Heredity, in press (1986). —— KAHLER, A. L. and WEHRHAHN, C. F.: Associations between quantitative traits and enzyme loci in the F2 population of a maize hybrid. Theor. Appl. Genet. **72**, 15—26 (1986). —— KATZ, M. and GOUX, J.-M.: The statistical properties of genetic absolute distance. Biom. J. **28**, 729—739 (1986). —— LEDIG, F. T., GURIES, R. P. and BONNEFELD, B. A.: The relation of growth to heterozygosity in pitch pine. Evolution **37**, 1227—1238 (1983). —— LERNER, I. M.: Genetic homeostasis. Oliver and Boyd, Edinburgh (1954). —— MCMILLIN, D. E., ALLAN, R. E. and ROBERTS, D. E.: Association of an isozyme locus and strawbreaker foot resistance derived from *Aegilops ventricosa* in wheat. Theor. Appl. Genet. **72**: 743—747 (1986). —— MEJNARTOWICZ, L. E.: Changes in genetic structure of Scots pine (*Pinus sylvestris* L.) population affected by industrial emission of fluoride and sulfur dioxide. Genetica Polonica **24**, 41—50 (1983). —— MEJNARTOWICZ, L. E.: Enzymatic investigations on tolerance in forest trees. In: KOZIOL, M. J. and WHATELY, F. R. (eds.): Gaseous air pollutants and plant metabolism. Butterworths, London, 381—398 (1984). —— MITTON, J. B. and GRANT, M. C.: Observations on the ecology and evolution of quaking aspen, *Populus tremuloides*, in the Colorado Front Range. Am. J. Bot. **67**, 202—209 (1980). —— MÜLLER-STARCK, G.: Sexually asymmetric fertility selection and partial self-fertilisation 2. Clonal gametic contributions to the offspring of a Scots pine seed orchard. Silvae Fennica **16**, 99—106 (1982). —— MÜLLER-STARCK, G.: Genetic differences between "tolerant" and "sensitive" beeches (*Fagus sylvatica* L.) in an environmental stressed adult forest stand. Silvae Genetica **34**, 241—247 (1985). —— NICHOLS, E. A. and RUDDLE, F. H.: A review of enzyme polymorphism, linkage electrophoretic conditions for mouse and somatic cell hybrids in starch gels. Histochem. and Cytochem. **21**: 1066—1081 (1973). —— POULIK, M. D.: Starch gel electrophoresis in a discontinuous system of buffers. Nature **180**, 1477—1478 (1957). —— RABE, R. and KREEB, K, H.: Wirkungen von SO_2 auf die Enzymaktivität in Pflanzenblättern. Z. Pflanzenphysiol. **97**, 215—226 (1980). —— RUDIN, D.: Inheritance of glutamate-oxalate-transaminases (GOT) from needles and endosperms of *Pinus sylvestris*. Hereditas **80**, 296—300 (1975). —— RUDIN, D.: Leucine-amino-peptidases (LAP) from needles and macrogametophytes of *Pinus sylvestris* L. Inheritance of allozymes. Hereditas **85**, 219—226 (1977). —— SCANDALIOUS, J. G.: Genetic control of multiple molecular forms of enzymes in plants: A review. Biochem. Gen. **3**, 37—79 (1969). —— SCHOLZ, F. and BERGMANN, F.: Selection pressure by air pollution as studied by isozyme-gene-systems in Norway spruce exposed to sulphur-dioxide. Silvae Genetica **33**, 238—241 (1984). —— SHAW, C. R. and KOEN, A. L.: Starch gel electrophoresis of enzymes. In: SMITH, I. (ed.): Chromatographic and electrophoretic techniques. Vol. II., Pittman Press, England, 325—364 (1968). —— SHAW, C. R. and PRASAD, R.: Starch gel electrophoresis of enzymes - a compilation of recipes. Biochem. Gen. **4**, 297—320 (1970). —— SZMIDT, A. E. and YAZDANI, R.: Electrophoretic studies of genetic polymorphism of shikimate and 6-phosphogluconate dehydrogenases in Scots pine (*Pinus sylvestris* L.). Arboretum Kornickie **29**, 63—71 (1984).

Some Aspects of the Population Structure of Black Spruce in Central New Brunswick

By T. J. B. BOYLE*) and E. K. MORGENSTERN

Dept. of Forest Resources, University of New Brunswick, Fredericton, New Brunswick, E3B 6C2, Canada

(Received 9th December 1985)

Summary

Genetic diversity within and genetic differentiation among six populations of black spruce (*Picea mariana* [MILL.] B.S.P.) in central New Brunswick were investigated using allozyme frequencies at 12 loci. Average heterozygosity ranged from 0.192 to 0.253 in the different populations. On the average, the percent heterozygous loci/individual ranged from 17.08 to 22.29, the average number of alleles/locus from 2.17 to 2.50 and the effective number of alleles/locus from 1.24 to 1.38. These values are comparable with those from other conifer species though they might be biased upwards somewhat due to the limited number of polymorphic, independently segregating loci sampled. Other measures of genetic diversity produced similar results.

Partitioning the observed variation into within- and among-population components by using F-statistics led to an estimate of within-population variation amounting to 99 per cent of total variation. These results suggest that emphasis be placed on intra-population sampling in black spruce for tree improvement. The relationship between geographic distance and Nei's genetic distance appeared to be weak, suggesting that isolation by distance may not be responsible for the observed differentiation.

Key words: *Picea mariana*, isozymes, population structure, gene diversity, genetic distance, heterozygosity.

Zusammenfassung

Innerhalb und zwischen 6 Populationen von *Picea mariana* im mittleren New Brunswick wurde die genetische Diversität durch elektrophoretische Methoden untersucht. Es wurden die Genhäufigkeiten an 12 Loci ermittelt. Der durchschnittliche Heterozygotiegrad lag in den verschiedenen Populationen zwischen 0,192 und 0,253. Die Durchschnittswerte lagen in folgenden Bereichen: % der heterozygoten Loci pro Individuum, 17,08—22,29; Zahl der Allele pro Locus, 2,17—2,50; und Zahl der effektiven Allele pro Locus, 1,24—1,38. Diese Werte sind mit solchen anderer Koniferen vergleichbar, können allerdings etwas hoch eingeschätzt sein, da sie auf einer geringen Zahl von polymorphen, unabhängig aufspaltenden Loci beruhen. Andere Messungen der genetischen Diversität ergaben ähnliche Resultate.

