# HOW MANY MACROGAMETOPHYTES PER INDIVIDUAL SHOULD BE ANALYZED TO GENOTYPE CONIFER SEED TREES AT GENETIC MARKERS?

**J. M. Pastorino[1] & H.-R. Gregorius[2]**

[1)] Unidad de Genética Forestal, INTA EEA Bariloche, CC 277, 8400 S.C. de Bariloche, R. N., Argentina, e-mail: mpastorino@bariloche.inta.gov.ar

[2)] Institut für Forstgenetik und Forstpflanzenzüchtung, Universität Göttingen, Büsgenweg 2, D- 37077 Göttingen, Germany, e-mail: hgregor@gwdg.de

*Received March 22, 2001; accepted March 12, 2002*

## ABSTRACT

Each macrogametophyte in gymnosperms, usually called endosperm, represents a single meiotic event in the parental plant, and it is genetically identical with the ovule of the corresponding seed. This particularity of conifer seeds is typically used to prove the mode of inheritance of isozyme markers and subsequently to genotype seed trees for those markers to be used in population studies. In population studies several macrogametophytes must be analyzed for each single tree in order to deduce heterozygosity (if different haplotypes are found) or infer homozygosity (if only one haplotype is found). Sampling of macrogametophytes of a tree entails statistical error in genotyping. Yet, the probability of misclassification of homozygosity can be controlled. It is well known that this probability is a function of the number of macrogametophytes surveyed, and formulae are available for the single-locus case. The multiple-locus case still awaits explicit analysis. The present paper shows how in this case the error probability depends on the number of analyzed loci. Finally, the multiple-locus multiple-tree case is analyzed. For each of the three cases the proper minimum sample size of macrogametophytes to genotype conifer seed trees is specified. The relevance of the results is demonstrated for several examples taken from the literature.

**Key words:** isozymes, macrogametophytes, conifer seed trees, genotype classification

## INTRODUCTION

A particular feature of conifer seeds allows to directly observe, instead of infer, the segregation of both alleles in a heterozygous individual. The primary endosperm of the gymnosperms is not the result of double fertilization as in angiosperms. Instead, this haploid parenchyma develops mitotically from the single survivor of the tetrad which arises after meiosis from the spore mother cell. Thus, the endosperm tissue of each seed, properly called macrogametophyte, represents a single meiotic event in the parental plant, and it is genetically identical with the ovule of the corresponding seed.

This particularity of conifer seeds has been profited since the early '70s to prove the mode of inheritance of isoenzymatic phenotypes and subsequently to identify genotypes of conifer seed trees (i.e. BARTELS 1971, BERGMANN 1971 and 1973, TIGERSTEDT 1973). While no variation is expected in the macrogametophytes of a homozygous seed tree, segregation into two classes of equal frequency is expected for a heterozygous seed tree according to MENDEL's first law of inheritance.

To prove the mode of inheritance of a banding pattern, deviation of the observed segregation from the MENDELian expectation is statistically tested. For this analysis usually some 20 to more than 100 macrogametophytes per individual from some few trees are electrophoresed. Thus, this method of inheritance analysis dispenses with the classic troublesome method of controlled crosses.

The conceptual basis of this approach towards an analysis of inheritance that rests on observations of haploid phases (gamete equivalents) and infers from these observations genetic characteristics of the diploid phase was addressed by GILLET (1996). This author developed the conceptual basis into a computarized method (HAPLOGEN, available at www.uni-forst.gwdg.de/forst/ fg/index.htm of qualitative inheritance analysis of arbitrarily complex electrophoresis banding patterns.

