Managing Cold Hardiness in a Tree Breeding Program while Balancing Gain and Diversity

Trevor D. Walker[1] and Fikret Isik[1]

[1]Cooperative Tree Improvement Program, Department of Forestry and Environmental Resources, NC State University, Raleigh, NC; *trevor_walker@ncsu.edu

Abstract: In addition to the balance between gain and diversity, forest tree breeding programs must also ensure that adaptability is maintained when choosing the trees and crosses to breed. Maintaining adaptability can be a challenge when there are undesirable genetic correlations, such as for cold hardiness and growth rate in *Pinus taeda*. Breeding zones have traditionally been used to maintain adaptability, but rigid demarcations might result in a sub-optimal crossing plan when the adaptability variation is clinal. For example, a strict breeding zone strategy would exclude crosses among trees that may have similar provenance origins but are separated by the zone boundary, such as the upper coastal plain and lower piedmont. Further, provenance origins become difficult to define after a few cycles of breeding, especially when trees from wide crosses are selected for breeding. A better approach may be to permit crosses among all individuals but apply a negative weight against wide crosses in a mate selection algorithm. Using the NC State University *Pinus taeda* breeding program as an example, we demonstrate this approach and evaluate the gain, diversity, and adaptability consequences for different scenarios for managing cold hardiness. The scenarios were evaluated using the differential evolution algorithm as applied by MateSel, which has grown in popularity among livestock, aquaculture, and plant breeding programs. The flexible breeding zones scenario provided the most desirable solution but required close monitoring of the results. Further research should include simulations over multiple generations to evaluate the long-term consequences and identify rules-of-thumb for choosing the appropriate weighting values.

INTRODUCTION

Tree breeding programs often use breeding zones to maintain adaptability within a species. A breeding zone is the set of environments for which improved varieties are being developed, and corresponds to a particular geographic region (White et al. 2007). Each breeding zone essentially acts as a distinct population, so there is an upfront cost of maintaining many breeding zones. There is also an opportunity cost of reduced gain per unit time, as managing multiple zones distracts from the number of crosses, progeny tested, and selections that can be afforded in any particular zone. However, too few breeding zones risks the loss of adaptability in part of the deployment range.

Because the inclusion of non-adapted material in a breeding zone will reduce overall suitability, breeders are challenged to define how wide any single zone can be before incurring a substantial loss of local adaptability (Namkoong et al. 1988).

Provenance-progeny trials (also known as common garden experiments) are critical for defining breeding zones. By testing every seed source in a replicated experimental design on multiple sites across the range of deployment environments, these trials allow evaluation of the important adaptability traits within a species. These trials also allow evaluation of the genetic correlations between adaptability traits and the traits targeted for selection. For example, provenance trials for *Pinus taeda* indicated that cold hardiness is the primary trait affecting adaptability, and that genotypes with warmer-source origins tend to have faster growth (Schmidtling 2001; Farjat et al. 2017). Using provenance trials and climatic date, Schmidtling (1994) presented a seed transfer distance metric for *P. taeda* that is calculated as the difference between the mean minimum winter temperature (MWT) of the genetic source origin (e.g., where the parent tree was selected) and the MWT of the planting site. It was found that genetic sources from warmer climates often outperformed the local source for growth rate whenever the seed transfer distance was 5°F or less, yet transfer distances greater than 10°F incurred cold damage that offset the gains in growth. Following the observations from provenance trials and progeny tests, the NC State University Cooperative Tree Improvement Program organized three breeding zones that represented the Coastal, Piedmont, and Northern deployment regions of the southeastern United States in their third and fourth cycles (McKeand and Bridgwater 1998; Isik and McKeand 2019). The purpose of these breeding zones was to ensure that cold hardiness is maintained within the Piedmont and Northern zones, because there is a strong undesirable genetic correlation between cold hardiness and growth, which results in a temptation to use less cold hardy material when designing mating plans.

