# Marker Genetics in Broad-leaved Species

**MARI RUSANEN**
*Finnish Forest Research Institute*
*Box 18, 01301 Vantaa, Finland*
*mari.rusanen@metla.fi*

**Rusanen M.** 2000. Marker Genetics in Broad-leaved Species. *Baltic Forestry,* 6 (2): 62–66.

Marker genetics provides useful tools for both genetic research and practical tasks in tree breeding and seed legislation. The possibilities to apply markers in broad-leaved trees has increased recently when new molecular marker techniques have been developed. This paper illustrates the various applications by summarizing some recent studies where markers have been applied to broad-leaved trees.

**Key words:** marker genetics, allozymes, microsatellites, broad-leaved trees

## Introduction

Marker genetics is a useful tool to study biological processes and genetic structures in natural populations, but also a tool to answer practical questions concerning e.g. the origin of forest reproduction material or genetic variability in a seed lot. Traditionally, marker genetics of forest trees has progressed more rapidly with conifers; terpenes were one of the first markers applied to forest trees and later on the haploid endosperm of coniferous seed has provided forest geneticists with methods that could not be applied to deciduos trees. However, during the past ten years molecular markers of broad-leaved trees have developed rapidly, thus leading to both deeper knowledge of the biological processes in natural stands and to practical applications for tree breeding and genetic conservation.

Marker genetics is a broad subject and this paper will limit its scope to the use of markers to solve problems in population genetics, reviewing some of the recent progress in this field. Helpful introductions to the techniques of marker genetics in plant genetic resources are e.g. Karp et al. (1997) and Ayad et al. (1997); and more specifically for forest trees a compendium of the EU research project 'Development, optimization and validation of molecular tools for assessment of biodiversity in forest trees' (Gillet, 1999).

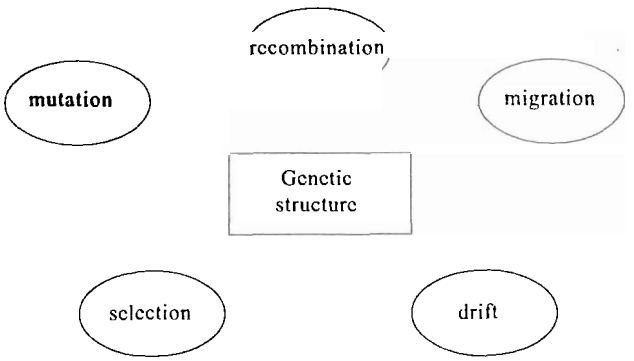
Some of the commonly used markers and their descriptors are summarized in Table 1. The choice of a marker depends first of all on the objective of the research and the target species. When studying within-population structures, markers with high variability and thus high resolution capacity are needed. On the other hand, highly variable markers may not be good for studies where the main interest lies at among-population level. The final decision is a combination of the economic limits one has in his or her project, the facilities and skill available. For example, one has to consider weather to use the already available markers or to make effort to develop new markers with higher resolution capacity.

**Table 1.** Some markers and their properties. Modified from Rafalski (1993), Kalendar et al. (1999), Ridout(1999) and Manninen (2000)

| | Isoenz | RFLP | RAPD | SSR | AFLP | REMAP |
|---|---|---|---|---|---|---|
| Principle | Enzyme morphs with Different Charges | Restriction Southern Blotting | DNA amplification with random primers | PCR of simple sequence repeats | Restriction Ligation of adapters Selective PCR | PCR of DNA between retrotransposons and SSR |
| Level of polymorphism | Low | High | Medium | Very high | Medium | High |
| Dominance | Codom. | Codom. | Dom. | Codom. | Dom. | Dom. |
| Squence Information required? | No | No | No | Yes | No | Yes |
| Development costs | Low | High | Low | High | Medium | Medium |
| Running cost /assay | Low | Medium | Low | Medium | Medium | Low |
| Repeatability | Very high | Very high | Fair | Very high | Very high | High |
| Ease of use | Easy | Labour intensive | Easy | Easy | Difficult initially | Easy |

Basically, most of the genetic research using markers measures the effect of one of the evolutionary forces or processes on a population and its genetic structure. The genetic structure is a result of these processes, the magnitude of these forces relative to each other and their interaction (Figure 1). With neutral mark-

**Figure 1.** The evolutionary forces that form the genetic structure of a population.




ers, which by definition are not affected by selection, one can study forces that influence the whole genome as such. These forces are migration, recombination and genetic drift.

