

Breeding Theory and Genetic Testing

T L White, University of Florida, Gainesville, FL, USA

© 2004, Elsevier Ltd. All Rights Reserved.

Introduction

Tree breeding programs, also called tree improvement programs, create genetically improved varieties for reforestation and afforestation. As with breeding programs of agricultural crops and animals, these programs aim to change allele frequencies for a few key traits of a given species through repeated cycles of activities such as selection, breeding, and genetic testing. Today, tree improvement is an integral component of most plantation programs in the world with the ultimate goal being to increase the economic and social value of the planted forest. Unlike annual crops, trees are long-lived and difficult to work with due to their size. This means that, while the concepts of tree breeding and crop breeding are similar, the details differ greatly.

Owing to differences in species' biology, silviculture, product goals and economic considerations, tree improvement programs for distinct species differ dramatically in both design and intensity. Nevertheless, there are basic concepts and activities common to most programs. The aim here is to describe these common activities and the underlying theory on which they are based.

Principles of Recurrent Selection

Recurrent selection means repeated cycles of selection and breeding aimed at gradual genetic improvement of a few key traits in a single species. The benefits of breeding are cumulative in that each cycle or generation of improvement builds upon advances made in prior generations. Genetic gain in selected traits results from changes in frequencies of alleles at loci controlling expression of those traits, with favorable alleles increasing in frequency. Since most commercially important traits are polygenic (i.e., controlled by many loci), gene frequencies change slowly at any single locus and these changes are generally unknown to breeders. Rather, progress is measured by mean performance for target traits (e.g., greater harvest yield, reduction in disease incidence, or increased wood density). Performance for non-target traits should change little, if at all, as long as those traits are controlled by different loci than the target traits.

With the rare exception of programs involving radiation or chemical mutagenesis, tree breeding

programs do not create new genetic variation; rather, they utilize naturally occurring variation in starting populations of the species of interest. Then, through recurrent selection and breeding, the existing variation is repackaged into individuals containing higher frequencies of favorable alleles for target traits. Some crop breeding programs began over 10 000 years ago and today's domesticated varieties do not even resemble their original progenitors, thus demonstrating the power of recurrent selection. Tree breeding programs are much less advanced. In fact, for most forest tree species, there have been three or fewer generations or cycles of recurrent selection and breeding. This means that today's genetically improved varieties of forest trees are essentially the native species, and the only changes might be faster growth rate, better disease resistance, straighter stems, higher wood density, or other small quantitative changes in a few key traits.

The oldest form of recurrent selection is called simple recurrent selection (SRS) and each cycle of breeding involves only two steps: (1) mass selection of individuals based solely on their outward, phenotypic appearance; and (2) intermating these selections to produce the offspring for the next generation. This was the method used over 10 000 years ago by ancient farmers to improve their field crops; seed from selected plants was retained for next year's crop. SRS is rarely used in tree breeding programs today, because it is less efficient at achieving genetic gains than forms of recurrent selection that incorporate genetic testing and pedigree control.

Almost all tree breeding programs employ recurrent selection for general combining ability (RS-GCA) in which genetic testing follows selection. This entails planting identified offspring from all the selections in randomized, replicated experiments. Selections with high GCA values for any trait are those that produce top-performing offspring. After genetic testing, selections with low GCA values are discarded, while those with high GCAs (or offspring of high GCA parents) are included in future cycles of breeding. Genetic testing greatly increases the genetic gain above that from mass selection and is especially effective for traits with low heritabilities (as for most traits of trees).

Nearly all tree breeding programs rely on RS-GCA involving recurrent cycles of selection, breeding, and genetic testing. In these breeding programs, only additive effects of alleles (whereby offspring tend to be intermediate to their parents) accumulate in the breeding population, meaning that genetic effects due to dominance and epistasis do not contribute to the cumulative genetic progress in the breeding population. In the early cycles (say the first 10), genetic variability for the selected traits changes little

from that in the initial founding population, and genetic gain is similar in each successive cycle given similar selection intensities and breeding practices. That is, it takes many cycles of recurrent selection to lead to a plateau after which selection is ineffective (through fixation of favorable alleles or other causes). Beginning with larger founding populations and infusing unrelated material into the population extend the number of cycles before a selection plateau is reached.

Activities and Populations of a Typical Breeding Cycle

The activities and population types of tree improvement programs are summarized using a model called the breeding cycle (Figure 1). The activities of selection (to choose the selected population of genotypes) and breeding (to form the offspring that comprise the next generation's base population) are core activities that must occur each cycle or generation of breeding. The other activities and population types may or may not occur. Programs differ widely in how these activities are implemented, in the size and occurrence of the various types of populations, and in program intensity. Further, the cycles of breeding may not be discrete but rather overlapping with several staggered sets of activities all occurring simultaneously. Here, to stress the concepts, we explain each

population type as if it were physically distinct and occurred each cycle of improvement.

Base Populations

The base population of a given cycle of improvement consists of all available candidate trees that could be selected. The base population is very large consisting of many thousands of genetically distinct individuals. At the beginning of a program, the base population consists of all trees available for selection growing in natural stands and possibly plantations within the defined breeding zone (i.e., the geographical area for which an improved variety is being developed).