Once the mode of inheritance of a banding pattern is established in this way, macrogametophytes can be used to genotype adult individuals. This task requires a much smaller number of macrogametophytes per tree to be surveyed. When two macrogametophyte classes are observed at a certain (already proved) locus in a sample of seeds of a given tree, heterozygosity of that tree (with respect to that locus) is deduced irrespective of the frequency of each class. In fact, when sampling seed sequentially, the first occurrence of two classes confirms heterozygosity at the studied locus. On the

other hand, when all of the electrophoresed macrogametophytes of a tree belong to the same class, homozygosity is assumed at the locus. Homozygosity can actually be inferred only to the degree to which the sample is likely to detect all classes potentially produced by the tree. Inference of homozygosity thus is the crucial step in genotyping, and sample strategies as well as sample sizes have to be decided upon with this fact in mind.

Considering a fixed amount of experimental effort, the use of vegetative (diploid) tissue in order to perform a genetic inventory seems to be the best option because in this case just one lane in the electrophoresis gel is enough to genotype an individual instead of the several required by the macrogametophyte method. Nevertheless, there are multiple and different cases where the analysis of endosperm is preferable; proof of this is the large quantity of isoenzymatic studies found in the literature based on the analysis of macrogametophytes. An acceptable good electrophoretical resolution is normally harder to get with the use of vegetative tissue because much more secondary compounds which hinder the enzymatic processes are found in such tissues. Leaves are an usual alternative (the younger the better), but they must be kept fresh till the electrophoretical analysis and this is not always possible, sometimes the lapse between field collection and laboratory analysis is too large. Buds are a good option because they are not so delicate as leaves, but unfortunately some species do not form buds, as it is the case of *Austrocedrus chilensis* (D.Don) Florin et Boutelje. The analysis of embryos is a further alternative, but this is a different generation. Someone could be interested in analyzing the adult generation exclusively. The marker loci determined in endosperm may be not active in different vegetative tissues (e.g. in *A. chilensis* one out of the 12 determined markers is not active in embryos, and two more are not active in leaves). Finally, zymograms become much more complex in diploid tissue and sometimes the overlapping of different zones makes impossible to identify the alleles. The existence of null alleles force the use of haploid tissue because the loci with such alleles present dominant inheritance (e.g. one of the two alleles found at the *Got2* locus in *A. chilensis* is null), and in fact their existence can never be ruled out when genotypic structures of new populations are being determined. Particularly in view of the widespread utilization of DNA-markers showing dominance (such as RAPD, AFLP), it is hard to see how reliable estimates of genotypic and allelic frequencies can be obtained at all without reference to tissue exhibiting gametic equivalence.

Although some authors prefer to work with more macrogametophytes (i.e. BERGMANN 1978: 9, EL-KASSABY *et al.* 1982: 8, YEH 1988: 10, BEAULIEU & SIMON 1994: 8, PASTORINO 2001: 8, PUGLISI *et al.* 1999: 12), population studies based on genotypes determined by means of six or even less macrogametophytes per individual are very common (i.e. FINS & LIBBY 1982: 3, LEDIG *et al.* 1983: 6 to 8, STRAUSS & CONKLE 1986: 6 to 8, CONKLE 1981 and 1987: 6, MORGANTE & VENDRAMIN: 1991: 6, MILLAR & MARSHAL 1991: 5 to 8).

The aim of the present paper is to determine the proper minimum size of the sample of macrogametophytes to genotype conifer seed trees for population genetic studies at arbitrary numbers of gene loci. Herewith, a "fixed total amount of experimental effort" might be the criterion that determines the sampling strategy. In this case, a reduction of the number of macrogametophytes surveyed per tree would allow an increment of the number of surveyed trees per population (MORRIS & SPIETH 1978). The decision on an optimal relation between the number of macrogametophytes per tree and number of trees will then be governed by a tolerable risk of misclassifying the trees' genotypes. In view of the frequently felt discontent with this topic, the present paper will be organized such that it starts out by recalling the principles underlying the well established single-locus single-tree case and proceeds to arbitrary numbers of loci and trees by formulating and applying the general probabilistic basis of this principle.

