ADVANCING BLACK SPRUCE (*PICEA MARIANA*) BREEDING THROUGH GENOMIC SELECTION: A COMPARATIVE ANALYSIS OF MODELS USING PEDIGREE AND GENOMIC MARKER INFORMATION

by

Jesse Edward Milani, RPF

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

FACULTY OF NATURAL RESOURCE MANAGEMENT
(Forestry)

LAKEHEAD UNIVERSITY

THUNDER BAY, ONTARIO

2025

LIBRARY RIGHTS STATEMENT

In presenting this thesis in partial fulfillment of the requirements for the MScF degree at Lakehead University in Thunder Bay, I agree that the University will make it freely available for inspection.

This thesis is made available by my authority solely for the purpose of private study and may not be copied or reproduced in whole or in part (except as permitted by the Copyright Laws) without my written authority.

Signature: _____

Date: _____

Abstract

Milani, J. E. 2025. **Keywords:** genomic selection, black spruce, tree improvement, half-sibs, growth traits, genetic markers, genetic gain, heritability.

Long generation times in forest trees constrain the pace of genetic improvement necessary to sustain productivity under climate change. Genomic selection offers a promising approach to accelerate breeding gains in long-lived species like black spruce (*Picea mariana*). In this study, we evaluated genomic selection models using data from a long-term half-sib progeny trial in the Lake Nipigon West breeding zone of northern Ontario. A subset of 1194 trees from 70 families was genotyped using two platforms: a SNP array (16,217 SNPs) and a genotyping-by-sequencing approach based on RADseq (10,626 SNPs). Growth traits—including height, diameter at breast height (DBH), growth rate, and volume—were measured at multiple time points.

We compared three animal models differing in their relationship matrices: pedigree-based (ABLUP), genomic-based (GBLUP), and a hybrid model integrating both pedigree and genomic information (HBLUP). The HBLUP model consistently produced the most accurate heritability estimates and the smallest prediction errors for key growth traits such as volume and DBH, likely due to its ability to incorporate both genotyped and ungenotyped individuals. Genomic models (GBLUP and HBLUP) outperformed pedigree-based models, highlighting the value of genomic information for improving selection efficiency.

While early height has traditionally served as a proxy for long-term growth, its low heritability in this study suggests caution in its use as a sole selection criterion. Instead, height may be better incorporated as part of multi-trait selection indices to capture its environmental responsiveness, particularly during early testing stages.

Among genotyping platforms, SNP chips consistently outperformed RADseq, indicating their preference when budget allows, though RADseq remains a cost-effective alternative that could benefit from complementary strategies such as imputation or hybrid integration.

Overall, our findings support the practical integration of genomic selection into black spruce breeding programs. By aligning genotyping strategies and model choice with specific trait characteristics and breeding goals, programs can accelerate genetic gain, reduce breeding cycle time, and enhance forest adaptability under future environmental challenges.

CONTENTS

List of Figures

List of Tables

List of abbreviations

BLUP	best linear unbiased prediction
BV	breeding value
CV	cross-validation
DBH10	diameter at breast height at 10 years
EBV	estimated breeding value
GBLUP	genomic best linear unbiased predictor
GBS	genotyping-by-sequencing; a class of reduced-representation sequencing methods
GEBV	genomic estimated breeding value
GR	growth rate
GS	genomic selection
GxE	genotype-by-environment interaction
HBLUP	hybrid best linear unbiased prediction
HS	half-sibs
HT	height
HT5	height at 5 years
HT10	height at 10 years
LNW	Lake Nipigon West Breeding Zone
LD	linkage disequilibrium
MAF	minor allele frequency
OP	open-pollinated
PA	predictive ability
PACC	prediction accuracy

RADseq	restriction site-associated DNA sequencing; a specific type of GBS using restriction enzymes
SNP	single-nucleotide polymorphism
VOL10	volume at 10 years

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to everyone who has supported me throughout the journey of completing this thesis.

This work was conducted on the traditional and ancestral territories of the Anishinaabe peoples, including the Red Rock Indian Band and other communities of the Robinson-Superior Treaty. I acknowledge this land with respect and gratitude, recognizing the enduring relationship Indigenous peoples have with these territories, and the responsibility I share in honouring and learning from that relationship.

First and foremost, I am deeply grateful to my thesis advisors, Dr. Ashley Thomson and Dr. Patrick Lenz. Dr. Thomson, your expertise, patience, and thoughtful guidance have been integral to the development of this work. Your encouragement and constructive feedback pushed me to think critically and refined my research in ways I never could have done on my own. Dr. Lenz, your insights and suggestions have significantly shaped my approach, and I truly appreciate your dedication to helping me grow as a researcher. I am incredibly fortunate to have had both of you as mentors.

I would also like to thank Jean-Philippe Laverdière for his assistance with the analysis portion of this thesis. Your willingness to collaborate and provide guidance was invaluable, and I am grateful for the time and effort you put into helping me refine the analytical aspects of this work.

A heartfelt thank you to my partner, Ashlyn Kopanski, whose unwavering love and support have meant the world to me. You were there for me through all the highs and lows of this journey, always offering encouragement, patience, and a sense of perspective when things felt overwhelming. I could not have done this without you, and I am so grateful for everything you've done.

Finally, I want to acknowledge my dog, Wesley, whose playful energy and constant companionship were a much-needed source of comfort and joy during long work hours. His presence was a reminder to take breaks and to enjoy the little moments along the way.

INTRODUCTION

In most boreal tree species, slow growth and delayed trait expression impede breeding progress, limiting genetic gain and the capacity to adapt to climate change. Traditional breeding approaches rely heavily on long-term field trials and pedigree-based selection, which are both time-consuming and costly to implement.

Genomic selection (GS) offers a powerful alternative to traditional breeding by enabling early selection of superior individuals based on genome-wide marker data. First proposed by Meuwissen, Hayes, and Goddard (2001), GS is based on the principle that genome-wide markers capture the combined effects of loci underlying complex traits through linkage disequilibrium (LD) (Meuwissen et al., 2001). This allows for the prediction of genomic estimated breeding values (GEBVs) without direct measurement of phenotypes. GS is particularly valuable for traits that are difficult, expensive, or time-consuming to assess in field trials, such as drought tolerance or pest resistance, and has the potential to greatly increase selection intensity and genetic gain per unit time (Grattapaglia & Resende, 2011) .

Among GS methods, genomic best linear unbiased prediction (GBLUP) is the most widely used (Liu et al., 2022). It replaces the pedigree-based relationship matrix (A-matrix) used in traditional best linear unbiased prediction (ABLUP) with a marker-derived realized genomic relationship matrix (G-matrix), which captures the actual proportion of alleles shared between individuals. This enables GBLUP to account for Mendelian sampling variance and more accurately estimate additive genetic relationships (Strandén & Garrick, 2009; VanRaden, 2008). Empirical studies in forest trees, such as white spruce (*Picea glauca*) and western redcedar (*Thuja plicata*) have

shown that GBLUP frequently outperforms ABLUP in predicting breeding values for traits such as growth, wood density, and phenology (Beaulieu et al., 2014; Gamal El-Dien et al., 2022).

Despite these advantages, the cost of genotyping all individuals in a breeding population remains a barrier to large-scale implementation of genomic selection. Hybrid best linear unbiased prediction (HBLUP) addresses this limitation by integrating pedigree and genomic information into a single H-matrix, enabling breeding value estimation for both genotyped and ungenotyped individuals (Ratcliffe et al., 2017; Simiqueli et al., 2023). This approach is particularly valuable in conifer breeding programs that rely on open-pollinated families, where assumptions of half-sib relatedness are often violated. In practice, such families may include full-sibs, half-sibs, and selfed individuals, which can bias genetic parameter estimates when using pedigree information alone (El-Kassaby et al., 2024). By directly quantifying allele sharing, genomic data can correct for pedigree error and better capture both recent and historical relatedness (Godbout et al., 2017).

To support GS, several genotyping technologies can be used to generate dense genome-wide marker data, including single-nucleotide polymorphism (SNP) chips and genotyping-by-sequencing (GBS) approaches such as restriction site-associated DNA sequencing (RADseq) (Scheben et al., 2017). RADseq is widely adopted in forestry due to its relatively low cost and ability to generate thousands of SNPs without prior genomic resources (Andrews et al., 2016; Davey & Blaxter, 2010; Tong et al., 2020; Ulaszewski et al., 2021). SNP chips, in contrast, offer consistent and targeted genomic coverage but are typically more expensive and less adaptable to species with limited genomic resources (Kim et al., 2022; You et al., 2018). The choice of genotyping

platform can significantly influence model performance, data quality—such as marker density, missing data rates, and genotyping accuracy—and downstream applications, such as diversity analysis and genomic prediction (Ma et al., 2022; Roorkiwal et al., 2018). These studies highlight differences in marker density, imputation accuracy, and genomic coverage that can affect predictive ability and the efficiency of selection.

Black spruce (*Picea Mariana* [Mill.] B.S.P) is a strong model for genomic selection (GS) in forestry due to its annotated reference genome and demonstrated genetic relatedness within populations (Lo et al., 2023). It is widely distributed across North America and exhibits substantial genetic variation in both growth and adaptive traits, such as growth and phenology (Moreau et al., 2020; Thomson et al., 2009) . It's a predominantly outcrossed mating system and wind-dispersed pollen promote random mating and maintain large effective population sizes (Isabel et al., 1995). Although black spruce tree improvement programs have been active since the 1980s, progress has been limited, particularly in northwestern Ontario, where most programs have only reached the second generation of selection (Thomas et al., 2024). Breeding cycles often exceed 30 years due to the need for 15 to 20 years of field testing to evaluate mature traits (Chang et al., 2019; Mullin et al., 2011).

Implementing GS in black spruce offers an opportunity to overcome this bottleneck. A proof-of-concept study by Lenz et al. (2017) demonstrated that GS can achieve promising predictive accuracy for growth and wood quality traits using full-sib families from Quebec genotyped with a SNP chip. However, no GS studies have yet been conducted in Ontario, where breeding programs are typically based on open-pollinated (half-sib) families and operate under more limited funding. In breeding contexts with diverse or half-sib germplasm, GBS is often preferred over SNP arrays

due to its cost-effectiveness, broader variant detection, and avoidance of ascertainment bias, making it better suited for genetically diverse populations (Badenes et al., 2016). Differences in family structure and genotyping platform may influence the accuracy and utility of genomic prediction models, underscoring the need to evaluate GS performance under conditions that reflect operational realities (Lenz et al., 2017; Werner et al., 2020). This study represents the first application of genomic selection (GS) in Ontario black spruce, evaluating the performance of prediction models using different relationship matrices: pedigree-based (A), genomic-based (G), and hybrid (H). It also compares two genotyping methods, microarray and GBS, offering practical insights into the feasibility of GS under resource-constrained breeding conditions.

In this study, the relative effectiveness of different approaches for estimating breeding values for growth traits in black spruce was evaluated, with particular attention to traditional pedigree-based models (ABLUP) and genomic methods (GBLUP and HBLUP). It was expected that genomic approaches, particularly GBLUP and HBLUP, would provide more accurate and reliable estimates of breeding values than the pedigree-based method. High-throughput genotyping technologies, including restriction site-associated DNA sequencing (RADseq) and array-based SNP genotyping, were also expected to enhance the estimation of genetic gain and improve the overall efficiency of breeding strategies. By comparing the performance of prediction models and genotyping platforms, this research aims to provide practical insights that support the integration of genomic tools into operational breeding programs for black spruce and other commercially important forest tree species.