Determining breeding zone boundaries is a critical decision in tree improvement programs, because each breeding zone has a separate improvement program with its own distinct base, selected, breeding and propagation populations. For example, the program for Douglas-fir (*Pseudotsuga menziesii*) in the north-western USA defined 80 first-cycle breeding zones each with an average size of 40 000 ha of native forests. Thus, there were 80 distinct first-generation improvement programs, each with its own breeding cycle, different set of population types, and improved variety being developed. In contrast, the program for *Pinus elliotii* (slash pine) in the southeastern USA defined one breeding zone consisting of the entire natural range of the species (approximately 4 000 000 ha of timberlands). The differences in size and number of breeding zones between these two programs reflect both the more homogeneous climate in the south-eastern USA and differences in breeding philosophy.

Advanced-generation base populations (i.e., after the first complete breeding cycle) consist of genetically improved trees formed by intermating members of the breeding population and planting their offspring in genetic test plantations. All trees in these genetic tests are available to be chosen for advanced-generation selected populations.

Selection and Breeding

Selection and breeding are applied sequentially during each cycle of improvement (Figure 1). For most tree improvement programs the selected population in any given cycle contains between 100 and 1000 selected trees for a single breeding zone. In the first cycle of improvement, trees are selected from natural stands and plantations based solely on their superior phenotypic appearance, and this is called mass selection (Figure 2a). Advanced-generation selections are made from pedigreed populations growing in genetic tests, and selection effectiveness is increased by using all information available about a candidate's progeny, relatives, and ancestors (Figure 2b).

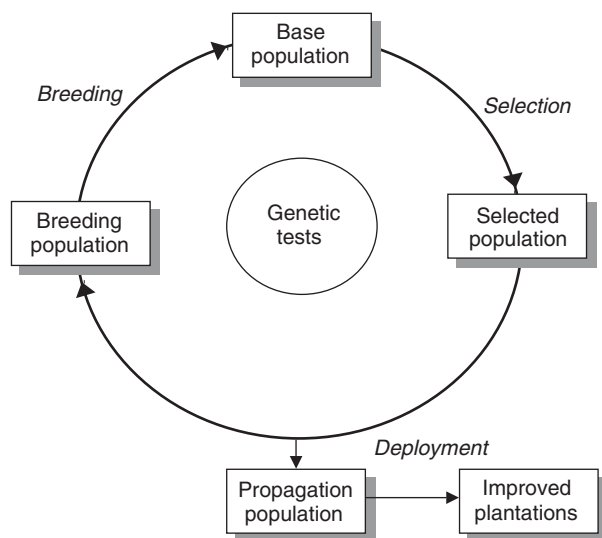


Figure 1 Schematic diagram of the breeding cycle of tree improvement programs showing the population types of genotypes that are formed (inside boxes) and the activities that are conducted (in italics) in a single cycle of improvement. Selection and breeding must be conducted each cycle, while the other activities may or may not occur. Starting at the top, the cycle turns a single revolution for each cycle or generation of selection and breeding in a recurrent improvement program. Adapted with permission from White TL (1987) *New Forests* 4: 325–342.



Figure 2 Forming the selected population means choosing superior individuals from the base population. (a) First-generation selection of *Pinus taeda* growing in the base population consisting of natural forests in Arkansas, USA. (Photograph courtesy of E.J. Jokela.) (b) Advanced-generation selection of *Eucalyptus grandis* growing in a base population consisting of pedigreed, randomized, replicated genetic tests of Cartón de Colombia.

Genetic gain in a given trait is achieved only if the selected population has a higher frequency of favorable alleles than the base population from which the selections were made. Genetic gain is greater if the selection is very intensive (only the very best individuals are selected) and if the trait is under strong genetic control (i.e., has a high heritability) with appreciable genetic variation. Allele frequencies differ between the base and selected populations both by intent and by chance. The breeder selects superior individuals and therefore intentionally alters allele frequencies for the target traits. In addition, allele frequencies for all traits (not just those included in the selection criteria) may change by chance due to sampling (i.e., choosing a subset of trees from a larger population). Some very rare alleles present in the base population can be absent from the selected population; however, with hundreds of individuals in the selected population, allele loss or large random changes in allele frequencies are unlikely.

After selection, some or all of the selections are included in that cycle's breeding population and are intermated to regenerate genetic variability through recombination of alleles during sexual reproduction. Intermating to produce full-sib families involves controlled pollination among selections. Female flowers on some selections are bagged (and emasculated if needed) to prevent contaminant pollination, and then pollen from other parents is injected into the bags (Figure 3). When two superior parents are mated, not all their offspring are superior, because some offspring

in a family receive more favorable alleles from their parents than others. Intermating results in a large amount of genetic variation both among and within the families planted in the genetic tests that form the new base population. Selection of superior trees from among these newly created progeny is the basis for making continuing genetic progress from recurrent cycles of selection and intermating.

Deployment of Genetically Improved Varieties

The ultimate goal of all tree improvement programs is to create improved varieties of trees to use for reforestation and afforestation of new plantations. This represents a tangible economic and social gain: higher-yielding, healthier planted forests. In each cycle of improvement, the propagation population (also called the production population or deployment population) is formed from some or all of the members of the selected population. Usually only a small subset of the very best selections is chosen to produce a sufficient quantity of genetically improved plants to meet the annual needs of the operational forestation program. The trees used for reforestation are a genetically improved variety (also called a breed), and the activity of mass propagation and planting of an improved variety is deployment.

The propagation population is distinguished from the central core of the breeding cycle in Figure 1, because core activities focus on maintaining a broad genetic base and achieving genetic gains over many generations of improvement. In contrast, the propagation population produces a commercial variety to maximize genetic gain in operational plantations in the short term. Seed orchards, a common type of propagation population, are often formed by grafting the very best members of the selected population into a single location that is managed intensively for the production of seed (Figure 4). The genetically improved seed from open pollination among the grafted trees is then grown by the nursery manager and the seedlings used for forestation.

