CALEDONIAN SCOTS PINE: ORIGINS AND GENETIC STRUCTURE

By BOHUN B. KINLOCH[1,*], R. D. WESTFALL[2]
AND G. I. FORREST[1]

[1] *Forestry Commission, Northern Research Station, Roslin, Midlothian, Scotland
and* [2] *Pacific Southwest Forest and Range Experiment Station, Box 245,
Berkeley, CA 94701, USA*

SUMMARY

Monoterpene and isozyme loci, used as markers to study the genetic structure of Scots pine (*Pinus sylvestris* L.) native to Scotland, showed that the endemic populations are not genetically impoverished, in spite of severe contraction in range and numbers as a result of both natural and anthropogenic causes. On the contrary, variability in the relict populations is almost the highest of any plant species studied, with average heterozygosities of 0·33 for monoterpenes (five loci) and 0·30 for isozymes (16 loci). The overwhelming proportion of this variability (> 95 %) was within populations, even though significant differences in gene frequencies of many individual loci existed among populations. Multiple-locus comparison of gene frequencies among populations, resolved by canonical variate analysis, showed no coherent geographic pattern of differences from population to population or region to region, with one major exception: certain populations in northwestern Scotland (Wester Ross) were distinct from all others and each other. The pattern of variability of the biochemical markers was consistent with that of metrical and physiological traits reported in the literature. These traits, in turn, show relatively little genetic affinity between contemporary Scottish and continental European populations. The genetic evidence, together with the anomalous distribution of pine pollen in the British Isles during the Holocene, suggests that the Caledonian race of Scots pine originated endemically from more than one refugium after the last glaciation.

Key words: Isozymes, monoterpene inheritance, multivariate analysis, *Pinus sylvestris*, Pleistocene.

INTRODUCTION

At the western extremity of the distribution of one of the most widespread and variable forest trees in the world, the Caledonian race of Scots pine (*Pinus sylvestris* L.) survives as relict populations scattered throughout the highlands of Scotland. These populations occupy sites of great climatic and edaphic diversity: from mild, wet, maritime climates near sea level in the west to continental, subalpine conditions at 615 m in the Cairngorm Mountains; annual precipitation ranges from 890 to 2800 mm; and soils range from deep peats to coarse sands and gravels (Steven & Carlisle, 1959). Stand density varies from closed canopies with several age classes to scattered, open woodlands of veteran trees. Few populations cover more than 1000 ha, and one is a grove of less than 30 trees.

Formerly, Scots pine had a much wider distribution in the British Isles, including Ireland, England and Wales. It was a main component, with birch, of

---

*Present address: Pacific Southwest Forest and Range Experiment Station, Box 245, Berkeley, CA 94701, USA.

the early forest that colonized the relatively bare, base-rich soils left by the recession of the glaciers and periglacial processes at the opening of Holocene (postglacial) time 10000 years ago. In the Scottish Highlands, population density of pine increased gradually at first, rose abruptly at around 8000 B.P. and reached a maximum from about 6500 to 5000 B.P. (O'Sullivan, 1977). But by this time, its decline was already evident in England, Wales and Ireland, where progressive climatic and soil amelioration and successive migrations of additional tree taxa made it unable to compete with elements of the mixed deciduous forest (oak, elm, hazel, alder, ash, linden) on favourable sites (Bennett, 1984). During the mid-Holocene, starting about 6000 B.P., the climate may have reverted to cooler, moister, more 'oceanic' conditions. Blanket-bog formation became widespread, especially from about 4000 B.P. in the northwest Highlands. This encroachment obliterated large areas of pine forest, fragmenting them into relict communities (H. H. Birks, 1975; H. J. B. Birks, 1977).

Up to this point, the early Holocene history of Scots pine recapitulates previous interglacial periods (Godwin, 1975a; West, 1980), when pine and birch were pre-eminent during pretemperate and again in posttemperate stages along with spruce and fir. In the interval, it was largely excluded from all but the most marginal sites by competition from broad-leaved species (Godwin, 1975a; Bennett, 1984). But in the Holocene, roughly coincident with the climatic deterioration that began about 6000 B.P., man began influencing the shape of the landscape, and from the first forest clearings by Neolithic farmers, the relative influence of man and climate on forests became increasingly confounded. However, Carlisle (1977) considered the earliest impact (from grazing, tilling the land and accidental fires) fell most heavily on the lowland, mixed deciduous forests, with only local and minimal consequences for the pine forests of the Highlands. In fact, until almost the end of the sixteenth century, the Highlands were largely unknown and were left alone '... by reason of the savageness of the inhabitants thereabout...' (Acts of Parliament; in Steven & Carlisle, 1959, p. 57). But with the exhaustion of lowland forests, serious exploitation of the pine woods began, culminating in the huge demands for wood imposed by the Napoleonic wars. At this time, the larger woods in Deeside, Speyside and the Black Wood of Rannoch were most severely affected. Had extraction from the more inaccessible glens not been so difficult, it is possible that the native woodlands would have disappeared altogether. However, felling slackened for about a century before it finally resurged during the two world wars of this century.

While timber was being extracted, regeneration within and at the margins of the forest was progressively declining. Again, climatic and anthropogenic influences are inseparable, and primary causes are controversial. But there is no question that increased grazing pressure and fire were important. Numbers of sheep and cattle increased greatly over recent centuries, as did deer after the extinction of wolves. Sometimes, both domestic and wild animals were enclosed in the woodlands. Occasional fires that had formerly promoted regeneration within the forest by reducing competing understorey vegetation and exposing mineral soil were suppressed, while deliberate and frequent burning of moors for sheep and grouse prevented the expansion of the forest along its margins.

What remains today are a few dozen small woodlands covering less than 11000 ha, of which less than a fifth is adequately stocked (Goodier & Bunce, 1977). In short, Caledonian Scots pine has suffered and survived as many disasters as any forest tree species still existing in at least a semi-natural state, including regional

extinction from large parts of its former range (Ireland, Wales and probably England), local fragmentation into isolated populations (accompanied by great reduction in numbers) and intensive artificial (and presumably natural) selection. What are the genetic consequences of this history?

This question is of basic importance, because continued exploitation of other commercial forest tree species will have similar demographic effects. How to protect the genetic integrity of populations and minimize genetic erosion are problems of increasing urgency. Solutions depend first of all on knowledge of the genetic structure of populations, which can be gained by selected case studies. We feel that Caledonian Scots pine is particularly appropriate, because its history has included so many different kinds of catastrophe, both natural and man-made. Practical motives are equally compelling. Scots pine remains the only indigenous softwood resource, even though its importance in British forestry has declined in favour of exotic species from western North America during the last century. At the same time, the historical, scientific and recreational importance of the native woods has increased immensely. The 'Caledonia Silva' was identified on Ptolemy's map, from descriptions by Roman invaders and by Greek explorers even earlier (Steven & Carlisle, 1959). The present remnants are refuges for native fauna and represent the last semblance of an indigenous forest ecosystem in the British Isles. Their aesthetic and emotional values cannot be quantified, but were epitomized by Steven & Carlisle (1959) in words approaching the mystical: 'To stand in them is to feel the past'.
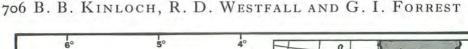
Our primary objective was to characterize the amount and distribution of genetic variability in the native pines of Scotland. We used allozymes and monoterpenes as markers because of their relative efficiency. We were also interested in comparing the variability observed in these two sets of traits, representing primary and secondary gene products, respectively, with the variability in physiological and morphological traits described in the literature. The unusual pattern of variation that emerged led us to re-examine traditional concepts of the genetic origins of Caledonian Scots pine.

MATERIALS AND METHODS

Population sampling

The natural distribution of Scots pine in Scotland was documented by Steven & Carlisle (1959), who described in detail about 36 authentically native populations, assigned to eight geographic regional groups. Our sampling was confined to these populations, with one exception. Strath Oykell, in the Northern group, probably originated as a plantation in the eighteenth century (Steven & Carlisle, 1959, p. 206), although the seed almost certainly derived from nearby native sources. We did not sample all the populations, but we further subdivided several (Loch Maree, Coulin, Glen Affric, Abernethy, Glenmore and Glentanar) to obtain estimates of microgeographic variation, making our total population sample 40 for monoterpenes and 14 for isozymes (Fig. 1, Appendix 1).

Most of the native woodlands were sampled for cortex monoterpenes over a two-year period (Forrest, 1980). Numbers of trees ranged from 29 at Glen Avon, where every tree was sampled, to over 200 in several of the larger populations (Appendix 3). A single branch, usually from the lower crown, was taken from each tree. The age of the tree was not considered, although most ranged from mature to very old.
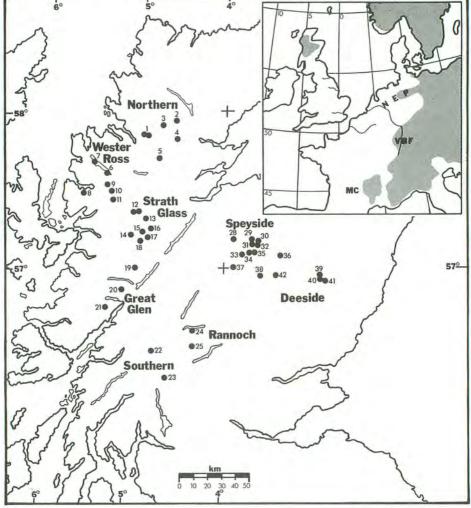
Fig. 1. Locations of Caledonian Scots pine (*Pinus sylvestris*) populations sampled (see Appendix 1 for details). Inset: natural distribution of Scots pines in Western Europe (after Critchfield & Little, 1966). NEP, Northern European Plain; V, Vosges Mountains; BF, Black Forest; MC, Massif Central.

The 14 populations sampled for isozyme analyses represented a range of sites and geographical groups; one to several cones were collected from 24 to 44 trees in each (Appendix 1). Seed of Balochbuie and Glentanar origins was obtained from grafts of phenotypically selected trees in the Forestry Commission's clone bank at Newton.

