


# Genomic Tools in Applied Tree Breeding Programs: Factors to Consider

Ross W. Whetten <sup>1,\*</sup> , Keith J. S. Jayawickrama <sup>2</sup>, W. Patrick Cumbie <sup>3</sup> and Gustavo S. Martins <sup>4</sup>

<sup>1</sup> Department of Forestry & Environmental Resources, North Carolina State University, Raleigh, NC 27606, USA

<sup>2</sup> Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331, USA

<sup>3</sup> Arborgen LLC, 2011 Broadbank Ct, Ridgeville, SC 29472, USA

<sup>4</sup> Riau Andalan Pulp and Paper, Pangkalan Kerinci, Riau 23800, Indonesia

\* Correspondence: ross\_whetten@ncsu.edu

**Abstract:** The past three decades have seen considerable research into the molecular genetics and genomics of forest trees, and a variety of new tools and methods have emerged that could have practical applications in applied breeding programs. Applied breeders may lack specialized knowledge required to evaluate claims made about the advantages of new methods over existing practices and are faced with the challenge of deciding whether to invest in new approaches or continue with current practices. Researchers, on the other hand, often lack experience with constraints faced by applied breeding programs and may not be well-equipped to evaluate the suitability of the method they have developed to a particular program. Our goal here is to outline social, biological, and economic constraints relevant to applied breeding programs to inform researchers, and to summarize some new methods and how they may address those constraints to inform breeders. The constraints faced by programs breeding tropical species grown over large areas in relatively uniform climates with rotations shorter than 10 years differ greatly from those facing programs breeding boreal species deployed in many different environments, each with relatively small areas, with rotations of many decades, so different genomic tools are likely to be appropriate.

**Keywords:** applied tree breeding; forest products; forest tree genomics



**Citation:** Whetten, R.W.; Jayawickrama, K.J.S.; Cumbie, W.P.; Martins, G.S. Genomic Tools in Applied Tree Breeding Programs: Factors to Consider. *Forests* **2023**, *14*, 169. <https://doi.org/10.3390/f14020169>

Academic Editors: Gancho Slavov and Jaroslav Klápště

Received: 8 November 2022

Revised: 20 December 2022

Accepted: 6 January 2023

Published: 17 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

“Applied tree breeding” is a generic term that includes a variety of breeding efforts, depending on the biology of the target tree species and the context for the breeding program [1]. For this short perspective, we define applied tree breeding to include long-term efforts to make or allow crosses with selected individual trees, measure a set of target traits in the progeny from those crosses, and choose a new set of selected individuals for the next round of crosses based on analysis of those measurements. Such programs typically have two main goals. The first is to estimate the genetic value of specific genetic entries, which may be clonal propagules, full-sibling families from controlled crosses, or open-pollinated families from known seed parents. Typical tree breeding programs assess thousands to tens of thousands of progeny to determine genetic value, over testing cycles ranging from a few years to several decades depending on the tree species. The second goal is to improve the average value of the target traits in the breeding population relative to the entire population of the species. Land managers use estimated genetic values in deciding which genetic material from the breeding program to use in establishing plantations. Plantation establishment and maintenance require substantial investment, and such investments are often made with the expectation of an economic return when the plantation is harvested, so there is good reason for a cautious, conservative approach to assessing the genetic value of propagules used for deployment. In contrast, decisions about which trees to cross do not carry the same expectation of a direct economic return, so breeders can be more flexible about testing new methods.

Tree breeding programs occur in different organizational contexts; some are proprietary programs within individual for-profit companies; others are collaborative programs with members representing both private-sector and public-sector entities, and others are entirely in the public sector. Despite these differences, many aspects of these diverse programs are similar. They are subject to biological constraints based on the life history of target tree species, to economic constraints that limit what can be done to achieve breeding objectives, and to social constraints that may limit what approaches are considered acceptable for implementation in tree improvement and forest management operations. In some cases, tree breeders work with diverse groups of stakeholders and must build consensus among those stakeholders to change established practices. Each of these constraints will be considered in more detail below, as all can affect whether a specific genomic technology is cost-effective in a particular breeding program.

Researchers pursuing fundamental understanding are often funded by government agencies through competitive grants and are charged with finding new knowledge and new applications for existing knowledge. This focus creates a strong incentive to seek the novel and to interpret the possible benefits of their research in the most favorable way. Researchers working to develop new products and services for service provider companies are also motivated to present their new research results in the most positive light. Researchers are encouraged to take risks, the cost of failure is relatively low, and there are few constraints to drastic changes in approach or strategy from experiment to experiment and grant to grant. This flexibility is important, but is inconsistent with the long-term planning and investment of applied breeders, which can lead to problems. For example, during the days of QTL mapping experiments, large full-sibling family block plantings were desirable experimental materials, and some breeders established such plantings at the request of researchers. By the time the plantings were large enough for phenotypic measurements, however, the research community had moved from QTL mapping in families to association mapping in diverse populations, and the researchers were no longer interested in the family-based materials. The net effect was that the breeders' time and efforts were wasted. Such experiences greatly reduce the enthusiasm of applied breeders to embrace cutting-edge methods and encourages a "wait and see" approach to determine which new technologies will stand the test of time before making any investment to test a new method.