Strict adherence to breeding zones is simple and effective but has limitations (in addition to the upfront and opportunity costs of managing multiple populations). The rigid demarcations between zones do not have a biological justification for clinal traits like cold hardiness. The zones arbitrarily prevent crosses among parents with genetic origins near the border, even though they may be similar in terms of adaptability. Further, in practicality, breeding programs have a history of experimenting with "wide" crosses (matings between parents from different zones), and selections made from these crosses will not have a straightforward zone assignment. An example of such a selection with a wide pedigree is shown in Figure 1, where the founders (represented by the top row of circles) come from disparate climates. The figure also demonstrates how MWT is calculated for advanced-generation selections using mid-parent values. This approach assumes polygenic inheritance of cold hardiness, which is a reliable premise given our current understanding of forest trees (Howe et al. 2003; Wisniewski et al. 2018; De La Torre et al. 2021).

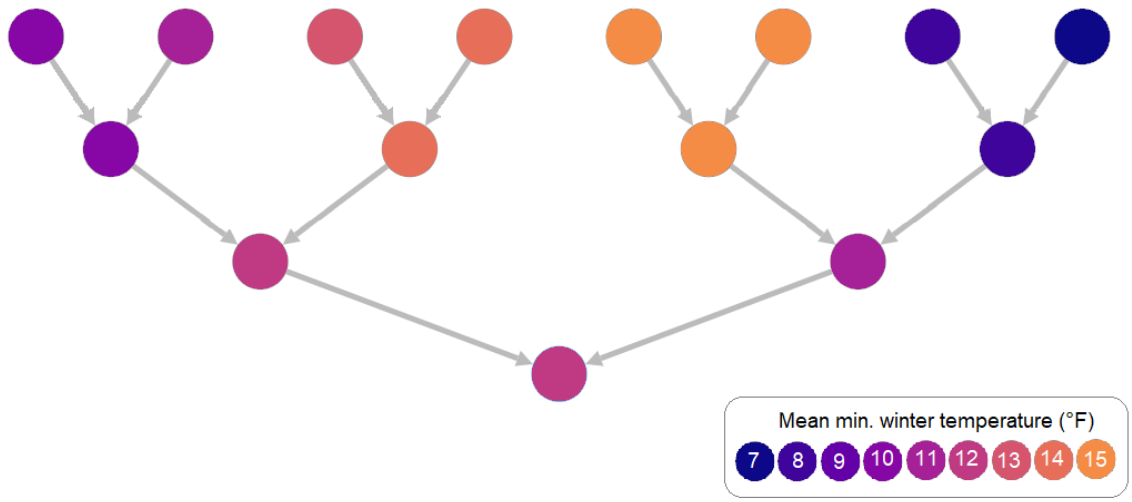

Figure 1. Illustration of the mean minimum winter temperature calculation for an advanced cycle selection with a particularly diverse pedigree of founders across the range of cold hardiness.

When developing a mating plan, one quickly encounters the "breeder's dilemma": the conflict between maximizing the short-term response to selection and maintaining diversity to avoid inbreeding and allow continued extraction of gain over the long-term (Lindgren and Mullin 1997). Modern breeding programs use mate selection algorithms (such as optimum contribution selection) to design their crossing plans to balance short term gain and preserve diversity (Mullin 2017; Yamashita et al. 2018; Isik and McKeand 2019). Computer programs are a necessity after the first two generations because pedigrees become too complex to design matings manually. The differential evolution algorithm is a popular method in livestock, aquaculture, and plant breeding (Kinghorn 2011). The objective function uses a weighting between 0 and 90 degrees that corresponds to the emphasis placed on short-term gain and diversity in the progeny population, with lower angles indicating a higher emphasis on short-term gain. The commercial software MateSel applies this algorithm to design mating plans while also incorporating many other features to accommodate the complex logistics of a breeding plan, such as biological limits on mating frequency per parent, accounting for crosses and selections already made but not yet reproductively mature, and logistical constraints on mating animals from different farms (Kinghorn and Kinghorn 2021). In this study, we investigate using the latter feature to introduce a concept of "flexible breeding zones", where the breeding population is managed as one, but a negative weighting is used to discourage making crosses among parents from different zones. We expect that managing the population as one will result in more short-term gain and adequate diversity compared to strict breeding zones, and we compare a few scenarios here to evaluate.