Die Aufteilung der beobachteten Variation in Komponenten „innerhalb" und „zwischen" Populationen durch F-Statistiken ergab einen Beitrag von 99% der Komponente „innerhalb" zur Gesamtvariation. Die Plusbaumauswahl sollte deshalb hauptsächlich auf Auslese innerhalb der Populationen beruhen. Die Korrelation zwischen geographischer Entfernung der Populationen und Nei's genetischer Distanz schien schwach zu sein, sodaß Distanzisolation wenig Einfluß auf die beobachtete Differenzierung haben mag.

Introduction

The population structure of a species has been defined by RIEGER *et al.* (1976) as: "The sum of all factors governing the pattern by which gametes from various individuals unite with each other during fertilization". BROWN (1978) recognizes two groups of evolutionary forces that affect the values which various measures of population genetic structure may assume. These are firstly, the interaction with other aspects of the population structure affecting, for

*) Present adress: Canadian Forestry Service, Petawawa Nat. for. Inst., Chalk River, K0J 1J0 Ontario, Canada.

example, the degree of isolation and therefore the level of genetic drift and migration and secondly, the spatial and temporal distribution of selection intensities.

A knowledge of the population genetic structure of a species is important for a number of reasons. In genetic conservation, for example, estimates of the amount of genetic diversity within populations, and the range and distribution of this diversity among populations, is necessary in order to devise an optimum sampling strategy (Marshall and Brown, 1975). Just as genetic conservation is basically a problem of sampling efficiency, so too are genetic improvement programs, since the prime objective is to maximise genetic gain with minimum effort. Thus, the ideal sampling strategy for any species will depend on its population genetic structure (Guries and Ledig, 1977). An intensive investigation of population structure should therefore preferably precede the initiation of a tree improvement program. This is not usually the case but flexible programs will allow adjustments as more information becomes available (Danzmann and Buchert, 1983).

In measuring genetic diversity in natural populations, it would be advantageous to be able to identify and enumerate genotypes coding for commercially valuable traits such as growth and form. However, due to strong environmental influences on these characters and their polygenic inheritance, such an approach is impossible (Brown, 1978; Brown and Moran, 1981). Isozymes though, being usually codominantly expressed and independent of environmental influence, provide suitable genetic markers, especially as a large number of loci, which may represent a random sample of the genome, can be simultaneously evaluated on one individual (Lewontin, 1974).

The population structure of black spruce *(Picea mariana* [Mill.] B.S.P.) in central Ontario was investigated by Morgenstern (1969) using genetic intra-class correlations based on variance components of 13 characters for populations, sub-populations, open -pollinated families and trees within families. He found that for eight characters associated with germination and early growth, 65 per cent of the variation could be accounted for by variation among families. However, for total height at age two and for four characters associated with phenology, variation among populations (site regions) accounted for 82 per cent of the observed variation. This, together with the very small amount of variation among subpopulations within site regions, led to the conclusion that the variation in such characters is primarily clinal with little evidence for ecotypes.

Table 1. — Enzyme systems assayed.

Enzyme system	Enzyme Commission No.	Abbr.	No. of Loci
Aldolase	E.C.4.1.2.13	ALD	1
Aspartate amino-transferase	E.C.2.6.1.1	AAT	1
Fumarase	E.C.4.2.1.2	FUM	1
Glucose-6-phosphate dehydrogenase	E.C.1.1.1.49	G6P	1
Glutamate dehydrogenase	E.C.1.4.1.3	GDH	1
Isocitrate dehydrogenase	E.C.1.1.1.42	IDH	1
Malate dehydrogenase	E.C.1.1.1.37	MDH	3
Phosphoglucomutase	E.C.2.7.5.1	PGM	1
6-Phosphogluconate dehydrogenase	E.C.1.1.1.44	6PGD	2
Phosphoglucose isomerase	E.C.5.3.1.9	PGI	1
Shikimic acid dehydrogenase	E.C.1.1.1.25	SDH	1

In this paper we report on the amount and pattern of gene diversity based on allozyme frequencies determined for a relatively small area in central New Brunswick. This forms part of a larger study on the mating system and population structure of black spruce in central New Brunswick (Boyle, 1985).

Material and Methods

Sampling

The area sampled constitutes part of the Eastern Lowlands Section of the Acadian Forest Region (Rowe, 1972). Six natural stands (populations), part of a network of stands reserved for seed production across the Province, were investigated *(Figure 1)*. The maximum distance between stands is 52 km *(Table 6)* and there are no major gradients or differences in soils, climate or elevation throughout the area.

Forty dominant or co-dominant trees within an area of about 2 ha were felled in each stand and up to 200 cones per tree collected. Twelve gametophyte-embryo pairs were assayed per tree. The genotype of the mother tree was inferred from the gametophytes. Since the megagametophyte tissue is haploid, the probability that every megagametophyte from a heterozygous mother tree carries the same allele is given by $0.5^{(k-1)}$, where k is the number of megagametophytes assayed, assuming a 1:1 segregation ratio. Thus the probability that a heterozygote is misclassified as a homozygote was 0.5^{11} (< 0.0005), or in other words, fewer than one tree in 2000 would be incorrectly genotyped.

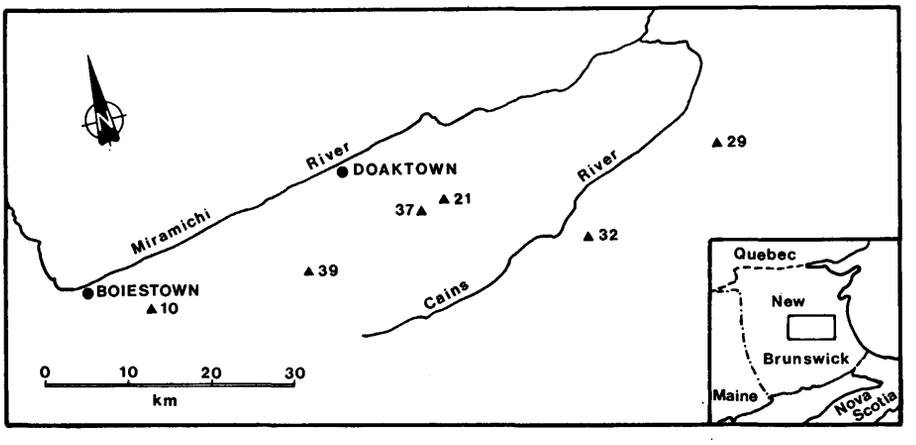

Figure 1. — Map showing location of six populations.

Electrophoresis methods

Horizontal starch gel electrophoresis was used to separate the isozymes at eleven enzyme systems. These are listed in *Table 1*. Details of the procedure and staining recipes used have been given elsewhere (Boyle and Morgenstern, 1985).

Statistical analysis

Measures of gene diversity

Four methods have been used to quantify the amount of gene diversity within populations. These included the following.