Given the differences between fundamental research and applied breeding, it is not surprising that applied breeders working to achieve specific breeding goals often view the claims made by researchers as exaggerated. This situation seems unlikely to change unless the incentive structures created by government funding agencies and market forces change. Some funding agencies are now requiring evidence of industrial uptake or partnerships in specific funding programs, and opportunities for applied breeders and researchers to work together toward common goals do exist.

The overall value of a genomic technology to a tree breeding program depends on the characteristics of that breeding program and on the capabilities of the genomic tool. "Application" of genomic tools to breeding also means different things to different people. To researchers, the fact that SNP genotyping arrays are used to generate data from breeding populations may represent application, while to breeders the data may be considered a research result rather than an application unless the resulting genotyping data produce better selections, less work, or lower expenses than previous practices. Collaborations between applied breeding programs and researchers can help identify challenges faced by individual breeding programs and find genomic tools to address those challenges. These insights require participation from both researchers and breeders, development of cost-effective genomic tools to improve breeding outcomes, and appreciation of the different reward systems for researchers and breeders. Several factors affect the costs and benefits of application of genomic tools in breeding. Our goal in this short perspective is to review those factors and help breeders identify appropriate criteria to use in assessing the suitability of genomic tools for their own programs. This is not intended to be a

comprehensive review of the literature, but rather an expression of our opinions about the current state of the art, with references to examples as an entry point for further reading.

## 2. Constraints on Applied Breeding Programs

### 2.1. Social Constraints

The genetic diversity of planted forests is an issue of much greater social concern in some regions than in others [2]. In some regions of the tropics, a relatively small number of different clonally propagated elite genotypes of an exotic species can be planted over hundreds of thousands of hectares. In contrast, in boreal regions, dozens of genetic entries of a native species are typically planted in any planted forest to satisfy social expectations about the genetic diversity of forests. The expectation of genetic diversity is not necessarily a constraint on the nature of planting stock; genetically diverse plantations can be established with clonal propagules by including a large number of different clonal genotypes, as is proposed for Norway spruce in Sweden [3]. The larger the area that can be planted with a single genetic entry, the more the investment in producing that genetic entry can be amortized with other establishment costs. However, the stakeholders for some breeding programs may be opposed to using methods considered to be “un-natural”. Genome editing is likely to face this barrier in some regions, but even controlled crossing or vegetative propagation and deployment of elite genotypes are considered unacceptable by some.

Another social constraint is the diversity of opinions in large stakeholder groups, particularly if those groups hold some formal power to control the breeding program. Responsible tree breeders consider the larger context of social needs and preferences in their work. Nonetheless, breeders working in private-sector proprietary programs have different stakeholders than breeders working in public–private cooperative organizations or in public-sector agencies. In the private sector, the economic return from applying genomic tools can be an important consideration. In contrast, cooperative and public-sector breeders may be unable to point to any direct increase in revenue from the cost of costly genomic tools, which therefore makes the decision to use such tools more difficult to justify.

### 2.2. Biological Constraints

The reproductive biology of the target tree species is important, particularly with respect to the efficiency of controlled pollination and production of seeds and seedlings. Some species, such as commercial species of the genus *Acacia*, have flowering biology that makes seed production from controlled pollination extremely inefficient [4], although open-pollinated seed can be produced easily. Such species are typically bred using an open-pollinated mating design, but analyses of genetic value are much more powerful using a full-sibling model with both parents known [5]. It is now possible to use single-nucleotide polymorphism (SNP) or other genetic markers derived from genome analysis to identify the pollen parent for most seedlings from open-pollinated seed lots, particularly for species grown as exotics where the level of pollen contamination from surrounding trees is low [6,7]. The ability to vegetatively propagate the target species and the typical propagation strategy for deployment must also be considered. Species that can be cloned when juvenile but not after maturation can benefit from the application of genomic prediction methods that allow identification of high-value individuals when they are still young enough to propagate efficiently.

Additional biological factors include the size of the deployment region for each breeding population, the number of breeding populations, the availability of data and biological materials from previous cycles of breeding and testing, genome size and complexity of the target tree species, typical tree age at harvest, perceived risks of climate change and potential for novel pests or diseases, and genetic diversity of the species or species pool in the breeding population(s) [8]. Many of these biological factors relate to testing, selection, and deployment, rather than to breeding per se. The size of the deployment area affects the amortization of costs for development of elite genetic entries used. For example, consider

a deployment area of one million hectares. In the boreal region, this land area might be divided into 20 deployment zones, each with its own breeding population, e.g., Sweden, see [9]. Such zones stratify the region based on differences in elevation, latitude, or other environmental factors. The costs of using genomic tools for 20 breeding populations will be much higher than the costs for a single breeding program in a tropical region that can deploy the same genetic material over the full one million hectares.