Laboratory analyses

Details of the monoterpene analysis are described in Forrest (1980). Briefly, resin was extracted by gentle suction from the cortical resin canals of the previous year's internodes into glass capillary tubes, which were stored temporarily in small glass vials at $-20\,^{\circ}\mathrm{C}$. Resin was dissolved in n-pentane, and about $0.5\,\mu\mathrm{l}$ of the solution was injected into a Pye Unicam Series 104 gas chromatograph with a

heated, dual-flame ionization detector. The column was 2·13 m glass, packed with 10% polyethylene glycol, 20 M, on 100 to 120 mean Diatomite C–AW. Oven temperature was 130 °C isothermal.

For isozyme analysis, seeds were germinated on moist filter paper until the radical emerged. The embryo and seed coat were then removed, and the haploid female gametophyte tissue was prepared for electrophoresis. Usually, from four to six gametophytes per tree were used, although larger numbers were used when necessary to confirm Mendelian segregation of allozyme bands. Conventional laboratory procedures for starch gel electrophoresis were according to Conkle *et al.* (1982). We assayed 17 different enzyme systems, of which nine (comprising 16 loci) were used in the analysis; the rest were unusable because they resolved poorly, or because some populations were not represented (Appendix 2). Loci were labelled sequentially from the anode, as were different allozymes within loci. For each tree, allozyme genotypes were recorded for each locus, except when stains were unreadable. Thus, different loci had varying sample sizes for any given population.

Analysis of the data

Gene diversity statistics (Nei, 1975), based on average heterozygosity, were computed for allozyme and monoterpene gene frequencies to assess the total amount of genetic variation (Ht) and the relative distribution of this variation within (Hs) and among (Gst) populations.

The geographic pattern of variation was analyzed by Nei's (1975) standard genetic distance,

$$D = -\log_e I,$$

where the genetic identity

$$I = \frac{\Sigma x_i y_i}{\Sigma x_i^2 y_i^2},$$

and x_i/y_i are the frequencies of the ith allele in population x and y, respectively. Two populations are identical at a locus ($I = 1$, $D = 0$) if their allele frequencies are equal; if they share no common alleles, $I = 0$ (and D is infinite). D values between populations represent averages over all the loci, and they were used to cluster stands by the unweighted pair-group method (UPGMA, Sneath & Sokal, 1973; equivalent to group averaging, Dunn & Everitt, 1982).

For greater resolution of differences among populations, we also used a canonical variate analysis (CVA) on both allozyme and monoterpene gene frequencies according to standard protocols (Cooley & Lohnes, 1971; Dunn & Everitt, 1982). Briefly, this procedure identifies a series of vectors, i.e. linear combinations of variables (in our case, allele frequencies), such that differences among groups are maximized, and variation within groups are minimized. Although individual gene loci are discrete variables and not normally distributed, the distribution of canonical scores of multiple locus vectors rapidly approaches multivariate normality as the number of loci increase (Smouse, Spielman & Park, 1982). The first vector accounts for the greatest proportion of the variance and may show substantial differences among groups, even if none of the alleles do individually. The next vector (the scores of which are uncorrelated with the first) is aligned in the direction of the next greatest variation, and so on for successive vectors. The number of vectors extracted are equal to the lesser of the number of variables in the analysis or the number of groups minus one (that is, the among-group degrees of freedom). The position of each group along any vector is determined by

canonical scores. Since there are far more alleles than populations in the allozyme data, alleles must be selected which maximally differentiate the groups. Variable selection was accomplished by the Statistical Analysis System (SAS) procedure STEPDISC, while the canonical analysis was conducted by CANDISC (SAS Institute, 1982).

The number and kinds of groups critically affect the outcome of the analysis. The number must be specified, but populations either can be clustered phenetically or assigned to some preconceived classification. We examined differences among groups using both criteria. For the first, four groups were designated on the basis of phenetic data, using a K-means clustering procedure (FASTCLUS, SAS Institute, 1982). For the second, populations were assigned to eight regional geographic groups according to Steven & Carlisle (1959).

RESULTS

Monoterpene inheritance

The average percentages for the several monoterpenes were reported previously for each of the populations (Forrest, 1980), but the range in values for many monoterpenes among individual trees within populations was great. When the frequencies of trees in the populations are arrayed by percentile classes, distinct bimodal and trimodal distributions are evident for several of the monoterpenes. Composite histograms pooling all the populations are given in Figure 2 for the six monoterpenes that comprised 95% of the total detected: [α-pinene (10%), β-pinene (13%), Δ^3carene (45%), myrcene (19%), limonene (4%) and β-phellandrene (4%)].

Inheritance of individual monoterpenes was inferred from the modal patterns expressed and from the close correspondence of these patterns with those of other pine species for which monoterpene inheritance has been determined. Squillace, Wells & Rockwood (1980) found similar bimodal peaks and ranges in cortex resin of *Pinus taeda* L. progenies for α-pinene, myrcene, limonene and β-phellandrene. Segregation ratios of full-sib and selfed families of *P. taeda* confirmed that inheritance was controlled at a single dimorphic locus in each case, with the allele for high monoterpene dominant to low. The same conclusions were drawn for β-pinene and myrcene in *Pinus pinaster* Ait. (Baradat *et al.*, 1972, 1975) and 3-carene in *Pinus monticola* Dougl. (Hanover, 1966) and Scots pine (Hiltunen *et al.*, 1975).

Our data show clear bimodal distributions for β-pinene, myrcene, limonene and β-phellandrene, with peaks at relatively high and very low percentages of these monoterpenes. We interpret each of these distributions to reflect genetic control at a single locus with two alleles, with high monoterpene dominant to low. Frequencies of the allele for low monoterpene were derived by taking the square root of the proportional frequency of trees with low phenotypes; the frequency of the alternate allele (high monoterpene) was obtained by subtracting this value from 1.

The trimodal distribution of 3-carene suggests additivity of the alleles at this locus. Since, in this case, the phenotypic frequencies of trees with low (< 10%), intermediate (10 to 60%) and high (> 60%) amounts of 3-carene are equivalent to genotypic frequencies of homozygous low, heterozygous and homozygous high, respectively, the model can be tested for conformity to the Hardy–Weinberg formula for genes at equilibrium frequencies: $q^2 + 2q(1-q) + (1-q)^2$, where q and
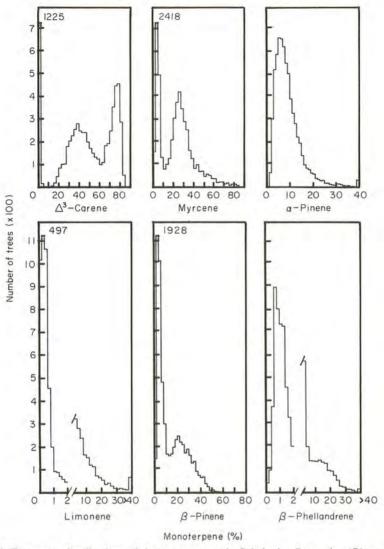
Fig. 2. Frequency distributions of six monoterpenes in Caledonian Scots pine (*Pinus sylvestris*).

$1-q$ are alternate alleles for low and high 3-carene. In each population, q was derived by dividing the number of trees in the low 3-carene class plus half of those in the intermediate class by the total number of trees. Results of chi-square tests showed that observed genotypic frequencies did not deviate significantly from expectation in any of the 40 populations.

We have not attempted genetic interpretation of the remaining six monoterpenes (Forrest, 1980); most are either at a very low concentration or unimodally distributed, or both. Terpinolene is trimodally distributed, but is highly correlated with 3-carene ($r = 0.98$) and is thus statistically redundant. Alpha-pinene, although a major component and highly variable among individual trees, is essentially unimodally distributed (Fig. 2) and cannot be explained by a simple genetic model. Similar patterns of variation in other species preclude genetic interpretation

of this monoterpene (Squillace, 1971; Hiltunen et al., 1975; Squillace et al., 1980). Strauss & Critchfield (1982) suggested that α-pinene was indirectly controlled by β-pinene, to which it responded in a complementary fashion.

Apportionment of genetic variation

Allele frequencies for the five monoterpene and 16 allozyme loci analyzed are given in the Appendices 3 and 4. All the loci are polymorphic, although variants of PGM2 and IDH are rare. The average number of allozyme alleles over all the

Table 1. *Estimates of gene diversity parameters in Caledonian Scots pine* (Pinus sylvestris)

Locus	Ht[1]	Hs	Gst	$\overline{D}^2$	Significance level[3]
		Allozymes (14 populations)			
ACP2	0·330	0·323	0·022	0·009	*
ADH2	0·492	0·462	0·059	0·063	**
GDH	0·329	0·319	0·029	0·013	+
GOT1	0·025	0·024	0·040	0·001	*
GOT2	0·613	0·598	0·024	0·037	+
GOT3	0·359	0·354	0·015	0·007	n.s.
IDH	0·011	0·011	0·036	0·000	**
LAP1	0·505	0·496	0·018	0·013	+
LAP2	0·150	0·146	0·026	0·003	*
MDH1	0·059	0·056	0·047	0·002	***
MDH2	0·078	0·076	0·021	0·001	n.s.
MDH3	0·412	0·404	0·021	0·049	n.s.
MDH4	0·606	0·595	0·019	0·028	n.s.
PGM2	0·021	0·030	0·017	0·000	n.s.
6PGD1	0·530	0·518	0·022	0·021	**
6PGD2	0·457	0·444	0·029	0·022	*
Mean	0·311	0·303	0·028		
		Monoterpenes (40 populations)			
3-Car.	0·482	0·454	0·059	0·101	**
Myr.	0·421	0·395	0·062	0·100	***
β-Pin.	0·362	0·353	0·025	0·012	***
Lim.	0·257	0·248	0·037	0·007	***
β-Phel.	0·208	0·204	0·023	0·004	***
Mean	0·346	0·330	0·045		

[1] Ht, total gene diversity; Hs and Gst, within and among population diversity, respectively (Nei, 1975).
[2] Genetic distance (Nei, 1975) averaged over all pairs of populations for the locus indicated.
[3] Significance of heterogeneity chi-square value, testing the null hypothesis that population frequencies for alleles at the locus indicated are equal; $+$, *, **, *** indicate differences at $P \leqslant 0·1$, $0·05$, $0·01$ and $0·001$, respectively; n.s., not significant. Computational formula from Workman & Niswander (1970).

loci and populations is 2·6, ranging from 2·2 at Mar to 2·9 at nearby Glentanar. Two loci (MDH4 and 6PGD1) have up to six alleles in some populations. More than two alleles for monoterpenes (if they exist) could not be detected by our method of analysis.