(a) The expected heterozygosity at each locus, which is defined as:

$$h_e = 1 - \sum_i p_i^2$$

where p_i is the frequency of the i^{th} allele (Nei, 1975). The average heterozygosity per population (H_e) is then the arithmetic mean of the h_e values over all loci. Sampling variances of average heterozygosity have been derived by Nei and Roychoudhury (1974).

(b) The percentages of heterozygous loci per individual (Yeh and O'Malley, 1980). It should be noted that since only about 27 per cent of the total genetic variation is detectable by electrophoresis (Shaw, 1965), the estimate of the percentage of heterozygous loci per individual should be adjusted accordingly.

(c) The average number of alleles per locus (Yeh and O'Malley, 1980). However, some of these alleles will occur at very low frequency and contribute very little to average heterozygosity or genetic variance (Crow and Kimura, 1970).

(d) The effective number of alleles per locus (n_e) is therefore more useful than the actual number (Crow and Kimura, 1970). It is defined as:

$$n_e = 1/\sum_i p_i^2$$

where p_i is as defined above. The effective number of alleles is maximised when allele frequencies at any locus are equal. Thus it can be considered as the number of alleles of equal frequency required to provide the same level of heterozygosity as that expected from the actual allele frequencies. The average effective number of alleles per population is then the geometric mean of the n values for individual loci (Lundkvist, 1979).

Genetic differentiation

Having quantified the amount of genetic diversity within a population, it is then of interest to examine how this diversity is distributed — whether there is genetic differentiation among populations and, if so, whether this differentiation is in an organised pattern or at random. There are several methods of approaching this question.

(a) A contingency χ^2 test. This tests the null hypothesis that the observed allele frequencies in the different populations are a random sample from a single population (Workman and Niswander, 1970). A significant value indicates therefore that the amount of differentiation between populations is greater than can be expected by chance alone, but gives no information as to how this differentiation is organised.

(b) F-statistics. Wright (1965) developed a set of statistics based on his "fixation index" (F), which measures the net deviation of genotype frequencies from Hardy-Weinberg expectations, in order to apportion genetic diversity to the various levels of a hierachial population. $\bar{F}_{IS}$ is the weighted average fixation index over all populations. F_{ST} represents the correlation among random gametes within populations relative to the total population, and therefore provides a measure of the extent of genetic differentiation among populations. For multiple allelic systems, F_{ST} is equivalent to Nei's (1973) G_{ST}. F_{IT} represents the overall fixation index, and can be estimated either by the "direct" method, as the correlation among gametes over the total population, or by the "indirect" method as:

$$F_{IT} = \bar{F}_{IS} + (1-\bar{F}_{IS})\, F_{ST}$$

Comparison of the estimates provided by the two methods gives information on whether the observed differentiation is random (Linhart *et al.*, 1981). In order to obtain an accurate estimate of the amount of differentiation among populations, a large number of loci forming a random sample of the genome and including both polymorphic and monomorphic loci should be used (Nei, 1973). Chakraborty (1974) derived an expression for the sampling variance of G_{ST} which can be used for tests of significance of F_{ST}.

(d) Genetic distance. Nei (1972) devised a measure of genetic differentiation (D) based on the "gene identity" between populations. Again, a large random sample of both polymorphic and monomorphic loci is desirable (Nei, 1975). Sampling variances of D were derived by Nei and Roychoudhury, (1974).

Results

Fourteen zones of activity were scored for the eleven enzyme systems. However, for two of these (SDH and MDH-1) the genetic origin of the observed variation could not be confirmed.

Since loci with high levels of variability are the most useful for the estimation of mating system parameters, there may be a slight bias in the selection of loci scored in favour of more variable loci. Thus the sample of loci may not be completely random (for example only MDH-2 was monomorphic in all populations). However, since to a large extent comparisons of the relative variability are considered, the estimates of genetic differentiation should not be greatly affected (Shaw, 1982).

Genetic diversity

(a) Allele frequencies and heterozygosities. *Table 2* shows the observed allele frequencies at all the loci for each population, along with the overall mean allele frequencies, observed and expected heterozygosities and average heterozygosity. Two loci (PGM and 6PGD-1) had a mean expected heterozygosity over all populations of greater than 0.5. At both of these loci the same two alleles predominated in all populations with a third allele present at a much lower frequency. Two more loci (G6P and PGI) produced mean expected heterozygosities between 0.25 and 0.5. One allele dominated these loci in all populations. At G6P two other alleles were present at much lower frequencies except for populations 37 and 39 where a fourth allele was recorded. Similarly, at PGI a total of three alleles was present in all populations except 29 and 32 where only two were recorded. Four loci, namely 6PGD-2, GDH, AAT and MDH-3 gave mean heterozygosities between 0.1 and 0.25. In general all these loci had two predominant alleles per population with one allele at a frequency consistently greater than 0.83. A third allele was recorded at 6PGD-2 in populations 10,37 and 39, AAT in population 32 and at MDH-3 in 29 and 39. The final three polymorphic loci (i.e. those with the frequency of the most common allele < 0.99) — ALD, IDH and FUM — had mean expected heterozygosities of less than 0.1. Aldo-

Table 2. — Allele frequencies, expected and observed heterozygosities and overall population means for all 11 polymorphic loci.