MATERIALS AND METHODS

Pine breeding population

The breeding population used in this study consisted of 369 selections made for the Cooperative Tree Improvement Program's 5th-Cycle breeding plan, which were topgrafted at the Arrowhead Breeding Center near Hawkinsville, Georgia, United States and in members' orchards across the southeastern United States. These selections were made from 4th-Cycle progeny tests planted across the southeastern United States (Figure 2 (a)). The progeny tests were grouped into series (CYCLE4C, CYCLE4P, and CYCLE4N) that correspond to the set of families from the corresponding breeding zone (Coastal, Piedmont, and Northern, respectively), but there is overlap both in the geography of test sites and in the families tested. Around a third of the families in the CYCLE4P tests are from the Coastal and Northern breeding zones, and around a third of the families in CYCLE4N are from the Piedmont and Coastal breeding zones, although care was taken to exclude maladapted families. The overlap in families was done to connect the series so that breeding values for all three populations could be estimated on the same scale.

The 5th-cycle selections were made based on an index calculated from individual-tree breeding values for stem volume, straightness, and disease resistance using weights of 60%, 20%, and 20%, respectively, from a standard normal transformation for each trait (where bigger values are more desirable). The MWT has a positive correlation (r = 0.32) with the index breeding value (Figure 2 (b)), driven by the correlation between MWT and volume breeding values.

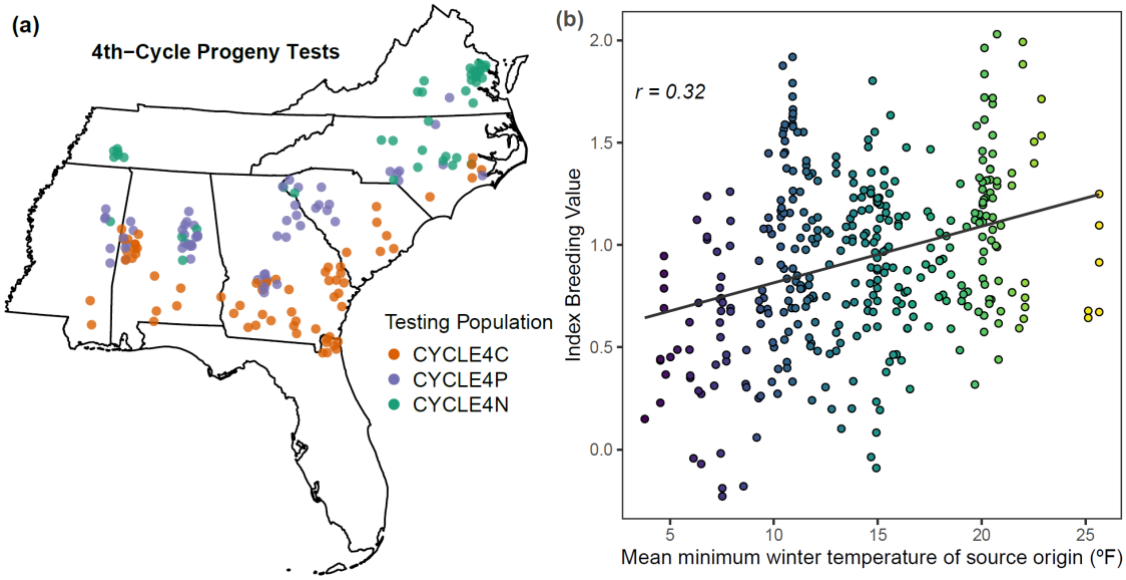

Figure 2. (a) Locations of 4th-cycle progeny tests in the southeastern United States. (b) Relationship between the index breeding value and mean minimum winter temperature for the 369 5th-cycle selections considered in this analysis.

The 369 selections varied in their cold hardiness, and the distribution of MWT was selected to reflect the distribution of area in the deployment region, ranging from 3.8°F to 25.7°F (Figure 3).