There are other options for forming the propagation population. For example, clonal forestry can be achieved for some tree species either through rooted cuttings or somatic embryogenesis. With this option, the top performing clones are identified through genetic testing and comprise the propagation population. These 20 or so clones are mass propagated to create enough trees to meet the annual forestation demand. The clones deployed to operational plantations are the improved variety.

It is common to include only the very best selections in the propagation population; for example, the selected population might contain several hundred genotypes, while the propagation population

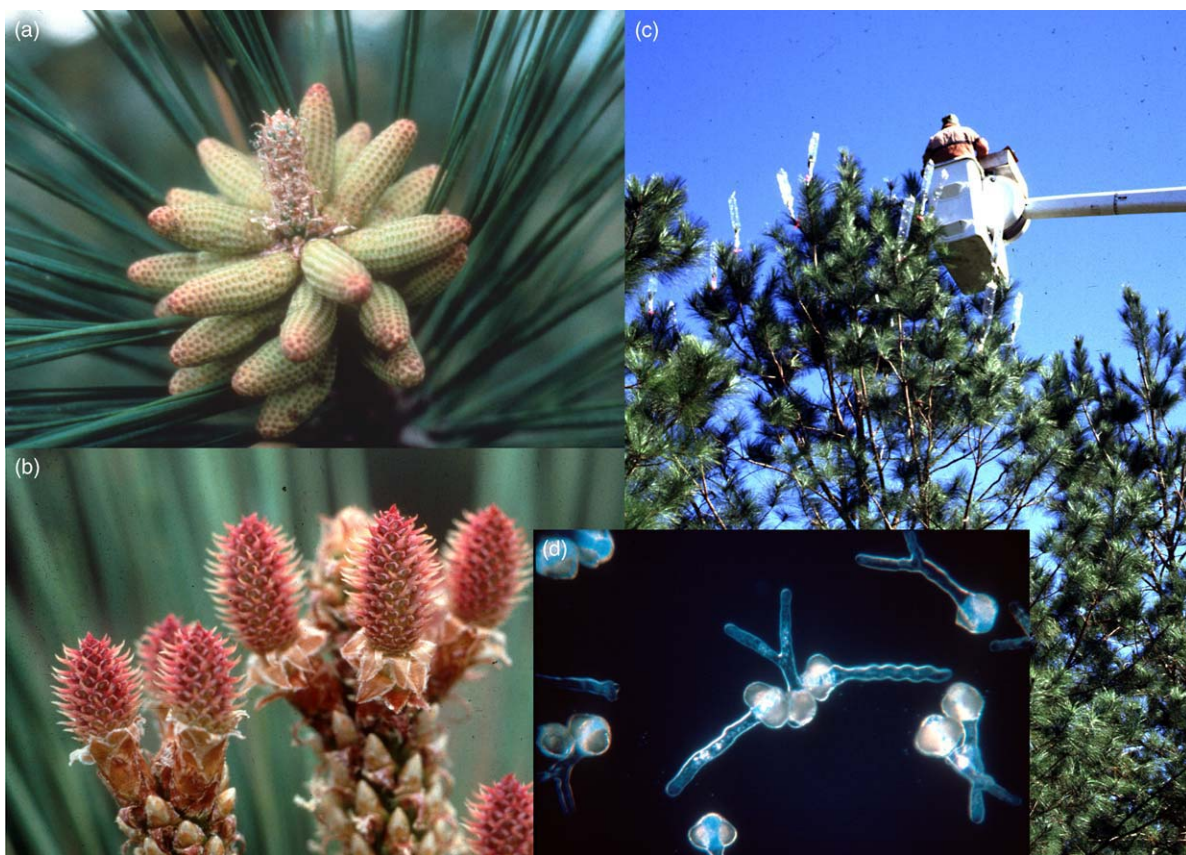


Figure 3 *Pinus taeda* is monoecious meaning that pollen catkins (a) and female strobili (b) are borne in different structures on the same tree. Controlled pollination of two parents for breeding takes 18 months and entails: (1) collecting catkins and drying them to extract pollen from the male parent; (2) bagging (c) the female strobili in the spring of year 1 to prevent wind pollination by other males; (3) injecting the pollen from the male parent into the bags covering the strobili of the female parent; (4) allowing pollen tubes (d) to begin growing in the first year and effect fertilization in spring of the second year; and (5) collecting full-sib seed in the fall of year 2, some 18 months after pollination.

might consist of the best 20 to 50. This increases the genetic gain expected from the operational variety that is planted, but also reduces its genetic diversity. It is also common to continually upgrade the genetic quality of the propagation population even during a single cycle of improvement. As information becomes available from genetic tests, genetically inferior selections can be removed from the propagation population, while superior selections not originally included can be added. This means that the improved variety being planted may change in its genetic composition and expected gain within a single cycle of improvement. This dynamic, rolling-front nature of improved varieties in forestry may be the reason why varieties are not given names as are cultivars in crops and breeds in animals.

Infusions from External Populations

Most tree improvement programs take advantage of opportunities to acquire new genetic material and infuse it into the breeding population. The infusions can be aimed at improving a specific trait. For

example, the *Pinus elliottii* program in the southeastern USA selected nearly 500 disease-free trees in stands highly infected with fusiform rust (*Cronartium quercuum* f. sp. *fusiforme*) to increase the frequency of resistance alleles in the breeding population. Another reason to include infusions is to broaden the genetic diversity existing in the program. For example, the *P. taeda* (loblolly pine) program in the southeastern USA made 3300 new selections to broaden the genetic diversity of several breeding populations spanning multiple breeding zones. In advanced-generation programs, it is often desirable to obtain infusions as proven selections from other programs working on the same species. Breeding programs for *Eucalyptus grandis* in many different countries sometimes exchange material for this purpose.