METHODS

Experimental Design and Phenotypic Data Acquisition

This study focused on a subset of trees from a first-generation open-pollinated progeny trial established in 1988 as part of the Ontario Tree Improvement Board's black spruce tree improvement program in the Lake Nipigon West (LNW) breeding zone of northwestern Ontario. The trial at Block #3 (48.91°N, -89.95°W) includes 400 open-pollinated (half-sib) families derived from plus-trees selected across the LNW breeding zone. The site was planted using a randomized complete block design with 32 replicates. Each replicate (block) was divided into four quadrants, and the 400 families were randomly assigned to four sets of 100 families (A–D), such that one tree from each family was represented in every block, totalling 32 trees per family.

Phenotypic data were collected in 1993 and 1998. Height at five years (H5) was measured in 1993, while height (H10) and diameter at breast height (DBH10) were recorded in 1998 (Fu, 2000). Growth rate (GR) was calculated as the average annual increase in height from age 5 to 10. Volume at age 10 (VOL10) was estimated using a standard volume equation for black spruce (Honer, 1967) using a form factor of 0.45, appropriate for young black spruce (Fradette et al., 2021).

$$\mathrm{VOL10} = \frac{\pi}{4} * (\frac{DBH10}{2})^2 * H10 * Form\ Factor \qquad (1)$$

For this study, a subset of 70 open-pollinated families was selected from the 400 represented in the progeny trial Block #3 (Figure 1). The full set of 400 families originated from "plus-tree" selections in wild stands. The 70 families analyzed in this study were selected from the families previously chosen for the LNW second-generation breeding population, based on earlier analysis by Fu (2000). These families generally

exhibited average or above-average breeding values for growth traits, as they were drawn from the top-performing group in the original evaluation. Due to financial constraints, it was not feasible to genotype all families in the trial; therefore, this subset was chosen to balance cost considerations with representation across the spectrum of observed performance.

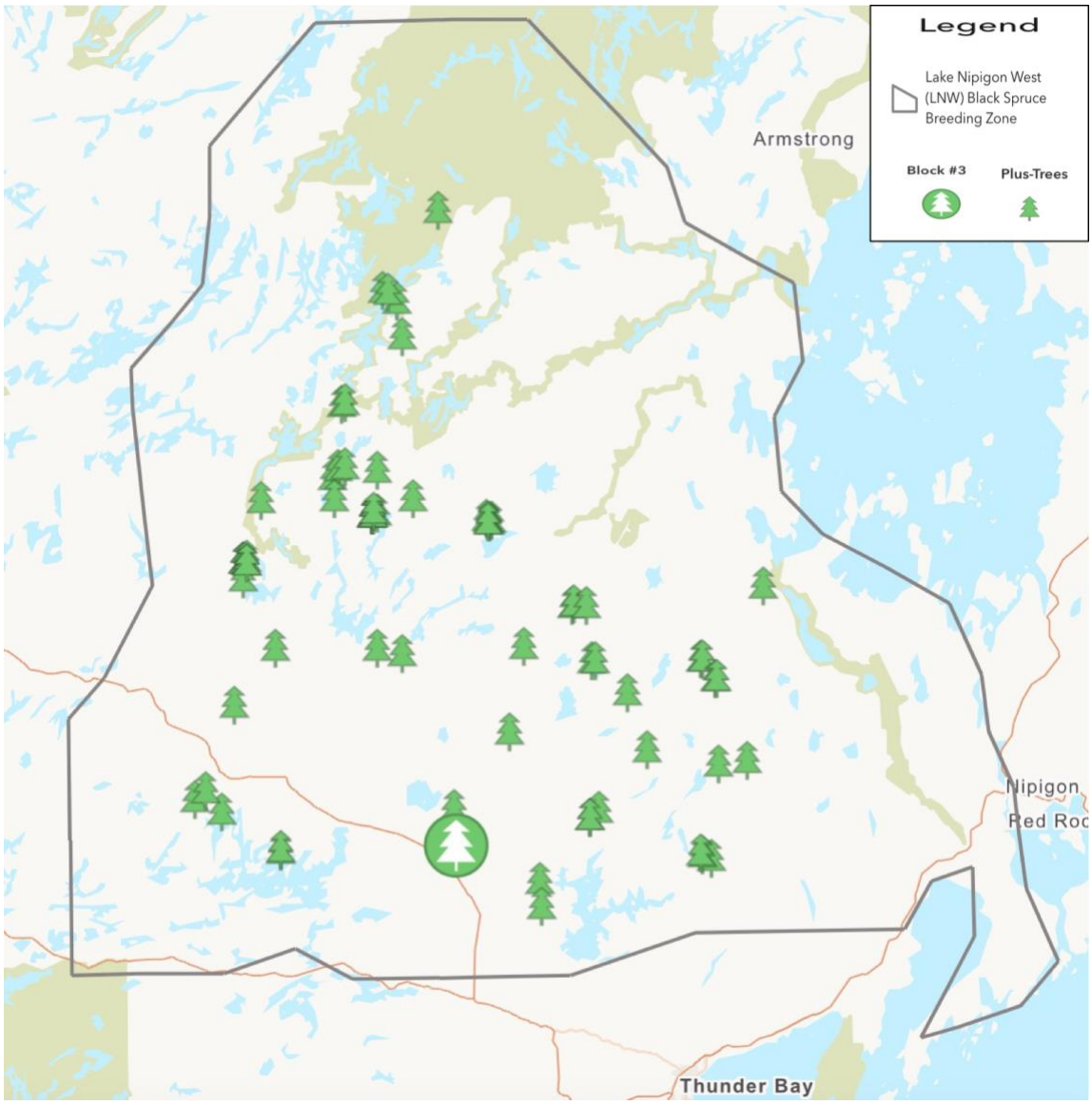

Figure 1 The Lake Nipigon West (LNW) Black Spruce Breeding Zone with test site Block #3 and subset 70 families plus-tree origin locations.

Genotypic Data Acquisition and Processing

Newly flushed bud tissue was collected from the upper crown of each tree in summer 2020 using an extendable pruning pole. Buds were immediately placed in silica beads to desiccate the tissue and preserve DNA. A total of 700 individuals were genotyped, representing an average of 10 trees per family for each of the 70 sampled families. Genomic DNA was extracted using the Macherey-Nagel NucleoSpin Plant II kit, with quality and concentration assessed via NanoDrop spectrophotometry and Qubit fluorometer, respectively. Samples were normalized to 20 ng/µL and shipped in 96-well plates to the Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, for library preparation using a genotyping-by-sequencing (GBS) method based on RADseq.

To evaluate enzyme efficiency, a pilot set of 96 samples was prepared using SbfI/MspI and PstI/MspI restriction enzyme pairs. Libraries were barcoded, multiplexed, and sequenced on an Illumina NovaSeq 6000 S4 lane (paired-end 100 bp). Based on superior SNP yield and genomic coverage, PstI/MspI was selected for the remaining 604 samples, which were sequenced across two NovaSeq lanes. Library preparation was done at IBIS, and sequencing at the McGill Genome Centre (Genome Québec). SNP calling and filtering were conducted by the IBIS bioinformatics team using STACKS v2.62 and custom scripts to remove low-quality SNPs, redundant samples, and potential paralogs (details in Supplemental Methods S1).

In addition to GBS, a subset of 700 individuals was genotyped using a newly developed 25k Infinium iSelect chip that combines SNPs for black spruce (Pavy et al. 2016) and newly detected SNPs for red spruce based on exome capture and sequencing (Gerardi et al., in prep.). Genotyping was conducted at the Genome Quebec Centre d'expertise et de services in Montreal. For both genotyping methods, quality control

involved removing individuals with more than 20% missing genotype data and SNPs with more than 15% missing data. Additionally, SNPs were filtered based on a minor allele frequency (MAF) threshold of <0.01 and an inbreeding coefficient (|FIS|) greater than 0.5. SNPs with high error rates in control genotypes (error rate >0.05; n = 30) were removed from the chip dataset. No control genotypes were available for the RADSeq dataset. Finally, the RADSeq dataset yielded 10,626 SNPs and the SNP chip dataset 16,217 SNPs

Following quality filtering, 612 individuals were retained in both the RADseq and SNP chip datasets. To enhance genomic coverage and leverage the complementary strengths of these two genotyping platforms, the marker matrices were integrated by concatenating SNP data column-wise. Prior to merging, each genomic relationship matrix (G matrix) derived from the RADseq and SNP chip datasets was carefully evaluated to remove individuals that could introduce bias, such as those markers providing inconsistent information. To ensure comparability between datasets, the G matrices were standardized by aligning their trace to that of the pedigree-based relationship matrix (A), normalizing scale and variance before blending. The resulting combined matrix comprised 26,843 unique SNPs across the 612 individuals and formed the foundation for all subsequent genomic prediction analyses. This approach follows best practices recently applied in forest tree genomic studies which emphasize rigorous filtering, standardization of relationship matrices, and dataset integration to improve marker density and prediction accuracy (Aguirre et al., 2024; Tumas et al., 2024)

Construction of Relationship Matrices

Three types of relationship matrices were constructed to model additive genetic effects: a pedigree-based matrix (A), a genomic matrix (G), and a hybrid matrix (H) (Figure 2).

The pedigree-based relationship matrix (A matrix) was generated using the Amatrix function from the AGHmatrix package (R. R. Amadeu et al., 2023), based on known family identities. Its variance structure can be expressed as:

$$\mathrm{Var(a)} = \mathrm{A}\sigma_e^2 \quad (2)$$

where a is the vector of additive genetic effects and σ_e^2 is the additive genetic variance. Its inverse (A^{-1}) was computed using the Ainverse() function in ASReml-R (Butler et al., 2023) to facilitate efficient mixed model computations.

The genomic relationship matrix (G matrix) was constructed using the G.matrix() function in the ASRgenomics R package (R. R. Amadeu et al., 2023), which employs the AGHmatrix methodology (R. R. Amadeu et al., 2016), based on the combined SNP dataset from both RADseq and SNP chip platforms. Initially, separate G matrices were created for the RADseq and SNP chip datasets using the VanRaden method (VanRaden, 2008), which estimates realized genomic relationships by incorporating allele frequencies and marker genotypes. The VanRaden formula used to compute G is:

$$G = \frac{(\mathrm{M}-2\mathrm{P})(\mathrm{M}-2\mathrm{P})^T}{2\sum p_j(1-p_j)} \quad (3)$$

where M is the matrix of marker genotypes coded as 0, 1, or 2 for the number of reference alleles, $\mathrm{p_j}$ is a matrix of allele frequencies, and p_j is the allele frequency at marker j.

Because marker-based relationship matrices derived from different platforms may differ in scale and variance structure, each G matrix was tuned prior to integration. Specifically, RADseq- and SNP chip-derived G matrices were aligned to the pedigree-based matrix (A) by matching the trace of G to A, standardizing their scale and reducing bias due to platform differences (Gezan et al., 2022). Individuals with unusual allele frequencies or extreme values were removed before merging. The standardized G matrices were then blended into a single genomic relationship matrix reflecting the combined information from both marker platforms.