### *Population genetic research*

**Allozymes** have been used widely to measure the genetic variation on species- and population level; excellent discussions on the method and reviews on the measurements of variation are eg. Hamrick et al. (1991) and Hamrick et al. (1992). Young, et al. (1993) addressed the specific question of forest fragmentation and its effect on the population structure of sugar maple (*Acer saccharum*) by comparing allozyme variation in eight fragmented forest patches to variation in eight population samples from an extensive continuous forest. Their hypotheses was that patch populations would have less genetic variation than continuous forest populations, owing to founder effects, increased drift and/or inbreeding, and also that the patch populations would be more genetically differentiated because of increased population isolation. The genetic data was collected from 50 1-year seedlings in each patch or sample, scoring 13 polymorphic loci.

Comparison of polymorphism, allelic diversity and individual heterozygosity between patch and control populations showed that forest fragmentation had not led to the expected loss of genetic variation at the population or individual level. There had been no apparent increase in inbreeding in patches. Similarly, based on genetic distances, fixation coefficients and allele frequency comparisons, the amount of differentiation among patch populations had not increased. However, forest fragmentation appeared to have affected the overall genetic variation in the forest system, with the eight patch populations having six fewer alleles than the eight control populations.

More recently the development of DNA-markers have made it possible to measure the within-stand structure and gene flow more accurately. **Microsatellites** are short, repetitive DNA sequences that consist of tandemly repeated short units. They are evenly distributed throughout the nuclear genome, highly variable and codominant with the exception of some null alleles. The high number of alleles provide a means to fingerprint individuals, using only a reasonable amount of loci and thus measure the mode of reproduction in a natural stand. So far only a few temperate broad-leaved species have been screened for microsatellite loci (Lefort et al. 1999). One of the first studies where microsatellites have been applied to natural broad-leaved stands is by Dow et Ashley (1996). They analyzed 64 adult trees and 100 saplings in a bur oak (*Quercus macrocarpa*) stand. As few as 4 loci, with the number of alleles per loci varying from 13 to 20, was sufficient to determine the parentage of the saplings by exclusion. The advantage of this method, compared to maximum likelihood or fractional paternity likelihood methods, is that you avoid making assumptions about population isolation or equal mating success of the adults. The result of this study is in accordance with the earlier indications of fairly effective migration, compared to selection and drift. Dow and Ashley detected high levels of long-distance pollination and no spatial clustering of pollen donors around maternal trees in the studied stand. The gene flow through acorns was notably lower. Of 100 saplings studied only six had both parents outside the stand and when corrected for cryptic gene flow this figure would mean 14 saplings having no parent in the stand. Cryptic gene flow is the possibility that a parent genotype that was interpreted to grow in the stand, is actually from outside; the correction factor is calculated based on allele frequencies.

In a more recent paper Dow and Ashley (1999) have used the same microsatellite markers to estimate if the genetic structure of the stand is changing. They divided the stand into spatial groups and compare the relatedness within the groups for the adults to relatedness for the saplings. The relatedness among the adults was always negative, while the relatedness among the saplings was substantially higher. These results suggested that while the adults do not represent clusters of offspring from a few seed parents, the saplings do to some extent. Their conclusion was that the genetic structure of the stand is changing.

Molecular markers on **non-nuclear DNA** (chloroplast- or mitochondrial-DNA) provide a tool to study variation on an evolutionary scale. Ferris et al. (1998) have studied evolutionary migration routes of *Quercus robur* and *Quercus petraea* using universal primers for chloroplast DNA. Cp- DNA is conservative and maternally inherited in oaks and thus well suited for studying intraspecific geographical structuring. Ferris et al. found four cytotypes: eastern, central, western, east-anglian. The three main types divide the species range into three longitudinal zones, extending from south to north, which supports the idea of three separate refugia during the glaciation.

During range expansion, gene flow to new areas is exclusively by seed, and therefore largely similar for the nuclear and cytoplasmic genomes. Following a contact of effectively continuous populations, the exchange of maternally inherited genes will, however, be effectively impeded. Even though long-term acorn transfer is not uncommon, the presence of local acorn pool will effectively swamp any cytoplasmic gene flow to an already occupied territory. As a consequence, the distribution of a cytoplasmic marker will be effectively frozen to a situation at the time of the contact. By contrast, pollen dispersal is less leptokurtic, and nuclear gene exchange between populations is likely to be much more common.

Ferris et al. (1998) also found that the cytotype distribution is independent of the conventional morphological species boundaries. This indicates efficient cytoplasmic gene exchange between these two species during or after the refugial isolation. This also means that cytoplasmic marker distribution does not make a reasonable basis for racial subdivisions, since it reflects poorly the history and composition of the nuclear genomes of the trees.

