009-2064

Mrode R, Pocrnic I (2024) Linear Models for the Prediction of the Genetic Merit of Animals, CABI, 4th edition.

Müller BSF, Neves LG, de Almeida Filho JE, Resende MFR Jr, Muñoz PR et al. (2017) Genomic prediction in contrast to a genome-wide association study in explaining heritable variation of complex growth traits in breeding populations of *Eucalyptus*. *BMC Genomics* 18:524. https://doi.org/10.1186/s12864-017-3920-2

Neves LG, Davis JM, Barbazuk WB, Kirst M. (2014) A high-density gene map of Loblolly pine (*Pinus taeda* L.) based on exome sequence capture genotyping. *G3 Genes|Genomes|Genetic* 4:29–37. https://doi.org/10.1534/g3.113.008714

Niu S, Li J, Bo W, Yang W, Zuccolo A et al. (2022). The Chinese pine genome and methylome unveil key features of conifer evolution. *Cell* 185:204-217.e14. https://doi.org/10.1016/j.cell.2021.12.006

Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin Y-C et al (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature* 497:579–584. https://doi.org/10.1038/nature12211

Papin V, Gorjanc G, Pocrnic I, Bouffier L, Sanchez L (2024) Unlocking genome-based prediction and selection in conifers: the key role of within-family prediction accuracy illustrated in maritime pine (*Pinus pinaster* Ait.). *Annals of Forest Science* 81:52. https://doi.org/10.1186/s13595-024-01269-0

Pavy N, Gagnon F, Rigault P, Blais S, Deschênes A et al. (2013) Development of high-density SNP genotyping arrays for white spruce (*Picea glauca*) and transferability to subtropical and nordic congeners. *Molecular Ecology Resources* 13:324-336. https://doi.org/10.1111/1755-0998.12062

Plomion C, Bartholomé J, Lesur I, Boury C, Rodríguez-Quilón I et al. 2016. High-density SNP assay development for genetic analysis in maritime pine (*Pinus pinaster*). *Molecular Ecology Resources*, 16:574-587. https://doi.org/10.1111/1755-0998.12464

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR et al. (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics* 81:559–575. https://doi.org/10.1086/519795

R Core Team (2022) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Scott AD, Zimin AV, Puiu D, Workman R, Britton M et al. (2020) A reference genome sequence for giant sequoia. *G3 Genes|Genomes|Genetics* 10:3907-3919. https://doi.org/10.1534/g3.120.401612

Self SG, K-Y Liang (1987) Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests under Nonstandard Conditions. *Journal of the American Statistical Association* 82:605-610. https://doi.org/10.1080/01621459.1987.10478472

Swiger L, Harvey WR, Everson D, Gregory K (1964) The variance of intraclass correlation involving groups with one observation. *Biometrics* 20:818-826. https://doi.org/10.2307/2528131

VanRaden PM (2008). Efficient methods to compute genomic predictions. J*ournal of Dairy Science* 91:4414-4423. https://doi.org/10.3168/jds.2007-0980

Thistlethwaite FR, Ratcliffe B, Klápště J, Porth I, Chen C et al. (2017) Genomic prediction accuracies in space and time for height and wood density of Douglas-fir using exome capture as the genotyping platform. *BMC Genomics* 18:930. https://doi.org/10.1186/s12864-017-4258-5

Tumas H, Ilska JJ, Gérardi S, Laroche J, A'Hara S et al. (2024) High-density genetic linkage mapping in Sitka spruce advances the integration of genomic resources in conifers. *G3 Genes|Genomes|Genetics* 14:kae020. https://doi.org/10.1093/g3journal/jkae020

Walker TD, Cumbie WP, Isik F (2022) Single-step genomic analysis increases the accuracy of within-family selection in a clonally replicated population of *Pinus taeda* L. Forest Science 68:37–52. https://doi.org/10.1093/forsci/fxab054

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. *Genetics Selection Evolution*, 50:1-15. https://doi.org/10.1186/s12711-018-0438-2

Wright, S (1922) Coefficients of inbreeding and relationship. *The American Naturalist* 56:330-338. https://doi.org/10.1086/279872

Xiong X, Gou J, Liao Q, Li Y, Zhou Q et al. 2021. The Taxus genome provides insights into paclitaxel biosynthesis. *Nature Plants* 7:1026–36. https://doi.org/10.1038/s41477-021-00963-5

Zhou L, Chen Z, Olsson L, Grahn T, Karlsson B. et al. (2020) Effect of number of annual rings and tree ages on genomic predictive ability for solid wood properties of Norway spruce. *BMC Genomics* 21:323. https://doi.org/10.1186/s12864-020-6737-3

Zimin A, Stevens KA, Crepeau MW, Holtz-Morris A, Koriabine M et al. (2014). Sequencing and Assembly of the 22-Gb Loblolly Pine Genome. *Genetics* 196:875–890. https://doi.org/10.1534/genetics.113.159715

Supplementary Information 1. Weighting of Small Clear Specimens (SCS) to obtain cross-sectional area (CSA) measures of MOE, MOR, and DEN.

The CSA measures for estimates of MOE, MOR, and DEN were a weighted average of the individual specimens. If z is the CSA measure of a trait for an individual with m SCS then $z = \sum_{j=1}^{m} w_j x_j$, where x_j and w_j are the trait and weight for the j^{th} SCS. The weights were allocated according to the position of the SCS in the section and are shown in Table S1.1.

Table S1.1. Weights (w) attached to SCS as a function of the number of SCS obtained per section (n) used for averaging, together with their sum. The first SCS will be in the pith, and subsequent pairs correspond to the same sequence position along opposing radii. In some sections, there was sufficient symmetry to obtain an equal number of rings along opposing radii (so n is odd) while in others one radius had one more sample than the other (so n is even).