## THE SINGLE-LOCUS SINGLE-TREE CASE

As was pointed out in the introduction, sequential sampling of gametophytes may end up with a definite decision in favor of heterozygosity, but such a decision is not possible for the assignment of homozygosity to a tree. There always remains a risk of misclassifying actual heterozygosity of an individual as homozygosity. This is the price that has to be paid for the simplicity of the method applicable in conifers. However, the probability of misclassification of homozygosity can be calculated and then handled in a way to minimize it.

As it was first calculated by TIGERSTEDT (1973), the probability of considering a seed tree as homozygous that is actually heterozygous is a function of the number of macrogametophytes surveyed. The probabilistic reasoning is based on the assumption that macrogametophytes as gamete equivalents are produced according to the MENDELian laws, so that there is a probability of ½ that a macrogametophyte carries either one of the two alleles of a heterozygous tree. If two macrogametophytes of the universal infinite set of macrogametophytes of a particular heterozygous tree are analyzed,

the probability that both of them carry the same allele is the product of both single probabilities because they are independent events. Sampling with replacement from the thus defined totality of macrogametophytes yields a probability

$$P = \left(\frac{1}{2}\right)^n \qquad [1]$$

of sampling $n$ macrogametophytes all carrying the same allele. The same probability is obtained for the other allele, so that the overall probability to $n$ times sample the same of either of the two alleles equals

$$P_e = \left(\frac{1}{2}\right)^n + \left(\frac{1}{2}\right)^n = \left(\frac{1}{2}\right)^{n-1} \qquad [2]$$

The formal prerequisites underlying this derivation are the mutual exclusivity of events (that is, if one allele occurs the other one does not) and KOLMOGOROV's pertaining third axiom of additivity. Equation [2] specifies the probability of misclassifying a heterozygote as homozygote, and the index $e$ of $P$ addresses the fact that it is an error probability. This is the formula usually found in the literature (e.g. BERGMANN 1978, CONKLE 1981, RITLAND & EL-KASSABY 1985, EL-KASSABY *et al.* 1994, BEAULIEU & SIMON 1994, PASTORINO & GALLO 1998). By applying the "complement theorem" to equation [2], the probability of correct classification of a heterozygote now becomes

$$P_{ne} = 1 - P_e = 1 - \left(\frac{1}{2}\right)^{n-1} \qquad [3]$$

where the index $ne$ in $P_{ne}$ indicates "no error". Solving equation [3] for $n$ yields

$$n = 1 + \frac{\ln(1 - P_{ne})}{\ln\frac{1}{2}} \qquad [4]$$

which specifies the number of macrogametophytes which are necessary to be analyzed per tree in order to avoid misclassification of homozygosity of the tree at a single locus at the error level $P_e$. For the usually accepted 5 % error probability (i.e. $P_e = 0.05$ or $P_{ne} = 0.95$), $n$ equals 5.32, that is, six macrogametophytes would be necessary.

Error probabilities may of course increase if segregation distortion leads to substantial deviations from the ideal Mendelian proportions. If $s$ is the frequency of the rarer of two alleles of a heterozygote in the gametic output ($s < 0.5$), application of the above line of reasoning immediately leads to

$$P_e = s^n + (1 - s)^n$$

so that

$$P_{ne} = 1 - s^n - (1 - s)^n$$

This does not allow for an analytical solution for the sample size $n$. It can however be used in standard computerized approximation algorithms. The essential and trivial message is that, as $s$ approaches zero, the sample size required to ensure a certain probability $P_{ne}$ of correct classification tends to infinity. Hence, there can be no a priori security of correct classification. However, since large segregation distortions (small $s$) imply that an actually heterozygote individual effectively functions as a homozygote during reproduction, no essential error is made if its heterozygosity is not detected. To avoid unnecessary complications, the following considerations of multiple loci will therefore be based on regular segregation.