To integrate pedigree and genomic data, a hybrid relationship matrix (H) was constructed by blending the genomic relationship matrix (G) with the pedigree-based relationship matrix (A) at a 98:2 ratio (R. Amadeu & Ferrao, 2025). The inverse hybrid matrix (H^{-1}) was calculated using the Ginverse() function in the ASRgenomics package. This calculation incorporates the inverse of both matrices and enables single-step genomic prediction (HBLUP) across genotyped and non-genotyped individuals. The inverse hybrid relationship matrix (H^{-1}) was calculated following (Legarra et al., 2009):

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix} \quad (4)$$

where A^{-1} is the inverse of the pedigree-based relationship matrix for all individuals, G^{-1} is the inverse of the genomic relationship matrix for genotyped individuals, and A_{22}^{-1} is the inverse of the pedigree-based relationship matrix for genotyped individuals only.

Statistical Model

Six genomic prediction models were used to evaluate the accuracy of pedigree-based (ABLUP), genomic (GBLUP), and hybrid (HBLUP) approaches. All models were

based on the following general linear mixed model structure (Aguilar et al., 2010; Henderson, 1975; VanRaden, 2008):

$$y = X\mu + Z_1 b + Z_1 a + e \quad (5)$$

In this formulation, *y* is the vector of observed phenotypes (e.g., H5, H10), and μ is the overall mean, modeled as a fixed effect and linked to the observations via the design matrix *X*. The vector *b* represents random block effects, assumed to follow a normal distribution with mean zero and variance σ_b^2 with the design matrix Z_1 linking blocks to observations. In this study, block effects were nested within fixed replication effects, reflecting the hierarchical trial structure. The term *a* denotes the additive genetic effects (breeding values), with Z_2 being the corresponding incidence matrix. Depending on the model used, *a* was assumed to follow a normal distribution with variance structured by different relationship matrices:) $a \sim N(0, \sigma_a^2 A)$ for ABLUP using the pedigree-based matrix *A*; $a \sim N(0, \sigma_a^2 G)$ for GBLUP using the genomic relationship matrix *G*; and $a \sim N(0, \sigma_a^2 H)$ for HBLUP using the hybrid matrix *H*, which integrates both pedigree and genomic information. The residual term *e* was assumed to follow $e \sim N(0, \sigma_a^2 I)$ representing uncorrelated error.

Variance components were estimated using a linear mixed model implemented in ASReml-R (Butler et al., 2023). The model used to calculate these components, also serving as the base model for ABLUP, was:

$$y = \mu + \text{Rep} + a + e \quad (6)$$

Where *y* is the vector of phenotypic observations, μ is the overall mean, Rep is the fixed effect of replication (i.e., quadrant within block), $a \sim N(0, \sigma_a^2 A)$ is the vector of additive

genetic effects (breeding values) based on the pedigree relationship matrix A, $e \sim N(0, \sigma_e^2 I)$ is the residual error.

In GBLUP and HBLUP, breeding values were predicted by solving the mixed model equations, typically expressed as:

$$\hat{a} = G^{-1} Z_2^T R^{-1}(y - X\mu - Z_1 b) \quad (7)$$

In this equation, â is the vector of predicted additive genetic effects (breeding values); G^{-1} is the inverse of the genomic relationship matrix (or H^{-1} in HBLUP); Z_2 is the design matrix for genetic effects; R^{-1} is the inverse of the residual covariance matrix; and the terms $X\mu$ and $Z_1 b$ represent the contributions of fixed replicate effect and random block effects, respectively.

Heritability Estimation

Narrow-sense heritability (h^2_{ind}) was calculated from variance components as follows (Hill & Mackay, 2004):

$$h^2_{ind} = \frac{\sigma 2\text{additive}}{\sigma 2\text{additive} + \sigma 2\text{residual}} \quad (8)$$

where $\sigma^2_{\text{additive}}$ is the variance of the additive genetic effects and $\sigma^2_{\text{residual}}$ is the residual variance.

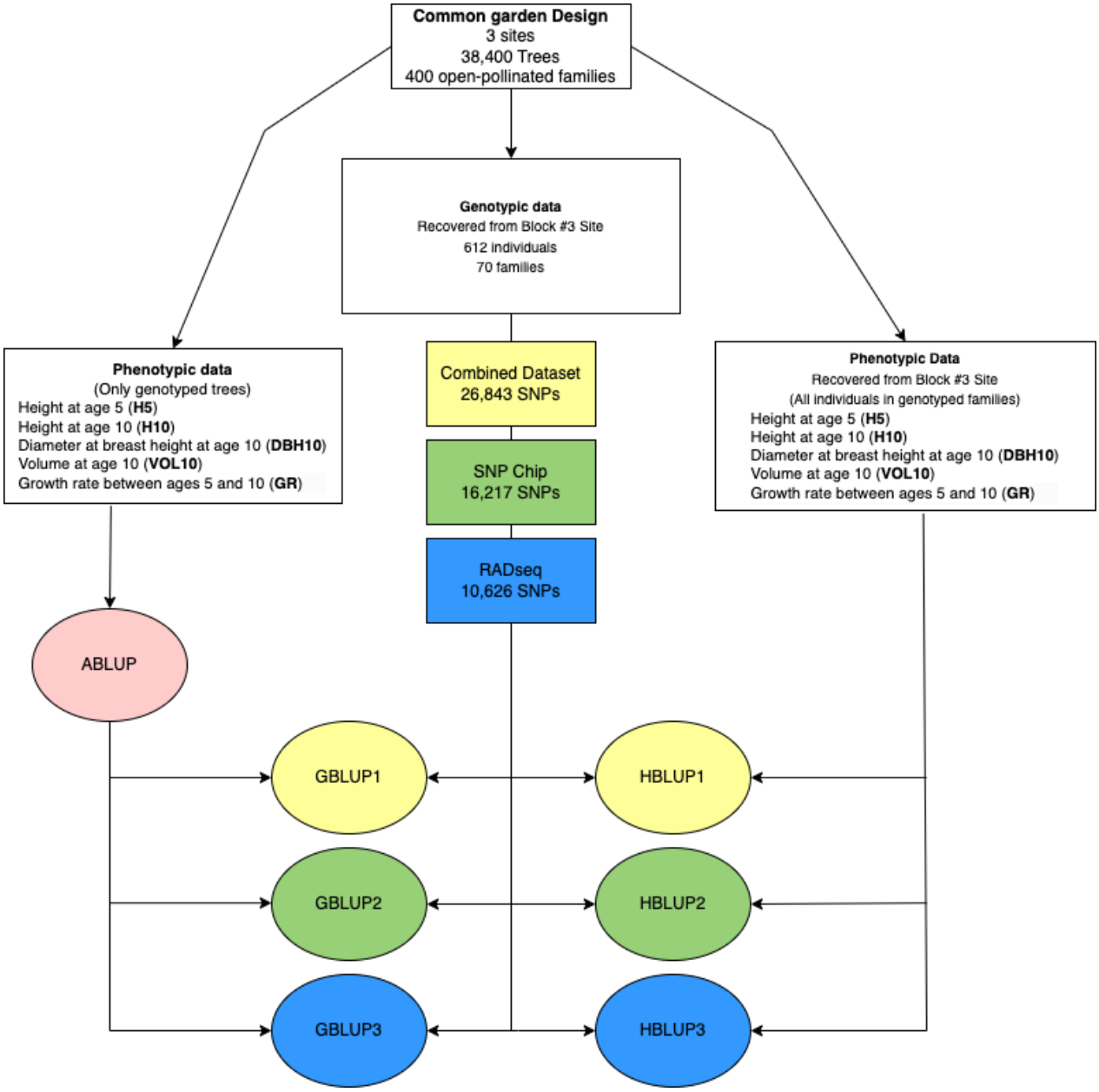

Figure 2 Flowchart illustrating the construction and integration of genetic matrices, including pedigree-based (A matrix), genomic (G matrix), and blended (H matrix) relationship matrices, used in breeding value estimation and genetic analyses.

Assessment of Predictive Ability and Accuracy

To support the evaluation of GS model performance, a custom function was developed using the Tidyverse package in R (Wickham et al., 2019) for generating cross-validation sets. Cross-validation is a resampling technique used to assess the predictive performance of statistical models by partitioning the data into training and validation subsets (Kohavi, 1995). The function enables flexible partitioning of the data into training and validation subsets based on a specified number of folds and repetitions.

Two cross-validation schemes were employed to assess the performance of pedigree-based and genomic prediction models. In the first scheme (CV1, within-family), individual trees were randomly assigned into folds while ensuring that each family was represented in every fold. In the second scheme (CV2, across-family), entire families were assigned to folds, such that each fold contained a unique subset of families, with no overlap between the training and validation datasets. Each cross-validation scheme used 10 folds and was repeated 10 times, and results were averaged across repetitions to ensure robust estimates.

Predictive ability (PA), defined as the Pearson correlation between predicted breeding values (PA_{BV}) and adjusted phenotypic values, where the phenotypes have been corrected for experimental design factors such as block effects, and prediction accuracy (PACC), defined as the correlation between predicted breeding values and the true breeding values, were assessed for each of the three model types: the pedigree-based model (ABLUP), the genomic-based model (GBLUP), and the hybrid model (HBLUP) (Estaghvirou et al., 2013). Predictive ability was calculated as the correlation between PA_{BV} and phenotypic values adjusted to remove non-genetic variation. Prediction accuracy of breeding value estimates ($PACC_{BV}$) was then derived as the ratio

of predictive ability to the square root of the individual-tree narrow-sense heritability (Dekkers, 2007):

$$PACC_{BV}=\frac{PA_{BV}}{\sqrt{h2ind}} \quad (9)$$

This ratio provides an estimate of the correlation between predicted and true breeding values and is commonly used to compare the reliability of genetic evaluation models (Daetwyler et al., 2013).

Comparative Genetic Gain from Conventional and Genomic Selection

Genetic gain was estimated to compare the long-term effectiveness of conventional and genomic selection (GS) approaches. For each trait, the mean of the top 5% genomic-estimated breeding values (GEBVs) was calculated based on predictions from the cross-validation procedure. The values were then averaged across the ten cross-validation repetitions to provide a stable estimate of selection gain under each model. To account for differences in the time required to complete a breeding cycle, we also calculated expected genetic gain per year. A breeding cycle length of 28 years was assumed for conventional selection, while a shorter cycle of nine years was assumed for GS, consistent with the acceleration enabled by early selection using genomic information. Genetic gain per year was calculated for each trait and model combination using the formula (Lenz et al., 2017):

$$Genetic\ Gain\ per\ unit = \frac{Mean\ of\ top\ 5\%\ GEBVs}{Breeding\ Cycle\ Length} \quad (10)$$

RESULTS

Trait Heritabilities

Heritability estimates (h^2_{ind}), varied across traits and models (Table 1, Figure 3), reflecting differences in the proportion of phenotypic variation explained by additive genetic effects. Overall, heritability was lowest for early height traits (H5 and H10) and relatively higher for later-stage traits such as DBH10 and VOL10, although all estimates remained low.

Height at 5 years (H5) exhibited negligible heritability across all models. All models except GBLUP3 produced heritability estimates of zero. GBLUP3, based on the RADseq dataset, yielded a slightly higher estimate of $h^2_{ind} = 0.07$, although the associated standard error was relatively large, indicating high uncertainty. In general, standard errors for H5 were large across models, further reflecting the low precision of these estimates. Height at 10 years (H10) similarly showed low heritability, with most models estimating values near zero. Slightly higher estimates were obtained from GBLUP1 and HBLUP1 ($h^2_{ind} = 0.07$), though these still indicate only a weak genetic contribution to trait variation.