	Radius in cm													
n	1	3	5	7	9	11	13	15	17	19	21	23	25	Sum
3	0.04	0.32	0.64											1
4	0.02	0.16	0.33	0.49										1
5	0.01	0.10	0.20	0.30	0.40									1
6	0.01	0.07	0.13	0.20	0.26	0.33								1
7	0.01	0.05	0.09	0.14	0.19	0.24	0.28							1
8	0.00	0.04	0.07	0.11	0.14	0.18	0.21	0.25						1
9	0.00	0.03	0.06	0.08	0.11	0.14	0.17	0.19	0.22					1
10	0.00	0.02	0.04	0.07	0.09	0.11	0.13	0.16	0.18	0.20				1
11	0.00	0.02	0.04	0.05	0.07	0.09	0.11	0.13	0.15	0.16	0.18			1
12	0.00	0.02	0.03	0.05	0.06	0.08	0.09	0.11	0.12	0.14	0.15	0.17		1
13	0.00	0.01	0.03	0.04	0.05	0.06	0.08	0.09	0.10	0.12	0.13	0.14	0.15	1

Supplementary Information 2. Assessment of imputation.

The accuracy of imputation was by two methods following Ilska et al. (2023): (i) varying the threshold for the posterior probability for calling a genotype and calculating the call rate, and (ii) a trial involving the masking of assigned genotypes accepted as known and assessing the posterior distribution of that genotype after imputation. The term assigned genotype used here implies a genotype obtained as described in the Materials & Methods prior to imputation. The fraction of genotypes assigned across all linkage groups and all offspring was 0.946.

Method 1. This summarised all the posterior genotype probabilities provided by AlphaPeel for the offspring by assuming a genotype to be called if the posterior probability for a genotype was greater than p and varying p over the range 0.5 to 1 (hence the call is unique). These were then summarised for a given value of p by considering: (a) what fractions of SNPs would be excluded if a given fraction of the offspring were required to be called for that SNP, and (b) what fraction of offspring would be excluded if a given fraction of SNPs were required to be called for that individual. The outcomes are shown for (i) prior to imputation, (ii) p=0.7, and (iii) p=0.9 for SNPs failing to meet a required call rate over offspring in Figure S2.1a and for individuals failing to meet a required call rate over SNPs in S2.1b.

The curves will tend to 1 as the required call rate tends to 1 as the constraint on including a SNP or offspring becomes stronger, 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 as the threshold value of p decreases as the constraint on calling a genotype becomes weaker. The value of the imputation is greatest when the asymptote is approached without a large decrease in p. When the threshold was set to p=0.9, 91% of SNPs had call rates exceeding 95% over all offspring (from Figure S2.1a), and 96% of offspring had call rates exceeding 95% over all SNPs (from Figure S2.1b). Without imputation, the corresponding values were 75% of SNPs call rates exceeding 95% over offspring and 84% of offspring had call rates exceeding 95% over SNPs. The results indicate substantial recovery of information obtained from the imputation.

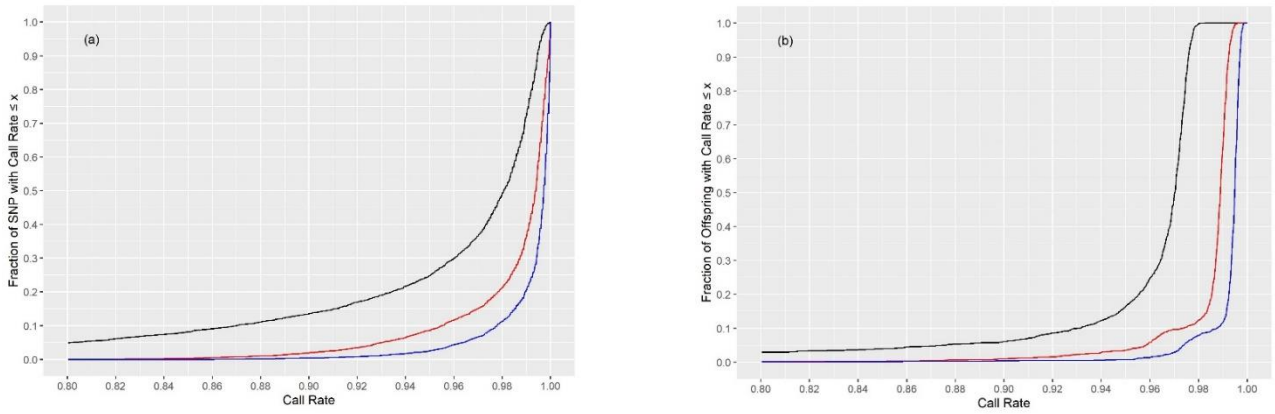


Figure S2.1. A summary of imputation success rates for offspring obtained from AlphaPeel: (a) the fraction of SNPs failing to meet a required call rate over individuals (x), and (b) the fraction of individuals failing to meet a required call rate over SNPs (x). These are shown for no imputation (black), with genotypes assigned with posterior probability >0.7 (blue), and genotypes assigned with posterior probability >0.9 (red).

Method 2. This was conducted in two stages. Firstly, 3 SNPs with assigned genotypes were randomly sampled from each linkage group for each offspring, giving a total of 48276 sampled SNPs. These prior assignments were then compared to the posterior probability for the genotype after imputation without masking the sampled genotypes, i.e., after estimating the potential error rate. The mean posterior probability was 0.981. Secondly, the sampling protocol was repeated, and these assignments were then masked prior to imputation (i.e., as if missing genotypes). The outcomes were compared as before and the mean posterior probability of the masked genotypes was lower 0.918. These results give an indication of genotype uncertainty after imputation by weighting the trial outcomes with the overall fractions of SNP genotypes obtained (0.946) and missing (0.054), i.e., $0.946 \times (1-0.981) + 0.054 \times (1-0.918) = 0.022$.

References

Ilska, J, Tolhurst, D, Tumas, H et al. (2023) Additive and non-additive genetic variance in juvenile Sitka spruce (Picea sitchensis Bong. Carr). *Tree Genetics & Genomes* 19:53. https://doi.org/10.1007/s11295-023-01627-5

Supplementary Information 3. Theoretical background to the evaluation models used.