### 2.3. Economic Constraints

A key issue is whether use of genomic tools adds costs to the breeding program, or if funds are simply shifted from one expense to another. If using genomic tools adds costs, it becomes important when those costs must be incurred and when the return on investment is received [10]. For example, a company in the seed orchard and nursery business can expect a return on investment from the sale of improved seedlings, without the need to wait for plantation establishment, growth, and harvest. On the other hand, if the return from investment is received only after the harvest of improved trees, which may be several decades for boreal species, then an unreasonably low discount rate may be required to yield a positive net present value for that investment. If no new expenses are incurred, and funds are simply shifted from one purpose to another, this issue is eliminated, so using genomic tools in ways that reduce other expenditures is a key to making them cost-effective [10].

## 3. Possible Applications of Genomic Tools in Practical Breeding

### 3.1. Marker-Assisted Selection

Traits controlled by a few genes have been targets for identifying markers linked to alleles with favorable effects on the trait. For example, marker-aided selection has often been tested in traits such as wood specific gravity [11–13] and resistance to fusiform rust disease in loblolly pine [14–16] or white pine blister rust disease in white pines [17,18]. These traits are often highly heritable, and simple phenotypic selection is generally effective for obtaining gains. The main advantage that markers can provide is the ability to select earlier and without the expense of direct phenotypic measurements. When the benefit does not justify the expense of marker genotyping, or if the breeding program needs to be able to make selections outside of the specific pedigree in which the marker-trait association was identified, then there is little justification for the use of markers in selection. One exception to this generalization might be the use of markers linked to rare alleles conferring disease resistance to aid in ‘pyramiding’ resistance genes or combining multiple genes from different backgrounds into a single individual [19]. Markers are better than phenotypic measurements in this case because one allele conferring resistance can mask the presence of another resistance allele unless a comprehensive collection of pathogen strains with different combinations of known virulence alleles is available for testing. Such resources are sometimes available for annual crop disease resistance breeding but are not common for forest trees.

### 3.2. Clonal Identity Validation in Seed Orchards and Breeding Programs

A small number of 10 to 40 informative markers can eliminate misidentified ramets and increase the genetic purity of harvested seed lots [20]. This application has immediate value to the customers purchasing family seed lots or seedlings, particularly those purchasing the most elite families, where any contaminants will degrade the average performance of the seed lot. Organizations that implement such quality control measures also benefit from the increase in consumer confidence in the genetic quality of planting stock [21].

### 3.3. Parentage Validation for Mass-Produced Control-Pollinated Seedlots

Depending on the relatedness levels between founding genotypes, a relatively small number of between 40 to 200 informative markers can be used to evaluate the genetic purity of control-pollinated seed lots [22,23]. As for clonal validation, this is valuable to

the purchaser of the seed lots or seedlings and to organizations that benefit from increased consumer confidence.

### 3.4. Parentage Reconstruction

A larger number of informative markers (>100) may be required to identify parentage, depending on whether the seed lots have any available information on seed-parent ancestry [6,7,22,23]. Another key factor is the number of different possible parents and whether all possible parents are available for genotyping—the fewer the possible parents and the greater the availability of possible parents for genotyping, the fewer the markers needed to achieve adequate confidence in parental assignments [24]. The value of parentage reconstruction is to allow full pedigree determination of progeny planted from open-pollinated seed lots. Doing so can increase the accuracy of genetic parameter estimates and allow the detection of non-additive effects, such as dominance or epistasis, that cannot be analyzed in open-pollinated progeny tests [5]. Furthermore, it may be the only way full-sibling family models can be applied to estimate genetic parameters for species where controlled crossing is inefficient or impossible, e.g., *Acacia*. The advantages of full-sibling genetic models over open-pollinated models are significant, so parentage reconstruction can be an important application of genomic tools in breeding [25].

### 3.5. Construction of Realized Genomic Relationship Matrices

Best linear unbiased prediction (BLUP) analysis of progeny test results can use a numerator relationship matrix based on pedigree records to estimate genetic covariances among individuals [26]. With pedigree data, the degree of resemblance between two individuals is measured using the coefficient of co-ancestry, which is simply the probability that two alleles at the same locus in the two individuals are identical by descent (IBD). However, it is possible to construct relationship matrices with molecular markers. The method most widely employed in forestry studies is that of Van Raden [27]. This uses SNP marker information exclusively, and estimates are computed through cross-products of marker genotypes deviated from mean allele frequencies and divided by the total heterozygosity at the locus. Van Raden's estimator reflects the actual proportion of marker alleles shared by two individuals and represents an identity-by-state (IBS) measure. Alternative methods have been developed for estimating additive genomic relationships [28], dominance genomic relationships [29], and IBD-based genomic relationship measures taking into account the IBD process, such as marker order and position within the genome and the segmental nature of DNA inheritance and linkage disequilibrium (LD) [30]. A marker-based realized genomic relationship matrix provides more accurate estimates of genetic parameters [31,32]. Thus, selection accuracy can be improved when genotype information is used with measured progeny phenotypes. Non-genotyped individuals can be included in the calculation of the relationship matrix if they have known relationships with genotyped individuals [33]. Tests of these methods with forest trees have reported significant improvement in the accuracy of breeding values estimated for the genotyped individuals but no significant improvement in the accuracy of breeding values for the non-genotyped individuals [25,34,35].