For growth rate (GR), greater differentiation among models was observed. ABLUP produced the lowest estimate ($h^2_{ind} = 0.02$), while HBLUP1 (based on the combined genotypic dataset) captured the strongest genetic signal ($h^2_{ind} = 0.15$). GBLUP models showed relatively low heritabilities for GR, ranging from 0.05 (GBLUP3) to 0.09 (GBLUP1).

Diameter at breast height at age 10 (DBH10) had the highest heritability estimates in the ABLUP and GBLUP2 models, both at $h^2_{ind} = 0.15$. Other models,

including GBLUP1, GBLUP3, and all HBLUP variants produced lower but still moderate estimates ranging from 0.05 (GLUP1 and GBLUP3) to 0.09 (HBLUP1).

Volume at 10 years (VOL10) exhibited the highest heritability overall, with HBLUP1 showing the strongest genetic influence ($h^2_{ind} = 0.17$). GBLUP1 and GBLUP2 also demonstrated relatively high heritability ($h^2_{ind} = 0.13$), whereas GBLUP3 showed the lowest estimate ($h^2_{ind} = 0.06$). The HBLUP2 and HBLUP3 models yielded moderate estimates of 0.09 and 0.10, respectively.

Table 1 Heritability estimates for models and traits and their corresponding standard errors.

Model	Type	Traits				
		H5	GR	H10	DBH10	VOL10
ABLUP	h^2_{ind}	0.00	0.02	0.00	0.15	0.11
	SE	0.09	0.09	0.09	0.10	0.10
GBLUP1	h^2_{ind}	0.00	0.09	0.07	0.05	0.13
	SE	0.09	0.10	0.10	0.10	0.10
GBLUP2	h^2_{ind}	0.00	0.08	0.01	0.15	0.13
	SE	0.08	0.09	0.08	0.09	0.09
GBLUP3	h^2_{ind}	0.07	0.05	0.00	0.05	0.06
	SE	0.07	0.07	NA	0.07	0.07
HBLUP1	h^2_{ind}	0.00	0.15	0.07	0.09	0.17
	SE	NA	0.08	0.08	0.08	0.08
HBLUP2	h^2_{ind}	0.00	0.08	0.01	0.08	0.09
	SE	0.05	0.05	0.04	0.05	0.05
HBLUP3	h^2_{ind}	0.00	0.07	0.00	0.06	0.10
	SE	NA	0.05	0.04	0.05	0.05

Note: SE values reported as NA indicate that the model did not converge for that trait
Model Notes: GBLUP1 combined genotypic datasets, GBLUP2 SNP chip genotypic dataset, GBLUP3 Radseq genotypic dataset, HBLUP1 combined genotypic datasets, HBLUP2 SNP chip genotypic dataset, HBLUP3 Radseq genotypic dataset.

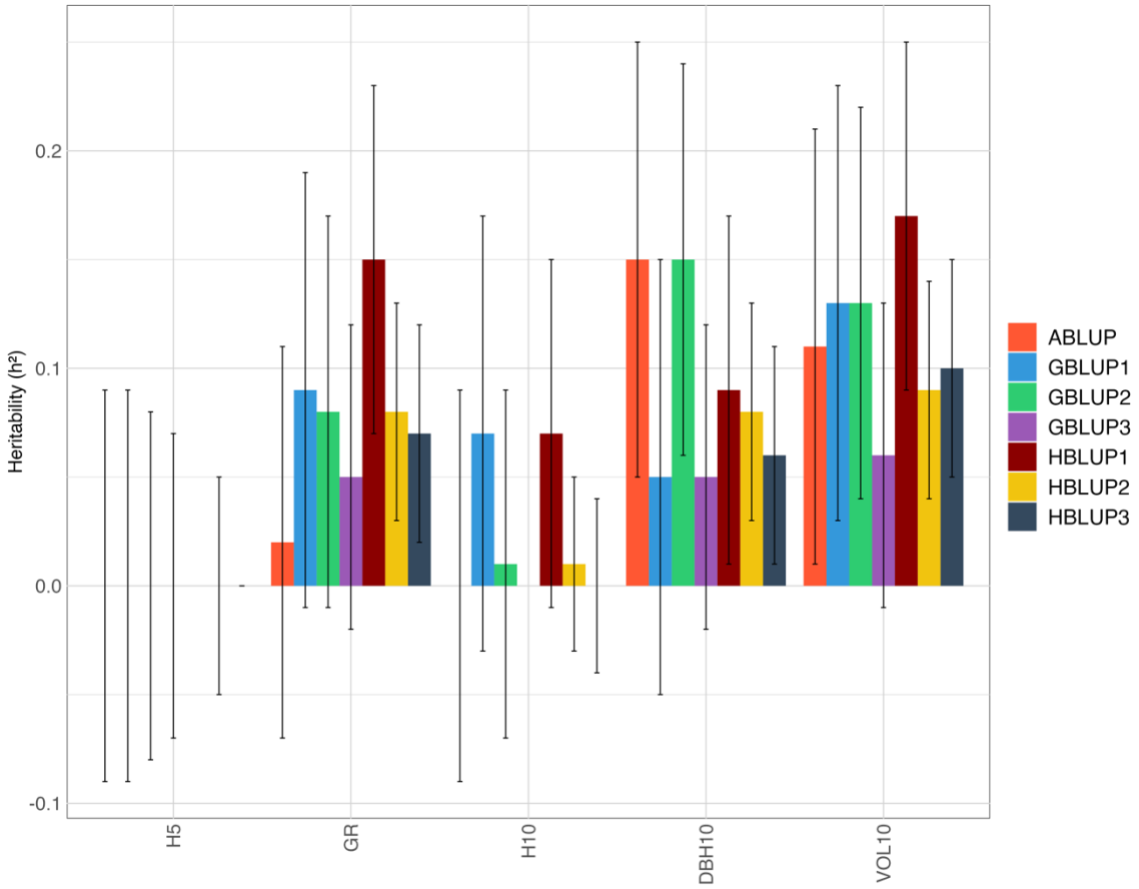

Figure 3. Heritability estimates for models and traits and their corresponding standard errors.

Predictive Ability and Accuracy

Across all models, predictive ability (PA) was uniformly low, with near-zero or negative values across most traits, reflecting weak correlations between predicted breeding values and observed phenotypes (Table 4, Figure 4). In contrast, predictive accuracy (PACC) revealed meaningful differences in model performance. The hybrid model HBLUP2 performed best overall, with PACC ranging from 0.61 for H5 to 0.69 for H10, while HBLUP3 showed a similar range (0.61–0.68). Among the GBLUP models, GBLUP2 (SNP chip) had the highest mean PACC (0.58), with accuracy values ranging from 0.52 for H5 to 0.60 for H10. GBLUP3 (RADseq) performed comparably (mean PACC = 0.56). GBLUP1, based on the combined genotyping dataset, had the weakest overall performance, with a mean PACC of 0.49 and trait-level values ranging from 0.43 to 0.51. The pedigree-based ABLUP model showed moderate accuracy (mean PACC = 0.55), with values ranging from 0.52 for H5 to 0.58 for GR.

Table 2 Accuracies of selection models based on half-sib families from a single-site analysis, using SNP chip, RADseq, and combined marker datasets. Standard errors are shown in brackets.

Model	Type	Traits					
		H5	GR	H10	DBH10	VOL10	Average
ABLUP	PA	-0.05 (0.11)	0.01 (0.12)	-0.05 (0.13)	0.08 (0.13)	0.06 (0.13)	0.06
	PACC	0.52 (0.07)	0.58 (0.06)	0.54 (0.07)	0.55 (0.07)	0.56 (0.10)	0.56
GBLUP1	PA	-0.08 (0.12)	0.04 (0.13)	0.02 (0.12)	0.01 (0.12)	0.06 (0.15)	0.06
	PACC	0.43 (0.10)	0.49 (0.10)	0.50 (0.10)	0.50 (0.09)	0.51 (0.11)	0.51
GBLUP2	PA	-0.08 (0.11)	0.05 (0.13)	-0.01 (0.14)	0.09 (0.13)	0.08 (0.14)	0.08
	PACC	0.52 (0.07)	0.59 (0.08)	0.60 (0.07)	0.59 (0.07)	0.59 (0.10)	0.59
GBLUP3	PA	-0.12 (0.11)	0.03 (0.12)	-0.05 (0.12)	0.03 (0.13)	0.05 (0.15)	0.05
	PACC	0.46 (0.09)	0.58 (0.07)	0.55 (0.07)	0.59 (0.08)	0.60 (0.11)	0.60
HBLUP1	PA	-0.03 (0.09)	0.06 (0.09)	0.03 (0.09)	0.04 (0.10)	0.07 (0.10)	0.07
	PACC	0.49 (0.06)	0.49 (0.07)	0.50 (0.06)	0.52 (0.06)	0.50 (0.08)	0.50
HBLUP2	PA	-0.06 (0.08)	0.06 (0.09)	0.00 (0.09)	0.06 (0.10)	0.07 (0.09)	0.07
	PACC	0.61 (0.04)	0.66 (0.04)	0.69 (0.04)	0.68 (0.04)	0.67 (0.06)	0.67
HBLUP3	PA	0.06 (0.08)	0.05 (0.09)	-0.01 (0.08)	0.04 (0.09)	0.06 (0.09)	0.06
	PACC	0.61 (0.05)	0.66 (0.04)	0.68 (0.04)	0.67 (0.04)	0.65 (0.07)	0.65

Model Notes: GBLUP1 combined genotypic datasets, GBLUP2 SNP chip genotypic dataset, GBLUP3 Radseq genotypic dataset, HBLUP1 combined genotypic datasets, HBLUP2 SNP chip genotypic dataset, HBLUP3 Radseq genotypic dataset.
PA: Predictive ability; PACC: Predictive accuracy.