Model I. This uses Wright's numerator relationship matrix, **A** (Wright, 1922). The matrix can be derived from IBD from the base population, which is the parental generation in this study. The use of **A** is based upon the pedigree-based inheritance model that for individual *i* its breeding value (a_i) is the average of parental breeding values plus the Mendelian sampling deviation (r_i), i.e.:

$$a_i = \tfrac{1}{2}(a_{s(i)} + a_{d(i)}) + r_i \qquad \text{[Eqn S3.1]}$$

If the additive genetic variation in the base generation is v_g then the variance of r_i is $\tfrac{1}{2}v_g[1-\tfrac{1}{2}(F_{s(i)} + F_{d(i)})]$. This model is equivalent to a ridge regression model with the set of covariates given by the scaled contributions (by descent) of each individual to all individuals with regression parameters given by scaled breeding values for base generation individuals, and scaled Mendelian terms for all other individuals, where the scaling is used to give a distribution to the parameters of MVN(0, v_g**I**). Note that the contribution of an individual to itself is 1, and its contribution to unrelated individuals or ancestors is 0. Therefore, in this data set the underlying model fitted required parameters for the *n* offspring plus *t* parents, so required the estimation of (*n*+*t*) parameters. The form of the model in Eqn 2 of the Materials & Methods using **A** calculates the breeding values directly.

Model II. This uses the genomic relationship matrix of VanRaden Model 1, $\mathbf{G}_{VR1}$ (VanRaden, 2008). $\mathbf{G}_{VR1}$ can be derived from the genome-based model for breeding values

$$a_i = \mathbf{w}_i^T\mathbf{m} \qquad \text{[Eqn S3.2]}$$

for individual *i* as a function of its centred vector of marker genotypes $\mathbf{w}_i$ and the effects of the *m* markers, **m**. The model with $\mathbf{G}_{VR1}$ is equivalent to a ridge regression model on the marker allele counts as covariates with the marker effects as parameters and assumed to be MVN(0, v_b**I**) where v_b is the variance of marker effects. Therefore, the underlying model requires the estimation of *m* parameters, which are equivalent to effects of allelic substitution for the markers, and from these both breeding values and Mendelian terms involved in the pedigree model can be estimated via Eqn S3.2, e.g., the parent average for *i* being $\tfrac{1}{2}(\mathbf{w}_{s(i)}^T + \mathbf{w}_{d(i)}^T)\hat{\mathbf{m}}$. The form of the model using $\mathbf{G}_{VR1}$ (as in Eqn 2 in the Materials & Methods) calculates the breeding values directly.

867 **Model III.** The model using the matrix relationship $\mathbf{G}_{FG}$ (Fernando & Grossman, 1989; Luan et al.
868 2012) can be viewed as embedding a genome-based haplotype model into the pedigree: $a_{i,k}^{s(i)} = a_{s(i),k}^{x}$
869 and $a_{i,k}^{d(i)} = a_{d(i),k}^{x}$ where $a_{i,k}^{g}$ is the value of the DNA segment for individual *i* at locus *k* for the
870 corresponding gamete, where *g* denotes paternal or maternal gamete, and in this simple pedigree
871 with only parents and offspring $x = p$ or *m*. For *n* offspring from *t* parents all genotyped with *k* markers,
872 the underlying model requires (in principle) to estimate 2*tk* parameters, the genetic differences
873 among the marked DNA chromosomal segments of the parents, but the limited amount of
874 recombination will greatly reduce the number of DNA segments that have to be estimated. The
875 whole-genome breeding value of an individual is then the sum of all DNA segments across all loci
876 and both haplotypes. Note that for offspring *i*,

$$a_{i,k} = \tfrac{1}{2}(a_{s(i),k}^{1} + a_{s(i),k}^{2}) + \tfrac{1}{2}(a_{d(i),k}^{1} + a_{d(i),k}^{2}) \pm \tfrac{1}{2}(a_{s(i),k}^{1} - a_{s(i),k}^{2}) \pm \tfrac{1}{2}(a_{d(i),k}^{1} - a_{d(i),k}^{2}) \quad \text{[Eqn S3.3]}$$

877 where the latter two terms are the Mendelian sampling term $r_{i,k}$ for locus *k* and are determined by the
878 true sample of gametes carried by *i*, e.g., if gamete 1 is sampled from *s(i)* and 2 from *d(i)* then the
879 3rd term in Eqn S3.3 is '+' and the 4th term is '-'. These notes show the connection of this breeding
880 value model to the pedigree-based model, where inheritance of alleles is not observed, and this
881 genome-based model, where alleles are observed although the IBS of the parental allele that is
882 inherited by an individual is not incorporated into the model. Within this model, there is an implicit
883 assumption that the genetic variance in the base and Mendelian sampling variance follows the same
884 relationship as in Model 1 where *F* is the observed inbreeding coefficient derived from tracking the
885 segregation from parent to offspring over generations. For the data in this study, with the parents
886 forming the base, F=0.
887 **Model IIIA.** Decomposition of $\mathbf{G}_{FG}$ into the estimation of the parental breeding values $a_{s(i)}$ and $a_{d(i)}$,
888 and the estimation of r_i the Mendelian terms using the genomic information introduces two
889 relationship matrices with associated components. Viewed as a ridge regression, the model requires
890 more parameters to be estimated as the chromosomal segments still need estimation as in $\mathbf{G}_{FG}$ and
891 also the parental transmitting abilities need estimation. However, there are two benefits for
892 estimation: (i) the parental breeding values, which explain half the genetic variance in the offspring,
893 are estimated directly by many fewer parameters and does not assume the markers will capture all

894 the variation, and (ii) there is no restriction that the Mendelian sampling variance is ½ of the total
895 genetic variance.
896
897 **References**
898 Fernando RL, Grossman M (1989) Marker assisted selection using best linear unbiased prediction.
899 *Genetics Selection Evolution* 21:467-477. https://doi.org/10.1051/gse:19890407
900 Luan T, Woolliams JA, Ødegård J, Dolezal M, Roman-Ponce SI et al. (2012) The importance of
901 identity-by-state information for the accuracy of genomic selection. *Genetics Selection Evolution*
902 44:28. doi:10.1186/1297-9686-44-28
903 VanRaden PM (2008) Efficient methods to compute genomic predictions. *Journal of Dairy Science*
904 91:4414-4423. https://doi.org/10.3168/jds.2007-0980
905 Wright, S (1922) Coefficients of inbreeding and relationship. *The American Naturalist* 56:330-338.
906 https://doi.org/10.1086/279872

Supplementary Information 4. Assessment of genotype by environment interaction (GxE) for the traits of interest.