Ferris et al. (1998) detected two cytotypes present in the narrow distribution of oak in southern Finland, which indicated that oak has colonized Finland twice, independently. This finding was confirmed with a survey of 48 Finnish populations (166 trees) with a restriction digest assay that separates the eastern cytotype from the others. Results from allozyme study of the same populations prove that this difference between the two migration routes of oak to Finland can not be detected with allozymes (Vakkari, unpublished).

**A comparison between markers,** applying different types of markers on the same material has been done by Streiff et al. (1998). They used three kinds of markers to describe the spatial distribution of genotypes in a stand of oaks (*Q. robur* and *Q. petraea*) with 355 mature trees. The markers they compared were isozymes (4 loci), microsatellites interpreted as unordered alleles (6 loci) and microsatellites interpreted as alleles ordered according to the size (6 loci). As expected, the microsatellite loci were much more variable than the isozyme loci, having average 21.7 and 4.3 alleles/loci, respectively. Consequently, the observed heterozygosity levels ($H_o$) were 0.81 when averaged over microsatellite loci and 0.25 averaged over isozyme loci.

Streiff et all. estimated the genetic differentiation among 144 squares within the stand by calculating $F_{st}$ values for allozymes and unordered microsatellites and the respective $R_{st}$ values for ordered microsatellites. In general, they observed the same tendency of significant, but low, spatial genetic structure for different markers and statistics. $F_{st}$ values were higher for isozymes than for microsatellites, but the larger mean values were also associated with larger variance and eventually the differentiation analyzed with unordered microsatellites was always more significant compared to differentiation measured with isozymes. Ordering of the microsatellite alleles did not improve the resolution of spatial organization of diversity in this study.

We are currently acquiring more and more knowledge of the ways different markers can benefit the understanding of genetic structure and behavior of the forest populations. A challenge for the near future will be **to compare the variation in neutral markers to that of morphological or quantitative traits**. Few studies on this field have been published, one of the first is a comparison of population differentiation ($F_{st}$) on allozymes and on both continuous and discrete morphological traits on an annual herb *Clarkia dudleyana* (Podolsky and Holtsford, 1995). Their finding was that the distribution of genetic variation within and between populations differed for the three types of traits. Overall, the allozymes had the lowest Fst estimates, indicating relatively little population differentiation. The continuous morphological traits had estimates similar to those from allozymes and the discrete morphological traits tended to have the highest estimates. The pattern of differentiation for the discrete morphological traits strongly reflected the geographic distribution of the populations, whereas patterns for the continuous traits and allozymes did not. These results suggest that selection has been occurring on the discrete morphological traits, selecting toward a common optimum within each geographic group, and optimal differing among geographic groups (Podolsky and Holtsford, 1995). With broad-leaved trees, an attempt to compare variation in

neutral markers to variation in adaptive characters is currently made in a SNS-funded, project 'Importance of life-history traits for gene conservation', where studies on marker-genetics, phenotypic plasticity and quantitative variation are joint to study selected stands of insect-pollinated Norway maple and wind-pollinated silver birch.

### *Practical applications*

Marker genetics is currently providing us with better understanding of the genetic structure and breeding system of natural broad-leaved forests. Naturally, the scientific progress will in the long run benefit also practical forestry, especially forest regeneration and production of regeneration material. However, new applications of marker genetics will also serve directly tasks like seed certification and **control of the trade of forest reproductive material**. These applications may be used to quantify genetic variation, verify of geographic origins or estimate the level of inbreeding.

An advanced approach to further develop the practical use of molecular markers is made in Austria where they have tested methods of microsatellite data analysis that require no genotype information on the parental population samples (Lexer et al. 1999). The aim of the pilot study was to make it possible to use microsatellites for analysing genetic variation in small, anonymous commercial seed lots. In particular, they used microsatellites to monitor the number of different seed parents and the number of pollen donors in a seed lot. In Austria marketable seedlots of oak have to be collected from a minimum of 20 different trees per stand. Small, anonymous seedlot samples must be shipped to the Federal Forest Research Centre by the companies in a way that samples from different seed parents are supplied separately. Control of trade includes monitoring these samples to define genetic relationship within and among the seedlot samples; to detect possible seed contamination and to infer the number of different seed parents in a lot. In accordance with the pilot study by Lexer et al. (1999), microsatellites are a promising tool to study the genetic composition of small seed lot samples even if the genetic information of the parents is missing.

