

Further Reading

- Burley J and Styles BT (eds) (1976) *Variation, Breeding and Conservation of Tropical Forest Trees*. London: Linnean Society.
- Frankel R and Galun E (1977) *Pollination Mechanism, Reproduction and Plant Breeding*. New York: Springer-Verlag.
- Hattermer HH and Melchior GH (1993) Genetics and its application to tropical forestry. In: Pancel L (ed.) *Tropical Forestry Handbook*, pp. 333–380. New York: Springer-Verlag.
- Hayward MD, Bosemark NO, and Romagosa I (eds) (1993) *Plant Breeding: principles and prospects*. London: Chapman & Hall.
- Koski V (1970) A study on pollen dispersal a mechanism of gene flow in conifers. *Communicationes Instituti Forestalis Fenniae* 70.4: 1–78.
- Koski V (1973) On self-pollination, genetic load, and subsequent inbreeding in some conifers. *Communicationes Instituti Forestalis Fenniae* 78.10: 1–42.
- Mandal AK and Gibson GL (eds) (1998) *Forest Genetics and Tree Breeding*. New Delhi: CBS Publishers.
- Owens JN (1993) Pollination biology. In: Bramlett DL, Askew GR, Blush TD, Bridgwater FE, and Jett JB (eds) *Advances in Pollen Management*, Agriculture Handbook no. 698, pp. 1–13. Washington, DC: US Department of Agriculture Forest Service.
- Owens JN and Blake MD (1985) *Forest Tree Seed Production: A Review on the Literature and Recommendations for Future Research*. Petawawa National Forest Institute, Canadian Forestry Service.
- Rehfeldt GE and Lester DT (1969) Specialization and flexibility in genetic systems of forest trees. *Silvae Genetica* 18(4): 118–123.
- Sarvas R (1962) Investigations on the flowering and seed crop of *pinus silvestris*. *Communicationes Instituti Forestalis Fenniae* 53: 1–198.
- Stebbins GL (1950) *Variation and Evolution in Plants*. New York: Columbia University Press.
- Stebbins GL (1974) *Flowering Plants: Evolution above the species level*. Cambridge, MA: Harvard University Press.
- Sybenga J (1992) *Cytogenetics in Plant Breeding*. New York: Springer-Verlag.

provenances) or individuals that can be used as seed parents or as clones for mass propagation. Where selection of individuals is involved the genetic improvement amounts to breeding, in which cumulative genetic gain is sought over successive generations. Efficient breeding is dependent on an understanding of the factors governing response to selection, in both the short term and the long term. Prediction of response can indicate what genetic gain is achievable and, when applied to alternative selection scenarios, it can be used to indicate how best to achieve the gain.

The classical tool for predicting response to selection, and optimizing various selection procedures, is quantitative genetics. This is based upon a model of individual gene action. It embodies the neo-Darwinian synthesis, which reconciles the usual pattern of continuous variation with Mendel's discovery that units of heredity represent discrete factors. Practical implementation is usually based on assuming that each trait is controlled by large numbers of genes ('polygenes'), each of very small effect, at widely dispersed sites or loci in the genome. This treatment will often be a major oversimplification of reality. Nevertheless, it is typically a powerful and remarkably robust framework for predicting response to selection and illustrating various guiding principles.

After a basic exposition of the genetic model, important topics are:

- how response to selection is governed by the different parameters, which need to be known at least reasonably well
- how the same principles can be extended to responses to progeny testing
- the factors governing the efficiency of various forms of indirect selection for breeding goals
- how the principles can be applied to multitrait selection
- how genotype $\times$ environment interaction can affect response to selection, appropriate structuring of a breeding program, and appropriate deployment of genetic material
- factors governing longer-term response to selection, as opposed to short-term, which need to be considered in structuring populations.

Complementary information needed by the breeder (but often very imperfectly known) includes the cost structures for the various breeding operations, and the economic worth functions for metric values of various traits. Such information allows efficient allocation of effort, and enables the breeder to decide on the appropriate emphasis to place on different traits.