Table S4.1 shows the evidence for GxE in some form for the traits studied and the potential to simplify the model from a bivariate model to a univariate model. This simplification has 4 requirements: (i) the genetic correlation between sites (r_g) equals 1; (ii) the genetic variances at the two sites are equal; (iii) block variances at the two sites are equal; and (iv) the environmental residuals at the two sites are equal. The first two of these may be regarded as GxE either by re-ranking or by scale differences. These requirements were fitted by constraining the model described in Eqn 1 in 4 progressive steps: (i) fixing r_g=0.999 in the bivariate model; (ii) constraining $v_{g,1} = v_{g,2}$; (iii) constraining $v_{c,1} = v_{c,2}$; and (iv) constraining $v_{e,1} = v_{e,2}$. At each step log L was recorded, and the final model was compared to a univariate model as a check. Test statistics (X^2) were calculated from twice the decrease in log L at each step. The test statistics were compared to χ^2 quantiles. The constraint r_g=0.999 is essentially on a boundary, so the initial decrease in X^2 is distributed under H_0 as $½\delta(0) + ½\chi_1^2$ where $\delta(0)$ is a point probability mass at X^2=0 (Self & Laing, 1987) and the total decrease over all 4 steps in X^2 is distributed under H_0 as $½\chi_3^2 + ½\chi_4^2$. Other quantiles were taken from χ_1^2.

There was consistent evidence for rejecting H_0: r_g=1 for B_{11}, V_{23} and MOR_{23}, i.e., across all relationship matrices. Taking requirements (i) and (ii) together results in a test statistic X^2 which under H_0: r_g =1 & $v_{g,1} = v_{g,2}$ is distributed $½\chi_1^2 + ½\chi_2^2$ which has a critical value of 5.140 for a type 1 error of 0.05. This H_0 was rejected using all relationship matrices for B_{11} and S_{11}, but was always accepted for H_{06}, D_{11}, MOE_{23} and DEN_{23}, while for other traits values close to the critical value were observed but not consistently above or below it. Across all the relationship matrices, there was no evidence to reject the fitting of a univariate model for H_{06} and MOE_{23} for these sites.

References

Self SG, K-Y Liang (1987) Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests under Nonstandard Conditions. *Journal of the American Statistical Association* 82:605-610. https://doi.org/10.1080/01621459.1987.10478472

Table S4.1. The log-likelihood of the full model and the increments of the test statistic X^2 in successive simplifications of the full model (*i*=0) in 4 steps (*i*=1,…,4) from full bivariate to a univariate model. The increments are described by the null hypothesis being tested by the simplified model, and the increment in $X^2 = 2*(\log L_{i-1} - logL_i)$. Under the null hypotheses, X^2 values are distributed as χ_1^2 except for $H_0 : r_g = 0.999$ and total X^2, which are mixtures $½\delta(0) + ½\chi_1^2$ and $½\chi_3^2 + ½\chi_4^2$ respectively, for which the 95%-iles are 2.705 and 8.763. Grey shading indicates P<0.05. Sub-tables (a), (b) and (c) are respectively from pedigree (**A**), genomic IBS ($\mathbf{G}_{VR1}$), and genomic IBD ($\mathbf{G}_{FG}$) information.

(a) **A**		Progressive model simplifications →				
	Full *log L₀*	$H_0 : r_g = 0.999$	$H_0 : v_{g,1} = v_{g,2}$	$H_0 : v_{c,1} = v_{c,2}$	$H_0 : v_{e,1} = v_{e,2}$	Total X^2
H_{06}	294.494	0.018	1.833	1.083	1.005	3.939
D_{11}	-1390.852	†-0.002	1.079	1.952	26.121	29.150
P_{11}	-2036.519	0.339	3.429	8.504	41.620	53.892
B_{11}	-856.000	10.095	15.165	1.263	0.159	26.682
S_{11}	-808.704	0.806	81.012	0.000	40.253	122.071
V_{23}	2372.995	4.143	0.446	0.016	0.019	4.624
MOE_{23}	-344.163	0.136	3.892	2.860	1.400	8.289
MOR_{23}	-3184.943	3.808	0.011	0.025	19.069	22.912
DEN_{23}	-5227.899	1.578	0.701	2.736	1.347	6.362

(b) $\mathbf{G}_{VR1}$		Progressive model simplifications →				
	Full *log L₀*	$H_0 : r_g = 0.999$	$H_0 : v_{g,1} = v_{g,2}$	$H_0 : v_{c,1} = v_{c,2}$	$H_0 : v_{e,1} = v_{e,2}$	Total X^2
H_{06}	295.691	†-0.001	4.025	0.954	0.580	5.558
D_{11}	-1387.333	†-0.001	1.685	1.804	27.644	31.132
P_{11}	-2035.071	0.959	4.726	9.084	40.806	55.576
B_{11}	-854.659	13.395	16.801	1.293	0.214	31.704
S_{11}	-814.259	1.424	76.936	0.001	42.699	121.061
V_{23}	2377.613	4.729	0.452	0.014	0.015	5.211
MOE_{23}	-343.490	0.000	3.253	2.270	2.083	7.605
MOR_{23}	-3182.082	4.826	0.150	0.006	18.999	23.980
DEN_{23}	-5219.196	2.664	0.020	3.412	4.416	10.511

(c) $\mathbf{G}_{FG}$		Progressive model simplifications →				
	Full *log L₀*	$H_0 : r_g = 0.999$	$H_0 : v_{g,1} = v_{g,2}$	$H_0 : v_{c,1} = v_{c,2}$	$H_0 : v_{e,1} = v_{e,2}$	Total X^2
H_{06}	294.835	0.016	1.961	0.956	0.966	3.900
D_{11}	-1390.971	†-0.001	0.822	1.863	25.564	28.247
P_{11}	-2037.678	0.546	4.291	8.662	41.240	54.740
B_{11}	-853.525	11.144	15.779	1.390	0.141	28.454
S_{11}	-809.089	0.686	84.580	0.000	40.484	125.751
V_{23}	2376.752	5.637	0.228	0.006	0.044	5.915
MOE_{23}	-335.949	0.305	3.795	1.851	1.913	7.863
MOR_{23}	-3177.568	4.951	0.505	0.226	17.316	22.998
DEN_{23}	-5216.385	2.898	0.278	4.013	2.341	9.530

† Negligible negative values arise due to ASReml-R binding r_g in the unconstrained model to a value marginally less than 0.999.