With infusions, the new material should be evaluated for target traits and adaptability through genetic tests to ensure that gains are not made in some traits with inadvertent losses in others. Tested infusions of sufficient merit are then intermated in the breeding population as part of the regular



Figure 4 Aerial view of a 12-year-old grafted seed orchard of *Pinus elliottii* showing the wide spacing and intensive management (mowing and weed control in the rows) aimed at maximizing seed production. The 50 or so superior selections grafted into this orchard are each represented by many grafted trees (called ramets) and together these ramets from the same selection form a clone (i.e., have the same genotype). Ramets from the same clone are separated from each other and the 50 superior clones are allowed to wind pollinate to produce seed for operational reforestation. This seed is a genetically improved variety with the amount of genetic gain depending directly on the genetic superiority of the clones grafted into the seed orchard.

breeding program. This maintains broad genetic diversity in the breeding population.

Advanced-Generation Tree Breeding Programs

Many tree breeding programs are entering the second or third cycle of breeding. These programs differ widely in breeding strategy, program design, and program intensity; yet, all programs face common challenges and sets of issues. In particular, all programs seek to achieve near-optimal short-term genetic gains in a few traits, while maintaining sufficient genetic diversity in the breeding population to ensure near-optimal long-term genetic gains, flexibility to changing conditions, and gene conservation. Common issues facing all programs addressing these goals are conveniently grouped for discussion below according to the population types of the breeding cycle.

Breeding Zones, Base Populations, and Selected Populations

Most advanced-generation tree breeding programs are opting for relatively large breeding zones encompassing sizeable portions of their plantation estate. Even if there is moderate genotype \times environment interaction present within a breeding zone, many programs are choosing to breed for broad adaptability by making selections that perform well across all edaphoclimatic conditions within the zone.

Minimizing the number of breeding zones has two distinct advantages: (1) the breeding and testing programs in each zone can be larger and more intensive, thereby achieving larger gains; and (2) costs are reduced, because there are fewer breeding populations and testing programs to manage.

Tree breeding programs with multiple breeding zones for the same species sometimes opt for overlapping zones that share selections among neighboring zones to reduce costs and increase gains. For example, the second-cycle *P. taeda* program in the southeastern USA defined multiple overlapping breeding zones, and each zone recruits new selections from zones on both sides. That is, excellent selections made in one zone are included in the selected populations of the neighboring zones. It is also possible to recruit material from other breeding programs of the same species in other regions or countries. Thus, advanced-generation breeding programs can capitalize on top genetic material from many sources provided the material is well adapted to the breeding zone.

With breeding zones defined, the next steps are to decide on how many selections to make (i.e., the size of the selected population), which traits to select for, and how to make the selections to achieve both genetic gain and diversity. Many advanced-generation tree breeding programs include 300 to 400 selections in the selected population for a given breeding zone in a single cycle. This number is sufficient to sustain near-maximum long-term gains over several cycles of recurrent breeding, even when there are several traits and when intensive efforts are used to achieve large gains in the first few generations. However, quite large populations (1000 or so) are required to ensure that almost all rare alleles are maintained in the population for many generations if the breeding population is also serving a gene-conservation role.

In all tree breeding, it is important to focus on very few traits. Inclusion of too many traits (say more than five) seriously dilutes the gains made on any one trait. Volume growth is almost always one of the important traits because it transcends product goals and technologies. Other important traits often include resistance to important edaphoclimatic stresses or fungal diseases. Finally, wood quality traits have become more important in recent years. When considering the candidate suite of traits, priority should go to those that:

- have high economic importance under a wide variety of future ownership, market, and technological scenarios
- are under moderate to strong genetic control

- are expressed at young tree ages and easily measured
- do not have unfavorable genetic correlations with other target traits.

The next step is to make selections and in advanced-generation programs these usually come from pedigreed genetic tests planted in randomized, replicated designs established on multiple sites in the breeding zone. Very often, data from the genetic tests are highly unbalanced, with not all families planted on all sites, varying test quality among sites and, sometimes, varying mating and field designs. The first step is to analyze the data properly using an analytical technique, called best linear unbiased prediction (BLUP), which incorporates all data from all sources to produce a unified set of predicted genetic values for each trait.

Next, a selection index is developed to weight each trait according to economic or other criteria, such that a single genetic worth is predicted for each pedigreed tree (which aggregates the BLUP predictions for each trait into a single weighted average for each tree). If there are say 50 000 pedigreed trees in genetic tests, and the breeder aims for a selected population with 300 individuals, the process is: (1) all 50 000 candidates are ranked according to their predicted genetic worth; and (2) the breeder begins at the top of list and selects winners subject to penalties for, or constraints on, relatedness. Normally, the penalty for relatedness means that only a certain number of selections can be made within a given family, and more are allowed from better families. Many programs include top performing selections from multiple generations, such as grandparents, parents, and offspring.

Use of a selection index often identifies candidates that are above average for all traits in the index but not outstanding for any single trait. In other words, the index-identified trees are most suitable for maximizing genetic worth of the aggregate index but do not maximize gain for each trait. Therefore, breeders usually include additional selections that are outstanding for each of the target traits.