Locus	Allele	Population 10	21	29	32	37	39	Mean
ALD	1	0.013	0.013	0.000	0.013	0.013	0.025	0.013
	2	0.987	0.987	1.000	0.975	0.987	0.937	0.979
	3	0.000	0.000	0.000	0.012	0.000	0.038	0.008
Expected h		0.025	0.025	0.000	0.049	0.025	0.119	0.041
Observed h		0.025	0.025	0.000	0.050	0.025	0.125	0.042
AAT	1	0.000	0.000	0.000	0.013	0.000	0.000	0.002
	2	0.987	0.937	0.912	0.912	0.975	0.912	0.939
	3	0.013	0.063	0.088	0.075	0.025	0.088	0.059
Expected h		0.025	0.117	0.160	0.162	0.049	0.160	0.114
Observed h		0.025	0.125	0.175	0.175	0.050	0.075	0.104
FUM	1	0.038	0.013	0.050	0.025	0.025	0.038	0.032
	2	0.962	0.987	0.950	0.975	0.975	0.962	0.968
Expected h		0.072	0.025	0.095	0.049	0.049	0.072	0.062
Observed h		0.075	0.025	0.100	0.050	0.050	0.075	0.063
G6P	1	0.063	0.175	0.187	0.138	0.188	0.162	0.152
	2	0.100	0.013	0.088	0.100	0.050	0.088	0.073
	3	0.000	0.000	0.000	0.000	0.050	0.025	0.013
	4	0.837	0.812	0.725	0.762	0.712	0.725	0.762
Expected h		0.285	0.309	0.432	0.390	0.452	0.440	0.391
Observed h		0.250	0.325	0.350	0.400	0.500	0.375	0.367
GDH	1	0.013	0.025	0.163	0.050	0.150	0.138	0.090
	2	0.987	0.975	0.837	0.950	0.850	0.862	0.910
Expected h		0.025	0.049	0.272	0.095	0.255	0.237	0.163
Observed h		0.025	0.050	0.325	0.050	0.300	0.275	0.171
IDH	1	0.050	0.025	0.050	0.000	0.013	0.025	0.027
	2	0.950	0.975	0.950	1.000	0.987	0.975	0.973
Expected h		0.095	0.049	0.095	0.000	0.025	0.049	0.053
Observed h		0.100	0.050	0.100	0.000	0.025	0.050	0.054
MDH-3	1	0.000	0.000	0.013	0.000	0.000	0.013	0.004
	2	0.950	0.912	0.862	0.937	0.962	0.837	0.910
	4	0.050	0.088	0.125	0.063	0.038	0.150	0.086
Expected h		0.095	0.160	0.240	0.117	0.072	0.276	0.164
Observed h		0.100	0.125	0.275	0.125	0.075	0.175	0.146
PGM	1	0.587	0.550	0.500	0.462	0.512	0.500	0.519
	2	0.063	0.125	0.138	0.138	0.138	0.163	0.128
	3	0.350	0.325	0.362	0.400	0.350	0.337	0.353
Expected h		0.528	0.576	0.600	0.607	0.596	0.610	0.590
Observed h		0.525	0.625	0.575	0.625	0.550	0.450	0.558
6PGD-1	1	0.400	0.363	0.475	0.450	0.387	0.413	0.415
	2	0.562	0.637	0.487	0.537	0.600	0.587	0.568
	3	0.038	0.000	0.038	0.013	0.013	0.000	0.017
Expected h		0.522	0.462	0.535	0.508	0.490	0.485	0.505
Observed h		0.475	0.475	0.575	0.525	0.550	0.475	0.513
6PGD-2	1	0.037	0.025	0.088	0.050	0.062	0.050	0.052
	2	0.950	0.975	0.912	0.950	0.925	0.937	0.942
	3	0.013	0.000	0.000	0.000	0.013	0.013	0.006
Expected h		0.096	0.049	0.160	0.095	0.140	0.118	0.110
Observed h		0.100	0.050	0.125	0.100	0.150	0.125	0.108
PGI	1	0.013	0.012	0.000	0.000	0.013	0.013	0.009
	2	0.825	0.900	0.750	0.825	0.775	0.737	0.802
	4	0.162	0.088	0.250	0.175	0.212	0.250	0.189
Expected h		0.293	0.182	0.375	0.289	0.354	0.393	0.321
Observed h		0.350	0.200	0.350	0.300	0.400	0.325	0.321
Average h		0.223 ±0.086	0.192 ±0.088	0.252 ±0.095	0.221 ±0.095	0.233 ±0.093	0.253 ±0.090	

lase was monomorphic in population 29 with two alleles recorded in 10, 21 and 37 and three in 32 and 39. Similarly, IDH was monomorphic in one population (32) with two alleles present in all the others, the more common allele having a frequency of at least 0.95 in every case. Fumarase was polymorphic in every population with allele frequencies very similar to those at IDH.

Table 4. — The effective number of alleles for each locus in each population.

	10	21	29	32	37	39
ALD	1.03	1.03	1.00	1.05	1.03	1.14
AAT	1.03	1.13	1.19	1.19	1.05	1.19
FUM	1.08	1.03	1.10	1.05	1.05	1.08
G6P	1.40	1.45	1.76	1.64	1.83	1.78
GDH	1.03	1.05	1.38	1.10	1.34	1.31
IDH	1.10	1.05	1.10	1.00	1.03	1.05
MDH-3	1.10	1.19	1.32	1.13	1.08	1.38
PGM	2.12	2.36	2.50	2.55	2.48	2.56
6PGD-1	2.10	1.86	2.15	2.04	1.96	1.94
6PGD-2	1.11	1.05	1.19	1.10	1.16	1.14
PGI	1.41	1.22	1.60	1.41	1.55	1.65

Comparison of the observed and expected heterozygosities shows that in most cases there is very close agreement, especially at the less variable loci, as would be expected. Looking at the mean expected and observed heterozygosities over all populations, again there is very good agreement at most loci. Only at four loci, namely PGM, AAT, G6P and MDH-3, is there any apparent deficiency in observed heterozygosities. Interestingly in all these cases, the major contribution to the deficiency is from population 39.

(b) Percentage of heterozygous loci per individual. Estimates of the percentage of heterozygous loci per individual are given in *Table 3*. It must be remembered that because the selection of loci for this study tended to favour variable loci for the purposes of estimating mating system parameters, the estimates are likely to be upper limits. As might be expected, since observed and expected heterozygosities agreed so closely, those populations with low average heterozygosities also have fewer heterozygous loci per individual.

(c) Average number of alleles per locus. *Table 3* also shows the average number of alleles per locus, for all twelve loci. There appears to be little relationship between the average number of alleles per locus and excepted heterozygosity, reflecting the observation of Crow and Kimura (1970) that often alleles may be present at very low frequencies which contribute very little to the heterozygosity.

(d) Effective number of alleles per locus. The effective numbers of alleles for all loci in each population are given in *Table 4*, with the geometric means in *Table 3*. Since both expected heterozygosity and effective number of alleles are functions of the square of observed allele frequen-

Table 3. — Percentage of heterozygous loci per individual, average number of alleles per locus and effective number of alleles per locus for each population.

	Population 10	21	29	32	37	39
Percent heterozygous loci/individual	17.08±10.2	17.29±10.2	24.58± 9.4	20.00± 9.2	22.29±11.3	21.04±10.5
Average number of alleles/locus	2.33± 0.65	2.17± 0.58	2.17± 0.72	2.25± 0.75	2.42± 0.79	2.50± 0.80
Effective number of alleles/locus	1.25	1.24	1.38	1.29	1.32	1.38