948

949 **Supplementary Information 5. Phenotypic and additive genetic variances from models.**

950 Tables S5.1 and S5.2 shows the phenotypic and additive genetic variances obtained from Models I,
951 II, and III when using all available data. These values are those used in the calculation of the
952 heritabilities shown in Table 4. Inspection shows that the estimates from using $\mathbf{G}_{VR1}$ are lower than
953 when using **A** when the heritabilities are greater, e.g., for the quality traits.

954

955 **Table S5.1**. The estimated phenotypic variances at each site for the traits considered estimated from
956 the three different relationship matrices using pedigree (**A**), genomic IBS ($\mathbf{G}_{VR1}$), and genomic IBD
957 ($\mathbf{G}_{FG}$) information used in Models I, II, and III using all available data. The genetic variance included
958 in the estimate is the variance within the elite population, so for site *j* the estimate is
959 $\hat{v}_{p,j} = k_G \hat{v}_{g,j} + \hat{v}_{c,j} + \hat{v}_{e,j}$ where $k_G = \overline{diag(\mathbf{G})} - \overline{\mathbf{G}}$ and **G** is the relevant relationship sub-matrix. The
960 S.E.s are in parentheses.

	Brecon			Kintyre		
Trait	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$
H_{06}	0.232 (0.013)	0.232 (0.013)	0.232 (0.013)	0.265 (0.016)	0.263 (0.016)	0.264 (0.016)
D_{11}	3.813 (0.236)	3.798 (0.233)	3.796 (0.233)	3.148 (0.257)	3.108 (0.249)	3.117 (0.251)
P_{11}	11.15 (0.72)	10.99 (0.68)	10.94 (0.67)	7.27 (0.59)	6.998 (0.46)	6.930 (0.46)
B_{11}	1.302 (0.074)	1.304 (0.075)	1.306 (0.075)	1.476 (0.098)	1.493 (0.097)	1.485 (0.098)
S_{11}	0.949 (0.053)	0.946 (0.052)	0.945 (0.052)	1.936 (0.154)	1.848 (0.120)	1.876 (0.130)
V_{23}	0.011 (0.001)	0.011 (0.001)	0.011 (0.001)	0.010 (0.001)	0.010 (0.001)	0.010 (0.001)
MOE_{23}	0.761 (0.065)	0.676 (0.042)	0.714 (0.048)	0.936 (0.089)	0.817 (0.053)	0.869 (0.061)
MOR_{23}	48.08 (4.11)	44.66 (2.95)	46.59 (3.25)	63.02 (4.91)	59.85 (3.81)	62.30 (4.29)
DEN_{23}	1441 (135)	1283 (84)	1322 (91)	1258 (118)	1127 (76)	1155 (80)

961

962 Model IIIA (see Eqn 4 in the Materials & Methods) was also fitted to all the available data and was
963 used to estimate the Mendelian sampling variance estimated from the markers with the variances of
964 the transmitting abilities of the parents. In particular the fractions $\mathbf{V}_M[s,s]/(4\mathbf{V}_T[s,s])$ were calculated
965 where s was Brecon and Kintyre. Given the parameterisation of Eqn 4 these ratios are expected to
966 be 1 with classical assumptions, as the Mendelian sampling variance is expected to contribute the
967 same amount of additive genetic variance as the parental transmitting abilities with random mating.
968 The denominator was chosen as being the more reliably estimated. Table S5.3 shows the estimates
969 of the ratios along with the 'z-ratios' for the components (i.e., $\hat{v} / S.E.(\hat{v})$). At Brecon, the Mendelian
970 sampling variance was not detected for H_{06} and S_{11}. The pooling of values over the traits followed

DerSimonian & Laird (1986) but excluded the undetectable variances, so potentially biasing the results upwards at Brecon. The pooled values were 0.686 (S.E. 0.194) and 0.653 (S.E. 0.154) for Brecon and Kintyre respectively, and pooled over sites the estimate was 0.666 (S.E. 0.121), over 2 S.E.s from the classical value of 1.

Table S5.2. The estimated genetic variances at each site for the traits considered estimated from the three different relationship matrices using pedigree (**A**), genomic IBS ($\mathbf{G}_{VR1}$), and genomic IBD ($\mathbf{G}_{FG}$) information used in Models I, II, and III using all available data. The genetic variance included in the estimate is the variance within the elite population, so for site *j* the estimate is $k_G\hat{v}_{g,j}$ where $k_G = \overline{diag(\mathbf{G})} - \overline{\mathbf{G}}$ and **G** is the relevant relationship sub-matrix. The S.E.s are in parentheses.

	Brecon			Kintyre		
Trait	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$
H_{06}	0.031 (0.013)	0.025 (0.009)	0.029 (0.012)	0.054 (0.019)	0.055 (0.017)	0.051 (0.017)
D_{11}	0.500 (0.198)	0.422 (0.150)	0.442 (0.166)	0.717 (0.227)	0.656 (0.175)	0.607 (0.173)
P_{11}	2.50 (0.91)	1.87 (0.61)	1.79 (0.63)	3.90 (1.08)	2.98 (0.61)	2.89 (0.67)
B_{11}	0.167 (0.074)	0.178 (0.073)	0.189 (0.079)	0.485 (0.153)	0.492 (0.128)	0.498 (0.14)
S_{11}	0.081 (0.044)	0.068 (0.038)	0.063 (0.036)	1.104 (0.284)	0.756 (0.155)	0.906 (0.195)
V_{23}	0.003 (0.001)	0.003 (0.001)	0.003 (0.001)	0.003 (0.001)	0.003 (0.001)	0.003 (0.001)
MOE_{23}	0.492 (0.125)	0.264 (0.049)	0.374 (0.068)	0.705 (0.172)	0.364 (0.062)	0.522 (0.088)
MOR_{23}	30.11 (7.76)	19.76 (3.73)	24.67 (4.68)	32.99 (9.00)	22.94 (4.77)	31.28 (6.40)
DEN_{23}	1145 (263)	643 (101)	804 (128)	992 (231)	668 (99)	745 (117)

Table S5.3. The estimates of Mendelian sampling variance as a fraction of the variance explained by the sires and dams obtained from Model IIIA together with S.E. in parentheses, and the z-ratios (i.e., $\hat{v} / S.E.(\hat{v})$) for the components at each site.