Variability is high in all populations, with average expected heterozygosities (Ht) for allozyme loci ranging from 27 % (Cougie) to 33 % (Glentanar), and for monoterpene loci from 24 % (Ardgour) to 42 % (Coulin-A). Standard errors were usually within 16 to 20 % of the estimates in both cases. Average heterozygosity estimated directly from individual allozyme genotypes (*H*) is 30 %, nearly the same

as that expected (31 %) under Hardy–Weinberg conditions of random mating and equilibrium gene frequencies. (The Hardy–Weinberg estimate of Ht for monoterpene loci in the 40-population sample is 33 %.) Allozyme genotype frequencies showed no indication of heterozygote excess or deficiency. Only 11 out of 224 population × locus combinations (i.e. 5 %) differed significantly from Hardy–Weinberg expectations by chi-square analysis, using the pooling method for rare alleles (Swofford & Selander, 1981a, b). From this, we conclude that allozyme genotypes are in equilibrium, and, by extrapolation, that monoterpene genotypes are also in equilibrium. Neither the sparse population at Glen Falloch ($H = 31\%$, allozymes; Ht = 34 %, monoterpenes), nor the grove of 29 trees at Glen Avon (Ht = 35 %, monoterpenes only) showed a reduction in variability or heterozygote deficiency.

The distribution of genetic variability was analyzed by Nei's (1975) gene diversity parameters, wherein total gene diversity (Ht) is subdivided into components within (Hs) and among (Dst) populations, and the proportion of variation among populations relative to the total variability is expressed by the coefficient of gene differentiation (Gst). Estimates of Ht, Hs and Gst are given in Table 1 for each of the 16 allozyme and five monoterpene loci. The overall diversity is large and varies greatly among loci, from 0·01 (IDH) to 0·61 (MDH4). The slightly higher values of both Ht and Hs for monoterpenes than for allozyme loci may reflect bias, since only monoterpenes with large ranges or bimodal distributions were analyzed. Although the mean Gst value of monoterpenes (0·045) is larger than that for allozymes (0·028), the overwhelming proportion of variability for both sets of loci is within populations. Depending on the locus, only 2 to 6 % of the variability is attributable to differences among populations.

Population differentiation

   Despite small Gst values, there are significant differences in allele frequencies among populations for half of the allozyme and all of the monoterpene loci (Table 1). Particularly conspicuous are differences in frequencies of the common alleles for 3-carene, myrcene and ADH2 (Appendices 3 and 4) between Shieldaig (and, to a lesser extent, other Wester Ross populations) and all the other populations.

   Genetic distances (*D*; Nei, 1975) between populations, computed for all possible pairs, are presented for each population (and regional group) averaged over all the others in Appendices 3 (for monoterpene loci) and 4 (for allozyme loci). The mean genetic distance for all the allozyme loci over all the populations is 0·013, and that for monoterpenes is 0·024, reflecting the generally low values between individual populations in each case. Mean *D* values of populations are relatively low and homogeneous, both within and among regional groups, with the exception of the Wester Ross group. For monoterpenes, these six populations not only have a higher mean value (0·055) but are also much more heterogeneous than any other group, with *D* values ranging from 0·016 (Achnashellach) to 0·144 (Shieldaig). Distances between populations within this group were generally much greater than those within any other group. Shieldaig also has the highest *D* value for allozyme loci (0·021), but Loch Maree, the only other Wester Ross population sampled, is average (0·013).

   Clustering of populations by the unweighted pair-group method (UPGMA) shows these relationships more clearly. In the 40-population sample for monoterpenes alone (Fig. 3), Loch Maree and and Coulin-A cluster together and are relatively close to Shieldaig, but all three of these populations are distinct from

Fig. 3. Clustering of 40 populations of Caledonian Scots pine (*Pinus sylvestris*) based on Nei's (1975) standard genetic distance. Population numbers and regional groups as in Appendix 1.

the rest, including their near neighbours, the island population in Loch Maree (Eilean Ruairidh Mor) and Coulin-B. For allozymes alone, Shieldaig (but not Loch Maree) separates into a cluster by itself, while in an analysis combining both allozymes and monoterpenes (13 populations), Shieldaig and Loch Maree cluster together and are distinctly different from all the other populations. In fact, except for the anomalous positions of these few Wester Ross populations, no consistent geographically coherent or meaningful pattern emerges from any of the cluster analyses.

Nei's (1975) genetic distance has been the most widely used parameter of population differentiation for forest trees (e.g. Guries & Ledig, 1982; Wheeler & Guries, 1982; Hiebert & Hamrick, 1983) and its use here is appropriate for comparison with other studies. A problem with this parameter is its relative

Table 2. *Canonical variate analysis of allozyme and monoterpene loci in Caledonian Scots pine* (Pinus sylvestris) : *coefficients of multiple correlation and percentage of the trace associated with the first two canonical vectors*

Clustering criterion	Trace[1]	Canonical vector	R^2	Percentage of trace
Allozymes (14 populations)				
Regional groups (4)[3]	20·9	1	0·969*	75
		2	0·899 n.s.	20
Phenetic clusters (4)	313·4	1	0·997***	53
		2	0·996***	40
Monoterpenes (40 populations)				
Regional groups (8)	3·1	1	0·771***	48
		2	0·687*	29
Phenetic clusters (4)	12·4	1	0·958***	91
		2	0·722***	9

[1] Trace $[\mathrm{Tr}(BW^{-1})]$ is the sum of the diagonal elements of the matrix product of the between-group sums of squares and cross-products matrix (B) and the inverse of the within-group sums of squares and cross-products matrix (W^{-1}). It is the multivariate equivalent of the ratio of the between-group sums of squares to the within-group sums of squares in the general linear model.

[2] The correlation between a linear combination of the allelic variables and the general linear model (the multivariate equivalent of R in analysis of variance).

[3] To avoid regional groups represented by single populations, Amat was assigned to the Strath Glass, and Glen Falloch to the Rannoch group (see Appendix 1).

*, ** and ***, significant at $P \leqslant 0.05$, 0.01 and 0.001, respectively; n.s., not significant.

insensitivity. Nei (1981) emphasized that the gene diversity statistics developed by him (and related to D) are more appropriate for estimating the distribution of genetic variation (i.e. the relative magnitude of variability between *vs* within populations) than for classifying populations. Since D is an average, effects of individual loci with large interpopulational differences are diluted, and cumulative effects of multiple loci, each with small or modest contributions, are not expressed. Furthermore, there is no adequate statistical test of D.

These limitations are largely removed or avoided by CVA. Here, differences among groups result from additive effects of multiple loci, where relatively small differences in allele frequencies over many loci have equal or greater resolving power than large differences in a few loci have with Nei's genetic distance. In contrast to the latter, distances among groups measured by CVA are Euclidean, and their significance can be tested.

The results of CVAs for allozyme and monoterpene loci by the two grouping criteria used — preassigned geographic regional groups and phenetic clustering — are shown in Tables 2 and 3 and Figure 4. For regional groups, canonical correlation coefficients of the first vector are significant for both monoterpene and allozyme loci (Table 2). The first vector accounts for nearly half, and the first two vectors for nearly all of the weighted, among-group variation in most analyses. Canonical scores of individual populations for the first vector are listed in Appendix 3 for monoterpenes and in Appendix 4 for allozymes. The monoterpene and allozyme loci most highly correlated with the first canonical vector for each grouping criterion (Table 3) are also those with the highest genetic diversity statistics (Table 1).

Table 3. *Canonical variate analysis of allozyme and monoterpene loci in Caledonian Scots pine* (Pinus sylvestris) : *correlations between allele frequencies and the first canonical vector*

Allele	Total canonical structure[1]
Allozymes (14 populations)	
Regional groups (4)[2]	
ADH22	0·49
MDH43	0·57
MDH41	−0·26
MDH32	0·51
MDH44	0·21
6PGD11	−0·42
PGM	−0·27
Phenetic clusters (4)	
ADH21	0·45
ADH22	−0·45
GOT21	−0·71
LAP11	0·02
MDH11	0·45
MDH46	0·63
PGM	0·08
MDH33	0·21
Monoterpenes (40 populations)	
Regional groups (8)	
Myr.	−0·94
β-Pin.	−0·28
3-Car.	0·86
Lim.	0·24
β-Phel.	−0·56
Phenetic clusters (4)	
3-Car.	0·97
β-Pin.	−0·44
Lim.	−0·26
Myr.	−0·95
β-Phel.	−0·70

[1] The correlation between population frequencies of the allele indicated and the first canonical vector.
[2] To avoid regional groups represented by single populations, Amat was assigned to the Strath Glass and Glen Falloch to the Rannoch group (see Appendix 1).

The isolation of Shieldaig into a solitary phenetic cluster far removed from the others is the most conspicuous result, with either monoterpene or allozyme loci. For monoterpene loci [Fig. 4(a)], two other Wester Ross populations, Coulin-A and Loch Maree, also form a separate cluster, closer to Shieldaig than the rest but nevertheless distinct from it and from their nearly contiguous neighbours in Wester Ross, Coulin-B and Eilean Ruairidh Mor. The remaining 37 populations split into two large clusters of approximately equal size relatively close to each other. Neither these two clusters, nor the three phenetic clusters for allozymes [other than Shieldaig, Fig. 4(c)] fall into any coherent geographic pattern. Results of *K*-means clustering [Fig. 4(a)] closely parallel those from UPGMA clustering based on genetic distances (Fig. 3). Grouping by geographic regions [Fig. 4(b), (d)] also fails to reveal differences among the groups, with the outstanding exception of Wester Ross [the other apparent exception, Dee, in Fig. 4(d), is not significant].
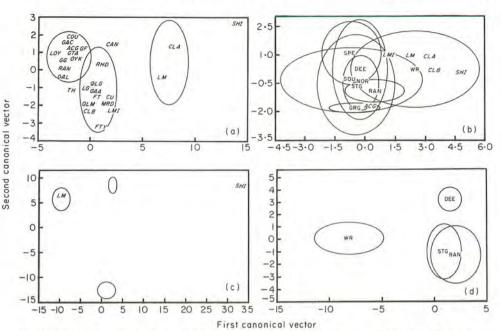
Fig. 4. Canonical variate analysis of Caledonian Scots pine (*Pinus sylvestris*): plots of the first two canonical vectors of monoterpene [(a), (b)] and allozyme [(c), (d)] loci. Clustering of populations is by two criteria: phenetic clusters [(a), (c)] and geographic regional groups [(b), (d)]. The 95 % concentration elipse is drawn for each cluster. Individual population means are shown in italics; all Wester Ross populations [(a), (b), (c)] and a representative sample of populations (a). Regional group means are shown at the centres of their elipses in normal type. To avoid regional groups represented by single populations in the allozyme analysis, Amat was assigned to the Strath Glass and Glen Falloch to the Rannoch group. Other population and regional group symbols as in Appendix 1.