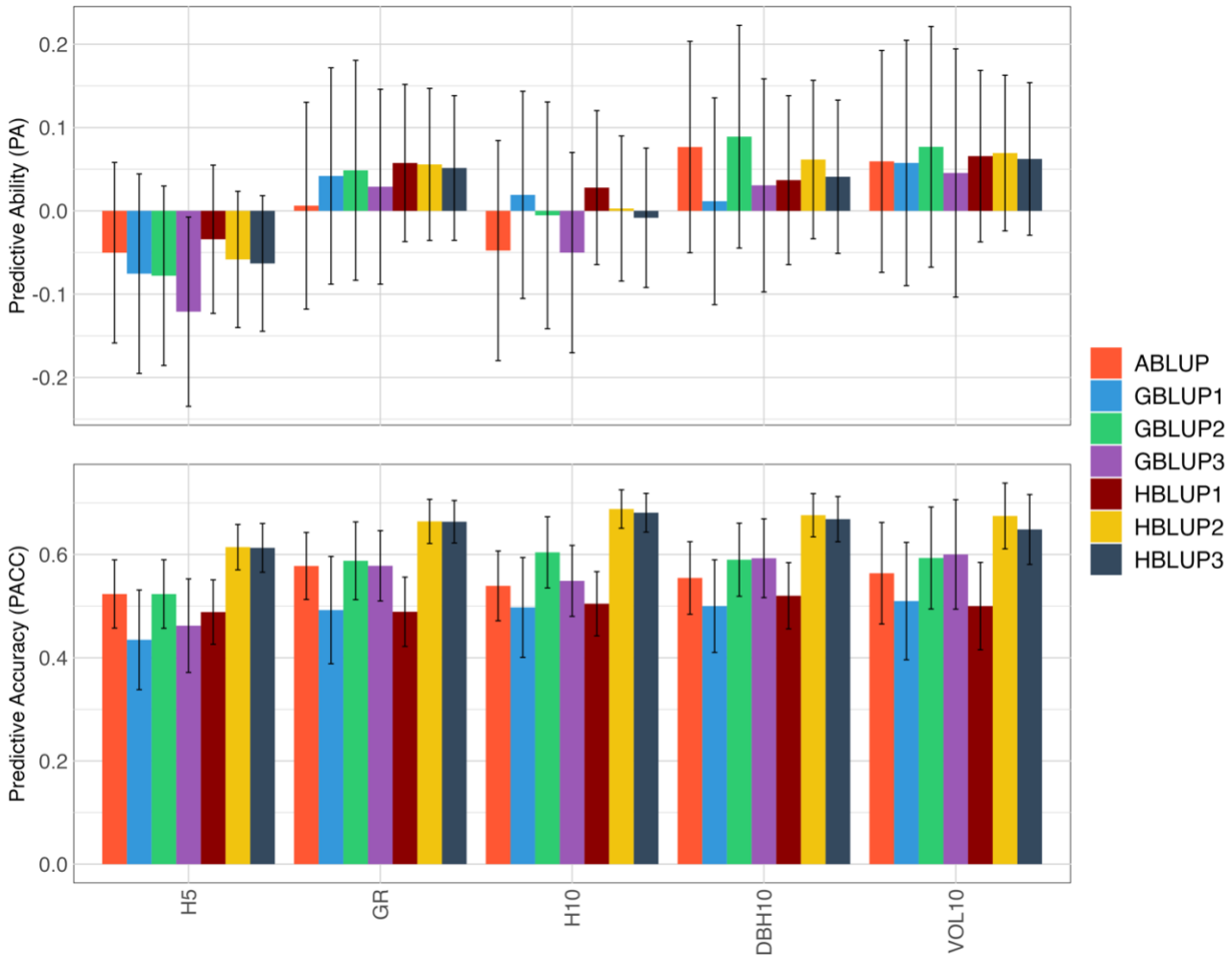

Figure 4 Predictive abilities (a) and predictive accuracies (b) of genomic selection models and the pedigree-based ABLUP model based on half-sib families in a single-site analysis, using microarray, RADseq, and combined marker datasets.

Genetic Gains

Genetic gains, expressed as absolute values per unit time, were slightly higher for marker-based methods than for pedigree-based approaches (Table 5, Figure 5). While no single model consistently outperformed all others across all traits, HBLUP1 (pedigree and combined genomic dataset) demonstrated the most consistently moderate gains, achieving the highest values for VOL10 (0.08), GR (0.07), and a shared top gain for H10 (0.05). GBLUP2 (SNP chip data) produced the highest predicted gain for DBH10 (0.07), while GBLUP1 (combined genomic data) shared the top gain for H10 (0.05). However, it is important to note that these differences in gain are small in absolute terms, reflecting the generally low additive genetic contribution detected in this dataset. Notably, genetic gain for H5 was zero across all models, consistent with the absence of detectable genetic control or heritability for this trait. For height at age 10 (H10), modest but variable genetic gain is expected, ranging from 0.0 (ABLUP and GBLUP3) to 0.05 (GBLUP1 and HBLUP1).

Table 3 Genetic gains for models and traits using pedigree and marker-based methods.

Model	Type	Cycle Length	Traits				
			H5	GR	H10	DBH10	VOL10
ABLUP	Pedigree	28 years	0.0	0.01	0.00	0.02	0.02
GBLUP1	Markers	9 years	0.00	0.06	0.05	0.04	0.07
GBLUP2	Markers	9 years	0.00	0.05	0.01	0.07	0.07
GBLUP3	Markers	9 years	0.00	0.04	0.00	0.04	0.04
HBLUP1	Markers	9 years	0.00	0.07	0.05	0.06	0.08
HBLUP2	Markers	9 years	0.00	0.05	0.02	0.05	0.05
HBLUP3	Markers	9 years	0.00	0.05	0.00	0.05	0.06

Model Notes: GBLUP1 combined genotypic datasets, GBLUP2 SNP chip genotypic dataset, GBLUP3 Radseq genotypic dataset, HBLUP1 combined genotypic datasets, HBLUP2 SNP chip genotypic dataset, HBLUP3 Radseq genotypic dataset

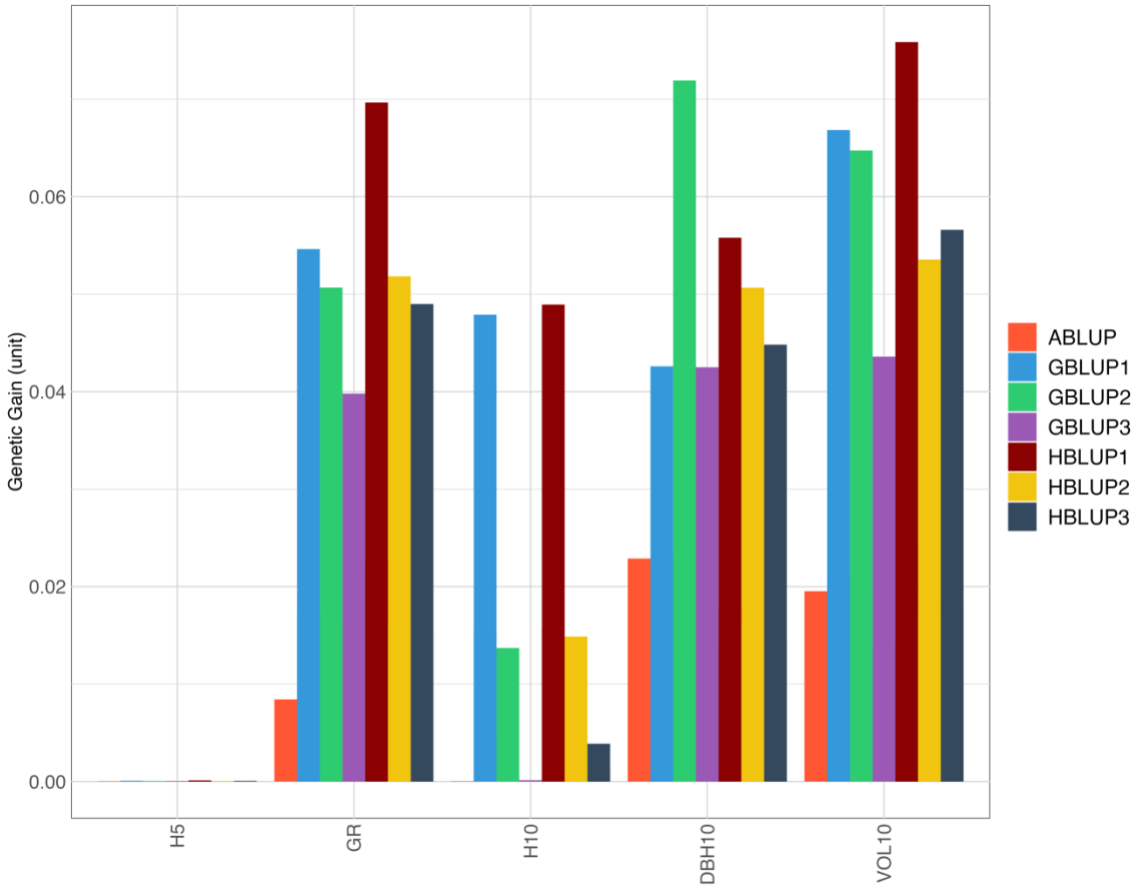

Figure 5 Genetic gain per year for each model and trait.
Bars represent the average genetic gain per year based on the mean of the top 5% predicted breeding values across ten cross-validation replicates. Gains for ABLUP reflect conventional selection assuming a 28-year breeding cycle. Gains for GBLUP and HBLUP models reflect genomic selection assuming a 9-year cycle.

DISCUSSION

Heritability Patterns and Trait-Level Insights

This study revealed that heritability in black spruce varies substantially by trait and by modeling approach, reflecting differences in genetic architecture, trait expression, and sensitivity to environmental variation. Across models, hybrid approaches, particularly HBLUP1, consistently produced higher heritability estimates than pedigree-based (ABLUP) or marker-only (GBLUP) methods. This supports findings from Callister et al. (2021), Ratcliffe et al. (2017), and Simiqueli et al. (2023), confirming the value of combining genomic and pedigree information to improve additive variance partitioning, especially in open-pollinated species with uncertain relatedness (Callister et al., 2021; Ratcliffe et al., 2017; Simiqueli et al., 2023).

Growth traits such as diameter at breast height (DBH10), growth rate (GR), and volume at 10 years (VOL10) exhibited moderate individual heritability (h^2_{ind} = 0.15–0.17), indicating stable genetic control and making them promising targets for selection. These results are consistent with findings from Lenz et al. (2017), who observed similar heritability ranges in full-sib black spruce families in Quebec. Importantly, our results extend this evidence to a more operational Ontario context, showing that moderate heritability can still be recovered despite lower replication and family size.

In contrast, early height traits (H5, H10) displayed relatively low individual-tree heritability in our analyses, particularly under ABLUP, where estimates approached zero. This stands in contrast to Fu (2000), who reported higher individual-tree heritability estimates for the same traits in Block 3 (0.07 for H5 and 0.09 for H10), highlighting a notable discrepancy. The lower estimates observed in our study may stem

from differences in pedigree depth, smaller family sizes, reduced replication, smaller dataset or greater environmental noise, which can attenuate additive genetic signals at the individual level.

However, these findings also diverge from a consistent trend observed across many studies in black spruce and other conifers, where height generally exhibits higher heritability than DBH (Beaulieu et al., 2020; Cappa et al., 2022; Lu & Charrette, 2008). For instance, Lu & Charrette (2008) reported narrow-sense heritability values for height of 0.19 at age 6, compared to lower values for DBH, based on a large dataset of over 42,000 trees. Such results have reinforced the perception of height as a more reliable selection trait in operational breeding. Although our results suggest early height may be a poor proxy for long-term performance under the current trial design, this interpretation must be tempered by the extensive literature showing height's strong genetic basis.

The apparent discrepancy may reflect maternal effects and microenvironmental heterogeneity (Laverdière et al., 2022), which are particularly pronounced in early stages and may obscure additive genetic variance when using less robust models. Therefore, the utility of early height traits should be reconsidered in light of broader empirical evidence and the limitations of our current dataset.

Predictive Ability, Accuracy, and Genetic Gain

Genomic models (GBLUP and HBLUP) consistently outperformed ABLUP in predictive ability (PA), accuracy (PACC), and estimated genetic gain across all traits. The greatest improvements were observed for traits with higher heritability estimates, such as GR, DBH10, and VOL10, where marker-based models more effectively captured additive variance and Mendelian sampling effects. These patterns reaffirm the theoretical advantages of genomic selection (Meuwissen et al., 2001; Strandén &

Garrick, 2009) and echo empirical findings in black spruce (Lenz et al., 2017) and other conifers such as loblolly pine (Resende et al., 2012).