Breeding Populations

After selection, the chosen trees are usually grafted into a convenient location for breeding (Figure 3). If all, say 300 to 400, selections are grafted and bred upon, then the breeding and selected populations are identical. Unlike most first-generation tree breeding programs that employed unstructured breeding populations in which all selections were bred in similar mating designs, advanced-generation programs structure the breeding population for at least three reasons.

First, many programs stratify the breeding population into two or more levels according to predicted genetic value with the top stratum being called the elite population. The elite population might contain the top 10% of the selections in the breeding population (e.g., the top 30 selections if the breeding population has 300 selections). The goal is to make more rapid and larger genetic gains in this small subset of the breeding population through accelerated breeding and more intensive genetic testing. The rest of the breeding population is allocated to the larger main population which is bred and tested less intensively to minimize costs, maintain genetic diversity, and achieve long-term genetic gains. As new selections are made in subsequent generations of breeding, they can be allocated to the elite, main, or both. This gene flow between the elite and main breeding populations increases genetic gain and diversity in the elite population.

Second, when the breeding population is subdivided for the purpose of applying different selection criteria to different segments, the subdivisions are called multiple populations or breeds. For example, the breeding population might be divided into three breeds with one being bred mainly for disease resistance for deployment to high-hazard sites, one as a multipurpose breed for general use, and the third for solidwood, high-value products. There might be some traits in common to all three breeds (such as rapid growth), but others that are only important in one breed or another (such as lumber quality in the solidwood population). The goal is to make more rapid gains for each of the breeds by minimizing the number of traits being improved. There can be few or many breeds, and there is no specific control on relatedness among them. Top selections and their relatives that possess many desirable traits may occur in more than one of the breeds.

The third and final type of structure to the breeding population aims to control the pattern of relatedness in the breeding population and produce noninbred offspring for operational deployment. Use of a single, unified breeding population eventually builds up relatedness among members of the breeding population and, if related parents are crossed to produce offspring for operational plantations, the plantations suffer from inbreeding depression. To avoid this problem, the breeding population is divided into sublines (sometimes called breeding groups). All breeding is conducted among selections in the same subline, and each subsequent generation's selections are assigned to the same subline as their progenitors. There is no breeding or relatedness among groups. All sublines are bred for the same objectives using the same selection criteria. Over generations of breeding,

relatedness builds up within each subline; however, members from different sublines are always unrelated, even after many generations of breeding. Therefore, any mating between members from different sublines always results in outcrossed progeny, precluding inbreeding depression. So, by choosing selections from different sublines to form the propagation population, outcrossed offspring are guaranteed for operational plantations.

In real tree breeding programs all three types of structure described above are used simultaneously. In addition, mating designs and field designs vary. This makes for a wide variety of breeding strategies, population structures, and program intensities in advanced generations.

Propagation Populations and Deployment

Most advanced-generation programs still employ open-pollinated (OP) seed orchards, as previously described, in which 15 to 40 top, unrelated selections are grafted into a single location to interpollinate and produce improved seed for operational forestation. Seed is often collected and deployed by OP family; this is called family forestry. For example, if there are 20 to 50 ramets of each of 30 clones in the orchard, then seed from all ramets of each clone is bulked together to produce 30 seedlots, one for each clone. Each seedlot is kept separate for storage, nursery production, and plantation establishment. This approach reduces genetic diversity only slightly, and allows families to be deployed to sites to which they are most suited based on genetic testing information of the maternal parent (e.g., disease-resistant families to high-hazard sites).

Family forestry is sometimes based on full-sib families created by controlled pollination (CP) (Figure 3). Top performing parents, based on performance of their offspring in genetic tests, are crossed together. If the CP process is efficient enough, the CP seed can be delivered directly to the nursery for growing seedlings for operational forestation. More commonly, the CP seed is too expensive for this option, and the seed is used to create hedges or stool beds in the nursery. Shoot tips from such plants are then rooted in either greenhouses or outdoor nurseries, and the resulting rooted cuttings from top full-sib families are planted operationally. This technique is being widely used for radiata pine (*Pinus radiata*) in Australia, New Zealand, and Chile.

In addition to family forestry, some intensive tree improvement programs are employing clonal forestry in which 10 to 30 well-tested clones are planted operationally. Clonal forestry means that a single genotype may be planted across many hectares (Figure 5). This achieves maximum genetic gain if



Figure 5 Nine-year-old clonal plantation of *Eucalyptus grandis* belonging to Mondi Forests in South Africa. All trees are a single genotype, a clone, and hence there is no genetic diversity within this plantation. Genetic diversity on a landscape scale is achieved by planting several clones in a patchwork in any one year and limiting the total area in which a given clone can be planted commercially.

top performing clones are planted, but eliminates genetic diversity within a single plantation. Several clones are often planted in patches (also called monoclonal mosaics) across the landscape, and most organizations limit the area of land that can be planted with a single clone. Clonal forestry is completely operational in intensive programs planting various species of *Eucalyptus*, *Salix*, and *Populus* that can be propagated easily as rooted cuttings. Conifers have proven more challenging since it is not possible, in general, to produce rooted cuttings from selection-age trees. Tissue culture techniques hold promise for achieving clonal forestry in conifers.

Genetic Testing

Genetic tests consist of pedigree, labeled offspring, or clonal propagules (ramets) usually established in randomized, replicated experiments in field locations on forest sites. Genetic tests are fundamental to all tree improvement programs and a single series of tests can serve any or all of the following objectives:

1. Evaluate relative genetic quality of selections to allow better selections to be favored in breeding and deployment.
2. Estimate genetic parameters such as heritabilities, genetic correlations, and genotype \times environment interactions to facilitate programmatic decisions and development of breeding strategies.
3. Plant offspring from breeding to create a base population of new genotypes from which to make the next cycle of selections.
4. Quantify or demonstrate genetic gains made by the program.