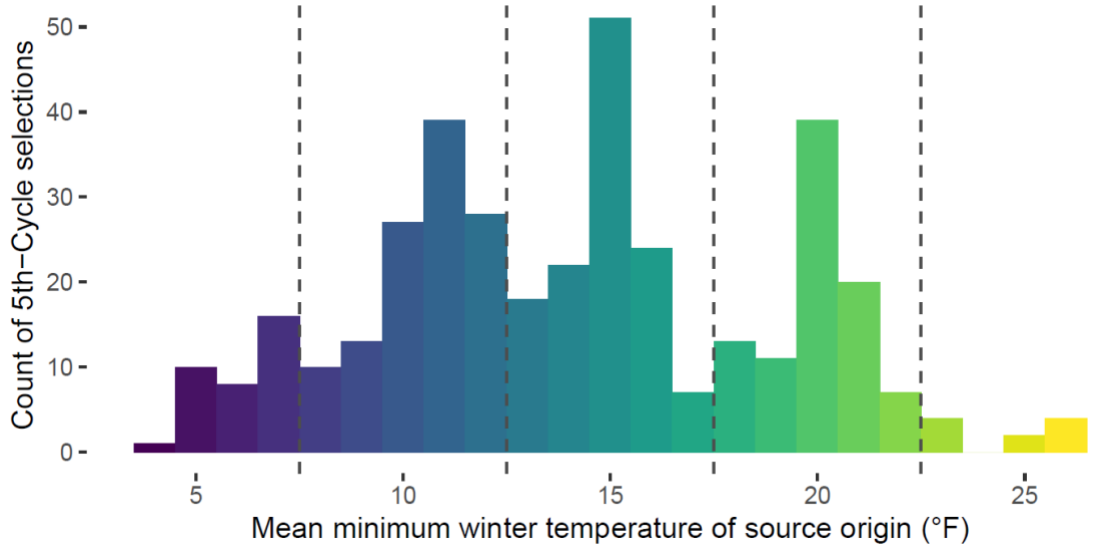

Figure 3. Distribution of mean minimum winter temperatures for 369 5th-Cycle selections used in the study. Vertical lines represent bins of 5°F that were used as mating groups in the flexible breeding zone scenario.

Mating scenarios

Three mating scenarios were compared: 1) no breeding zones, 2) strict breeding zones, and 3) flexible breeding zones. The first scenario represents a do-nothing approach and illustrates the consequences of ignoring the need for cold hardiness. The second scenario represents the traditional approach, and selections were assigned to three separate breeding zones based on MWT breaks of 8°F and 13°F. The third scenario represents an attempt to generalize the breeding zone scenario to match the clinal nature of cold hardiness.

For the flexible breeding zones scenario, we utilized the mating group functionality of MateSel, which was originally designed to place logistical restrictions on mating livestock from different farms. In our application, selections were assigned to five MWT bins with widths of 5°F that corresponded to mating groups (Figure 3). A target number of matings within each mating group was provided. A mating permission matrix was designed that allowed crosses among adjacent mating groups (MWT bins), but not more distant groups. For example, matings between selections from the <7.5°F and 7.5°F – 12.5°F bins are allowed, but not between selections from the <7.5°F and 12.5°F – 17.5°F bin. The exception is that the permission matrix allowed matings between the

third group (12.5°F - 17.5°F) and the fifth group (MWT > 22.5°F), because the geographic region corresponding to these MWT bins can be considered a single breeding zone (the Coastal region). A weighting parameter is used by the algorithm to indicate how much deviation from the target number of matings within each group. A weight of 0 places no emphasis on the targets and is equivalent to the no constraints scenario (scenario 1). Preliminary analysis indicated that a weighting factor of 0.70 provided a desirable solution, so that was the value used in this analysis.

In all three scenarios, 500 crosses were requested and the number of crosses per breeding zone/mating group was made proportional to the number of selections in that group. The target degrees parameter was set at 35 (a typical value for the Cooperative's breeding program) and selections were allowed to be used as both female and male. No more than 9 matings per selection were allowed.

Metrics considered

Each scenario produced a mating list (list of trees to cross and their mating pairs). To evaluate the cold hardiness of the resulting mating lists, the mid-parent MWT was calculated for each cross, and the distributions were evaluated in a figure. This can also be considered the distribution of the progeny. To evaluate short-term gain, the mid-parent index value for each cross was calculated, and the distributions were evaluated. To evaluate the diversity of each mating list, we reported the number of parents used in the mating, the number of founders represented, and the mean coancestry coefficient among progeny. Also presented is the status number, an estimate of the effective population size (Lindgren et al. 1996).