### 3.6. Genomic Prediction of Progeny Breeding Value

Some livestock and annual crop breeders routinely use marker information and a trained statistical model, without progeny phenotypes, to predict the breeding value of selection candidates at a very early age. Genomic prediction can shorten the breeding cycle, but relies heavily on LD between the genotyped markers and the unknown genetic variants that underlie phenotypic variation [36]. Such strong LD is rare in typical populations of forest trees but can be created in specialized breeding populations [37]. While genomic prediction uses marker data to advance breeding, it differs from marker-assisted selection in that no test for statistical significance of association between marker loci and trait phenotype(s) is performed. This strategy avoids the Type II error common to genome-wide



association studies (GWAS), which use very stringent statistical tests to control Type I (false-positive) errors at the cost of allowing high Type II (false-negative) error rates [28]. Instead, genomic prediction models are tested by cross-validation, using the model to predict genetic values of individuals not used in training the model, but which have both genotypes and known genetic values available [38].

These example applications differ in the costs of sampling, DNA extraction, genotyping, and the value returned to the breeding program. The value obtained from genomic tools depends on the size and structure of the breeding program and its social, biological, and economic constraints. In addition to understanding potential applications, breeders must understand the resources and tools required beyond genotyping. There is generally a gap between the analytical capabilities of researchers and what is available to practical breeders. To incorporate genomic tools, breeding programs may need to add personnel who can analyze and manage large genomic datasets. There are opportunities to outsource genomic data analysis, but this brings questions as to whether the methodology for analysis truly aligns with the breeding program and whether the analytical method can be fully understood or defended by the breeding program staff. The ideal situation requires internal genomic data capabilities to build confidence and maintain control of the breeding program within the organization.

#### 4. Example Scenarios for Application of Genomic Tools

##### 4.1. Tropical Species

Several examples of applications of genomic tools have been described in *Eucalyptus* species and hybrids, including progress toward genomic prediction [39,40], as well as use of realized relationship matrices in quantitative genetic analysis [41,42]. Another example comes from *Acacia crassicarpa* grown in the tropics by Riau Andalan Pulp and Paper with unpublished and published results [43]. This example illustrates key ideas, so will be explained in more depth. Like several *Acacia* species, *A. crassicarpa* is insect-pollinated, with reproductive biology that makes controlled crossing difficult [4]. Breeding was based on testing open-pollinated families to identify seed parents with desirable progeny phenotypes. Vegetative propagation is cost-effective but possible only with juvenile material. By the time phenotypes can be measured for assessment of genetic value, the selected tree can only be propagated by grafting. The deployment region consists of hundreds of thousands of hectares of relatively uniform climate conditions, so the ability to identify elite germplasm during the juvenile stage would allow the deployment over large areas. Identifying parents of progeny in open-pollinated families would permit the establishment of full-sibling genetic trials to enable the estimation of additive and non-additive effects, and specific and general combining ability. Given these advantages, the company decided to use genomic tools for marker discovery to allow parentage reconstruction and genomic prediction. The first step in this process was to support development of a chromosome-scale genome assembly for the species, followed by a population survey to identify SNP variants and estimate allele frequencies.

The company then synthesized three genotyping panels: a low-density panel of a few dozen SNPs for fingerprinting and parentage reconstruction, a mid-density panel of several hundred SNPs for imputation, and a higher-density panel containing tens of thousands of SNPs for constructing realized genomic relationship matrices and whole-genome regression. A large set of open-pollinated families with known seed parents was grown from seed collected from an elite seed orchard with a limited number of parents. A population of 100,000 seedlings in a nursery setting were genotyped with the low-density SNP panel. This allowed reconstruction of parentage for about two-thirds of the progeny and identification of at least 80 full-sibling families with enough progeny to establish a multi-site replicated field trial. This number of full-sibling families is more than half of the parental combinations possible in a half-diallel mating design. The seedlings with inferred parentage were used to establish a multi-site replicated progeny field test series. The field tests were designed to estimate specific combining effects and non-additive genetic

variance components. All progeny established in the trials were further genotyped with the mid-density panel for imputation of the high-density genotypes of the parents. The seed orchard that produced these progeny has a useful life of decades, while the rotation age for plantations is less than 10 years, so knowledge about the genetic value of parents in the orchard can be exploited for multiple rotations. Identifying progeny with high genetic value during the juvenile phase will allow use of a deployment strategy that should increase the yield and quality of product harvested from plantations within 12 to 15 years. The social environment in which the company operates is permissive to these plans. If genomic prediction allows within-family selection, then the gains in productivity and quality will be even higher, but just using parental reconstruction will provide substantial benefit relative to the open-pollinated breeding and deployment strategy used previously.