### *Future prospects*

In the future, the technical challenges will be to further develop the laboratory procedures so that more laboratories would have access to automated procedures with high repeatability from one laboratory to another. Especially the development of universal primers will promote application of the microsatellite markers to more species and wider geographic areas. Another challenge is to develop statistical analysis to improve the cost/benefit ratio of the marker applications. One interesting approach in this field is a comparison of dominant AFLP-markers and co-dominant microsatellites to be used for parentage analysis, done by Gerber et al. (1999). Better understanding of different types of variation will be gained, when there are more studies comparing variation in markers to that in morphological and quantitative characters. There is also a request for markers to study traits that are affected by selection and even markers that are capable to differentiate between stochastic processes (e.g. mating system) and selection-processes.

Marker genetics have greatly contributed to our understanding of genetic processes currently influencing the natural populations as well as the forces that have formed the genetic structures in an evolutionary time scale. In forest tree breeding, biochemical markers have for long served purposes like controlling grafts in a seed orchard and verifying crossings. For broad-leaved trees, the presently available molecular markers enhance the possibilities to utilize markers in practical tasks such as seed certification and in long-term breeding. Currently both molecular laboratory techniques and statistical data analysis are developing rapidly, thus providing good foundation for future reasearch.

## References

**Ayad W. G., Hodgkin T., Jaradat A. and Rao V. R.** 1997. Molecular genetic techniques for plant genetic resources. Series Molecular genetic techniques for plant genetic resources . International Plant Genetic Resources Institute, Rome. 137 p.

**Dow B. D. and Ashley M. V.** 1996. Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. Molecular Ecology, 5: 615-627.

**Dow B. D. and Ashley M. V.** 1999. Changing genetic structure of savanna bur oak population. Forest Genetics, 6: 29-39.

**Ferris C., King R. A., Väinölä R. and Hewitt G. M.** 1998. Chloroplast DNA recognizes three refugial sources of European oaks and suggests independent eastern and western immigrations to Finland. Heredity, 80: 584-593.

**Gerber S., Mariette S., Streiff S., Bodénès C. and Kremer A.** 1999. Comparison of microsatellites and AFLP markers for parentage analysis. In: E.M. Gillet (ed.). Which marker for which purpose? European Union DGXII. URL http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm.

**Gillet E. M.** (ed.). 1999. Which marker for which purpose? European Union DGXII. URL http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htmp.

**Hamrick J. L., Godt, M. J. W., Murawski, D. A. and Loveless, M. D.** 1991. Correlations between species traits and allozyme diversity: implications for conservation biology. In: D.A. Falk and K.E. Holsinger (eds.). Genetics and Conservation of Rare Plants. p. 75-86.

**Hamrick J. L., Godt M. J. W. and Sherman-Broyles S.L.** 1992. Factors influencing levels of genetic diversity in woody plant species. New Forests, 6: 95-124.

**Kalendar R., Grop T., Regina M., Suoniemi A. and Schulman A. H.** 1999. IRAP and REMAP: Two new retrotransposon-based DNA fingerprinting techniques. Theoretical and Applied Genetics, 98: 704-711.

**Karp A., Kresovich S., Bhat K. V., Ayad W. G. and Hodgkin, T.** 1997.Molecular tools in plant genetic resources conservation: a guide to technologies. Series Molecular tools in plant genetic resources conservation: a guide to technologies. International Plant Genetic Resources Institute, Rome. 47 p.

**Lefort F., Echt C., Streiff R. and Vendramin G. C.** 1999. Microsatellite sequences: a new generation of molecular markers for forest genetics. Forest Genetics, 6: 15-20.

**Lexer C., Heinze B., Steinkellner H., Kampfer S., Ziegenhagen B. and Glössl J.** 1999. Microsatellite analysis of maternal half-sib families of *Quercus robur*, pedunculate oak; detection of seed contaminations and inference of the seed parents from the offspring. Theoretical and Applied Genetics, 99: 185-191.

**Manninen O.** 2000. Genetic mapping of traits important in barley breeding. Department of Biosciences. Series Genetic mapping of traits important in barley breeding . University of Helsinki. Helsinki.

**Podolsky, R.H. and Holtsford T.P.** 1995. Population structure of morphological traits in *Clarcia dudleyana* I. Comparison of $F_{ST}$ between allozymes and morphological traits. Genetics, 140: 733-744.

**Rafalski J. A. and Tingey S. V.** 1993. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends in Genetics, 9: 275-280.

**Ridout C. J. and Donini P.** 1999. Use of AFLP in cereal research. Trends in Plant Sciences, 4: 76-79.

**Streiff R., Labbe T., Bacilieri R., Steinkellner H., Glössl J. and Kremer A.** 1998. Within-population genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. Molecular Ecology, 7: 317-328.

**Young A. G., Merriam H. G. and Warwick I.** 1993. The effects of forest fragmentation on genetic variation in *Acer saccharum* Marsh. (sugar maple) populations. Heredity, 71: 277-289.

Received July 1, 2000