RESULTS

Maintenance of cold hardiness

Ignoring the adaptability requirements (no breeding zones) resulted in a mating list where all but one progeny/cross had a mid-parent MWT greater than 8°F (Figure 4). The strict breeding zones and flexible breeding zones scenarios had similar distributions that were trimodal. The minimum mid-parent MWT for the strict breeding zones and flexible breeding zones were the same (both were 5.3°F). However, the mean for the strict breeding zones was lower (14.3°F) compared to that of the flexible breeding zones (15.2°F), which was actually slightly higher than the mean for the no breeding zones scenario (15.1°F). In all three scenarios, the distribution of cross MWT values had a warmer minimum than that of the breeding candidates, although the scenarios with strict and flexible breeding zones had appropriate values for the target deployment range.

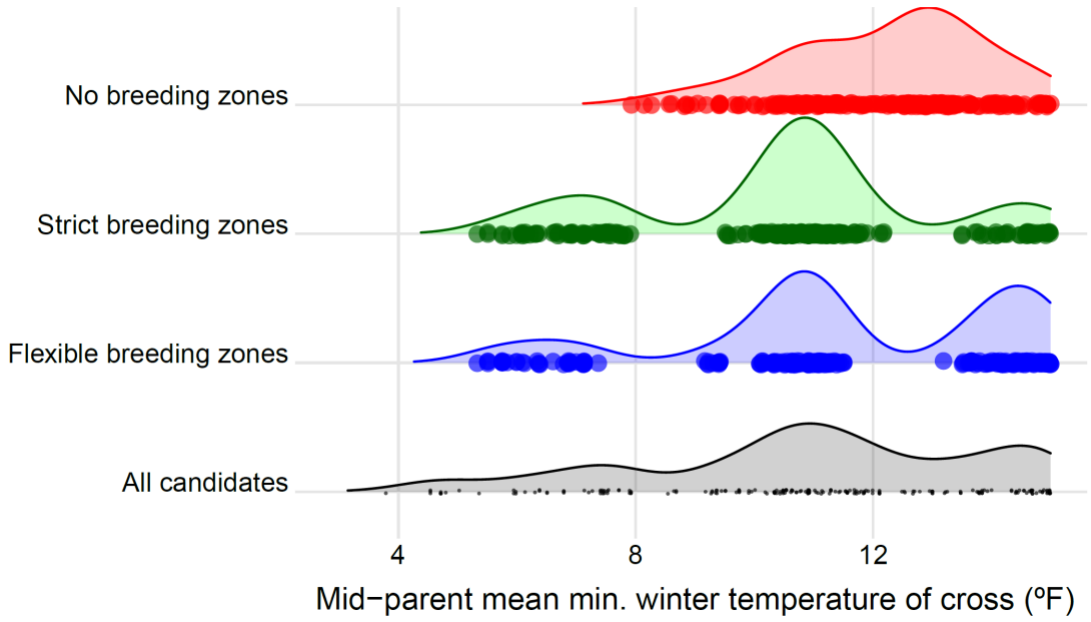

Figure 4. Distribution of progeny/cross (mid-parent) mean minimum winter temperature values for three mating scenarios. The distribution of all candidates for mating is shown on the bottom in black.

Short-term gain

Improvement in mean progeny index breeding value was highest for the scenario without breeding zones and lowest for the strict breeding zones scenario (Figure 5). The flexible breeding zones scenario was intermediate. All of the scenarios made considerable gain over the population from which the parents were selected.