#### 4.2. Temperate-Region and Boreal Species

For some hardwood and conifer breeding programs, genetic markers are routinely used for validation of clone identity in seed orchards and parents in controlled crosses. The cost of genotyping is justified by the added value of correcting the identity of mis-identified parents in seed orchards and correcting pedigree errors from previous mistakes [44,45]. When existing genomic information is available for the species, individual breeding programs do not have to bear the cost of developing the marker panels or optimizing the genotyping procedures.

Widespread use of genomic prediction does not yet appear to be cost-effective for most temperate and boreal forest tree species. Progeny testing is currently less expensive than genotyping for most breeding programs, but this may change with time. Costs of progeny testing continue to rise, and genotyping costs may be reduced by technical advances, so the economic balance may shift in the future, and genotyping has the advantage of allowing earlier selection. In addition to genotyping cost, another significant barrier to widespread use of genomic prediction is the cost of collecting tissue samples from thousands of test trees and extracting DNA. The accuracy of genomic prediction may well be lower than phenotypic selection but could still provide more genetic gain per year if breeding cycle length can be reduced substantially [10]. Use of genomic predictions for making deployment decisions is likely to require more time and testing than for breeding because of the much higher economic stakes involved. Many forestland managers are now very comfortable deploying seed from selections that were rigorously tested in the field for a quarter to a third of a rotation period but may be less confident in planting thousands of hectares from selections made based on a genomic prediction model.

Within-family selection is one case where genomic selection could add genuine value to applied tree improvement in conifers. In this case, reduced costs of progeny testing could offset the costs of DNA extraction and genotyping. The best trees within each family are often indistinguishable in terms of the best linear unbiased prediction (BLUP) based on data from field trial measurements. Without genomic tools, breeders have two options. One option is to make multiple selections from the best families, and then test them all in the next cycle of progeny tests to increase the likelihood of finding a true winner. An alternative option is to select only one offspring per family and accept that the selected tree may not be the best in that family. The use of the realized relationship matrix plus progeny phenotypes increases the accuracy of within-family selection, allowing selection of a single offspring per family while maintaining confidence that an elite individual was identified. This approach could reduce the number of different families that must be tested, and therefore reduce the cost of progeny testing enough to cover the costs of genotyping. Selective genotyping of progeny within selected families would reduce genotyping costs, and joint analysis of progeny from genotyped and non-genotyped families together has been shown to improve selection accuracy for genotyped individuals [25,34,35].

The prospects for successful use of genomic selection seem highest within a breeding program where personnel are familiar with both genomic data analysis and the long-term needs and priorities of the program. Financial support is best derived from funds that

support operational breeding, rather than from competitive grants. Efforts in spruce have shown promise with the application toward long-term gains in longer-rotation species like lodgepole pine [46] and in spruce programs [47,48]. This model has worked well in integrating genomic prediction and other advanced methods in crop breeding by large germplasm companies. Grant-funded fundamental research is important as a means of developing new methods and testing new ideas at pilot-scale, but generally cannot carry projects through the long periods required to validate outcomes in advanced generations of tree breeding programs.

## 5. Conclusions

A variety of social, biological, and economic factors affect the potential costs and benefits from using genomic tools in applied breeding programs. These factors should be considered by breeders and by researchers interested in collaborating on genomic methods for applied tree breeding. Collaborations between applied breeders and fundamental researchers can be very helpful in overcoming obstacles to the implementation of genomic tools.

Due to the long-term nature of tree breeding programs, care is needed in making changes and modifications to these programs. Some barriers to change will be economic, and others will involve the details of practical application. Breeders may perceive that they have to abandon proven methods to incorporate new technology. This perception can create an “either-or” situation where they feel they must either continue with a current method that works or adopt a risky new technology. It is best to avoid this dichotomy and supplement existing methods with genomic tools where possible. To incorporate new tools, breeders must be willing to adapt. Incentives to use genomic tools may come from economic drivers to reduce cycle time, improve selection efficiency, or increase gain. These incentives are balanced by the limits imposed by both the biological constraints of the target species and the constraints imposed by the society where the breeding program operates [49]. Public-sector breeding programs that are not primarily driven by economic gain may find sufficient justification for use of genomic tools in the social pressures for better management of forest diversity and resilience to climate change. Research in those environments should focus on demonstration and economic valuation of those types of benefits, in addition to any economic benefits that can also be captured [50]. Breeding programs that are primarily focused on economic gain may evaluate the use of genomic tools primarily in terms of cost versus genetic gain but should not overlook the value of diversity and resilience factors as well. Breeders in all kinds of breeding programs must also be willing to reevaluate how their breeding program works and where money is spent.