## THE MULTIPLE-LOCUS SINGLE-TREE CASE

When the genotype of an individual is to be identified at more than a single locus, the relevant probability of misclassification is that for at least one of the surveyed loci with focus on its complement, the probability of correct classification at all loci. The probability of correct classification at all of $l$ loci will be denoted by $P_{ne}(l)$, so that $P_{ne}(1) = P_{ne}$. The presumably easiest way to derive this probability for a number of $l$ loci is to consider among the set of all samples of size $n$ those with correct classification at $l - 1$ loci. The probability of this set is $P_{ne}(l-1)$. Under the assumption of free recombination among all loci (which will be returned to later) one has stochastic independence among the loci in the gametes, so that the probability of correct classification of the $l$-th locus, given correct classification at the $l - 1$ loci, equals $P_{ne}$ as given in equation [3]. Hence, the probability $P_{ne}(l)$ of correct classification at $l$ loci becomes $P_{ne}(l) = P_{ne}(l-1).P_{ne}$, and by backward iteration of this transition equation one obtains the desired result

$$P_{ne}(l) = (P_{ne})^l = \left[1 - \left(\frac{1}{2}\right)^{n-1}\right]^l \qquad [5]$$

As in the single-locus case, the sample size required to guarantee a given probability of correct classification

results from solving equation [5] for $n$, which yields

$$n = 1 + \frac{\ln(1 - \sqrt[l]{P_{ne}(l)})}{\ln\frac{1}{2}} \qquad [6]$$

The assumption of free recombination is indeed conservative in the sense that it marks the worst case scenario for the classification of heterozygosity. To see this consider the extreme case of complete linkage between two heterozygous loci. Correct classification of heterozygosity for one locus from a sample of macrogametophytes implies correct classification of the other locus. Hence, the possibility of misclassification of one locus and correct classification of the other is ruled out as one of the possibilities to misclassify at least one locus. This reduces the overall probability of misclassification and thus increases the probability of correct classification. It is easily verified that this principle extends to arbitrary numbers of pairs of loci linked and degrees of incomplete linkage. Therefore, $P_{ne}(l)$ as given by equation [5] specifies the lowest limit of this probability, and, consequently, the sample size $n$ given in equation [6] applies to all situations where no prior information on the linkage relationships is available.

## SEPARATE SAMPLING FOR EACH GENE LOCUS

The probability $P_{ne}(l)$ of correct classification given in equation [5] applies to electrophoretic surveys in which all $l$ loci are scored for each of $n$ macrogametophytes. The usual method of sampling when working with isozymes on horizontal starch gel is to survey the $l$ loci in a unique set of macrogametophytes per tree (TIGERSTEDT 1973 can be cited as an exception). This method of sampling will be termed "joint sampling" in the following. As an advantage of this electrophoretic technique, several enzyme systems are revealed in each single sample just by using the same homogenate in two (or even more) gels and by slicing horizontally each gel in up to four slices. Besides, more than one locus can be found in each enzyme system. This fact leads to the situation that the haplotype of a single macrogametophyte can be identified at more than 20 loci in the great majority of conifers just in a totally normal routine.

If this is not feasible technically (i.e. too small macrogametophytes), it might be necessary at the extreme to take a separate sample of macrogametophytes for each of $l$ loci to be classified. The method will be termed "separate sampling". Denoting by $n_i$ the sample size for the $i$-th locus and sampling each locus independently of all others, the probability $P_{ne}(l)$ of correct classification of heterozygosity at all $l$ loci again results from multiplication of the single-locus probabilities $P_{ne}$ given in equation [3], i.e.

$$P_{ne}(l) = \prod_{i=1}^{l}\left[1 - \left(\frac{1}{2}\right)^{n_i - 1}\right] \qquad [7]$$

If all single-locus sample sizes are the same, i.e. $n_i = m$ for $i = 1,...,l$, then this equation is formally identical to equation [5] for joint sampling. This raises the question as to the difference in efficiency between the two methods of sampling, where efficiency can be considered with respect to overall sample sizes, to the experiment effort and to the accuracy in the identification of genotypes.