The principles are far from specific to forest tree breeding, but certain features of forest tree breeding,

Quantitative Genetic Principles

R D Burdon, New Zealand Forest Research Institute Ltd, Rotorua, New Zealand

© 2004, Elsevier Ltd. All Rights Reserved.

Introduction

Response to selection is the basic prerequisite for successful genetic improvement. Within a species, the entities that are selected can be populations (or

notably the amount of physical resources entailed, have favored much explicit use of the quantitative genetic model.

The Basic Model

It is appropriate to state the basic role of genetic influences in governing the phenotype (P), the phenotype being the tree as one sees it and can measure it for any particular trait. We have the relationship

$$P = G + e = A + NA + e$$

where G denotes the effect of the genotype, and e denotes the effect of the environment. The latter is typically a local environmental effect, e.g., an effect of poor planting of the individual tree or a patch of ground with missing topsoil, as opposed to measurable effects of the general environment.

It can be seen that G has two components: (1) A , which results from additive gene effects, whereby offspring tend to be intermediate between their parents, and which are the effects that form the basis of cumulative genetic gain over generations, and (2) NA , which results from nonadditive gene effects, whereby offspring tend to depart from intermediacy between their parents, and which cannot be recaptured and accumulated over successive generations. Such effects can include dominance of expression of one allele over another at a single locus, or epistasis, whereby the effect of one allele at a locus can be conditional upon what allele(s) is/are represented at one or more other loci.

Note that $A/2$ represents the general combining ability (GCA) of the parent when it is crossed with the rest of the population to produce a half-sib family.

Complications that can arise, but are generally disregarded here, are epigenetic effects. Such effects can often masquerade as true genetic effects. They can include maternal effects which, in the case of seed-weight effects, are usually transient. However, clonally propagated material is subject to 'C-effects,' which typically reflect the state of the material at time of propagation and often can be erased only by sexual reproduction (seed production).

Key Genetic Parameters

Variances In terms of variances (which represent the squares of standard deviations, and for which usual notation is either V or σ^2), the basic model means that

$$\sigma_P^2 = \sigma_G^2 + \sigma_e^2 = \sigma_A^2 + \sigma_{NA}^2 + \sigma_e^2$$

where σ_P^2 is the phenotypic variance, and so on.

Heritability Heritability, the percentage of phenotypic variance that is heritable, is a key parameter, with bounds 0 and 1. It has two forms:

- narrow-sense, given by the ratio σ_A^2/σ_P^2 and usually denoted h^2 , which is applicable to propagation by seed
- broad-sense, given by the ratio σ_G^2/σ_P^2 and usually denoted H^2 ($\geq h^2$), which is applicable to mass propagation of selected clones.

A heritability is specific to a trait, and, in some measure, to the population and the macroenvironment (e.g., site). Narrow-sense heritabilities tend to be low (0.2 or less) for growth rate traits, but much higher (≥ 0.5) for wood properties.

Genetic correlation A genetic correlation (denoted r_A or r_G) between traits is a measure of the degree of common genetic control for the two traits concerned. Its bounds of 1 and -1 relate to complete common control to perfect inverse control respectively. Genetic correlations, if they are adverse, impose severe constraints on the genetic gain that is simultaneously achievable in the traits concerned. Common examples of adverse correlations in forest trees include wood density and diameter growth, or growth potential and hardness. On the other hand, favorable or even neutral genetic correlations can provide major opportunities for the breeder.

Some Simple Expectations

Assuming the usual quantitative genetic model of polygenic control we have the following expectations for composition of variances (ignoring certain minor components of nonadditive gene effects):

- among half-sib families (each parent in question mating with a random sample of a large population) $-\frac{1}{4}\sigma_A^2$; this amounts to GCA variance
- within such families $-\frac{3}{4}\sigma_A^2 + \sigma_{NA}^2 + \sigma_e^2$
- among full-sib families (crosses between random pairs of parents) $-\frac{1}{2}\sigma_A^2 + \frac{1}{4}\sigma_{NA}^2$
- within such families $-\frac{1}{2}\sigma_A^2 + \frac{3}{4}\sigma_{NA}^2 + \sigma_e^2$.