The gains for early height traits were comparatively modest, even under genomic models, due to the low heritability estimated for these traits (ranging from 0.00 to 0.07 across models). While HBLUP and GBLUP improved over ABLUP, the absolute gains remained constrained. This confirms that predictive performance is closely tied to trait heritability and that genomic selection is most impactful when applied to traits with stronger genetic determinism.

Notably, HBLUP models offered a significant operational advantage by enabling predictions for ungenotyped individuals. This feature is particularly valuable in resource-constrained breeding programs and supports the feasibility of implementing GS at scale without genotyping every individual. Within genomic models, SNP chip-based approaches (GBLUP2, HBLUP2) outperformed RADseq-based models (GBLUP3, HBLUP3), highlighting the importance of marker quality and genome-wide coverage in achieving robust predictions. While our findings and previous studies (Kim et al., 2022; Ma et al., 2022) highlight the importance of marker density for improving prediction accuracy, it is also possible that the distribution and properties of the markers, such as those obtained through RADseq, play a significant role, potentially by enabling better separation of additive and dominance alleles, an aspect that warrants further investigation in future studies.

Limitations and Sampling Considerations

Several limitations must be acknowledged when interpreting the current findings. First, the study was conducted within a single breeding zone (LNW), limiting the generalizability of results across the broader range of black spruce. Second, the use of

open-pollinated families introduces pedigree uncertainty, which undermines the accuracy of relatedness estimates in ABLUP, GBLUP, and HBLUP. This issue is well-documented in conifers (El-Kassaby et al., 2024) and can bias estimates of heritability and breeding values. The relatively small number of families (n = 70) and reduced replication further constrained our ability to accurately estimate genetic parameters, particularly for environmentally sensitive traits such as early height (Perron et al., 2013). These factors likely contributed to the low heritability and limited gain estimates observed for some traits.

As Beaulieu et al. (2014) and others have noted, sampling design plays a critical role in the reliability of GS models. Moreover, the comparison of genotyping platforms revealed that reduced-representation methods like RADseq may fall short in providing the dense, uniform genome-wide coverage needed for high-resolution prediction, particularly under low-replication conditions. While cost-effective, such platforms may not be ideal for routine operational deployment unless complemented with imputation or hybrid approaches. However, imputation introduces its own challenges, as inferred genotypes can be a source of error and uncertainty. Gamal El-Dien et al. (2015) examined how different imputation methods influence genomic selection accuracy, highlighting the trade-offs involved in relying on imputed data (Gamal El-Dien et al., 2015).

Future Research Directions

Future genomic selection studies in black spruce should adopt multi-site trial designs to explicitly evaluate genotype-by-environment (G×E) interactions. Given the species' wide ecological range, capturing environmental heterogeneity is essential for developing broadly applicable prediction models and understanding trait plasticity under

varying climatic conditions. High-density genotyping approaches such as exome capture or whole-genome sequencing should be explored as sequencing costs continue to decline. These platforms could improve marker resolution and trait-locus associations, potentially increasing prediction accuracy, especially for traits under low heritability. There is also significant potential in incorporating machine learning techniques and multi-trait genomic prediction models. These methods may offer potential benefits for traits where additive genetic variance is limited or where non-additive effects play a significant role. However, accurately detecting and utilizing non-additive variance requires specifically designed experiments, such as those involving controlled crosses (Nadeau et al., 2023). Additionally, future studies should systematically assess how different sampling structures, such as family number, individual replication, and relatedness, affect prediction accuracy and parameter estimation. Simulation-based tools could help optimize these designs before implementation.

Management Implications and Operational Relevance

The integration of genomic selection into black spruce breeding programs is both feasible and beneficial, particularly for high-value growth traits like DBH and volume that exhibit consistent genetic control. Our findings suggest that genomic models, especially HBLUP, can substantially accelerate breeding progress by improving selection efficiency while also accommodating ungenotyped individuals, making it a pragmatic tool for resource-limited programs. Although early height has historically served as a proxy for long-term growth, its low heritability in this study suggests that it should be used cautiously. Rather than excluding height from selection programs, it may be more effective to treat it as a component in multi-trait indices. This would help

account for how height responds to different environmental conditions, especially in the context of early testing.

The performance difference between genotyping platforms has direct operational implications. SNP chips consistently delivered superior results, indicating they are the preferred platform when budgets allow. RADseq remains a cost-effective alternative but may require supplemental strategies, such as imputation or hybrid integration with chip data, to be competitive. While combined-platform approaches showed some promise, their cost-effectiveness and logistical feasibility for routine application remain uncertain.

Ultimately, this study reinforces the operational value of genomic tools in accelerating black spruce breeding. By aligning model choice and genotyping strategy with specific trait characteristics and breeding objectives, programs can achieve more reliable genetic gain, reduce cycle time, and increase the adaptability of forest populations to future environmental challenges.

Literature Cited

Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., & Lawlor, T. J. (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*(2), 743–752. https://doi.org/10.3168/jds.2009-2730

Aguirre, N. C., Villalba, P. V., García, M. N., Filippi, C. V., Rivas, J. G., Martínez, M. C., Acuña, C. V., López, A. J., López, J. A., Pathauer, P., Palazzini, D., Harrand, L., Oberschelp, J., Marcó, M. A., Cisneros, E. F., Carreras, R., Martins Alves, A. M., Rodrigues, J. C., Hopp, H. E., … Marcucci Poltri, S. N. (2024). Comparison of ddRADseq and EUChip60K SNP genotyping systems for population genetics and genomic selection in Eucalyptus dunnii (Maiden). *Frontiers in Genetics*, *15*, 1361418. https://doi.org/10.3389/fgene.2024.1361418

Amadeu, R., & Ferrao, L. (2025). *AGHmatrix: Relationship Matrices for Diploid and Autopolyploid Species* (p. 2.1.4) [Dataset]. https://doi.org/10.32614/CRAN.package.AGHmatrix

Amadeu, R. R., Cellon, C., Olmstead, J. W., Garcia, A. A. F., Resende, M. F. R., & Muñoz, P. R. (2016). AGHmatrix: R Package to Construct Relationship Matrices for Autotetraploid and Diploid Species: A Blueberry Example. *The Plant Genome*, *9*(3), plantgenome2016.01.0009. https://doi.org/10.3835/plantgenome2016.01.0009

Amadeu, R. R., Garcia, A. A. F., Munoz, P. R., & Ferrão, L. F. V. (2023). AGHmatrix: Genetic relationship matrices in R. *Bioinformatics*, *39*(7), btad445. https://doi.org/10.1093/bioinformatics/btad445

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92. https://doi.org/10.1038/nrg.2015.28

Badenes, M. L., Fernández I Martí, A., Ríos, G., & Rubio-Cabetas, M. J. (2016). Application of Genomic Technologies to the Breeding of Trees. *Frontiers in Genetics*, *7*. https://doi.org/10.3389/fgene.2016.00198

Beaulieu, J., Doerksen, T., Clément, S., MacKay, J., & Bousquet, J. (2014). Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. *Heredity*, *113*(4), 343–352. https://doi.org/10.1038/hdy.2014.36

Beaulieu, J., Nadeau, S., Ding, C., Celedon, J. M., Azaiez, A., Ritland, C., Laverdière, J., Deslauriers, M., Adams, G., Fullarton, M., Bohlmann, J., Lenz, P., & Bousquet, J. (2020). Genomic selection for resistance to spruce budworm in white spruce and relationships with growth and wood quality traits. *Evolutionary Applications*, *13*(10), 2704–2722. https://doi.org/10.1111/eva.13076

Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J., & Thompson, R. (2023). ASReml-R Reference Manual Version 4.2. *VSN International Ltd.*

Callister, A. N., Bradshaw, B. P., Elms, S., Gillies, R. A. W., Sasse, J. M., & Brawner, J. T. (2021). Single-step genomic BLUP enables joint analysis of disconnected breeding programs: An example with *Eucalyptus globulus* Labill. *G3 Genes|Genomes|Genetics*, *11*(10), jkab253. https://doi.org/10.1093/g3journal/jkab253

Cappa, E. P., Ratcliffe, B., Chen, C., Thomas, B. R., Liu, Y., Klutsch, J., Wei, X., Azcona, J. S., Benowicz, A., Sadoway, S., Erbilgin, N., & El-Kassaby, Y. A. (2022). Improving lodgepole pine genomic evaluation using spatial correlation

structure and SNP selection with single-step GBLUP. *Heredity*, *128*(4), 209–224. https://doi.org/10.1038/s41437-022-00508-2

Chang, W.-Y., Gaston, C., Cool, J., & Thomas, B. R. (2019). A financial analysis of using improved planting stock of white spruce and lodgepole pine in Alberta, Canada: Genomic selection versus traditional breeding. *Forestry: An International Journal of Forest Research*, *92*(3), 297–310. https://doi.org/10.1093/forestry/cpz011

Daetwyler, H. D., Calus, M. P. L., & Pong-Wong, R. (2013). Genomic Prediction in Animals and Plants: Simulation of Data, Validation, Reporting, and Benchmarking. *Genetics*, *193*, 347–365.

Davey, J. W., & Blaxter, M. L. (2010). RADSeq: Next-generation population genetics. *Briefings in Functional Genomics*, *9*(5–6), 416–423. https://doi.org/10.1093/bfgp/elq031

Dekkers, J. C. M. (2007). Prediction of response to marker-assisted and genomic selection using selection index theory. *Journal of Animal Breeding and Genetics*, *124*(6), 331–341. https://doi.org/10.1111/j.1439-0388.2007.00701.x

El-Kassaby, Y. A., Cappa, E. P., Chen, C., Ratcliffe, B., & Porth, I. M. (2024). Efficient genomics-based 'end-to-end' selective tree breeding framework. *Heredity*, *132*(2), 98–105. https://doi.org/10.1038/s41437-023-00667-w

Estaghvirou, S. B. O., Ogutu, J. O., Schulz-Streeck, T., Knaak, C., Ouzunova, M., Gordillo, A., & Piepho, H.-P. (2013). Evaluation of approaches for estimating the accuracy of genomic prediction in plant breeding. *BMC Genomics*, *14*.

Fradette, O., Marty, C., Tremblay, P., Lord, D., & Boucher, J.-F. (2021). Allometric Equations for Estimating Biomass and Carbon Stocks in Afforested Open

Woodlands with Black Spruce and Jack Pine, in the Eastern Canadian Boreal Forest. *Forests*, *12*(1), 59. https://doi.org/10.3390/f12010059

Fu, Y.-B. (2000). Lake Nipigon West Breeding Zone Black Spruce Family Tests: 1998 assessment and calculation of family and individual breeding values. *Ontario Tree Improvement Board*, 1–142.

Gamal El-Dien, O., Ratcliffe, B., Klápště, J., Chen, C., Porth, I., & El-Kassaby, Y. A. (2015). Prediction accuracies for growth and wood attributes of interior spruce in space using genotyping-by-sequencing. *BMC Genomics*, *16*(1), 370. https://doi.org/10.1186/s12864-015-1597-y

Gamal El-Dien, O., Shalev, T. J., Yuen, M. M. S., Stirling, R., Daniels, L. D., Breinholt, J. W., Neves, L. G., Kirst, M., Van Der Merwe, L., Yanchuk, A. D., Ritland, C., Russell, J. H., & Bohlmann, J. (2022). Genomic selection reveals hidden relatedness and increased breeding efficiency in western redcedar polycross breeding. *Evolutionary Applications*, *15*(8), 1291–1312. https://doi.org/10.1111/eva.13463

Gezan, S. A., de Oliveira, A. A., Galli, G., & Murray, D. (2022). User's Manual for ASRgenomics v. 1.1.0 An R package with complementary genomic functions. *VSN International*.