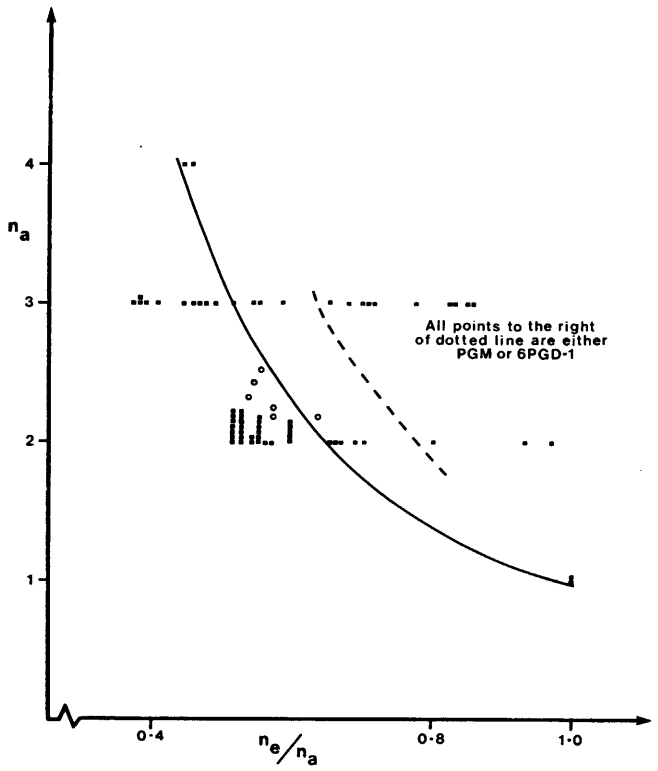

Figure 2. — Relationship between the actual number of alleles (n_a) and the ratio of effective number to actual number (n_e/n_a). Solid line is theoretically expected relationship. Square dots mark individual loci, circles the population average.

cies and both are maximised when allele frequencies are equal, there is obviously a very close relationship between the two.

Kirby and Halliday (1973) suggested that the ratio of the effective number of alleles to the actual number can be used to examine whether the isozyme loci involved are selectively neutral. By using the neutral allele model of Kimura and Crow (1964), the theoretical relationship between this ratio and the actual number of alleles can be calculated. *Figure 2* shows the theoretical curve, together with the plotted points from this study.

Table 5. — Results of the X^2 contingency tests and F-statistics for each locus. Totals are given for X^2, df and F_{ST}, means for F_{IS}, F_{IT} (direct) and F_{IT} (indirect).

	χ^2	df	$\bar{F}_{IS}$	F_{IT} (direct)	F_{IT} (indirect)	F_{ST}
ALD	13.20	10	-0.012	-0.024	0.004	0.016
AAT	12.63	10	0.045	0.088	0.060	0.016
FUM	2.27	5	-0.030	-0.016	-0.024	0.005
G6P	28.05*	15	0.046	0.061	0.060	0.014
GDH	22.50**	5	-0.013	-0.049	0.035	0.047
IDH	6.09	5	-0.024	-0.019	-0.011	0.013
MDH-3	14.31	10	0.046	0.110	0.068	0.023
PGM	5.86	10	0.045	0.054	0.050	0.005
6PGD-1	10.58	10	-0.025	-0.016	-0.017	0.008
6PGD-2	6.76	10	-0.005	0.018	0.003	0.007
PGI	11.92	10	-0.037	0.000	-0.016	0.020
TOTAL/AVERAGE	134.17	100	0.003	0.019	0.019	0.010

Significance levels: * 5%, ** 1%.

Genetic differentiation

(a) Contigency χ^2 tests. Results of the contingency χ^2 tests for differences in allele frequencies are given in *Table 5*. The values should be compared with a χ^2 distribution with (s-1) (k-1) degrees of freedom, where s is the number of populations and k is the total number of alleles recorded at the locus. When considered individually, there are only two loci at which allele frequencies differ significantly between populations. Differences in allele frequencies at G6P are significant at the 5 per cent level, whilst at GDH the differences are significant at the 1 per cent level. Chi-square values can be summed and compared with a distribution with degrees of freedom equal to the sum of the individual degrees of freedom (Workman and Niswander, 1970). The total χ^2 over all loci is 134.17, with 100 d.f., which is significant at the 1% level.

Table 6. — Distance among populations. Above the diagonal — genetic distance (D), with standard errors calculated by the method of Nei and Roychoudhury (1976). Below the diagonal — geographic distances (km).

	Population					
	10	21	29	32	37	39
10		.0023 ±.0011	.0038 ±.0018	.0031 ±.0024	.0021 ±.0011	.0036 ±.0020
21	26.2		.0084 ±.0049	.0041 ±.0022	.0033 ±.0024	.0042 ±.0025
29	52.1	26.2		.0023 ±.0009	.0025 ±.0020	.0025 ±.0015
32	39.3	14.3	13.7		.0017 ±.0009	.0023 ±.0013
37	24.4	1.4	27.8	15.7		.0008 ±.0003
39	13.4	12.8	38.7	25.9	11.2	

(b) F-statistics. The values of $\bar{F}_{IS}$, F_{IT} (direct and indirect) and F_{ST} are given in *Table 5* for individual loci, together with the mean values for all loci. The value of F_{ST} gives an indication of the relative amount of variation between populations compared with the total observed variation. Thus between population variation accounts for only 1 per cent of the total overall, whilst at individual loci this varies between 0.5 per cent at FUM to 4.7 per cent at GDH.

(c) Genetic distance. The matrix of genetic distances among populations is given in *Table 6*. As explained above, since monomorphic loci are probably under-represented, these estimates of genetic distance are probably upper limits. If the observed genetic distances between populations are due to isolation by distance, it would be expected that a strong relationship should exist between genetic and geographic distance (Nei 1975). The product-moment correlation between genetic and geographic distances is 0.064.

As there is no general method for specifying the degrees of freedom for pairwise comparisons of two matrices of intercorrelated populations (JORDE, 1980), the significance level of this correlation coefficient cannot be determined.

Discussion

An ideal measure of genetic diversity should produce a single value which accounts for the observed allele frequencies and number of alleles per locus. On these grounds, the expected heterozygosity at each locus and average heterozygosity over all loci have been criticized by BERGMANN and GREGORIUS (1979) since they do not take into account the number of alleles at a locus and, in the latter case, the number of loci used in the calculation. Nevertheless, expected and average heterozygosities are among the most frequently reported statistics for characterizing genetic diversity.

Many of these analyses assume a random sample of independently segregating loci (NEI, 1973, 1975). Several of the loci included in this study are either linked or in linkage disequilibrium (BOYLE and MORGENSTERN, 1985). However, as a large sample of common loci is desirable, especially for comparisons among species, it was considered that the bias involved in including some non-independent loci would be less serious than that resulting from a much smaller sample of independent loci.