Note: It is very convenient for forest geneticists and tree breeders that, with pollination by wind, seed collections from individual trees often approximate closely to half-sib families.

Thus a heritability (or repeatability) of half-sib family means (h_{HS}^2), with n individuals per family, is given by:

$$h_{HS}^2 = \frac{1}{4}\sigma_A^2 / [\frac{1}{4}\sigma_A^2 + (\frac{3}{4}\sigma_A^2 + \sigma_{NA}^2 + \sigma_e^2)/n]$$

Intercrossing among parents can be done in a range of systematic mating designs, which almost all give

some forms of both half- and full-sib family information.

Applying the same approach to heritability of clonal means (H_C^2), we have the expectation

$$H_C^2 = \sigma_G^2 / (\sigma_G^2 + \sigma_e^2/n)$$

which will tend to be much higher for given n than h_{HS}^2 , illustrating how clonal material can give information with a much better signal-to-noise ratio than seedling families.

Also of interest is the variance attributable to interaction between parents, reflected in the component of family performance that cannot be predicted from the additive genetic merit or GCA values of the two parents; such variance is specific combining ability (SCA) variance, of composition $\leq 1/4\sigma_{NA}^2$.

Response to Selection

Direct Selection

For a single trait, expected response to selection ($E(R)$) is the product of selection differential (D) which is the difference between the mean for the selected individuals and that of the population from which they are selected, the heritability (h^2 , or H^2 for clonal selection), and the phenotypic standard deviation (S), such that:

$$E(R) = D \times h^2 \times S$$

If the distribution is normal, such that equal (or symmetrical) responses can be expected to equivalent selection for high or low values of the trait, D can be expressed in terms of number (i) of phenotypic standard deviations of the population in which selection is done. Thus

$$D = i \times \sigma_P$$

The formula for expected genetic gain ($E(R)$) can be expanded and manipulated into various forms, but the following features may be noted:

1. If h^2 is low, with σ_e^2 fixed and large, gain will vary in proportion to σ_A^2 .
2. If σ_A^2 (or σ_G^2) is fixed, but h^2 varies through variation in σ_e^2 , $E(R)$ will vary in proportion to h , and will therefore be reduced far less than in proportion to an associated drop in h^2 (σ_e^2 may be inflated by measurement error).
3. While i (number of phenotypic standard deviations, or selection intensity) always increases as the number screened for each one saved increases, it does so non-linearly, according to the law of diminishing returns (Table 1).

Table 1 Selection intensity (standardized selection differential, or i) in relation to number screened per individual saved, for global proportion and within finite subgroups

Trees screened per tree saved	Global proportion		Finite subgroups
	i	Marginal i per tree screened	i
2	0.798	0.399	0.564
5	1.400	0.201	1.163
10	1.755	0.071	1.539
100	2.665	0.010	2.542
1000	3.367	0.00078	3.241
10 000	3.958	0.000068	3.852
100 000	4.479	0.0000058	4.384
1 000 000	4.948	0.000000052	4.863
10 000 000	5.380	0.0000000048	5.301

4. If selection is done within finite subgroups (e.g., the best out of every two individuals instead of the top 50% of a very large population), i is less, especially when the subgroups are small (Table 1).
5. Percentage gain obtainable is dependent on the coefficient of variation (phenotypic standard deviation divided by population mean) as well as i and h^2 .
6. If, with cost constraints, there is a choice between screening more trees cheaply or fewer trees thoroughly, there is effectively a trade-off between i and h^2 .

Indirect Selection

It is often possible to select for a 'target' trait (y) that represents part or all of a breeding goal by using another ('index') trait (x) as a proxy, in what is termed indirect selection. Examples may include stem volume as y and stem diameter as x , or a mechanical property of wood as y and wood density as x , or else harvest-age performance as y and early performance as x , or even a DNA marker or an identified gene (x) that is strongly associated with desirable expression of a trait (y).