Considering equal overall sample sizes $n$ for both methods of sampling, then $\sum_i n_i = n$ and

$$\prod_i\left[1 - \left(\frac{1}{2}\right)^{n_i - 1}\right] < \prod_i\left[1 - \left(\frac{1}{2}\right)^{n - 1}\right] = \left[1 - \left(\frac{1}{2}\right)^{n_i - 1}\right]^l$$

Hence, as expected, for separate sampling the probability of correct classification is always smaller than for joint sampling. Separate sampling thus requires a higher overall sample size to reach the same probability of correct classification as joint sampling. To reach equality in the probability of correct classification between both sampling methods, it is required for separate sampling to sample for each locus the same number of macrogametophytes as in the single sample taken for joint sampling. Hence, the total number of macrogametophytes required in separate sampling to reach the same probability of correct classification as in joint sampling is $l$ times the number of macrogametophytes used in joint sampling. This is true in the case of free recombination between all surveyed loci. Otherwise, as it was demonstrated above, if linkage exists among any of the surveyed loci the probability of correct classification is still greater for joint sampling because separate sampling means stochastic independence among the samples of each locus, no matter linkage exists or not.

Usage of the same set of macrogametophytes to survey all of the studied loci is thus strongly recommendable, because the involved total amount of experimental effort is considerably smaller, and in case of linkage among the loci this sampling method would even mean a smaller error risk.

In case of starch gel electrophoresis, separate sampling

may in fact mean a series of joint samples due to the fact that more than one locus can often be surveyed in each enzyme system. Therefore, the probability of correct classification results as the product of the corresponding joint sampling probabilities.

## THE MULTIPLE-LOCUS MULTIPLE-TREE CASE

In population genetic studies, the object of study is not a single tree but rather a collection of trees whose genetic structure is to be estimated. Thus, the attempt is typically made to identify the genotypes of 10 to 30 individuals, which are assumed to be a representative sample of a given collection or deme, at several loci jointly. In this situation estimation problems exist at two levels, the level of genotype identification and the level of genotypic frequencies. In this paper we are concerned with the former level. Making use of the above results, the problem can be formulated as one of determining the probability of correct classification of heterozygosity at each of $l$ loci for each of $a$ trees on the basis of samples of macrogametophytes. This probability will be denoted by $P_{ne}(l,a)$. Given $a$ independent samples of equal size $n$, the probability $P_{ne}(l,a)$ results directly as the product of the single tree probabilities given in equation [5], i.e.

$$P_{ne}(l,a) = (P_{ne}(l))^a = \left(\left[1 - \left(\frac{1}{2}\right)^{n_i - 1}\right]^l\right)^a \qquad [8]$$

From this, one obtains the number of macrogametophytes required per tree to guarantee with probability $P_{ne}(l,a)$ correct classification of the $l$-locus genotypes of all of $a$ trees as

$$n \quad 1 \quad \frac{\ln(1 \quad \sqrt[l.a]{P_{ne}})}{\ln\frac{1}{2}} \qquad [9]$$

An interesting statement in this equation is to be seen in the fact that $n$ depends on the number of loci and the number of trees only via their product $l \cdot a$. Thus, large numbers of trees and small numbers of loci may require the same $n$ as small numbers of trees and large numbers of loci.

## DISCUSSION

Each of the calculated formulae are true for its own regard level. To choose one of them implies a decision that should be based on the objectives of a particular study, but also should look at some practical considerations.

It is difficult to imagine a research objective to which the first level applies, that is the single-locus single-tree level.