Comparison of allozyme and monoterpene loci

An assumption underlying population genetic studies is that the marker loci used are an unbiased sample of the genome. This assumption can be challenged for allozyme loci because of their nearly exclusive involvement in intermediary metabolism. A more controversial argument is that variants of these loci may be neutral with respect to fitness and therefore are not sensitive to environmental gradients. Thus, any set of loci other than allozymes is valuable for comparison.

In our data, similarities between allozyme and monoterpene loci overshadow differences; estimates of total gene diversity (Ht) are comparable, as is the distribution of variation within (Hs) and among (Gst) populations (Table 1). Although average genetic distances are greater for monoterpenes than for allozymes, high correlation between some monoterpene loci inflate this statistic with redundant information. For example, because the correlation between 3-carene and myrcene is -0.92, using both loci has nearly the same effect as using either locus twice. Since allozymes, on the other hand, tend to have low interlocus correlations, each contributes more independently to multilocus differentiation. This is seen in the higher trace values for allozymes (Table 2).

DISCUSSION

The amount of variability measured in these relict populations of the Caledonian pine forest is surprisingly high, with total gene diversity (Ht = 0·31 to 0·35)

Table 4. *Comparative genetic diversity statistics for Scots pine* (Pinus sylvestris) *in Europe*

Country	No. of Populations	Loci	Ht	Gst	Reference
Sweden	9	11	0·31	0·02	Gullberg *et al.*, 1985
Sweden	3	3	0·39	0·02	Rudin *et al.*, in Hamrick, 1983b.
Central Europe	19	3	0·36	0·09	Mejnartowicz, 1979
Poland	Seed orchard of select trees	12	0·36	—	Krzakowa *et al.*, 1977
Scotland	14	16	0·31	0·03	This study
Scotland	40	5	0·35	0·05	This study

exceeding that of most other conifers (Hamrick, 1983a,b). As a group, conifers are among the most variable organisms, but they span the range from the lowest (Ht zero or near zero for *Pinus torreyana* Parry and *Pinus resinosa* Ait.; Fowler & Morris, 1977; Ledig & Conkle, 1983) to the highest (Ht = 0·33 for *Pinus longaeva*; Bailey, Hiebert & Hamrick, 1983). As in most other conifers, the overwhelming proportion of this variability is among trees within populations. Only a few studies of allozyme variability have been made in other parts of the range of Scots pine, usually with fewer loci and populations, but results are consistent with ours (Table 4).

The amount and distribution of variability are consistent for both monoterpene and allozyme loci, suggesting that these loci may typify the genome as a whole. Steven & Carlisle (1959) studied a number of qualitative and metrical traits in the native pinewoods, ranging from the internal anatomy of leaves to the habit type of whole trees (Table 5). Even though many of their traits were not analyzed statistically, the main pattern is clear. Wide variation is evident for most of the traits examined, and this variation is distributed predominantly among trees in populations; less than a third of the traits also exhibit moderate to strong variation at the population or regional group level. Thus, in broad outline at least, genes controlling specific enzymes and monoterpenes, as well as those which determine complex morphological and anatomical traits, all appear to reflect a similar amount and pattern of diversity.

Whatever the causes of extinction of the Scots pine in large parts of its former range in the British Isles, they evidently have not had a serious impact on its genetic variability, which remains very high even in populations severely reduced in numbers. The hypothesis of heterozygote advantage, when considered in the historical context of the ecological role the Scots pine played during the Pleistocene, fits the observed data. Lerner (1954) proposed that heterozygotes may have an adaptive advantage over homozygotes by reason of superior biochemical efficiency and regulation; diploid individuals with alternate alleles at a locus may be endowed with greater physiological homeostasis to fluctuating environmental change. Presumably, the degree of this buffering is proportional to overall heterozygosity. Heterozygosity ensures greater flexibility in adapting to inevitable long-term environmental changes. Ledig, Guries & Bonefield (1983) recently reviewed the increasing evidence in support of Lerner's hypothesis from a wide variety of organisms and presented new evidence associating heterozygosity with

Table 5. *Distribution of morphological variability in Caledonian Scots pine* (Pinus sylvestris *; from data in Steven & Carlisle, 1959*)

Trait	Predominant source of variation[1]	Remarks
Habit type	T	Six main types
Bark type	T, (G?)	Twelve types, two predominant
Bud colour	T	Brown (909 trees), red (1090 trees)
Leaf length	—	Site most important determinant
Leaf colour	T	Glaucous-silvery to yellow-green
Leaf persistence, years	T, G*	Two to five years
Leaf resin canals, number	T*, P*, G*	Two to 15
Leaf resin canals, position	T	Marginal to stelar
Leaf sclerenchyma, around resin canals	T*	One to two layers
Leaf sclerenchyma, around stele	T*, G*	Well to poorly developed
Hypodermis, thickness	T*	One to three cell layers
Cone colour	T	Four main types
Cone length	P*	2·0 to 6·5 cm
Cone form	T, G	Three types: broad to elongate
Apophysis, shape	T**, P**, G**	Three types: flat to prominently raised
Seed length, length:width	—	Varies within trees
Seed colour	T	Four types: light to dark
Seed wing, form	—	Varies within cone and tree
Seed wing, colour	T*, G*	Four types: pale to dark brown
Seed wing, striping	T	Three types: unstriped to well-defined stripes
Pollen catkin, colour	T	Three types: yellow to violet-red

[1] T, P and G indicate that the observed variability occurred primarily among trees within populations, among populations, or among regional groups, respectively. Single and double asterisks indicate observed differences were statistically significant at $P \leqslant 0.05$ and 0.01, respectively.

a fitness trait (growth rate) in pitch pine (*Pinus rigida* Mill.). Scots pine has been present in Britain during all the recorded interglacial and interstadial periods, being prominent to dominant in the early and late stages but giving way to competition from broad-leaved trees in the interval (West, 1980; Bennett, 1984). During the prolonged cycle from cool to warm–temperate and back again, homeostatic regulation would seem to be of adaptive advantage.

It has long been presumed that the Scots pine could not have survived glacial conditions in the British Isles (McVean, 1964; Godwin, 1975a). It follows from this that contemporary populations in Scotland, as well as those that formerly occupied Ireland and England, must have derived from continental sources by migration over connecting land bridges after the climate had sufficiently ameliorated. The main arguments against survival during the full glacial have been the obvious severity of the climate, together with the paucity of fossil evidence of pine during the late glacial. While it is certain that forests could not have remained,

the evidence against the survival of small groups of trees or even individuals in sheltered habitats is not compelling. On the contrary, neither the genetic evidence nor the geographic pattern of the spread of pine populations indicated by fossil pollen data at early Holocene sites supports the traditional concept of continental origin of Scots pine in Britain, with the probable exception of southern England. The two lines of evidence are complementary; we will attempt to show that they are more consistent with an endemic origin of the Caledonian Scots pine from more than one glacial refugium.

The genetic evidence is implicit in the separate taxonomic status long accorded to Caledonian Scots pine as var. *scotica*. Some of the many named varieties of Scots pine in Europe, although originally based on phenotypic distinctions, have since been shown to have a genetic basis (Ruby, 1967).

Provenance trials have brought the distinctiveness of different seed sources of Scots pine into sharper focus. Early tests in Great Britain, established between 1929 and 1942, compared Scottish seed sources with those of continental origin ranging from northern Scandinavia to the Sierra Nevada in Spain, and from near the Atlantic coast to the Ural Mountains (Lines & Mitchell, 1964). About half of the Scottish sources were from authentic native populations, and the rest came from established plantations (often derived from native populations or pollinated by them, or both). Certain trends emerged clearly, even though limitations in the experimental design of some of these tests precluded statistical analysis. For example, sources from northern Scandinavia, the Urals and southern Spain consistently failed or grew poorly on all sites. Sources from the Northern European Plain and central Germany were clearly superior in growth to Scottish sources in southeastern England, on a site more closely representative of a continental climate, but inferior on a range of sites in Scotland. None of the continental sources appeared to be closely related to Scottish sources, based on one or more parameters of growth, survival, stem-form or branch characteristics, and most were clearly different.

The most comprehensive genetic tests with experimental designs adequate to make comparisons of many different provenances were established in the United States by Wright and his colleagues. Over 160 seed sources from Europe and Asia, many from natural populations, were studied in nursery and plantation environments in many different locations (see Ruby & Wright, 1976, for a general description and references). From results of nursery tests, Wright & Bull (1963) recognized 14 ecotypes, based on a large number of morphological, physiological and phenological traits. Values of each trait were combined in a 'summation-of-differences' index, a multivariate parameter invented by the authors. The Scottish sources comprised one of these ecotypes, distinct from all other continental sources, but showing the least difference from sources in the Massif Central of France. Subsequent measurement of the same material for growth and a variety of other traits after more than 10 years in different Michigan plantations corroborated the earlier classifications with only minor changes, but this time incorporated the various provenances into a revision of previously recognized varieties. While retaining var. *scotica* in this scheme, Ruby & Wright (1976) nevertheless concluded that 'the Scottish population resembles south European varieties so much as to indicate it is a post-Pleistocene derivative of southern rather than central or northern stock'. By central and northern were meant all sources north of the Alps. Their 'southern stock' presumably referred to sources from the Massif Central, since all others from the south (Spain, Italy, Greece, etc.) were

so obviously different from both the Massif Central and Scottish sources in many of the traits compared. But the alleged similarity of Scottish with Massif Central sources is not convincing by their own data: large discontinuities are evident in several traits, including the date of first year bud-set and vulnerability to certain pests. Previously, Ruby (1967) had measured significant differences in 10 out of 19 cone, seed and leaf characters between these two sources. On the basis of monoterpene composition, by contrast, Tobolski & Hanover (1971) considered the Scottish sources to have more genetic affinity with those of middle Europe and southern Scandinavia. They thought the near absence of 3-carene in southern European populations, compared to its relative abundance in almost all northern populations, indicated that northern populations had never come into contact with southern populations in their retreat from advancing glaciers, and therefore that southern populations had no influence on the reforestation of northern Europe. Forrest (1982), on the other hand, perceived genetic influences from Scandinavia, as well as both middle and southern Europe, on various Scottish populations.