Godbout, J., Tremblay, L., Levasseur, C., Lavigne, P., Rainville, A., Mackay, J., Bousquet, J., & Isabel, N. (2017). Development of a Traceability System Based on a SNP Array for Large-Scale Production of High-Value White Spruce (Picea glauca). *Frontiers in Plant Science*, *8*, 1264. https://doi.org/10.3389/fpls.2017.01264

Grattapaglia, D., & Resende, M. D. V. (2011). Genomic selection in forest tree breeding. *Tree Genetics & Genomes*, *7*(2), 241–255. https://doi.org/10.1007/s11295-010-0328-4

Henderson, C. R. (1975). Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometrics*, *31*(2), 423. https://doi.org/10.2307/2529430

Hill, W. G., & Mackay, T. F. C. (2004). D. S. Falconer and Introduction to Quantitative Genetics. *Genetics*, *167*(4), 1529–1536. https://doi.org/10.1093/genetics/167.4.1529

Honer, T. G. (1967). Standard volume tables and merchantable conversion factors for the commercial tree species of central and eastern Canada. *Department of Forestry and Rural Development, Forest Management Research and Services Institute, Information Report FMR-X-5.*

Isabel, N., Beaulieu, J., & Bousquet, J. (1995). Complete congruence between gene diversity estimates derived from genotypic data at enzyme and random amplified polymorphic DNA loci in black spruce. *Proceedings of the National Academy of Sciences*, *92*(14), 6369–6373. https://doi.org/10.1073/pnas.92.14.6369

Kim, K.-W., Nawade, B., Nam, J., Chu, S.-H., Ha, J., & Park, Y.-J. (2022). Development of an inclusive 580K SNP array and its application for genomic selection and genome-wide association studies in rice. *Frontiers in Plant Science*, *13*, 1036177. https://doi.org/10.3389/fpls.2022.1036177

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, 1137–1143.

Laverdière, J., Lenz, P., Nadeau, S., Depardieu, C., Isabel, N., Perron, M., Beaulieu, J., & Bousquet, J. (2022). Breeding for adaptation to climate change: Genomic selection for drought response in a white spruce multi-site polycross test. *Evolutionary Applications*, *15*(3), 383–402. https://doi.org/10.1111/eva.13348

Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science*, *92*(9), 4656–4663. https://doi.org/10.3168/jds.2009-2061

Lenz, P. R. N., Beaulieu, J., Mansfield, S. D., Clément, S., Desponts, M., & Bousquet, J. (2017). Factors affecting the accuracy of genomic selection for growth and wood quality traits in an advanced-breeding population of black spruce (Picea mariana). *BMC Genomics*, *18*(1), 335. https://doi.org/10.1186/s12864-017-3715-5

Liu, H., Xia, C., & Lan, H. (2022). An efficient genomic prediction method without the direct inverse of the genomic relationship matrix. *Frontiers in Plant Science*, *13*, 1089937. https://doi.org/10.3389/fpls.2022.1089937

Lo, T., Coombe, L., Gagalova, K. K., Marr, A., Warren, R. L., Kirk, H., Pandoh, P., Zhao, Y., Moore, R. A., Mungall, A. J., Ritland, C., Pavy, N., Jones, S. J. M., Bohlmann, J., Bousquet, J., Birol, I., & Thomson, A. (2023). Assembly and annotation of the black spruce genome provide insights on spruce phylogeny and evolution of stress response. *G3: Genes, Genomes, Genetics*, *14*(1), jkad247. https://doi.org/10.1093/g3journal/jkad247

Lu, P., & Charrette, P. (2008). Genetic parameter estimates for growth traits of black spruce in northwestern Ontario. *Canadian Journal of Forest Research*, *38*(12), 2994–3001. https://doi.org/10.1139/X08-133

Ma, J., Cao, Y., Wang, Y., & Ding, Y. (2022). Development of the maize 5.5K loci panel for genomic prediction through genotyping by target sequencing. *Frontiers in Plant Science*, *13*, 972791. https://doi.org/10.3389/fpls.2022.972791

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics*, *157*(4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Moreau, G., Chagnon, C., Auty, D., Caspersen, J., & Achim, A. (2020). Impacts of Climatic Variation on the Growth of Black Spruce Across the Forest-Tundra Ecotone: Positive Effects of Warm Growing Seasons and Heat Waves Are Offset by Late Spring Frosts. *Frontiers in Forests and Global Change*, *3*, 613523. https://doi.org/10.3389/ffgc.2020.613523

Mullin, T., Andersson, B., Bastien, J., Beaulieu, J., Burdon, R., Dvorak, W., King, J., Kondo, T., Krakowski, J., Lee, S., McKeand, S., Pâques, L., Raffin, A., Russell, J., Skrøppa, T., Stoehr, M., & Yanchuk, A. (2011). Economic Importance, Breeding Objectives and Achievements. In C. Kole (Ed.), *Genetics, Genomics and Breeding of Conifers*. Science Publishers. https://doi.org/10.1201/b11075-3

Nadeau, S., Beaulieu, J., Gezan, S. A., Perron, M., Bousquet, J., & Lenz, P. R. N. (2023). Increasing genomic prediction accuracy for unphenotyped full-sib families by modeling additive and dominance effects with large datasets in white spruce. *Frontiers in Plant Science*, *14*, 1137834. https://doi.org/10.3389/fpls.2023.1137834

Perron, M., DeBlois, J., & Desponts, M. (2013). Use of resampling to assess optimal subgroup composition for estimating genetic parameters from progeny trials.

Tree Genetics & Genomes, *9*(1), 129–143. https://doi.org/10.1007/s11295-012-0540-5

Ratcliffe, B., El-Dien, O. G., Cappa, E. P., Porth, I., Klápště, J., Chen, C., & El-Kassaby, Y. A. (2017). Single-Step BLUP with Varying Genotyping Effort in Open-Pollinated *Picea glauca*. *G3 Genes|Genomes|Genetics*, *7*(3), 935–942. https://doi.org/10.1534/g3.116.037895

Resende, M. F. R., Muñoz, P., Resende, M. D. V., Garrick, D. J., Fernando, R. L., Davis, J. M., Jokela, E. J., Martin, T. A., Peter, G. F., & Kirst, M. (2012). Accuracy of Genomic Selection Methods in a Standard Data Set of Loblolly Pine (*Pinus taeda* L.). *Genetics*, *190*(4), 1503–1510. https://doi.org/10.1534/genetics.111.137026

Roorkiwal, M., Jarquin, D., Singh, M. K., Gaur, P. M., Bharadwaj, C., Rathore, A., Howard, R., Srinivasan, S., Jain, A., Garg, V., Kale, S., Chitikineni, A., Tripathi, S., Jones, E., Robbins, K. R., Crossa, J., & Varshney, R. K. (2018). Genomic-enabled prediction models using multi-environment trials to estimate the effect of genotype × environment interaction on prediction accuracy in chickpea. *Scientific Reports*, *8*(1), 11701. https://doi.org/10.1038/s41598-018-30027-2

Scheben, A., Batley, J., & Edwards, D. (2017). Genotyping-by-sequencing approaches to characterize crop genomes: Choosing the right tool for the right application. *Plant Biotechnology Journal*, *15*(2), 149–161. https://doi.org/10.1111/pbi.12645

Simiqueli, G. F., Resende, R. T., Takahashi, E. K., De Sousa, J. E., & Grattapaglia, D. (2023). Realized genomic selection across generations in a reciprocal recurrent selection breeding program of Eucalyptus hybrids. *Frontiers in Plant Science*, *14*, 1252504. https://doi.org/10.3389/fpls.2023.1252504

Strandén, I., & Garrick, D. J. (2009). Technical note: Derivation of equivalent computing algorithms for genomic predictions and reliabilities of animal merit. *Journal of Dairy Science*, *92*(6), 2971–2975. https://doi.org/10.3168/jds.2008-1929

Thomas, B. R., Stoehr, M., Schreiber, S. G., Benowicz, A., Schroeder, W. R., Soolanayakanahally, R., Stefner, C., Elliott, K. A., Philis, N., Roubal, N., Périnet, P., Perron, M., Simpson, D., Fullarton, M., Sherrill, J., Myers, M., Steeves, D., Bockstette, S., English, B., & Kort, J. (2024). Tree Improvement in Canada – past, present and future, 2023 and beyond. *The Forestry Chronicle*, *100*(1), 59–87. https://doi.org/10.5558/tfc2024-004

Thomson, A. M., Riddell, C. L., & Parker, W. H. (2009). Boreal forest provenance tests used to predict optimal growth and response to climate change: 2. Black spruce. *Canadian Journal of Forest Research*, *39*(1), 143–153. https://doi.org/10.1139/X08-167

Tong, C., Yao, D., Wu, H., Chen, Y., Yang, W., & Zhao, W. (2020). High-Quality SNP Linkage Maps Improved QTL Mapping and Genome Assembly in *Populus*. *Journal of Heredity*, *111*(6), 515–530. https://doi.org/10.1093/jhered/esaa039

Tumas, H., Ilska, J. J., Gérardi, S., Laroche, J., A'Hara, S., Boyle, B., Janes, M., McLean, P., Lopez, G., Lee, S. J., Cottrell, J., Gorjanc, G., Bousquet, J., Woolliams, J. A., & MacKay, J. J. (2024). High-density genetic linkage mapping in Sitka spruce advances the integration of genomic resources in conifers. *G3 Genes|Genomes|Genetics*, *14*(4), jkae020. https://doi.org/10.1093/g3journal/jkae020

Ulaszewski, B., Meger, J., & Burczyk, J. (2021). Comparative Analysis of SNP Discovery and Genotyping in Fagus sylvatica L. and Quercus robur L. Using RADseq, GBS, and ddRAD Methods. *Forests*, *12*(2), 222. https://doi.org/10.3390/f12020222

VanRaden, P. M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*, *91*(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

Werner, C. R., Gaynor, R. C., Gorjanc, G., Hickey, J. M., Kox, T., Abbadi, A., Leckband, G., Snowdon, R. J., & Stahl, A. (2020). How Population Structure Impacts Genomic Selection Accuracy in Cross-Validation: Implications for Practical Breeding. *Frontiers in Plant Science*, *11*, 592977. https://doi.org/10.3389/fpls.2020.592977

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

You, Q., Yang, X., Peng, Z., Xu, L., & Wang, J. (2018). Development and Applications of a High Throughput Genotyping Tool for Polyploid Crops: Single Nucleotide Polymorphism (SNP) Array. *Frontiers in Plant Science*, *9*, 104. https://doi.org/10.3389/fpls.2018.00104

SUPPLEMENTAL MATERIALS

. GBS Data Processing and Filtering.

Raw RADseq data were processed using STACKS v2.62, following the standardized stacks_workflow v2.62 pipeline. Adapter trimming and quality filtering were performed with Cutadapt v1.18, allowing up to 20% mismatches and requiring a minimum read length of 50 base pairs. Reads were demultiplexed using process_radtags with quality control settings and enzyme specifications matching the library preparation protocol (PstI and MspI).