	Brecon	**Kintyre**	**Brecon**		**Kintyre**	
	$\hat{v}_M / (4\hat{v}_T)$	$\hat{v}_M / (4\hat{v}_T)$	$\hat{v}_M / S.E.(\hat{v}_M)$	$\hat{v}_T / S.E.(\hat{v}_T)$	$\hat{v}_M / S.E.(\hat{v}_M)$	$\hat{v}_T / S.E.(\hat{v}_T)$
H_{06}	0.000 (-)*	1.283 (1.221)	(-)*	2.48	1.19	2.76
D_{11}	0.965 (1.763)	0.291 (0.676)	0.57	2.49	0.44	3.12
P_{11}	3.188 (1.904)	0.336 (0.396)	2.47	2.56	0.90	3.56
B_{11}	1.928 (2.229)	1.019 (0.777)	0.98	2.17	1.52	3.12
S_{11}	0.000 (-)*	0.369 (0.370)	(-)*	1.82	1.07	3.81
V_{23}	0.644 (0.801)	1.220 (0.836)	0.85	3.33	1.68	3.22
MOE_{23}	1.117 (0.496)	0.704 (0.345)	3.14	3.67	2.58	3.85
MOR_{23}	0.557 (0.368)	1.204 (0.556)	1.74	3.76	2.95	3.55
DEN_{23}	0.547 (0.281)	0.730 (0.302)	2.39	4.07	3.21	4.06
Pooled	0.686 (0.194)	0.653 (0.154)				

* REML constrained the Mendelian sampling variance to a negligibly small value with no S.E.

986

987 **References**

988 DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Controlled Clinical Trials*
989 7:177-188. https://doi.org/10.1016/0197-2456(86)90046-2

Supplementary Information 6. Accuracy of genomic and pedigree predictions with increased resources.

Formulae to predict the accuracy of EBV have been developed and refined in several studies (e.g., Dekkers (2007), Goddard et al. 2007; Daetwyler et al. 2008; Elsen, 2016, 2017; Dekkers et al. 2021; Cuyabano et al. 2024) and already used in tree breeding settings (Grattapaglia & Resende, 2011; Papin et al. 2024). Here the approach used to predict the potential future accuracy ($r_{\hat{g},g}$) of a new unrecorded individual from within a population (e.g., a sapling with unknown parents) will follow Dekkers (2007) and Daetwyler et al. (2008).

Genomic prediction. This will focus on prediction using $\mathbf{G}_1$ which is based on ridge regression on marker counts, and consider how an initial accuracy of r_0 with N_0 records will increase *k* times as many records. The increase of records is assumed to be take place in the same population. An approach to the accuracy of genomic prediction for the breeding value ($r_{\hat{g},g}$) of a new unrecorded individual from within a population, e.g., a sapling with unknown parents, is offered by. The starting point is given by considering the trait as combination of a part of the breeding value that is accessible for prediction by the markers *g(M)* and the residual part (Dekkers, 2007). Then

$$r_{\hat{g},g} = r_{\hat{g},g(M)} r_{g(M),g} \qquad \text{[Eqn S6.1]}$$

where $r_{g(M),g}$ is the maximum accuracy with the markers, with $r^2_{g(M),g}$ being the fraction of the genetic variance captured by the markers; and $r_{\hat{g},g(M)}$ is the accuracy of predicting the accessible component. It is clear that the maximum accuracy achievable will be $r_{g(M),g}$ and in this study the estimates from Supplementary Information 5 estimates the maximum accuracy as $r_{g(M),g} = \sqrt{0.833} = 0.913$ with the current marker set.

Daetwyler et al. (2008) assume $r_{g(M),g} = 1$ the accuracy is given by

$$r_{\hat{g},g} = \sqrt{(h^2 N / M_e) / [1 + (h^2 N / M_e)]} \qquad \text{[Eqn S6.2]}$$

where h^2 is the heritability of the trait, N is the number of records used to parameterise the marker effects in the genomic prediction, and M_e is termed the number of independent segments of the population genome. M_e is a parameter that accounts for the length of the genome and the degree of linkage disequilibrium among the loci across the genome but it is difficult to estimate, particularly for a species which has been domesticated for only two generations with a declining effective population size. The term h^2N is a measure of the genetic information available for the prediction and the h^2N / M_e is a measure of information per independent segment.

Eqns S6.1 and S6.2 can be combined and transformed to give

$$r_{\hat{g},g}^{-2} = r_{g(M),g}^{-2} + (M_e / h^2)N^{-1} \qquad \text{[Eqn S6.3]}$$

which can be parameterised from the existing data: $r_{g(M),g}^2 = 0.833$ and if the accuracy with existing data with N_0 records is r_0 then $(M_e / h^2)N_0^{-1} = r_0^{-2} - 1.20$. With kN_0 records then $r_{\hat{g},g}^{-2} = 1.20 + \left(r_0^{-2} - 1.20\right) / k$: e.g., if $r_0 = 0.5$ with N_0 records, then doubling the number of records, i.e., k=2 predicts $r_{\hat{g},g}^{-2}$ = 1.2 + (4-1.2)/2 = 2.6 and $r_{\hat{g},g}$ = 0.62.