A consideration for any population proposed as a progenitor for another, in addition to genetic similarities, is the feasibility of potential migration routes. Broad-scale immigration of pine into East Anglia and southern England began no earlier than a few centuries after the abrupt rise in temperature that marked the opening of the Holocene at about 10000 B.P. (Godwin, 1975a; Huntley & Birks, 1983; Bennett, 1984). Since mean sea-level at this time was probably no more than 30 m below the present level (Devoy, 1979), access to Britain was restricted to a relatively narrow land bridge across the Straits of Dover and parts of the North Sea south of the Dogger Bank (approximately between 51° and 54° N). The simplest hypothesis is that the earliest immigrants derived from ancestors of contemporary populations closest to this area (see Fig. 1), which presently occupy parts of the Vosges mountains and the Black Forest (both areas included in var. *haguenesis*; Ruby & Wright, 1976) or the Northern European Plain (vars *hercynica* and *polonica*). But this is inconsistent with the genetic evidence, which shows little affinity between these and Scottish populations (Wright & Bull, 1963; Lines & Mitchell, 1964; Ruby & Wright, 1976).

Problems in rationalizing more distant continental origins include how migrating populations could precede more proximate populations or pass through or by them in the bottleneck connecting Britain to the continent while still maintaining their own racial integrity. A further complication is that the pattern of variation for allozymes and monoterpenes found in this study requires at least two such origins, one for Wester Ross populations and one for the rest. Selection can always be invoked to explain anomalous patterns of variation, but this is much more difficult to prove.

The palynological evidence in fact strongly suggests that pine did not reach Scotland from England after the last glaciation. Godwin's (1975a) isopol maps show the migration of pine from the south coast of England and East Anglia at the beginning of the Holocene (about 10000 B.P.; pollen zone IV) progressively north and west to a line approximately between the Tweed and western Wales by 8000 B.P. (pollen zone VI; see also Bennett, 1984, p. 142). In the ensuing zone VIIa, pine rapidly disappeared from England and was virtually extinct by historical times. Meanwhile, two widely separated and apparently independent foci of development and centrifugal expansion of pine appeared; one in northwest Scotland and the other in west central Ireland, only a relatively short

time (zone V) after the appearance of significant amounts of pine pollen in England. This general pattern, based on pollen-zone chronologies, has since been confirmed and amplified in detail at many sites by radiocarbon dating (H. H. Birks, 1972; Godwin, 1975b; H. J. B. Birks, 1977, 1985). The pine isopol maps of Huntley & Birks (1983) differ in some detail but lead to the same conclusion. Both sets of maps show a gap over 150 km wide between the northern advance of pine in England and the southern limit reached in Scotland. The problem of moving pine from the continent to Ireland is even greater: its nearly simultaneous appearance in southeastern England and both eastern and west central Ireland in the very early Holocene period requires a short, direct route for rapid migration. But the depth of the St George's Channel in the Irish Sea (and waters south) make any such land connection highly improbable by the time temperatures had warmed enough to allow rapid and extensive plant migration. Devoy (1985) concluded that if any functional land bridge existed between Britain and Ireland, it could only have been between northern Ireland and southwest Scotland.

The idea that pine survived the last glacial episode in the British Isles is not new but has been clouded with ambiguity. Godwin's (1975a) statement, 'It is hard to believe that any trees save perhaps birches survived in this country south of the ice' is difficult to reconcile with his comment elsewhere in the same year (Godwin, 1975b), '...if this gap [the 150 to 300 km separating the furthest measured advances of the English and Scottish populations]... can be shown to be real it will strengthen the case for considering that the tree reinvaded from the northwest as well as from the south, a fact possibly to be associated with the separate taxonomic status of the Scottish native pine as *P. sylvestris* var. *scotica*'. His cryptic allusion to the northwest was not elaborated but perhaps implied a pine refugium somewhere in the Hebrides. Huntley & Birks (1983, p. 308) considered that '...the early Holocene behaviour of pine in Ireland and Scotland suggests that refugia may have existed in or near western Britain', and their map (p. 650) indicated the Outer Hebrides, the northwestern tip of Skye and two peninsulas in the extreme west central part of Ireland as possible locations, although evidence for these was not documented. Bennett (1984, p. 149) invoked rare events of long-distance seed dispersal to explain the discontinuous distribution and expansion of Scots pine from different foci in England, Scotland and Ireland, but he also allowed the possibility of one or more glacial refugia on the continental shelf to the west.

The precise origins of contemporary Scottish populations may never be known, but important clues can be found in patterns of genetic variation. From the main dichotomy between most Wester Ross populations and virtually all the others, at least two founder populations can be inferred. The division is based on multivariate statistical parameters representing differences in allozyme and monoterpene gene frequencies (Fig. 4; Tables 3 and 4). It is also strongly supported by differences in physiological traits observed in earlier field trials. The most dramatic of these was an experiment at Glen Trool, a harsh site in southwest Scotland selected to assess the adaptability of different Scottish provenances to climates in western Britain (Forestry Commission, 1965 and unpublished data). Only six native populations were included, but these represented strong comparisons between Wester Ross and more easterly sources. Trees from Shieldaig and Loch Maree were consistently less affected by winter blast than any other source, and after 19 years were considered to be the only ones capable of making a crop on the site. Trees from Achnashellach, a source that did not group with other Wester Ross

populations for monoterpene gene frequencies, were intermediate in blast susceptibility.

In fact, the heterogeneity among these populations in western Scotland contrasted strongly with the relatively small and random differences among populations in most other parts of the country. These western populations possibly derived from more than one founder population or represent the result of eastern and western founders coming together. The fragmentation of populations and redistribution of their genetic variation into transient races was one of the major impacts of the Pleistocene identified by Critchfield (1984) on several boreal conifers of North America. According to this hypothesis, remnants of populations destroyed during the full glacial emerge from refugia at the beginning of each interglacial period, rapidly expand, coalesce to varying degrees and thereby undergo considerable genetic restructuring. The anomalous patterns of variation generated may thus largely reflect historical accidents. Although the same general sequence of events recurs in each glacial–interglacial cycle, characteristics of refugial populations, their number, size, location, etc. vary, restricting the life of any race to one cycle. The genetic data of Caledonian Scots pine fit this hypothesis well.

The hypothesis of endemic origins of the Caledonian Scots pine gets scant support from the known palaeo-climatic facts. During the height of the last glacial period, the Gulf Stream was offset to the south, flowing on a course due east to the Iberian peninsula. The temperature gradient marking its northern boundary at 42° N was very steep, and August surface temperatures of the ocean surrounding Britain were as low as 1 to 3 °C (CLIMAP, 1976). Unglaciated sites may have been few in the north and subject to permafrost or solifluction.

Yet, it is easier, in our view, to rationalize these difficulties than to reconcile a coherent theory of immigration of the Scots pine from the continent with the existing genetic and fossil evidence. It further must be acknowledged that many gaps remain in our understanding of Pleistocene climates and potential glacial refugia. In the dissected topography of mountainous regions, local climates are more variable and often less severe than the climate in general and are perhaps the controlling factors in the survival, re-immigration, and evolution of forest vegetation (Frenzel, 1968).

REFERENCES

BARADAT, PH., BERNARD-DAGAN, C., FILLAN, C., MARPEAU, G. & PAULY, G. (1972). Les terpènes du pin maritime: aspects biologiques et génétiques. II. Hérédité de la tenur en monoterpènes. *Annales de Science Forestière*, **29**, 307–334.

BARADAT, PH., BERNARD-DAGAN, C., PAULY, G. & ZIMMERMANN-FILLON, C. (1975). Les terpènes du pin maritime: aspects biologiques et génétiques. III. Hérédité de la teneur en myrcène. *Annales de Science Forestière*, **32**, 29–54.

BENNETT, K. D. (1984). The post-glacial history of *Pinus sylvestris* in the British Isles. *Quaternary Science Reviews*, **3**, 133–155.

BIRKS, H. H. (1972). Studies in the vegetational history of Scotland. III. A radiocarbon-dated pollen diagram from Loch Maree, Ross and Cromarty. *New Phytologist*, **71**, 731–754.

BIRKS, H. H. (1975). Studies in the vegetational history of Scotland. IV. Pine stumps in Scottish blanket peats. *Philosophical Transactions of the Royal Society of London B*, **240**, 182–226.

BIRKS, H. J. B. (1977). The Flandrian forest history of Scotland: a preliminary synthesis. In: *British Quaternary Studies — Recent Advances* (Ed. by F. W. Shotton), pp. 119–135. Clarendon Press, Oxford.

BIRKS, H. J. B. (1985). A pollen mapping project in Norden for 0–13000 B.P. NORDMAP 1. *University of Bergen Botanical Institute Report*, **38**, 1–41.

CARLISLE, A. (1977). The impact of man on the native pinewoods of Scotland. In: *Native Pinewoods of Scotland*. Proceedings of the Aviemore Symposium, 1975 (Ed. by R. G. H. Bunce & J. N. R. Jeffers), pp. 70–77. Institute of Terrestrial Ecology, Cambridge.

CLIMAP PROJECT MEMBERS. (1976). The surface of the ice age earth. *Science*, **191**, 1131–1137.

CONKLE, M. T., HODGSKISS, P. D., NUNNALLY, L. B. & HUNTER, S. C. (1982) *Starch Gel Electrophoresis of Conifer Seeds: A Laboratory Manual*. US Department of Agriculture, Forest Service, General Technical Report PSW-64.

COOLEY, W. W. & LOHNES, P. R. (1971). *Multivariate Data Analysis*. John Wiley, New York.

CRITCHFIELD, W. B. (1984). Impact of the Pleistocene on the genetic structure of North American conifers. In: *Proceedings of the Eighth North American Forest Biology Workshop, Logan, Utah, July 30–August 1, 1984* (Ed. by R. M. Lanner), pp. 70–118. Utah State University, Logan.