Genotype frequencies are needed in order to infer previous processes or current dynamics of the analyzed populations. Therefore, the third level appears to be the appropriate in such cases. However, the increase of experimental effort at this level is drastic, and the implied increase in accuracy of genotype identification is in fact hardly needed for population genetic purposes. For example, when 30 trees are to be genotyped at 10 loci, 300 genotype identifications are to be performed, and then, some few mistakes (say three out of these 300, i.e., 1 % error) can surely be tolerated because they can hardly be relevant. Thus, for the majority of the population genetic studies the probability of making not a single mistake at the third level becomes a too strict condition. Instead, the multiple-locus single-tree level may in many cases be adequate.

Some remarks with respect to the intermediate multiple-locus single-tree level are still in order. Formula [5] applies to a worst case scenario in which all of the surveyed loci are truly heterozygous and recombine freely. Therefore, monomorphic loci must obviously not be considered and $l$ must be defined as the number of the surveyed polymorphic loci.

Since the number of actually heterozygous loci is unknown before the identification of genotypes, the worst case scenario should be applied. However, a more real scenario could be considered if the polymorphism type of each locus were known for the species that is being studied. In this case, virtual monomorphic loci (as called by LEWONTIN 1985, that is, polymorphisms with a predominant allele escorted by one or more extremely rare ones[1]) could also be excluded from $l$ because they are in fact extremely unlikely to occur in heterozygous form. With the same criterion, one could further reduce the number of considered loci fixing a certain threshold that excludes the lower portion of the minor polymorphisms. One must, however, keep in mind that each reduction lowers the accuracy of the genotype identification.

This same concept was treated by ALEXANDER *et al.* (1995), however without taking into account the multilocus case. By means of a BAYESian analysis, they showed that the probability of misclassification of

[1] A frequency of 2 % is argued by FINKELDEY (1992) as a possible threshold between adaptively significant and deleterious alleles, and virtual monomorphism is defined on the basis of this threshold by PASTORINO (2001).

diploid genotypes using haploid genotypes decreases with the increment of homozygosis of the surveyed locus in the sampled population. Thus, prior (and hardly available) knowledge about the genetic structures or the inbreeding coefficients of the studied populations would give orientation about a misclassification risk level.

In a population genetic study on *Austrocedrus chilensis* (*Cupressaceae*) (PASTORINO 2001), the genotype of each individual was identified at 12 polymorphic isoenzymatic loci, six of which being virtually monomorphic, one a minor polymorphism and the remaining five major polymorphisms.

Eleven of these loci were genotyped with a set of eight macrogametophytes. The 12th marker (one of the major polymorphisms) was determined after finishing the genotyping at the other 11 loci. Therefore, new macrogametophytes were electrophoresed to identify genotypes just at this locus. In this case, six macrogametophytes were analyzed. The sampling method is thus separate sampling consisting of one joint sample for 11 loci and one single-locus sample. According to the explanations following equation [7], the probability of making no errors at any of the loci is

$$P_{ne} = \left[1 - \left(\frac{1}{2}\right)^{8-1}\right]^{11} \times \left[1 - \left(\frac{1}{2}\right)^{6-1}\right] = 0.889$$

However, excluding the virtual monomorphic loci this probability rises to

$$P_{ne} = \left[1 - \left(\frac{1}{2}\right)^{8-1}\right]^{5} \times \left[1 - \left(\frac{1}{2}\right)^{6-1}\right] = 0.931$$

where possibly existing linkage is not taken into account. In other words, the probability of misclassifying a heterozygous individual at at least one of the surveyed loci (that is, at least one mistake out of 12 genotype determinations) is smaller than 7 %.

For an accepted error probability of 5 % ($P_{ne}(l)$ = 0.95), the number of necessary macrogametophytes could have been calculated before a genetic inventory by

$$n = 1 + \frac{\ln(1 - \sqrt[6]{0.95})}{\ln\frac{1}{2}} = 7.88$$

Thus, the original intuitive choice of eight macrogametophytes was adequate to survey those 12 loci in that species.