In this study average heterozygosities ranged from 0.192 to 0.253 for the six populations of black spruce with a mean of 0.229 *(Table 2)*. These estimates are similar to those obtained for other conifer species. For example, in Douglas fir *(Pseudotsuga menziesii* (MIRB.) FRANCO), estimates have been reported of 0.155 (YEH and O'MALLEY, 1980) and 0.388 (YANG *et al.*, 1977). In Sitka spruce (*Picea sitchensis* (BONG.) CARR.), YEH and EL-KASSABY (1980) found an average heterozygosity of 0.147, whilst in eleven populations of Norway spruce *(P. abies* KARST.) in Sweden, LUNDKVIST and RUDIN (1977) obtained estimates ranging from 0.32 to 0.41. In lodgepole pine *(Pinus contorta* ssp. *latifolia* CRITCHFIELD) YEH and LAYTON (1979) reported a figure of 0.160 compared with an estimate of 0.146 in jack pine *(P. banksiana* LAMB.) (DANZMANN and BUCHERT, 1983). In a review of population genetic studies on a wide array of plant species, HAMRICK *et al.* (1979) list observed heterozygosities for conifer species ranging from zero (for *Pinus resinosa* AIT.) to 0.43 (for Norway spruce).

In most of these studies, considerable inter-locus heterogeneity has been noted in expected heterozygosities, though characteristically with a modal value approaching zero (YEH and O'MALLEY, 1980). This heterogeneity results in large standard errors of average heterozygosity, as indeed was the case in this study, with standard errors ranging from 36 per cent to 46 per cent of the estimates. For this reason (YEH (1981) recommends a large sample of loci. YEH and O'MALLEY (1980) considered that the much higher estimate of average heterozygosity obtained for Douglas fir by YANG *et al.* (1977) was in fact an artifact due to the much smaller number of loci used in the latter study. This theory would appear to be reasonable, since the high estimates reported for Norway spruce by LUNDKVIST and RUDIN (1977) were also based on a small number of loci (4). Twelve loci were assayed in this study, which were selected for the estimation of mating system parameters. For this reason, the sample was somewhat biased towards more variable loci (SHAW, 1982). The estimates reported here are therefore likely to be somewhat biased upwards.

As noted previously, the effective number of alleles per locus *(Table 4)* is highly correlated with expected heterozygosity *(Table 2)* since their response to allele frequencies is so similar. Therefore, the same comments apply as for expected heterozygosity. However, since the effective number of alleles does give an indication of the number of alleles at a locus, it should be included in any description of genetic diversity. As can be seen from *Figure 2*, there is generally quite good agreement between the observed values of n_a and n_e/n_a, implying that the isozyme loci involved may be selectively neutral. The values for population averages lie very close to the theoretical curve. All of the extreme points to the right of the curve are accounted for by two loci — PGM and 6PGD-1. There may therefore be some indication of non-neutrality at these loci. However, using some other approaches to examine selective neutrality, BOYLE (1985) found no evidence for selection at these two loci. Methods for the detection of selective neutrality among enzyme polymorphisms still require further refinement, as evidenced by the differing results from various approaches. KIRBY and HALLIDAY (1973) caution that any interpretation of the relationship between n_a and n_e/n_a must be tentative. Thus, the selective neutrality of any of these loci cannot be determined with certainty. The other two measures of genetic diversity used, namely the number of alleles per locus and the percentage of heterozygous loci per individual are comparable to those reported in other widely distributed, predominatly outcrossing conifers (e.g. YEH and O'MALLEY, 1980; YEH and EL-KASSABY, 1980). Standard errors for both measures were so large that no significant differences among the populations could be detected.

Only two loci (G6P and GDH) out of eleven showed any significant differentiation in allele frequencies among the six populations *(Table 5)*. In contrast, in ponderosa pine *(Pinus ponderosa* LAWS.), differentiation among six clusters within a single stand of about 2 ha was detectable at two out of seven loci (LINHART *et al.* 1981). GURIES and LEDIG (1977) studied four populations of pitch pine *(Pinus rigida* MILL.) in New Jersey which were separated by distances comparable to those in this study. They found significant differentiation at nine out of 15 loci, and considerable variation in χ^2 values, which they attributed to individual loci responding independently to different factors causing the variation. Therefore, they concluded that environmental heterogeneity was at least partly responsible. Similar results were obtained from a larger study of the same species involving eleven populations from North Carolina to Quebec (GURIES and LEDIG, 1981).

If the genetic differentiation among populations is random, the direct and indirect methods of estimating F_{IT} should give identical results (LINHART *et al.*, 1981). *Table 5* shows that at some loci for example G6P, PGM and 6PGD-1, the agreement is quite close but for others the two methods give very different estimates. The overall averages for the two methods are identical, suggesting that the observed differentiation may indeed be random.

Values of F_{ST} in most conifer species studies have tended to be less than 0.1 (summaries in GURIES and LEDIG 1981 and YEH 1981). This is interpreted as meaning that more than 90 per cent of the total variation observed is attributable to variation within populations and less than 10% to variation among populations. The results obtained for black spruce *(Table 6)* suggest that as much as 99 per cent of the variation is contained within populations with 1 per cent among populations, at least within distances of up to

52 km. These figures are larger than for most species, but comparable with balsam fir *(Abies balsamea* (L.) MILL.) in New Hampshire (NEALE, 1978). YEH (1981) has attributed low F_{ST} values to the ecological amplitude of the species investigated, their breeding system and to the lack of effective barriers to gene flow. The obvious implications of low F_{ST} values are that intensive efforts to sample a large number of populations within a region are not worthwhile; it would be far more efficient to concentrate on intra-population sampling.

Comparisons of genetic distance should be made with caution since the absolute values are affected by the number of monomorphic loci included (LINHART *et al.*, 1981). Also, as NEI (1972) points out, the method is not reliable for values near zero and the large standard errors calculated indicate that many of the distances are not significantly different from zero. Nevertheless, these results are clearly typical of many conifer species. LUNDKVIST (1979), using only polymorphic loci, obtained distances between four Swedish populations of Norway spruce ranging from 0.011 to 0.042. On the other hand, YEH and O'MALLEY (1980), with three monomorphic loci out of 21 assayed, found genetic distances among eleven populations of Douglas fir ranging from 0.0002 to 0.0082. As noted earlier, a strong correlation between genetic and geographic distance is expected if isolation by distance is responsible for the observed differentiation. The near-zero correlation coefficient found here suggests that some other factor may be responsible for the differentiation. Of course, due to the large standard errors of genetic distances, the correlation coefficient may merely be reflecting the fact that many of the distances are approximately zero.

Black spruce readily hybridizes with red spruce *(Picea rubens* SARG.), which forms a major component of the natural forest in New Brunswick. It might be expected therefore that differences in the degree of introgression of the two species would inflate the estimates of between-population variation. For this reason, although black spruce appears to be typical of many conifer species in terms of the amount and distribution of genetic diversity, it is perhaps surprising that larger values were not obtained for G_{ST} and D. On the other hand, the populations sampled have been selected as phenotypically superior stands and may therefore be purer black spruce than the average New Brunswick population.