CRITCHFIELD, W. B. & LITTLE, JR, E. L. (1966). *Geographic Distribution of the Pines of the World*. US Department of Agriculture, Forest Service, Miscellaneous Publication 991, Washington.

DEVOY, R. J. (1979). Flandrian sea level changes and vegetation history of the lower Thames estuary. *Philosophical Transactions of the Royal Society of London B*, **285**, 355–410.

DEVOY, R. J. (1985). The problem of a late quaternary landbridge between Britain and Ireland. *Quaternary Science Reviews*, **4**, 43–58.

DUNN, G. & EVERITT, B. S. (1982). *An Introduction to Mathematical Taxonomy. Cambridge Studies in Mathematical Biology, 5*. Cambridge, London.

FORESTRY COMMISSION (1965). *Report on Forest Research for the Year Ended March, 1965*. HMSO, London.

FORREST, G. I. (1980). Genotypic variation among native Scots pine populations in Scotland based on monoterpene analysis. *Forestry*, **53**, 101–128.

FORREST, G. I. (1982). Relationship of some European Scots pine populations to native Scottish woodlands based on monoterpene analysis. *Forestry*, **55**, 109–37.

FOWLER, D. P. & MORRIS, R. W. (1977). Genetic diversity in red pine: evidence for low heterozygosity. *Canadian Journal of Forest Research*, **7**, 341–347.

FRENZEL, B. (1968). The Pleistocene vegetation of northern Eurasia. *Science*, **161**, 637–649.

GODWIN, H. (1975a). History of the natural forests of Britain: establishment, dominance, and destruction. *Philosophical Transactions of the Royal Society of London B*, **271**, 47–67.

GODWIN, H. (1975b). *History of the British Flora*, 2nd Edn. Cambridge University Press, Cambridge.

GOODIER, R. & BUNCE, R. G. H. (1977). The native pinewoods of Scotland: the current state of the resource. In: *Native Pinewoods of Scotland*. Proceedings of the Aviemore Symposium, 1975 (Ed. by R. G. H. Bunce & J. N. R. Jeffers), pp. 78–87. Institute of Terrestrial Ecology, Cambridge.

GULLBERG, U., YAZDONI, R., RUDIN, D. & RYANNEN, N. (1985). Allozyme variation in Scots pine (*Pinus sylvestris* L.) native to Sweden. *Silvae Genetica*, **34**, 193–201.

GURIES, R. P. & LEDIG, F. T. (1982). Genetic diversity of population structure in pitch pine (*Pinus rigida* Mill.). *Evolution*, **36**, 387–402.

HAMRICK, J. L. (1983a). The distribution of genetic variation within and among natural plant populations. In: *Genetics and Conservation* (Ed. by C. M. Schonewald-Cox, S. M. Chambers, S. M. MacBryde & W. L. Thomas), pp. 335–348. Benjamin/Cummings, London.

HAMRICK, J. L. (1983b). Appendix 4. The distribution of genetic variation within and among natural plant populations. In: *Genetics and Conservation* (Ed. by C. M. Schonewald-Cox, S. M. Chambers, S. M. MacBryde & W. L. Thomas), pp. 500–508. Benjamin/Cummings, London.

HANOVER, J. W. (1966). Inheritance of 3-carene concentration in *Pinus monticola*. *Forest Science* **12**, 447–450.

HIEBERT, R. D. & HAMRICK, J. L. (1983). Patterns and levels of genetic variation in Great Basin Bristlecone pine *Pinus longaeva*. *Evolution*, **37**, 302–311.

HILTUNEN, R., TIGERSTEDT, P. M. A., JUVONEN, S. & POHJOLA, J. (1975). Inheritance of 3-carene quantity in *Pinus sylvestris* L. *Farmaseuttinen Aikakauslehti*, **84**, 69–72.

HUNTLEY, B. & BIRKS, H. J. B. (1983). *An Atlas of Past and Present Pollen Maps for Europe: 0–13000 years ago*. Cambridge University Press, Cambridge.

KRZAKOWA, M., SZWEYKOWSKI, J. & KOROZYK, A. (1977). Population genetics of Scots pine (*Pinus sylvestris* L.) forests. Genetic structure of plus-trees in Bolwice near Poznan (West Poland). *Bulletin de L'Académie Polonaise des Sciences*, série des sciences biologiques, **25**, 583–590.

LEDIG, F. T. & CONKLE, M. T. (1983). Gene diversity and genetic structure in a narrow endemic, Torrey pine (*Pinus torreyana* Parry ex Carr.). *Evolution*, **37**, 79–85.

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

LERNER, L. M. (1954). *Genetic Homeostasis*. Wiley & Sons, New York.

LINES, R. & MITCHELL, A. F. (1964). Results of some older Scots pine provenance experiments. *Report on Forest Research for the Year Ended March, 1964*, pp. 171–194. HMSO, London.

McVEAN, D. N. (1964). Pre-history and ecological history. In: *The Vegetation of Scotland* (Ed. by J. H. Burnett), pp. 561–567. Oliver & Boyd, Edinburgh & London.

MEJNARTOWICZ, L. (1979). Genetic variation in some isozyme loci in Scots pine (*Pinus sylvestris* L.) populations. *Arboretum Kornickie Rocznik*, **24**, 91–104.

NEI, M. (1975). *Molecular Population Genetics and Evolution*. North-Holland, Amsterdam.

NEI, M. (1981). Gene differences between human races. *American Naturalist*, **117**, 88–89.

O'SULLIVAN, P. E. (1977). Vegetation history and the native pinewoods. In: *Native Pinewoods of Scotland* (Ed. by R. G. H. Bunce & J. N. R. Jeffers), pp. 60–69. Proceedings of the Aviemore Symposium, 1975. Institute of Terrestrial Ecology, Cambridge.

RUBY, J. L. (1967). The correspondence between genetic, morphological, and climatic variation patterns in Scotch pine. I. Variations in parental characters. *Silvae Genetica*, **16**, 50–56.

RUBY, J. L. & WRIGHT, J. W. (1976). A revised classification of geographic variation in Scots pine. *Silvae Genetica*, **25**, 169–175.

SAS INSTITUTE (1982). SAS *Users Guide: Statistics*. SAS Institute, Cary, North Carolina.

SMOUSE, P. E., SPIELMAN, R. S. & PARK, M. H. (1982). Multiple locus allocation of individuals to groups as a function of the genetic variation within and differences among human populations. *American Naturalist*, **119**, 445–463.

SNEATH, P. H. & SOKAL, R. R. (1973). *Numerical Taxonomy*. W. H. Freeman, San Francisco.

SQUILLACE, A. E. (1971). Inheritance of monoterpene composition in cortical oleoresin of slash pine. *Forest Science*, **17**, 381–387.

SQUILLACE, A. E., WELLS, O. O. & ROCKWOOD, D. L. (1980). Inheritance of monoterpene composition in cortical oleoresin of loblolly pine. *Silvae Genetica*, **29**, 141–151.

STEVEN, H. M. & CARLISLE, A. (1959). *The Native Pinewoods of Scotland*. Oliver & Boyd, London.

STRAUSS, S. H. & CRITCHFIELD, W. B. (1982). Inheritance of β-pinene in xylem oleoresin of knobcone × Monterey pine hybrids. *Forest Science*, **28**, 687–696.

SWOFFORD, D. L. & SELANDER, R. B. (1981a). Biosys-1: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. *Journal of Heredity*, **72**, 281–283.

SWOFFORD, D. L. & SELANDER, R. B. (1981b). *Biosys-1. A Computer Program for the Analysis of Allelic Variation in Genetics*. Users manual. Department of Genetics and Development, University of Illinois, Urbana.

TOBOLSKI, J. F. & HANOVER, J. W. (1971). Genetic variation in the monoterpenes of Scotch pine. *Forest Science*, **17**, 293–299.

WEST, R. G. (1980). Pleistocene forest history in East Anglia. *New Phytologist*, **85**, 571–622.

WHEELER, N. C. & GURIES, R. P. (1982). Population structure, genic diversity, and morphological variation in *Pinus contorta* Dougl. *Canadian Journal of Forest Research*, **12**, 595–606.

WORKMAN, P. L. & NISWANDER, J. D. (1970). Population studies on southwest indian tribes. II. Local genetic differentiation in the Papago. *American Journal of Human Genetics*, **22**, 24–49.

WRIGHT, J. W. & BULL, W. F. (1963). Geographic variation in Scotch pine. Results of a 3-year Michigan study. *Silvae Genetica*, **12**, 1–25.

Appendix 1. *Caledonian Scots Pine* (Pinus sylvestris) *populations sampled for monoterpenes and allozymes*(*)

Population no.	Code	Name†	Grid reference	Elevation (m)
Northern (NOR)				
1	RHD	Rhidorroch	NH 239930	150
2	OYK	Strath Oykell	NC 462016	90
3	EIN	Glen Einig	NH 367987	90
4	AM*	Amat	NH 466895	135
5	STV	Strath Vaich	NH 342765	300
Wester Ross (WR)				
6	LM*	Loch Maree	NG 995654	25
7	LMI	Eilean Ruairidh Mor	NG 897735	45
8	SHI*	Shieldaig	NG 825519	15
9	CLA	Coulin	NG 990575	120
10	CLB	Coulin	NG 015525	215
11	ACG	Achnashellach	NH 019471	185
Strath Glass (STG)				
12	SF	Glen Strathfarrar	NH 165377	335
13	CAN	Glen Cannich	NH 255339	380
14	GAA*	Glen Affric (West)	NH 158219	250
15	GAB*	Glen Affric (Central)	NH 225237	250
16	GAC*	Glen Affric (East)	NH 282279	260
17	GUI*	Guisachan	NH 248238	330
18	COU*	Cougie	NH 235209	300
Great Glen (GRG)				
19	GG	Glengarry	NN 177989	215
20	LOY	Glen Loy	NN 092842	150
21	ARD	Ardgour	NM 967713	190
Southern (SOU)				
22	BM	Black Mount	NN 285417	230
23	GF*	Glen Falloch	NN 365229	215
Rannoch Moor (RAN)				
24	RAN*	Rannoch	NN 572563	230
25	MEG*	Meggernie	NN 563454	335
Speyside (SPE)				
28	DAL	Dulnan	NH 845194	335
29	LG	Abernethy	NH 976185	230
30	FT	Abernethy	NJ 015175	275
30	FTY	Abernethy	NJ 015175	275
31	TH	Abernethy	NH 998167	265
32	CU	Abernethy	NJ 012163	290
33	ROT	Rothiemurchus	NH 918093	275
34	QLM	Queen's (Loch Morlich)	NH 957097	320
35	QGL	Queen's (Glenmore)	NH 995102	365
36	GAV	Glen Avon	NJ 171074	455
37	FES	Glen Feshie	NN 855999	365
Deeside (DEE)				
38	MRD*	Mar	NO 040940	455
39	GTT*	Glentanar (Tanar)	NO 451942	260
40	GTG	Glentanar (Gairney)	NO 462922	305
41	GTA	Glantanar (Allachy)	NO 485915	305
42	BB*	Ballochbuie‡	NO 200895	405

† Population and regional group names after Steven & Carlisle (1959).
‡ Not sampled for monoterpenes.