Mating designs and field designs, specifying how the genetic material is created and arranged in the experiment, are discussed below for this range of objectives.

Mating Designs for Genetic Tests

Incomplete-pedigree mating designs When many parents are allowed to intermate with no control on pedigree and a single bulk lot of seed is collected, then there is no knowledge of maternal or paternal identity. Bulk collections are most useful for comparing means of different populations such as in realized gains trials (objective 4 mentioned above). For example, consider 200 selections in a selected population with the top 20 being used in the propagation population to produce seed for operational plantations. Two genetic entries (a bulk seed collection from the 200 and one from the 20) could be planted in randomized, replicated designs across several test locations to compare genetic gain differences between the two populations. Unpedigreed, bulk collections are not used for any of the other three objectives of genetic tests.

In OP mating designs, seed is collected from each of the parents in the population and kept separate by parent for planting as OP families. For each tree planted, the female parent is known but the male parent is unknown. If there are 200 parents in the selected population, a test of all parents would involve planting 200 genetic entries or treatments (i.e., 200 OP families). OP mating designs are widely used for a variety of angiosperm and gymnosperm species, and are especially useful when many different male parents pollinate each family (i.e., when the OP families are highly outcrossed). When this occurs, OP families can be used very effectively for ranking selections and estimating most genetic parameters (objectives 1 and 2), and these designs are efficient and cost-effective.

Use of OP mating designs for creating a base population from which to make selections has been criticized on the basis that the male parentage is unknown meaning that two selections from different OP families could have the same male parent. Subsequent intermating of these selections in a propagation population (e.g., if both selections were grafted in the same seed orchard) could lead to inbreeding depression and hence reduced gain in the operational plantations. However, more recently, OP mating designs have been recommended for creating the base population based on theoretical gains calculations, empirical evidence of substantial genetic gains, and logistical ease coupled with low cost. Thus, OP mating designs have a role in some advanced-generation programs. For example, a

program could opt to breed the elite portion of the breeding population using full-sib, complete pedigree designs (see next section), but use OP mating for rapid, repeated cycling of the main portion of the breeding population.

In pollen mix (PM) or polymix mating designs, controlled pollination is used to pollinate each female parent with a mixture of pollen from a number of male parents (Figure 3). As with OP designs, there is one family for each parent being tested (i.e., 200 PM families for 200 parents in a selected population), and progeny trees from each PM family are labeled according to their female parent. PM designs are very cost-effective for estimating most genetic parameters and for ranking selections based on performance of their offspring but not for making fully pedigreed selections. Disadvantages of PM compared to OP designs are increased cost and time associated with controlled pollination.

Complete pedigree (full-sib family) designs There are various mating designs that employ full-sib (FS) families and hence maintain complete pedigree of all progeny planted. These designs share the following characteristics: (1) controlled pollination is used to create the seedling offspring; (2) identification of progeny trees in the field requires naming both parents; (3) the maximum number of unrelated forward selections is one-half the number of parents assuming parents are unrelated; and (4) each parent should be mated with four or more other parents to assure precise parental rankings.

Some FS designs can be prohibitive to implement owing to the large number of controlled crosses. For example, for a selected population of 200 parents, there are 19 900 possible pairs of matings ($200 \times 199/2$) in what is called a half-diallel without selfs. Even large tree improvement programs limit the number of total crosses for a given breeding zone to no more than several hundred FS families per cycle. The discussion below does not describe the variety of designs available, but rather focuses on one design that is feasible, cost-effective, and efficient for all four objectives of genetic tests listed above.

In a partial diallel mating design only some of the possible pairs of crosses are made, and there are many variants of partial diallels that make more or fewer crosses. More crosses per parent mean more precision for estimating genetic parameters (objective 2), for ranking the parents (objective 1), and for achieving more gain from making selections from the offspring planted (objective 3). Four or five crosses per parent seem optimal for most purposes, if all crosses are successful. For example, if there are 200 parents in the selected population and each parent is crossed

with four other parents, there would be a total of 400 full-sib families to establish in genetic tests.

In partial diallels and other mating designs, it is best to create the full-sib families such that there is a genetic linkage, called connectedness, among all parents. Even if two parents are never mated together directly, they can still be connected indirectly (e.g., crosses $A \times B$, $B \times C$, and $C \times D$ connect A with C and D). In this way, all crosses provide information about all parents. Disconnected designs with small groups of parents separated from others are less efficient for ranking parents.

Full-sib families are sometimes generated and planted over a period of years in what is called a rolling-front approach. As trees in the breeding population begin to flower, they are crossed with others that are flowering. Instead of waiting for all crosses to be completed and planted in a single series of genetic tests, the crosses are planted as seed becomes available. There are several series of genetic tests planted over a series of years. This creates unbalanced data, since each series contains only a partial set of the full-sib families; however, the data can be analyzed with BLUP if proper connectedness among the series is maintained.

Field Designs for Genetic Tests

Defining the field design of any series of genetic tests means specifying: (1) plot shape and number of trees per plot; (2) statistical design at each location including randomization scheme and number of replicates; (3) number and location of field sites; and (4) inclusion of other seedlings or clones as checklots, borders, and fillers. Specification of an optimal field design involves statistical, genetic, logistical, and economic issues. Most important are the objectives of the tests and, just as for mating designs, some field designs achieve certain objectives better than others. Here two extreme field designs are discussed, one appropriate for simultaneously achieving objectives 1 to 3 and the other suitable for estimating realized gains (objective 4).