If the multilocus case is not considered, $n$ would equal 5.32, as said, and six macrogametophytes would have been electrophoresed per tree. Calculating now the probability of no mistakes with this sample size but for the multilocus case, and again excluding the virtual monomorphic loci, $P_{ne}(l)$ equals 0.826, so that the corresponding error probability of 17.4 % considerably exceeds the accepted 5 %.

Some additional examples from the literature may be useful to demonstrate the relevance of the multilocus consideration. Although the sampling method (joint or separate) is usually not specified in any paper, joint sampling is assumed when horizontal starch gel electrophoresis is performed, because as said, it is the usual sampling method for this technique.

The genetic variation of 34 natural populations of *Sequoiadendron giganteum* (Lindl.) Buch. was studied on the basis of the genotypes of mother trees identified at four major polymorphic allozyme loci by means of the analysis of three macrogametophytes per tree (FINS & LIBBY 1982). With this sampling size, the probability of misclassifying a heterozygous individual at at least one of the surveyed loci was 68 %.

A similar population genetic study was performed on *Chamaecyparis lawsoniana* (A. Murr.) Parl (MILLAR & MARSHALL 1991). Here, genotypes were identified at 26 polymorphic allozyme loci utilizing five to eight macrogametophytes per tree. With the lowest boundary of this sampling size range, the probability of a mistake at at least one of the surveyed loci rises to 81 % as compared with the last example. This is due to the fact that considerably more loci were investigated with only a slightly increased sample size.

A third study dealt with the genetic variation of *Cupressus macrocarpa* Hartw. (CONKLE 1987). In this case, genotypes of seed trees were identified at 17 polymorphic isozyme loci analyzing six macrogametophytes per tree. The corresponding $P_e$ was 42 %, which is again far above the accepted maximum error level.

Populations are composed by individuals whose genomes can hardly be characterized by the independent view of single loci. A multilocus approach should always be considered in a population genetic study. The identification of genotypes is a basic step of such studies. If these data suffer from a certain unreliability, the information derived from them could result in mere speculation. A relevant value should be given to that tedious but fundamental task.

## REFERENCES

ALEXANDER, H. M., SLADE, N. A. & GOMULKIEWICZ, R. 1995: A Bayesian approach to the inference of diploid genotypes using haploid genotypes. *Theor. Appl. Genet.* **91**: 1284–1287.

BARTELS, H. 1971: Genetic control of multiple esterases from needles and macrogametophytes of *Picea abies*. *Planta (Berl.)* **99**: 283–289.

BEAULIEU, J. & SIMON, J.-P. 1994: Inheritance and linkage relationships of allozymes in *Pinus strobus* L. *Silvae Genetica* **43**: 253–261.

BERGMANN, F. 1971: Genetische Untersuchungen bei *Picea abies* mit Hilfe der Isoenzym-Identifizierung. I. Möglichkeiten für genetische Zertifizierung von Forstsaatgut. *Allg. Forst- u. J.- Ztg.* **142**: 278–280.

BERGMANN, F. 1973: GenetischeUntersuchungen bei *Picea abies* mit Hilfe der Isoenzym-Identifizierung. II. Genetische Kontrolle von Esterase- und Leucin-aminopeptidase-Isoenzymen im haploiden Endosperm ruhender Samen. *Theor. Appl. Genet.* **43**: 221–225.

BERGMANN, F. 1978: The allelic distribution at an acid phosphatase locus in Norway spruce (*Picea abies*) along similar climatic gradients. *Theor. Appl. Genet.* **52**: 57–64.

CONKLE, M. T. 1981: Isozyme variation and linkage in six conifer species. *In:* Proc. Symp. Isoz. North. Am. For. Trees and For. Ins. July 27, 1979. (ed. M.T. Conkle). Berkeley, California: 11–17.