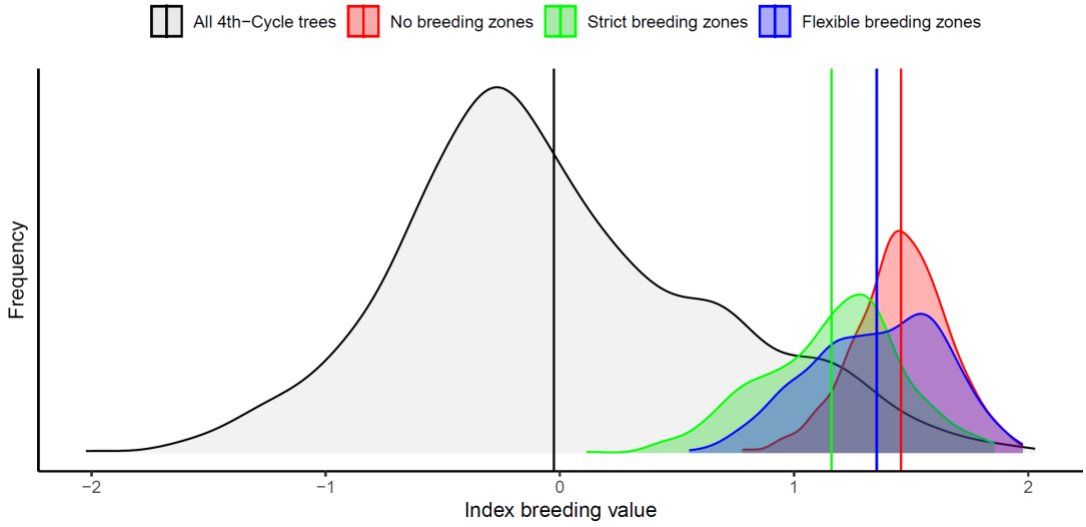

Figure 5. Distribution of progeny/cross (mid-parent) index breeding values for three mating scenarios. The distribution the 4th-Cycle population is shown in gray. Index breeding value units are on the standard normal scale.

Diversity

For all of the diversity metrics considered, the flexible breeding zones scenario was intermediate between the no breeding zones scenario (least diverse) and the strict breeding zones scenario (most diverse). For the number of parents used and the number of founders represented, the flexible zones tended to be more similar to the strict breeding zones scenario. For the status number and mean relationship among progeny, the flexible zones scenario tended to be more similar to the no breeding zones scenario.

Table 1. Diversity metrics for three mating scenarios.

Scenario	Number of parents used	Number of founders represented	Mean coancestry among progeny	Status number
No breeding zones	126	171	0.01469	31.9
Strict breeding zones	236	290	0.00656	66.3
Flexible breeding zones	191	253	0.01250	37.1

DISCUSSION

Ignoring cold hardiness in this population of *Pinus taeda* will result in breeding plans that choose warmer-source parents and, ultimately, the loss of adaptability for the colder part of the deployment range. The use of breeding zones is an effective tool to maintain cold hardiness, but it has a cost of short-term gain and increased breeding load (more parents are required to participate in breeding). The work we present here sheds light on quantifying these costs.

The flexible breeding zones approach produced a solution that met the cold hardiness needs and appeared to strike a good balance between short-term gain and diversity needs. However, this work did not elaborate on the choice of the weighting parameter on the target number of crosses per breeding zone. The value used (0.70) was "backed into to" by running several preliminary analyses. Breeders that use the flexible zones approach need to carefully re-evaluate this weight to ensure the mating list solution is adequate. Another limitation of our analysis was that it included just one cycle of breeding. Future research should evaluate the scenarios over multiple breeding cycles to better understand their impact. Further, we only considered the cold hardiness, short-term gain, and diversity impacts on the breeding population. Future research should also evaluate the

impacts on gain and cold-hardiness on the best set of unrelated parents, as the deployment population will consist of a much smaller set of parents from the breeding population.

There has been considerable research to develop methods for defining appropriate breeding zone boundaries (Raymond et al. 1990; Johnson 1997; Thomson et al. 2010; Ukrainetz et al. 2018; Yu et al. 2022). Owing to the fact that most tree breeding programs are in their infancy, with few cycles of breeding completed, there is little research on improving the use of breeding zones in advanced cycles. More research is needed to generalize the breeding zone strategy when adaptability traits are clinal.

ACKNOWLEDGEMENTS

These analyses were possible because of the establishment, management, and measurement of tests by members of the Cooperative Tree Improvement Program at North Carolina State University. We are grateful for their many contributions and their financial support. Funding was also provided by the Department of Forestry and Environmental Resources in the College of Natural Resources at North Carolina State University and by USDA National Institute of Food and Agriculture McIntire-Stennis Project NCZ04149.