Appendix 2. *Isozyme systems assayed in Caledonian Scots pine* (Pinus sylvestris)

Code	Name	EC Reference	Buffer system*	No. of loci
(A) Used in population analysis					
ACP	Acid phosphatase	3.1.3.2	B	1†
ADH	Alcohol dehydrogenase	1.1.1.1	A	1†
GDH	Glutamate dehydrogenase	1.4.1.3	B	1
GOT	Glutamate-oxalacetate transaminase	2.6.1.1	B	3
IDH	Isocitrate dehydrogenase	1.1.1.42	C	1
LAP	Leucine aminopeptidase	3.4.11.1	A	2
MDH	Malate dehydrogenase	1.1.1.37	C	4
PGM	Phosphoglucomutase	2.7.5.1	A	1†
6PGD	6–Phosphogluconate dehydrogenase	1.1.1.44	C	2
(B) Assayed, but not used in population analysis‡					
ACO	Aconitase	4.2.1.3	C	1
EST	Alpha esterase	3.1.1.1	A	3–4
CAT	Catalase	1.11.1.6	B	2
G6PD	Glucose 6–phosphate dehydrogenase	1.1.1.49	B	2
PEP	Peptidase	3.4.13.1	A	3
DIA	Diaphorase	1.6.4.3	A	3
ALD	Aldolase	4.1.2.13	C	2
PGI	Phosphoglucose isomerase	5.3.1.9	A	2

* Gel/electrode buffers (Conkle *et al.*, 1982): A, Tris citrate (pH 8·3)/lithium borate (pH 8·3); B, Tris citrate (pH 8·8)/sodium borate (pH 8·0); C, Tris citrate (pH 6·2)/(same as gel buffer).

† Additional loci detected for these systems, relative to the main locus, were as follows: ACP, one anodal; ADH, one anodal and one cathodal; PGM, one cathodal.

‡ Staining was poor to inconsistent in these systems, except for ACO and PEP, for which population samples were incomplete.

Appendix 3. *Frequency of the most common allele for five monoterpene loci in 40 populations of Scots pine, average heterozygosity (H), average genetic distance (D) of each population (and regional group) from all the others and canonical scores (Cs) of the first canonical variate for each population (and regional group)*

Population†	n	3-Carene	Myrcene	β-Pinene	Limonene	β-Phellandrene	H	D	Cs‡ (1)	(2)
Northern		0·611	0·690	0·772	0·888	0·895	0·328	0·013	−0·22	−0·31
01 RHD	288	0·548	0·646	0·741	0·886	0·892	0·347	0·014	−0·05	−1·21
02 OYK	92	0·652	0·691	0·767	0·897	0·891	0·325	0·014	−0·53	−1·32
03 EIN	54	0·611	0·694	0·817	0·913	0·933	0·299	0·014	0·08	−0·37
04 AM*	284	0·656	0·744	0·822	0·843	0·886	0·319	0·015	0·25	−1·34
05 STV	92	0·587	0·676	0·715	0·903	0·872	0·348	0·013	−0·85	0·27
Wester Ross		0·431	0·543	0·768	0·772	0·840	0·364	0·059	2·50	4·92
06 LM*	216	0·340	0·396	0·677	0·822	0·869	0·378	0·065	2·22	6·82
07 LMI	213	0·465	0·643	0·729	0·748	0·914	0·378	0·021	1·14	2·22
08 SHI*	264	0·153	0·261	0·731	0·841	0·769	0·333	0·144	4·80	14·04
09 CLA	101	0·361	0·545	0·764	0·751	0·731	0·419	0·047	3·18	8·13
10 CLB	85	0·594	0·694	0·861	0·584	0·847	0·381	0·038	3·41	−0·05
11 ACG	123	0·674	0·721	0·846	0·888	0·910	0·294	0·016	0·26	−1·63
Strath Glass		0·637	0·731	0·783	0·886	0·900	0·308	0·017	−0·46	−1·11
12 SF	132	0·689	0·754	0·830	0·921	0·925	0·274	0·018	−0·53	−2·30
13 CAN	185	0·562	0·666	0·841	0·825	0·841	0·353	0·015	1·70	2·04
14 GAA*	161	0·549	0·605	0·645	0·885	0·906	0·362	0·020	−1·11	0·31
15 GAB*	90	0·544	0·715	0·699	0·810	0·925	0·356	0·015	−0·59	−0·12
16 GAC*	77	0·701	0·797	0·789	0·912	0·853	0·299	0·020	−1·18	−2·11
17 GUI*	139	0·738	0·844	0·853	0·898	0·933	0·243	0·026	−0·89	−3·89
18 COU*	123	0·678	0·738	0·826	0·954	0·915	0·272	0·018	−0·65	−1·68
Great Glen		0·724	0·798	0·854	0·902	0·932	0·254	0·020	−0·54	−3·35
19 GG	123	0·691	0·831	0·879	0·841	0·920	0·268	0·023	0·21	−2·53
20 LOY	92	0·722	0·752	0·828	0·938	0·956	0·253	0·021	−0·90	−3·51
21 ARD	210	0·759	0·811	0·854	0·926	0·921	0·241	0·026	−0·92	−4·00
Southern		0·724	0·801	0·802	0·866	0·819	0·307	0·027	−0·70	−2·37
22 BM	33	0·788	0·938	0·792	0·871	0·778	0·274	0·044	−1·97	−3·92
23 GF*	50	0·660	0·663	0·812	0·860	0·860	0·340	0·015	0·58	−0·81
Rannoch		0·637	0·718	0·820	0·836	0·910	0·320	0·013	0·47	−1·16
24 RAN*	258	0·679	0·736	0·820	0·860	0·930	0·299	0·014	−0·06	−2·46
25 MEG*	192	0·595	0·699	0·819	0·812	0·889	0·341	0·016	1·00	0·15
Speyside		0·580	0·715	0·707	0·830	0·886	0·356	0·017	−0·69	−0·33
28 DAL	333	0·679	0·771	0·760	0·877	0·915	0·305	0·016	−1·20	−2·77
29 LG	109	0·597	0·691	0·691	0·835	0·867	0·370	0·013	−0·80	−0·52
30 FT	135	0·544	0·661	0·725	0·794	0·898	0·372	0·014	0·32	0·57
30 FTY	68	0·477	0·581	0·618	0·804	0·939	0·380	0·028	−0·48	1·11
31 TH	76	0·651	0·752	0·778	0·858	0·903	0·320	0·014	−0·53	−1·77
32 CU	90	0·494	0·624	0·675	0·806	0·876	0·390	0·019	0·08	1·93
33 ROT	262	0·565	0·752	0·713	0·796	0·865	0·367	0·014	−0·48	0·12
34 QLM	118	0·555	0·736	0·664	0·834	0·897	0·360	0·016	−1·36	−0·24
35 QLG	121	0·562	0·692	0·744	0·828	0·896	0·356	0·012	0·04	0·37
36 GAV	29	0·638	0·851	0·669	0·851	0·809	0·350	0·025	−2·33	−1·36
37 FES	158	0·620	0·759	0·738	0·842	0·879	0·342	0·013	−0·81	−1·07
Deeside		0·593	0·704	0·749	0·859	0·869	0·344	0·017	−0·26	0·13
38 MRD*	142	0·503	0·682	0·697	0·717	0·814	0·414	0·021	0·77	2·11
39 GTT*	181	0·541	0·643	0·725	0·910	0·910	0·337	0·015	−0·44	1·14
40 GTG	109	0·652	0·729	0·772	0·904	0·867	0·323	0·014	−0·71	−1·02
41 GTA	157	0·675	0·762	0·802	0·903	0·885	0·301	0·016	−0·67	−1·72
Total: 5765 Mean:		0·594	0·699	0·763	0·849	0·882	0·332	0·024		

† Population numbers and abbreviations as in Appendix 1 and Figure 1; regional groups are according to Steven & Carlisle (1959). Populations with both allozyme and monoterpene data are indicated by asterisks.

‡ Cs from clustering by (1) regional groups and (2) K-means clusters.