Pedigree accuracy. With a pedigree model using **A,** a similar approach can be taken to predict the accuracy with an increase in resources. However, it should be noted that if the number of records is increased only by increasing the number of parents, then no increase in accuracy will be achieved. The increase in accuracy is achieved by increasing the number of offspring per parent. So here N_0 is the number of offspring, and a k-fold increase in resources is achieved by the number of offspring increasing to kN_0. An established formula (Gorjanc et al. 2015; Mrode & Pocrnic 2024) for the accuracy of progeny testing a parent ($r_{\hat{g},g}[parent]$) is

$$r_{\hat{g},g}[parent] = \sqrt{\lambda N / (1 + \lambda N)} \qquad \text{[Eqn S6.4]}$$

where $\lambda = ¼h^2 / (1 - ¼h^2)$ which is a measure of signal-to-noise ratio for the information on the parent and N the number of offspring. In this study the information on the parent is not a classical

progeny test but includes both full-sibs and half-sibs. Nevertheless, the form of Eqn [S6.4] justifies the assumption that for some λ

$$r_{\hat{g},g}^{-2}[parent] = 1 + (1/\lambda)N^{-1} \qquad \text{[Eqn S6.5]}$$

and that the full genetic variance is captured by the pedigree so its maximum accuracy is 1. If the accuracy of the parent is $r_0[parent]$ with N_0 records, then following the argument above with kN_0 records $r_{\hat{g},g}^{-2}[parent] = 1 + (r_0^{-2}[parent] - 1)/k$. This will increase to 1 as k increases. However, when predicting a new offspring as in validation, the offspring is predicted from the average of its parents. With random mating

$$r_{\hat{g},g}^{2} = \tfrac{1}{4}(r_{\hat{g},g}^{2}[sire] + r_{\hat{g},g}^{2}[dam]) = \tfrac{1}{2} r_{\hat{g},g}^{2}[parent] \qquad \text{[Eqn S6.6]}$$

assuming approximately equal information per parent as in this study. Therefore, the maximum possible accuracy of the validation is $\sqrt{\tfrac{1}{2}} = 0.707$ and $r_{\hat{g},g}^{-2} = 2[1 + (\tfrac{1}{2} r_0^{-2} - 1)/k]$. If with pedigree, the validation accuracy of an offspring is $r_0 = 0.5$ with N_0 offspring for a parent, doubling the number of offspring gives an accuracy of 0.577.

References

Cuyabano BCD, Boichard D, Gondro C (2024) Expected values for the accuracy of predicted breeding values accounting for genetic differences between reference and target populations. *Genetics Selection Evolution* 56:15. https://doi.org/10.1186/s12711-024-00876-9

Daetwyler HD, Villanueva B, Woolliams JA (2008). Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS One* 3:10. https://doi.org/10.1371/journal.pone.0003395

Dekkers JCM (2007) Prediction of response to marker-assisted and genomic selection using selection index theory. *Journal of Animal Breeding and Genetics* 124:331-341. https://doi.org/10.1111/j.1439-0388.2007.00701.x

Dekkers JCM, Su H, Cheng J (2021) Predicting the accuracy of genomic predictions. Genetics Selection Evolution 53:55. https://doi.org/10.1186/s12711-021-00647-w

Elsen J-M (2016) Approximated prediction of genomic selection accuracy when reference and candidate populations are related. *Genetics Selection Evolution* 48:18. https://doi.org/10.1186/s12711-016-0183-3

Elsen J-M (2017) An analytical framework to derive the expected precision of genomic selection. *Genetics Selection Evolution* 49:95. https://doi.org/10.1186/ s12711-017-0366-6

Goddard ME, Hayes BJ. (2007) Genomic selection. *Journal of Animal Breeding and Genetics* 124:323-330. https://doi.org/10.1111/j.1439-0388.2007.00702.x

Gorjanc G, Bijma P, Hickey JM (2015) Reliability of pedigree-based and genomic evaluations in selected populations. *Genetics Selection Evolution* 47:65. https://doi.org/10.1186/s12711-015-0145-1

Grattapaglia D, Resende MDV (2011) Genomic selection in forest tree breeding. *Tree Genetics and Genomes* 7:241–255. https://doi.org/10.1007/ s11295-010-0328-4

Mrode R, Pocrnic I (2024) Linear Models for the Prediction of the Genetic Merit of Animals, CABI, 4th edition.
Papin V, Gorjanc G, Pocrnic I, Bouffier L, Sanchez L (2024) Unlocking genome-based prediction and selection in conifers: the key role of within-family prediction accuracy illustrated in maritime pine (Pinus pinaster Ait.). *Annals of Forest Science* 81:52. https://doi.org/10.1186/s13595-024-01269-0

Supplementary Information 7. Estimates of bias.

The bias was estimated as the slope of the regression of recorded phenotype on EBV and a slope of 1 indicates no bias in the EBVs for estimating differences in true breeding values: a value <1 indicates over-estimation of differences in true breeding values by the EBVs, and >1 indicates under-estimation of differences in true breeding values by the EBVs. Table S7.1 shows very little evidence of bias in the EBVs with the mean slope pooled over blocks within 1.5 S.E. of 1, bar a single exception (S_{11} at Brecon in Model IIIA, which used $\mathbf{M}_{FG}$). When pooled over sites, shown in Table S7.2, all estimates were within 1.5 S.E. of 1.

Table S7.1. Bias of EBVs estimated as the slope of the regression of recorded phenotype on EBV by cross validation using four different relationship matrices with pedigree (**A**), genomic IBS ($\mathbf{G}_{VR1}$), and genomic IBD ($\mathbf{G}_{FG}$ and $\mathbf{M}_{FG}$) information for each site. The S.E.s are given in parentheses.