Field design for breeding and base population genetic tests Often in advanced-generation breeding programs, a single series of full-sib tests is aimed at estimating genetic parameters (e.g., heritability and genotype \times environment interaction for key traits), ranking the parents in the breeding population and making the next-generation selections from the progeny planted (objectives 1 to 3). Nearly always, there are numerous families (50 to 400) to accommodate, and an appropriate field design is: (1) single-tree plots (STP) meaning that each full-sib family is

represented by a single progeny tree as its own plot within each complete replicate; (2) an incomplete block design in which each complete replicate is further subdivided into smaller units called incomplete blocks; (3) fifteen to 20 complete replicates at each location (meaning 15 to 20 offspring per family planted at each location); and (4) four to eight field sites (each with this same design) that span the range of edaphoclimatic conditions in the breeding zone. For a test series with 200 full-sib families, 15 replicates per site and five sites, there would be 15 000 progeny trees planted ($200 \times 15 = 3000$ per site) and each family would be represented by 75 trees ($15 \text{ replicates} \times 5 \text{ sites} \times 1 \text{ tree per replicate}$).

STPs assure higher statistical precision (i.e., better genetic parameter estimates and better rankings of parents, families, and progeny trees) than do experiments with row plots or rectangular plots that contain multiple trees per family. This is for two reasons: (1) STPs allow more replicates for a given effort, so each family samples more of the microsite variability on a site instead of being clumped together in multiple-tree plots on fewer microsites; and (2) STPs mean that the size of each replicate is smaller for a given number of entries, so replicates are more homogeneous. STPs facilitate testing large numbers of families or clones in many replicates.

Even with STPs, each replicate can occupy a large area. For example, with a typical planting density of 1200 trees per hectare, a test with 200 families in STPs means that each replicate is 0.17 ha ($200/1200$). In most places, replicates smaller than 0.1 ha are advisable. This is the reason to further subdivide each complete replicate into incomplete blocks. Each incomplete block might contain 15 to 30 trees planted in a more uniform microenvironment than families in the rest of the same complete replicate. The incomplete blocks add another blocking factor used to adjust entry means for microsite differences among incomplete blocks. This reduces the experimental error and increases the precision of the rankings of genetic entries by removing sources of environmental noise from comparisons of genetic entries. Use of incomplete block designs requires special computer programs to design the layout and analyze the data; however, these designs increase precision with no additional costs of test implementation, maintenance, or measurement.

There is a trade-off between the number of complete replicates per test site and number of sites in a test series. The total number of progeny needed per full-sib family across all sites ranges from 50 to 150 depending on several factors. So, with more replicates per site, fewer sites are needed. However, enough sites are needed to adequately sample the

edaphoclimatic zones within the breeding zone. For example, a less intensive breeding program with relatively uniform conditions across a small planting zone might opt to plant a test series on three different site locations with 20 complete replicates per site (60 total trees per family), while a more intensive program with a large breeding zone spanning several soil types and climates might opt to establish a test series on eight sites with 15 replicates per site (120 trees per family). These numbers of sites are inflated if the risk of catastrophic loss of entire test sites is appreciable.

Field design for quantifying realized gains Estimating progress from a tree breeding program involves field tests of material of distinct levels of genetic improvement that reflect different stages of program development over time. For example, a single test series might aim to estimate the genetic means of the following nine entries for a few key traits: (1) an unimproved bulk seedlot that represents the operational plantations established before tree breeding began; (2) two bulk seedlots collected from a first-generation seed orchard before and after roguing of inferior clones; (3) three single-family seedlots of families planted operationally; and (4) three operational clones. Tests might also compare first-, second- and third-generation commercial breeds to quantify progress.

Usually, the mating and field designs of realized gains tests (also called yield trials) simulate genetic and stand-level growing conditions of operational plantations so that the estimates of gain are appropriate for subsequent harvest scheduling and economic analysis (e.g., to justify the expenditures on the breeding program). For these reasons, common features of tests to estimate realized gain include:

1. The tested materials simulate operational varieties or breeds from the past, present or future.
2. Unimproved seedlots or unimproved clones are included to have a baseline for comparing newer varieties.
3. Rectangular plots of each genetic entry are used to approximate competitive conditions existing in operational plantations.
4. Tests are of long duration (half to full rotation length) to quantify gains in harvest yield and product quality.

Points 3 and 4 argue for using plots large enough to permit normal silvicultural activities (fertilization, weed control, and thinning) for an entire rotation period. Therefore, square or rectangular plots, each plot containing 25 to 100 trees of the same genetic

entry (e.g., same seedlot, family, or clone) are recommended for realized-gains tests. Rectangular plots are preferred because they simulate stand conditions in operational plantations, and so produce unbiased estimates of yield on a per unit area basis. Inter-tree competition intensifies as trees age, and genotypes that start slowly may be disadvantaged. Therefore, means from small plots (e.g., STPs described above) may be biased with results favoring genetic entries that are strong competitors, especially at early ages. Rectangular plots avoid these biases since inter-tree competition is among trees of the same genetic entry (same bulk mix, family, or clone).

The disadvantage of rectangular plots is that replicate sizes are large, so only a small number of different genetic entries can be included (normally fewer than 10 different entries in a given test series) and the number of replicates is limited on a given site. A test series is normally established on several site locations both to accumulate enough replicates to produce sufficient precision for comparing the genetic entries and to ensure that inferences about genetic gains and program progress truly apply to the entire program. Incomplete block designs are very useful to increase precision of comparing the genetic entries, but still the number of entries needs to be small.