For indirect selection to be more efficient than direct selection there must be at least a fairly strong genetic correlation between the target and index traits. Other conditions, of which one or more needs to be met are:

- markedly higher heritability for the index trait than for the target trait, either as an inherent heritability or through more precise measurement
- much cheaper determination of the index trait, allowing more intensive selection than is possible with direct selection

- earlier expression of the index trait, allowing a shorter generation interval and thence more rapid gain per unit time.

The roles of these conditions are illustrated by the equation for the relative efficiency of indirect selection $E(x/y)$ compared with direct selection. In a single generation it is given by

$$E(x/y) = (i_x/i_y)(h_x/h_y)r_{G_{xy}}$$

where i_x and i_y denote the intensities of selection achieved for the respective traits, h_x^2 and h_y^2 are the corresponding heritabilities, and $r_{G_{xy}}$ is the genetic correlation between the two traits.

Multiplying the right-hand side by the ratio t_y/t_x , for the generation intervals under the two forms of selection, gives $E(x/y)$ in terms of gain per unit time.

Cases of indirect selection with forest trees include using wood density, which is usually very heritable yet not very expensive to determine, as the selection criterion where strength and/or side hardness may figure in the underlying breeding goal. In the case of *Pinus radiata*, selecting for closely spaced branch clusters, a highly heritable feature, has been used as a selection criterion to control branch size which is very subject to environmental influences. DNA markers will have the advantage of perfect heritability and can be determined very early in the life cycle, but may not be well correlated genetically with the trait of interest. Efficient early selection is very widely sought with forest trees.

Using Information on Relatives

Progeny Testing

The essence of progeny testing (and clonal testing) is to improve the effective heritability. Selection wholly on progeny test results, which often amounts to reselection of parents, can be termed backwards selection. For instance, with half-sib families, the heritability of family means (formulated above), will increase with the number of individuals per family. However, the number of parents, will be limited by the number of selections made from the preceding generation. This will often impose a practical constraint on i , even though trading off smaller family size and thence lower h_{HS}^2 for higher i may have a theoretical advantage.

Note that gain expected from selection on half-sib family information for seed orchards has a coefficient of 2 inserted in the adapted form of the equation for response to selection; this is to take account of the fact that both pollen and seed parents will be selected.

More General Use of Information on Relatives

Selection purely on the basis of progeny test performance is a special case of using information on relatives, in this case placing 100% reliance on information from offspring. It is in principle possible to use information from almost any class of relatives, including parents and siblings in selecting individual offspring. A simple but common case in forest tree breeding is selecting individual offspring on both sib-family and individual information, weighting individual information heavily if heritability is high and family information heavily if heritability is low, in a combined family-plus-individual selection index.

Information from multiple classes of relatives, with varying representation of relatives among candidates, can be used to estimate the genotypic merit of candidates by best linear unbiased prediction (BLUP).

Genotype $\times$ Environment Interaction

This phenomenon represents differential performance of genotypes with respect to each other in different environments. To accommodate it, the basic genetic model can be expanded to:

$$P = G + E + GE + e$$

where E is the effect of the macroenvironment (e.g., a site category), which can be allowed for, and GE is the genotype $\times$ environment interaction.

The interaction can have two components: that tending to cause changes in genotypic ranking among environments (which is usually of prime interest), and that reflecting differential expression of genetic difference among the environments. As above, G is composed of $A + NA$, and (in principle, at least) GE of $AE + NAE$.

Applying this extended model to variance components, and partitioning off macroenvironmental effects, we have

$$\sigma_P^2 = \sigma_G^2 + \sigma_{GE}^2 + \sigma_e^2 = \sigma_A^2 + \sigma_{NA}^2 + \sigma_{AE}^2 + \sigma_{NAE}^2 + \sigma_e^2$$

where $\sigma_G^2 + \sigma_{GE}^2$ denotes σ_G^2 and $\sigma_A^2 + \sigma_{AE}^2$ denotes σ_A^2 , within a single site.