Applied breeding programs will need to consider the adoption of genomic tools with a good understanding of the risk involved. To manage risk, it may be helpful to apply these tools in specific populations until the value can be assessed. Results from academic research may not translate directly to breeding applications, but the effort to incorporate new technologies in test populations could generate new ideas to apply genomics successfully in other aspects of the breeding effort.

**Author Contributions:** Conceptualization, R.W.W.; writing—original draft preparation, R.W.W. and K.J.S.J.; writing—review and editing, R.W.W., K.J.S.J., W.P.C. and G.S.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors received no external funding to write this review. Some of the data described were produced by company researchers supported by internal funds from Riau Andalan Pulp and Paper.

**Data Availability Statement:** All data described are from published articles with the exception of the *Acacia crassiparva* data which are the property of Riau Andalan Pulp and Paper. Readers interested in access to those proprietary data are encouraged to contact Gustavo S. Martins to discuss their interest.

**Conflicts of Interest:** The authors declare no conflict of interest.



## References

1. Zobel, B.J.; Talbert, J. *Applied Tree Improvement*; Blackburn Press: Caldwell, NJ, USA, 1984.
2. Ivetić, V.; Devetaković, J. Concerns and evidence on genetic diversity in planted forests. *Reforest* **2017**, *3*, 196–207. [\[CrossRef\]](#)
3. Rosvall, O.; Bradshaw, R.H.W.; Egertsdotter, U.; Ingvarsson, P.K.; Mullin, T.J.; Wu, H. Using Norway spruce clones in Swedish forestry: Implications of clones for management. *Scand. J. For. Res.* **2019**, *34*, 390–404. [\[CrossRef\]](#)
4. Griffin, A.R.; Vuong, T.D.; Harbard, J.L.; Wong, C.Y.; Brooker, C.; Vaillancourt, R.E. Improving controlled pollination methodology for breeding *Acacia mangium* Willd. *New For.* **2010**, *40*, 131–142. [\[CrossRef\]](#)
5. Lynch, M.; Walsh, B. *Genetics and Analysis of Quantitative Traits*; Oxford University Press: London, UK, 1998.
6. Lambeth, C.; Lee, B.-C.; O'Malley, D.; Wheeler, N. Polymix breeding with parental analysis of progeny: An alternative to full-sib breeding and testing. *Theor. Appl. Genet.* **2001**, *103*, 930–943. [\[CrossRef\]](#)
7. El-Kassaby, Y.A.; Lstibůrek, M. Breeding without breeding. *Genet. Res.* **2009**, *91*, 111–120. [\[CrossRef\]](#) [\[PubMed\]](#)
8. White, T.L.; Adams, W.T.; Neale, D.B. *Forest Genetics*; CABI Publishing: Wallingford, UK, 2007.
9. Danell, Ö. Survey of past, current and future Swedish forest tree breeding. *Silva Fenn.* **1991**, *25*, 241–247. [\[CrossRef\]](#)
10. Fugeray-Scarbel, A.; Ben-Sadoun, S.; Bouchet, S.; Lemarié, S. Analyzing the economic effectiveness of genomic selection relative to conventional breeding approaches. In *Genomic Prediction of Complex Traits: Methods and Protocols*, Nourollah Ahmadi, N.; Bartholomé, J., Ed.; Springer Nature: London, UK, 2022; pp. 619–643. [\[CrossRef\]](#)
11. Brown, G.R.; Gill, G.P.; Kuntz, R.J.; Langley, C.H.; Neale, D.B. Nucleotide diversity and linkage disequilibrium in loblolly pine. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15255–15260. [\[CrossRef\]](#)
12. Devey, M.E.; Groom, K.A.; Nolan, M.F.; Bell, J.C.; Dudsinski, M.J.; Old, K.M.; Matheson, A.C.; Moran, G.F. Detection and verification of quantitative trait loci for resistance to *Dothistroma* needle blight in *Pinus radiata*. *Theor. Appl. Genet.* **2004**, *108*, 1056–1063. [\[CrossRef\]](#)
13. Williams, C.; Neale, D.B. Conifer wood quality and marker-aided selection: A case study. *Can. J. For. Res.* **2011**, *22*, 1009–1017. [\[CrossRef\]](#)