	Brecon				Kintyre			
Trait	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$	$\mathbf{M}_{FG}$	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$	$\mathbf{M}_{FG}$
H_{06}	1.037 (0.164)	1.039 (0.165)	1.034 (0.167)	0.988 (0.161)	0.951 (0.148)	0.935 (0.136)	0.958 (0.152)	0.966 (0.149)
D_{11}	0.922 (0.137)	0.859 (0.133)	0.923 (0.134)	0.864 (0.142)	0.974 (0.098)	1.016 (0.093)	0.998 (0.099)	1.048 (0.104)
P_{11}	0.858 (0.110)	0.873 (0.130)	0.872 (0.129)	0.815 (0.134)	1.027 (0.074)	1.031 (0.072)	1.035 (0.074)	1.047 (0.077)
B_{11}	0.948 (0.189)	0.932 (0.176)	0.928 (0.175)	0.973 (0.193)	0.962 (0.095)	0.942 (0.092)	0.962 (0.092)	0.918 (0.099)
S_{11}	0.765 (0.199)	0.800 (0.205)	0.710 (0.206)	0.672 (0.205)	1.021 (0.083)	1.040 (0.081)	1.027 (0.083)	1.060 (0.084)
V_{23}	0.955 (0.086)	0.937 (0.082)	0.950 (0.089)	0.955 (0.090)	0.951 (0.117)	0.954 (0.106)	0.987 (0.113)	1.125 (0.098)
MOE_{23}	0.953 (0.066)	0.981 (0.066)	0.946 (0.061)	0.943 (0.067)	0.984 (0.066)	1.002 (0.079)	0.969 (0.065)	1.015 (0.062)
MOR_{23}	0.942 (0.064)	0.961 (0.061)	0.939 (0.061)	0.945 (0.064)	1.044 (0.092)	1.041 (0.088)	1.017 (0.083)	1.052 (0.089)
DEN_{23}	0.991 (0.053)	1.004 (0.053)	0.996 (0.052)	1.007 (0.055)	1.037 (0.060)	1.024 (0.050)	1.042 (0.062)	1.090 (0.064)

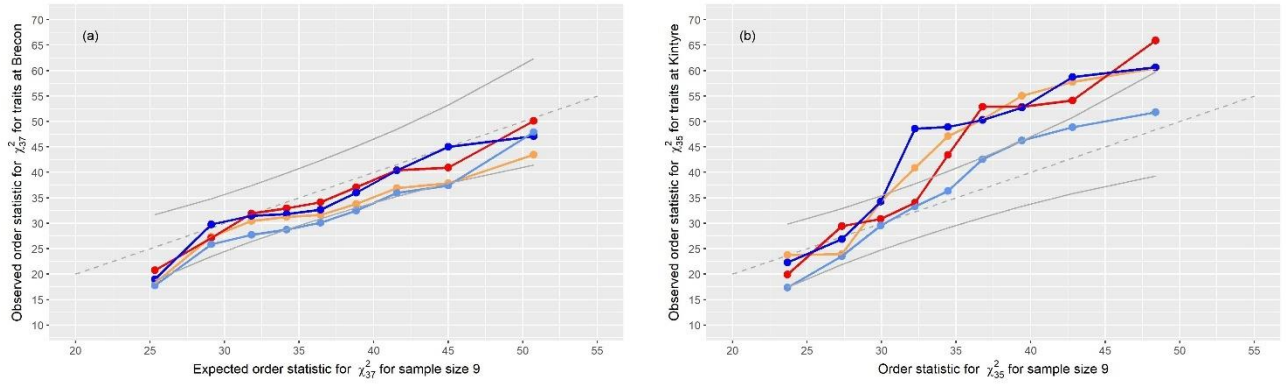
In the cross validation used to produce Table S7.1, the slopes were estimated for each block across both sites. The meta-analysis, following DerSimonian & Laird (1986), produced the pooled estimates shown and standard errors, together with an assessment of variation between blocks over and above the sampling variance for a block. For Brecon there was no evidence of such variance, however there was evidence for such variance at Kintyre. This was most evident for H_{06} and S_{11} where significant ($P<0.05$) variation was indicated with all 4 relationship matrices, but some relationship matrices also identified variance in V_{23}, MOE23 and MOR23. The critical value for each relationship

matrix for each site was not amended for carrying out 9 tests over the 9 traits. The plots of the order statistics for the 9 chi-squared values at Brecon and Kintyre against the expected order statistic is shown in Fig S7.1, which only approximates a Q-Q plot primarily because the 9 traits are treated as independent whereas they are genetically and environmentally correlated. The plots reflect the observation of variation among blocks at Kintyre (for at least one trait), but not at Brecon.

Table S7.2. Bias of EBVs estimated as the slope of the regression of recorded phenotype on EBV by cross validation using four different relationship matrices with pedigree (**A**), genomic IBS ($\mathbf{G}_{VR1}$), and genomic IBD ($\mathbf{G}_{FG}$ and $\mathbf{M}_{FG}$) information pooled over sites. The S.E.s are given in parentheses.

Trait	**A**	$\mathbf{G}_{VR1}$	$\mathbf{G}_{FG}$	$\mathbf{M}_{FG}$
H_{06}	0.981 (0.110)	0.967 (0.104)	0.985 (0.112)	0.968 (0.108)
D_{11}	0.956 (0.080)	0.964 (0.076)	0.972 (0.080)	0.983 (0.084)
P_{11}	0.972 (0.061)	0.971 (0.064)	0.975 (0.063)	0.957 (0.068)
B_{11}	0.959 (0.085)	0.940 (0.082)	0.955 (0.082)	0.928 (0.086)
S_{11}	0.990 (0.074)	1.014 (0.073)	0.992 (0.074)	1.014 (0.075)
V_{23}	0.952 (0.071)	0.945 (0.066)	0.969 (0.071)	1.032 (0.066)
MOE_{23}	0.969 (0.045)	0.995 (0.052)	0.957 (0.044)	0.983 (0.045)
MOR_{23}	0.994 (0.053)	0.996 (0.051)	0.976 (0.049)	0.995 (0.052)
DEN_{23}	1.014 (0.038)	1.014 (0.036)	1.018 (0.038)	1.048 (0.040)

Figure S7.1. The order statistic of the 9 chi-squared values obtained from DerSimonian and Laird (1986) for the 9 traits, plotted against the expected of the order statistic for a sample of 9 chi-squared values with the same degrees of freedom: (a) for Brecon, with χ^2_{37}, and (b) for Kintyre, with χ^2_{35}. The colours are: pedigree (**A**), tan; genomic IBS ($\mathbf{G}_{VR1}$), red; genomic IBD ($\mathbf{G}_{FG}$), blue; and genomic IBD ($\mathbf{M}_{FG}$), light blue. The dashed grey line is y=x, and the solid grey lines show the 5%-ile and 95%-ile for the corresponding order statistic.



References

DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Controlled Clinical Trials* 7:177-188. https://doi.org/10.1016/0197-2456(86)90046-2