Hence the heritability (narrow-sense) for selecting an individual for performing in its particular environment is given by

$$h^2 = \sigma_A^2 + \sigma_{AE}^2 / (\sigma_A^2 + \sigma_{AE}^2 + \sigma_e^2)$$

Deleting σ_{AE}^2 from the numerator gives the heritability for selecting the individuals for performing across the various environments. Substituting

$\sigma_G^2 + \sigma_{GE}^2$ for $\sigma_A^2 + \sigma_{AE}^2$ gives H^2 . This extended model can readily be applied to expected response to selection.

It may be noted that the ratio $\sigma_A^2/(\sigma_A^2 + \sigma_{AE}^2)$ is actually a genetic correlation between the performance from one environment to another, if corresponding variances are all expressed equally in all the environments. Such a correlation can be calculated either for a set of environments, or between a pair of environments. A favorable test environment is one that shows both high genetic correlations for tree performance with other important growing environments and a high heritability.

There are important implications for whether or not to institute regional subdivisions in a breeding program. Key points are:

1. A high ratio σ_{AE}^2/σ_A^2 favors regionalization, particularly if fairly discrete site categories can be identified.
2. However, even with quite substantial interaction, if progeny testing can be well spread across environments it may be possible to produce a set of selections that are near-optimal for all the individual environments.
3. That said, failure to screen in an environment that is strongly interactive with respect to the others can sacrifice much potential gain in that environment.

Genotype $\times$ environment interaction is the subject of a voluminous literature. Much of that, however, focuses on characterizing the interactive behavior of specific plant genotypes (which are often cultivar varieties). For forest tree breeding, the main interest often lies in the roles of environments in generating interaction.

Multitrait Selection

Selection for more than one trait usually dilutes the gains obtainable for individual traits (Table 2). However, if several traits are uncorrelated but of equal economic worth and equal heritability, it may be better to spread the selective effort among several traits, especially if large numbers of candidates can be easily screened for each of several traits. This is also illustrated in Table 2. In practice, traits will differ markedly in heritability, economic worth, and cost of evaluation per individual, which means that selection needs to be focused on just a very few key traits.

Genetic correlations between traits are crucial. Adverse correlations, if strong, severely restrict the gains that can be achieved simultaneously in the

Table 2 Selection intensities (standardized selection differentials, or i), for varying numbers of individuals screened for each one saved, assuming either a single trait or three uncorrelated traits for selected at equal culling rates. All cases relate to selection within single large population

Trees screened per tree saved	i		
	Single trait	Three traits	
		Per trait	Summed over traits
8	1.65	0.798	2.39
30	2.22	1.114	3.34
125	2.72	1.400	4.20
1000	3.37	1.755	5.26
1 000 000	4.95	3.367	10.10

traits concerned. In forest tree breeding such correlations may exist between growth potential and hardiness, or between wood density and stem volume production. Conversely, if intercorrelations are favorable, selecting for additional traits may dilute gains achievable in individual traits little if at all.

Four ways of selecting for multiple traits in any one generation are:

1. Independent culling levels, setting thresholds of acceptability for each trait.
2. Sequential culling (sometimes called tandem selection), first for one trait and then another, and so on. This may be cost-efficient if evaluation costs differ widely among traits, so one might, for example, select first for stem diameter and only then evaluate wood properties and cull for them.
3. Tandem selection which, in the true sense of the term, involves selecting for different traits in successive generations. For forest tree breeding, deliberate adoption of this method from the outset is unlikely to be realistic, although elements of it may be practiced as a fortuitous result of changed perceptions of breeding goals.
4. Multitrait index selection, attaching weights to the values for individual traits, like the partial regression coefficients in multiple regression. Thus an indifferent ranking for one trait may be offset by a candidate being outstanding for one or more other traits. Various bases exist for weighting. A theoretical optimum takes account of economic weights of the respective traits and variance-covariance matrices that encapsulate the heritabilities of all traits and all the phenotypic and genotypic intercorrelations among those traits. There can thus be elements of both direct and

indirect selection. In practice, this optimality depends on various assumptions, notably concerning cost structures and the quality of the information on heritabilities and all the intercorrelations between traits. Various modifications of the multitrait selection index are possible, including use of elements of independent culling levels, and setting restrictions on expected gains in one or more traits.