14. Wilcox, P.L.; Amerson, H.V.; Kuhlman, E.G.; Liu, B.H.; O'Malley, D.M.; Sederoff, R.R. Detection of a major gene for resistance to fusiform rust disease in loblolly pine by genomic mapping. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 3859–3864. [\[CrossRef\]](#)
15. Cumbie, W.P.; Huber, D.A.; Steel, V.C.; Rottmann, W.; Cannistra, C.; Pearson, L.; Cunningham, M. Marker associations for fusiform rust resistance in a clonal population of loblolly pine (*Pinus taeda*, L.). *Tree Genet. Genomes* **2020**, *16*, 86. [\[CrossRef\]](#)
16. Lauer, E.; Isik, F. Major QTL confer race-nonspecific resistance in the co-evolved *Cronartium quercuum* f. sp. *Fusiforme*—*Pinus taeda* pathosystem. *Heredity* **2021**, *127*, 288–299. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Snieszko, R.A.; Smith, J.; Liu, J.-J.; Hamelin, R.C. Genetic resistance to fusiform rust in southern pines and white pine blister rust in white pines—A contrasting tale of two rust pathosystems—Current status and future prospects. *Forests* **2014**, *5*, 2050–2083. [\[CrossRef\]](#)
18. Liu, J.-J.; Snieszko, R.A.; Zamany, A.; Williams, H.; Wang, N.; Kegley, A.; Savin, D.P.; Chen, H.; Sturrock, R.N. Saturated genic SNP mapping identified functional candidates and selection tools for the *Pinus monticola* Cr2 locus controlling resistance to white pine blister rust. *Plant Biotechnol. J.* **2017**, *15*, 1149–1162. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Dormatey, R.; Sun, C.; Ali, K.; Coulter, J.A.; Bi, Z.; Bai, J. Gene Pyramiding for sustainable crop improvement against biotic and abiotic stresses. *Agronomy* **2020**, *10*, 1255. [\[CrossRef\]](#)
20. Nybom, H.; Weising, K.; Rotter, B. DNA fingerprinting in botany: Past, present, future. *Investig. Genet.* **2014**, *5*, 1. [\[CrossRef\]](#)
21. White, T.L.; Davis, J.; Gezan, S.; Hulcr, J.; Jokela, E.; Kirst, M.; Martin, T.A.; Peter, G.; Powell, G.; Smith, J. Breeding for value in a changing world: Past achievements and future prospects. *New For.* **2014**, *45*, 301–309. [\[CrossRef\]](#)
22. Marshall, T.C.; Slate, J.; Kruuk, L.E.B.; Pemberton, J.M. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* **1998**, *7*, 639–655. [\[CrossRef\]](#)
23. Jones, A.G.; Ardren, W.R. Methods of parentage analysis in natural populations. *Mol. Ecol.* **2003**, *12*, 2511–2523. [\[CrossRef\]](#)
24. Jones, A.G.; Clayton, M.S.; Paczolt, K.A.; Ratterman, N.L. A practical guide to methods of parentage analysis. *Mol. Ecol. Resour.* **2010**, *10*, 6–30. [\[CrossRef\]](#)
25. Ratcliffe, R.; El-Dien, O.G.; Cappa, E.P.; Porth, I.; Klápšte, J.; Chen, C.; El-Kassaby, Y.A. Single-Step BLUP with varying genotyping effort in *Picea glauca*. *Genes Genom. Genet.* **2017**, *7*, 935–942.
26. Henderson, C.R. *Applications of Linear Models in Animal Breeding*; University of Guelph: Guelph, ON, Canada, 1984.
27. VanRaden, P.M. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **2008**, *91*, 4414–4423. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Yang, J.; Benyamin, B.; McEvoy, B.P.; Gordon, S.; Henders, A.K.; Nyholt, D.R.; Madden, P.A.; Heath, A.C.; Martin, N.G.; Montgomery, G.W.; et al. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **2010**, *42*, 565–569. [\[CrossRef\]](#)
29. Vitezica, Z.G.; Varona, L.; Legarra, A. On the additive and dominant variance and covariance of individuals within the genomic selection scope. *Genetics* **2013**, *195*, 1223–1230. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Han, L.; Abney, M. Using identity by descent estimation with dense genotype data to detect positive selection. *Eur. J. Hum. Genet.* **2013**, *21*, 205–211. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Hayes, B.J.; Visscher, P.M.; Goddard, M.E. Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res.* **2009**, *91*, 47–60. [\[CrossRef\]](#)