Conclusions

Estimates of genetic diversity should take into account both the number of alleles at a locus and their relative frequency. Therefore both average heterozygosity and the average number, are required to characterise the diversity. Since there is a large amount of inter-locus heterogeneity, a large, random sample of loci is required in order to provide accurate estimates. Unfortunately, since five of the twelve loci assayed in this study were either linked to or formed gametic associations with other loci, there were only six independent, polymorphic loci and a single monomorphic locus. Standard errors of the various measures of diversity were therefore large. Despite this, black spruce appears to be typical of many conifer species in terms of levels of genetic diversity, especially when the upward bias due to the small number of loci is considered.

Similarly, although the various estimates of population differentiation suggest that somewhat more of the total variation in black spruce resides within populations than in some other species, the figures obtained are not atypical. An intense within-population sampling strategy is therefore recommended for black spruce tree improvement in New Brunswick.

The degree of differentiation among populations proved to be non-significant based on the contingency analysis. Thus the present rangewide provenance tests probably adequately sample the overall pattern of variation in the species. What differentiation does exist between populations is apparently not the result of isolation by distance since genetic and geographical distances seem to be only weakly related. However, this interpretation can only be tentative since the standard errors of genetic distance were so large.

Acknowledgements

We wish to thank RAYMOND LEBLANC for field and laboratory assistance, Drs. D. P. FOWLER and Y. S. PARK for advice on analytical procedures and manuscript review and especially to Dr. W. M. CHELIAK and an anonymous reviewer for many helpful suggestions for improvement of the manuscript.

References

BERGMANN, F. and GREGORIUS, H. R.: Comparison of the genetic diversities of various populations of Norway spruce *(Picea abies).* Proc. Conf. on Biochemical Genetics of Forest Trees, Umea, Sweden, 1978. p. 99—111 (1979). — BOYLE, T. J. B.: The mating system and population structure of black spruce in central New Brunswick and its influence on tree improvement stategy. Ph. D. dissertation. Univ. of New Brunswick, Frederiction, N. B. (1985). — BOYLE, T, J. B. and MORGENSTERN, E. K.: Inheritance and linkage relatioships of some isozymes of black spruce in New Brunswick. Can. J. For. Res. **15**: 992—996 (1985). — BROWN, A. H. D.: Isozymes, plant population genetic structure and genetic conservation. Theor. Appl. Genet. **52**: 145—157 (1978). — BROWN, A. H. D. and MORAN, G. F.: Isozymes and the genetic resources of forest trees. *In:* M. T. CONKLE (ed.). Proc. Symp. on Isozymes of N. American Forest Trees and Forest Insects, Berkeley, Ca. (1979). USDA Gen. Tech. Rep. PSW-48, p. 1—10 (1981). — CHAKRABORTY, R.: A note on Nei's measure of gene diversity in a substructured population. Humangenetik **21**: 85—88 (1974). — CROW, J. F. and KIMURA, M.: An introduction to population genetics theory. Harper and Rowe, New York (1970). — DANZMANN, R. G. and BUCHERT, G. P.: Isozyme variability in central Ontario jack pine. Proc. 28th Northeast. For. Tree Improv. Conf., 1982, p. 232—248 (1983). — GURIES, R. P. and LEDIG, F. T.: Analysis of population structure from allozyme frequencies. Proc. 14th South. Forest Tree Improv. Conf., Gainsville, Fla., p. 246—253 (1977). — GURIES, R. P. and LEDIG, F. T.: Genetic structure of populations and differentiation in forest trees. *In:* M. T. CONKLE (ed.). Proc. Symp. on Isozymes of N. American Forest Trees and Forest Insects, Berkeley, Ca., (1979). USDA Gen. Tech. Rep. PSW-48, p. 42—47 (1981). — HAMRICK, J. L., LINHART, Y. B. and MITTON, J. B.: Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. Ann. Rev. Ecol. Syst. **10**: 173—200 (1979). — JORDE, L. B.: The genetic structure of subdivided human populations. A review. *In:* J. H. MIELKE and M. H. CRAWFORD (eds.). Current developments in anthropological genetics, Vol. 1. Plenum Press, New York p. 135—208 (1980). — KIMURA, M. and CROW, J. F.: The number of alleles that can be maintained in a finite population. Genetics **49**: 725—738 (1964). — KIRBY, G. C. and HALLIDAY, R. B.: Another view of neutral alleles in natural populations. Nature **241**: 463—464 (1973). LEWONTIN, R.: The genetic basis of evolutionary change. Columbia Univ. Press, New York (1974). — LINHART, Y. B., MITTON, J. B., STURGEON, K. B. and DAVIS, M. M.: Genetic variation in space and time in a population of ponderosa pine. Heredity **46**: 407—426 (1981). — LUNDKVIST, K.: Allozyme frequency distributions in four Swedish populations of Norway spruce *(Picea abie* K.). II. Estimations of genetic variation within and among populations, genetic linkage and a mating system parameter. Hereditas **90**: 127—143 (1979). — LUNDKVIST, K. and RUDIN, D.: Genetic variation in eleven populations of *Picea abies* as determined by isozyme analysis. Hereditas **85**: 67—74 (1977). — MARSHALL, D. R. and BROWN, A. H. D.: Optimum sampling strategies in genetic conservation. *In:* O. H. FRANKEL and J. G. HAWKES (eds.) Crop genetic resources for today and tomorrow. Cambridge Univ. Press, Cambridge. p. 53—80 (1975). — MORGENSTERN, E. K.: Genetic variation in seedlings of *Picea mariana* (MILL.) B.S.P. II. Variation patterns. Silvae Genet. **18**: 161—167