In fact, it is possible in principle to combine multitrait information with information from various classes of relatives, and unequal representation of relatives among the candidates, to obtain a multitrait BLUP solution.

Long-Term Gain

Assuming polygenic control, pair-crosses will continue to show genetic segregation, each containing roughly half the base population additive genetic variance. This segregational variation is expected to persist over generations, although it will decay over time, especially if populations are small and/or selection is intense, but it will tend to be replenished by mutation. It is the key to cumulative genetic gain over generations of crossing and selection, which lead to increases in frequencies of favorable genes (alleles) at the various loci. Some genes of quite large effect can be present without causing behavior that is obviously different from that expected with polygenic inheritance. However, the use of DNA technology to recognize such genes offers greater selection efficiency.

Initial response to selection will involve mainly the loci of intermediate gene frequencies, which contribute most to expressed genetic variation. If favorable alleles approach 100% frequencies at such loci, these loci will contribute little to continued genetic gain. However, initially rare favorable alleles at other loci may increase in frequency to the point where they become the prime basis of response to selection.

There are some situations where the quantitative genetic model may not suffice for the tree breeder. In breeding for disease resistance there may be genes of large effect, and some can mask the expression of resistance genes at other loci. These genes of large effect need to be recognized and managed carefully, in order to select efficiently and ensure durability of resistance against mutation and genetic shifts in the pathogens. DNA technology offers a general means of recognizing and capturing favorable genes of significant individual effects, thus going beyond the classical polygenic model.

Inbreeding

With normally outbreeding organisms, which include almost all forest trees, the quantitative genetic model may not be straightforwardly applicable when significant inbreeding occurs. With outcrossing, the genetic load, which mainly represents genes that are individually rare but very deleterious in the homozygous state, contributes almost nothing to the expressed genetic variation. Inbreeding, however, allows such genes to be expressed strongly, thus contributing a different element of expressed genetic variation.

See also: **Genetics and Genetic Resources:** Genecology and Adaptation of Forest Trees; Genetic Systems of Forest Trees; Population, Conservation and Ecological Genetics. **Tree Breeding, Principles:** A Historical Overview of Forest Tree Improvement; Breeding Theory and Genetic Testing; Conifer Breeding Principles and Processes; Economic Returns from Tree Breeding; Forest Genetics and Tree Breeding; Current and Future Signposts; Genetics and Improvement of Wood Properties; *Pinus Radiata* Genetics.

Further Reading

- Baker RJ (1986) *Selection Indices in Plant Breeding*. Boca Raton, FL: CRC Press.
- Bulmer MG (1985) *The Mathematical Theory of Quantitative Genetics*. Oxford, UK: Clarendon Press.
- Cotterill PP and Dean CA (1990) *Successful Tree Breeding with Index Selection*. East Melbourne, Victoria: CSIRO Publications.
- Crow JF (1986) *Basic Concepts in Population, Quantitative and Evolutionary Genetics*. New York: WH Freeman.
- Falconer DS and Mackay TFC (1996) *Introduction to Quantitative Genetics*, 4th edn. Harlow, UK: Longman.
- Fins L, Friedman ST, and Brotschol JV (1992) *Handbook of Forest Genetics*. Dordrecht, The Netherlands: Kluwer.
- Kang MS and Gauch HG (eds) (1996) *Genotype-by-Environment Interaction*. Boca Raton, FL: CRC Press.
- Lindgren D and Nilsson J-E (1985) *Calculations Concerning Selection Intensity*. Umeå, Sweden: Swedish University of Agricultural Sciences, Department of Forest Genetics and Plant Physiology.
- Mrode RA (1996) *Linear Models for the Prediction of Animal Breeding Values*. Wallingford, UK: CAB International.
- Namkoong G, Kang H-C, and Brouard JS (1988) *Tree Breeding: Principles and Strategies*. New York: Springer-Verlag.
- White TL, Neale DB, and Adams WT (2003) *Forest Genetics*. Wallingford, UK: CAB International.