Locus assembly and SNP calling were carried out using the STACKS pipeline, which included the ustacks, cstacks, sstacks, tsv2bam, gstacks, and populations modules. The populations module was run with parameters to retain loci present in at least four populations and in at least 60% of individuals, producing a minimally filtered VCF file.

SNPs were further filtered using a custom script that retained loci with a minimum read depth of 4×, no more than 40% missing data per group, and rare alleles present in at least three individuals. Samples with more than 20% missing data (n = 73) were removed, and the filtering was repeated. Additional filtering removed three individuals with high relatedness and five with unusually low heterozygosity, followed by a final round of filtering using the same thresholds.

To minimize genotyping errors, a modified HD plot method was used to identify loci likely affected by paralogy or tag over-merging. Only high-confidence (canonical) SNPs were retained, and SNPs in strong linkage disequilibrium were pruned by keeping one SNP per cluster. Missing genotypes were imputed based on overall allele frequencies.

Population structure was analyzed using ADMIXTURE v1.3.0, testing values of K from 1 to 20. Based on cross-validation error and visual inspection of admixture plots, K = 1 was selected. A final round of genotype imputation was performed per SNP by randomly sampling alleles according to population-level frequencies.

Supplemental Table S1. Predictive ability (PA) and prediction accuracy (PACC) of breeding value estimates for five growth traits across pedigree-based (ABLUP), genomic (GBLUP), and hybrid (HBLUP) models using 10-fold cross-validation. Values shown are means with standard errors in parentheses.

Trait	Model	PA (SE)	PACC (SE)
DBH10	ABLUP	0.08 (0.13)	0.55 (0.07)
VOL10	ABLUP	0.06 (0.13)	0.56 (0.10)
GR	ABLUP	0.01 (0.12)	0.58 (0.06)
H10	ABLUP	-0.05 (0.13)	0.54 (0.07)
H5	ABLUP	-0.05 (0.11)	0.52 (0.07)
DBH10	ABLUP_FULL	0.07 (0.02)	0.82 (0.01)
VOL10	ABLUP_FULL	0.07 (0.02)	0.82 (0.01)
GR	ABLUP_FULL	0.06 (0.02)	0.83 (0.01)
H5	ABLUP_FULL	0.07 (0.02)	0.82 (0.01)
H10	ABLUP_FULL	0.07 (0.02)	0.82 (0.01)
VOL10	GBLUP1	0.06 (0.15)	0.51 (0.11)
GR	GBLUP1	0.04 (0.13)	0.49 (0.10)
DBH10	GBLUP1	0.01 (0.12)	0.50 (0.09)
H5	GBLUP1	-0.08 (0.12)	0.43 (0.10)
H10	GBLUP1	0.02 (0.12)	0.50 (0.10)
DBH10	GBLUP2	0.09 (0.13)	0.59 (0.07)
VOL10	GBLUP2	0.08 (0.14)	0.59 (0.10)
GR	GBLUP2	0.05 (0.13)	0.59 (0.08)
H5	GBLUP2	-0.08 (0.11)	0.52 (0.07)
H10	GBLUP2	-0.01 (0.14)	0.60 (0.07)
VOL10	GBLUP3	0.05 (0.15)	0.60 (0.11)
DBH10	GBLUP3	0.03 (0.13)	0.59 (0.08)
GR	GBLUP3	0.03 (0.12)	0.58 (0.07)
H5	GBLUP3	-0.12 (0.11)	0.46 (0.09)
H10	GBLUP3	-0.05 (0.12)	0.55 (0.07)
VOL10	HBLUP1	0.07 (0.10)	0.50 (0.08)
GR	HBLUP1	0.06 (0.09)	0.49 (0.07)
DBH10	HBLUP1	0.04 (0.10)	0.52 (0.06)
H5	HBLUP1	-0.03 (0.09)	0.49 (0.06)
H10	HBLUP1	0.03 (0.09)	0.50 (0.06)
VOL10	HBLUP2	0.07 (0.09)	0.67 (0.06)
DBH10	HBLUP2	0.06 (0.10)	0.68 (0.04)
GR	HBLUP2	0.06 (0.09)	0.66 (0.04)
H5	HBLUP2	-0.06 (0.08)	0.61 (0.04)

H10	HBLUP2	0.00 (0.09)	0.69 (0.04)
VOL10	HBLUP3	0.06 (0.09)	0.65 (0.07)
GR	HBLUP3	0.05 (0.09)	0.66 (0.04)
DBH10	HBLUP3	0.04 (0.09)	0.67 (0.04)
H5	HBLUP3	-0.06 (0.08)	0.61 (0.05)
H10	HBLUP3	-0.01 (0.08)	0.68 (0.04)

*ABLUP represented the phenotypic data for the genotyped trees, ABLUP_FULL represented the full phenotypic data from the entire trial, GBLUP1 combined genotypic datasets, GBLUP2 SNP chip genotypic dataset, GBLUP3 Radseq genotypic dataset, HBLUP1 combined genotypic datasets, HBLUP2 SNP chip genotypic dataset, HBLUP3 Radseq genotypic dataset.

Supplemental Table S2. Narrow-sense heritability estimates (± standard errors) for five growth traits estimated using pedigree-based (ABLUP), genomic (GBLUP), and hybrid (HBLUP) models.

Trait	ABLUP	ABLUP_FULL	GBLUP1	GBLUP2	GBLUP3	HBLUP1	HBLUP2	HBLUP3
DBH10	0.15 (0.10)	0.05 (0.01)	0.05 (0.10)	0.15 (0.09)	0.05 (0.07)	0.09 (0.08)	0.08 (0.05)	0.06 (0.05)
GR	0.02 (0.09)	0.04 (0.01)	0.09 (0.10)	0.08 (0.09)	0.05 (0.07)	0.15 (0.08)	0.08 (0.05)	0.07 (0.05)
H5	0.00 (0.09)	0.05 (0.01)	0.00 (0.09)	0.00 (0.08)	0.00 (0.07)	0.00 (NaN)	0.00 (0.05)	0.00 (0.00)
H10	0.00 (0.09)	0.05 (0.01)	0.07 (0.10)	0.01 (0.08)	0.00 (NaN)	0.07 (0.08)	0.01 (0.04)	0.00 (0.04)
VOL10	0.11 (0.10)	0.06 (0.01)	0.13 (0.10)	0.13 (0.09)	0.06 (0.07)	0.17 (0.08)	0.09 (0.05)	0.10 (0.05)

*ABLUP represented the phenotypic data for the genotyped trees, ABLUP_FULL represented the full phenotypic data from the entire trial, GBLUP1 combined genotypic datasets, GBLUP2 SNP chip genotypic dataset, GBLUP3 Radseq genotypic dataset, HBLUP1 combined genotypic datasets, HBLUP2 SNP chip genotypic dataset, HBLUP3 Radseq genotypic dataset.

Supplemental Table S3. Variance Component Estimates (± SE) for Block, Residual, and Additive Genetic Effects Across Growth Traits Using ABLUP, GBLUP, and HBLUP Models.

Trait	Component	ABLUP	ABLUP_FULL	GBLUP1	GBLUP2	GBLUP3	HBLUP1	HBLUP2	HBLUP3
DBH10	Block	21.63 (8.91)	3.22 (0.86)	21.42 (8.87)	21.79 (8.96)	21.64 (8.95)	17.30 (5.25)	17.27 (5.24)	17.34 (5.26)
	Residual (units!R)	121.72 (15.92)	111.03 (1.43)	135.24 (16.36)	120.90 (14.59)	133.70 (14.19)	111.62 (9.81)	113.24 (7.71)	114.50 (7.59)
	Additive (vm(TreeID, ...))	21.76 (15.20)	5.62 (1.03)	7.77 (13.83)	22.15 (13.58)	7.64 (9.64)	11.47 (9.34)	9.42 (6.53)	7.33 (5.64)
GR	Block	7.69 (3.55)	1.37 (0.39)	7.90 (3.62)	7.78 (3.58)	7.80 (3.58)	9.46 (2.99)	9.40 (2.97)	9.41 (2.98)
	Residual (units!R)	83.06 (8.66)	98.59 (1.24)	76.81 (9.93)	78.21 (8.55)	79.73 (8.50)	69.45 (7.00)	74.28 (5.20)	74.35 (5.12)
	Additive (vm(TreeID, ...))	1.75 (7.26)	4.02 (0.84)	7.53 (8.65)	6.51 (7.43)	3.97 (5.77)	11.80 (6.98)	6.49 (4.47)	5.76 (3.93)
H5	Block	56.87 (29.75)	27.40 (7.34)	56.87 (29.76)	56.87 (29.76)	56.87 (29.79)	59.78 (21.37)	59.78 (21.36)	59.78 (21.37)
	Residual (units!R)	973.13 (101.14)	892.32 (11.58)	973.13 (108.70)	973.14 (96.78)	973.14 (99.83)	851.34 (35.33)	851.34 (51.70)	851.34 (35.33)
	Additive (vm(TreeID, ...))	0.00 (NA)	47.59 (8.45)	0.00 (NA)	0.00 (NA)	0.00 (NA)	0.00 (NA)	0.00 (NA)	0.00 (NA)
H10	Block	463.02 (204.44)	105.10 (28.52)	474.29 (208.00)	463.71 (204.62)	463.02 (204.38)	513.84 (159.52)	509.78 (158.65)	509.24 (158.57)
	Residual (units!R)	4246.88 (437.13)	4366.63 (56.45)	3937.86 (485.31)	4222.30 (411.41)	4246.91 (246.46)	3696.51 (324.00)	3943.77 (235.67)	3967.87 (234.73)
	Additive (vm(TreeID, ...))	0.00 (NA)	225.38 (40.90)	290.32 (414.32)	23.87 (326.71)	0.00 (NA)	285.36 (305.72)	26.28 (171.14)	1.78 (150.56)
VOL10	Block	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	Residual (units!R)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	Additive (vm(TreeID, ...))	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)

Model Key (Column Header Definitions):

- ABLUP – Pedigree-based model using genotyped trees only
- ABLUP_FULL – Pedigree-based model using the full phenotypic dataset
- GBLUP1 – Genomic BLUP using combined genotypic datasets
- GBLUP2 – Genomic BLUP using SNP chip data
- GBLUP3 – Genomic BLUP using RADseq data
- HBLUP1 – Hybrid BLUP using pedigree and genomic data (combined genotypes)
- HBLUP2 – HBLUP using SNP chip genotypes
- HBLUP3 – HBLUP using RADseq genotypes

Component Descriptors

- Block
  Represents the variance attributable to spatial or experimental design factors, such as field blocks or environmental differences across replicate plots. Modeled as a random effect to account for structure in the trial layout.
- Residual (units!R)
  Denotes the unexplained variance (error variance) within each observation after accounting for block and genetic effects. This captures noise due to micro-environmental variation, measurement error, or individual-specific factors not modeled elsewhere.
- Additive (vm(TreeID, gen_matrix))
  Represents the additive genetic variance associated with individual trees, modeled using a variance-covariance structure defined by the relationship matrix (gen_matrix). The vm(TreeID, gen_matrix) syntax indicates a variance model where the effect of each tree is modeled using a pedigree-based or genomic-based relationship matrix, depending on the model (ABLUP, GBLUP, or HBLUP).

Supplemental Table S4. Summary statistics for selected traits.

Trait	Min	Max	Average
DBH10	4	89	32.8
GR	0.4	73.8	38.9
H5	3	256	123.0
H10	95	540	317.5
VOL10	0.000002	0.002684	0.000371