32. Jurcic, E.J.; Villalba, P.V.; Pathauer, P.S.; Palazzini, D.A.; Oberschelp, G.P.J.; Harrand, L.; Garcia, M.N.; Aguirre, N.C.; Acuña, C.V.; Martínez, M.C.; et al. Single-step genomic prediction of *Eucalyptus dunnii* using different identity-by-descent and identity-by-state relationship matrices. *Heredity*. **2021**, *127*, 176–189. [\[CrossRef\]](#)
33. Legarra, A.; Aguilar, I.; Misztal, I. A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* **2009**, *92*, 4656–4663. [\[CrossRef\]](#)
34. Ukrainetz, N.K.; Mansfield, S.D. Prediction accuracy of single-step BLUP for growth and wood quality traits in the lodgepole pine breeding program in British Columbia. *Tree Genet. Genomes* **2020**, *16*, 64. [\[CrossRef\]](#)
35. Walker, T.D.; Cumbie, W.P.; Isik, F. Single-Step Genomic Analysis Increases the Accuracy of Within-Family Selection in a Clonally Replicated Population of *Pinus taeda* L. *For. Sci.* **2022**, *68*, 37–52. [\[CrossRef\]](#)
36. Hickey, J.M.; Chiurugwi, T.; Ian Mackay, I.; Powell, W. Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nat. Genet.* **2017**, *49*, 1297–1303. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Grattapaglia, D.; Resende, M.D.V. Genomic selection in forest tree breeding. *Tree Genet. Genomes* **2011**, *7*, 241–255. [\[CrossRef\]](#)
38. Daetwyler, H.D.; Calus, M.P.L.; Pong-Wong, R.; de los Campos, G.; Hickey, J.M. Genomic Prediction in Animals and Plants: Simulation of Data, Validation, Reporting, and Benchmarking. *Genetics* **2013**, *193*, 347–365. [\[CrossRef\]](#)
39. Durán, R.; Isik, F.; Zapata-Valenzuela, J.; Balocchi, C.; Valenzuela, S. Genomic predictions of breeding values in a cloned *Eucalyptus globulus* population in Chile. *Tree Genet. Genomes* **2017**, *13*, 74. [\[CrossRef\]](#)
40. Tan, B.; Grattapaglia, D.; Martins, G.S.; Ferreira, K.Z.; Sundberg, B.; Ingvarsson, P.K. Evaluating the accuracy of genomic prediction of growth and wood traits in two *Eucalyptus* species and their F1 hybrids. *BMC Plant Biol.* **2017**, *17*, 110. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Lima, B.M.; Cappa, E.P.; Silva-Junior, O.B.; Garcia, C.; Mansfield, S.D.; Grattapaglia, D. Quantitative genetic parameters for growth and wood properties in *Eucalyptus “urograndis”* hybrid using near-infrared phenotyping and genome-wide SNP-based relationships. *PLoS ONE* **2019**, *14*, e0218747.
42. Callister, A.N.; Bradshaw, B.P.; Elms, S.; Gillies, R.A.W.; Sasse, J.M.; Brawnner, J.T. Single-step genomic BLUP enables joint analysis of disconnected breeding programs: An example with *Eucalyptus globulus* Labill. *Genes Genomes Genet.* **2021**, *11*, jkab253. [\[CrossRef\]](#)
43. Martins, G.S.; Yuliarto, M.; Antes, R.; Sabki; Prasetyo, A.; Unda, F.; Mansfield, S.D.; Hodge, G.R.; Acosta, J.J. Wood and Pulp Properties Variation of *Acacia crassicaarpa* A.Cunn. ex Benth. and Sampling Strategies for Accurate Phenotyping. *Forests* **2020**, *11*, 1043. [\[CrossRef\]](#)
44. Grattapaglia, D.; Diener, P.S.D.A.; Dos Santos, G.A. Performance of microsatellites for parentage assignment following mass controlled pollination in a clonal seed orchard of loblolly pine (*Pinus taeda* L.). *Tree Genet. Genomes* **2014**, *10*, 1631–1643. [\[CrossRef\]](#)
45. Plomion, C.; LeProvost, G.; Pot, D.; Vendramin, G.; Gerber, S.; Decroocq, S.; Brach, J.; Raffin, A.; Pastuszka, P. Pollen contamination in a maritime pine polycross seed orchard and certification of improved seeds using chloroplast microsatellites. *Can. J. For. Res.* **2001**, *31*, 1816–1825. [\[CrossRef\]](#)
46. Cappa, E.P.; Chen, C.; Klutsch, J.G.; Sebastian-Azcona, J.; Ratcliffe, B.; Wei, X.; Da Ros, L.; Ullah, A.; Liu, Y.; Benowicz, A.; et al. Multiple-trait analyses improved the accuracy of genomic prediction and the power of genome-wide association of productivity and climate change-adaptive traits in lodgepole pine. *BMC Genom.* **2022**, *23*, 536. [\[CrossRef\]](#)
47. Beaulieu, J.; Doerksen, T.; Clément, S.; MacKay, J.; Bousquet, J. Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. *Heredity* **2014**, *113*, 343–352. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Lenz, P.R.N.; Nadeau, S.; Azaiez, A.; Gérardi, S.; Deslauriers, M.; Perron, M.; Isabel, N.; Beaulieu, J.; Bousquet, J. Genomic prediction for hastening and improving efficiency of forward selection in conifer polycross mating designs: An example from white spruce. *Heredity* **2020**, *124*, 562–578. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Howe, G.T.; Jayawickrama, K.J.S.; Cherry, M.; Johnson, G.R.; Wheeler, N.C. Breeding Douglas-fir. *Plant Breed. Rev.* **2006**, *27*, 245–353.
50. Porth, I.; Bull, G.; Ahmed, S.; El-Kassaby, Y.A.; Boyland, M. Forest genomics research and development in Canada: Priorities for developing an economic framework. *For. Chron.* **2015**, *91*, 60–70. [\[CrossRef\]](#)

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.