CONKLE, M. T. 1987: Electrophoretic analysis of variation in native Monterey cypress (*Cupressus macrocarpa* Hartw.). *In:* Proc. Conf. Conservation and Management of Rare and Endangered Plants. Native Plants Society. Sacramento, CA: 249–256

EL-KASSABY, Y. A., RUSSELL, J. & RITLAND, K. 1994: Mixed mating in an experimental population of western red cedar, *Thuja plicata*. *J. Hered.* **85**: 227–231.

EL-KASSABY, Y. A., YEH, F. C. & SZIKLAI, O. 1982: Inheritance of allozyme variants in coastal Douglas-fir (*Pseudotsuga menziesii* var. *menziesii*). *Can. J. Genet. Cytol.* **24**: 325–335.

FINKELDEY, R. 1992:Auswahlkriterien und Anlage genetischer Ressourcen bei der Fichte (*Picea abies* (L.) Karst.). *Forstarchiv* **63**: 25–32.

FINS, L. & LIBBY, W. J. 1982: Population variation in *Sequoiadendron*: seed and seedling studies, vegetative propagation, and isozyme variation. *Silvae Genetica* **31**: 102–110.

GILLET, E. M. 1996: Qualitative inheritance analysis of isozymes in haploid gametophytes: Principles and a computarized method. *Silvae Genetica* **45**: 8–16.

LEDIG, F. T., GURIES, R. P. & BONEFELD, B. A. 1983: The relation of growth to heterozygosity in pitch pine. *Evolution* **37**: 1227–1238.

LEWONTIN, R. C. 1985: Population genetics. *Ann. Rev. Gent.* 19: 81–102.

MILLAR, C. I. & MARSHALL, K. A. 1991: Allozyme variation of "Port-Orford cedar" (*Chamaecyparis lawsoniana*): implications for genetic conservation. *For. Sci.* **37**: 1060 –1077.

MORGANTE, M. & VENDRAMIN, G. G. 1991: Genetic variation in Italian populations of *Picea abies* (L.) Karst. and *Pinus leucodermis* Ant. *In:* Genetic Variation in European Populations of Forest Trees. (eds. Müller-Starck, G, Ziehe, M.). pp. 205–227. Sauerländer's, Frankfurt a.M.

MORRIS, R. W. & SPIETH, P. T. 1978: Sampling strategies for using female gametophytes to estimate heterozygosity in conifers. *Theor. Appl. Genet.* **51**: 217–222.

PASTORINO, M. J. 2001: Genetic variation and reproduction system of *Austrocedrus chilensis* (D. Don) Florin et Boutelje, a cypress endemic to the Andean-Patagonian Forest. Ph. D. thesis, Faculty of Forest Sciences and Forest Ecology, Georg-August University of Göttingen. Cuvillier Verlag, Göttingen, 164 pp.

PASTORINO, M. J. & GALLO, L. A. 1998: Inheritance of isozyme variants in *Austrocedrus chilensis* (D. Don) Florin et Boutelje. *Silvae Genetica* **47**: 15–20.

PUGLISI, S., LOVREGLIO, R., CIFARELLI, S. & LEONE, V. 1999: Allozyme variation in five populations of *Pinus halepensis* Mill. in Southern Italy. *Forest Genetics* **6**: 241–24 6.

RITLAND, K. & EL-KASSABY, Y. A. 1985: The nature of inbreeding in a seed orchard of Douglas fir as shown by an efficient multilocus model. *Theor. Appl. Genet.* **71**: 375–384.

STRAUSS, S. H. & CONKLE, M. T. 1986: Segregation, linkage, and diversity of allozymes in knobcone pine. *Theor. Appl. Genet.* **72**: 483–493.

TIGERSTEDT, P. M. A. 1973: Studies on isozyme variation in marginal and central populations of *Picea abies*. *Hereditas* **75**: 47–60.

YEH, F. C. 1988: Isozyme variation of *Thuja plicata* (Cupressaceae) in British Columbia. *Biochem. Syst. and Ecol.* **16**: 373–377.