Appendix 4. *Sample size (N), allele frequencies, average heterozygosities (H), alleles per locus (A), percentage of loci polymorphic (P), average genetic distance (D) and canonical scores of the first canonical variate (Cs) for 16 allozyme loci in 14 populations of Scots pine*

Locus/allele	Northern	Wester Ross		Strath Glass					Southern	Rannoch		Deeside			Mean
	(4)* AM	(6) LM	(8) SHI	(14) GAA	(15) GAB	(16) GAC	(17) GUI	(18) COU	(23) GF	(24) RAN	(25) MEG	(38) MRD	(39) GTT	(42) BB	
ACP2 (N)	35	24	25	41	43	32	25	44	34	44	30	27	21	13	13
1	0·843	0·771	0·796	0·756	0·795	0·656	0·740	0·885	0·838	0·828	0·833	0·889	0·810	0·846	0·806
2	0·144	0·125	0·122	0·134	0·108	0·250	0·240	0·081	0·118	0·137	0·117	0·092	0·119	0·038	0·128
3	0·043	0·104	0·082	0·110	0·097	0·063	0·020	0·034	0·044	0·035	0·050	0·019	0·048	0·077	0·059
4	—	—	—	—	—	0·031	—	—	—	—	—	—	0·024	0·038	0·007
ADH2 (N)	23	23	19	24	20	21	23	27	25	25	24	24	13	9	
1	0·457	0·422	0·711	0·334	0·536	0·405	0·445	0·226	0·260	0·280	0·459	0·468	0·520	0·471	0·428
2	0·543	0·578	0·289	0·644	0·464	0·595	0·555	0·774	0·740	0·720	0·541	0·532	0·480	0·529	0·570
3	—	—	—	0·022	—	—	—	—	—	—	—	—	—	—	0·002
GDH (N)	30	24	25	29	35	28	25	34	32	33	27	27	21	13	
1	0·183	0·125	0·143	0·345	0·209	0·182	0·245	0·223	0·238	0·246	0·186	0·092	0·333	0·154	0·207
2	0·817	0·875	0·857	0·655	0·791	0·818	0·755	0·777	0·762	0·754	0·814	0·908	0·667	0·846	0·793
GOT1 (N)	30	24	25	33	37	29	25	27	32	34	27	26	21	13	
1	0·983	1·000	1·000	0·970	1·000	1·000	1·000	0·982	0·984	1·000	1·000	1·000	0·905	1·000	0·987
2	—	—	—	0·030	—	—	—	—	0·016	—	—	—	0·048	—	0·008
3	0·017	—	—	—	—	—	—	0·018	—	—	—	—	0·048	—	0·005
GOT2 (N)	37	24	21	32	36	27	24	26	30	32	25	19	18	13	
1	0·227	0·333	0·048	0·143	0·171	0·241	0·362	0·173	0·134	0·157	0·080	0·270	0·195	0·291	0·202
2	0·302	0·188	0·357	0·349	0·272	0·241	0·191	0·269	0·250	0·359	0·300	0·325	0·278	0·250	0·281
3	0·471	0·479	0·595	0·509	0·557	0·518	0·447	0·558	0·616	0·484	0·620	0·405	0·527	0·458	0·517
GOT3 (N)	30	24	25	33	35	29	25	27	34	34	27	27	21	13	
1	0·183	0·292	0·306	0·213	0·222	0·263	0·160	0·186	0·333	0·165	0·203	0·241	0·214	0·269	0·232
2	0·817	0·708	0·694	0·787	0·778	0·737	0·840	0·814	0·667	0·835	0·797	0·759	0·762	0·731	0·766
3	—	—	—	—	—	—	—	—	—	—	—	—	0·024	—	0·002
IDH (N)	30	23	24	23	34	27	23	28	32	33	18	25	21	13	
1	1·000	1·000	1·000	1·000	0·970	1·000	1·000	1·000	0·953	1·000	1·000	1·000	1·000	1·000	0·995
2	—	—	—	—	0·030	—	—	—	—	—	—	—	—	—	0·002
3	—	—	—	—	—	—	—	—	0·047	—	—	—	—	—	0·003

Appendix 4. Sample size (N), allele frequencies, average heterozygosities (H), alleles per locus (A), percentage of loci polymorphic (P), average genetic distance (D) and canonical scores of the first canonical variate (Cs) for 16 allozyme loci in 14 populations of Scots pine

Locus/allele	Northern	Wester Ross		Strath Glass					Southern	Rannoch		Deeside			Mean
	(4)* AM	(6) LM	(8) SHI	(14) GAA	(15) GAB	(16) GAC	(17) GUI	(18) COU	(23) GF	(24) RAN	(25) MEG	(38) MRD	(39) GTT	(42) BB	
LAP1 (N)	31	24	25	34	37	29	25	35	32	34	28	27	21	13	
1	0·050	—	—	0·044	—	0·018	0·020	0·057	0·031	0·045	—	0·037	—	—	0·022
2	0·590	0·604	0·592	0·717	0·713	0·666	0·520	0·711	0·635	0·632	0·643	0·629	0·667	0·615	0·638
3	0·066	—	0·061	0·044	0·014	0·071	0·040	—	0·016	0·103	0·035	—	0·024	—	0·034
4	0·017	—	0·020	—	—	0·018	—	0·014	—	—	0·054	0·019	0·048	—	0·014
5	0·278	0·396	0·327	0·194	0·274	0·228	0·420	0·217	0·317	0·221	0·268	0·315	0·262	0·385	0·293
LAP2 (N)	31	24	25	32	34	29	25	34	32	34	28	27	21	13	
1	0·919	0·979	0·857	0·875	0·970	0·929	0·940	0·941	0·952	0·808	0·910	0·963	0·976	0·885	0·920
2	0·081	0·021	0·122	0·079	0·030	0·071	0·060	0·059	0·031	0·162	0·072	0·037	0·024	0·038	0·063
3	—	—	0·020	0·046	—	—	—	—	0·016	0·015	0·018	—	—	0·038	0·011
4	—	—	—	—	—	—	—	—	—	0·015	—	—	—	0·038	0·004
MDH1 (N)	35	24	25	34	40	31	24	43	34	42	30	24	21	13	
1	0·942	0·979	1·000	1·000	1·000	0·951	0·959	1·000	0·985	1·000	0·983	0·958	0·857	0·962	0·970
2	0·058	0·021	—	—	—	0·049	0·041	—	0·015	—	0·017	0·042	0·143	0·038	0·030
MDH2 (N)	33	23	25	32	40	31	24	41	32	38	28	22	20	13	
1	0·925	0·956	0·980	0·969	0·962	1·000	0·936	0·964	0·906	0·948	0·964	1·000	1·000	0·923	0·960
2	0·075	0·044	0·020	0·031	0·038	—	0·064	0·036	0·094	0·052	0·036	—	—	0·077	0·040
MDH3 (N)	34	23	25	32	39	31	23	42	32	38	28	23	20	13	
1	0·701	0·826	0·776	0·813	0·718	0·704	0·805	0·723	0·734	0·776	0·696	0·644	0·700	0·731	0·739
2	0·284	0·130	0·102	0·156	0·231	0·164	0·130	0·144	0·172	0·157	0·286	0·356	0·250	0·231	0·200
3	—	0·044	0·061	0·015	0·025	0·049	0·021	0·048	0·031	0·027	—	—	0·025	0·038	0·027
4	0·015	—	0·041	0·015	0·025	0·066	0·043	0·085	0·063	0·040	0·018	—	0·025	—	0·031
5	—	—	0·020	—	—	0·017	—	—	—	—	—	—	—	—	0·003
MDH4 (N)	35	24	25	34	41	31	24	43	34	42	30	23	21	13	
1	0·029	—	0·040	—	0·012	0·017	—	—	—	—	0·033	—	0·024	—	0·011
2	0·114	0·229	0·120	0·090	0·087	0·197	0·125	0·256	0·136	0·147	0·183	0·022	0·119	0·115	0·139
3	0·257	0·250	0·220	0·388	0·395	0·344	0·417	0·232	0·319	0·306	0·333	0·369	0·310	0·423	0·326
4	—	—	—	—	—	—	—	0·024	—	0·012	—	—	0·024	—	0·004
5	—	—	—	—	—	—	—	—	—	0·012	0·017	—	—	—	0·002
6	0·600	0·521	0·620	0·522	0·506	0·442	0·458	0·488	0·545	0·523	0·433	0·608	0·524	0·462	0·518
PGM2 (N)	29	24	24	30	26	27	23	30	28	34	27	22	21	13	
1	0·965	1·000	1·000	0·983	0·963	0·981	1·000	1·000	1·000	1·000	1·000	0·981	0·976	1·000	0·989
2	0·035	—	—	0·017	0·037	0·019	—	—	—	—	—	0·019	0·024	—	0·011

		(N)	24	23	24	28	30	27	16	34	34	34	19	10	19	12	
6PGD1	1		0·679	0·800	0·638	0·563	0·644	0·679	0·645	0·658	0·603	0·676	0·565	0·700	0·702	0·709	0·662
	2		0·086	0·022	0·043	0·110	0·153	0·037	0·130	0·059	0·088	0·103	0·128	0·100	0·027	0·083	0·084
	3		0·086	—	0·085	0·054	0·034	0·057	0·032	0·014	0·176	0·030	0·128	—	—	0·083	0·056
	4		0·021	0·089	—	0·110	0·067	0·037	0·097	0·030	—	0·030	0·026	—	0·054	0·041	0·043
	5		0·064	—	0·021	—	—	0·019	—	0·014	—	0·015	—	—	0·054	—	0·013
	6		0·064	0·089	0·213	0·164	0·102	0·170	0·097	0·224	0·132	0·147	0·154	0·201	0·162	0·083	0·143
6PGD2	(N)		24	23	23	28	32	28	16	36	34	35	19	10	19	12	
	1		0·727	0·696	0·822	0·630	0·604	0·655	0·774	0·569	0·765	0·681	0·659	0·600	0·783	0·739	0·693
	2		0·205	0·304	0·156	0·278	0·302	0·236	0·129	0·375	0·162	0·232	0·268	0·400	0·190	0·217	0·247
	3		0·046	—	0·022	0·018	0·047	0·109	—	—	0·074	0·029	0·048	—	—	—	0·028
	4		—	—	—	0·018	0·031	—	—	0·028	—	—	—	—	—	—	0·006
	5		0·023	—	—	0·055	0·016	—	0·097	0·028	—	0·057	0·025	—	0·027	0·043	0·027
H			0·318	0·284	0·290	0·325	0·310	0·329	0·312	0·291	0·311	0·310	0·314	0·287	0·329	0·316	0·309
A			2·8	2·2	2·6	2·8	2·7	2·8	2·5	2·7	2·6	2·8	2·7	2·2	2·9	2·5	2·6
P			94	81	75	88	88	81	81	81	94	75	81	81	88	81	84
D			0·009	0·013	0·021	0·012	0·010	0·010	0·014	0·015	0·014	0·012	0·010	0·014	0·012	0·010	
Cs(1)†			0·025	6·934	9·308	−2·384	−1·099	−0·991	−0·224	−0·919	−1·429	−1·251	−3·450	−1·347	−2·136	−1·035	
Cs(2)			3·170	−9·241	34·104	1·236	2·158	−9·532	−11·275	0·801	−0·136	2·879	1·363	2·519	−9·758	−8·288	

* Population numbers and abbreviations as in Appendix 1 and Figure 1; regional groups are according to Steven & Carlisle (1959).
† Clustering criteria for Cs (1) and (2) are by regional groups and *K*-means clusters, respectively.