Conclusions

Tree breeding programs develop genetically improved varieties for forestation through repeated cycles of activities such as selection, breeding, and genetic testing. Selection aims to increase the frequency of favorable alleles of a few important traits, while breeding re-creates and repackages genetic variation so that still better selections can be made in the next cycle of selection. Nearly all programs rely on a breeding strategy involving recurrent selection for general combining ability, which means that genetic tests are used during each cycle of breeding to rank the parents in the breeding population and increase gains from selection. Thus, designing, implementing, and analyzing genetic tests are critical aspects of tree breeding programs that must be done well to maximize genetic gain per unit time at minimum cost. Tree breeding programs vary widely in design, breeding strategy, and intensity owing to differences in species' biology, breeding objectives, and economic considerations. Nevertheless, tree breeding is an integral part of operational silviculture for most plantation forestry programs in the world. New technologies such as tissue culture to achieve operational planting of tested clones, marker-assisted breeding to increase gain from selection, and genetic modification to insert novel genes all have promise to

enhance gains from tree breeding, but must be tested for effectiveness, safety, and public acceptance.

See also: Genetics and Genetic Resources: Genecology and Adaptation of Forest Trees; Genetic Systems of Forest Trees; Propagation Technology for Forest Trees; Quantitative Genetic Principles. **Tree Breeding, Practices:** A Historical Overview of Forest Tree Improvement; Breeding for Disease and Insect Resistance; Economic Returns from Tree Breeding; Forest Genetics and Tree Breeding; Current and Future Signposts; *Pinus Radiata* Genetics.

Further Reading

- Ahuja MR and Libby WJ (1993) *Genetics and Biotechnology*, vol. 1, *Clonal Forestry*. New York: Springer-Verlag.
- Eldridge K, Davidson J, Harwood C, and Van Wyk G (1994) *Eucalypt Domestication and Breeding*. Oxford, UK: Clarendon Press.
- Lindgren D and Mullin TJ (1997) Balancing gain and relatedness in selection. *Silvae Genetica* 46: 124–129.
- Mandal AK and Gibson GI (1998) *Forest Genetics and Tree Breeding*. Daryaganj, New Delhi, India: CBS Publishing.
- Namkoong G, Kang HC, and Brouard JS (1988) *Tree Breeding: Principles and Strategies*. New York: Springer-Verlag.
- Williams ER, Matheson AC, and Harwood CE (2002) *Experimental Design and Analysis for Tree Improvement*. Collingwood, Australia: CSIRO.
- White TL (1987) A conceptual framework for tree improvement programs. *New Forests* 4: 325–342.
- White TL (1992) Advanced-generation breeding populations: size and structure. In *Proceedings of the IUFRO Resolving Tropical Forestry. Resource Concerns through Tree Improvement, Gene Conservation, and Domestication of New Species*, October 9–18, Cali, Colombia, pp. 208–222.
- White TL, Neale DB, and Adams WT (2003) *Forest Genetics*. Wallingford, UK: CAB International.
- Young A, Boshier D, and Boyle T (eds) (2000) *Forest Conservation Genetics*. Collingwood, Australia: CSIRO.
- Zobel BJ and Talbert J (1984) *Applied Forest Tree Improvement*. New York: John Wiley.

the long term, on scales varying from that of an individual enterprise to the entire extent of a species' use, in terms of each individual strategic and technical decision or of their cumulative effect, and in purely financial or in wider economic terms. The gains realized by particular stages or elements of tree breeding can be so dramatic that only the most cursory economic evaluation is necessary to substantiate them; conversely, strategic and technological options may be so complex, and realization of benefits so contingent on particular assumptions, that sophisticated economic analyses are necessary to inform investment decisions.

As for many forestry activities, economic analyses of tree breeding investments are variously complicated by long investment and rotation cycles, uncertainties about costs and benefits over these long time horizons, and by the challenges of accounting for nonmarket benefits and costs. However, there is both a long history of, and an increasing focus on, economic analyses of tree breeding investment decisions, which have contributed significantly to the design and development of tree breeding strategies and programs.

Economic returns from tree breeding are determined by species- and program-specific combinations of the following key parameters:

- the genetic characteristics of the population subject to breeding—reflecting inherent levels of genetic variation in a species, the extent to which that variation has been sampled in the population subject to breeding, and the stage of breeding of the population
- the breeding strategies and technologies employed, and the breeding objectives specified
- the value of the products and services, and the scale of deployment, of improved populations
- the institutional arrangements for breeding and benefit sharing.

Our discussion of the topic is structured around these parameters. In general, they are better characterized for longer-established, advanced industrial tree breeding programs, such as those for loblolly pine (*Pinus taeda*) in the southeast USA, radiata pine (*P. radiata*) in the southern hemisphere, or *Eucalyptus* species in continents other than North America. Industrial tree breeding programs are typically more advanced and data-rich, and have markets and benefit regimes that are generally better defined, than are programs for the breeding of trees for nonindustrial uses. Consequently, while the latter also have the demonstrated potential to make very significant economic contributions to people's

Economic Returns from Tree Breeding

P J Kanowski, Australian National University, Canberra, Australia

N M G Borralho, Forest and Paper Research Institute (RAIZ), Alcoentre, Portugal

© 2004, Elsevier Ltd. All Rights Reserved.

Introduction

The economic returns from tree breeding can be estimated over time-frames ranging from the short to