(1969). — NEALE, D.: Allozyme studies in balsam fir. M.Sc. thesis, Univ. of New Hampshire, Durham, N. H. (1978). — NEI, M.: Genetic distance between populations. Amer. Nat. **106**: 283—292 (1972). — NEI, M.: Analysis of gene diversity in subdivided populations. Proc. Nat. Acad. Sci. No. **70**, Washington, D.C. p. 3321—3323 (1973). — NEI, M.: Molecular population genetics and evolution. *In:* A. NEUBERGER and E. L. TATUM (eds.). Frontiers of Biology Vol. **40**. Am. Elsevier, New York (1975). — NEI, M. and ROYCHOUDHURY, A. K.: Sampling variances of heterozygosity and genetic distance. Genetics **76**: 379—390 (1974). — RIEGER, R., MICHAELIS, A. and GREEN, M. M.: Glossary of genetics and cytogenetics. Springer-Verlag, Berlin (1976). — ROWE, J. S.: Forest regions of Canada. Can. For. Serv. Pub. 1300 (1972). — SHAW, C. R.: Electrophoretic variation in enzymes. Science **149**: 936—943 (1965). — SHAW, D. V.: The mating system and breeding structure of Douglas fir *(Pseudotsuga menziesii)*. Ph. D. dissertation. Univ. of California, Davis, Ca. (1982). — WORKMAN, P. L. and NISWANDER, J. D.: Population studies on southwestern Indian tribes. II. Local genetic differentiation in the Papago. Am. J. Hum. Genet. **22**: 24—49 (1970). — WRIGHT, S.: Interpretation of population structure by F-statistics with special regard to systems of mating. Evolution **19**: 395—420 (1965). — YANG, J. C., CHING, T. M. and CHING, K. K.: Isoenzyme variation of coastal Douglas fir. I. A study of geographic variation in three enzyme systems. Silvae Genet. **26**: 10—18 (1977). — YEH, F. C.: Analysis of gene diversity in some species of conifers. *In:* M. T. CONKLE (ed.). Proc. Symp. on Isozymes of N. American Forest Trees and Forest Insects, Berkeley, Ca., 1979. USDA Gen. Tech. Rep. PSW-48. p. 48—52 (1981). — YEH, F. C. and EL-KASSABY, Y. A.: Enzyme variation in natural populations of Sitka spruce *(Picea sitchensis)*. I. Genetic variation patterns among trees from 10 IUFRO provenances. Can. J. For. Res. **10**: 415—422 (1980). — YEH, F. C. and LAYTON, C.: The organisation of genetic variability in central and marginal populations of lodgepole pine *(Pinus contorta* ssp. *latifolia)*. Can. J. Genet. Cytol. **21**: 487—503 (1979). — YEH, F. C. and O'MALLEY, D. M.: Enzyme variations in natural populations of Douglas fir *(Pseudotsuga menziesii* (MIRB.) FRANCO) from British Columbia. I. Genetic variation patterns in coastal populations. Silvae Genet. **29**: 83—92 (1980).

Identifizierung von Hybridlärchen mit Hilfe chemischer Merkmale

Von G. WEISSMANN*) und S. RECK**)

Bundesforschungsanstalt für Forst- und Holzwirtschaft Hamburg, Institut für Holzchemie und chemische Technologie des Holzes*) und Institut für Forstgenetik und Forstpflanzenzüchtung**)

(Eingegangen 8. Dezember 1986)

Zusammenfassung

Eine Unterscheidung zwischen Hybriden und artreinen Lärchen ist in der Forstwirtschaft und in der Züchtung außerordentlich wichtig, nach dem äußeren Erscheinungsbild jedoch nicht immer möglich. Ziel der Untersuchungen war es, chemische Merkmale auf ihre Eignung zur Identifizierung von Hybridlärchen zu prüfen. Dazu wurden Nadelöle und Harzbalsame aus definierten Beständen europäischer *(Larix decidua)* und japanischer Lärche (*L. kaempferi)* sowie Hybridlärchen (*L. decidua* × *L. kaempferi* und *L. kaempferi* × *L. decidua*) gewonnen und mit Hilfe von GC und GC/MS untersucht. Die neutralen, nicht flüchtigen Bestandteile der Harzbalsame enthalten hauptsächlich bicyclische Diterpenalkohole neben Harzaldehyden. Ein monocyclischer Diterpenalkohol (Thunbergol) tritt nur im Balsam von *L. kaempferi* und in den Hybriden auf, ist aber im Balsam der europäischen Lärche nicht nachzuweisen und damit für eine Unterscheidung zwischen reinen europäischen Lärchen und Hybridlärchen geeignet.

Summary

In forestry and forest tree breeding a differentiation between hybrids and pure larches is very important but according to the exterior appearance not always possible. The aim of this investigation was to test chemical markers for the identification of larch hybrides. Needle oils and oleoresins from authentic stands of European (*Larix decidua*) and Japanese larch (*L. kaempferi*) as well as hybrids (*L. decidua* × *L. kaempferi, L. kaempferi* × *L. decidua*) were isolated and analyzed by GLC and GLC/MS. The neutral non-volatile fractions of oleoresins contain mainly bicyclic diterpene alcohols aside from resin aldehydes. A monocyclic diterpene alcohol (thunbergol) is present only in oleoresins from *L. kaempferi* and hybrids but could not be identified in *L. decidua* oleoresin and is therefore suitable for a differentiation between pure European larch and hybrids.

Anschrift: *) Leuschnerstraße 91, D-2050 Hamburg 80
**) Sieker Landstraße 2, D-2070 Großhansdorf 2

Key words: Forest tree breeding, *L. decidua, L. kaempferi*, Identification of hybrids, Composition of needle oils, Separation and composition of oleoresins, Gas-liquid chromatography, Mass-spectrometry.

Einleitung

Hybriden zwischen europäischen und japanischen Lärchen (*Larix eurolepis* HENRY) können in der Natur nach spontaner Kreuzbefruchtung zwischen beiden Arten entstehen. Da beide Lärchenarten in beiden Richtungen miteinander leicht kreuzbar sind, ist in allen Gebieten, in denen europäische und japanische Lärchen gemeinsam vorkommen, mit dem Auftreten zufällig entstandener Arthybriden zu rechnen (LANGNER 1951/52). In der Forstpflanzenzüchtung werden Hybridlärchen durch künstliche Zwischenartkreuzungen erzeugt oder aus „Hybridsamenplantagen" nach freier Befruchtung gewonnen mit dem Ziel, durch die Hybridisierung eine erhöhte Wuchsleistung auszulösen („Heterosis) oder erwünschte arttypische Eigenschaften miteinander zu kombinieren. In allen Fällen sind Kriterien für eine sichere Identifizierung von Hybriden erforderlich. Für die Beurteilung von Lärchenbeständen auf ihre Eignung als Saatguterntebestände ist von Bedeutung, ob und mit welchem Anteil Hybriden enthalten sind. In der Hybridlärchenzüchtung ist die Identifizierung von Hybriden in den Ausgangspopulationen und in den gezüchteten Nachkommenschaften wesentlicher Teil der Züchtungskontrolle.

Da eine sichere Unterscheidung zwischen Hybriden und artreinen Lärchen nach dem äußeren Erscheinungsbild nicht immer möglich ist, sollten chemische Merkmale auf ihre Eignung zur Identifizierung von Hybridlärchen geprüft werden. Als relativ leicht zugängliches Untersuchungsmaterial kommen die Nadelöle und die Harzbalsame in Betracht, die aus definierten Beständen europäischer (*Larix decidua*) und japanischer Lärchen (*L. kaempferi*) sowie der Hybridlärchen (*L. decidua* × *L. kaempferi* und *L. kaempferi* × *L. decidua*) gewonnen wurden.