REFERENCES

De La Torre, A. R., B. Wilhite, D. Puiu, J. B. St. Clair, M. W. Crepeau, S. L. Salzberg, C. H. Langley, B. Allen, and D. B. Neale. 2021. Dissecting the Polygenic Basis of Cold Adaptation Using Genome-Wide Association of Traits and Environmental Data in Douglas-fir. *Genes*. 12(1):110.

Farjat, A. E., A. K. Chamblee, F. Isik, R. W. Whetten, and S. E. McKeand. 2017. Variation among loblolly pine seed sources across diverse environments in the southeastern United States. *For. Sci.* 63(1):39–48.

Howe, G. T., S. N. Aitken, D. B. Neale, K. D. Jermstad, N. C. Wheeler, and T. H. Chen. 2003. From genotype to phenotype: unraveling the complexities of cold adaptation in forest trees. *Can. J. Bot.* 81(12):1247–1266.

Isik, F., and S. E. McKeand. 2019. Fourth cycle breeding and testing strategy for Pinus taeda in the NC State University Cooperative Tree Improvement Program. *Tree Genet. Genomes*. 15(5):70.

Johnson, G. 1997. Site-to-site genetic correlations and their implications on breeding zone size and optimum number of progeny test sites for coastal Douglas-fir. *Silvae Genet.* 46(5):280–285.

Kinghorn, B., and S. Kinghorn. 2021. MateSel. Available online at: https://www.matesel.com/content/documentation/MateSelInstructions.pdf?v=20220117; last accessed March 15, 2022.

Kinghorn, B. P. 2011. An algorithm for efficient constrained mate selection. *Genet. Sel. Evol.* 43(1):4.

Lindgren, D., L. Gea, and P. Jefferson. 1996. Loss of genetic diversity monitored by status number. *Silvae Genet.* 45(1):52–58.

Lindgren, D., and T. Mullin. 1997. Balancing gain and relatedness in selection. *Silvae Genet.* 46(2):124–128.

McKeand, S., and F. Bridgwater. 1998. A strategy for the third breeding cycle of loblolly pine in the Southeastern US. *Silvae Genet.* 47(4):223–234.

Mullin, T. J. 2017. OPSEL 2.0: a computer program for optimal selection in tree breeding. *Arbetsrapport Från Skogforsk*. (954–2017).

Namkoong, G., H.-C. Kang, and J. S. Brouard. 1988. *Tree breeding: principles and strategies*. Springer-Verlag, New York, New York. 178 p.

Raymond, C. A., G. Namkoong, and others. 1990. Optimizing breeding zones: genetic flexibility or maximum value. *Silvae Genet.* 39(3–4):110–3.

Schmidtling, R. C. 2001. Southern pine seed sources. *Gen Tech Rep SRS-44 Asheville NC US Dep. Agric. For. Serv. South. Res. Stn. 25 P*.

Schmidtling, R. C. 1994. Use of provenance tests to predict response to climate change: loblolly pine and Norway spruce. *Tree Physiol.* 14(7-8–9):805–817.

Thomson, A. M., K. A. Crowe, and W. H. Parker. 2010. Optimal white spruce breeding zones for Ontario under current and future climates. *Can. J. For. Res.* 40(8):1576–1587.

Ukrainetz, N. K., A. D. Yanchuk, and S. D. Mansfield. 2018. Climatic drivers of genotype–environment interactions in lodgepole pine based on multi-environment trial data and a factor analytic model of additive covariance. *Can. J. For. Res.* 48(7):835–854.

Wisniewski, M., A. Nassuth, and R. Arora. 2018. Cold Hardiness in Trees: A Mini-Review. *Front. Plant Sci.* 9 Available online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.01394; last accessed October 25, 2023.

Yamashita, M., T. J. Mullin, and S. Safarina. 2018. An efficient second-order cone programming approach for optimal selection in tree breeding. *Optim. Lett.* 12(7):1683–1697.

Yu, Y., S. N. Aitken, L. H. Rieseberg, and T. Wang. 2022. Using landscape genomics to delineate seed and breeding zones for lodgepole pine. *New Phytol.* 235(4):1653–1664.